

# REVISTA DE LA ACADEMIA COLOMBIANA de Ciencias Exactas, Físicas y Naturales

LA ACADEMIA ES ÓRGANO CONSULTIVO DEL GOBIERNO NACIONAL

VOLUMEN XXXII

DICIEMBRE DE 2008

NÚMERO 125

DIRECTOR DE LA REVISTA: PEDRO PRIETO



## CONTENIDO - CONTENTS

	Pág.		Pág.
<b>Botánica (Botany)</b>		<b>Historia y Filosofía de la Ciencia</b> (History and Philosophy of Science)	
Macrofauna en suelos de Bosque y Pajonal de la Reserva Natural Pueblo Viejo, Nariño, Colombia [Macrofauna in soils of forest and grassland of the Pueblo Viejo Nature Reserve, Nariño, Colombia] <i>Patricia Cerón, Silvia Montenegro &amp; Elkin Noguera</i> .....	447	Hacia una Ontología Relacional del Espacio-Tiempo [Towards a Relational Spacetime Ontology] <i>Favio Ernesto Cala Vitery</i> .....	527
Novedades en los géneros <i>Espeletia</i> Mutis Ex Humb. & Bonpl. y <i>Espeletiopsis</i> cuatrec. (Asteraceae, Heliantheae, Espeletiinae) [Gender novelties in <i>Espeletia</i> Mutis Ex Humb. & Bonpl. y <i>Espeletiopsis</i> cuatrec. (Asteraceae, Heliantheae, Espeletiinae)] <i>Santiago Díaz Piedrahita &amp; Betsy Viviana Rodríguez</i> .....	455	<b>Química (Chemistry)</b>	
Frecuencia de las sigatokas negra y amarilla en plátano ( <i>Musa</i> spp.) de acuerdo a los estados de desarrollo de la planta [Frequency of black and yellow sigatokas on plantain ( <i>Musa</i> spp.) according to the development stages of the plant] <i>Mayary Torrado-Jaime &amp; Jairo Castaño-Zapata</i> .....	465	On the behaviour of the Ca-SO <sub>4</sub> -H <sub>2</sub> O system [Sobre el comportamiento del sistema Ca-SO <sub>4</sub> -H <sub>2</sub> O] <i>Iván Berdugo, Enrique Romero, Maarten Saaltink &amp; María Albis</i> .....	545
<b>Ciencias de la Tierra (Earth Sciences)</b>		Esquemas de reacción para licuefacción y gasificación hidrotérmica de macromoléculas orgánicas. Una revisión. [Reaction pathways for both hydrothermal liquefaction and gasification of organic macromolecules: A review] <i>Alejandro Amadeus Castro V., Luis Ignacio Rodríguez V., José de Jesús Díaz V.</i> .....	559
Geotecnia en suelos no saturados [Unsaturated soils in geotechnical] <i>Miguel Angel Alfaro Soto</i> .....	471	<b>Zoología (Zoology)</b>	
Estimación del espectro multifractal para series de precipitación horaria en los Andes Tropicales de Colombia [Multifractal spectrum estimatives for hourly rainfall at Colombia tropical Colombian Andes] <i>Jesús D. Gómez &amp; Germán Poveda</i> .....	483	Histología comparada del sistema reproductivo de cuatro especies de rayas caribeñas (Chondrichthyes: Batoidea) [Comparative histology of the reproductive system of four caribbean ray species (Chondrichthyes: Batoidea)] <i>Arturo Acero P., Marcela Grijalba-Bendeck, Fabián Moreno, Kelly Acevedo &amp; Esperanza González</i> ....	571
<b>Física (Physics)</b>		<b>Índice General del Volumen XXXII, año 2008</b> (General Index of volume XXXII, year 2008).....	587
Non-gaussianity and loop corrections in a quadratic two-field slow-roll model of inflation. Part I [No gaussianidad y correcciones de lazo en un modelo inflacionario de rodadura lenta con potencial escalar cuadrático de dos componentes. Parte I] <i>Yeinzon Rodríguez</i> .....	503	<b>Índice de autores</b> (Index of Authors).....	589
Non-gaussianity and loop corrections in a quadratic two-field slow-roll model of inflation. Part II [No gaussianidad y correcciones de lazo en un modelo inflacionario de rodadura lenta con potencial escalar cuadrático de dos componentes. Parte II] <i>Heiner R. S. Cogollo, Yeinzon Rodríguez, &amp; César A. Valenzuela-Toledo</i> .....	515	<b>Constitución de la Academia</b> (Members of the Academy).....	591
		<b>Publicaciones de la Academia</b> (Publications of the Academy).....	593

# ACADEMIA COLOMBIANA DE CIENCIAS EXACTAS, FÍSICAS Y NATURALES

APARTADO 44763 - BOGOTÁ, D.C., COLOMBIA  
SEDE DE LA ACADEMIA: CRA. 28A No. 39A-63  
Líneas de teléfono y FAX No. (571) 2443186, 2683290, 2682846  
*Correo electrónico: revista@accefyn.org.co*

DIRECTOR DE PUBLICACIONES: PEDRO PRIETO

## COMITÉ EDITORIAL Período 2006-2008

VÍCTOR SAMUEL ALBIS GONZÁLEZ  
DIÓGENES CAMPOS ROMERO  
JAIME CANTERA KINTZ  
ÁLVARO CAMPO CABAL

JAIRO MOJICA CORZO  
SANTIAGO DÍAZ PIEDRAHITA  
ARMANDO ESPINOSA

GONZALO CORREAL URREGO  
MICHEL HERMELÍN  
ELENA STASHENKO

## COMITÉ CIENTÍFICO INTERNACIONAL

JAIME BERMEJO BARRERA  
PEDRO JOSEPH-NATHAN  
SANTIAGO GAVIRIA-MELO  
ABRAHAM KRIKORIAN

JAIME DE PORTA VERNET  
JOHN J. PIPOLY  
GASTÓN GUZMÁN HUERTA  
JORGE SAHADE

JOHN HORVATH  
REINHARD SCHNETTER  
HELMUT STURN  
HANS WEBER



## EMBLEMA DE LA ACADEMIA MATRIZ ESPAÑOLA

Esta revista es reseñada en: Historia Mathematica, Math Sci (on line data base),  
Current Mathematical Publications, Index Kewensis,  
Mathematical Reviews, Zoological Record.  
PERIÓDICA (Índice de Revistas Latinoamericanas en Ciencias)  
Actualidad Iberoamericana (Índice Internacional de revistas en idioma castellano)

## ARTÍCULO 32 DE LOS ESTATUTOS

**Del Director de publicaciones.** Al Director de publicaciones le corresponde la selección de los trabajos que se publicarán en la Revista y en las diferentes colecciones de libros y otras publicaciones de la Academia. Estará asesorado por un comité editorial designado por la Junta Directiva para un período de dos (2) años.

La responsabilidad de las ideas emitidas en la Revista corresponde a sus autores.

No se devuelve la colaboración espontánea ni se mantiene correspondencia sobre ella.

*República de Colombia*  
**MINISTERIO DE EDUCACIÓN NACIONAL**



Libertad y Orden

Esta Publicación se ha financiado mediante la transferencia de recursos del Gobierno Nacional a la Academia Colombiana de Ciencias Exactas, Físicas y Naturales. El Ministerio de Educación Nacional no es responsable de las opiniones aquí expresadas.

Armada digital, impresión y acabados:

EDITORA GUADALUPE S. A. - Bogotá, D.C. - Colombia.

Cra. 42B No. 10A-57 Tels.: 2690788 - 5627250. Correo electrónico: ediguada@yahoo.es

# MACROFAUNA EN SUELOS DE BOSQUE Y PAJONAL DE LA RESERVA NATURAL PUEBLO VIEJO, NARIÑO, COLOMBIA

Por

Patricia Cerón<sup>1</sup>, Silvia Montenegro<sup>2</sup> & Elkin Noguera<sup>2</sup>

## Resumen

**Cerón P., S. Montenegro & E. Noguera:** Macrofauna en suelos de Bosque y Pajonal de la reserva natural Pueblo Viejo, Nariño, Colombia. *Rev. Acad. Colomb. Cienc.* **32**(125): 447-453, 2008. ISSN 0370-3908.

Se compara la densidad y biomasa de la macrofauna de suelos con cobertura en bosque en ladera (Bosque 1), bosque en vertiente del cauce de quebrada (Bosque 2) y Pajonal, a similar altitud (3300-3350 m). La densidad es más alta en Bosque 2, seguida de Bosque 1 y Pajonal, con un promedio de 1.301, 581 y 293 ind.\*m<sup>2</sup>, respectivamente; se expresan diferencias estadísticas significativas entre Bosque 2 y Pajonal en el caso de densidad de Gasteropoda y Diplopoda. La biomasa es más alta en Pajonal (51%) donde Oligochaeta aporta el 68% del peso.

**Palabras clave:** edafofauna, páramo, Bosque Altoandino, Nariño.

## Abstract

The density and biomass of the edaphofauna are compared between ladder woodlands (Forest 1), woodlands in the drainage area of a water source (Forest 2) and grassland, which are located at the same rank of heights (3.300-3.350 meters). The density is higher in Forest 2 than in Forest1 and the lowest is in Grassland soil, with medias of 1301, 581 and 293 ind/m<sup>2</sup>. In the case of density of Gasteropoda and Diplopoda, there are significant statistical differences among Forest 2 and Grassland. The biomass is higher in grassland soil (51%) where Oligochaeta are a 68% of the total weight.

**Key words:** macrofauna, páramo, Highland Andes Forest, Nariño.

1 Docente Universidad de Nariño, Calle 23 N. 4 este 11 manzana K, casa 2, Campiña de oriente, Pasto, Nariño; Correo electrónico: patriciacl@hotmail.com.

2 Estudiantes de Biología, Universidad de Nariño. Correos electrónicos: elkalexno@yahoo.com, silvimomu@gmail.com, A.A. 1174

## Introducción

Los suelos de alta montaña albergan una fauna que, al parecer, contiene particularidades en función de que se halle en páramo o bosque alto andino, e incluso dentro de cada uno de estos ecosistemas se presentan variaciones dependiendo de las condiciones e interrelaciones edáficas, climáticas, geográficas y antrópicas.

Al comparar la edafofauna entre páramo y bosque altoandino de los Andes colombianos se encuentran algunas diferencias. El número total de individuos y de grupos taxonómicos suele ser más alto en el bosque en relación con el páramo. No obstante, se hace necesario observar con detalle el nivel de comparación, pues se ha encontrado que la abundancia de Oribatei (Acari), es mayor en el páramo (68%) que en el bosque (32%), mientras que en el caso de Enchytraeidae son similares (Rangel-Ch. y Sturm, 1994). El páramo presenta un índice de riqueza en los principales taxa edáficos de 0,72, más alto que bosque alto andino con 0,49 (Chamorro y Zuluaga, 1998). Tratándose de coleópteros, el índice de diversidad es mayor en bosque altoandino (0,229) que en páramo (0,193), mientras que el índice de riqueza es más alto en páramo (0,333) que en bosque (0,271), (Camero y Chamorro, 1997).

En lo concerniente a la variación interna en los ecosistemas de alta montaña, los estudios evidencian que las características del suelo, como por ejemplo el contenido de materia orgánica y la estructura, conllevan una heterogeneidad vertical de la densidad tanto de un mismo como de diferentes grupos taxonómicos. Cambios en las condiciones climáticas causan alteraciones en la fauna, tal es el caso de las quemadas y la cantidad de precipitación que hace diferir la densidad, biomasa y diversidad de la población. En cuanto a la geografía, se ha encontrado que la altitud incide en la abundancia de los organismos; en el caso de la lombriz de tierra en el páramo, la población disminuye a medida que aumenta la altitud sobre el nivel del mar (Chamorro, 1989, García y Chamorro, 1995, Unigarro *et al.*, 2005). Las perturbaciones en las condiciones naturales debidas a la actividad antrópica, como el uso y manejo del suelo para actividades agropecuarias, provoca cambios en la biomasa, densidad, diversidad y riqueza de morfotipos (Casasbuenas y Estupiñán, 2007; Chamorro 1989; Coral y Bonilla, 2002; Morales y Sarmiento, 2002; Unigarro *et al.*, 2005).

Estudios comparativos entre la macrofauna de páramo y bosque altoandino de áreas silvestres no son frecuentes. Las investigaciones previas permiten prever que existen particularidades en la edafofauna, lo que hasta ahora se ha podido explicitar a partir de la abundancia de algu-

nos individuos y grupos taxonómicos, así como de los índices de diversidad y riqueza. En este contexto, se realizó un estudio exploratorio con el fin de comparar la densidad y biomasa de la macrofauna del pajonal y bosque altoandino en ladera y vertiente del cauce de agua, situados a altitud similar, en la reserva natural Pueblo Viejo.

## Metodología

La Reserva Natural Pueblo Viejo se localiza en el suroeste de Colombia, en el Departamento de Nariño, municipio de Mallama, en la vereda Pueblo Viejo. Las coordenadas geográficas extremas son 1°3'33,7'' y 1°1'39,8'' Norte y 77°48'3,3'' y 77°46'17,2'' Oeste. Tiene una extensión de 5,8 km<sup>2</sup>, que se distribuye entre los 2.950 y 3.800 msnm, (Navas, *et al.* 2008).

El Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM, 2007) en la estación El Paraíso, municipio de Túquerres, registra una precipitación total anual de 961 mm, con un régimen de lluvias bimodal, siendo junio - septiembre los meses más secos y octubre-diciembre los más lluviosos; la temperatura promedio es 10,9°C, presentando el valor mínimo (9,3°C) en el mes de agosto y el máximo (12,3°C) en abril. De acuerdo con el Instituto Geográfico Agustín Codazzi (IGAC, 2004), los suelos se clasifican como Andosoles, hacen parte de la consoación Acrudoxic Melanudands que ocupa la posición de coladas de lava en alturas comprendidas entre los 3000 y 3700 msnm. Los suelos se han desarrollado a partir de cenizas volcánicas de espesor variable que recubren rocas volcánicas andesitas producto de erupciones volcánicas.

Pajonal hace parte del páramo con vegetación herbácea formadora de macollas, dominada por la familia Poaceae, especialmente del género *Calamagrostis* asociada a otros géneros como *Rhynchospora* y *Castilleja*. Comprende algunas especies como *Hypericum junipericum* Kunth, *Hypericum laricifolium* Juss, *Lachemilla mutisi* Roth, *Hypoacheris sessiliflora* Kunth, *Cortaderia nitida* (Kung) Pilg., *Monnina aestuans* (LF) DC., *Bidens andicola* H.B.K., *Rhynchospora macrochaeta* L. y, *Pernettya postrata* (Cav). D.C. El bosque contiene arbustos y arbolitos, entre los que se encuentran *Gynoxis sanctii-antonii* Cuatr., *Lepichinia vulcanicola* Word, *Zanthoxylum quinduensi* (Tul), *Baccharis buddlejoides* H.B.K., *Hesperomeles* sp, *Oreopanax bogotensis* Cuat., *Euphorbia laurifolia* Lam., *Morella parvifolia* Benth.

Las muestras de suelo y macrofauna se tomaron en el mes de agosto. Se establecieron tres sitios: Bosque 1, Bosque 2 y Pajonal, ubicados entre los 3300 y 3350 m de altitud; de cada sitio se tomaron tres réplicas separadas

por, al menos, cinco metros, para un total de nueve muestras. Bosque 1 está ubicado en ladera y Bosque 2 en la vertiente del cauce de la cuenca alta de la quebrada Honda, que se encontraba sin agua en el momento. Para caracterizar la macrofauna se extrajeron tres monolitos de 25 cm. x 25 cm. x 30 cm. de volumen de suelo, para cada uno de los sitios seleccionados (Bosque 1, Bosque 2 y Pajonal). A cada monolito se le trazó una zanja a su alrededor, y posteriormente se tomaron submuestras: hojarasca, 0-10 cm., 10-20 cm. y 20-30 cm. de profundidad, que fueron colocados en bolsas etiquetadas. Las muestras fueron llevadas al Laboratorio de Entomología de la Universidad de Nariño donde se procedió a extraer manualmente los organismos. Éstos se seleccionaron, se preservaron en alcohol al 70%, se identificaron hasta orden y se midieron su densidad y su biomasa. Las muestras de suelo se tomaron alrededor de los monolitos a una profundidad de 0-20 cm. Los datos se trataron con pruebas no paramétricas en el programa Statistica, el análisis de varianza se realizó con el estadístico Kruskal-Wallis mediante la significación exacta por tratarse de muestras pequeñas, después se estimó una comparación múltiple. También se procesó una correlación de Pearson con todas las variables y un Análisis de Componentes Principales (ACP) en el programa SAS a partir de las siguientes variables: arena, densidad aparente, densidad real, pH, C, CIC, Al, P, Ca, Mg, K, Mn, Cu, Zn, Fe, biomasa de macrofauna, densidad de artrópodos, de gasterópodos y de oligoquetos.

## Resultados

Tomando como referencia los niveles generales de interpretación de análisis de suelos para cultivos, aportados por el Instituto Colombiano Agropecuario (ICA, s.f.), se puede anotar que en Pajonal el suelo es muy fuertemente ácido (4,8), presenta muy alta CIC (46 Cmol\*Kg), alta proporción de carbono orgánico (12%), nitrógeno total (0,68%) y aluminio (2 Cmol\*Kg), nivel ideal de potasio (0,38 Cmol\*Kg); medio de hierro (35 Mg\*kg); contenido bajo de fósforo (12 Mg\*kg), calcio (2,7 Cmol\*Kg) y magnesio (0,33 Cmol\*Kg), déficit de manganeso (1,17 Mg\*kg), cobre (0,33 Mg\*kg), zinc (0,47 Mg\*kg) y boro (0,13 Mg\*kg). La textura es Arenosa franca con una proporción de arena del 75%, porosidad de 69%, densidad real y aparente de 2,2, y 0,7 Mg\*m<sup>3</sup>, respectivamente.

Bosque 1 muestra un suelo fuertemente ácido (5,3), con alta proporción de CIC (29 Cmol\*Kg), contenido ideal de C (7%), Al (0,9 Cmol\*Kg) y K (0,37 Cmol\*Kg); nivel medio de N (0,48%), Ca (3,9 Cmol\*Kg), Fe (30,7 Mg\*kg); bajo en P (11 Mg\*kg) y Mg (0,6 Cmol\*Kg), déficit de Mn (4,4 Mg\*kg),

Cu (0,19 Mg\*kg), Zn (0,38 Mg\*kg) y B (0,08 Mg\*kg). La textura es arenosa franca, con promedio de arena del 73%, porosidad del 66%, densidad real y aparente de 2,4 y 0,8 Mg\*m<sup>3</sup>, respectivamente. Bosque 2 se caracteriza porque el suelo es fuertemente ácido (5,2), con nivel ideal de P (34 Mg\*kg), Al (0,3 Cmol\*Kg) y CIC (11 Cmol\*Kg), proporción media de C (3%) y N (0,25%); contenido bajo de Ca (2,07 Cmol\*Kg), Mg (0,27 Cmol\*Kg) y K (0,09 Cmol\*Kg); deficiente en Fe (19,3 Mg\*kg); Mn (1,8 Mg\*kg), Cu (0,2 Mg\*kg), Zn (0,2 Mg\*kg) y B (0,1 Mg\*kg); textura arenosa, con promedio de arena de 89%, porosidad de 55%, densidad real 2,5 y aparente de 1,1 Mg\*m<sup>3</sup>.

La macrofauna encontrada pertenece a tres phylla: Annelida, Arthropoda y Mollusca. Annelida está constituida por la clase Oligochaeta, Mollusca por la clase Gastropoda y Arthropoda está conformado por Chilopoda, Diplopoda, Arachnida, Malacostraca e Insecta. Respecto a la clase Insecta, en los tres sitios sobresalen los coleópteros, tanto en biomasa (93%) como en densidad (83%); entre las familias se encontraron Curculionidae, Elateridae, Silphidae, Staphylinidae, Scarabaeidae y Tenebrionidae.

La densidad total de la edafofauna presenta un promedio de 2176 ind.\*m<sup>2</sup>. Es más alta en Bosque 2, seguida de Bosque 1 y más baja en Pajonal, tendencia contraria a los valores de C, N, CIC y Al, mayores en Pajonal, intermedios en Bosque 1 y menores en Bosque 2, con diferencia estadística significativa entre Pajonal y Bosque 2. Gráfica 1. La densidad de artrópodos sigue la misma tendencia que el total de la macrofauna, con mayor promedio en Bosque 2, continúa en Bosque 1 y termina en Pajonal, aunque sólo se presenta diferencia estadística significativa en diplópodos entre Bosque 2 y Pajonal. La densidad de oligoquetos fue mayor en Pajonal, seguida de Bosque 2 y disminuye en Bosque 1; es de mencionarse que entre Pajonal y Bosque 1 se presentan diferencias estadísticas significativas entre Cu y Mn, el primero con mayor contenido en Pajonal y el segundo con menor contenido en dicho sitio. Los moluscos se encontraron principalmente en Bosque 2 y no se registraron en Pajonal, con diferencias significativas entre estos dos sitios. Tabla 1.

La biomasa presenta un promedio total de 35,4 (g.p.f.\*m<sup>2</sup>), y está siendo aportada principalmente por coleópteros (39%) y oligoquetos (38%). De hecho, la biomasa total se correlaciona con la biomasa de coleópteros (r: 0,88 p: 0,002) y la biomasa (r: 0,90 p: 0,001) y densidad de oligoquetos (r: 0,73 p: 0,026). Teniendo en cuenta los sitios, es más alta en Pajonal (51%), donde los valores de C, N, CIC, Al son más elevados y la densidad aparente es más baja, con diferencias estadísticas significativas de Bosque 2. Gráfica 2. En Pajonal, la lombriz de tierra aporta

**Tabla 1.** Densidad de macrofauna (individuos\*m<sup>2</sup>)

Clase	Orden	Bosque 1		Bosque 2		Pajonal		K-W
		No.	%	No.	%	No.	%	
Gastropoda	Pulmonada	27ab	5	491 <sup>a</sup>	38	0b	0	0,011
Oligochaeta		5 <sup>a</sup>	1	64 <sup>a</sup>	5	123 <sup>a</sup>	42	0,111
Arachnida	Araneae	16 <sup>a</sup>	3	11 <sup>a</sup>	1	5 <sup>a</sup>	2	0,357
Chilopoda	Geophilomorpha	91 <sup>a</sup>	16	85 <sup>a</sup>	7	0 <sup>a</sup>	0	0,046
Diplopoda	Polydesmida	101ab	17	363 <sup>a</sup>	28	0b	0	0,007
Malacostraca	Isopoda	5 <sup>a</sup>	1	0 <sup>a</sup>	0	0 <sup>a</sup>	0	1,000
Insecta	Coleoptera	240 <sup>a</sup>	41	213 <sup>a</sup>	16	112 <sup>a</sup>	38	0,311
	Dermaptera	11 <sup>a</sup>	2	0a	0	0 <sup>a</sup>	0	1,000
	Diptera	5 <sup>a</sup>	1	37 <sup>a</sup>	3	11 <sup>a</sup>	4	0,679
	Hemiptera	0 <sup>a</sup>	0	5 <sup>a</sup>	0,4	16 <sup>a</sup>	5	1,000
	Lepidoptera	21 <sup>a</sup>	4	5 <sup>a</sup>	0,4	0 <sup>a</sup>	0	0,464
Formas inmaduras sin identificar		59 <sup>a</sup>	10	27 <sup>a</sup>	2	27 <sup>a</sup>	9	0,725
Promedio total de densidad		581 <sup>a</sup>	27	1301 <sup>a</sup>	60	293 <sup>a</sup>	13	0,039

K-W: Kruskal-Wallis, significación exacta. En letras iguales no hay diferencia significativa en la comparación entre sitios ( $\alpha \leq 0.05$ )

el 68% de la biomasa. Específicamente, los mayores valores de biomasa para Bosque 1 se registraron en coleópteros (57%), formas inmaduras (17%) y oligoquetos (8%); para Bosque 2 en coleópteros (34%), diplópodos (24%) y gasterópodos (20%), mientras que en Pajonal, fue en lombriz de tierra (68%) y coleópteros (31%). No se presentaron diferencias estadísticas significativas. Tabla 2.

Considerando la distribución vertical de la edafofauna, cuantificando también las formas inmaduras que no se lograron identificar, se encontró que en los bosques, los anélidos estaban a 0-10 cm., a diferencia del Pajonal, donde el 87% se localizó entre 20 y 30 cm. En los tres sitios, los

artrópodos, se encontraron principalmente en hojarasca, seguido de los primeros 10 cm., aunque en Pajonal, un 20% profundiza a 10-20 cm. Los moluscos se albergan en la hojarasca de los bosques y un 20% llegó a los primeros 10 cm. en Bosque 1. Hay tendencia a que en los bosques la edafofauna habite especialmente la hojarasca y los primeros 10 cm. mientras que en Pajonal llegan a mayor profundidad. Tabla 3.

El análisis de componentes principales muestra que los dos primeros factores explican el 76% de la variabilidad. En el primer factor, con el 50%, las variables que más contribuyen en forma positiva son: CIC (0,98), C (0,97), Al (0,95)

**Tabla 2.** Biomasa de macrofauna (g.p.f.\*m<sup>2</sup>)

Clase	Orden	Bosque 1		Bosque 2		Pajonal		K-W
		No.	%	No.	%	No.	%	
Gastropoda	Pulmonada	0,02 <sup>a</sup>	0,3	1,62 <sup>a</sup>	20	0,0 <sup>a</sup>	0,0	0.100
Oligochaeta		0,74 <sup>a</sup>	8,0	0,51 <sup>a</sup>	6,2	12,2 <sup>a</sup>	68	0,336
Arachnida	Araneae	0,03 <sup>a</sup>	0,4	0,08 <sup>a</sup>	0,9	0,01 <sup>a</sup>	0,0	0,379
Chilopoda	Geophilomorpha	0,55 <sup>a</sup>	5,9	0,63 <sup>a</sup>	7,7	0,0 <sup>a</sup>	0,0	0,061
Diplopoda	Polydesmida	0,46 <sup>a</sup>	5,0	2,01 <sup>a</sup>	24	0,0 <sup>a</sup>	0,0	0,025
Malacostraca	Isopoda	0,05 <sup>a</sup>	0,6	0,00 <sup>a</sup>	0,0	0,0 <sup>a</sup>	0,0	1,000
Insecta	Coleoptera	5,35 <sup>a</sup>	57	2,80 <sup>a</sup>	34	5,58 <sup>a</sup>	31	0,511
	Dermaptera	0,07 <sup>a</sup>	0,7	0,00 <sup>a</sup>	0,0	0,0 <sup>a</sup>	0,0	1,000
	Diptera	0,10 <sup>a</sup>	1,1	0,39 <sup>a</sup>	4,8	0,02 <sup>a</sup>	0,1	1,000
	Hemiptera	0,00 <sup>a</sup>	0,0	0,01 <sup>a</sup>	0,1	0,05 <sup>a</sup>	0,3	1,000
	Lepidoptera	0,36 <sup>a</sup>	3,8	0,01 <sup>a</sup>	0,1	0,0 <sup>a</sup>	0,0	0,250
Formas inmaduras		1,56 <sup>a</sup>	17	0,16 <sup>a</sup>	1,9	0,04 <sup>a</sup>	0,2	0,064
Promedio total de biomasa		9,3 <sup>a</sup>	26	8,2 <sup>a</sup>	23	17,9 <sup>a</sup>	51	0,829

K-W: Kruskal-Wallis, significación exacta. En letras iguales no hay diferencia significativa en la comparación entre sitios ( $\alpha \leq 0.05$ )

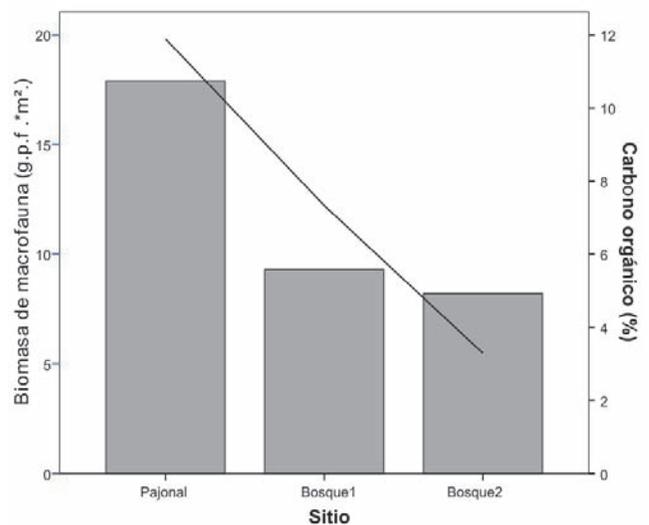
**Tabla 3.** Promedio de macrofauna por profundidad.

Phylum	Annelida			Arthropoda			Mollusca		Formas inmaduras			Total
	B1	B2	P	B1	B2	P	B1	B2	B1	B2	P	
Hojarasca	0	0	0	309	432	91	21	485			0	1397
%				(55)	(60)	(55)	(80)	(99)	37 (64)	21 (80)		(62)
0-10 Cm.	5	64	5	197	235		5	5			0	
%	(100)	(100)	(4)	(35)	(32)	32 (19)	(20)	(1)	11 (18)	5 (20)		565 (25)
10-20 cm.	0	0	11	32	43		0	0	5	0		
%			(9)	(6)	(6)	37 (23)			(9)		16 (60)	144 (6)
20-30 cm.	0	0	107	21	16	5	0	0	5	0		
%			(87)	(4)	(2)	(3)			(9)		11 (40)	165 (7)

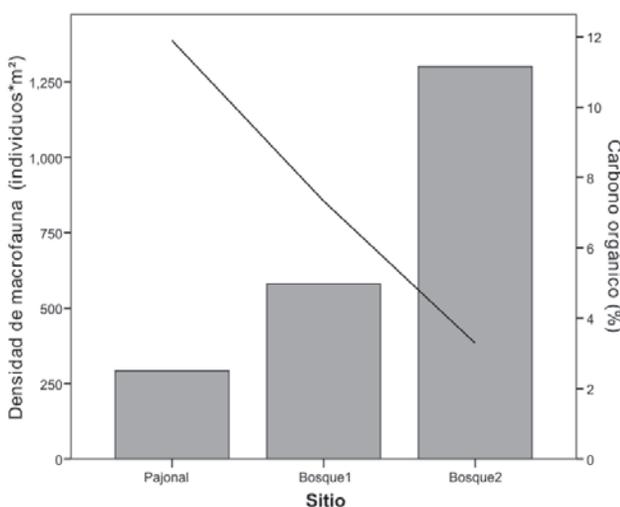
B1: Bosque 1, B2: Bosque 2, P: Pajonal.

y Fe (0,84) y en forma negativa son densidad aparente (-0,94), densidad real (-0,85), P (-8,45) y densidad de artrópodos (-0,77), valores asociados a las propiedades de Pajonal, en oposición a Bosque 2. El segundo factor que explica el 26% de la variabilidad, se relaciona con las propiedades de Bosque 1, manganeso (0,95), magnesio (0,85), calcio (0,83) y valores negativos en densidad de oligoquetos (-0,83).

La correlación de Pearson evidencia, por una parte, la asociación entre la biomasa de la macrofauna con Pajonal. De hecho, como ya se mencionó, la biomasa total se correlaciona con densidad y biomasa de oligoquetos, siendo que en Pajonal se registró el mayor número y peso de lombriz de tierra. A la vez, la biomasa de este anélido se correlaciona positivamente con CIC ( $r: -0,70$   $p: 0,036$ ) y Al ( $r: 0,83$   $p: 0,006$ ) y, negativamente con pH ( $r: -0,75$   $p: 0,019$ ) y densidad real ( $r: -0,79$   $p: 0,012$ ) propiedades cercanas a Pajonal. Por otra parte, la relación entre la densidad total



**Figura 2.** Biomasa de macrofauna en relación con carbono orgánico del suelo.



**Figura 1.** Densidad de macrofauna en relación con carbono orgánico del suelo.

de la edafofauna con los bosques, especialmente con Bosque 2, de manera que se correlaciona con densidad de gasterópodos ( $r: 0,90$   $p: 0,001$ ) y de artrópodos ( $r: 0,88$   $p: 0,002$ ), más abundantes en Bosque 2 seguido de Bosque 1.

**Discusión**

La macrofauna de Pajonal está constituida por insectos, oligoquetos y algunos arácnidos, lo que concuerda con estudios en páramos en los que se ha encontrado que Oligochaeta (Enchitreidae, Lumbricidae), Insecta (Colembolla, Coleoptera, Díptera) y Arachnida (Acarina) aparecen entre los taxa dominantes (Chamorro, 1989), aunque en el páramo también suelen aparecer miriápodos en la composición, como lo registran Casasbuenas y Estupiñán (2007) y Unigarro *et al.* (2005) en páramos de Colombia y, Morales y Sarmiento (2002) en zona de páramo no perturbado de los Andes de Venezuela, donde Chilopoda repre-

sentaba el 12% de los macroinvertebrados. Bosque 1 contiene un predominio de artrópodos (84%), similar a lo encontrado en el bosque alto andino de Colombia, donde se halló que el mayor porcentaje de individuos lo conforman ácaros e insectos, mientras que los no artrópodos tuvieron menor abundancia (**García y Chamorro**, 1995). En el caso de Bosque 2, aparece un 38% de caracoles, con diferencia estadística significativa de Pajonal, posiblemente influenciado por las condiciones de humedad, al estar ubicado en la vertiente de la fuente de agua. Es de anotarse que tanto en pajonal como en los bosques de Pueblo Viejo, comparados con otros lugares, difiere la abundancia de órdenes y familias dentro de las clases taxonómicas, lo cual también puede deberse a diferencias en las metodologías en la valoración de los organismos.

En lo concerniente con la densidad, se registró un mayor número de individuos en Bosque 2, que decrece en Bosque 1 y termina en Pajonal aunque sólo se expresan diferencias estadísticamente significativas en diplópodos y gasterópodos entre Bosque 2 y Pajonal. La tendencia a un mayor número, tanto en individuos como en grupos, en bosque con respecto al páramo, fue planteada por **Rangel-Ch. y Sturm** (1994), quienes proponen que dicha diferencia se debe a las condiciones climáticas, que son más extremas en el páramo, y a la ausencia en éste de una capa de hojarasca. También habría que tenerse en cuenta las propiedades edáficas particulares, lo que se evidencia con la tendencia inversa entre los valores de C, N, CIC y Al y la densidad de macrofauna, lo que puede indicar que en Bosque 2 hay más presencia de macrofauna por la mayor descomposición de la materia orgánica. Bosque 2 tiende a mayores promedios en densidad total, de diplópodos y de gasterópodos frente a Bosque 1, aunque no se registran diferencias estadísticas significativas. Probablemente se debe a la ubicación en la vertiente del cauce de agua del primero y en ladera del segundo. En consecuencia, se sugiere que la topografía puede ser otra variable importante en la estimación de macrofauna de los ecosistemas de alta montaña. De hecho, **Feijoo et al.** (1999) encontraron valores diferentes en biomasa y densidad de lombriz de tierra si la muestra es tomada en la zona baja, media o alta de la colina.

En biomasa, es de resaltarse la asociación con el Pajonal, donde a la vez, se incrementan los valores de C, N, CIC y disminuye la densidad aparente, lo que puede indicar mayor acumulación de materia orgánica. En dicho sitio, la lombriz de tierra aparece con más densidad y aporta buena parte de la proporción de biomasa. Este organismo suele abundar y tener bastante peso en suelos con vegetación herbácea que incluso puede ser similar o superar a la cobertura en bosque. **Coral y Bonilla** (2002) encontraron mayor densidad de

lombriz de tierra en un pastizal (4560 ind\*m<sup>2</sup>) que aportó más biomasa (67,1 g.p.f.\*m<sup>2</sup>) que en selva secundaria (1024 ind\*m<sup>2</sup>) con biomasa de (18,2 g.p.f.\*m<sup>2</sup>); por su parte, **Feijoo et al.** (1999) cuantificaron mayor densidad de lombriz de tierra en Kikuyo (22963 ind\*m<sup>2</sup>), en relación con la selva de más de 40 años (9440 ind\*m<sup>2</sup>), aunque la biomasa fue más alta en la selva que en el pastizal.

Para terminar, se sugiere que la macrofauna de los páramos y los bosques altoandinos, a similar altitud, presenta diferencias. En Pueblo Viejo, una mayor densidad de la macrofauna se asocia con la cobertura en bosques, especialmente con el bosque ubicado en la vertiente de la quebrada, aunque sólo se presentan diferencias estadísticas significativas en densidad de gasterópodos y diplópodos entre Bosque 2 y Pajonal. Mientras que la biomasa se relaciona con Pajonal, en buena parte debido a la lombriz de tierra.

### Agradecimientos

A la Vicerrectoría de Investigaciones de la Universidad de Nariño por el apoyo financiero al proyecto "Caracterización físico-biótica y antrópica de la Reserva Natural Pueblo Viejo"; a Benhur Cerón y Luis Navas, compañeros de investigación, a Mauricio Rodríguez y Diego Martínez por su apoyo con la identificación taxonómica.

### Bibliografía

- Camero E., Chamorro C.** 1997. Bioedafología del orden Coleoptera en tres regiones naturales de Colombia. *Suelos Ecuatoriales* 27: 228-230.
- Casasbuenas L., Estupiñán L.** 2007. Efecto del cultivo de papa sobre la fauna edáfica en el páramo de guerrero. *Revista U.D.C.A Actualidad & Divulgación Científica* 10(2): 31-42.
- Chamorro C., Zuluaga, D.** 1998. Comunidades bioedáficas de tres regiones naturales de Colombia. *Suelos Ecuatoriales* 28: 269-272.
- Chamorro, C.** 1989. Efecto del uso del suelo sobre la composición edafofaunística de los páramos que circundan la ciudad de Bogotá. *Suelos Ecuatoriales*, 19(1): 48-62.
- Coral, D., Bonilla, C.** 2002. Impacto de las prácticas agrícolas sobre la macrofauna del suelo en la cuenca alta del Lago Guamués, Pasto, Colombia. *Suelos Ecuatoriales* 32: 157-160.
- Feijoo, A., Ochoa, W., Amezcuita, E., Knapp, E.** 1999. Relación entre el ambiente físico del suelo y la fluctuación de lombrices de tierra en áreas de laderas andinas, Colombia. *Suelos Ecuatoriales* 29(1): 76-82.
- García M., Chamorro C.** 1995. Contribución al Conocimiento de la Dinámica Temporal de la Edafofauna en un Bosque Altoandino de la región de Monserrate. En: Mora-Osejo y Sturm (editores). *Estudios Ecológicos del Páramo y del Bos-*

- que Altoandino, Cordillera Oriental de Colombia. Tomo 2. Academia Colombiana de Ciencias Exactas, Físicas y Naturales. Bogotá, 619-630.
- Instituto Colombiano Agropecuario (ICA), s.f. Fertilización en diversos cultivos: quinta aproximación, manual de asistencia técnica No. 25. ICA. Bogotá.
- Instituto de Hidrología, Meteorología y Estudios Ambientales 2007. Sistema de información, estación El Paraíso, municipio de Túquerres. Enero 1987-diciembre 2006.
- Instituto Geográfico Agustín Codazzi 2004. Estudio general de suelos y zonificación de tierras, departamento de Nariño. IGAC. Bogotá.
- Morales J, Sarmiento L.** 2002. Dinámica de los Macroinvertebrados Edáficos y su relación con la vegetación en una sucesión secundaria en el páramo Venezolano. *Ecotrópicos* **15**(1): 99-110.
- Navas, L., Cerón, B., Ceron, P.** 2008. Caracterización físico-biótica y antrópica de la Reserva Natural Pueblo Viejo, Municipio de Mallama, departamento de Nariño. *Tendencias* **9**(1): 7-46.
- Rangel-Ch. O, Sturm H.** 1994. Consideraciones sobre la vegetación, la productividad primaria neta y la artropofauna asociada en regiones paramunas de la Cordillera Oriental. En: Mora Osejo y Sturm (Editores). Estudios ecológicos del páramo y del Bosque Altoandino, Cordillera Oriental de Colombia, Tomo 1. Academia Colombiana de Ciencias Exactas, Físicas y Naturales. Bogotá, 47-70.
- Unigarro, A., Burbano, H., Sánchez, M.** 2005. Evaluación de la calidad de suelos *Dystric Cryandept* en el Santuario de Flora y Fauna Galeras, Nariño. *Acta agronómica* **54**(4): 7-12.

Recibido: septiembre 12 de 2008.

Aceptado para su publicación: diciembre 9 de 2008.

# NOVEDADES EN LOS GÉNEROS *ESPELETIA* MUTIS EX HUMB. & BONPL. Y *ESPELETIOPSIS* CUATREC. (ASTERACEAE, HELIANTHEAE, ESPELETIINAE)

Por

Santiago Díaz-Piedrahita<sup>1</sup> & Betsy Viviana Rodríguez-Cabeza<sup>2</sup>

## Resumen

**Díaz-Piedrahita, S. & B. V. Rodríguez:** Novedades en los géneros *Espeletia* Mutis Ex Humb. & Bonpl. y *Espeletiopsis* Cuatrec. (Asteraceae, Heliantheae, Espeletiinae). Rev. Acad. Colomb. Cienc. **32**(125): 455-464, 2008. ISSN 0370-3908.

Se nomina, describe e ilustra una nueva especie de *Espeletiopsis* Cuatrec. recolectada en jurisdicción de Samacá, Departamento de Boyacá y se propone un nuevo estatus para una variedad de *E. jimenes-quezadae* (Cuatrec.) Cuatrec. También se describen e ilustran dos nuevas especies de *Espeletia* Mutis ex Humb. & Bonpl. provenientes del Santuario de Flora y Fauna de Guanentá en el municipio de Encino, Departamento de Santander y se propone el estatus de especie para una subespecie de *E. brachyaxiantha* S. Díaz.

**Palabras clave:** *Espeletia*, *Espeletiopsis*, Espeletiinae, Heliantheae, Asteraceae, nuevas especies, Colombia, Botánica sistemática.

## Abstract

Two new species of *Espeletia* Mutis ex Humb. & Bonpl. and a new species of *Espeletiopsis* Cuatrec. are named, described and illustrated. These new taxa were recorded in northern Colombia, Guanentá Flora and Fauna Sanctuary Paramo of Rabanal in Samaca-Boyacá. Additionally, new status for a variety of *E. jimenes-quezadae* (Cuatrec.) Cuatrec. and a subspecies of *E. brachyaxiantha* S. Díaz are proposed.

**Key words:** *Espeletia*, *Espeletiopsis*, Espeletiinae, Heliantheae, Asteraceae, new species, Colombia, Systematic botany.

<sup>1</sup> Miembro de número Academia Colombiana de Ciencias Exactas, Físicas y Naturales, Apartado 44763, Bogotá D.C., Colombia. Correo electrónico: santiagodiaz@etb.net.co

<sup>2</sup> Territorial Norandina, Parques Nacionales Naturales de Colombia, Ave. Quebrada Seca No. 30-12, Bucaramanga, Santander, Colombia. Correo electrónico: betsyviviana@gmail.com

Los géneros *Espeletia* Mutis ex Humb. & Bonpl. y *Espeletiopsis* Cuatrec. están ampliamente representados en los páramos de Colombia. Como resultado de nuevas recolecciones en áreas mal exploradas o poco conocidas desde el punto de vista florístico, el número de especies en ambos géneros se ha incrementado en los últimos años. La última publicación al respecto **Díaz Piedrahita, Rodríguez & Galindo** (2006) dio a conocer siete nuevas especies de *Espeletia*, una variedad de *E. curialensis* Cuatrec. y una especie de *Espeletiopsis*. En la presente publicación se describen dos nuevas especies de *Espeletia*, una más de *Espeletiopsis* y se propone un nuevo estatus para una subespecie de *Espeletia brachyaxiantha* S. Díaz y para una variedad de *Espeletiopsis jimenez-quesadae* (Cuatrec.) Cuatrec. Con estas novedades el número total de especies para el primer género se eleva a 75 de las cuales 65 prosperan en los páramos de Colombia. El número de especies para *Espeletiopsis* llega a 29.

Con base en 23 especies segregadas del género *Espeletia* Mutis ex Humb. & Bonpl., **Cuatrecasas** (1976) propuso el género *Espeletiopsis* Cuatrec. el cual se reconoce principalmente por presentar inflorescencias corimbiformes compuestas y de tipo monocasial que rematan en un capítulo; las hojas, ramas, pedúnculos y brácteas de estas sinflorescencias son siempre alternas. Adicionalmente, las hojas rosulares son notablemente rígidas y presentan indumento menos denso o muy escaso en ambas caras. **Cuatrecasas** (1987) fijó el número de especies en veinte, al excluir tres de las inicialmente asignadas, las cuales fueron transferidas a otros géneros y al reducir a *Espeletiopsis guacharaca* (S. Díaz) Cuatrec. a una variedad de *E. jimenez-quesadae* (Cuatrec.) Cuatrec. **Díaz-Piedrahita & Obando** (2004) dieron a conocer una nueva especie y **Díaz-Piedrahita, Rodríguez & Galindo-Tarazona** (2006) añadieron otra. En el presente escrito se propone una más; adicionalmente y luego de estudiar abundante material, se restablece el estatus de *E. guacharaca* (S. Díaz) Cuatrec, con lo cual el número de especies conocidas queda en veintinueve, aclarando que en esta cifra están incluidas dos especies algo irregulares y que con seguridad corresponden a híbridos, por lo que en una futura revisión del género deberán ser excluidas.

**Espeletiopsis guacharaca** (S. Díaz) Cuatrec., Phytologia 35(1):55. 1976. stat. nov.

*Espeletia guacharaca* S. Díaz, Caldasia 11: 19. 1975.

*Espeletiopsis jimenez-quesadae* var. *guacharaca* (S. Díaz) Cuatrec. Anales Jard. Bot. Madrid 54: 374. 1996.

Si bien es cierto que una de las especies más afines a ella es *Espeletiopsis jimenez-quesadae* (Cuatrec.) Cuatrec.,

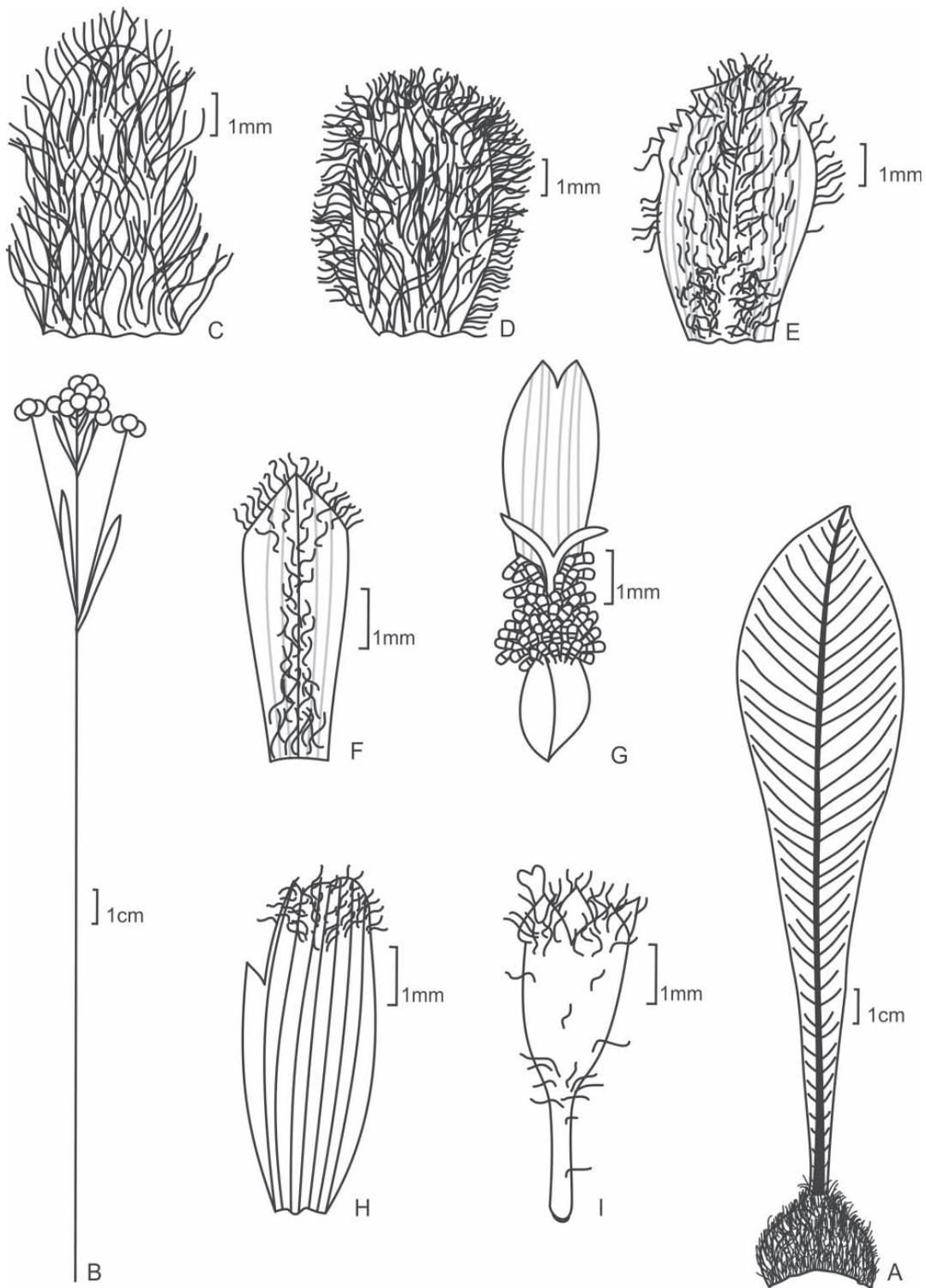
por la suma de características *Espeletiopsis guacharaca* (S. Díaz) Cuatrec. constituye una excelente especie, bien definida, con amplia distribución geográfica y con poblaciones abundantes y bastante estables, además de la localidad típica, en el páramo de la Rusia, en el páramo de la cortadera en jurisdicción de los municipios de Toca y Pesca, en el sector de las playas, Santuario de Flora y Fauna de Guanenta y en el Parque Nacional Natural de Pisba. Sus hojas, en relación con la especie a la cual fue adscrita, son más cortas y coriáceas, la lámina foliar difiere ampliamente, como difieren las sinflorescencias, que en longitud son menores que las hojas de la roseta y muestran los capítulos bastante aglomerados; las flores liguladas forman una corola corta pero bien definida que contrasta con las flores femeninas reducidas a la porción tubular de la corola. Además de las diferencias anotadas, cabe aclarar que *E. guacharaca* (S. Díaz) Cuatrec. muestra su mayor afinidad no con *E. jimenez-quesadae* (Cuatrec.) Cuatrec. sino con la especie que se describe a continuación.

**Espeletiopsis rabanalensis** S. Díaz & Rodríguez-Cabeza sp. nov. **Figs. 1 - 2a-b**

*Espeletiopsis guacharacae* (S. Díaz) Cuatrec. affinis a qua imprimis differt foliis obovatis minoribus vel minus rigidis, ramis floriferis longioribus et in partibus proximilibus inflorescentiarum aphyllis, flores foeminei ligulati vel flosculi hermaphroditi plurimi, corollis flores radiorum ligulatis pistillo longiores, lobis floribus discorum extus dense pilosibus.

**TYPUS:** COLOMBIA, BOYACÁ: Municipio de Samacá, Páramo del Rabanal. 3.412 m. de altitud, N 05°44' W 073°35'. 3-5 oct. 2007. *Betsy Viviana Rodríguez-Cabeza, Félix Márquez 1895*, Holotypus COL (528210), Isotypi COL (528211), UIS y por repartir.

“Caulirrósula de 1 m de altura, hasta 2 m. **Hojas:** Coriáceas. Verde lustroso en haz, indumento amarillo lustroso seríceo y no muy denso en envés. Venas central y secundarias con indumento denso seríceo amarillo lustroso. Las hojas más pequeñas se ubican hacia la parte externa de la roseta. **Vaina:** Haz, mitad inferior blanca, mitad superior verde limón. Envés, indumento denso seríceo café oscuro lustroso, que se va tornando dorado hacia el ápice. **Inflorescencias:** 9 por roseta. Pedúnculos, pedicelos y brácteas con indumento seríceo crema lustroso. **Capítulos:** 1 cm de ancho, liguladas amarillas. **Observaciones:** Pajonal-Frailejón con *Calamagrostis effusa*, *Halenia asclepiadea*, *Lupinus* sp. y *Valeriana* sp. Crece sobre laderas con suelo húmedo, no encharcado. En algunos sectores se observa mezclado con el frailejón BVR-1896. Abundante hacia la parte alta del páramo. Las únicas afec-



**Figura 1.** *Espeletiopsis rabanalensis* S. Díaz & Rodríguez-Cabeza sp. nov. **A.** Esquema de la rama florífera; **B.** Perfil foliar; **C-D.** Filarias intermedias vistas por el envés; **E.** Filaria interior vista por el envés; **F.** Palea de las flores femeninas; **G.** Flor femenina ligulada; **H.** Palea de los flósculos; **I.** Flósculo hermafrodita. Diseño digital de Betsy Viviana Rodríguez-C., elaborado con base en el ejemplar típico (*B.V. Rodríguez et al. 1895 COL*).

taciones observadas en la población fueron ocasionadas por la apertura de una carretera, pues aunque hubo quema hace tres años los frailejones siguieron su crecimiento. No hay evidencia de pastoreo en la zona”.

**Etimología:** El epíteto alude a la localidad de origen, páramo del Rabanal en el Municipio de Samacá, donde prospera una población con más de 600 individuos.

Tallos erectos arrosado de hasta 1 m. de altura, cubiertos por las vainas secas de las hojas marchitas. Vaina foliar de 2.2-3 cm. long. x 3.4-4.2 cm. lat., anchamente ovada, dorsalmente con indumento largo seríceo, fusco y lustroso que se va tornando aureo hacia el ápice; lámina foliar de 17-23.5 cm. long. x 3.5-6 cm. lat., coriácea, obovada, con ápice acuminado agudo, la base notablemente atenuada formando un largo pseudopetioló, haz verde lustroso con indumento seríceo amarillo lustroso, poco denso en envés, margen lisa, ligeramente revoluta, nervio medio ligeramente ensanchado hacia la base, cubierto por indumento seríceo amarillo lustroso, prominente en la haz, prominente en el envés, nervios secundarios 11-20 por decímetro, conspicuos por el envés y divergiendo en ángulo de 50-70°.

Ramas floríferas hasta 9 coetáneas, de 27-30 cm. de longitud, cubiertas con indumento seríceo espadiceo lustroso; porción proximal vegetativa de 18.5-24 cm. long., ebracteada; porción distal fértil, formada por un apretado racimo terminal con 14-23 capítulos; brácteas florales de 1.5-4 cm. long. x 0.2 - 0.4 cm. lat., lineales, con ápice agudo; pedúnculos florales 1-6 cm. long. Capítulos de 1 cm. de ancho; involucro hemisférico, filarias 10, dispuestas en 2 series, de 5 - 9 mm. long. x 2.5 - 5 mm. lat., obovadas, cara abaxial densamente cubierta por indumento largo seríceo, las más internas con una lista longitudinal de tricomas que se ensancha y se torna más densa hacia el ápice, quedando a cada lado un área glabra, cara adaxial glabra, ápice redondeado, margen ondulado y/o dentado hacia la mitad superior; paleas de las lígulas de 5-6 mm. long. x 2-3 mm. lat., angosto-obovadas, externamente con una lista longitudinal de tricomas que se ensancha y se torna más densa hacia el ápice, quedando a cada lado un área glabra, ápice agudo, margen generalmente entera, lisa, a veces sinuada y con pequeños dientecillos; paleas de los flósculos de 6-6.5 mm. long. x 1.8-2.5 mm. lat., obovadas, margen con ondulaciones y dientes en el ápice, abaxialmente con indumento seríceo en el ápice, con 7 nervios longitudinales; flores liguladas 32-36 por capítulo, de 5.4-6 mm. long., amarillas, dispuestas en 3-4 series; porción tubular de 1-2 mm. long., densamente cubierta por pelos pluricelulares largos; limbo de 4-4.5 mm. long. x 1.3-2 mm. lat., angosto-elíptico, ápice bidentado, dientes de 0.5 mm. long., ovados, agudos; pistilo de 2.5-4 mm.

long., estigma bifido, ramas estigmáticas de 1.5 mm. long.; flósculos 58-65 por capítulo, corola infundiliforme de 5.5 - 6.3 mm. long., amarilla; porción tubular de 3.3-4 mm. long., porción ensanchada pilosa hacia la base, con escasos pelos en el resto de su longitud y en la porción tubular; lóbulos triangulares de 0.8 mm. long., densamente pilosos en la cara externa, pelos pluricelulares hialinos; estambres exsertos, filamentos 4-4.5 mm. long., adnatos a la porción tubular de la corola, anteras singenésicas de 2-2.4 mm. long., apéndice apical angosto ovado de 0.4 mm. long.; pistilo de 5.8-7 mm. long. Aquenios de 2 mm. long. x 1.3-1.6 mm. lat., obovoideos, triquetros, glabros y con costillas prominentes.

*Espeletiopsis rabanalensis* S. Díaz & Rodríguez-Cabeza muestra su mayor afinidad con *Espeletiopsis guacharaca* (S. Díaz) Cuatrec., especie de la que difiere por presentar ramas floríferas más largas y carentes de brácteas estériles en su zona proximal o vegetativa; tanto las flores liguladas femeninas como los flósculos hermafroditas son más abundantes en número por capítulo, las flores femeninas tienen la lígula más desarrollada al punto que supera la longitud del pistilo y las flores del disco presentan los lóbulos densamente pilosos en su cara externa. Adicionalmente las hojas son obovadas, más pequeñas y menos rígidas.

En la clave propuesta por Cuatrecasas (1996) la nueva especie se puede incluir a partir del numeral 19 así:

- 19- Sinflorescencias 2-3 veces más largas que las hojas de la roseta. Corolas del disco hirsutas. Caulirrósula sésil o subsésil. .... **E. sclerophylla**
- 19'- Sinflorescencias similares en longitud a las hojas de la roseta. Caulirrósulas con el tallo alargado, erecto, de 2- 3 m de altura, densamente cubierto de hojas marcescentes ..... 20.
- 20- Porción proximal o vegetativa de las ramas floríferas con brácteas estériles ..... 21.
- 20'- Porción proximal o vegetativa de las ramas floríferas áfila. Lámina foliar 17-23.5 cm. x 3.5-6 cm. Flores femeninas con la lígula más larga que el pistilo, lóbulos de los flósculos densamente pilosos en su cara externa ..... **E. rabanalensis**
- 21- Lámina foliar (25) 36-58 (66) x 3.5-12 cm., obovado-oblonga, corolas del disco con limbo glabro o casi glabro. .... **E. jimenez-quesadae**
- 21'- Lámina foliar 20-40 x 3-6 cm., angosto-elíptica, flores femeninas con lígula más corta o igual que el pistilo, corolas del disco con limbo hirsuto, escasamente piloso en la cara externa de los lóbulos ..... **E. guacharaca**

***Espeletia pescana* (S. Díaz) S. Díaz stat. nov.**

*Espeletia brachyaxiantha* subsp. *pescana* S. Díaz, Mutisia 61:8. 1985.

Esta especie fue propuesta como una subespecie de *Espeletia brachyaxiantha* S. Díaz. El estudio de copiosos ejemplares provenientes de dos poblaciones con abundantes individuos, la típica ubicada en el Páramo de la Cortadera, (Municipios de Pesca y Toca), la otra en la Plaza del Diablo (Municipio de Siachoque) deja en claro que se trata de una especie bien definida netamente diferenciada por presentar sinflorescencias más largas y que sobrepasan en longitud a las hojas de la roseta, al tiempo que presentan mayor número de capítulos; las corolas de las flores liguladas son bastante más largas y las flores del disco presentan pelos bastante abundantes en la cara externa de los lóbulos. *E. pescana* (S. Díaz) S. Díaz muestra particular afinidad con la especie que se describe a continuación.

***Espeletia cachaluensis* Rodríguez-Cabeza & S. Díaz sp. nov. Figs. 2d-3**

*Espeletia pescanae* S. Díaz similis a qua imprimis differt foliis ellipticis nimis petiolatis, ramis floriferis sursum laxis et cum foliis bracteiformibus axillaribus maioribus, ligulis florum foeminarum item maioribus.

**Typus:** COLOMBIA, SANTANDER: Santuario de Fauna y Flora Guanentá Alto Río Fonce. Municipio de Encino, Vereda Avendaños, sector Los Cuadros. 3744 m. de altitud, N 05°49' W 073°06'. 12 oct 2007. *Betsy Viviana Rodríguez-Cabeza, Herber Palacios, Rafael Rivero, José Velasco 1897*, Holotypus COL (528207), Isotypi COL (528208), UIS y por repartir.

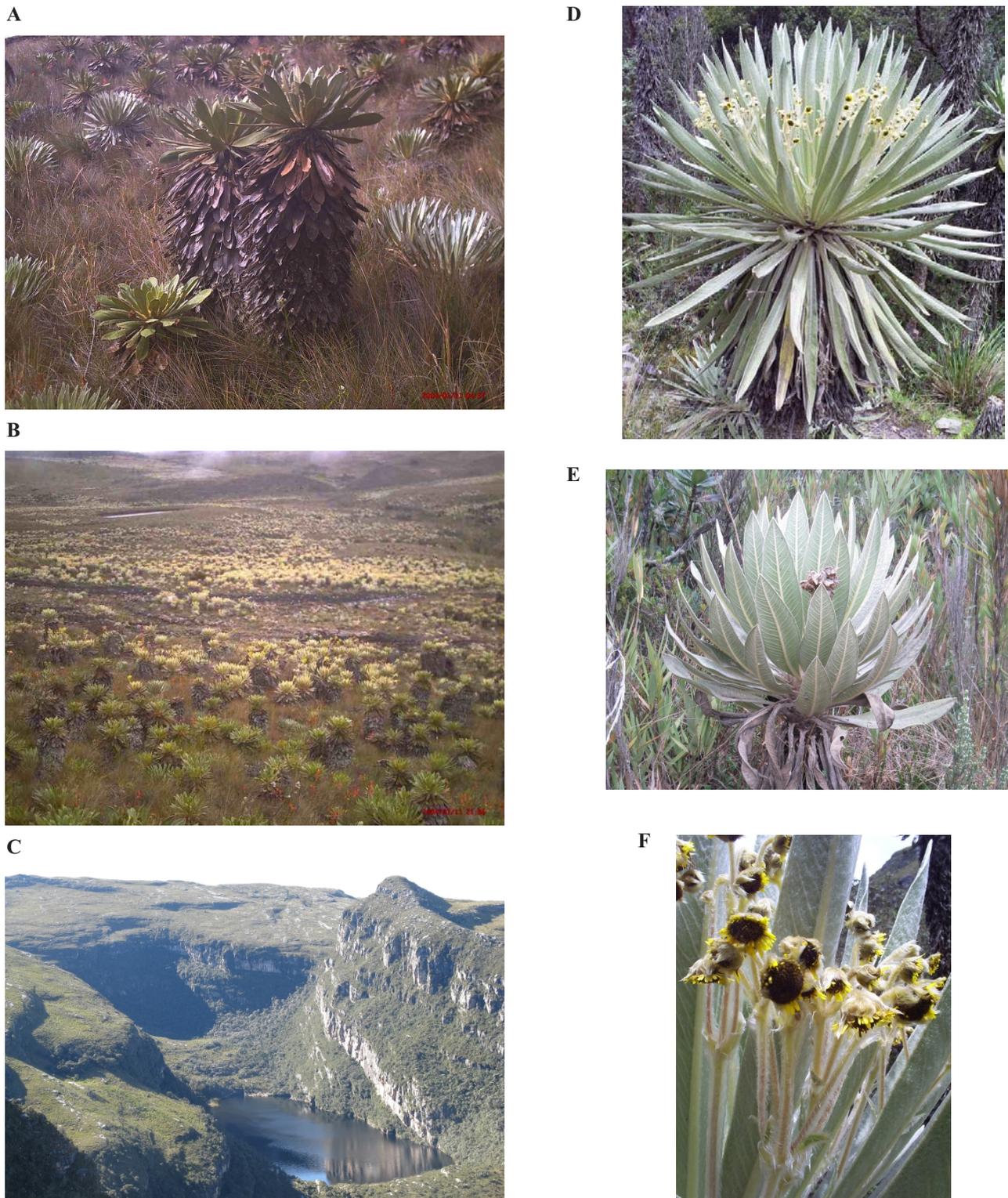
“Caulirrósula de 1,80 m de altura, hasta 6 m. **Hojas:** Haz, verde con indumento largo seríceo y largo flocoso blanco que se va tornando gris y se pierde en las hojas maduras. Envés, verde con corto indumento seríceo blanco. Venas central y secundarias con indumento seríceo amarillo claro visibles por el envés, ligeramente visibles en haz. **Vaina:** Haz, blanco con ápice verde limón e indumento largo seríceo blanco en las juveniles. Envés, indumento largo seríceo blanco y crema. **Inflorescencias:** 45-51 cm longitud, 6 por roseta, con 7-9 capítulos. Pedúnculos verdes, pedicelos verdes y vinotinto, todos ellos con indumento largo seríceo y flocoso blanco. Brácteas semejantes a las hojas de la roseta. **Capítulos:** 4-4,5 cm de ancho con liguladas extendidas. Flores liguladas amarillas, con lígula de 13 mm de long. Filarias verdes escasamente cubiertas por indumento café y largo flocoso blanco”. **Observaciones:** Arbustal-Frailejónal de hasta 3 m de altura, con *Weinmannia* sp.,

*Gynoxys* sp., *Polylepis quadrijuga*, *Hesperomeles* sp., asteráceas y chusque de hasta 1,50 m. Siempre fue observada en los peñascos o en la parte baja de estos en asociación con arbustos. La población tenía más de 300 individuos y a medida que estos van creciendo pierden gran parte de la necromasa, así en aquellos de 3 m de altura o más solo se observa la necromasa más cercana a la roseta. Es probable que esta especie sea la misma que se observa en las peñas de la laguna de Cachalú”.

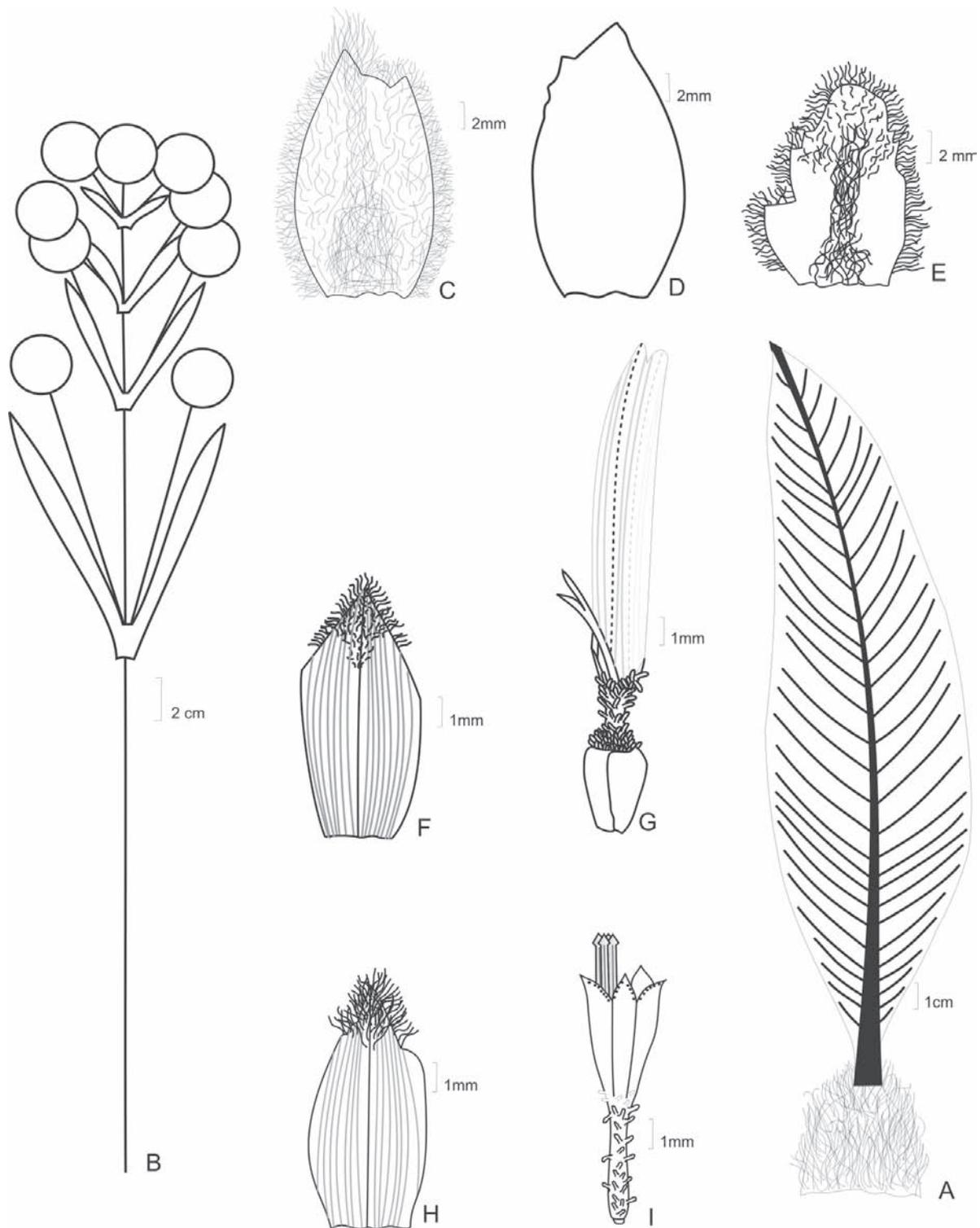
**Etimología:** El epíteto alude a la Laguna de Cachalú, en cuyos alrededores crece la especie, al abrigo de los peñascos.

Caulirrósulas de hasta 6 m. de altura. Vaina foliar de 5-5.5 cm. long. x 4-5 cm. lat., ovada, dorsalmente cubierta con indumento largo seríceo marrón muy claro, brillante, blanquecino; pecíolos cortos pero conspicuos, de 1-2.5 cm. long. x 1-1.2 cm. lat.; lámina foliar de 28-33.5 cm. long. x 7.5-8.8 cm. lat., elíptica, ligeramente papiráceo-coriácea, cubierta en la haz por indumento seríceo largo, flocoso, blanco que se torna gris y se pierde con la madurez, envés con indumento corto, seríceo, blanco; base foliar atenuada, margen lisa, ligeramente revoluta, ápice agudo, nervio medio ensanchado hacia la base, cubierto por indumento seríceo y largo flocoso cerino, impreso por la haz y prominente por el envés, nervios secundarios 11-13 por decímetro, conspicuos por el envés y divergiendo en ángulo de 50-70°.

Ramas floríferas 6 coetáneas, de 45-51 cm. de largo, cubiertas con indumento seríceo blanco al que se sobrepone un estrato flocoso de pelos largos; porción vegetativa de 23-25 cm. long., afila; porción distal un racimo con 7-9 capítulos; primer par de brácteas subtendientes de 11-15 cm. long. x 1.2-1.5 cm. lat., angosto lineales, las intermedias y superiores de 3-7 cm. long. x 1-2 cm. lat., angosto-ovadas, agudas en el ápice; pedúnculos florales verdes teñidos de vináceo, los inferiores de 7-12.5 cm. de long., los superiores de 2-6.5 cm long. Capítulos de 4-4.5 cm. de diámetro con las lígulas extendidas; involucreo hemisférico, filarias 6-7, dispuestas en 2 series, elípticas, las exteriores de 15-19 mm. long. x 10-14 mm. lat., cara abaxial cubierta totalmente por indumento seríceo largo y denso, cara adaxial con el cuarto superior cubierto de escaso indumento largo seríceo, el resto glabro, las interiores de 11-16 mm. long. x 5.5-10 mm. lat., cara abaxial cubierta con una lista longitudinal de tricomas que se ensancha hacia en ápice, quedando a cada lado un área glabra, cara adaxial con escasos tricomas hacia el ápice, ápice agudo, algunas filarias muestran la margen dentada; paleas de las lígulas de 8.5-13 mm. long. x 2.5-5.5 mm. lat., angosto elípticas o angosto



**Figura 2.** A. *Espeletopsis rabanalensis* S. Díaz & Rodríguez-Cabeza sp. nov.; B. Un aspecto del Páramo de Rabanal; al frente una población de la especie; C. Alrededores de la laguna de Cachalú, hábitat de *Espeletia cachaluensis* Rodríguez-Cabeza & S. Díaz sp. nov.; D. Detalle de un ejemplar de esta especie; E. *Espeletia chontalensis* Rodríguez-Cabeza & S. Díaz sp. nov.; F. Detalle de una planta con su rama florífera.



**Figura 3.** *Espeletia cachaluensis* Rodríguez-Cabeza & S. Díaz sp. nov. **A.** Esquema de la rama florífera; **B.** Perfil foliar; **C.** Filaria externa vista por el envés; **D.** Forma de las filarias exteriores; **E.** Filaria interior vista por el envés; **F.** Palea de las flores femeninas; **G.** Flor femenina ligulada; **H.** Palea de los flósculos; **I.** Flósculo hermafrodita. Diseño digital de Betsy Viviana Rodríguez-C., elaborado con base en el ejemplar típico (B.V. Rodríguez et al. 1897 COL).

-obovadas, externamente con una lista longitudinal de tricomas que se ensancha y se torna más densa hacia el ápice, quedando a cada lado un área glabra, ápice agudo; paleas de los flósculos de 7-9 mm. long. x 2-4 mm. lat., angosto-elípticas o elípticas, agudas en el ápice, abaxialmente con tricomas en el ápice y en la margen de la mitad superior y con 13 nervios longitudinales; flores femeninas liguladas 62-65 por capítulo, de 15-19 mm. long., amarillas; porción tubular de 2-2.2 mm. long., densamente cubierta por tricomas pluricelulares; limbo de 12.8-17 mm. long. x 1.8-2.5 mm. lat., angosto-elíptico, ápice con tres dientes de 0.7-1 mm. long., ovados, apicalmente obtusos; pistilo 6.5-7.5 mm. long., estigma bifido, ramas estigmáticas 1-2 mm. long.; flósculos 212-220 por capítulo, de 8-9 mm. long., infundibuliformes y amarillos; porción tubular de 3.5-4 mm. long., pilosa; lóbulos de la corola de 1 mm. long., triangulares, con ápice agudo y escasos tricomas hacia la margen; estambres exsertos, filamentos 6 mm long., adnatos a la porción tubular de la corola, anteras singénicas de 4 mm. long., con apéndice apical angosto ovado de 0.3 mm. long.; pistilo de 9 mm. long. Aquenios de 2.5-3 mm. long. x 2 mm. lat., obovoideos, triquetros, glabros y con costillas prominentes.

*Espeletia cachaluensis* Rodríguez-Cabeza & S. Díaz muestra su mayor afinidad con *E. pescana* (S. Díaz) S. Díaz, especie de la que difiere por la forma de la lámina foliar (elíptica vs. obovada), las hojas con pecíolo definido, aunque breve y las ramas floríferas, cuya porción distal o fértil es más laxa y abierta y con las brácteas foliáceas axilares mayores. Adicionalmente las ligulas son bastante más largas, triplicando la longitud del pistilo. La población típica supera los 500 individuos, algunos de los cuales son bastante altos al punto que sus tallos notablemente delgados se pueden doblar por el peso de la roseta.

***Espeletia chontalensis* Rodríguez-Cabeza & S. Díaz  
sp. nov. Figs. 2e-f-4**

Caulirrósula rotunda ad aspectum. Folia rosularia papiracea anguste-elliptica, utrinque attenuata, ad apicem acuta, basi in longo petiolo dilatata; rami floriferi cum foliorum rosula se aequantes vel paulo breviores ad basim foliorum paribus muniti; capitula pedunculata in cymis corimbiferis disposita, pedunculi capitulorum inaequales ipse deorsum longitudine aequantia simulans.

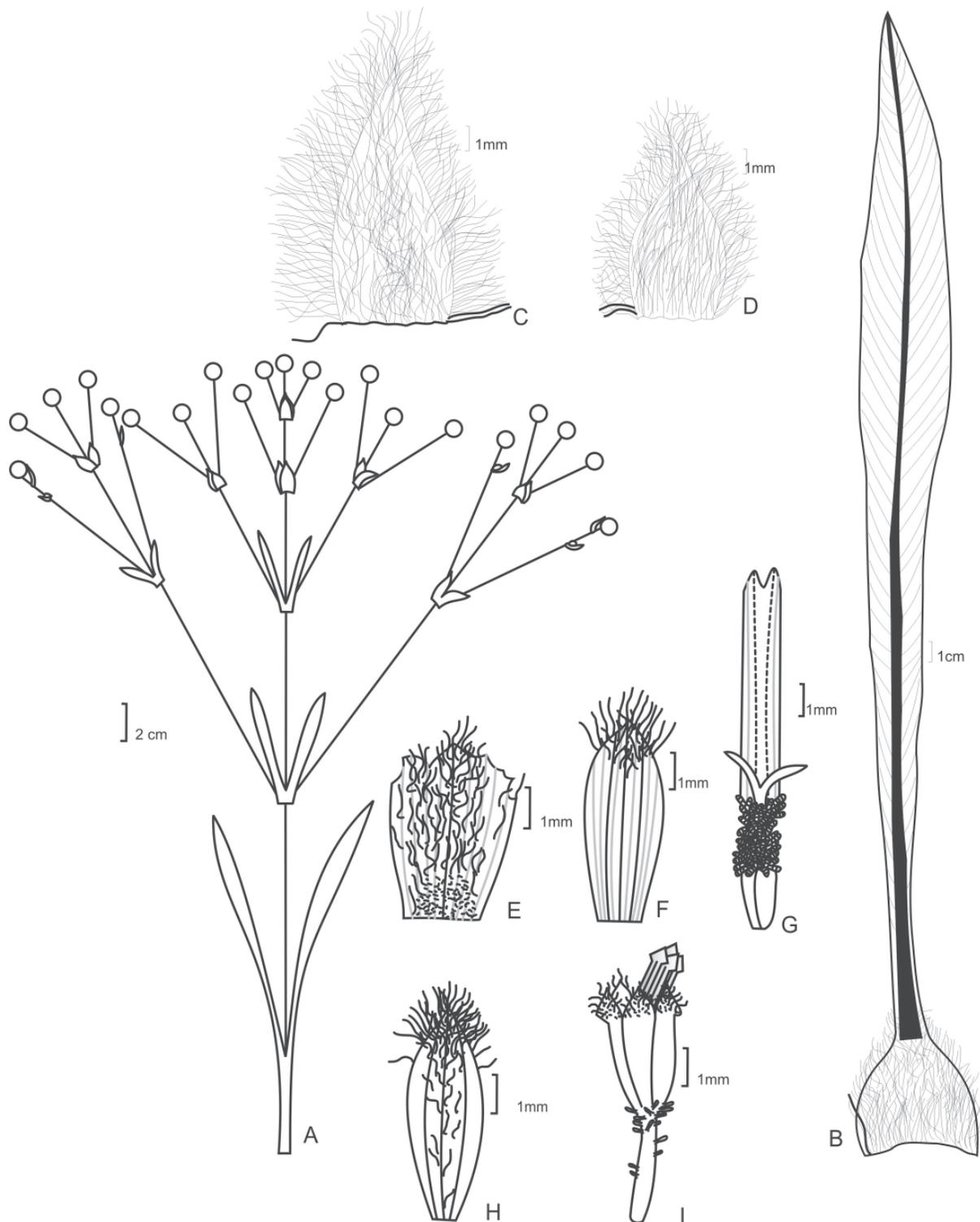
**Typus:** COLOMBIA, SANTANDER: **Santuario de Fauna y Flora Guanentá Alto Río Fonce.** Municipio de Encino, Vereda Río Negro, camino sector Chontales-Páramo de la Rusia. 28 jun. 2007. *Betsy Viviana Rodríguez-Cabeza, Herber Palacios 1874*, Holotypus COL (528209), Isotypi UIS y por repartir.

“Caulirrósula de 2 m de altura. **Hojas:** Cartáceas. Haz verde oscuro, Envés verde claro con indumento blanco-grisáceo claro que se torna oscuro en las adultas. Venas central y secundarias con indumento blanco-amarillo muy claro. **Vaina:** Haz blanco, hacia el ápice verde limón con indumento café lustroso. Envés indumento largo seríceo café lustroso. **Inflorescencias:** 47-51 cm de largo, 20-21 capítulos. Pedúnculos, pedicelos y brácteas de la inflorescencia con indumento largo seríceo blanco. Brácteas del capítulo con indumento seríceo café-amarillo lustroso. **Capítulos:** 2,5-3 cm de ancho con liguladas extendidas. Liguladas amarillas con ligula de 0,7 cm de largo. **Observaciones:** Pajonal-Frailejonal con *Calamagrostis effusa* y arbustitos aislados de *Vaccinium floribundum*, *Hypericum* sp., *Gaultheria anastomosans* y Asteraceae. Solo se encontró un individuo en floración, no se señala el tamaño de la población pues lo demás individuos no tienen inflorescencias, entonces se confunde con *Espeletia* BVR-1867. Acompañado por *Espeletia* BVR-1870 y *Espeletia* BVR-1867. Se observaron vacas, cabras y ovejas en el sector.”

**Etimología:** El epíteto alude al sector de Chontales en el Páramo de la Rusia, lugar de origen de la especie, área infortunadamente afectada por quemas frecuentes.

Tallo erecto arrosado de hasta 2 m. de altura; roseta con aspecto redondeado por la disposición de las hojas y de las ramas floríferas. Vaina foliar de 4.5-5.5 cm. long. x 5.2-5.5 cm. lat., ovada o ancho ovada, dorsalmente con indumento largo seríceo brúneo lustroso; lámina foliar de 49-52 cm. long. x 3.8-4.3 cm. lat., angosto elíptica, papirácea, haz verde oscuro, envés verde claro, cubierta en ambas caras con indumento argentado que se torna oscuro con la madurez; base largamente atenuada, pseudopeciada, margen lisa, ligeramente revoluta, ápice agudo; nervio medio ensanchado hacia la base, cubierto por indumento seríceo ebúrneo, impreso por la haz y prominente por el envés, nervios secundarios 8-9 por decímetro, conspicuos por el envés y divergiendo en ángulo de 20-30°.

Ramas floríferas de 47-52 cm. de largo total, cubiertas con indumento largo seríceo blanco; porción vegetativa de 19-24 cm. long., con un par de brácteas estériles de 19-20 cm. long. x 0.8-1 cm. lat., opuestas y ubicadas en la base de la porción vegetativa y con lámina similar a la de las hojas de la roseta; porción distal reproductiva tirsóidea, pero dando la apariencia de un corimbo, por alcanzar el tope todos los pedúnculos, con 20-21 capítulos; primer par brácteas subtendientes de 6.5-7.5 cm. long. x 0.7 cm. lat., angosto ovadas, las intermedias y superiores de 1.5-4 cm. long. x 0.4-0.5 cm. lat., angosto-ovadas, agudas en el



**Figura 4.** *Espeletia chontalensis* Rodríguez-Cabeza & S. Díaz sp. nov. **A.** Esquema de la rama florífera; **B.** Perfil foliar; **C.** Filaria exterior vista por el envés; **D.** Filaria interior vista por el envés; **E.** Palea de las flores femeninas con margen dentado; **F.** Palea de las flores femeninas con margen liso; **G.** Flor femenina ligulada; **H.** Palea de los flósculos hermafroditas; **I.** Flósculo hermafrodita. Diseño digital de Betsy Viviana Rodríguez-C., elabora con base en el ejemplar típico (*B.V. Rodríguez et al. 1874 COL.*).

ápice; pedúnculos florales de 2.5-16.5 cm. de long. Capítulos de 2.5-3 cm. de ancho, incluidas las lígulas extendidas; involucre hemisférico, filarias 9-10, dispuestas en 2 series, las exteriores de 8-10.5 mm. long. x 4-5.5 mm lat., las interiores de 5.5-7.5 mm. long. x 3.9-5.5 mm. lat., ovadas, cara abaxial cubierta totalmente por largo indumento seríceo, cara adaxial glabra, ápice agudo y alargado, margen lisa; paleas de las lígulas de 4.5-6 mm. long. x 2-4 mm. lat., obovado-elípticas, externamente con una lista longitudinal de tricomas que se ensancha y se torna más densa hacia el ápice, quedando a cada lado un área glabra o con indumento seríceo alargado solo en el ápice, agudas en ápice; paleas de los flósculos de 5-6 mm. long. x 2-2.4 mm. lat., obovadas o elípticas, con ápice agudo, abaxialmente con indumento seríceo denso en la porción apical y algunos tricomas en la parte central, con 3 nervios longitudinales; flores liguladas 58-60 por capítulo, de 7.5-9.7 mm. long., amarillas, dispuestas en 5-6 series; porción tubular de 1.8-2 mm. long., densamente cubierta por tricomas pluricelulares hasta de 0.4 mm. long.; limbo de 5.7-7.7 mm. long. x 1-1.2 mm. lat., angosto-elíptico, ápice con dos dientes de 0.5-0.7 mm. long., ovados, apicalmente agudos; pistilo 3.5-4.8 mm. long., estigma bifido, ramas estigmáticas 0.8-1.5 mm. long.; flósculos 127-176 por capítulo, de 5.2-6.7 mm. long., corola infundibuliforme, amarilla; porción tubular de 2-3 mm. long., escasamente pilosa en base y ápice del tubo; lóbulos de la de 0.8-1 mm. long., densamente pilosos; estambres exsertos, filamentos 3.5-4 mm. long., adnatos a la porción tubular de la corola, anteras singenésicas de 2.2-2.4 mm. long., con apéndice apical angosto ovado de 0.3-0.4 mm. long.; pistilo de 5.2-6 mm. long. Aquenios de 2-2.5 mm. long. x 0.8-1.2 mm. lat., obovoideos triquetros, glabros y con costillas prominentes.

Por la suma de sus características esta especie no muestra afinidad cercana con ninguna otra. Por la disposición de las hojas y de las ramas floríferas la caulirrósula semeja una estructura globosa; las hojas son notablemente largas, angosto elípticas, agudas en el ápice y pseudopeciadas en la base, las ramas floríferas son iguales o más cortas que las hojas y aparentan un corimbo dado que los capítulos de la panícula alcanzan alturas similares. El indumento es blanco argentado pero su densidad no alcanza a tapar,

en vivo, el color verdoso de las hojas y amarillo rojizo de las ramas floríferas que dan un hermoso contraste.

### Agradecimientos

La Unidad de Parques Nacionales Naturales de Colombia y en particular Nancy Rivera, Herber Palacios, Rafael Rivero y Jorge Villamil, facilitaron la recolección de material de estudio en el Santuario de Fauna y Flora Guantáná, Alto Río Fonce. CORPOBOYACÁ concedió apoyo financiero el proyecto: *Evaluación del estado de conservación de cuatro especies amenazadas de Boyacá* y la Fundación Ecosistema Colombia apoyo su ejecución. Los biólogos Félix Márquez, Felipe Rubio y Omar Lizarazo facilitaron la investigación en los páramos boyacenses e Isauro Cortez cooperó en la labor de campo. Robinson Galindo ha apoyado en forma permanente el estudio de los frailejones. Los curadores del Herbario Nacional Colombiano, Instituto de Ciencias de la Universidad Nacional de Colombia facilitaron la revisión de sus colecciones. Dejamos constancia de nuestra gratitud a todos por tan eficaz cooperación, merced a la cual sigue enriqueciéndose, con nuevas especies de asteráceas, la flora de Colombia con nuevas especies.

### Bibliografía citada

- Cuatrecasas, J. 1976. A new subtribe in the Heliantheae (Compositae) Espeletiinae. *Phytologia* **35**(1): 43-61.
- Cuatrecasas, J. 1996. Clave provisional de las especies del género *Espeletiopsis* Cuatrec. (Espeletiinae, Compositae). *Anales Jard. Bot. Madrid* **54**: 370-377.
- Díaz-Piedrahita, S. 1975. Aportes a la Flora de Colombia. Estudios en Compuestas II. *Caldasia* **11**(53): 19-25.
- Díaz-Piedrahita, S. & S. Obando. 2004. Una especie nueva de *Espeletiopsis* (Asteraceae, Heliantheae) de Colombia. *Rev. Acad. Colomb. Cienc.* **28**(108): 323-326.
- Díaz-Piedrahita, S., B.V. Rodríguez & R. Galindo-Tarazona. 2006. Interesantes novedades en Espeletiinae (Asteraceae-Heliantheae) de Colombia. *Rev. Acad. Colomb. Cienc.* **30**(116): 331-352.

Recibido: noviembre 16 de 2008

Aceptado para su publicación: diciembre 4 de 2008

# FRECUENCIA DE LAS SIGATOKAS NEGRA Y AMARILLA EN PLÁTANO (*MUSA* SPP.) DE ACUERDO A LOS ESTADOS DE DESARROLLO DE LA PLANTA

Por

Mayary Torrado-Jaime<sup>1</sup> & Jairo Castaño-Zapata<sup>1</sup>

## Resumen

**Torrado-Jaime, M., & J. Castaño-Zapata:** Frecuencia de las sigatokas negra y amarilla en plátano (*Musa* spp.) de acuerdo a los estados de desarrollo de la planta. Rev. Acad. Colomb. Cienc. 32(125): 465-470, 2008. ISSN 0370-3908.

En Colombia, las sigatokas negra y amarilla están entre las principales enfermedades que atacan al plátano. Aunque hay registros de su presencia, se desconoce la época en que aparecen de acuerdo a los estados fenológicos y condiciones ambientales. Se sembraron cormos de Dominico Hartón, África, FHIA-20 y FHIA-21, en un diseño de bloques al azar con cuatro repeticiones con 10 plantas por repetición. Se definieron cinco etapas de desarrollo: brotación y emergencia, plántula, prefloración, floración, y llenado del racimo y madurez fisiológica. Las sigatokas se presentaron desde etapa de plántula hasta madurez fisiológica. La mayor población de conidias en D. Hartón y África, se presentó en etapas de prefloración y llenado del racimo; en los híbridos FHIA-20 y FHIA-21, coincidió con la etapa de floración. Temperatura, precipitación y humedad foliar favorecieron la incidencia y severidad de las sigatokas.

**Palabras clave:** estados de desarrollo, *Mycosphaerella fijiensis*, *Mycosphaerella musicola*, plátano.

## Abstract

In Colombia, black and yellow Sigatokas are two of the major diseases that affect plantain cultivation. Although there is information about their presence, the time that they appear according to plant development and environmental conditions is unknown. Rhizomes of Dominico Hartón, África, FHIA-20 and FHIA-21 were sowed following a random block design with four repetitions

<sup>1</sup> Universidad de Caldas, Facultad de Ciencias Agropecuarias. Maestría en Fitopatología. Apartado Aéreo 275. Manizales, Colombia. Correo electrónico: fitotecnia@ucaldas.edu.co

and 10 useful plants per repetition. It was defined five development stages for the crop: emergency, seedling, pre-bloom, bloom, and physiologic maturity. The sigatokas appeared from seedling stage and stayed during the whole cultivation cycle. The highest conidia population in D. Hartón and Africa was presented in pre-bloom and filled bunch stages; in the hybrids FHIA-20 and FHIA-21, it coincided with the bloom stage. Temperature, precipitation and foliar humidity influenced diseases incidence and severity.

**Key words:** stages of development, *Mycosphaerella fijiensis*, *Mycosphaerella musicola*, plantain.

## Introducción

En Colombia, la Sigatoka negra (*Mycosphaerella fijiensis*, anamorfo *Paracercospora fijiensis*) y la Sigatoka amarilla (*Mycosphaerella musicola*, anamorfo *Pseudocercospora musae*) se encuentran entre las nueve enfermedades principales que afectan actualmente la producción de plátano en el país (Martínez *et al.*, 2006).

En el mundo en general, y en Colombia en particular, se conocen las principales enfermedades que atacan al cultivo del plátano, pero no se dispone de información que indique la presencia de éstas de acuerdo a los estados de desarrollo de la planta; tampoco se conoce la tasa de desarrollo de esas enfermedades a través del tiempo. Esta investigación permite determinar la época en que las sigatokas empiezan a atacar y cómo se desarrollan durante el primer ciclo del cultivo, de acuerdo a los elementos ambientales prevalentes en la región de Santágueda en el municipio de Palestina, Caldas. El conocimiento de la época en que las enfermedades aparecen en las cultivariedades de plátano África y Dominico Hartón, y en los híbridos FHIA-20 y FHIA-21, se complementa con curvas de incidencia de las enfermedades, lo cual permitirá diseñar estrategias efectivas de manejo en estos genotipos, antes de alcanzar niveles de daño económico.

## Metodología

El estudio se realizó en la granja Montelindo, propiedad de la Universidad de Caldas, ubicada en la región de Santágueda, en el municipio de Palestina departamento de Caldas, ubicado a 5°05' latitud norte y 75°40' longitud occidental, a una altitud de 1050 m., con temperatura media de 22.5°C, humedad relativa del 76%, precipitación anual de 2100 mm y brillo solar anual de 2010 h. El tipo de suelo es Inceptisol de la serie Montelindo (Tropoeps Dystropepts).

Se empleó como semilla cormos de Dominico Hartón, África, FHIA-20 y FHIA-21, provenientes de las plantaciones existentes dentro de la misma granja. No se realizó tratamiento a la semilla. La distancia de siembra fue de 3 x 2 m entre surcos y plantas, respectivamente. Se realiza-

ron las prácticas agronómicas convencionales de fertilización, manejo de arvenses, descoline (dejando dos colinos por planta), desbellote, resiembras y deshoje fitosanitario, según análisis de suelo y las recomendaciones técnicas establecidas para el cultivo del plátano en la región. La fecha de inicio del experimento fue el 26 de mayo de 2004, día en que se realizó la siembra, y tuvo una duración total de 18 meses.

Se establecieron parcelas independientes de los cuatro materiales, con 10 plantas cada una, en un diseño de bloques al azar con cuatro repeticiones, sembrando una planta por sitio. Semanalmente a cada una de las plantas se le tomó datos de número de hojas e hijos, y el tiempo (días) a emergencia del primer hijo (DEH). En floración se hicieron registros de altura de la planta (AP), diámetro del pseudotallo (DP) y días a floración (DF).

Para enmarcar las enfermedades evaluadas dentro de la fenología del cultivo se establecieron las etapas de desarrollo para cada uno de los materiales. Para esto se simplificaron las etapas establecidas por Aristizábal y Jaramillo (2005) para la cultivariedad Dominico Hartón. Se utilizaron los datos recolectados en el presente estudio de días para emisión de hijuelos y días a floración. Los datos fueron complementados con los de días a cosecha obtenidos por Herrera y Aristizábal (2003) y los de Aristizábal y Jaramillo (2005) para brotación y emergencia y duración de la floración.

Semanalmente se llevaron a cabo lecturas de severidad de las sigatokas en la hoja más joven enferma (HMJE) de cada una de las plantas. Para esto se hizo uso de los diagramas estándares de Stover (1971) y la escala de grados de desarrollo de las sigatokas (Aranzazu *et al.*, 2002). Al momento de floración se realizó nuevamente lectura de la HMJE y además la hoja más joven manchada (HMJM). La HMJE representa la hoja más joven en manifestar estrías claramente visibles desde el suelo (Orjeda, citado por Gómez & Castaño-Zapata, 2001) y la HMJM corresponde a la primera hoja totalmente abierta que presenta 10 o más lesiones discretas necrosadas y maduras o un área grande necrosada con 10 centros de color claro, contando

las hojas de arriba hacia abajo (Stover & Dickson, citados por Gómez & Castaño-Zapata, 2001). Adicionalmente, con los datos recolectados se calculó el tiempo de evolución de los síntomas (TES) y la tasa de desarrollo (r) de las sigatokas. El TES es el número de días desde la aparición de los primeros síntomas hasta el estado de mancha con centro necrosado (grado 5-6) (Molina & Castaño-Zapata, 2003). Para la tasa de desarrollo se tomaron los valores promedios de severidad semanal y se transformaron a proporción; también se usaron los valores correspondientes a las semanas de evaluación y la diferencia de días transcurridos desde la primera evaluación y la última (Castaño-Zapata, 2002). Para tal fin, se empleó la siguiente fórmula:

$$r = \frac{1}{t_1 - t_0} \left( \text{Log}_e \frac{X_1}{1 - X_1} - \text{Log}_e \frac{X_0}{1 - X_0} \right)$$

Donde:  $t_1$  = tiempo final de evaluación,  $t_0$  = tiempo inicial de evaluación,  $X_1$  = severidad final y  $X_0$  = severidad inicial.

Para determinar la predominancia de ambas sigatokas se realizó cada 2 meses el método de la impronta desarrollada por Aguirre *et al.* (2003). Para fines de este estudio se tomaron dos improntas en la misma hoja, una por la haz y la otra por el envés, de una planta por repetición.

Se tomaron datos de precipitación y humedad relativa de la estación meteorológica de la granja. Los registros de temperatura y humedad foliar se obtuvieron de una microestación climatológica de Spectrum Technologies.

Utilizando el programa estadístico SAS (*Statistical Analysis System*, SAS Institute) se sometieron los datos obtenidos a análisis de varianza según un diseño en bloques al azar con cuatro repeticiones. La separación de medias se efectuó mediante la prueba de Tukey y se realizaron análisis de correlación de Pearson.

### Resultados y discusión

El tiempo de floración de la cultivariedad África fue significativamente menor, entre 4 y 7 semanas, que el de los otros materiales, presentando un ciclo vegetativo más corto. En una investigación realizada anteriormente en la misma área de estudio, también se reporta al material África con un tiempo de floración significativamente menor que el de los otros materiales, caracterizándolo como un material precoz.

La duración de las etapas de crecimiento de los materiales, en las cuales se enmarcaron las enfermedades identificadas, se presenta en la Figura 1.

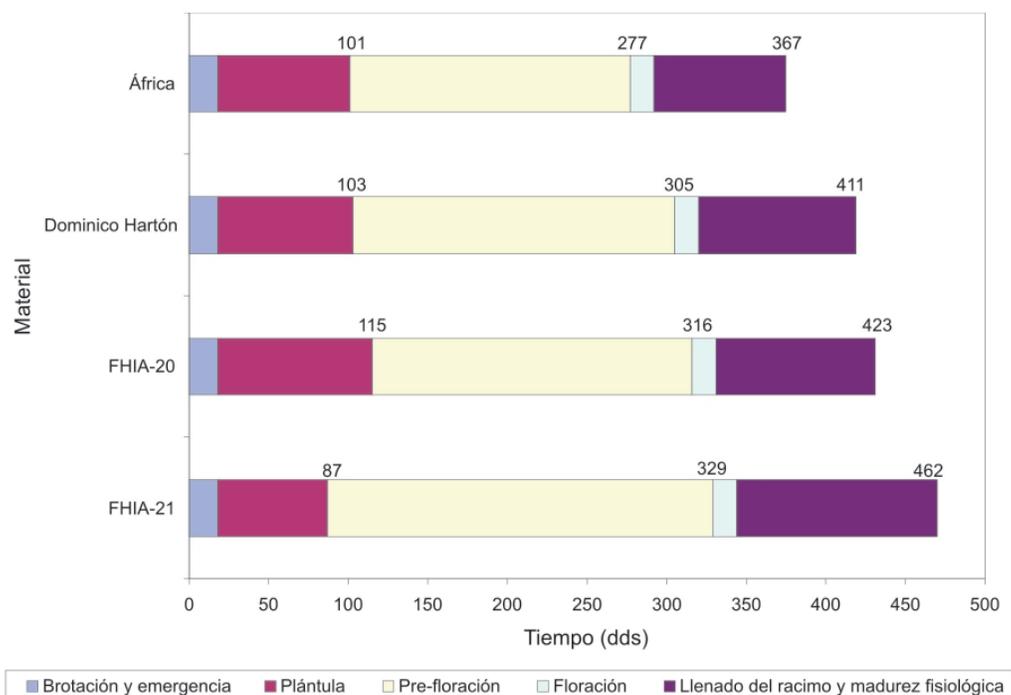


Figura 1. Duración de las etapas de crecimiento en cuatro materiales de plátano en la región Santágueda (Caldas).

La presencia de las sigatokas amarilla y negra se detectó en todos los materiales 70 días después de la siembra (**dds**), cuando las plantas se encontraban en etapa de plántula, presentando entre 4 y 5 hojas verdaderas. La presencia de las sigatokas en esta etapa, además de que el área de estudio presenta condiciones óptimas para el desarrollo de estas enfermedades, también se puede explicar por la existencia de hojas disponibles para permitir infección, ya que las hojas verdaderas son emitidas aproximadamente 30 **dds**, y el periodo de incubación (tiempo transcurrido desde la inoculación hasta la aparición de los primeros síntomas) que requieren estos hongos es de 29 días (**Belalcázar et al.**, 2000). Los valores de incidencia fueron mayores para África y Dominico Hartón, de 95 y 93% respectivamente, mientras que en FHIA-20 (84%) y FHIA-21 (72%) los valores fueron menores. La incidencia alcanzó su valor máximo 7 días después en Dominico Hartón y FHIA-21, en FHIA-20 tardó 14 días y en África después de 21 días ésta llegó al 100%. Estas incidencias se mantuvieron durante todo el ciclo del cultivo (Figura 2).

Dos semanas antes de ser detectadas las sigatokas se presentó un periodo de lluvias intensas. Posteriormente, las lluvias disminuyeron y surgió un periodo seco con incremento gradual en la temperatura desde 23.4°C hasta aproximadamente 26°C. La duración de la película de agua en las hojas igualmente mostró un aumento (de 7 a 10 h/día) en el tiempo que fueron detectadas las enfermedades, condiciones óptimas para la germinación de las conidias y ascosporas (**Craenen**, 1998) (Figura 2).

En general, los materiales presentaron un buen número de hojas funcionales a floración, **NHFF** (11 en promedio), teniendo en cuenta que se espera que las plantas tengan

un mínimo de ocho hojas funcionales en este estado (**Stover; Meredith; Martínez**, citados por **Gómez & Castaño-Zapata**, 2001). Los híbridos con 12 hojas y las cultivariedades África y Dominico Hartón con 9 y 10 hojas respectivamente, mostraron diferencias estadísticas altamente significativas. Estas diferencias posiblemente están relacionadas con la resistencia o susceptibilidad de los materiales a las sigatokas (Tabla 1).

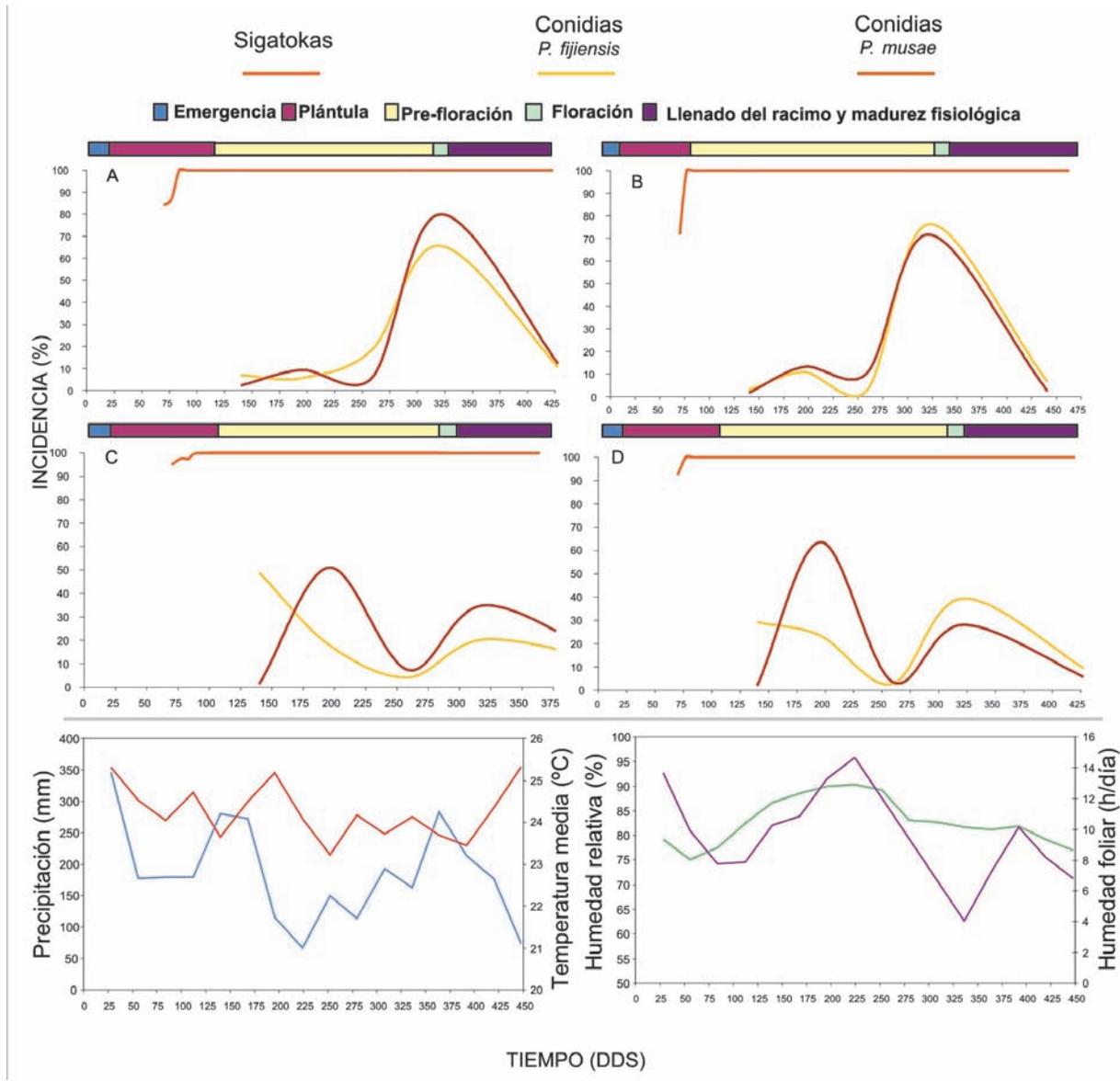
En las variables hoja más joven enferma, hoja más joven manchada y número de conidias de *P. fijiensis* y de *P. musae*, se aprecia claramente la resistencia a las sigatokas de los híbridos y la susceptibilidad de las cultivariedades. FHIA-20 presentó el valor más alto de hoja más joven enferma, **HMJE**, con diferencias significativas con las cultivariedades e igualmente junto a FHIA-21 obtuvieron los valores más altos de hoja más joven manchada, **HMJM**. En los híbridos se detectó un número significativamente menor de conidias, tanto de *P. fijiensis* como de *P. musae*, comparativamente con las cultivariedades lo cual reafirma la resistencia de estos materiales a las sigatokas. Así mismo, estos resultados corroboran el alto grado de susceptibilidad de África y Dominico Hartón, ya que presentaron una población de conidias con respecto a FHIA-21 y FHIA-20 en relación 4:1 y 5:1, respectivamente (Tabla 1).

Los datos obtenidos de tasa de desarrollo (**r**) y tiempo de evolución de los síntomas (**TES**) hacen más evidentes las diferencias entre los materiales en cuanto a su resistencia o susceptibilidad a las sigatokas. La tasa de desarrollo y el tiempo de evolución de los síntomas para Dominico Hartón y África presentaron diferencias significativas con los híbridos pero no entre sí. Para ambos materiales el desarrollo de las sigatokas fue rápido con valores

**Tabla 1.** Reacción a las sigatokas de cuatro materiales de plátano en la región Santágueda (Caldas).

Materiales	Variables de respuesta (Media)						
	NHFF	HMJE	HMJM	Conidias/cm <sup>2</sup> (Media)		r	TES
				<i>P. fijiensis</i>	<i>P. musae</i>		
FHIA – 20	12a*	6.3a	9a	8a	8a	0.81b	31b
FHIA – 21	12a	5.3ab	9a	14a	10a	0.34c	59c
África	9b	4.8b	7b	55b	31b	1.06a	27a
D. Hartón	10b	4.8b	8ab	52b	39b	1.12a	25a
D.M.S.(5%)	2	1.0	2.0			0.21	

\* Valores promedios en la misma columna seguidos por letras distintas denotan diferencias significativas según la prueba de comparación de medias de Tukey al 5%.



\* A) FHIA-20, B) FHIA-21, C) África, D) Dominico Hartón.

**Figura 2.** Fenología de cuatro materiales de plátano y la incidencia de las sigatokas en la región Santagueda (Caldas) con relación a las variables ambientales.

altos de tasa de desarrollo ( $r = 1.12$  y  $1.06$ , respectivamente) y periodos cortos de evolución de los síntomas ( $TES = 25$  y  $27$  días, respectivamente), reflejando una alta susceptibilidad a las sigatokas. Por otro lado, FHIA-20 con una tasa de desarrollo y un tiempo de evolución de síntomas con valores intermedios de  $0.81$  y  $31$  días, respectivamente, muestra cierta susceptibilidad a las sigatokas, lo cual podría indicar que su resistencia a estas enfermedades esta disminuyendo. Estas dos variables en el híbrido FHIA-21 son significativamente más lentas comparadas con las de

los otros materiales, con valores de  $r = 0.34$  y  $TES = 59$  días, lo cual pone de manifiesto su resistencia a dichas enfermedades (Tabla 1).

Aún cuando no se presentaron diferencias estadísticas significativas en las poblaciones de conidias a nivel de especie, *Paracercospora fijiensis* (estado anamorfo de *Mycosphaerella fijiensis*), con una población media de  $19$  conidia/cm<sup>2</sup> y *Pseudocercospora musae* (estado anamorfo de *Mycosphaerella musicola*) con  $13$  conidias/cm<sup>2</sup> (datos

transformados), la cantidad de conidias en las dos cultivariedades refleja que la Sigatoka negra es más agresiva en el área experimental. En Dominico Hartón el número medio de conidias de *P. fijiensis* fue de 52/cm<sup>2</sup> y de *P. musae* 39/cm<sup>2</sup>, y en África las poblaciones fueron de 55 conidias/cm<sup>2</sup> y 31 conidias/cm<sup>2</sup> para *P. fijiensis* y *P. musae*, respectivamente. Resultados similares han sido reportados por **Molina y Castaño-Zapata** (2003) en la misma área de estudio y por otros autores (**Cardona-Sánchez & Castaño-Zapata**, 2002; **Aguirre & Castaño-Zapata**, 2005) en una región cercana con condiciones similares. Ellos observaron la tendencia de la Sigatoka negra a desplazar la Sigatoka amarilla dada su mayor agresividad, basados en las poblaciones de conidias de materiales susceptibles a las Sigatokas como son el Dominico Hartón y el África, entre otros. El número de conidias de *P. fijiensis* fue significativamente mayor en el envés (838/cm<sup>2</sup>) que en la haz (200/cm<sup>2</sup>). Sin embargo, el valor medio de conidias de *P. musae* fue de 486 con./cm<sup>2</sup> en el envés y 224 con./cm<sup>2</sup> en la haz, valor significativamente mayor en el envés que en la haz, lo cual difiere de lo encontrado en la literatura (**APS**, 2005; **Gasparotto et al.**, 2006).

En la Figura 2, observando el comportamiento de los materiales susceptibles se aprecia claramente que, en general, en los periodos de menor pluviosidad (259 y 425 dds) el número de esporas disminuyó significativamente y en los periodos con incremento en la precipitación (192 y 322 dds) la población de esporas fue mayor. Sin embargo, esto no aplica para los materiales resistentes en los que la producción de conidias, si bien coincidió con época de precipitación, pareciera estar influenciada por el estado fenológico de la planta que coincide justamente con la época de floración. Ésto parece posible si se considera la etapa de floración como un periodo de gasto energético incrementado para la planta, lo cual podría en cierta forma presentarse como un factor predisponente. La temperatura y humedad relativa con promedios de 24.5°C y 83.7% favorecieron la producción de conidias. De acuerdo con **Craenen** (1998) temperaturas entre 20 y 35°C favorecen el desarrollo de las sigatokas y más aún cuando se presentan con humedades relativas mayores al 80% (**Merchán**, 1997). Estudios similares resaltan también la importancia de las condiciones ambientales, principalmente la precipitación, sobre la producción de conidias de *P. fijiensis* y *P. musae* (**Aguirre y Castaño-Zapata**, 2005; **Cardona-Sánchez y Castaño-Zapata**, 2002). De igual forma, la mayor producción de esporas, aún en materiales susceptibles como Dominico Hartón, fue obtenida en la época de pre-floración sin que ésta coincidiera con el mayor nivel de precipitación (**Cardona-Sánchez y Castaño-Zapata**, 2002).

## Bibliografía

- Aguirre, MC & J. Castaño-Zapata**. 2005. Epidemiología de *Mycosphaerella fijiensis* Morelet y *M. musicola* Leach, en siete genotipos de Musáceas. Fitopatología Colombiana 29 (1): 7-11.
- Aguirre, MC, J. Castaño-Zapata & LE Zuluaga**. 2003. Método rápido de diagnóstico de *Mycosphaerella musicola* Leach y *M. fijiensis* Morelet, agentes causales de las Sigatokas amarilla y negra. Revista Academia Colombiana de Ciencias 27 (105): 619-623.
- APS**. The American Phytopathological Society. 2005. Sigatoka negra, síntomas y signos <<http://www.apsnet.org/Education/LessonsPlantPath/BlackSigatokaEspanol/symptom.htm>>.
- Aranzazu, F. et al.** 2002. El cultivo de plátano: manual técnico. Manizales: Corpoica. 114 p.
- Aristizábal, LM & C. Jaramillo**. 2002. Identificación y descripción de las etapas de crecimiento del plátano (*Musa* AAB) Dominico hartón. Pp. 79-90 In Memorias II Seminario Internacional sobre Producción, Comercialización e Industrialización de Plátano. Manizales.
- Belalcázar, S. et al.** 2000. Plagas y enfermedades del plátano. In Boletín de Sanidad Vegetal. Santafé de Bogotá, 106 p.
- Cardona-Sánchez, CL & J. Castaño-Zapata**. 2002. Frecuencia de *Paracercospora fijiensis* y *Pseudocercospora musae* en plátano Dominico Hartón. Infomusa 11 (1): 9-13.
- Castaño-Zapata, J.** 2002. Principios básicos de fitoepidemiología. Manizales: Unicaldas, 398 p.
- Craenen, K.** 1998. Black Sigatoka disease of banana and plantain. A reference manual. Ibadan, Nigeria : IITA, 60 p.
- Gasparotto, L. et al.** 2006. Sigatoka-negra da bananeira. Manaus: Embrapa Amazônia Ocidental, 177 p.
- Gómez, LA & J. Castaño-Zapata**. 2001. Manejo integrado de las Sigatokas negra y amarilla en la cultivariedad de plátano África. Infomusa. 10(2): 3-7.
- Herrera, JW. & M. Aristizábal L.** 2003. Caracterización del crecimiento y producción de híbridos y cultivariedades de plátano en Colombia. Infomusa 12 (2): 22-24.
- Martínez, HJ, Y. Peña & C. Espinal**. 2006. La cadena de plátano en Colombia: una mirada global de su estructura y dinámica 1991-2005. Documento de trabajo No. 102. Bogotá (Colombia): Ministerio de Agricultura y Desarrollo Rural. <[http://www.agrocadenas.gov.co/platano/documentos/caracterizacion\\_platano.pdf](http://www.agrocadenas.gov.co/platano/documentos/caracterizacion_platano.pdf)>. 95 p.
- Merchán, VM.** 1997. Prevención y manejo de la Sigatoka negra. Boletín de Sanidad Vegetal 17. Santafé de Bogotá : ICA, 29 p.
- Molina, TOI & J. Castaño-Zapata**. 2003. Análisis de algunos componentes de resistencia en los híbridos de banano y plátano FHIA 01, FHIA 17 y FHIA 21 a las Sigatokas negra (*Mycosphaerella fijiensis* Morelet) y amarilla (*M. musicola* Leach). Revista Academia Colombiana de Ciencias 27 (103): 181-189.
- Stover, RH.** 1971. A proposed international scale for estimating intensity of banana leaf spot. Tropical Agriculture (Trinidad) 48; p. 185-196.

Recibido: junio 14 de 2007

Aceptado para su publicación: noviembre 3 de 2008

# GEOTECNIA EN SUELOS NO SATURADOS

Por

Miguel Angel Alfaro Soto<sup>1</sup>

## Resumen

**Alfaro Soto, M.A.:** Geotecnia en suelos no saturados. Rev. Acad. Colomb. Cienc. **32**(125): 471-481, 2008. ISSN 0370-3908.

Se presenta una síntesis sobre tópicos referentes a la mecánica de suelos no saturados con la finalidad de incentivar su adopción en la Geotecnia. Se muestran aspectos más relevantes relativos al tema tales como: características del suelo no saturado, conceptos de succión, variables de estados de tensión y algunas teorías para retratar algunas propiedades mecánicas como variación de volumen y resistencia al corte y por otro lado, propiedades hidráulicas como permeabilidad en suelos no saturados.

**Palabras clave:** geotecnia, suelo no saturado, succión.

## Abstract

In this paper, a synthesis on referring topics to unsaturated soils mechanics is presented to incentive their practice in geotechnical engineering. Some important aspects are presented as: unsaturated soil characteristics, suctions concepts, stress state variable and some theories to portray some mechanical properties like as volume change and shear strength and on the other hand, hydraulic properties as permeability in unsaturated soils.

**Key words:** geotechnical engineering, unsaturated soil, suction.

## Introducción

Conceptos de la Mecánica de Suelos clásica fueron desarrollados, en su mayoría, considerando el suelo saturado. Esto se debió probablemente a factores como: casi o total saturación de los suelos en los países donde esas teorías fueron concebidas (climas templados y fríos del hemisferio Norte), porque la saturación constituye la si-

tuación crítica para una diversidad de obras o debido a la simplificación de los modelos para explicar el comportamiento de esos sistemas bifásicos (solo y agua completamente ocupada en los vacíos del suelo).

Sin embargo, una gran parte de la población del mundo se encuentra implantada en centros urbanos en que predominan suelos con nivel freático profundos donde los

<sup>1</sup> Doctor, Laboratorio de Estudos de Bacias (LEBAC), Universidade Estadual Paulista (UNESP), Rio Claro-SP, Brasil, alfaro@rc.unesp.br.

vacíos del suelo no se encuentran totalmente ocupados por agua y sí con agua y aire. En esos lugares, son necesarias nuevas formulaciones para entender los comportamientos de esos suelos, así como, nuevas metodologías para ensayarlos (Wolle, 2004).

Las limitaciones de la Mecánica de Suelos tradicional son evidentes cuando se necesita explicar las deformaciones en suelos no saturados o de las estructuras apoyadas en esos suelos (pavimentos, taludes, cimentaciones entre otros) sujetos a cargas de servicio o en los estados de tensiones totales presentes *in situ*. Algunos trabajos (Parreira *et al.*, 2004; Alfaro Soto, 2004) han mostrado la influencia del grado de saturación en la deformación e resistencia de los suelos. Esta situación, puede ser más compleja aún, cuando nos encontramos con suelos de características expansivas o colapsables donde la estructura del suelo también influye en la variación del volumen del suelo tal como mostrado por Pereira & Pejón (2004), Teixeira *et al.* (2004), entre otros.

En todos los casos mencionados, el volumen del suelo no saturado se modifica por que se encuentra sometido a un estado de tensión diferenciado e influenciado por una componente denominada tensión de succión. La succión puede hacer variar el volumen del suelo, sin embargo, su magnitud, depende de la humedad del medio poroso como veremos más adelante pues, succión y humedad se relacionan íntimamente.

Problemas geotécnicos de estados límites (estabilidad de taludes, capacidad de carga de cimentaciones superficiales o profundas, estabilidad de contenciones, etc.), relativos a la resistencia al corte de suelos no son diferentes, y son analizados considerando el suelo saturado.

Podemos mencionar la estabilidad de taludes como un ejemplo típico de esta situación, donde los modelos utilizados para el análisis cuantitativo de los factores de seguridad fueron desarrollados para la situación más crítica (saturación). No obstante, la desestabilización es un problema de suelo no saturado que puede ocurrir después de largos periodos de precipitación debido a la disminución de las tensiones de succión (aumento del grado de saturación desde la humedad natural en que inicialmente se encontraba). El caso contrario a la desestabilización ocurre, con aumento de la resistencia al corte debido a los efectos benéficos del aumento de la succión (disminución de la humedad) ocasionando aumentos en el factor de seguridad. Diversos estudios en suelos no saturados entre ellos Santos & Vilar (2004) y Reis & Vilar (2004), han mostrado cómo la succión es responsable por la modificación de los parámetros mecánicos.

El flujo de agua en el suelo no saturado, es también una propiedad que es de interés en diferentes problemas geotécnicos como los citados anteriormente (infiltración en taludes, consolidación, resistencia entre otros). Sin embargo, uno de los problemas de actual interés es respecto a la transmisión de contaminantes a través de la zona no saturada, cuando en la superficie terrestre, ocurren derrames accidentales de productos químicos, aplicación de fertilizantes y pesticidas o derrames de tanques sépticos, de gasolina, o todos aquellos que contribuyen para la contaminación del agua subterránea. En relación a ese tema Alfaro Soto & Chang (2007), mostraron la variabilidad de la conductividad hidráulica en suelos no saturados sometidos a las condiciones atmosféricas.

Sean propiedades mecánicas o hidráulicas sus magnitudes se encuentran en función de la tensión de succión, que dependerá a su vez del grado de saturación del suelo. Esto sugiere que en lugares en que el suelo es parcialmente saturado la mecánica de suelos tradicional no necesariamente puede ser la más adecuada y puede llevar soluciones con diseños de emprendimientos sobre-dimensionados (para satisfacer una condición crítica) de alto costo y no representativos del medio en que se encuentran.

Este artículo pretende mostrar los aspectos más relevantes de la mecánica de suelos no saturados con la finalidad de incentivar a su utilización en la Geotecnia.

### Succión y estado tensional de suelos no saturados

Según Fredlund (1995), los suelos no saturados en el medio ambiente pueden ser esquematizados según la Figura 1. Las Figuras 1a, 1b y 1c muestran representaciones de suelos en situaciones general, en región árida y región húmeda respectivamente. En todos los casos la elipse contiene una línea horizontal que representa el nivel freático y que divide el suelo saturado (debajo de la línea) del suelo no saturado (arriba de la línea) también llamado zona vadosa.

Debajo del nivel freático las presiones intersticiales (o presión neutra) serán positivas. Inmediatamente arriba del nivel freático en la zona no saturada ocurre una capa capilar con un grado de saturación de aproximadamente 100% y que puede tener una espesura variable de hasta 10m en función del tipo de suelo (Figura 1d). Arriba de esa capa el suelo presenta presiones intersticiales o de poro negativas y que resultan de la diferencia de las presiones del aire ( $u_a$ ) y del agua ( $u_w$ ). La diferencia entre esas presiones ( $u_a - u_w$ ) es llamada de succión matricial y que corresponde a una de las dos variables de tensión que describen el comportamiento de un suelo no saturado.

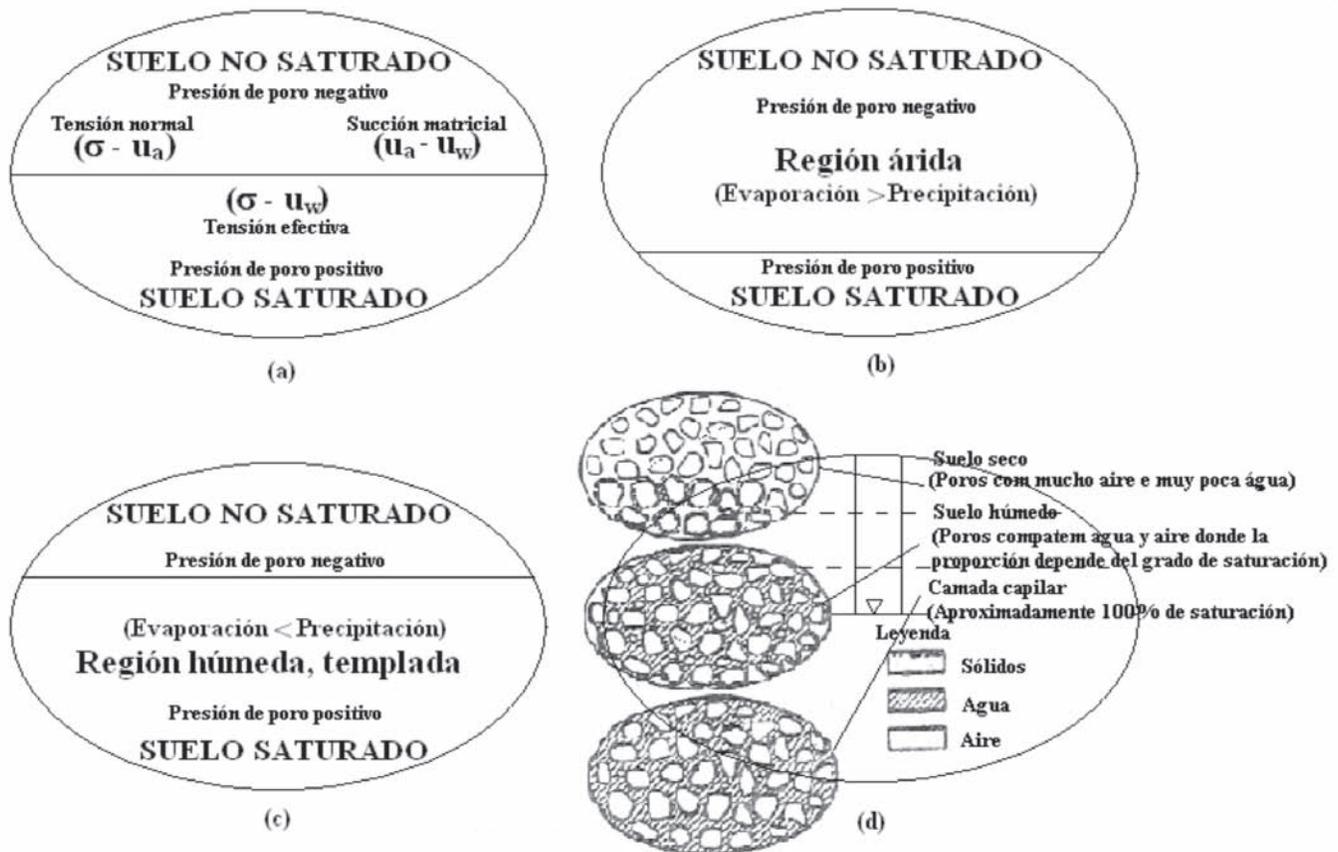


Figura 1. Componentes de un suelo para los problemas geotécnicos. a) Generalización; b) Situación en región árida; c) Situación en región templada y húmeda; d) Composición del medio poroso. (Modificado de Fredlund, 1995).

**Succión matricial**

La succión en suelos no saturados está compuesta por la succión matricial ( $S_m$ ) y la succión osmótica ( $S_{osm}$ ) y la suma de ambos componentes se denomina succión total ( $S_t$ ). En ese contexto, la magnitud de la succión total corresponde al trabajo total de las fuerzas de capilaridad, absorción y osmosis. La succión total, se representa en unidades de presión y puede ser expresa:

$$S_t = S_m + S_{osm} \tag{1}$$

La succión total, se define como la presión manométrica negativa, relativa a la presión externa de gas sobre el agua del suelo, que deberá ser aplicada a un reservorio de agua pura (a la misma cota y temperatura) de tal forma que se mantenga en equilibrio, a través de una membrana semipermeable (permite o flujo del agua, y no de solutos), entre el agua do reservorio y el agua del suelo. La Figura 2, ilustra los conceptos de succiones matricial, osmótica y total.

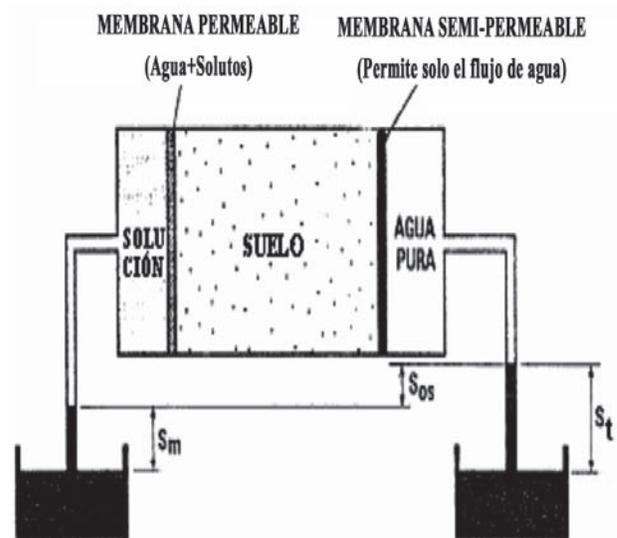
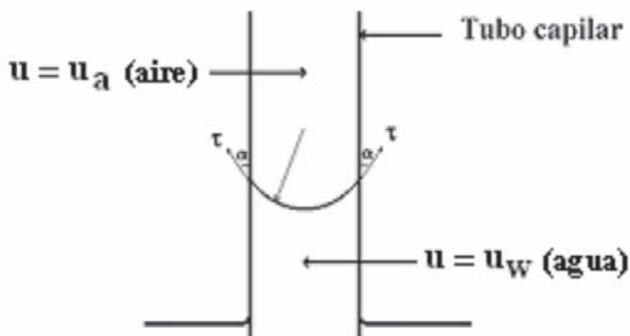


Figura 2. Representación de los conceptos de succiones matricial, osmótica y total.

La succión matricial ( $S_m$ ) es igual a la total cuando el agua del suelo es idéntica al agua padronizada (agua pura o solución con la misma composición del agua del suelo), quedando apenas el efecto de la matriz del suelo (capilaridad y adsorción). Los efectos matriciales provienen de las presiones desarrolladas por el menisco capilar y adsorción del agua debido a fuerzas ejercidas por las superficies de las partículas. Estas son cuantificadas de forma global debido a la dificultad de discriminarlas (**Jimenes Salas, 1993**).

La interfase aire-agua generada por los efectos capilares en el menisco que se forma entre las partículas de suelo adyacentes puede ser representado por el llamado modelo capilar mostrado en la Figura 3, (**Buckingham, 1907**).



**Figura 3.** Modelo capilar para análisis de la succión matricial (**Buckingham, 1907**).

A través del equilibrio de fuerzas en la interfase aire-agua se observa que la fuerza ejercida por el aire es igual a las contrarias ejercidas por el agua y que puede ser representada como:

$$(u_a - u_w) = \frac{2 \cdot \sigma_s}{r} \quad (2)$$

Donde:

r: radio do menisco [L];

$\sigma_s$ : tensión superficial del agua [MT<sup>-2</sup>].

Cuando en  $(u_a - u_w)$  la presión del aire ( $u_a$ ) corresponde a la atmosférica, la presión existente en las moléculas del agua es proporcional a la tensión superficial y al radio del capilar.

La tensión superficial ( $s_s$ ) es originada por la interacción de las fuerzas inter-moleculares producidas en las zonas de contacto entre las partículas del suelo, agua y aire, siendo esta responsable por la concavidad de la interface aire-agua y de la ascensión de la columna de agua en el tubo

capilar. El valor de tensión superficial es una característica del líquido y su valor depende de la temperatura, disminuyendo a la medida que la temperatura aumenta. El valor de tensión superficial del agua a 20°C es de 0,07275 N/m. (**Libardi, 1995**).

La succión osmótica ( $S_{osm}$ ), es igual a la succión total cuando el suelo se encuentra saturado, quiere decir, cuando la componente matricial no ocurre, actuando solo el efecto de la concentración de los solutos. En la Figura 2, el agua pura está en contacto con el suelo (con mayor concentración de solutos) a través de una membrana semipermeable que es permeable para las moléculas de agua y no para los solutos.

La mayor concentración del agua del suelo causa una atracción de las moléculas de agua pura y consecuentemente un flujo de estas a través de la membrana semipermeable.

El equilibrio es alcanzado cuando la presión hidrostática es suficiente para equilibrar las fuerzas osmóticas que producen el flujo de las moléculas del agua pura para el agua del suelo.

La importancia de la succión osmótica en Mecánica de Suelos parece estar más relacionada con los suelos dispersivos o expansivos, aunque se admita que el valor de la succión total corresponde casi integralmente al valor de succión matricial (**Fredlund & Rahardjo, 1993**). Consecuentemente, en la práctica es usual considerar a la succión del suelo como la succión matricial.

### Succión versus humedad del suelo

Las relaciones succión (total, matricial o osmótica) versus humedad (grado de saturación, humedad volumétrica o gravimétrica), son de importancia para la caracterización de los suelos no saturados. La representación gráfica de dicha relación es denominada “curva de retención del agua en el suelo” (**Croney & Coleman, 1960**).

Según **Fredlund et al.** (1994), la curva de retención del agua en el suelo puede ser definida como la variación de la succión con la capacidad de retención del agua en los macro y micro poros del suelo. Los valores corresponden al tipo de suelo, con una determinada densidad, y la naturaleza de esa relación está directamente asociada a la granulometría y mineralogía del suelo. De una forma general, la geometría de los poros, la magnitud y composición mineralógica de la fracción fina son determinantes en la posición relativa, forma e inclinación de la curva.

La curva de retención puede ser caracterizada con la finalidad de ecuacionarla. La caracterización es realizada a

partir de las trayectorias típicas obtenidas por procedimientos de secado o humedecimiento. La Figura 4 muestra características de las curvas de retención, obtenidas por humedecimiento y secado, pudiendo verificarse el fenómeno de histéresis (curvas no coincidentes por los procesos de humedecimiento y secado) asociada a las dos formas de obtención.

$\theta_s$ : humedad volumétrica de saturación para la curva obtenida por el proceso de secado;

$\theta'_s$ : humedad volumétrica de saturación para la curva obtenida por el proceso de humedecimiento;

$\theta_s - \theta'_s$ : contenido de aire residual;

Presión de entrada del aire: succión correspondiente al inicio de la entrada del aire;

$\theta_r$ : humedad volumétrica residual.

### Estado de tensiones

En suelos saturados la caracterización de las condiciones de tensión en que el suelo se encuentra sometido es dado por el conocido principio de las tensiones efectivas. Según **Terzagui** (1936) y posteriormente **Jennings & Burland** (1962); **Bishop & Blight** (1963) entre otros, las tensiones efectivas son las responsables por los efectos mecánicos en un determinado suelo, y son definidas por tensiones total y intersticial. Los efectos mecánicos son asociados a los cambios de volumen y resistencia al corte

de suelos saturados o secos. Este principio se representa con la siguiente expresión:

$$\sigma' = \sigma - u \quad (3)$$

Donde  $u$ , representa la presión neutra o presión en el fluido intersticial;  $\sigma$  es la tensión total o aplicada proveniente de la fuerza o sistema de fuerzas externas aplicadas y  $\sigma'$  la tensión efectiva de las fuerzas inter-granulares.

Sin embargo, a diferencia de los suelos saturados que son constituidos por dos fases (sólidos-aire o sólidos-agua), los suelos no saturados, están compuestos por tres fases (sólidos-aire-agua), y por lo tanto son necesarias nuevas formulaciones para el concepto de tensiones efectivas, debido a que los vacíos del suelo son ahora ocupados por las fases aire y agua. Las diferencias de presiones en esas fases dan origen a los efectos de succión debido a fuerzas de atracción entre las partículas (capilares y de adsorción).

Inicialmente, los conceptos de tensiones efectivas en suelos no saturados se crearon como extensión de los suelos saturados modificando la expresión (3). Diversos autores (**Bishop**, 1959; **Croney et al.**, 1958; **Lambe**, 1960; **Aitchison**, 1961; **Richards**, 1966, entre otros), tentaron retratar la tensión efectiva a partir de un equilibrio de fuerzas (actuando en el contacto inter-partículas). Para dichas ecuaciones la contribución de la succión, generalmente era afectada por un factor empírico con la finalidad de ser incluida dentro del estado de tensión (**Nelson & Miller**,

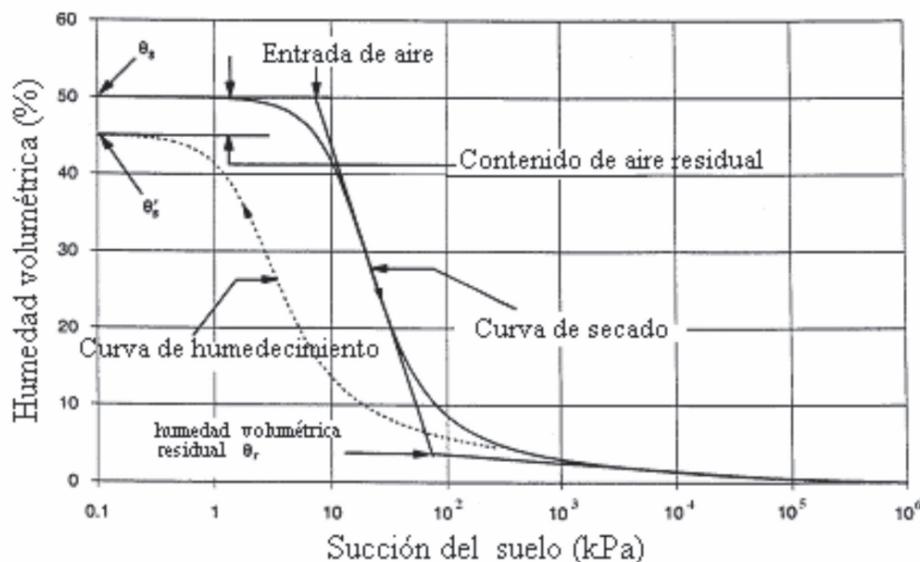


Figura 4. Elementos de la curva de retención según Fredlund *et al.* (1994).

1992). La Tabla 1, resume las diferentes ecuaciones y parámetros propuestos (Modificado de **Fredlund & Morgenstern, 1977**).

**Tabla 1.** Ecuaciones para tensiones efectivas en suelos no saturados (Modificado de **Fredlund & Morgenstern, 1977**)

Ecuación	Referencia
$\sigma' = \sigma - u_a + \chi (u_a - u_w)$	Bishop (1959)
$\sigma' = \sigma + p''$	Donald (1956)
$\sigma' = \sigma - \beta' \cdot u_w$	Croney, Coleman & Black (1958)
$\sigma' = \sigma + \Phi \cdot p''$	Aitchison (1961)
$\sigma' = \sigma + \beta p''$	Jennings (1961)
$\sigma' = \sigma - u_a + \chi_m (h_m + u_a) + \chi_s (h_s + u_a)$	Richards (1966)
$\sigma' = \sigma - u_a + \chi_m \cdot p''_m + \chi_s \cdot p''_s$	Aitchison (1973)
$\sigma' = \sigma - u_w \cdot Sr - (1 - Sr) \cdot u_a$	Öberg & Sällfors (1997)

Donde:

$\chi$ : parámetro relacionado con el grado de saturación;

$u_a, u_w$ : presión del aire y del agua respectivamente;

$p''$ : presión del agua debajo de la atmosférica;

$\beta'$ : factor de unión;

$\Phi$ : parámetro entre 0 y 1;

$\beta$ : factor estadístico para medir el área de contacto;

$\chi_m, \chi_s$ : parámetros de tensión efectiva para succión matricial y de soluto respectivamente;

$h_m, h_s$ : succión matricial y de soluto respectivamente;

$p''_m, p''_s$ : succión matricial y de soluto respectivamente;

Sr: grado de saturación.

Entre esas ecuaciones se destaca la expresión formulada por **Bishop** (1959), la misma que se encuadró de forma aceptable a los datos con que se contaban en esa época. Para esa formulación el parámetro  $\chi$  varía entre 0 para suelos secos y 1 para suelos saturados, y los valores intermedios dependen de la trayectoria de tensiones, de los ciclos de humedecimiento y secado y principalmente del grado de saturación (**Blight, 1967**). Posteriormente, estudios experimentales sobre ese tema, mostraron que la ecuación de las tensiones efectivas en suelos no saturados presentaba limitaciones

**Jennings & Burland** (1962), comprobaron que la expresión de **Bishop** (1959), no traducía el comportamiento (relación entre el índice de vacíos y las tensiones efectivas) de la mayoría de suelos con un grado de saturación inferior a un determinado valor crítico.

**Bishop & Blight** (1963), hacen una reflexión sobre la formulación y atribuyen que una variación de la succión no corresponde a una variación de la tensión intersticial. Otros autores cuestionaron la formulación cuando se consideran problemas de variación de volumen o resistencia al corte. De esta forma es evidenciada la dificultad de la obtención del factor  $\chi$  cuando es asociado a las deformaciones volumétricas, debido a que, caso ocurra reducción de las tensiones efectivas, podría ocurrir colapso o expansión del suelo, esto en función del término de la ecuación que sea reducido. Se solamente se reduce la succión podría ocurrir colapso, entretanto se solamente es reducida la tensión total, puede provocar expansión de la masa del suelo (**Alonso & Lloret, 1985**).

Frente a la dificultad de cuantificar el valor del parámetro  $\chi$ , **Aitchison** (1967), consideró que dada la variabilidad de ese parámetro solamente es posible obtener un valor apropiado de tensión efectiva considerándose una única trayectoria para cada término  $\sigma'$  y  $(u_a - u_w)$ , razón por la cual resulta innecesaria la cuantificación de  $\chi$ . Debido a las dificultades presentadas para cuantificar las tensiones efectivas en una ecuación como parámetro unitario independiente, la evaluación de las propiedades del suelo como variaciones de volumen y resistencia del suelo fueron realizadas utilizando los conceptos de variables de estado (**Matyas & Radakhrisma, 1968**).

Segundo **Porooshasb** (1961), las variables de estado se definen como las variables físicas e independientes del suelo, necesarias para definir su estado de tensiones. Para el caso de suelos arcillosos normalmente consolidados, se considera que el estado del suelo puede ser caracterizado por el estado de tensiones y el índice de vacíos (o humedad). De otro lado, para suelos no saturados son necesarios además del estado de tensiones y el índice de vacíos, el grado de saturación y la estructura del suelo (**Henkel, 1960; Porooshasb, 1961**).

El estado de un elemento de suelo puede ser representado gráficamente por un punto en un espacio tridimensional definido por un sistema de ejes coordenados que corresponden a los parámetros o variables de estado. Así, tensiones asociadas a cargas mecánicas o geostáticas pueden ser representadas por cualquiera de las variables ( $\sigma - u_a$ ) o ( $\sigma - u_w$ ). Las tensiones asociadas con la presión de poros se representa con la variable succión matricial ( $u_a - u_w$ ), y las

variaciones de volumen del suelo se asocian a los dos estados de tensiones independientes a través del índice de vacíos.

**Fredlund & Morguensten** (1977) establecieron que el estado de tensiones puede ser representado por dos variables independientes dentro de un grupo de tres variables tales como:

$$(4) \quad \begin{bmatrix} (\sigma_x - u_a) & \tau_{yx} & \tau_{zx} \\ \tau_{xy} & (\sigma_y - u_a) & \tau_{zy} \\ \tau_{xz} & \tau_{yz} & (\sigma_z - u_a) \end{bmatrix}$$

$$(5) \quad \begin{bmatrix} (\sigma_x - u_w) & \tau_{yx} & \tau_{zx} \\ \tau_{xy} & (\sigma_y - u_w) & \tau_{zy} \\ \tau_{xz} & \tau_{yz} & (\sigma_z - u_w) \end{bmatrix}$$

$$(6) \quad \begin{bmatrix} (u_a - u_w) & \tau_{yx} & \tau_{zx} \\ \tau_{xy} & (u_a - u_w) & \tau_{zy} \\ \tau_{xz} & \tau_{yz} & (u_a - u_w) \end{bmatrix}$$

Donde:

$\sigma_x, \sigma_y, \sigma_z$ : tensiones totales en las direcciones x, y, z.

De estos, los grupos de tensiones (4) y (6) son los más utilizados para los suelos no saturados.

### Formulaciones teóricas para los suelos no saturados

Formulaciones para representar las variaciones de volumen, resistencia al corte y flujo de agua en suelos no saturados son presentadas a continuación.

#### Cambios de volumen

Los cambios de volumen en suelos no saturados pueden ser asociados con las variables de estado de tensión usando relaciones apropiadas.

Según **Coleman** (1962), los cambios de volumen total (V) y del agua (V<sub>w</sub>) en un suelo sometido a cargas isotrópicas son dados por las expresiones siguientes:

$$-\frac{dV}{V} = C_{11}d(u_a - u_w) + C_{12}d(\sigma - u_a) \quad (7)$$

$$-\frac{dV_w}{V} = C_{12}d(u_a - u_w) + C_{22}d(\sigma - u_a) \quad (8)$$

Donde los índices C<sub>ij</sub> pueden ser positivos o negativos dependiendo del histórico de las tensiones y succiones a las que el suelo se encuentra sometido.

Posteriormente, **Bishop & Blight** (1963); **Burland** (1965), manifestaron algunas observaciones y sugirieron la adopción de  $(\sigma - u_a)$  y  $(u_a - u_w)$  como variables independientes relacionadas con el índice de vacíos (**Lloret & Alonso**, 1980).

**Fredlund** (1979), formula una relación constitutiva con base en la superficie de estado formada al plotar el logaritmo de las variables de tensión *versus* los parámetros de volumen (e) o humedad (Sr, w) representados por las siguientes ecuaciones:

$$e = e_o - Ct \cdot \log \frac{(\sigma - u_a)}{(\sigma - u_a)_o} - Cm \cdot \log \frac{(u_a - u_w)}{(u_a - u_w)_o} \quad (9)$$

$$w = w_o - Dt \cdot \log \frac{(\sigma - u_a)}{(\sigma - u_a)_o} - Dm \cdot \log \frac{(u_a - u_w)}{(u_a - u_w)_o} \quad (10)$$

Donde:

e, e<sub>o</sub>: Índices de vacíos final y inicial respectivamente;

Ct: Índice de compresión;

Cm: Índice de succión en función en función del índice de vacíos y succión matricial;

Dt: Índice del contenido de humedad en función de la tensión efectiva del suelo saturado;

Dm: índice del contenido de humedad en función de la succión matricial;

$(\sigma - u_a)$ ,  $(\sigma - u_a)_o$ : tensiones efectivas del suelo saturado final y inicial, respectivamente;

$(u_a - u_w)$ ,  $(u_a - u_w)_o$ : succión matricial final y inicial respectivamente.

#### Resistencia al corte

Según **Fredlund** (1995), la combinación de dos variables de estado como  $(\sigma - u_a)$ ,  $(\sigma - u_w)$  y  $(u_a - u_w)$  son capaces

de reproducir la resistencia de un suelo no saturado. Generalmente, son preferidos el uso de  $(\sigma - u_a)$  y  $(u_a - u_w)$  debido a que las variaciones en la tensión total y en la presión del agua pueden ser evaluadas separadamente y por que, con frecuencia, la presión del aire corresponde a la atmosférica que es igual a cero.

**Fredlund et al.** (1978), propusieron una ecuación para la resistencia al corte que se trata de una extensión del criterio de Mohr-Coulomb para suelos no saturados y que es representada por:

$$\tau = c' + (\sigma - u_a) \cdot \tan \phi' + (u_a - u_w) \cdot \tan \phi^b \quad (11)$$

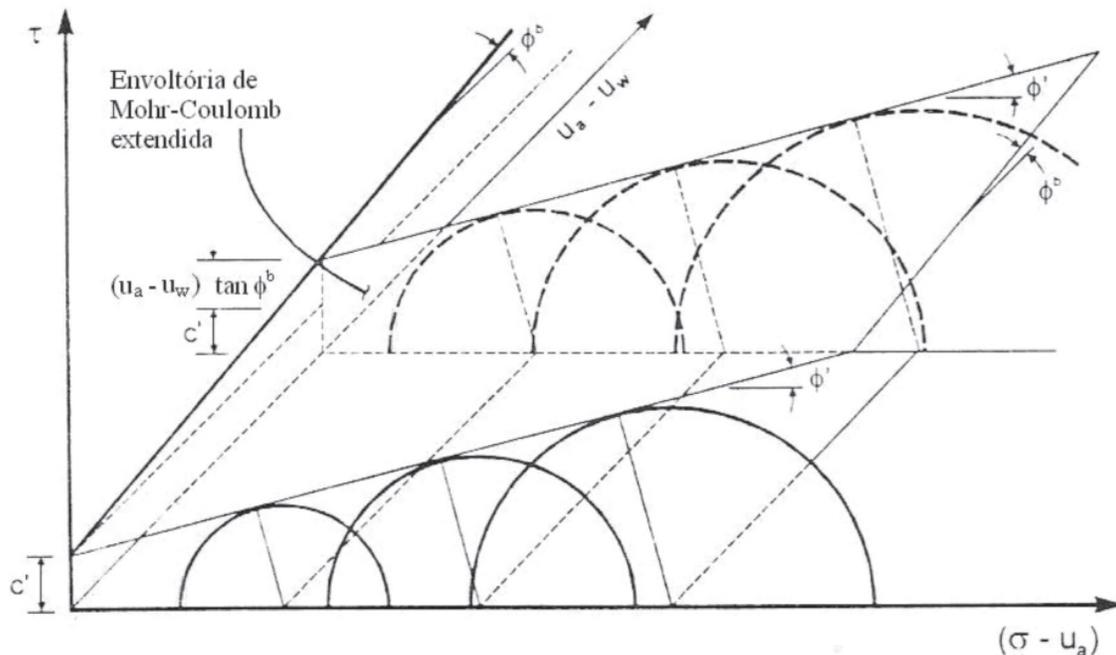
Donde  $c'$  es la intersección de la cohesión del suelo para la condición saturada,  $\phi'$  es el ángulo de fricción interna del suelo y  $\phi^b$  es el ángulo de fricción interna en relación a la succión matricial. La Figura (5) ilustra la propuesta original de **Fredlund et al.** (1978).

En esta figura se observa que, la influencia de la succión es representada por un tercer eje. De esta forma, un aumento de la succión  $(u_a - u_w)$  ocasionada por ejemplo, por la pérdida de humedad de un suelo sujeto a la evaporación, aumenta la magnitud de la resistencia en mayor o menor grado, según las características (peso específico, capacidad de retención del agua, textura, mineralogía, entre otros), del tipo de suelo.

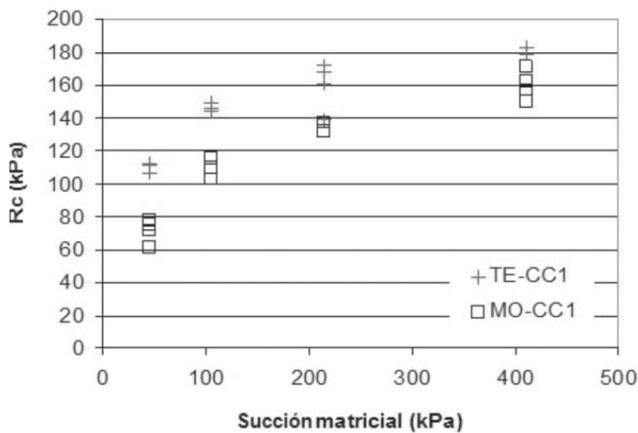
A partir de la ecuación (11), la influencia de la succión en la resistencia es reflejada por el aumento de la cohesión que se muestra lineal. Sin embargo, resultados experimentales por diversos autores (**Escarro & Sáez**, 1986 entre otros) muestran la no linealidad. **Fredlund et al.** (1987) admite la limitación de su propuesta inicial y propone una envolvente curva, con un tramo inicial determinado por el ángulo de fricción interna del suelo saturado. Se admite que, anteriormente a la retirada del agua (presión de entrada de aire) el principio de las tensiones efectivas es válido y el ángulo  $\phi^b$  es igual al ángulo de fricción interna drenado. Cuando la presión de aire es superada la resistencia es comandada por una relación no lineal con la succión, que tiende a un valor asintótico.

En la Figura (6) es mostrada la influencia de la succión en ensayos de resistencia en muestras de suelos no saturados realizados por **Alfaro Soto** (2004).

Para realización de estos ensayos se utilizaron para imposición de la succión, los métodos translación de ejes (TE) e osmótico (MO). En ambos casos, las succiones impuestas fueron de 45, 105, 215 y 410kPa. Después de la imposición de la succión, las muestras se llevaron a la ruptura en una máquina de compresión simple dentro de una sala climatizada (humedad relativa y temperatura controlada) para evitar cambios en la magnitud de la succión. Los



**Figura 5.** Envoltoria de resistencia para suelos no saturados (Fredlund et al., 1978)



**Figura 6.** Resultados de ensayos de compresión simple versus succión matricial, esta última obtenida por métodos de translación de ejes (TE) y osmótico (MO). (Alfaro Soto, 2004).

resultados muestran la evidencia del aumento de la resistencia ( $R_c$ ) con el aumento de la succión y que a diferencia de Fredlund *et al.* (1978), denotan una tendencia no lineal. Estudios sobre ese tema fueron presentados por diversos autores y entre ellos Vilar (2007), presenta un método para estimar la resistencia al corte en suelos no saturados.

### Flujo de agua en el suelo no saturado

El flujo de agua en suelo no saturado es el que ocurre en el suelo sujeto a cualquier condición de humedad menor que la de saturación; en la fase líquida de un suelo no saturado, tanto el agua como el aire se movilizan por la influencia de los potenciales presentes en el medio no saturado.

Se admite que el flujo no saturado es descrito por la ley de Darcy, con la diferencia de que la conductividad hidráulica no saturada  $K(u_a - u_w)$  (también representada por  $K(w)$ ), no es constante y que varía, predominantemente en función de la humedad ( $w$ ) y consecuentemente, de la succión matricial ( $u_a - u_w$ ).

Experiencias de Childs & Collis-George (1950), confirmaron que dicha ley puede ser aplicada satisfactoriamente a los suelos no saturados. La ley de Darcy para el flujo no saturado puede ser expresado como:

$$q = -K(u_a - u_w) \frac{\partial \phi}{\partial z} \quad (12)$$

Donde:

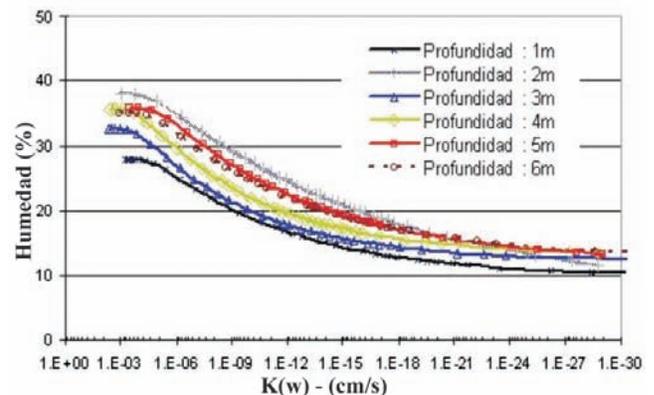
$K(u_a - u_w)$ : conductividad hidráulica no saturada (m/s);

$q$ : cantidad de flujo no saturado que equivale a la velocidad de descarga (m/s);

$z$ : coordenada de posición (m);

$\phi$ : potencial total do suelo no saturado (m).

Alfaro Soto *et al.* (2007), obtuvieron resultados de la conductividad hidráulica no saturada para un perfil de suelo situado en el interior del estado de São Paulo-Brasil. En estos estudios los autores utilizaron sondas TDR (*Time Domain Reflectometry*) para la medición de la humedad, instaladas a lo largo de un pozo de 7 metros de profundidad. Para determinación de  $K(w)$  fueron utilizados métodos indirectos que se apoyan en la curva de retención de agua del suelo. La Figura 7 muestra los resultados



**Figura 7.** Resultados de conductividad hidráulica no saturada ( $K(w)$ ) versus humedad, obtenida por métodos indirectos a partir da curva de retención de agua en el suelo (Alfaro Soto *et al.*, 2007).

Como observado en esta figura,  $K(w)$  es influenciado por la humedad del suelo, así cuando el suelo se encuentra saturado (succión igual a cero)  $K(w)$  pasa a tener un valor máximo y constante. Por otro lado, cuando la humedad del suelo disminuye por efectos naturales (evaporación, drenaje, etc.) o antrópicos (bombeo del agua subterránea, desmatamiento, etc.) aumenta la succión del suelo, consecuentemente  $K(w)$  es reducido drásticamente, pues es introducido un nuevo potencial que incluye los efectos matriciales e que influencia en el movimiento del agua del suelo. La presencia de este potencial significa reducción en la velocidad de percolación debido a que el drenaje ocurre inicialmente en los poros de mayor tamaño, así el agua remaneciente se concentra en forma de meniscos (efectos capilares y absorción) en contacto con las partículas. Luego, el flujo es transferido para los poros de menor tamaño siendo necesarios gradientes muy grandes y tiempos extensos para detectar el movimiento del agua en el suelo.

### Consideraciones finales

Se mostró una síntesis de una mecánica de suelos con las más relevantes teorías y formulaciones que son una extensión de la mecánica de suelos tradicional, desarrollada para retratar el comportamiento de suelos no saturados.

Tal vez la demora en su evolución se deba a que las teorías más genéricas se tengan enfrentado con grandes dificultades debido a la complejidad de representaciones matemáticas necesarias o por la complejidad de las determinaciones experimentales *in situ* o en laboratorio (Vilar, 1997). Adicionalmente, el hecho de que los suelos no saturados ocurren principalmente en regiones menos desarrolladas.

No obstante, son evidentes los esfuerzos obtenidos estos últimos años deseando mejorar los diseños de las obras inclusive de naturaleza ambiental. Aún no es suficiente, siendo necesaria su amplia difusión en universidades, congresos, simposios, etc., principalmente en los países donde su uso y puesta en práctica venga a beneficiarlos.

### Bibliografía

- Alfaro Soto, MA.** 2004. Comparação entre métodos de imposição e de controle da sucção em ensaios com solos não saturados. São Carlos, 2004. 206p. Tese (Doctorado). Escola de Engenharia de São Carlos - Universidade de São Paulo – USP.
- Alfaro Soto, MA, Chang HK.** 2007. Utilización de TDR en estudios de la dinámica del agua en suelo no saturado. In: 8° Iberoamerican Conference on Mechanic Engineering, Cuzco - Peru. Proceedings of the 8° Iberoamerican Conference on Mechanic Engineering. CD-ROM.
- Alfaro Soto, MA, Kumayama, DM, Chang HK.** 2007. Calibração de um Reflectômetro para estudos do Fluxo de Água em Solo Não Saturado. Geociências (São Paulo), v. 26, p. 357-368.
- Aitchison, GD.** 1961. Relationships of moisture stress and effective stress functions on unsaturated soils. Pore pressure and Suction in Soils, Butterworths, London.
- Aitchison, GD.** 1967. The separate roles of site investigation, quantification of soil properties and selection of operational environment in the determination of foundation design on expansive clay. 3<sup>rd</sup> Asian Conf. Soil Mechanics Found., Eng., Haifa.
- Alonso, EE. y Lloret, A.** 1985. Comportamiento de suelos parcialmente saturados. Revista de Obras Públicas, p. 435-461.
- Bishop, AW.** 1959. The principle of effective stress. Tecknisk Ukeblad 106(39), p. 859-863.
- Bishop, AW, Blight, GE** 1963. Some aspects of effective stress in saturated and partly saturated soils. Géotechnique, 13(3): 177-198.
- Blight, G.E.** 1967. Effective stress evaluation of unsaturated soils. Proceedings of the American Society of Civil Engineers, v. 93.
- Buckingham, E.** 1907. Studies on the movement of soil moisture. U.S. Dept. Agr. Bur. Soils, Bul., 38 (apud Alonso *et al.*, 1987).
- Burland JB.** 1965. Some aspects of the mechanical behavior of partly saturated soil. In Moisture Equilibria and Moisture Changes in the Soils Beneath Covered Areas. Australia: Butterworths p. 270-278.
- Childs, EC & Collins-George, N.** 1950. The permeability of porous materials. Roy. Soc. Proc., London, 201, p. 392-405.
- Coleman, JD.** 1962. Stress/strain relations for partly saturated soils. Geotechnique, 12(4): 348-350.
- Croney, D, Coleman, JD, Black, WPM** 1958. Studies of the movement and distribution of water in soil in relation to highway design and performance. Highway Research Board, Special Report 40, Washington D.C., p. 226-252.
- Croney, D. & Coleman, JD.** 1960. Pore pressure and suction in soil. Conference on Pore Pressure and Suction in Soils. London, Proceedings..., p. 31-37.
- Escario, V, Saez J.** 1986. The Shear Strength of Partly Saturated Soils. Geotechnique 36. No. 13.
- Fredlund, DG, Morgenstern, NR.** 1977. Stress State Variables for Unsaturated Soils. Journal of Geot. of Div. ASCE., v. 103 GT5, p. 447-466.
- Fredlund, DG et al.** 1978. The shear strength of unsaturated soil. Canadian Geotechnical Journal, v. 15, p. 313-321.
- Fredlund, DG.** 1979. Appropriate concepts and technology for unsaturated soils. Canadian. Geotech. Journal. 16, p. 121-139.
- Fredlund, DG, Rahardjo H, Gan, JKM.** 1987. Non linearity of shear strength envelope of unsaturated soils. In. Proceedings, 6<sup>th</sup> International Conference on Expansive Soils, New Delhi, v. 1, p. 49-54.
- Fredlund, DG, Rahardjo, H.** 1993. Soil mechanics for unsaturated soils. New York, John Wiley & Sons, Inc. 1993. p. 1-6.
- Fredlund, DG, Xing, A. & Huang, S.** 1994. Predicting the permeability functions for insaturated soil using the soil-water characteristic curve. Canadian Geotechnical Journal, 31(4): p. 533-546.
- Fredlund, DG.** 1995. The scope of unsaturated soil mechanics: an overview. In: Proceedings of the First International Conference on Unsaturated Soils, UNSAT 95, Paris-França, v. 3.
- Henkel, DJ.** 1960. The relationships between the effective stresses and water content in saturated clays. Géotechnique, Vol. 10, No 2.
- Jennings, JEB, Burland, JB.** 1962. Limitations to the use of effective stresses in partly saturated soils. Géotechnique, v.12, No. 84, (SM5), p. 125-144.
- Jimenes, SJ.** 1993. "Hacia una mecánica de Suelos no Saturada", Revista Ingeniería Civil v. 88 -Ira parte. p. 145-164.

- Lambe, TW.** 1960. A mechanistic picture of shear strength in clay. Research Conf. On Shear Strength of Cohesive Soils, Boulder, Colorado, p. 555-580.
- Libardi, PL.** 1995. Dinâmica da Água no Solo. Edição do Autor, ESALQ, Depto. de Física e Meteorologia, Piracicaba, p. 497.
- LLoret, A, Alonso, EE.** 1980. State surfaces for partially saturated soils. Proc. 11<sup>th</sup> ICSMFE., San Francisco, 2, p. 557-562.
- Matyas, EL, Radhakrishna, HS.** 1968. Volume change characteristics of partially saturated soil. Géotechnique 18 (4): 432-448.
- Nelson, JD, Miller, DJ.** 1992. Expansive soils- Problems and practice in foundation and pavement engineering. John Wiley & Sons, Inc., p.,259.
- Öberg, AL, Sällfors, GA.** 1997. Determination of shear strength parameters of unsaturated silts and sands based on the water retention curve. Geotechnical Testing Journal. v. 20, No. 1.
- Parreira, AB, Takeda, M. De C., Luz MP.** 2004. Avaliação da influência do período de imersão nos resultados do ensaio CBR de solos tropicais. V Simposio Brasileiro de Solos Não Saturados. São Carlos, v. 1, p. 383-388.
- Pereira, EM, Pejon, OJ.** 2004. Características de Materiais Não Saturados Expansivos mediante ensaios com soluções salinas. In: V Simpósio Brasileiro de Solos Não Saturados, 2004, São Carlos. V Simpósio Brasileiro de Solos Não Saturados, São Carlos, v. 1, p. 471-476.
- Poorooshasb, HB.** 1961. The properties of soil and other granular media in simple shear. Ph.D. Thesis, University of Cambridge.
- Reis, RM, Vilar, OM.** 2004. Resistência ao cisalhamento de dois solos residuais de gnaiss não saturados. V Simpósio Brasileiro de Solos não Saturados. v. 1, p. 109-114.
- Richards, BG.** 1966. The significance of moisture flow and equilibria in unsaturated soils in relation to the design of engineering structures built on shallow foundations in Australia. Symp. on Permeability and Capillarity of Soils, ASTM STP 417., Atlantic City.
- Santos, CRD, Vilar, OM.** 2004. Análise paramétrica da estabilidade de taludes em solos não saturados: a influência do tipo de solo. V Simpósio Brasileiro de Solos não Saturados, São Carlos. v. 1, p. 181-188.
- Teixeira, RS, Belincanta, A, Lopes, FF, Gutierrez, NHM, Branco, CJMC.** 2004. Avaliação do colapso do solo da camada superficial da cidade de Londrina/PR. V Simpósio Brasileiro de Solos Não Saturados, São Carlos. v. 1, p. 495-500.
- Terzaghi, K.** 1936. The shearing resistance of saturated soils and the angle between the planes of shear, In: International Conference on Soil Mechanics and Foundation Engineering, 1<sup>o</sup>, Proceedings....., Cambridge, International Society for Soil Mechanics and Foundation Engineering, 1: p. 54-56.
- Vilar, OM.** 1997. Introdução à mecânica de solos não saturados. Notas de aula.
- Vilar, OM.** 2007. An Expedite Method to Predict the Shear Strength of Unsaturated Soils. Soils and Rocks, v. 30, p. 51-61.
- Wolle, CM.** 2004. Algumas considerações sobre a utilização dos novos conceitos de comportamento dos solos não saturados em projetos de engenharia. 5<sup>o</sup> Simpósio Brasileiro de Solos Não Saturados, 25 a 27 de Agosto, São Carlos –SP, Brasil. v.1, p. 635-641.

Recibido: enero 19 de 2007

Aceptado para su publicación: diciembre 9 de 2008

# ESTIMACIÓN DEL ESPECTRO MULTIFRACTAL PARA SERIES DE PRECIPITACIÓN HORARIA EN LOS ANDES TROPICALES DE COLOMBIA

Por

Jesús D. Gómez<sup>1</sup> & Germán Poveda<sup>2</sup>

## Resumen

**Gómez, J. D. & G. Poveda:** Estimación del espectro multifractal para series de precipitación horaria en los Andes Tropicales de Colombia. Rev.Acad.Colomb.Cienc. **32**(125): 483-502, 2008. ISSN 0370-3908.

El espectro multifractal,  $f(\alpha)$ , fue estimado para registros de precipitación horaria en 47 estaciones ubicadas en los Andes tropicales de Colombia mediante cinco metodologías distintas. Cada una de las metodologías fue aplicada a series provenientes de medidas binomiales, de las cuales se conoce su espectro multifractal teórico. Los resultados permiten concluir que: (i) todas las metodologías presentan resultados satisfactorios en la estimación de  $f(\alpha)$  para las medidas binomiales sintéticas (ii) las cinco metodologías presentan estimaciones muy diferentes del espectro para las series temporales de lluvia a escala horaria (iii) se encuentran valores muy variables de la fuerza multifractal,  $\Delta\alpha$ , variando desde 0.66 a 7.4 (iv) el exponente de Renyi,  $\tau(q)$ , puede ser representado por un modelo simple, basado en la generalización de un modelo de cascadas multiplicativas (v) No se encontró una relación clara entre los parámetros encontrados en este trabajo y la elevación. Finalmente se presenta una discusión en torno a las características de los modelos y los problemas de estimación en series de lluvia a resolución horaria.

**Palabras clave:** Multifractales, Modelación, Precipitación, Andes tropicales, Colombia.

1 Escuela de Geociencias y Medio Ambiente, Universidad Nacional de Colombia, Medellín, Colombia. New Mexico Institute of Mining and Technology, Socorro, NM 87801, U. S. Correo electrónico: jdgomez@nmt.edu

2 Escuela de Geociencias y Medio Ambiente, Universidad Nacional de Colombia, Medellín, Colombia. Correo electrónico: gpoveda@unal.edu.co

### Abstract

The multifractal spectrum,  $f(\alpha)$ , was estimated for hourly rainfall records of 47 gauges located at the tropical Andes of Colombia using five different methodologies. All of these methodologies were applied to binomial measures, which have a well known theoretical multifractal spectrum. From the results, it is possible to conclude the following: (i) all methodologies have satisfactory results in the estimation of  $f(\alpha)$  for the synthetic binomial measures; (ii) the five methodologies showed different estimations for the rainfall time series spectrum; (iii) the multifractal strength,  $\Delta\alpha$ , showed a wide range of values, varying from 0.66 to 7.4; (iv) the Renyi exponent,  $\tau(q)$ , could be represented by a simple two-parameter model, which is based on a generalized version of the multiplicative cascade model; and (v) no clear relationship between the parameters used in this research and elevation was found. Finally, a discussion about the characteristics of the models and problems in the spectrum estimation for hourly time series is presented.

**Key words:** Multifractals, Modelling, Precipitation, Tropical Andes, Colombia.

### 1. Introducción

En el presente trabajo se pretende aplicar metodologías orientadas al estudio de las propiedades de escalamiento para series de tiempo de precipitación en los Andes tropicales de Colombia. En trabajos anteriores se han realizado estudios similares para otras regiones (Olsson, 1995; Harris et al., 1996; Svensson et al., 1996; de Lima & Grasman, 1999; Sivakumar, 2001; Kantelhardt et al., 2003), sin embargo este tipo de estudios no son fácilmente generalizados y cambian notablemente con las características de la región analizada, y, en particular, presentan una relación directa con el origen de la lluvia, ya sea convectivo o advectivo, siendo común la presencia de multiescalamiento en campos con génesis convectiva. En el caso de los Andes tropicales de Colombia se cuenta con un precedente de investigación orientada al entendimiento de la lluvia a partir de las series de precipitación horaria utilizadas en este trabajo, en el cual se realizó un diagnóstico del ciclo diurno promedio a largo plazo, así como durante los diversos meses del año, y la influencia que en éste se presenta a partir de la variabilidad altitudinal y ubicación entre los valles interandinos de los ríos Cauca y Magdalena. Además, se evaluó su comportamiento durante la ocurrencia de fenómenos macroclimáticos y astronómicos como el ENSO, la oscilación de Madden-Julian, las ondas tropicales del este (4-8 días), las fases de la luna y los huracanes en los océanos Atlántico y Pacífico, y se aplicaron teorías de escalamiento y colas pesadas, análisis de intermitencia, intensidades máximas y valores extremos (Agudelo et al., 2001; Álvarez & Toro, 2001; Vieira & Moreno, 2001; Poveda et al., 2005).

En las últimas cuatro décadas, la comunidad científica ha tratado de entender el comportamiento de la

atmósfera, y en particular de campos de precipitación a nivel espacial y temporal, con especial énfasis en el primero, por medio de modelos de cascada, que admiten una fuerte anisotropía e intermitencia de las escalas en dichos campos y por ende la presencia de campos multifractales altamente variables (Schertzer & Lovejoy, 1996; Tessier, 1993). El problema, es básicamente entender cómo y en qué escalas se presenta el proceso de cascada de los campos de precipitación.

La base teórica de la modelación de campos atmosféricos como medidas multifractales está fundada en que los flujos de agua y energía en la atmósfera son representados por procesos de cascadas multiplicativas (Peitgen et al., 1992), los cuales transfieren dichas cantidades de escalas mayores a menores (Gupta & Waymire, 1993; Olsson, 1995; Schertzer & Lovejoy, 1996; Schertzer et al., 1997). Esta hipótesis se fundamenta en teorías y experimentos. Teóricamente, el fenómeno de la cascada puede ser deducido a partir de las ecuaciones de la turbulencia hidrodinámica, las cuales se asumen aproximadamente válidas para la atmósfera, y empíricamente, las observaciones muestran una estructura jerárquica en los campos de lluvia (Olsson, 1995). El proceso de cascada es definido como remolinos que se parten en sub-remolinos, de los cuales cada uno recibe una parte de la masa (energía) del remolino padre. De esta forma, la parte principal del flujo es concentrado en porciones cada vez más pequeñas del dominio, dando como resultado un campo con extrema variabilidad e intermitencia.

La invarianza de las propiedades de escala y la multifractalidad en los procesos de la lluvia sobre una gran cantidad de escalas, permite el entendimiento de sus fluctuaciones fuertes e irregulares, lo cual no puede ser totalmente captado a partir de otras descripciones de la

dinámica compleja de la lluvia tropical. El principal objetivo de la teoría multifractal y sus modelos en el caso de la lluvia, es darle uso como una herramienta para la predicción en cuencas no instrumentadas (**Predictions in Ungaged Basins, PUB**) (**Schertzer et al.**, 2003) y la producción sintética de series de precipitación de alta resolución, las cuales puedan ser utilizadas en aplicaciones hidrológicas como la de los modelos lluvia-escorrentía, la generación de curvas de intensidad-frecuencia-duración (**IDF**), y el diseño hidrológico (**Flores**, 2004; **Douglas & Barros**, 2003).

La organización de este artículo es la siguiente: en la sección 2 se presenta una descripción de algunas herramientas para el estudio de las propiedades de escalamiento en series de tiempo y las metodologías para la estimación del espectro multifractal. En la sección 3 se describe la base de datos utilizada. Finalmente, en las secciones 4 y 5 se presentan el análisis de los resultados obtenidos y las conclusiones del estudio, respectivamente.

## 2. Análisis multifractal

En las últimas dos décadas se han desarrollado una gran variedad de metodologías orientadas a identificar comportamientos fractales, multifractales y caóticos en series de tiempo. Parte de dicho comportamiento en series de tiempo puede ser captado por medio de una descripción estadística estándar, como el espectro de potencias. Además, la función distribución de probabilidad empírica provee información importante acerca del comportamiento fractal y su tipo, monofractal o multifractal (**Sivakumar**, 2001).

**2.1. Función de distribución de la densidad de probabilidad empírica.** Si para umbrales de intensidad altos,  $x$ , la cola de la función de distribución de probabilidad de la intensidad de la lluvia,  $X$ , sigue una ley potencial de la forma

$$P(X > x) \propto x^{-q_{cr}}, \quad (1)$$

se dice que la serie se caracteriza por intermitencia hiperbólica, lo cual se da generalmente, pero no necesariamente, en un proceso multifractal (**Harris et al.**, 1996). Estas distribuciones se derivan cuando se simula la intermitencia y el proceso de cascada de un fenómeno, conservando la masa en el promedio (“cascadas canónicas”), las cuales no necesariamente implican la divergencia de los momentos (e.g., las distribuciones lognormales poseen colas hiperbólicas para

tamaños de muestra finitos, pero poseen momentos convergentes), pero pueden ser vistas como un soporte complementario en el uso de la caracterización multifractal y modelamiento de la lluvia a partir de “cascadas canónicas” (**Harris et al.**, 1996).

En este caso  $q_{cr}$  es el exponente de probabilidad crítico, el cual refleja los estadísticos de valores extremos e indica la escala crítica a partir de la cual los momentos son divergentes. Además, en la literatura se sugiere que valores de  $q_{cr} < 2$  indican que modelos monofractales son suficientes para caracterizar el comportamiento de la serie de tiempo, mientras que modelos multifractales son adecuados si  $q_{cr} > 2$  (**Olsson**, 1995; **de Lima & Grasman**, 1999; **Sivakumar**, 2001).

**2.2. Espectro de potencias.** El espectro de potencias describe la variación de la potencia de una señal (energía por unidad de tiempo) respecto a la frecuencia. Dicho espectro es una herramienta estándar para la investigación del comportamiento fractal de series geofísicas (**Olsson**, 1995). Si el espectro de potencias obedece a una ley de la forma

$$E(f) \propto f^{-\beta}, \quad (2)$$

donde  $f$  es la frecuencia y  $\beta$  es el exponente espectral, dicho exponente puede ser un indicativo de la ausencia de una escala de tiempo característica en el rango de la ley de potencias, y, por tanto se puede asumir la presencia de un comportamiento (multi-)fractal. Además, el espectro de potencias es particularmente usado para el estudio de oscilaciones de un proceso (**Sivakumar**, 2001). En general, para un proceso aleatorio, el espectro de potencias oscila aleatoriamente alrededor de un valor constante, indicando que una frecuencia no explica más de la varianza de la secuencia que otra. Para secuencias periódicas o casi-periódicas, sólo existen picos a ciertas frecuencias, el ruido medido le agrega un piso continuo al espectro. De esta manera, en el espectro, la señal y el ruido son fácilmente diferenciados. Señales caóticas pueden tener líneas espectrales muy notorias, pero siempre en la ausencia de ruido ellas serán una parte continua del espectro (**Sivakumar**, 2001).

El exponente  $\beta$  puede ser entendido como un indicador de suavidad de la serie de tiempo, la cual es directamente proporcional a él. La lluvia producida directamente por un proceso de cascadas multiplicativas, en general, es poco suave y debería poseer exponente  $\beta < 1$  (cascadas ilimitadas), por lo tanto desde el punto de vista del análisis multifractal estos campos son mas convenientes, pues su descripción es completamente especificada por la función de momentos y  $q_{cr}$  (**Harris et**

al., 1996). Los campos con  $\beta > 1$  (cascadas limitadas), generalmente, no muestran escalamiento de momentos y su análisis puede ser a partir de fluctuaciones (Tessier, 1993).

**2.3. Definición del espectro multifractal y metodologías para su estimación.** Considere la proporción de cajas no vacías en una  $\delta$ -partición de  $\Omega$  que tienen masa en el rango de  $R^D$  ( $\delta^{\alpha+d\alpha}, \delta^\alpha$ ). Esta proporción está dada por:

$$\frac{\text{Número de cajas con masa en } (\delta^{\alpha+d\alpha}, \delta^\alpha)}{\text{Número de cajas no vacías}} \quad (3)$$

donde el denominador es  $\sim \delta^D$  y  $D$  es la dimensión del conjunto donde la masa se encuentra distribuida (Peitgen et al., 1992). Definiremos una medida multifractal como aquella para la cual el numerador obedece: número de cajas con masa en  $(\delta^{\alpha+d\alpha}, \delta^\alpha) \sim \delta^{-f(\alpha)}$ . El exponente varía con  $\alpha$  y se conoce como el *espectro multifractal*, o *espectro de singularidades*, o sólo como *la curva*  $f(\alpha)$ . Las medidas multifractales se caracterizan por su curva  $f(\alpha)$ .

Aunque todavía no existe una definición estándar de lo que es un multifractal, podría decirse que una medida multifractal es aquella que tiene una curva  $f(\alpha)$  bien definida. La curva  $f(\alpha)$  puede entenderse en términos de la noción familiar de una densidad, aunque las medidas multifractales en sí mismas son demasiado singulares para tener densidades. Un rasgo básico de esta curva, es que permite identificar comportamientos de leyes potenciales del tipo  $\delta^\alpha$  con respecto a cambios en el tamaño del pixel,  $\delta$ , mientras otros tipos de dependencia para  $\delta$  (por ejemplo  $e^{-\delta}$ ) son ignorados. Aquí la curva  $f(\alpha)$  es una lente de leyes potenciales para observar medidas singulares.

A continuación se presentan algunos de los métodos para la estimación del espectro  $f(\alpha)$  de una medida multifractal, acompañados de sus pasos de implementación. En este trabajo se implementó cada uno de estos métodos para series de series de lluvia, partiendo de un proceso de verificación de los algoritmos planteados, el cual se llevó a cabo a partir de una expresión teórica del espectro de la medida binomial.

Sea  $\mu$  una medida binomial generada a partir de relaciones de partición  $p : (1-p)$ , con  $p < 1$ , para cada orden  $q$  ( $-\infty < q < \infty$ ), se tiene (véase la referencia Peitgen et al. (1992)):

$$\alpha(q) = - \left[ \frac{p^q}{p^q + (1-p)^q} \log_2(p) + \frac{(1-p)^q}{p^q + (1-p)^q} \log_2(1-p) \right] \quad (4a)$$

$$f(\alpha(q)) = - \left[ \frac{p^q}{p^q + (1-p)^q} \log_2 \left( \frac{p^q}{p^q + (1-p)^q} \right) + \frac{(1-p)^q}{p^q + (1-p)^q} \log_2 \left( \frac{(1-p)^q}{p^q + (1-p)^q} \right) \right] \quad (4b)$$

Dadas las ecuaciones (4a) y (4b), es posible estimar el error cometido en la estimación del espectro. Vale la pena tener presente que en todos los casos es posible estimar el error cometido al estimar  $\alpha$  y  $f(\alpha)$  en función de  $q$ .

**2.3.1. Método de los momentos.** Este método está basado en el formalismo multifractal, y tiene como una de sus hipótesis fundamentales la presencia de "cascadeo" multiplicativo. En series temporales de variables geofísicas, sólo se tiene la medida en una iteración  $k$ , si hacemos la analogía con una medida binomial. El proceso consiste en la reconstrucción de cualquier iteración anterior,  $h < k$ , a partir de agregaciones en intervalos de tamaño  $2^{-h}$ .

El método de los momentos se basa en una cantidad llamada *función de partición*, debido a su analogía con la función de partición en la teoría del equilibrio termodinámico (Peitgen et al., 1992). Esta función se define como

$$\chi_q(\epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q, \quad q \in \mathbb{R}, \quad (5)$$

donde  $N(\epsilon)$  es el número de cajas de tamaño  $\epsilon$  necesarias para cubrir la medida.

En el caso general, se puede reescribir la medida  $\mu_i$  en términos de las cajas de tamaño  $\epsilon$  como  $\mu_i = \epsilon_i^\alpha$ , luego  $\chi_q(\epsilon) = \sum_{i=1}^{N(\epsilon)} (\epsilon_i^\alpha)^q$ . Motivados por los resultados del análisis de la medida binomial, se denota por  $N_\epsilon(\alpha) d\alpha$  al número de cajas, entre las  $N(\epsilon)$  cajas totales, con un exponente de Hölder que satisface  $\alpha < \alpha_i < \alpha + d\alpha$ , además, se asume que existen constantes  $\alpha_{min}$  y  $\alpha_{max}$  tales que  $0 < \alpha_{min} < \alpha < \alpha_{max} < \infty$ , y que  $N_\epsilon(\alpha)$  es continuo. Luego, la contribución del subconjunto de cajas con  $\alpha_i$  entre  $\alpha$  y  $\alpha + d\alpha$  a  $\chi_q(\epsilon)$ , es  $N_\epsilon(\alpha) (\epsilon_i^\alpha)^q d\alpha$ .

En lugar de hacer la contribución por cada caja  $i$  separadamente, se realiza una integración sobre  $d\alpha$  adicionando la contribución cuyo exponente de Hölder se encuentre entre  $\alpha$  y  $\alpha + d\alpha$

$$\chi_q(\epsilon) = \int N_\epsilon(\alpha) (\epsilon_i^\alpha)^q d\alpha \quad (6)$$

Si  $N_\epsilon(\alpha) \sim \epsilon^{-f(\alpha)}$ , se tiene

$$\chi_q(\epsilon) = \int \epsilon^{q\alpha - f(\alpha)} d\alpha. \quad (7)$$

En el límite  $\epsilon \rightarrow 0$ , la contribución dominante en la integral viene de los valores de  $\alpha$  cercanos al valor que minimiza el exponente  $q\alpha - f(\alpha)$ . Si  $f(\alpha)$  es diferenciable, la condición necesaria para la existencia de un extremo es

$$\frac{\partial}{\partial \alpha} q\alpha - f(\alpha) = 0. \quad (8)$$

Para un valor dado de  $q$ , el extremo ocurre para el valor  $\alpha = \alpha(q)$  que satisface

$$\frac{\partial}{\partial \alpha} \Big|_{\alpha=\alpha(q)} = q \quad (9)$$

y este extremo es un mínimo sólo si

$$\frac{\partial^2}{\partial \alpha^2} \Big|_{\alpha=\alpha(q)} > 0. \quad (10)$$

Teniendo en cuenta sólo la contribución dominante en la ecuación (7) e introduciendo la transformada de Legendre (**Courant & Hilbert, 1953**)—

$$\alpha(q) = \frac{\partial \tau(q)}{\partial q} \quad (11a)$$

$$\tau(q) = q\alpha(q) - f(\alpha(q)) \quad (11b)$$

se encuentra

$$\chi_q(\epsilon) \sim \epsilon^{\tau(q)}. \quad (12)$$

Finalmente, es fácil estimar  $f(\alpha)$  a partir de las ecuaciones (11a) y (11b).

*Algoritmo de cálculo del método de los momentos.* A continuación se presenta el algoritmo para el cálculo del espectro multifractal de series de tiempo a partir del método de los momentos, el cual está basado en la teoría descrita anteriormente.

1. Se cubre la medida con  $N(\epsilon)$  cajas de tamaño  $\epsilon$  ( $[B_i(\epsilon)]_{i=1}^{N(\epsilon)}$ ), determinando la medida correspondiente a cada caja,  $\mu_i = \mu(B_i(\epsilon))$ .
2. Se estima la función de partición a partir de la ecuación (5) para varios valores de  $\epsilon$ .
3. Se verifica que los gráficos  $\log_{10}(\chi_q(\epsilon))$  vs.  $\log_{10}(\epsilon)$  sean líneas rectas. Solo si esto se cumple, la ecuación (12) es válida.
4. Se encuentra  $f(\alpha)$  con la transformada de Legendre (véanse las ecuaciones (11a) y (11b))

*2.3.2. Método de Chhabra y Jensen.* A finales de los 80's, Ashvin Chhabra y Roderick Jensen propusieron una metodología para la determinación directa de  $f(\alpha)$  en medidas multifractales basados en un método canónico para el cálculo de promedios termodinámicos (**Chhabra & Jensen, 1989; Chhabra et al., 1989**). La característica fundamental del método, aplicado a datos experimentales, es que no necesita el cálculo de

las dimensiones generalizadas,  $D_q$ , ni el uso de la transformada de Legendre, mencionadas anteriormente.

El formalismo multifractal, a través de la relación existente entre las dimensiones generalizadas,  $D_q$ , y el espectro de singularidades,  $f(\alpha)$ , provee una adecuada descripción de las propiedades de escalamiento de una medida multifractal. Sin embargo, esta relación es válida desde el punto de vista teórico, pero no necesariamente en la práctica, debido básicamente a dos razones:

- Los gráficos log-log involucrados en la estimación de los  $D_q$ , en vez de mostrar un comportamiento perfectamente lineal, tienden a ser modulados por oscilaciones y mostrar dispersión (usualmente por el número de datos).
- Se pueden presentar discontinuidades en la curva  $D_q$ , lo cual deja sin base teórica el formalismo multifractal, el cual se basa en curvas suaves. Estas discontinuidades, corresponden a transiciones de fase en el equilibrio de la mecánica estadística. Además, es bien conocido que estas transiciones ocurren frecuentemente en sistemas no-hiperbólicos y en sistemas hiperbólicos de altas dimensiones (e.g. el mapa logístico) (**Chhabra & Jensen, 1989; Chhabra et al., 1989; Ott et al., 1984; Grassberger et al., 1988**).

Si se busca describir una medida singular  $P(x)$ , una cantidad de interés es la dimensión de Hausdorff del soporte teórico de la medida,  $P(x)$ . Este es simplemente la ínfima dimensión de los conjuntos en los cuales toda la medida vive (los complementos de esos conjuntos tienen medida cero). Para clases especiales de medidas que surgen de procesos multiplicativos (descritos por probabilidades  $P_i$ ), se tienen varios teoremas que dan información de como calcular la dimensión del soporte teórico de la medida. En particular, se conoce que la entropía,  $S$ , de cada uno de los procesos está dada por

$$S = - \sum_i P_i \log P_i, \quad (13)$$

y que la dimensión de Hausdorff (**Barnsley, 1993**; de  $\mathcal{M}$ , que es el soporte teórico de la medida asociada con cada uno de los procesos, puede ser relacionada con la entropía por medio del teorema de Billingsley como (**Billingsley, 1965**))

$$d_h(\mathcal{M}) = - \lim_{N \rightarrow \infty} \frac{1}{\log N} \sum_{i=1}^N P_i \log P_i. \quad (14)$$

Si se divide la medida experimental considerada, de forma tal que  $P_i(l)$  corresponda a la probabilidad de un proceso multiplicativo con  $N \sim l^{-1}$ , la ecuación (14) provee una fórmula para el cálculo de la dimensión de Hausdorff del conjunto, el cual es el soporte teórico de la medida  $P(x)$ .

Es posible utilizar los resultados anteriores para evaluar  $f(\alpha)$  en una medida multifractal,  $P(x)$ . Esto se logra, inicialmente, por la construcción de una familia, de un solo parámetro, de medidas normalizadas,  $\mu(q)$ , donde las probabilidades en las cajas de tamaño  $l$  son

$$\mu(q, l) = \frac{[P_i(l)]^q}{\sum_j [P_j(l)]^q} \quad (15)$$

El parámetro  $q$  en la ecuación (15) es análogo a un microscopio que permite explorar diferentes regiones de la medida  $P$ . Para  $q > 1$ ,  $\mu(q)$  amplifica las regiones más intermitentes de  $P$ , para  $q < 1$  se acentúan las regiones menos intermitentes y para  $q = 1$  se replica la medida original. La dimensión de Hausdorff del soporte teórico de la medida de  $\mu(q)$  puede ser obtenida mediante la ecuación (14), de manera que

$$\begin{aligned} f(q) &= - \lim_{N \rightarrow \infty} \frac{1}{\log N} \sum_{i=1}^N \mu_i(q, l) \log[\mu_i(q, l)] \\ &= \lim_{l \rightarrow 0} \frac{\sum_i \mu_i(q, l) \log[\mu_i(q, l)]}{\log l} \end{aligned} \quad (16)$$

Además, es posible hallar el valor promedio del exponente de Hölder o fuerza de singularidad,  $\alpha_i = \frac{\log(P_i)}{\log(l)}$ , con respecto a  $\mu(q)$ , al evaluar

$$\begin{aligned} \alpha(q) &= - \lim_{N \rightarrow \infty} \frac{1}{\log N} \sum_{i=1}^N \mu_i(q, l) \log[P_i(l)] \\ &= \lim_{l \rightarrow 0} \frac{\sum_i \mu_i(q, l) \log[P_i(l)]}{\log l} \end{aligned} \quad (17)$$

Las ecuaciones (16) y (17) proveen una definición alterna para la estimación del espectro de singularidades, de forma directa, a partir de datos y sin usar la transformada de Legendre, evitando los problemas mencionados anteriormente.

*Algoritmo de cálculo.* A continuación se describe el algoritmo utilizado para la estimación del espectro multifractal a partir de la metodología propuesta por Chhabra y Jensen, y la cual se utiliza en la referencia (Chhabra et al., 1989).

1. Se cubre la medida experimental con cajas de tamaño  $l$  y se calcula la probabilidad (integración de la medida) en cada una de ellas,  $P_i(l)$ .

2. Se construye la familia de medidas normalizadas,  $\mu(q, l)$ , donde las probabilidades en las cajas de tamaño  $l$  están dadas por la ecuación (15).
3. Para cada valor de  $q$ , se grafica  $\sum_i \mu_i(q, l) \log_{10}[\mu_i(q, l)]$  vs.  $\log_{10}(l)$ , cuya pendiente corresponde a  $f(q)$  (véase la ecuación (16)).
4. Para cada valor de  $q$ , se grafica  $\sum_i \mu_i(q, l) \log_{10}[P_i(l)]$  vs.  $\log_{10}(l)$ , donde la pendiente de la recta corresponde a  $\alpha(q)$  (véase la ecuación (17)).
5. Finalmente se tiene para cada  $q$  un valor de  $f$  y su correspondiente  $\alpha$  ( $f(\alpha(q))$ ).

*2.3.3. Método del análisis multifractal de fluctuaciones.* La descripción de las metodologías que se mencionarán a continuación está basada en el trabajo realizado por Kantelhart y otros, en torno a las propiedades de escalamiento de registros históricos de series de caudal (Kantelhardt et al., 2003; Koscielny-Bunde et al., 2006). El objetivo fundamental de estas metodologías es la detección de correlaciones temporales de largo plazo y comportamiento multifractal en presencia de tendencias para series temporales de variables geofísicas.

En registros temporales de variables geofísicas con limitada extensión, es complicado distinguir entre memoria de largo plazo y jerarquías de memoria de corto plazo y/o tendencias. Lo anterior, debido a que las series de tiempo estacionarias y con correlaciones de largo plazo exhiben un comportamiento persistente y tendencia a permanecer en un valor momentáneamente. Esto causa desviaciones positivas o negativas respecto al valor medio en largos periodos de tiempo que pueden lucir como tendencias (Koscielny-Bunde et al., 2006).

En los últimos años se han desarrollado métodos como el análisis de fluctuaciones sin tendencias, para detectar correlaciones de largo plazo en presencia de tendencias. Sin embargo, para poder caracterizar la compleja dinámica involucrada en los procesos geofísicos, se hace necesario el análisis de todos los momentos estadísticos y así captar completamente la distribución de las fluctuaciones (análisis multifractal).

En el caso de las variables geofísicas, se trabajará con series a las cuales se les han removido las tendencias estacionarias periódicas (por ejemplo el ciclo anual o el diario). En el caso de series de tiempo con resolución

horaria, se removerá el ciclo en cada hora de los 365 días del año (por ejemplo *hora 1 de abril 1*).

Sea  $W_i$  el registro correspondiente a la hora  $i$ ,  $i = 1, 2, \dots, N$ . La nueva serie, sin tendencias estacionales se genera como

$$\phi_i = W_i - \overline{W}_i \quad (18)$$

donde  $\overline{W}_i$  corresponde a la media sobre todos los días  $i$  (por ejemplo sobre todos los registros *hora 1 de abril 1* en los años de registro).

En el análisis de fluctuaciones, se considera el “perfil”

$$z_n = \sum_{i=1}^n \phi_i, \quad n = 1, 2, \dots, N \quad (19)$$

y se estudian las fluctuaciones del perfil, en una ventana de tiempo de tamaño  $s$ . El perfil  $z_n$  se puede considerar como la posición de una caminata aleatoria de  $n$  pasos (Koscielny-Bunde et al., 2006).

*Análisis multifractal de fluctuaciones (MF-FA)*. En el caso análisis multifractal de fluctuaciones se obtienen funciones de fluctuación a partir de los valores del perfil en los extremos de cada segmento  $\nu$  (Kantelhardt et al., 2003; Koscielny-Bunde et al., 2006)

$$F_q(s) = \left\{ \frac{1}{N_s} \sum_{\nu=1}^{N_s} |z_{\nu s} - z_{(\nu-1)s}|^q \right\}^{\frac{1}{q}} \quad (20)$$

en este caso  $q$  puede tomar cualquier valor real excepto cero. La diferencia de este análisis de fluctuaciones, respecto al análisis estándar radica en la variación de la función de fluctuaciones con  $q$ , lo cual es un reflejo del comportamiento multifractal en los registros.

En general, la aproximación multifractal se introduce en la función de partición

$$Z_q(s) = \sum_{\nu=1}^{N_s} |z_{\nu s} - z_{(\nu-1)s}|^q \sim s^{\tau(q)}, \quad (21)$$

donde  $\tau(q)$  es el exponente de escalamiento de Renyi ( $\tau: \mathbb{R} \rightarrow \mathbb{R}$ ). La ecuación (21) implica

$$F_q(s) \sim s^{h(q)}, \quad (22)$$

donde

$$h(q) = \frac{[\tau(q) + 1]}{q}. \quad (23)$$

El exponente  $h(q)$  definido en la ecuación (22), está directamente relacionado con el exponente multifractal de escalamiento clásico,  $\tau(q)$ . Este exponente es denominado de “Hurst generalizado”, que en el caso de  $h(1)$  es idéntico al conocido exponente de Hurst (Mesa & Poveda, 1993).

Finalmente, es posible obtener el espectro de singularidades o multifractal a partir de  $\tau(q)$  vía la transformada de Legendre (Mandelbrot, 1989).

$$\alpha = \frac{d\tau(q)}{dq} = h(q) + q \frac{dh(q)}{dq} \quad (24a)$$

$$f(\alpha) = q\alpha - \tau(q) = q[\alpha - h(q)] + 1. \quad (24b)$$

El MF-FA descrito tiene los siguientes problemas:

- Tendencias monótonas en los registros pueden generar resultados espurios para el exponente de fluctuación  $h(q)$ , el cual a su vez, genera valores espurios para la estimación en el exponente de Renyi,  $\tau(q)$ .
- El comportamiento no estacionario, caracterizado por los exponentes  $h(q) \geq 1$  no puede ser detectado por el método y siempre aproxima a  $F_q(s) \sim s$ .

*Algoritmo de cálculo MF-FA*. A continuación se presenta el algoritmo para el cálculo del espectro multifractal de series de tiempo a partir de MF-FA, el cual está basado en la teoría descrita anteriormente.

1. Se remueven las tendencias estacionarias de la serie de tiempo. Este procedimiento conduce a la generación de una nueva serie denominada  $\phi_i$  (véase la ecuación (18)).
2. Se genera el perfil de la serie de tiempo a partir de la ecuación (19).
3. Con el perfil definido, se procede a realizar el análisis de fluctuaciones, definiendo un  $q$  constante y evaluando la función de fluctuaciones  $F_q(s)$  para varias resoluciones  $s$  (véase la ecuación (20)).
4. Con los valores encontrados de  $F_q(s)$  y  $s$  se realiza un ajuste lineal entre  $\log_{10}(F_q(s))$  vs.  $\log_{10}(s)$ , donde la pendiente de dicho ajuste corresponde al exponente de Hurst generalizado para el momento de orden  $q$ ,  $h(q)$  (véase la ecuación (22)).
5. Se repiten los pasos 3 y 4 para varios valores de  $q$ , obteniendo la curva  $h(q)$ .
6. Posterior al cálculo de la curva del exponente de Hurst generalizado,  $h(q)$ , se procede a utilizar la transformada de Legendre (ecuaciones (24a) y (24b)) para la estimación del espectro multifractal  $f(\alpha)$ .

*Análisis multifractal de fluctuaciones sin tendencias (MF-DFA).* Es posible identificar diferentes ordenes de MF-DFA, los cuales se caracterizan por la forma en la cual se remueven las tendencias de los datos. En el orden más bajo (MF-DFA1) se determina, para cada segmento  $\nu$ , el mejor ajuste *lineal* del perfil, y se identifican las fluctuaciones por medio de la varianza  $F^2(\nu, s)$  del perfil respecto a la línea recta. De esta forma, se elimina la posible influencia de tendencias lineales a escalas mayores a la del segmento. Es importante notar que al eliminar las tendencias en el perfil, se están eliminando en los datos.

En general, en el MF-DFA de orden  $n$ , se calculan las varianzas del perfil, respecto al mejor ajuste polinomial de orden  $n$ , lo cual elimina la influencia de posibles tendencias de orden  $n$  a escalas mayores que el tamaño del segmento.

Explícitamente, se ajusta el mejor polinomio  $y_\nu(i)$  al perfil en cada uno de los  $N_s$  segmentos  $\nu$  y se determina la varianza

$$F^2(\nu, s) \equiv \frac{1}{s} \sum_{i=1}^s [z_{(\nu-1)s+i} - y_\nu(i)]^2 \quad (25)$$

Finalmente, para estimar la fluctuación media se utiliza

$$F_q(s) = \left\{ \frac{1}{N_s} \sum_{\nu=1}^{N_s} [F^2(\nu, s)]^{\frac{q}{2}} \right\}^{\frac{1}{q}}. \quad (26)$$

A partir de este punto el procedimiento es similar al utilizado en el MF-FA, donde se estima el exponente de Hurst generalizado a partir de la relación mostrada en la ecuación (22) y, con este y la transformada de Legendre (ecuaciones (24a) y (24b)), se hace posible el cálculo del espectro multifractal.

*Algoritmo de cálculo MF-DFA.* A continuación se presenta el algoritmo para el cálculo del espectro multifractal de series de tiempo a partir de MF-DFA, el cual está basado en la teoría descrita anteriormente.

1. Se remueven las tendencias estacionarias de la serie de tiempo. Este procedimiento conduce a la generación de la serie  $\phi_i$  (véase la ecuación (18)).
2. Se genera el perfil de la serie de tiempo a partir de la ecuación (19).
3. Con el perfil definido, se procede a realizar el análisis de fluctuaciones, definiendo un  $q$  constante y analizando varios intervalos en el siguiente orden:

- a) En cada intervalo  $\nu$ , se ajusta un polinomio  $y_\nu$  de orden  $n$ , para el caso de MF-DFA.
  - b) Se calculan las varianzas del perfil respecto al ajuste (véase la ecuación (25)).
  - c) Con las varianzas de cada segmento se estima la función de fluctuaciones para el valor dado de  $s$  y  $q$  (véase la ecuación (26)).
4. Con los valores encontrados de  $F_q(s)$  y  $s$ , se realiza un ajuste lineal entre  $\log_{10}(F_q(s))$  vs.  $\log_{10}(s)$ , donde la pendiente de dicho ajuste corresponde al exponente de Hurst generalizado para el momento de orden  $q$ ,  $h(q)$  (véase la ecuación (22)).
  5. Se repiten los pasos 3 y 4 para varios valores de  $q$ , obteniendo la curva  $h(q)$ .
  6. Posterior al cálculo de la curva del exponente de Hurst generalizado,  $h(q)$ , se procede a utilizar la transformada de Legendre (ecuaciones (24a) y (24b)) para estimación del espectro multifractal  $f(\alpha)$ .

### 3. Descripción de los datos

Para este trabajo se contó con registros horarios de precipitación correspondientes a 47 estaciones pluviográficas ubicadas en las regiones cafeteras de Colombia (véase la figura 1 y Tabla 1). Dichos registros fueron suministrados por el Centro Nacional de Investigaciones del Café (CENICAFE) y por Empresas Públicas de Medellín (EPM).

Las estaciones poseen periodos de registro entre los 22 y 28 años, el porcentaje de datos faltantes es generalmente inferior al 5% y se encuentran entre  $1^{\circ}15'N - 10^{\circ}20'N$  y  $77^{\circ}29'W - 72^{\circ}40'W$ .

Los Andes Colombianos se encuentran divididos en tres ramales principales, llamados Cordillera Occidental, Cordillera Central y Cordillera Oriental, los cuales se encuentran separados por los valles interandinos de los ríos Cauca y Magdalena. Las estaciones en la Tabla 1 son numeradas y agrupadas de acuerdo a su localización a lo largo de las pendientes andinas (Poveda et al., 2005):

**Grupo A:** Vertiente occidental de la Cordillera Occidental (1-2).

**Grupo B:** Sur del "Macizo Colombiano" (3-5).

**Grupo C:** Vertiente oriental de la Cordillera Occidental perteneciente al valle del río Cauca (6-11).

**Grupo D:** Vertiente occidental de la Cordillera Central (12-30).

**Grupo E:** Vertiente oriental de la Cordillera Central perteneciente al valle del río Magdalena (31-37).

**Grupo F:** Flanco occidental de la Cordillera Oriental (38-44).

**Grupo G:** Flanco oriental de la Cordillera Oriental (45-46).

**Grupo H:** Estación en la Sierra Nevada de Santa Marta (47).

Las estaciones utilizadas reflejan un comportamiento unimodal, bimodal y multimodal en su ciclo diario y se hace confusa la definición de una hora común de máxima lluvia en todas las estaciones, en la cual se presenta la máxima precipitación durante el día.

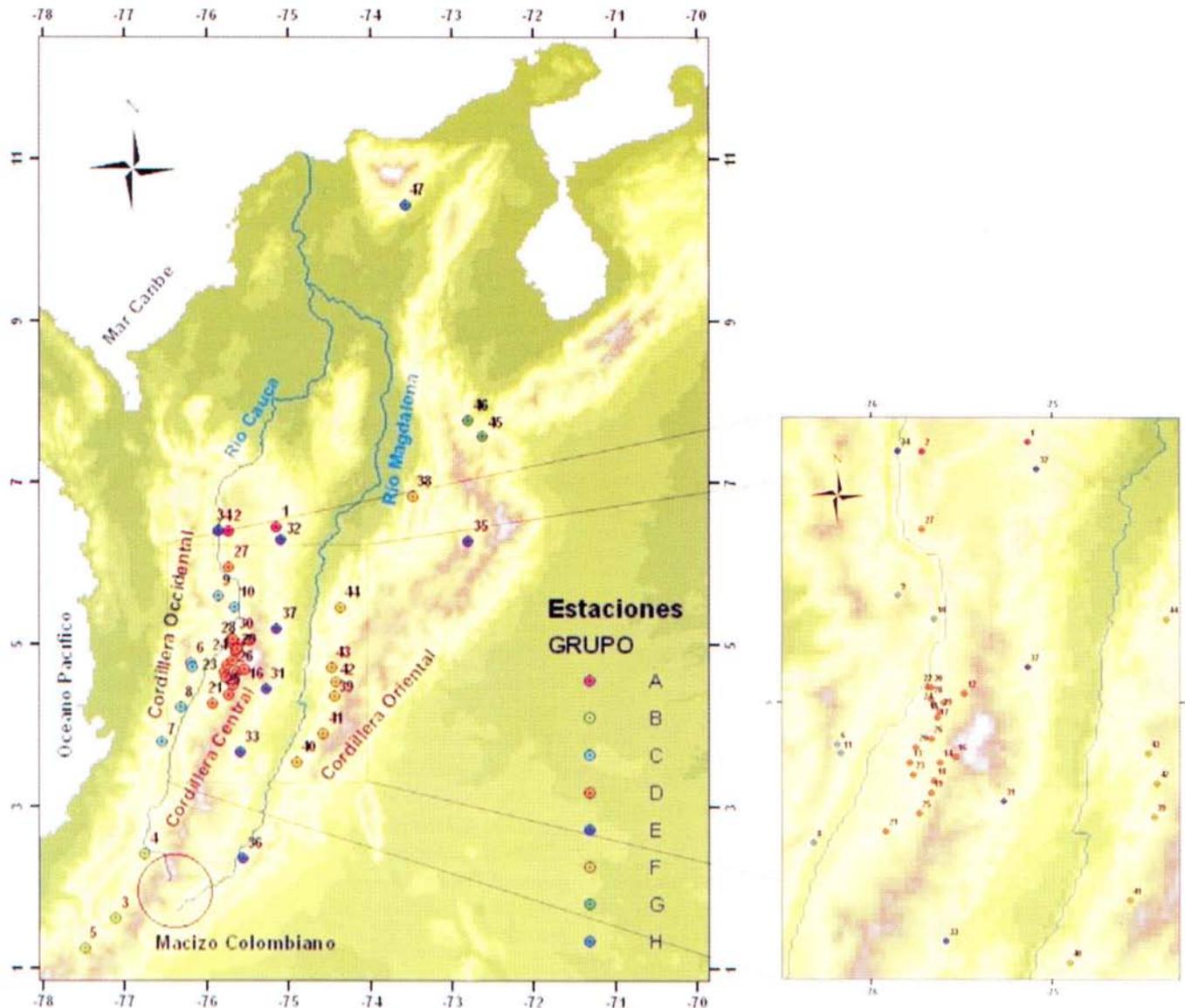


FIGURA 1. Ubicación geográfica de las 47 estaciones de precipitación. La escala de colores en las estaciones corresponde a su localización a lo largo de las pendientes andinas.

CUADRO 1. Localización y detalles de las 47 estaciones de precipitación.

ID	Estación	Lat (N) grados	Lon (W) grados	Elev (m)	Periodo Registro	Datos (%) Faltantes	Elevación desde el piso del valle (m)	Grupo
1	Mande	6.45	75.13	495	MAY78-ABR99	0	x	A
2	Santa Bárbara	6.4	75.72	2595	MAR78-DIC99	0	x	A
3	El Sauce	1.62	77.1	1610	ENE71-DIC93	12.4	x	B
4	Manuel Mejía	2.42	76.75	1700	ENE72-NOV93	2.1	x	B
5	Ospina Pérez	1.25	77.48	1700	ENE72-DIC93	0.7	x	B
6	Albán	4.77	76.18	1500	JUN73-DIC93	1.7	780	C
7	Julio Fernandez	3.8	76.53	1360	ENE72-DIC93	1	480	C
8	Manuel María Mallarino	4.22	76.32	1380	ENE73-DIC93	2.5	570	C
9	Miguel Valencia	5.6	75.85	1570	ENE71 - DIC93	1.9	x	C
10	Rafael Escobar	5.47	75.65	1320	ENE71 - DIC93	0.1	x	C
11	Santiago Gutiérrez	4.72	76.17	1550	ENE71-DIC93	5.7	820	C
12	Agronomía	5.05	75.48	2150	ENE72-DIC99	1.1	1480	D
13	Arturo Gómez	4.67	75.78	1320	ENE72-DIC93	1.5	590	D
14	Bremen	4.67	75.62	2040	MAR72-DIC93	1.7	1300	D
15	Cenicafe	5	75.6	1310	ENE72-DIC93	0.4	630	D
16	El Cedral	4.7	75.53	2120	ENE73-DIC93	5.8	1420	D
17	El Jazmín	4.92	75.63	1600	ENE72-DIC93	2.1	910	D
18	El Sena	4.57	75.65	1550	DIC71-DIC93	8.8	800	D
19	La Bella	4.5	75.67	1450	ENE72-DIC93	0.4	690	D
20	La Catalina	4.75	75.75	1350	NOV76-DIC93	1	630	D
21	La Sirena	4.28	75.92	1500	ENE71-DIC93	1	700	D
22	Luker	5.08	75.68	1020	MAR70-DIC99	2.7	360	D
23	Maracay	4.6	75.77	1450	ENE82-DIC93	0.6	710	D
24	Naranjal	4.98	75.65	1400	ENE71-DIC93	0.6	720	D
25	Paraguaicito	4.38	75.73	1250	ENE72 - DIC93	0.5	470	D
26	Planta de Tratamiento	4.8	75.67	1450	ENE69-DIC89	4	740	D
27	Rosario	5.97	75.72	1600	ENE71 - DIC93	0.9	1086	D
28	Santa Ana	5.02	75.67	1250	ENE72-DIC93	2.2	580	D
29	Santa Helena	4.95	75.62	1525	ENE80-DIC93	2.1	840	D
30	Santagueda	5.08	75.67	1010	ABR72-DIC99	1.2	350	D
31	Chapetón	4.45	75.27	1300	ENE71-DIC88	2.2	1010	E
32	El Bizcocho	6.3	75.08	1070	ABR71-DIC99	0	1120	E
33	El Limón	3.67	75.58	990	ENE71-DIC93	7.2	610	E
34	El Peñol	6.4	75.85	1880	ABR60-FEB99	0	1750	E
35	Inmarco	6.28	72.8	260	AGO68-AGO92	0	140	E
36	Jorge Villamil	2.37	75.55	1500	ENE72 - DIC93	1.1	x	E
37	Llanadas	5.2	75.13	1420	ENE72-DIC99	10	1220	E
38	Aguas Blancas	6.83	73.48	920	ENE72-DIC99	8.4	x	F
39	Granja Tibacuy	4.37	74.43	1550	ENE72-DIC93	1.4	1250	F
40	La Montaña	3.55	74.9	1260	ENE73 - DIC93	15.9	x	F
41	Luis Bustamante	3.9	74.57	1643	ENE72-DIC89	18.9	x	F
42	Misiones	4.55	74.42	1540	ENE77-DIC93	3	1263	F
43	Santa Inés	4.72	74.47	1250	ENE72-DIC93	1.9	990	F
44	Yacopí	5.47	74.37	1340	ENE72-DIC99	1.5	x	F
45	Blonay	7.57	72.62	1235	ENE72-DIC99	0.7	x	G
46	Francisco Romero	7.77	72.80	1000	ENE72 - DIC99	1.1	x	G
46	Pueblo Bello	10.42	73.57	1000	ENE72 - DIC99	3.4	x	G
47	Yacopí	5.46	74.36	1340	ENE72 - DIC99	1.5	x	H

#### 4. Resultados

##### 4.1. Función densidad de probabilidad empírica.

Se calculó la probabilidad de excedencia de un evento  $x$ ,  $P(X > x)$ , para cada serie de precipitación, encontrando una tendencia a dos zonas diferentes de comportamiento potencial en las distribuciones (véase la ecuación 1).

Para caracterizar el comportamiento potencial se trabajó en el espacio logarítmico,  $(\log(P(X > x)))$  vs.  $\log(x)$ , donde se ajustan rectas a las dos sub-regiones, dando como resultado pendientes  $q_{cr1}$  y  $q_{cr2}$ , que corresponden al negativo de la potencia a la cual

escala la distribución de probabilidades (véanse la figura 2 y la tabla 2). El criterio utilizado para encontrar el punto de quiebre es el siguiente:

1. En el paso  $i$ , se divide la muestra en dos sub-muestras con tamaño mínimo de 20 datos.
2. A cada sub-muestra se le ajusta una recta y se obtienen los coeficientes de determinación  $R_{i1}^2$  y  $R_{i2}^2$ .
3. Se define la cantidad  $\Delta_i = R_{i1}^2 + R_{i2}^2$ .
4. Se repite el proceso iterativamente para todos los  $i$  posibles.
5. El punto de quiebre se define como aquel que presente el  $\Delta_{max}$ .

Con el propósito de decidir si el comportamiento es diferente en las sub-muestras seleccionadas, se realiza una prueba de hipótesis que verifique si las pendientes son estadísticamente diferentes, a partir de los intervalos de confianza para la pendiente obtenida en una regresión lineal, el cual a su vez refleja la calidad total de la recta de regresión.

Bajo la hipótesis de que las observaciones están distribuidas de manera normal e independiente, el intervalo de confianza para la pendiente  $D$  del  $100(1-\alpha)$  por ciento en una regresión lineal simple es (Montgomery &

Runger, 1996):

$$\hat{D} - t_{\frac{\alpha}{2}, n-2} \sqrt{\frac{\widehat{\sigma^2}}{S_{xx}}} \leq D \leq \hat{D} + t_{\frac{\alpha}{2}, n-2} \sqrt{\frac{\widehat{\sigma^2}}{S_{xx}}}, \quad (27)$$

donde  $\sigma^2$  es la varianza del término de error  $\epsilon$  en el modelo de regresión, el cual refleja la variación aleatoria alrededor de la recta de regresión. Además,  $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$ . Teniendo el intervalo de confianza para  $q_{cr1}$ , se acepta la hipótesis de igualdad en las pendientes, si dicho intervalo contiene a  $q_{cr2}$ , caso en el cual se acepta que el comportamiento potencial es el mismo en todo el dominio.

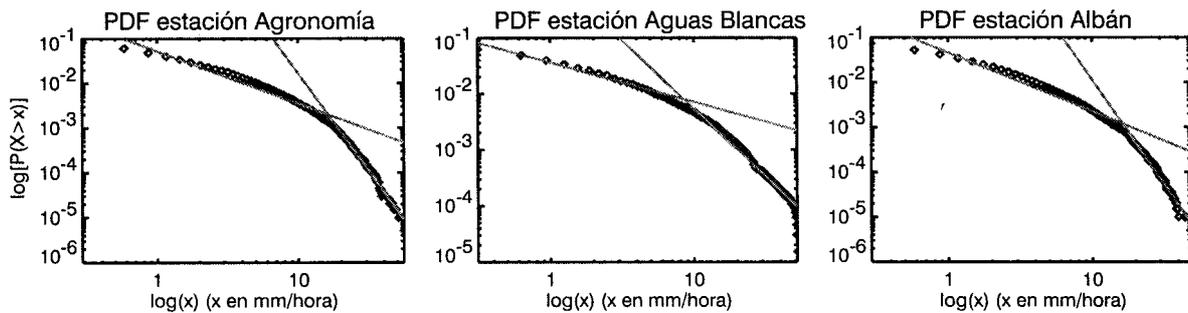


FIGURA 2. Distribución de probabilidad empírica para las estaciones Agronomía, Aguas Blancas y Albán. La línea roja refleja la recta ajustada.

Como se mencionó anteriormente, la distribución de probabilidad empírica describe la fractalidad de umbrales de intensidad para las fluctuaciones en la serie a determinadas escalas. La principal justificación para el uso de un modelo de colas pesadas, se describe en el estudio del ciclo diurno (véanse las referencias (Agudelo et al., 2001; Álvarez & Toro, 2001; Vieira & Moreno, 2001)), donde se encontró que las series de precipitación horarias utilizadas son altamente intermitentes, con predominancia de ceros (no lluvia) en los registros (en general, al menos el 85% de los datos). Basados en esto, se hace necesario el análisis de la función de distribución de probabilidades y en particular de los cuantiles superiores.

El ajuste potencial en la cola de la distribución muestra valores de  $q_{cr2}$  entre 2,4 y 5,8, los cuales son suficientes para justificar el uso de una aproximación multifractal en el estudio del comportamiento de estas series de lluvia (Olsson, 1995; Sivakumar, 2001). Por el contrario, a la izquierda del quiebre se encuentra que  $q_{cr1} < 2$  en todas las estaciones, lo cual sugiere un comportamiento monofractal para intensidades inferiores a

la del quiebre, cuyos valores presentan intensidades entre 14-18  $\frac{mm}{hora}$  en la mayoría de los casos, tomando los valores más bajos en 5.2  $\frac{mm}{hora}$  y máximos en 30.4  $\frac{mm}{hora}$ .

**4.2. Espectro de potencias.** Como se ha dicho anteriormente, el espectro de potencias es una herramienta estándar para el estudio del comportamiento fractal en series geofísicas. Dado lo anterior, se estimó dicho espectro para cada una de las 47 estaciones de precipitación horaria y se estimaron los parámetros del ajuste  $E(f) = Cf^{-\beta}$ , detectando tendencias y quiebres como se explicó anteriormente. Además, se realizaron pruebas de hipótesis orientadas a verificar la diferencia estadística entre dichas pendientes (véanse la figura 3 y la tabla 3).

En la Tabla 3 se presentan los resultados obtenidos en la estimación de la pendiente, puntos de quiebre y pruebas de hipótesis de las 47 estaciones. En particular, se puede distinguir seis grupos de quiebres importantes:

- Entre las 2 – 4 horas: Bremen, El Bizcocho, El Peñol, La Montaña, Planta de tratamiento,

CUADRO 2. Resultados de los ajustes de la forma  $P(X > x) \propto x^{-q_{er}}$ , realizados a las funciones densidad de probabilidad empírica. En todos los casos se encontraron dos regiones (1 - izquierda y 2 - derecha).

Estación	$C_1$	$q_{er1}$	$R_1^2$	$C_2$	$q_{er2}$	$R_2^2$	$x$ (mm/hora)	P(x)
Agronomía	0,052 ± 0,0020	-1,18 ± 0,020	0.94	404,1 ± 43,1	-4,4 ± 0,03	0.99	16.2	0.002
Aguas Blancas	0,036 ± 0,0005	-0,71 ± 0,015	0.97	1,5 ± 0,1	-2,4 ± 0,03	0.96	8.5	0.008
Albán	0,046 ± 0,0019	-1,31 ± 0,021	0.95	460,1 ± 71,3	-4,6 ± 0,04	0.98	16.3	0.001
Arturo Gómez	0,046 ± 0,0014	-1,01 ± 0,017	0.95	167,0 ± 18,1	-4,0 ± 0,03	0.98	16.2	0.003
Blonay	0,036 ± 0,0011	-1,08 ± 0,018	0.96	10,0 ± 0,9	-3,2 ± 0,03	0.98	14.4	0.002
Bremen	0,055 ± 0,0011	-0,76 ± 0,015	0.96	47,2 ± 6,3	-3,4 ± 0,04	0.94	12.4	0.008
Cenicafé	0,055 ± 0,0016	-0,91 ± 0,017	0.95	330,6 ± 42,9	-4,0 ± 0,04	0.97	16.8	0.004
Chapetón	0,040 ± 0,0008	-0,74 ± 0,018	0.96	7,5 ± 1,1	-3,0 ± 0,04	0.93	10.0	0.007
El Bizcocho	0,092 ± 0,0027	-0,86 ± 0,016	0.95	2,756,4 ± 476,8	-4,3 ± 0,04	0.95	19.3	0.007
El Cedral	0,082 ± 0,0038	-1,25 ± 0,021	0.93	2,138,7 ± 264,6	-4,6 ± 0,03	0.99	21.3	0.002
El Jazmín	0,057 ± 0,0015	-0,92 ± 0,014	0.96	137,7 ± 17,0	-3,7 ± 0,03	0.97	16.1	0.004
El Limón	0,067 ± 0,0022	-0,96 ± 0,017	0.96	213,1 ± 39,9	-3,7 ± 0,05	0.94	19.9	0.004
El Peñol	0,061 ± 0,0018	-0,90 ± 0,029	0.93	13,6 ± 1,6	-3,5 ± 0,04	0.96	7.9	0.010
El Sauce	0,048 ± 0,0019	-1,13 ± 0,021	0.94	155,9 ± 22,6	-4,0 ± 0,04	0.98	16.2	0.002
El Sena	0,051 ± 0,0009	-0,76 ± 0,013	0.96	13,4 ± 1,6	-3,0 ± 0,03	0.95	12.4	0.007
Francisco Romer	0,050 ± 0,0014	-0,82 ± 0,015	0.96	177,8 ± 30,0	-3,6 ± 0,04	0.94	19.3	0.004
Granja Tibacuy	0,023 ± 0,0003	-0,79 ± 0,015	0.97	1,7 ± 0,2	-2,8 ± 0,03	0.93	8.1	0.004
Inmarco	0,047 ± 0,0013	-0,93 ± 0,013	0.96	2,113,8 ± 461,9	-4,4 ± 0,05	0.94	22.7	0.003
Jorge Villamil	0,032 ± 0,0009	-1,08 ± 0,021	0.96	5,1 ± 0,4	-3,3 ± 0,03	0.98	10.2	0.003
Julio Fernandez	0,026 ± 0,0003	-0,71 ± 0,017	0.96	0,9 ± 0,1	-2,6 ± 0,03	0.95	6.5	0.007
La Bella	0,061 ± 0,0025	-1,11 ± 0,019	0.94	16,893,7 ± 3,908,1	-5,0 ± 0,06	0.96	24.4	0.002
La Catalina	0,048 ± 0,0014	-1,00 ± 0,016	0.94	688,0 ± 124,5	-4,3 ± 0,05	0.96	17.6	0.003
La Montaña	0,066 ± 0,0031	-1,23 ± 0,019	0.93	458,636,0 ± 151,512,2	-5,8 ± 0,08	0.97	30.4	0.001
La Sirena	0,044 ± 0,0012	-0,89 ± 0,015	0.95	186,9 ± 26,5	-3,9 ± 0,04	0.96	16.7	0.004
Llanadas	0,064 ± 0,0017	-0,92 ± 0,015	0.95	406,7 ± 67,9	-4,1 ± 0,05	0.95	15.5	0.005
Luis Bustamante	0,046 ± 0,0010	-0,94 ± 0,014	0.97	19,4 ± 2,0	-3,3 ± 0,03	0.97	12.9	0.004
Luker	0,053 ± 0,0015	-1,00 ± 0,016	0.96	244,0 ± 45,2	-4,1 ± 0,05	0.94	15.7	0.003
Mande	0,076 ± 0,0020	-0,91 ± 0,015	0.97	443,2 ± 84,3	-3,9 ± 0,05	0.95	17.3	0.006
Manuel María Ma	0,041 ± 0,0012	-1,02 ± 0,016	0.95	396,5 ± 60,4	-4,2 ± 0,04	0.97	17.9	0.002
Manuel Mejía	0,073 ± 0,0037	-1,33 ± 0,021	0.94	16,290,9 ± 3,579,6	-5,2 ± 0,06	0.98	24.8	0.001
Maracay	0,052 ± 0,0016	-0,90 ± 0,016	0.95	126,0 ± 13,3	-3,7 ± 0,03	0.98	16.8	0.004
Miguel Valencia	0,050 ± 0,0003	-0,67 ± 0,017	0.96	2,5 ± 0,3	-3,1 ± 0,04	0.93	5.2	0.017
Misiones	0,031 ± 0,0005	-0,94 ± 0,012	0.98	8,5 ± 1,4	-3,1 ± 0,05	0.92	13.0	0.003
Naranjal	0,064 ± 0,0022	-0,97 ± 0,018	0.95	1,444,6 ± 284,9	-4,4 ± 0,05	0.95	18.9	0.004
Ospina Pérez	0,040 ± 0,0014	-1,23 ± 0,021	0.94	45,1 ± 4,8	-3,9 ± 0,03	0.98	13.3	0.002
Paraguaicito	0,040 ± 0,0006	-0,73 ± 0,012	0.97	14,3 ± 2,0	-3,1 ± 0,04	0.92	11.8	0.007
Planta de Trata	0,049 ± 0,0007	-0,69 ± 0,012	0.97	8,4 ± 0,8	-2,9 ± 0,03	0.95	10.3	0.010
Pueblo Bello	0,030 ± 0,0006	-0,72 ± 0,011	0.97	28,3 ± 4,6	-3,2 ± 0,04	0.94	16.7	0.004
Rafael Escobar	0,061 ± 0,0027	-1,23 ± 0,020	0.93	22,873,6 ± 4,213,2	-5,3 ± 0,05	0.98	23.1	0.001
Rosario	0,062 ± 0,0021	-1,04 ± 0,018	0.96	142,9 ± 15,2	-3,7 ± 0,03	0.98	17.9	0.003
Santa Ana	0,077 ± 0,0035	-1,20 ± 0,020	0.94	33,142,9 ± 9,060,8	-5,2 ± 0,07	0.96	24.9	0.002
Santa Bárbara	0,048 ± 0,0003	-0,72 ± 0,015	0.97	2,2 ± 0,2	-3,0 ± 0,03	0.94	5.2	0.015
Santa Helena	0,079 ± 0,0027	-0,86 ± 0,018	0.96	245,9 ± 50,9	-3,5 ± 0,05	0.94	19.9	0.006
Santa Inés	0,034 ± 0,0009	-0,95 ± 0,014	0.96	74,9 ± 12,1	-3,6 ± 0,04	0.95	17.4	0.002
Santagueda	0,044 ± 0,0012	-0,88 ± 0,016	0.95	192,9 ± 36,0	-3,9 ± 0,05	0.94	15.5	0.004
Santiago Gutiérrez	0,056 ± 0,0028	-1,44 ± 0,022	0.92	57,740,1 ± 12,021,7	-5,8 ± 0,05	0.99	23.3	0.001
Yacopí	0,068 ± 0,0025	-1,04 ± 0,017	0.96	8,730,6 ± 1,754,1	-4,8 ± 0,05	0.97	22.7	0.003

Pueblo Bello, Santa Bárbara, Santa Inés y Santagueda.

- Entre las 4,5 – 7 horas: Aguas Blancas, Cenicafé, El limón, El Sauce, Granja Tibacuy, Llanadas, Luis Bustamante, Rafael Escobar, Santiago Gutiérrez y Yacopí.
- Entre las 7,5 – 11 horas: El Jazmín, El Sena, Inmarco, Julio Fernández, La Sirena, Mande, Naranjal, Paraguaicito y Santa Ana.
- Entre las 11,5 – 14,5 horas: Chapetón,, Jorge Villamil, Maracay y Luker.
- Entre las 15 – 18 horas: Rosario y Santa Elena.

- Entre las 23 – 24 horas: Luker, Miguel Valencia y Ospina Pérez.

Dichos grupos hacen referencia u oscilan alrededor de horas importantes del ciclo diurno como las 4, 6, 9, 12, 16 y 24 horas.

Para todos los casos en los cuales se presentaron quiebres en los espectros, la hipótesis de igualdad de pendientes se rechazó con un 95 % de significancia. El espectro de potencias de una serie de tiempo, a través de su exponente espectral,  $\beta$ , puede ser usado para obtener información importante acerca de la cascada que

resulta en un proceso, además, de la identificación de presencia o ausencia de escalas características de tiempo (naturaleza fractal). En este sentido, una condición necesaria, pero no suficiente, para que el proceso sea un resultado directo de una cascada sin límites es que  $\beta < 1$  (Sivakumar, 2001). Sin embargo, cascadas sin límites pueden producir procesos con  $\beta > 1$ , pero ellas no son multifractales en todos los casos (Sivakumar, 2001). En nuestro caso se encontró que la lluvia en los Andes tropicales de Colombia se puede modelar directamente por un proceso de cascadas multiplicativas ( $\beta < 1$ ), a excepción de tres estaciones, las cuales sin embargo están muy cercanas al umbral: El Peñol ( $1,07 \pm 0,01$ ), Julio Fernández ( $1,03 \pm 0,03$ ) y Mande ( $1,20 \pm 0,02$ ).

La concepción clásica considera la turbulencia como un campo anidado de remolinos de tamaños decrecientes, donde la energía cinética turbulenta “cascadea”

con despreciable disipación de los remolinos de mayor energía a remolinos de menor y menor energía, hasta alcanzar la escala de Kolmogorov (i.e., escala viscosa), donde los efectos de la viscosidad no pueden ser despreciados y las fluctuaciones turbulentas comienzan a ser suavizadas.

Sin embargo, en algunos casos dicha homogeneidad es inadecuada, dada la intermitencia de del fenómeno modelado, por lo tanto la hipótesis de Kolmogorov es insostenible pues dichos campos turbulentos son no homogéneos y corresponden a procesos dependientes de la escala, en escalas cada vez más pequeñas (Chorin, 1994; Seuront et al., 1999).

Cuando el exponente  $\beta$  del espectro de potencias es igual a  $5/3 \simeq 1,67$ , se presenta la turbulencia homogénea de Kolmogorov. En todas las estaciones se presentan valores del exponente  $\beta$  menores a  $5/3$ .

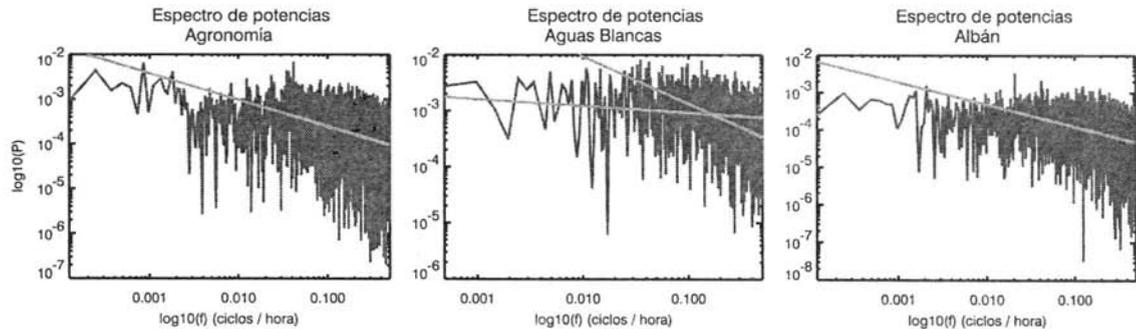


FIGURA 3. Espectro de potencias para las estaciones Agronomía, Aguas Blancas y Albán. La línea roja refleja la tendencia del espectro.

**4.3. Espectros multifractales.** En la figura 4 se presentan los espectros multifractales estimados por los cinco métodos, para tres de las estaciones analizadas. El método de los momentos presenta problemas para valores de  $q$  muy negativos, generando inestabilidad numérica y curvaturas irregulares en el espectro.

El método MF-FA, presenta inestabilidades muy fuertes en algunos casos, como se puede apreciar en la figura 4, sin embargo la remoción de tendencias (MF-DFA) supera estas inestabilidades y genera un espectro bien definido.

En general, para valores positivos de  $q$ , los métodos MF-FA, MF-DFA1 y MF-DFA4 estiman curvas similares, presentando divergencias muy fuertes a medida que  $q$  se hace negativo (i.e. valores grandes de  $\alpha$ ).

Es importante conocer el origen de la multifractalidad presente en las series de precipitación horaria. Al reordenar aleatoriamente los datos y estimar el espectro multifractal o verificar la existencia de multiescalamiento, es posible encontrar que los nuevos datos se comporten de forma monofractal, lo que indicaría que el comportamiento multifractal es debido a correlaciones de largo plazo en los registros, de lo contrario, si la multifractalidad persiste, se atribuirá la causa a singularidades en la distribución de los datos, es decir, a algún tipo de no linealidad estadística (Koscielny-Bunde et al., 2006). La figura 5, muestra los espectros estimados para la estación Agronomía con las cinco metodologías, tanto para la serie original, como para la serie reordenada aleatoriamente. En ambos casos se mantiene el comportamiento multifractal, lo cual atribuye el origen de

CUADRO 3. Resultados de los ajustes de la forma  $E(f) \propto f^{-\beta}$ , realizados a los espectros de potencia. En varios casos se encontraron dos regiones (1 - izquierda y 2 - derecha).

Estación	$C_1$	$\beta_1$	$C_2$	$\beta_2$	$f^*$ (ciclos/hora)	$P^*$	$t^*$ (horas)
Agronomía	-	-	0,00006 ± 0,000001	0,59 ± 0,01	-	-	-
Aguas Blancas	0,0007 ± 0,00009	0,12 ± 0,04	0,00019 ± 0,000011	0,84 ± 0,05	0,17	0,00087	6,0
Albán	-	-	0,00003 ± 0,000001	0,59 ± 0,01	-	-	-
Arturo Gómez	-	-	0,00004 ± 0,000001	0,53 ± 0,01	-	-	-
Blonay	-	-	0,00008 ± 0,000002	0,69 ± 0,01	-	-	-
Bremen	0,0002 ± 0,00004	0,28 ± 0,07	0,00010 ± 0,000003	0,62 ± 0,02	0,27	0,00023	3,7
Cenicafé	0,0004 ± 0,00009	0,15 ± 0,06	0,00021 ± 0,000007	0,52 ± 0,02	0,14	0,00056	6,9
Chapetón	0,0008 ± 0,00015	0,06 ± 0,05	0,00015 ± 0,000005	0,74 ± 0,02	0,09	0,00090	11,4
El Bizcocho	0,0002 ± 0,00005	0,33 ± 0,06	0,00014 ± 0,000003	0,72 ± 0,01	0,45	0,00025	2,2
El Cedral	-	-	0,00013 ± 0,000003	0,63 ± 0,01	-	-	-
El Jazmín	0,0004 ± 0,00009	0,25 ± 0,07	0,00021 ± 0,000006	0,49 ± 0,02	0,11	0,00062	9,3
El Limón	0,0010 ± 0,00024	0,17 ± 0,06	0,00046 ± 0,000015	0,58 ± 0,02	0,16	0,00131	6,1
El Peñol	0,0000 ± 0,00001	0,45 ± 0,08	0,00001 ± 0,000000	1,07 ± 0,01	0,29	0,00003	3,5
El Sauce	0,0003 ± 0,00004	0,27 ± 0,05	0,00021 ± 0,000013	0,41 ± 0,05	0,22	0,00039	4,5
El Sena	0,0013 ± 0,00034	0,24 ± 0,08	0,00056 ± 0,000029	0,63 ± 0,04	0,10	0,00235	10,0
Francisco Romer	-	-	0,00006 ± 0,000001	0,80 ± 0,01	-	-	-
Granja Tibacuy	0,0000 ± 0,00001	0,25 ± 0,06	0,00003 ± 0,000001	0,54 ± 0,01	0,19	0,00007	5,2
Inmarco	0,0002 ± 0,00003	0,14 ± 0,05	0,00004 ± 0,000001	0,74 ± 0,01	0,10	0,00023	10,0
Jorge Villamil	0,0001 ± 0,00003	0,11 ± 0,07	0,00002 ± 0,000001	0,69 ± 0,02	0,07	0,00015	14,3
Julio Fernandez	0,0011 ± 0,00017	0,06 ± 0,04	0,00011 ± 0,000004	1,03 ± 0,03	0,09	0,00129	11,1
La Bella	-	-	0,00007 ± 0,000001	0,53 ± 0,01	-	-	-
La Catalina	-	-	0,00013 ± 0,000003	0,46 ± 0,01	-	-	-
La Montaña	0,0001 ± 0,00001	0,13 ± 0,03	0,00010 ± 0,000005	0,28 ± 0,04	0,40	0,00012	2,5
La Sirena	0,0004 ± 0,00011	0,15 ± 0,07	0,00013 ± 0,000004	0,70 ± 0,02	0,12	0,00057	8,3
LLanadas	0,0006 ± 0,00012	0,19 ± 0,06	0,00044 ± 0,000022	0,32 ± 0,04	0,17	0,00079	6,0
Luis Bustamante	0,0006 ± 0,00009	0,24 ± 0,05	0,00038 ± 0,000022	0,47 ± 0,04	0,17	0,00088	5,8
Luker	0,0000 ± 0,00000	1,16 ± 0,02	0,00001 ± 0,000000	0,45 ± 0,05	0,15	0,00002	6,5
Mande	0,0003 ± 0,00008	0,34 ± 0,07	0,00005 ± 0,000002	1,20 ± 0,02	0,13	0,00059	7,5
Manuel María Ma	-	-	0,00006 ± 0,000001	0,69 ± 0,01	-	-	-
Manuel Mejía	-	-	0,00026 ± 0,000006	0,26 ± 0,01	-	-	-
Maracay	0,0011 ± 0,00026	0,02 ± 0,06	0,00028 ± 0,000008	0,57 ± 0,02	0,08	0,00118	13,0
Miguel Valencia	0,0013 ± 0,00035	-0,17 ± 0,07	0,00009 ± 0,000003	0,70 ± 0,02	0,04	0,00077	23,6
Misiones	-	-	0,00013 ± 0,000004	0,60 ± 0,02	-	-	-
Naranjal	0,0002 ± 0,00004	0,18 ± 0,04	0,00008 ± 0,000002	0,66 ± 0,02	0,10	0,00036	10,0
Ospina Pérez	0,0005 ± 0,00011	-0,17 ± 0,06	0,00003 ± 0,000001	0,71 ± 0,02	0,04	0,00029	23,5
Paraguaito	0,0002 ± 0,00007	0,19 ± 0,08	0,00007 ± 0,000002	0,77 ± 0,02	0,11	0,00037	9,2
Planta de Trata	0,0006 ± 0,00012	0,23 ± 0,07	0,00039 ± 0,000019	0,59 ± 0,04	0,35	0,00071	2,9
Pueblo Bello	0,0001 ± 0,00002	0,11 ± 0,06	0,00008 ± 0,000005	0,42 ± 0,04	0,35	0,00013	2,9
Rafael Escobar	0,0002 ± 0,00004	0,21 ± 0,06	0,00009 ± 0,000002	0,53 ± 0,01	0,15	0,00024	6,7
Rosario	0,0008 ± 0,00019	0,00 ± 0,06	0,00020 ± 0,000006	0,48 ± 0,02	0,06	0,00076	16,6
Santa Ana	0,0003 ± 0,00006	0,15 ± 0,06	0,00008 ± 0,000003	0,76 ± 0,02	0,13	0,00036	7,5
Santa Bárbara	0,0001 ± 0,00002	0,40 ± 0,07	0,00005 ± 0,000002	0,67 ± 0,02	0,30	0,00011	3,4
Santa Helena	0,0032 ± 0,00093	-0,02 ± 0,07	0,00033 ± 0,000010	0,78 ± 0,02	0,06	0,00307	17,1
Santa Inés	0,0000 ± 0,00000	0,37 ± 0,03	0,00003 ± 0,000001	0,68 ± 0,03	0,32	0,00006	3,1
Santagueda	0,0001 ± 0,00002	0,24 ± 0,06	0,00004 ± 0,000001	0,59 ± 0,01	0,28	0,00009	3,5
Santiago Gutiérrez	0,0001 ± 0,00003	0,11 ± 0,07	0,00004 ± 0,000001	0,61 ± 0,02	0,15	0,00013	6,8
Yacopí	0,0001 ± 0,00005	0,21 ± 0,07	0,00009 ± 0,000001	0,51 ± 0,01	0,19	0,00021	5,1

dicho comportamiento a una no linealidad estadística. En todas las estaciones se encontró un resultado similar.

Dada la discrepancia encontrada en los espectros estimados, se hace necesario el uso de una herramienta un poco más estándar como la fuerza multifractal,  $\Delta\alpha$  (Koscielny-Bunde et al., 2006), la cual se define como la diferencia entre los valores máximos y mínimos de  $\alpha$ :

$$\Delta\alpha = \alpha_{max} - \alpha_{min}. \quad (28)$$

Esta fuerza multifractal es una medida de que tan fuerte es la multifractalidad presente en un proceso. En este trabajo se estimó  $\Delta\alpha$  para cada estación y con las cinco metodologías para momentos de ordenes entre -4 y 4, tratando de evitar la divergencia de momentos en la mayoría de estaciones.

Los resultados muestran los valores más altos de  $\Delta\alpha$  con MF-FA, del orden de 2 a 3 veces más de lo estimado con las otras metodologías.

En general, todas las estaciones presentan valores superiores a 0,66 y con valores promedio entre 1,0 y 1,8.

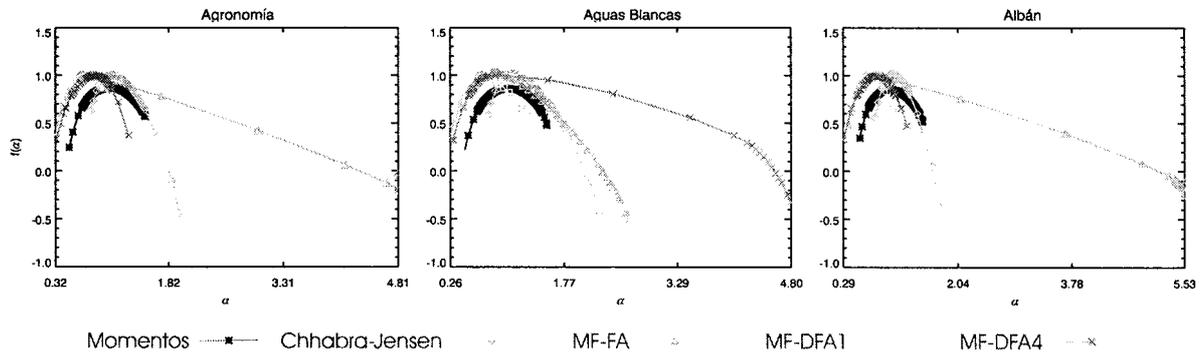


FIGURA 4. Espectros multifractales estimados para las estaciones Agronomía, Aguas Blancas y Albán, a partir de las cinco metodologías propuestas.

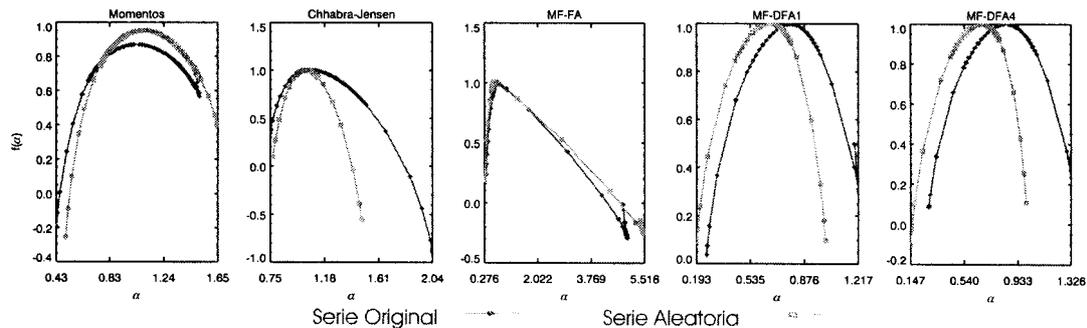


FIGURA 5. Espectros multifractales estimados para la estación Agronomía, a partir de las cinco metodologías propuestas. En azul se presenta el espectro de la serie original y en rojo la serie reordenada aleatoriamente.

Vale la pena tener presente la alta incertidumbre en  $\Delta\alpha$ , pues a medida que se aumenta el orden  $q$ , se incurre en estimaciones menos confiables debido a inestabilidad numérica, la aplicación de la transformada de Legendre y la posible presencia de transiciones de fase.

**4.4. Modelo para el exponente de Renyi.** En los trabajos de **Kantelhardt et al. (2003)** y **Koscielny-Bunde et al. (2006)** se presenta un modelo simple para representar el exponente de Renyi,  $\tau(q)$ , el cual se relaciona con el espectro multifractal,  $f(\alpha(q))$ , mediante la transformada de Legendre (véanse las ecuaciones 11a y 11b), aplicado a registros de caudal y lluvia. Este modelo permite caracterizar de forma “universal” el comportamiento multifractal de una medida mediante sólo dos parámetros independientes,  $a$  y  $b$ , los cuales podrían considerarse rasgos multifractales característicos del fenómeno estudiado, de tal forma que:

$$\tau(q) = -\ln(a^q + b^q)/\ln(2) \quad (29)$$

A partir del modelo de cascadas multiplicativas se deduce que  $a + b = 1$ ; sin embargo, **Kantelhardt et al. (2003)** amplía la condición, considerado que  $a + b \geq 1$  y  $0 < a < b < 1$ , con el objetivo de ajustarse más adecuadamente a los datos de caudal.

El espectro multifractal se puede estimar a partir de la transformada de Legendre como:

$$\alpha(q) = -\frac{1}{\ln(2)} \left[ \frac{a^q \ln(a) + b^q \ln(b)}{a^q + b^q} \right] \quad (30a)$$

$$f(\alpha(q)) = q\alpha(q) - \tau(q). \quad (30b)$$

Es importante entender el comportamiento de dichos modelos, pues sus parámetros serán rasgos característicos del fenómeno modelado. En particular, cuando  $a = b$  se encuentra un comportamiento monofractal, caracterizado por tener un exponente de Renyi que varía de forma lineal ( $\tau(q) = \eta q - 1$ ), con exponente Hölder ( $\alpha$ ) y espectro multifractal ( $f(\alpha)$ ) iguales a la dimensión de embebimiento  $D = 1$ . La pendiente  $\eta$  es 1 si  $a = b = 0,5$ ,

disminuye si los parámetros son mayores a 0,5 y aumenta en el caso contrario (véase la figura 6(a)).

A medida que  $a \neq b$ , el exponente  $\tau(q)$  comienza a variar de forma no lineal, cruzando siempre por el punto  $[q = -1, \tau = 0]$  y por lo tanto formando un espectro de singularidades bien definido, el cual es función de dichos parámetros, presentando una variación de la fuerza multifractal ( $\Delta\alpha = \alpha(-\infty) - \alpha(\infty)$ ) de la forma:

$$\Delta\alpha = \frac{\ln(\frac{b}{a})}{\ln(2)}, \quad (31)$$

mostrando que a medida que “a” sea más parecido a “b”, la fuerza multifractal se acercará a cero (caso monofractal) y a medida que se alejen aumenta la multifractalidad. Similarmente, es posible deducir una expresión para la posición del centro del espectro multifractal ( $\alpha(0)$ ):

$$\alpha(0) = \frac{\ln(b \cdot a)}{2 \ln(\frac{1}{2})}, \quad (32)$$

desplazando la curva  $f(\alpha)$  de derecha a izquierda a medida que “b · a” aumenta (véase la figura 6 (a) y (b)).

En este trabajo se ajustó el modelo descrito anteriormente con todas las metodologías, a excepción del método de Chhabra y Jensen, utilizando un algoritmo genético (Adeli & Hung, 1995) para encontrar los parámetros que minimicen el error cuadrático medio (RMSE) al adaptar el modelo en los valores del exponente de Renyi estimados (véase la figura 7). Con el método de los momentos ( $\mu_{a+b} = 1,05, \sigma_{a+b} = 0,05$ ), MF-DFA1 ( $\mu_{a+b} = 1,24, \sigma_{a+b} = 0,09$ ) y MF-DFA4 ( $\mu_{a+b} = 1,11, \sigma_{a+b} = 0,12$ ); se observa que para la mayoría de las estaciones se cumple  $a + b \geq 1$ , sin embargo con el MF-FA ( $\mu_{a+b} = 1, \sigma_{a+b} = 0,18$ ) se encuentra  $a + b < 1$  como el comportamiento general.

Cuando se analizaron los parámetros  $a$  y  $b$ , se encuentra que en general el parámetro característico es el primero, puesto que el segundo se mantiene aproximadamente constante en todas las metodologías y para todas las estaciones, presentando  $\mu_b = 0,84$  y  $\sigma_b = 0,02$  para MF-FA, MF-DFA1 y MF-DFA4; y  $\mu_b = 0,72$  y  $\sigma_b = 0,02$  el método de los momentos. En la siguiente sección se tratará de relacionar el parámetro  $a$  con la elevación desde el piso del valle en cada estación.

Finalmente, en todos los casos se tienen coeficientes de determinación mayores a 0.99, lo cual es un indicativo de un excelente ajuste de los datos al modelo, sin

embargo, desde el punto de vista del error cuadrático medio los peores ajustes del modelo se presentan con la metodología MF-FA ( $\mu_{RMSE} = 32, \sigma_{RMSE} = 0,13$ ). Similarmente, es posible ajustar el modelo a los datos del exponente de Hölder,  $\alpha(q)$  o al espectro de potencias,  $f(\alpha(q))$ , utilizando las expresiones (30a) y (30b), respectivamente. Sin embargo, el ajuste es menos acertado que en el caso del exponente de Renyi para el primer caso y prácticamente imposible en el segundo, lo cual, como se mencionó anteriormente, se le atribuye a inestabilidad numérica, la aplicación de la transformada de Legendre y la posible presencia de transiciones de fase.

**4.5. Elevación sobre el piso del valle.** En este caso no se encuentra una relación directa entre la elevación de las estaciones desde el piso del valle y los exponentes encontrados para la distribución de probabilidad empírica (exponente de probabilidad crítico), el espectro de potencias (exponente espectral) y el parámetro  $\alpha$  cuando se analizan todas las estaciones en cada vertiente, a diferencia de lo encontrado en el trabajo de **Harris et al., 1996**. Además, se definieron tres transectos los cuales descienden sobre la ladera oriental de la cordillera central, sin encontrar una relación clara (véase la figura 8). Con los anteriores análisis se esperaba encontrar una relación entre las variables mencionadas y una de las características más importantes de la lluvia en los Andes tropicales de Colombia, conocida como el óptimo pluviométrico (**Mesa et al., 1997**).

## 5. Conclusiones

La conclusión principal de este trabajo es el hallazgo de comportamiento multifractal en las series horarias de precipitación en los Andes tropicales de Colombia, lo cual abre un inmenso abanico de posibilidades de investigación, orientadas a comprender a fondo la climatología Colombiana y a mejorar la forma en la cual se modela gran cantidad de procesos hidrológicos, pues mejores predicciones requieren la identificación de la estructura de dichos procesos y la utilización de métodos no lineales. En el caso de fenómenos caóticos y altamente no lineales como lo es la hidroclimatología Colombiana, el determinismo es un camino “casi imposible” para su comprensión, lo cual potencializa la utilidad de descripciones estocásticas como la multifractal.

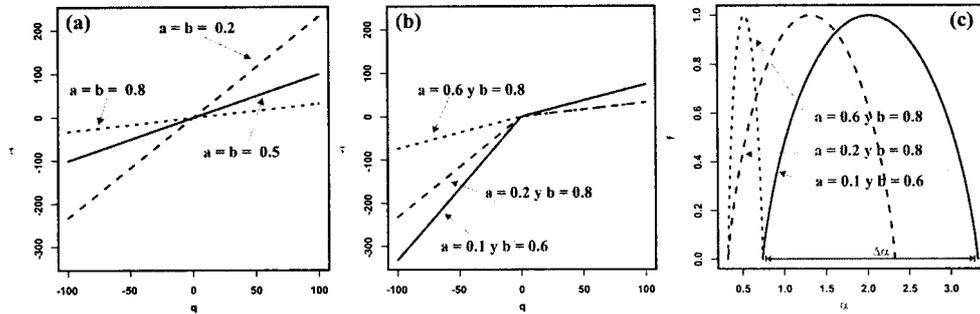


FIGURA 6. Modelo multifractal propuesto por Kantelhardt et al. (2003) y su variación en algunos casos particulares

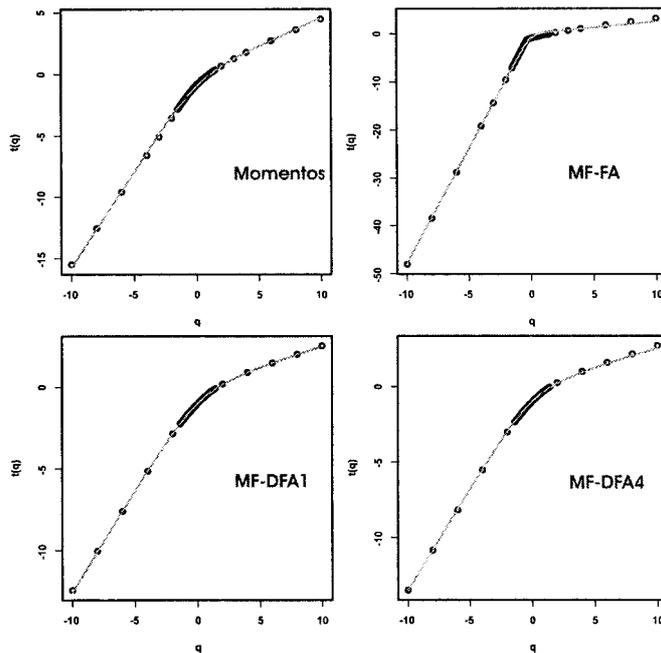


FIGURA 7. Modelos ajustados a los valores de  $\tau(q)$  estimados por las cuatro metodologías. La curva corresponde al ajuste

Los periodos sin precipitación (ceros en la serie de tiempo) juegan un papel fundamental en la estimación de espectro multifractal de la lluvia. Estos son un valor característico, pero en el caso de las mediciones, su presencia puede ser sobreestimada debido al truncamiento de los instrumentos, lo cual genera valores nulos donde no existen, jugando un papel fundamental en la estadística del análisis.

Se propuso un modelo potencial de colas pesadas para la distribución de probabilidad de la lluvia en cada estación ( $P(X > x) = Cx^{q_{cr}}$ ). En ambos casos fue posible dividir el dominio en dos regiones, para las cuales

se encontró que los valores extremos (región derecha) se acoplan adecuadamente al modelo multifractal, es decir,  $q_{cr} > 2$ , mientras para la otra región el monofractal es suficiente.

Los espectros de potencia mostraron un comportamiento de la forma  $E(f) \propto f^{-\beta}$ , donde para la mayoría de los casos se encontró que las series poseen  $\beta < 1$  y por tanto cumplen una de las condiciones para ser originadas a partir de una cascada multiplicativa (Sivakumar, 2001). En estos ajustes se encontraron dos regiones características para varias de las estaciones, en las cuales se tiene  $\beta < 1$ .

El proceso de calibración muestra que las metodologías presentaron buenos resultados en la estimación del espectro multifractal de las medidas binomiales generadas sintéticamente. En este proceso se resalta el método MF-FA como el más eficiente, debido a sus buenas estimaciones y simpleza algorítmica. Sin embargo, es de anotar que en el caso de series geofísicas la remoción de tendencias si juega un papel significativo.

El espectro multifractal de las series de precipitación horaria, encontrado con las cinco las metodologías difiere considerablemente en su estimación, lo cual deja

este tipo de descripción con un alto grado de incertidumbre, al menos desde el punto de vista cuantitativo, pues todas las metodologías convergen a evidenciar un escalamiento múltiple, pero con diferentes características.

En algunos casos se presentan valores de  $f(\alpha)$  menores que cero, sin embargo no parecen ser producto de las características propias de la serie (Mandelbrot, 1990, 2003), sino de inestabilidades numéricas por divergencia de los momentos, dada la cercanía al cero y a que corresponden a valores de momento mayores a  $|q_{cr}|$ .

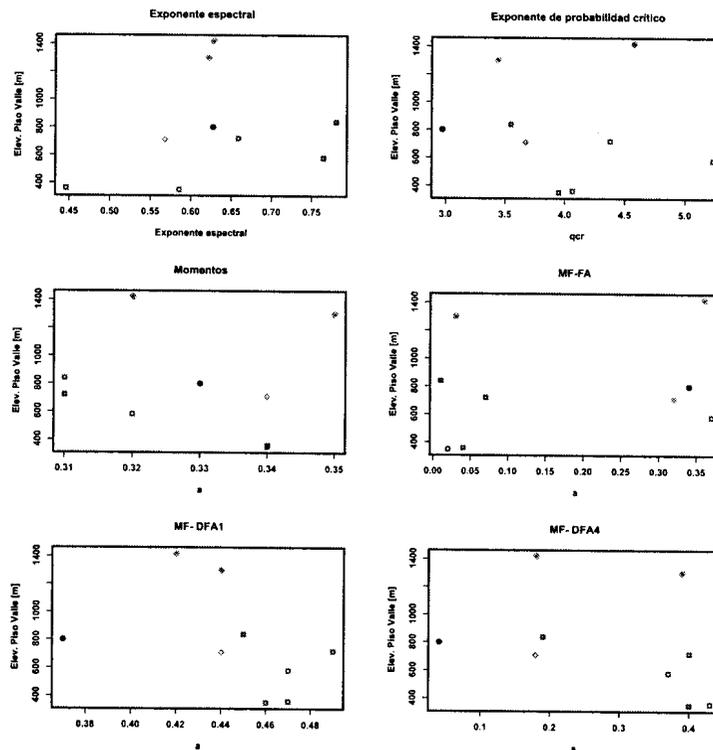


FIGURA 8. Análisis de la elevación sobre el piso del valle de las estaciones en los transectos: 1) El Cedral, Bremen y El Sena (círculos negros), 2) Santa Helena, Naranjal Santa, Ana, Santaguada y Luker (cuadros azules) y 3) El Cedral, Bremen y Maracay (diamantes rojos) y los parámetros  $\beta$ ,  $q_{cr}$  y  $a$ .

Se encuentra que la multifractalidad presente en las series de precipitación horaria se puede atribuir a una no linealidad estadística (Olsson, 1995) y no a correlaciones de largo plazo en los registros.

Los exponentes encontrados para la distribución de probabilidad empírica (exponente de probabilidad crítico), el espectro de potencias (exponente espectral) y el parámetro  $a$ , estimados para cada estación y para

cada metodología, no presentan una relación clara con la ubicación espacial, ni con la elevación sobre el piso del valle.

Finalmente, las metodologías utilizadas en este trabajo para la estimación del espectro multifractal tienen falencias y problemas al trabajar con series geofísicas como la lluvia, por lo tanto se hace necesaria la investigación de otras metodologías que permitan hacer un

estudio cuantitativo más confiable y por ende la aplicación del multiescalamiento de la lluvia en problemas tan importantes como el de los modelos lluvia-escorrentía, la generación de curvas de intensidad frecuencia y duración (IDF), y el diseño hidrológico, entre otros. Ante este panorama, se potencializa el uso del modelo mencionado en la sección 4.4, dejando de lado la transformada de Legendre y sus restricciones ante posibles transiciones de fase, y utilizando las definiciones matemáticas para modelar espectro de singularidades del fenómeno.

**Agradecimientos.** El trabajo de **G. Poveda** hace parte del Programa de Investigación del Grupo GRECIA, financiado por COLCIENCIAS. Al Centro Nacional de Investigaciones del Café (CENICAFE) y a las Empresas Públicas de Medellín (EPM), por el suministro de la información.

#### Referencias

- Adeli, H. & S. L. Hung** (1995), *Machine learning. Neural networks, genetic algorithms and fuzzy systems*, John Wiley and Sons, Canada.
- Agudelo P., P. Arias, & L. Salazar** (2001), *Caracterización del ciclo diario de precipitación en los Andes tropicales de Colombia. Región Centro*, Tesis Ingeniería Civil, Universidad Nacional de Colombia, Facultad de Minas.
- Álvarez, F. & V. Toro** (2001), *Caracterización del ciclo diario de precipitación en los Andes tropicales de Colombia. Región Sur*, Tesis Ingeniería Civil, Universidad Nacional de Colombia, Facultad de Minas.
- Barnsley, M. F.** (1993), *Fractals Everywhere*, 2<sup>nd</sup> ed., Academic Press, San Diego.
- Billingsley, P.** (1965), *In Ergodic Theory and Information*, Wiley, New York.
- Braun, E.** (1996), *Caos, fractales y cosas raras*. Fondo de Cultura Económica, México.
- Chhabra, A. & R. Jensen** (1989), Direct determination of the  $f$  singularity spectrum, *Physical Review Letters*, 62(12), 1372–1330.
- Chhabra, A., C. Meneveau, R. Jensen, & K. R. Sreenivasan** (1989), Direct determination of the  $f$  singularity spectrum and its applications to fully developed turbulence, *Physical Review A*, 40(9), 5284–5294.
- Chorin, A.** (1994), *Vorticity and Turbulence*, Springer-Verlag, New York.
- Courant R. & D. Hilbert** (1953), *Methods of Mathematical Physics*, Vol. 1, Interscience Publishers, New York.
- de Lima M. I. P & J. Grasman** (1999), Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. *Journal of Hydrology*, 220, 1–11.
- Douglas, E. M. & A. P. Barros** (2003), Probable maximum precipitation estimation using multifractals: Application in the eastern United States, *Journal of Hydrometeorology*, 1012–1024.
- Flores, C.** (2004), Multiplicative cascade models for rain in hydro-meteorological disasters risk management, ASTIN-Kolloquium, Bergen, Norway.
- Grassberger, P., R. Badii, & A. Politi**, (1988), Scaling laws for invariant measures on hyperbolic and nonhyperbolic attractors, *Physics and Astronomy*, 51, 135–178.
- Gupta V. K. & E. Waymire** (1993), A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32, 251–267.
- Harris, D., M. Menabde, A. Seed & G. Austin** (1996), Multifractal characterization of rain fields with a strong orographic influence, *Journal of Geophysical Research*, D21, 101, 405–26.
- Kantelhardt, J. W., D. Rybsky, S. Zschiegner, P. Braun, E. Koscielny-Bunde, V. Livina, S. Havlin, & A. Bunde** (2003), Multifractality of river runoff and precipitation: comparison of fluctuation analysis and wavelet methods, *Physica A*, 330, 240–245.
- Koscielny-Bunde, E., J. Kantelhardt, P. Braun, A. Bunde, & S. Havlin** (2006), Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies, *Journal of Hydrology*, 322 (1–4), 120–137.
- Mandelbrot, B.** (1989), Multifractal measures especially for the geophysicist, *PAGEOPH*, 131(12), 38.
- Mandelbrot, B.** (1990), Negative fractal dimensions and multifractals, *Physica A*, 163, 10.
- Mandelbrot, B.** (2003), Multifractal power law distributions: Negative and critical dimensions and other anomalies, explained by a simple example, *Journal of Statistical Physics*, 110, 739–774.
- Mesa O. J. & Poveda G.** (1993), The Hurst effect: The scale of fluctuation approach, *Water Resources Research*, 29, 3995–4002.
- Mesa O., G. Poveda & L. Carvajal** (1997), *Introducción al Clima de Colombia*. Universidad Nacional de Colombia, Facultad de Minas.
- Montgomery, D. & G. Runger** (1996), *Probabilidad y estadística aplicadas a la ingeniería*, McGraw-Hill, Mexico. p. 1000.
- Moreno, H. & G. Poveda** (2004), Colas pesadas en el análisis probabilístico de la lluvia y exponente de Hurst durante las fases del ENSO, *XVI Seminario Nacional de Hidráulica e Hidrología*.
- Olsson, J.** (1995), Limits and characteristics of the multifractal behavior of a high-resolution rainfall time series, *Nonlinear Processes in Geophysics*, 2, 23–29.
- Olsson, J. & J. Niemczynowicz** (1996), Multifractal analysis of daily spatial rainfall distributions, *J. Hydrol.* 187, 29–43.
- Ott, E., W. D. Withers, & J. A. Yorke**, (1984), Is the dimension of chaotic attractors invariant under coordinate

- changes?, *Journal of Statistical Physics*, 36, 687–697.
- Peitgen, H., H. Jürgens, & D. Saupe** (1992), *Chaos and Fractals New Frontiers of Science (Appendix B. Multifractal Measures)*, Hamilton Printing Co., New York.
- Poveda, G., O. Mesa, L. Salazar, P. Arias, H. Moreno, S. Viera, P. Agudelo, V. Toro, & J. Álvarez** (2005), The diurnal cycle of precipitation in the tropical Andes of Colombia, *Monthly Weather Review*, 133, 228–240.
- Schertzer, D. & S. Lovejoy** (1996), Notes and correspondence. universal multifractals do exist!: Comments on statistical analysis of mesoscale rainfall as a random cascade, *Journal of Applied Meteorology* 136, 1296–1303.
- Schertzer, D., P. Hubert & S. Lovejoy** (2003), Scaling, Multifractals and Predictions in Ungaged Basins: Where We've Been, Where We're Going?, in *Prediction of Ungaged Basins*, An IAHS Initiative, Eds. P. Hubert et al., IAHS Press, Wallingford UK.
- Schertzer, D., S. Lovejoy, F. Schmitt, Y. Chigirinskaya, & D. Marsan** (1997), Multifractal cascade dynamics and turbulent intermittency. *Fractals*, 5, 427–471.
- Seuront, L., F. Schmitt, Y. Lagadeuc, D. Schertzer & S. Lovejoy** (1999), Universal Multifractal Analysis as a tool to characterize multiscale intermittent patterns: example of phytoplankton distribution in turbulent coastal waters, *Journal of Plankton Research*, 21, 877–922.
- Sivakumar, B.** (2001), Is a chaotic multi-fractal approach for rainfall possible?, *Hydrological Processes*, 15, 943–955.
- Svensson, C., J. Olsson & R. Berndtsson** (1996), Multifractal properties of daily rainfall in two different climates, *Water Resources Research*, 32, 2463–2472.
- Tessier, Y., S. Lovejoy, & D. Schertzer** (1993), Universal multifractals: Theory and observations for rain and clouds, *Journal of Applied Meteorology*, 223–250.
- Vieira S. & H. Moreno** (2001), *Caracterización del ciclo diurno de precipitación en los Andes tropicales de Colombia. Región Sur*, Tesis Ingeniería Civil, Universidad Nacional de Colombia, Facultad de Minas.

Recibido: abril 19 de 2007

Aceptado para publicación: diciembre 3 de 2008

# NON-GAUSSIANITY AND LOOP CORRECTIONS IN A QUADRATIC TWO-FIELD SLOW-ROLL MODEL OF INFLATION. PART I

By

Yeinzon Rodríguez<sup>1,2</sup>

## Abstract

**Rodríguez, Y.:** Non-Gaussianity and Loop Corrections in a Quadratic Two-Field Slow-Roll Model of Inflation. Part I. Rev. Acad. Colomb. Cienc. **32** (125): 503-513, 2008. ISSN 0370-3908.

I am showing in this paper that it is possible to attain very high, *including observable*, values for the level of non-gaussianity  $f_{NL}$  in a particular quadratic two-field slow-roll model of inflation with canonical kinetic terms. This is done by taking care of loop corrections both in the spectrum  $P_\zeta$  and the bispectrum  $B_\zeta$  of the primordial curvature perturbation  $\zeta$ . Sizable values for  $f_{NL}$  arise even if  $\zeta$  is generated during inflation. Five issues are considered when constraining the available parameter space: 1. we must ensure that we are in a perturbative regime so that the  $\zeta$  series expansion, and its truncation, are valid. 2. we must apply the correct condition for the (possible) loop dominance in  $B_\zeta$  and/or  $P_\zeta$ . 3. we must satisfy the spectrum normalisation condition. 4. we must satisfy the spectral tilt constraint. 5. we must have enough inflation to solve the horizon problem.

**Key words:** Primordial curvature perturbation, non-gaussianity, slow-roll inflationary models.

<sup>1</sup> Centro de Investigaciones, Universidad Antonio Nariño, Cra. 3 Este No. 47A-15, Bogotá D.C., Colombia, Correo electrónico: yeinzon.rodriguez@uan.edu.co

<sup>2</sup> Escuela de Física, Universidad Industrial de Santander, Ciudad Universitaria, Bucaramanga, Colombia, Correo electrónico: yrodrig@uis.edu.co

## Resumen

Muestro en este artículo que es posible obtener valores altos, *incluso observables*, para el nivel de no gaussianidad  $f_{NL}$  en un particular modelo inflacionario del tipo slow-roll con un potencial escalar cuadrático de dos componentes y términos cinéticos canónicos. Lo anterior se hace teniendo en cuenta correcciones de lazo tanto en el espectro  $P_\zeta$  como en el bispectro  $B_\zeta$  de la perturbación primordial en la curvatura  $\zeta$ . Se obtienen valores grandes para  $f_{NL}$  incluso si  $\zeta$  es generada durante inflación. Se tienen en cuenta cinco restricciones que reducen la ventana de parámetros disponible: 1. debemos estar seguros de estar trabajando en un régimen perturbativo de tal manera que la expansión en serie de  $\zeta$ , y su truncamiento, sean válidas. 2. debemos aplicar la condición correcta acerca del (posible) dominio de las correcciones de lazo en  $B_\zeta$  y/o  $P_\zeta$ . 3. debemos satisfacer la condición de normalización del espectro 4. debemos satisfacer el índice espectral observado. 5. debemos asegurar el monto mínimo de inflación para resolver el problema de horizonte.

**Palabras clave:** Perturbación primordial en la curvatura, no gaussianidad, modelos inflacionarios del tipo slow-roll.

## 1 Introduction

The primordial curvature perturbation  $\zeta$  (Dodelson, 2003; Liddle & Lyth, 2000; Mukhanov, 2005; Weinberg, 2008), and its  $\delta N$  expansion (Lyth, Malik, & Sasaki, 2005; Lyth & Rodríguez, 2005a; Sasaki & Stewart, 1996; Starobinsky, 1985), was the subject of study in two recent papers (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b). The authors were interested in how well the convergence of the  $\zeta$  series was understood, and if the traditional naive arguments to cut out the  $\zeta$  series at second order (Lyth & Rodríguez, 2005a; Zaballa, Rodríguez, & Lyth, 2006), keeping only the tree-level terms to study the statistical descriptors of  $\zeta$  (Alabidi, 2006; Battfeld & Easther, 2007; Byrnes, Choi, & Hall, 2008; Byrnes, Sasaki, & Wands, 2006; Seery & Lidsey, 2007; Vernizzi & Wands, 2006; Yokoyama, Suyama, & Tanaka, 2007; Yokoyama, Suyama, & Tanaka, 2008a; Yokoyama, Suyama, & Tanaka, 2008b), were reliable<sup>3</sup>. The authors argued that a previous study of the  $\zeta$  series convergence, the viability of a perturbative regime, and the relative weight of the loop contributions against the tree-level terms, were completely necessary and in some cases surprising. For instance, the levels of non-gaussianity  $f_{NL}$  and  $\tau_{NL}$  in the bispectrum  $B_\zeta(k_1, k_2, k_3)$  and trispectrum  $T_\zeta(k_1, k_2, k_3, k_4)$  of  $\zeta$  respectively, for slow-roll inflationary models with canonical kinetic terms (Liddle & Lyth, 2000; Lyth, 2008; Lyth & Riotto, 1999), are usually thought to be of order  $\mathcal{O}(\epsilon_i, \eta_i)$  (Battfeld & Easther, 2007; Seery & Lidsey, 2007; Seery, Sloth, & Vernizzi, 2008; Vernizzi & Wands, 2006; Yokoyama, Suyama, & Tanaka, 2007),

were  $\epsilon_i$  and  $\eta_i$  are the slow-roll parameters with  $\epsilon_i, |\eta_i| \ll 1$  (Lyth & Riotto, 1999). However, in order to reach such a conclusion, only the tree-level terms were considered without giving a satisfactory explanation of why the loop contributions are comparatively suppressed, neither why  $\zeta$  is indeed represented by the  $\delta N$  expansion, nor why the truncated  $\delta N$  expansion may be used. A couple of papers (Alabidi, 2006; Byrnes, Choi, & Hall, 2008) show that large, *and observable*, non-gaussianity in  $B_\zeta$  is indeed possible for certain classes of *slow-roll* models with *canonical* kinetic terms and special trajectories in field space, relying only on the tree-level terms. Nonetheless, although the resultant phenomenology from these two papers is very interesting, a satisfactory argument about the tree-level dominance over the loop corrections was still lacking. In Refs. (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b), the authors addressed these issues, and showed how important the requirements to guarantee the  $\zeta$  series convergence and the existence of a perturbative regime are. Supported in Ref. (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a), I show in this paper that for a particular quadratic two-field *slow-roll* inflationary model with *canonical* kinetic terms, the one-loop correction to  $B_\zeta$  might be much bigger than the tree-level terms, giving as a result large, *and observable*, non-gaussianity parameterised by  $f_{NL}$ . Based on the same grounds, the level of non-gaussianity  $\tau_{NL}$  for the same slow-roll model studied here will be the subject of study in a companion paper (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008c).

<sup>3</sup>In this paper, I follow the terminology of Ref. (Byrnes, Koyama, Sasaki, & Wands, 2007) to identify the tree-level terms and the loop contributions, within the  $\delta N$  formalism, in a diagrammatic approach. The associated diagrams are called *Feynman-like diagrams*.

## 2 Basic definitions and observation

Given the probability distribution function  $f(\zeta)$ , for the primordial curvature perturbation  $\zeta(\mathbf{x}, t)$ , there are an infinite number of standardized moments that work as statistical descriptors of  $\zeta(\mathbf{x}, t)$ :

$$\text{the mean value : } m_\zeta(1) \equiv \langle \zeta \rangle = \int \zeta f(\zeta) d\zeta, \quad (1)$$

$$\text{the variance : } m_\zeta(2) \equiv \int (\zeta - \langle \zeta \rangle)^2 f(\zeta) d\zeta, \quad (2)$$

$$\text{the skewness : } m_\zeta(3) \equiv \frac{\int (\zeta - \langle \zeta \rangle)^3 f(\zeta) d\zeta}{[m_\zeta(2)]^{3/2}}, \quad (3)$$

$$\text{the kurtosis : } m_\zeta(4) \equiv \frac{\int (\zeta - \langle \zeta \rangle)^4 f(\zeta) d\zeta}{[m_\zeta(2)]^2}, \quad (4)$$

and so on.

Departures from the exact gaussianity come either from non-vanishing odd standardized moments  $m_\zeta(n)$  with  $n \geq 3$ , in which case the probability distribution function is non-symmetric around the mean value, or from higher ( $n \geq 4$ ) even standardized moments different to products of the variance, in which case the probability distribution function continues to be symmetric around the mean value although its “peakedness”<sup>4</sup> is bigger than that for a gaussian function, or from both of them.

Working in momentum space, the standardized moments of the probability distribution function have a direct connection with the correlation functions for the Fourier modes  $\zeta_{\mathbf{k}} \equiv \int d^3k \zeta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}$  defined in flat space. As the  $n$ -point correlators of  $\zeta_{\mathbf{k}}$  are generically defined in terms of spectral

functions of the wavevectors involved<sup>5</sup>:

$$\begin{aligned} \text{two - point correlator} &\rightarrow \text{spectrum } P_\zeta : \\ \langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \rangle &\equiv (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2) P_\zeta(k), \end{aligned} \quad (5)$$

$$\begin{aligned} \text{three - point correlator} &\rightarrow \text{bispectrum } B_\zeta : \\ \langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle &\equiv (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times \\ &\quad \times B_\zeta(k_1, k_2, k_3), \end{aligned} \quad (6)$$

$$\begin{aligned} \text{four - point correlator} &\rightarrow \text{trispectrum } T_\zeta : \\ \langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \zeta_{\mathbf{k}_4} \rangle &\equiv (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4) \times \\ &\quad \times T_\zeta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4), \end{aligned} \quad (7)$$

and so on,

the standardized moments of the distribution are then written in terms of momentum integrals of the spectral functions for the modes  $\zeta_{\mathbf{k}}$ :

$$\text{the variance : } m_\zeta(2) = \int \frac{d^3k}{(2\pi)^3} P_\zeta(k), \quad (8)$$

$$\text{the skewness : } m_\zeta(3) = \frac{\int \frac{d^3k_1 d^3k_2}{(2\pi)^6} B_\zeta(k_1, k_2, k_3)}{\left[ \int \frac{d^3k}{(2\pi)^3} P_\zeta(k) \right]^{3/2}}, \quad (9)$$

$$\text{the kurtosis : } m_\zeta(4) = \frac{\int \frac{d^3k_1 d^3k_2 d^3k_3}{(2\pi)^9} T_\zeta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)}{\left[ \int \frac{d^3k}{(2\pi)^3} P_\zeta(k) \right]^2}, \quad (10)$$

and so on.

Non-gaussianity in  $\zeta$  is, therefore, associated with non-vanishing higher order spectral functions, starting from the bispectrum  $B_\zeta$ .

Now I will parametrize the spectral functions of  $\zeta$  in terms of quantities which are the ones for which observational bounds are given. Because of the direct connection between these quantities and the standardized moments of the probability distribution function  $f(\zeta)$ , I may also call these

<sup>4</sup>Higher even standardized moments different to products of the variance mean more of the variance is due to infrequent extreme deviations, as opposed to frequent modestly-sized deviations.

<sup>5</sup>Statistical homogeneity, i.e., invariance of the correlation functions under translations, requires the presence of the Dirac delta functions (Ackerman, Carroll, & Wise, 2007; Carroll, Tseng, & Wise, 2008; Dimopoulos, Lyth, & Rodríguez, 2008). Statistical isotropy, i.e., invariance of the correlation functions under rotations, requires that the spectrum  $P_\zeta$  and bispectrum  $B_\zeta$  are functions of the wavenumbers only (Ackerman, Carroll, & Wise, 2007; Carroll, Tseng, & Wise, 2008; Dimopoulos, Lyth, & Rodríguez, 2008). For the trispectrum  $T_\zeta$  and the other higher order spectral functions, the momentum dependence also involves the direction of the wavevectors.

quantities as the statistical descriptors for  $f(\zeta)$ . The spectrum  $P_\zeta$  is parametrized in terms of an amplitude  $\mathcal{P}_\zeta^{1/2}$  and a spectral index  $n_\zeta$  which measures the deviation from an exactly scale-invariant spectrum (**Liddle & Lyth, 2000; Weinberg, 2008**):

$$P_\zeta(k) \equiv \frac{2\pi^2}{k^3} \mathcal{P}_\zeta \left( \frac{k}{aH} \right)^{n_\zeta - 1}, \quad (11)$$

where  $a$  is the global expansion parameter and  $H = \dot{a}/a$  is the Hubble parameter, with the dot meaning a derivative with respect to cosmic time. The bispectrum  $B_\zeta$  and trispectrum  $T_\zeta$  are parametrized in terms of products of the spectrum  $P_\zeta$ , and the quantities  $f_{NL}$  and  $\tau_{NL}$  respectively<sup>6</sup> (**Boubekeur & Lyth, 2006; Maldacena, 2003**):

$$B_\zeta(k_1, k_2, k_3) \equiv \frac{6}{5} f_{NL} \left[ P_\zeta(k_1)P_\zeta(k_2) + \text{cyclic permutations} \right], \quad (12)$$

$$T_\zeta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{2} \tau_{NL} \left[ P_\zeta(k_1)P_\zeta(k_2)P_\zeta(|\mathbf{k}_1 + \mathbf{k}_4|) + \text{cyclic permutations} \right]. \quad (13)$$

Higher order spectral functions would be parametrized in an analogous way. Given the present observational state-of-the-art,  $n_\zeta$ ,  $f_{NL}$ , and  $\tau_{NL}$  are the statistical descriptors that discriminate among models for the origin of the large-scale structure once  $\mathcal{P}_\zeta^{1/2}$  has been fixed to the observed value. Since non-vanishing higher order spectral functions such as  $B_\zeta$  and  $T_\zeta$  imply non-gaussianity in the primordial curvature perturbation  $\zeta$ , the statistical descriptors  $f_{NL}$  and  $\tau_{NL}$  are usually called the levels of non-gaussianity.

The COBE satellite<sup>7</sup> provided us with a reliable value for the spectral amplitude  $\mathcal{P}_\zeta^{1/2}$  (**Bunn & White, 1997**):  $\mathcal{P}_\zeta^{1/2} = (4.957 \pm 0.094) \times 10^{-5}$  which is usually called the COBE normalisation. As regards the spectral index, the latest data release and analysis from the WMAP satellite<sup>8</sup> shows that  $n_\zeta = 0.960 \pm 0.014$  (**Komatsu et. al., 2008**) which rejects exact scale invariance at more than  $2\sigma$ . Such a result has been extensively used to constrain inflation model building (**Alabidi & Lidsey, 2008; Alabidi & Lyth, 2006**), and although several classes of inflationary models have been ruled out through the spectral index, lots of models are still allowed; that is why it is so important an appropriate knowledge of the statistical descriptors  $f_{NL}$  and

$\tau_{NL}$ . Present observations show that the primordial curvature perturbation  $\zeta$  is almost, but not completely, gaussian. The level of non-gaussianity  $f_{NL}$  in the bispectrum  $B_\zeta$ , after five years of data from NASA's WMAP satellite, is in the range  $-9 < f_{NL} < 111$  at  $2\sigma$  (**Komatsu et. al., 2008**). There is at present no observational bound on the level of non-gaussianity  $\tau_{NL}$  in the trispectrum  $T_\zeta$  although it was predicted that COBE should either measure  $\tau_{NL}$  or impose the lower bound  $|\tau_{NL}| \lesssim 10^8$  (**Boubekeur & Lyth, 2006; Okamoto & Hu, 2002**). It is expected that future WMAP data releases will either detect non-gaussianity or reduce the bounds on  $f_{NL}$  and  $\tau_{NL}$  at the  $2\sigma$  level to  $|f_{NL}| \lesssim 40$  (**Komatsu & Spergel, 2001**) and  $|\tau_{NL}| \lesssim 2 \times 10^4$  (**Kogo & Komatsu, 2006**) respectively. The ESA's PLANCK satellite<sup>9</sup> (**The Planck Collaboration, 2006**), whose launch is currently scheduled in the spring of 2009, promises to reduce the bounds to  $|f_{NL}| \lesssim 10$  (**Komatsu & Spergel, 2001**) and  $|\tau_{NL}| \lesssim 560$  (**Kogo & Komatsu, 2006**) at the  $2\sigma$  level if non-gaussianity is not detected. In addition, by studying the 21-cm emission spectral line in the cosmic neutral Hydrogen prior to the era of reionization, it is also possible to know about the levels of non-gaussianity  $f_{NL}$  and  $\tau_{NL}$ ; the 21-cm background anisotropies capture information about the primordial non-gaussianity better than any high resolution map of cosmic microwave background radiation: an experiment like this could reduce the bounds on the non-gaussianity levels to  $|f_{NL}| \lesssim 0.2$  (**Cooray, 2006; Cooray, Li, & Melchiorri, 2008**), and  $|\tau_{NL}| \lesssim 20$  (**Cooray, Li, & Melchiorri, 2008**) at the  $2\sigma$  confidence. Finally, it is worth stating that there have been recent claims about the detection of non-gaussianity in the bispectrum  $B_\zeta$  of  $\zeta$  from the WMAP 3-year data (**Yadav & Wandelt, 2008**). Such claims, which report a rejection of  $f_{NL} = 0$  at more than  $2\sigma$  ( $26.9 < f_{NL} < 146.7$ ), are based on the estimation of the bispectrum while using some specific foreground masks. The WMAP 5-year analysis (**Komatsu et. al., 2008**) shows a similar behaviour when using those masks, but reduces the significance of the results when other more conservative masks are included allowing again the possibility of exact gaussianity.

<sup>6</sup>There is actually a sign difference between the  $f_{NL}$  defined here and that defined in Ref. (**Maldacena, 2003**). The origin of the sign difference lies in the way the observed  $f_{NL}$  is defined (**Komatsu & Spergel, 2001**), through the Bardeen's curvature perturbation (**Bardeen, 1980**):  $\Phi^B = \Phi_L^B + f_{NL}(\Phi_L^B)^2$  with  $\Phi^B = (3/5)\zeta$ , and the way  $f_{NL}$  is defined in Ref. (**Maldacena, 2003**), through the gauge invariant Newtonian potential:  $\Phi^N = \Phi_L^N + f_{NL}(\Phi_L^N)^2$  with  $\Phi^N = -(3/5)\zeta$  (**Komatsu, 2008**).

<sup>7</sup>NASA's COBE mission homepage: <http://lambda.gsfc.nasa.gov/product/cobe/>.

<sup>8</sup>NASA's WMAP mission homepage: <http://wmap.gsfc.nasa.gov/>.

<sup>9</sup>ESA's PLANCK mission homepage: <http://planck.esa.int/>.

### 3 The model

According to the classification of inflationary models proposed in Ref. (Dodelson, Kinney, & Kolb, 1997), the small-field models are those of the form that would be expected as a result of spontaneous symmetry breaking, with a field initially near an unstable equilibrium point (usually taken to be at the origin) and rolling toward a stable minimum  $\langle \phi \rangle \neq 0$ . Thus, inflation occurs when the field is small relative to its expectation value  $\phi \ll \langle \phi \rangle$ . Some interesting examples are the original models of new inflation (Albrecht & Steinhardt, 1982; Linde, 1982), modular inflation from string theory (Dimopoulos & Lazarides, 2006), natural inflation (Freese, Frieman, & Olinto, 1990), and hilltop inflation (Boubekeur & Lyth, 2005). As a result, the inflationary potential for small-field models may be taken as

$$V = \sum_i \Lambda_i \left[ 1 - \left( \frac{\phi_i}{\mu_i} \right)^p \right], \quad (14)$$

where the subscript  $i$  here denotes the relevant quantities of the  $i$ th field,  $p$  is the same for all fields, and  $\Lambda_i$  and  $\mu_i$  are the parameters describing the height and tilt of the potential of the  $i$ th field.

While Ref. (Ahmad, Piao, & Quiao, 2008) studies the spectrum of  $\zeta$  for general values of the parameter  $p$  and an arbitrary number of fields, assuming  $\zeta$  series convergence and tree-level dominance, I will specialize to the  $p = 2$  case for two fields  $\phi$  and  $\sigma$ :

$$V = V_0 \left( 1 + \frac{1}{2} \eta_\phi \frac{\phi^2}{m_P^2} + \frac{1}{2} \eta_\sigma \frac{\sigma^2}{m_P^2} \right), \quad (15)$$

where I have traded the expressions

$$\Lambda_1 + \Lambda_2 \quad \text{for} \quad V_0, \quad (16)$$

$$\frac{\Lambda_1}{\mu_1^2} \quad \text{for} \quad -V_0 \frac{\eta_\phi}{2m_P^2}, \quad (17)$$

and

$$\frac{\Lambda_2}{\mu_2^2} \quad \text{for} \quad -V_0 \frac{\eta_\sigma}{2m_P^2}, \quad (18)$$

and defined  $m_P$  as the reduced Planck mass. On doing this, and assuming that the first term in Eq. (15) dominates,  $\eta_\phi < 0$  and  $\eta_\sigma < 0$  become the usual  $\eta$  slow-roll parameters associated with the fields  $\phi$  and  $\sigma$ .

I have chosen for simplicity the  $\sigma = 0$  trajectory (see Fig. 1) since in that case the potential in Eq. (15) reproduces for some number of e-folds the hybrid inflation scenario (Linde, 1994) where  $\phi$  is the inflaton and  $\sigma$  is the waterfall field. Non-gaussianity in the bispectrum  $B_\zeta$  of  $\zeta$  for such a model has been studied in Refs. (Alabidi, 2006; Byrnes, Choi,

& Hall, 2008; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Enqvist & Väihkönen, 2004; Lyth & Rodríguez, 2005a; Lyth & Rodríguez, 2005b; Väihkönen, 2005; Zamballa, Rodríguez, & Lyth, 2006); in particular, Ref. (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a) shows that the one-loop correction dominates over the tree-level terms, generating in this way large values for  $f_{NL}$  even if  $\zeta$  is generated during inflation. Refs. (Alabidi, 2006; Byrnes, Choi, & Hall, 2008), in contrast, work only at tree-level with the same potential as Eq. (15) but relaxing the  $\sigma = 0$  condition, finding that large values for  $f_{NL}$  are possible for a small set of initial conditions. Ref. (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b) studies the trispectrum  $T_\zeta$  of  $\zeta$  in this model considering dominant loop corrections with  $\sigma = 0$ ; its results show that large values for  $\tau_{NL}$  are generated even if  $\zeta$  is generated during inflation.

The slow-roll conditions for single-field inflationary models with canonical kinetic terms read

$$\dot{\phi}^2 \ll V(\phi), \quad (19)$$

$$|\ddot{\phi}| \ll |3H\dot{\phi}|, \quad (20)$$

where  $\phi$  is the inflaton field and  $V(\phi)$  is the scalar field potential. On defining the slow-roll parameters  $\epsilon$  and  $\eta_\phi$  as (Liddle & Lyth, 2000)

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \quad (21)$$

$$\eta_\phi \equiv \epsilon - \frac{\ddot{\phi}}{H\dot{\phi}}, \quad (22)$$

the slow-roll conditions in Eqs. (19) and (20) translate into strong constraints for the slow-roll parameters:  $\epsilon, |\eta_\phi| \ll 1$ . Multifield slow-roll models may also be characterized by a set of slow-roll parameters which generalize those in Eqs. (21) and (22) (Lyth & Riotto, 1999):

$$\epsilon_i \equiv \frac{m_P^2}{2} \left( \frac{V_i}{V} \right)^2, \quad (23)$$

$$\eta_i \equiv m_P^2 \frac{V_{ii}}{V}. \quad (24)$$

By writing the slow-roll parameters in terms of derivatives of the scalar potential, as in the last two expressions, we realize that the slow-roll conditions require very flat potentials to be met.

Since I am considering a slow-roll regime for the model given by the potential in Eq. (15), the evolution of the fields in such a case is given by

$$\phi(N) = \phi_* \exp(-N\eta_\phi), \quad (25)$$

and

$$\sigma(N) = \sigma_* \exp(-N\eta_\sigma), \quad (26)$$

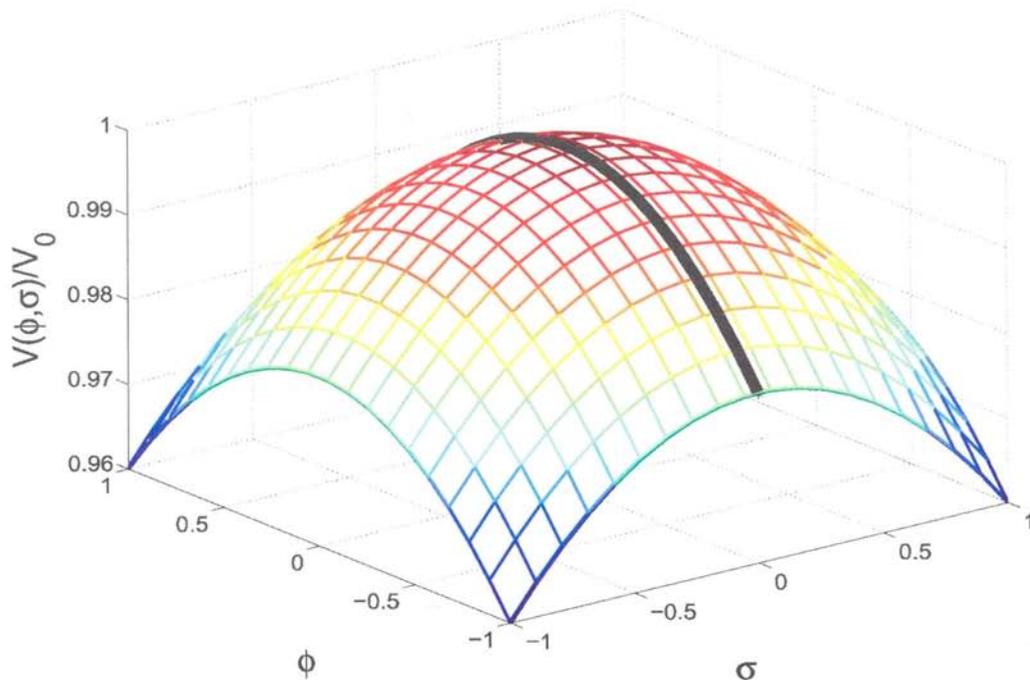


Figure 1: My small-field slow-roll potential of Eq. (15) with  $\eta_\phi, \eta_\sigma < 0$ . The inflaton starts near the maximum and moves away from the origin following the  $\sigma = 0$  trajectory depicted with the solid black line. (This figure has been taken from Ref. (Alabidi, 2006)).

in terms of the amount of inflation  $N$  since horizon exit, and the field values  $\phi_*$  and  $\sigma_*$  at the time when the relevant cosmological scales exit the horizon. Such expressions, together with Eq. (15), seed the  $\delta N$  formalism in order to calculate the spectrum and the bispectrum of the curvature perturbation including the tree-level (see Fig. 2) and the one-loop contributions (see Fig. 3)<sup>10</sup> (see the respective calculational details in Refs. Byrnes, Koyama, Sasaki, & Wands, 2007; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Lyth &

Rodríguez, 2005a; Sasaki & Stewart, 1996):

$$\mathcal{P}_\zeta^{tree} = \frac{1}{\eta_\phi^2 \phi_*^2} \left( \frac{H_*}{2\pi} \right)^2, \quad (27)$$

$$\begin{aligned} \mathcal{P}_\zeta^{1-loop} &\simeq \frac{\eta_\sigma^2}{\eta_\phi^4 \phi_*^4} \exp[4N(|\eta_\sigma| - |\eta_\phi|)] \times \\ &\times \left( \frac{H_*}{2\pi} \right)^4 \ln(kL), \end{aligned} \quad (28)$$

$$B_\zeta^{tree} = -\frac{1}{\eta_\phi^3 \phi_*^4} \left( \frac{H_*}{2\pi} \right)^4 4\pi^4 \left( \frac{\sum_i k_i^3}{\prod_i k_i^3} \right), \quad (29)$$

$$\begin{aligned} B_\zeta^{1-loop} &\simeq \frac{\eta_\sigma^3}{\eta_\phi^6 \phi_*^6} \exp[6N(|\eta_\sigma| - |\eta_\phi|)] \times \\ &\times \left( \frac{H_*}{2\pi} \right)^6 \ln(kL) 4\pi^4 \left( \frac{\sum_i k_i^3}{\prod_i k_i^3} \right), \end{aligned} \quad (30)$$

where  $L$  is the infrared cutoff chosen so that the quantities are calculated in a minimal box (Lyth, 2007), i.e.  $\ln(kL) \sim$

<sup>10</sup>For an adequate explanation of the Feynman-like diagrams in cosmology, and their application within the  $\delta N$  formalism, see Ref. (Byrnes, Koyama, Sasaki, & Wands, 2007).

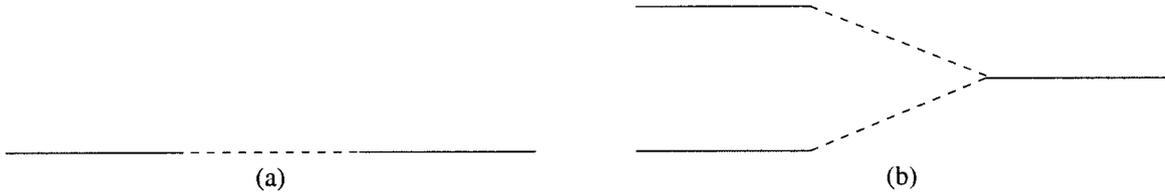


Figure 2: Tree-level Feynman-like diagrams for (a)  $P_\zeta$ , and (b)  $B_\zeta$ . The internal dashed lines correspond to two-point correlators of field perturbations.



Figure 3: One-loop Feynman-like diagrams for (a)  $P_\zeta$ , and (b)  $B_\zeta$ . The internal dashed lines correspond to two-point correlators of field perturbations.

$\mathcal{O}(1)$ , and  $k_1 \sim k_2 \sim k_3$ .

The important factor in the loop corrections is the exponential. This exponential function is directly related to the quadratic form of the potential with a leading constant term. It will give a large contribution if  $|\eta_\sigma| > |\eta_\phi|$ . I have chosen the concave downward potential in order to satisfy the spectral tilt constraint, which makes either  $\eta_\phi < 0$ , if  $\mathcal{P}_\zeta \simeq \mathcal{P}_\zeta^{tree}$ , or  $\eta_\sigma < 0$ , if  $\mathcal{P}_\zeta \simeq \mathcal{P}_\zeta^{1-loop}$ , while keeping  $|\eta_\sigma| > |\eta_\phi|$ .

## 4 Constraints to have a reliable parameter space

### 4.1 Convergence of the $\zeta$ series and existence of a perturbative regime

It has been proved (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b) by means of a non-perturbative approach, that there exist what are called two “coupling constants”  $x$  and  $y$  for the potential in Eq. (15). Such coupling constants allow us to obtain a necessary condition for the convergence of the  $\zeta$  series and work in a perturbative regime if they are much less

than one:

$$\begin{aligned}
 |x| &\equiv \left| \frac{\delta\phi_\star}{\phi_\star} \right| \approx \left( \frac{H_\star}{2\pi} \right) \frac{1}{\phi_\star} \ll 1, & (31) \\
 |y| &\equiv \left\{ \frac{\eta_\sigma^3 \delta\sigma_\star^2}{\eta_\phi^3 \phi_\star^2} \exp[2N(\eta_\phi - \eta_\sigma)] \right\}^{1/2} \\
 &\approx \left\{ \frac{\eta_\sigma^3}{\eta_\phi^3} \left( \frac{H_\star}{2\pi} \right)^2 \frac{1}{\phi_\star^2} \exp[2N(\eta_\phi - \eta_\sigma)] \right\}^{1/2} \ll 1. & (32)
 \end{aligned}$$

### 4.2 Tree-level or loop dominance

Because of the exponential factors in Eqs. (28) and (30) it might be possible that the loop corrections dominate over  $\mathcal{P}_\zeta$  and/or  $B_\zeta$ . There are three possibilities in complete connection with the position of the  $\phi$  field when the relevant scales are exiting the horizon. Here I will consider only the intermediate  $\phi_\star$  region, corresponding to the case when  $B_\zeta$  is dominated by one-loop corrections and  $\mathcal{P}_\zeta$  is dominated by the tree-level term, because this is the only possibility which gives interesting and observationally relevant results.

### $B_\zeta$ dominated by one-loop corrections and $\mathcal{P}_\zeta$ dominated by the tree-level term: the intermediate $\phi_*$ region

Looking at Eqs. (27) and (30) I require in this case that

$$\frac{\eta_\sigma^2}{\eta_\phi^2} \exp[4N(|\eta_\sigma| - |\eta_\phi|)] \ll \frac{1}{\frac{1}{\phi_*^2} \left(\frac{H_*}{2\pi}\right)^2}, \quad (33)$$

$$\frac{\eta_\sigma^3}{\eta_\phi^3} \exp[6N(|\eta_\sigma| - |\eta_\phi|)] \gg \frac{1}{\frac{1}{\phi_*^2} \left(\frac{H_*}{2\pi}\right)^2}, \quad (34)$$

which combine to give

$$\begin{aligned} \frac{r\mathcal{P}_\zeta \eta_\sigma^2}{8 \eta_\phi^2} \exp[4N(|\eta_\sigma| - |\eta_\phi|)] &\ll \left(\frac{\phi_*}{m_P}\right)^2 \\ &\ll \frac{r\mathcal{P}_\zeta \eta_\sigma^3}{8 \eta_\phi^3} \exp[6N(|\eta_\sigma| - |\eta_\phi|)], \end{aligned} \quad (35)$$

where the definition for the tensor to scalar ratio  $r$  (Lyth, 2008) has been employed:

$$r \equiv \frac{\mathcal{P}_T}{\mathcal{P}_\zeta} = \frac{\frac{8}{m_P^2} \left(\frac{H_*}{2\pi}\right)^2}{\mathcal{P}_\zeta}. \quad (36)$$

In the latter expression,  $\mathcal{P}_T^{1/2}$  represents the amplitude of the spectrum for primordial gravitational waves.

### 4.3 Spectrum normalisation condition

Since I am considering  $\zeta$  being generated during inflation, I must satisfy the appropriate spectrum normalisation condition. According to Eq. (27) if  $\mathcal{P}_\zeta$  is dominated by the tree-level term, I have

$$\mathcal{P}_\zeta^{tree} = \frac{1}{\eta_\phi^2 \phi_*^2} \left(\frac{H_*}{2\pi}\right)^2 = \frac{1}{\eta_\phi^2} \left(\frac{m_P}{\phi_*}\right)^2 \frac{r\mathcal{P}_\zeta}{8} = \mathcal{P}_\zeta, \quad (37)$$

which reduces to

$$\left(\frac{\phi_*}{m_P}\right)^2 = \frac{1}{\eta_\phi^2} \frac{r}{8}. \quad (38)$$

Notice that in such a situation, the value of the  $\phi$  field when the relevant scales are exiting the horizon depends exclusively on the tensor to scalar ratio, once  $\eta_\phi$  has been fixed by the spectral tilt constraint.

### 4.4 Spectral tilt constraint

The current observed value for the spectral tilt is  $n_\zeta - 1 = -0.040 \pm 0.014$  (Komatsu et. al., 2008), and again I will

consider only the case when  $\mathcal{P}_\zeta$  is dominated by the tree-level term. That means that the usual spectral index formula (Sasaki & Stewart, 1996) applies:

$$n_\zeta - 1 = -2\epsilon - 2m_P^2 \frac{\sum_{ij} V_i N_j N_{ij}}{V \sum_i N_i^2}, \quad (39)$$

giving the following result once the derivatives of  $N$  with respect to  $\phi_*$  and  $\sigma_*$  have been calculated:

$$n_\zeta - 1 = -2\epsilon + 2\eta_\phi. \quad (40)$$

The effect of the  $\epsilon$  parameter may be discarded in the previous expression since  $\epsilon$  is much less than  $|\eta_\phi|$ :

$$\epsilon = \frac{m_P^2}{2} \frac{V_\phi^2 + V_\sigma^2}{V^2} = |\eta_\phi| \left[ \frac{1}{2} |\eta_\phi| \left(\frac{\phi}{m_P}\right)^2 \right] \ll |\eta_\phi|, \quad (41)$$

according to the prescription that the potential in Eq. (15) is dominated by the constant term. Thus, using the central value for  $n_\zeta - 1$ , I get

$$\eta_\phi = -0.020. \quad (42)$$

### 4.5 Amount of inflation

It is well known that the number of e-folds of expansion from the time the cosmological scales exit the horizon to the end of inflation is presumably around but less than 62 (Dodelson, 2003; Liddle & Lyth, 2000; Weinberg, 2008). The slow-roll evolution of the  $\phi$  field in Eq. (25) tells us that such an amount of inflation is given by

$$N = \frac{1}{|\eta_\phi|} \ln \left( \frac{\phi_{end}}{\phi_*} \right) \lesssim 62, \quad (43)$$

where  $\phi_{end}$  is the value of the  $\phi$  field at the end of inflation. Because of the characteristics of the inflationary potential in Eq. (15), there is no a definite mechanism to end inflation in this model. It could not be by means of the violation of the  $\epsilon < 1$  condition since this would imply extrapolating our results to a region where the potential in Eq. (15) is no longer dominated by the constant term which, in addition, would spoil the large non-gaussianity generated and could send the model to an unknowable quantum gravity regime. Keeping in mind the results of Ref. (Armendariz-Picon, Fontanini, Penco, & Trodden, 2008) which say that the ultraviolet cut-off in cosmological perturbation theory could be a few orders of magnitude bigger than  $m_P$ , I will therefore assume that inflation comes to an end when  $|\eta_\phi| \phi^2 / 2m_P^2 \sim 10^{-2}$ . This allows me to be in a safe side (avoiding large modifications to the potential coming from ultraviolet cutoff-suppressed non-renormalisable terms, and keeping the potential dominated by the constant  $V_0$  term), leaving the implementation of a

mechanism to end inflation for a future work<sup>11</sup>. Coming back to Eq. (43), I get then

$$N = \frac{1}{|\eta_\phi|} \ln \left( \frac{m_P}{\phi_*} \right) \lesssim 62, \quad (44)$$

which leads to

$$\frac{\phi_*}{m_P} \gtrsim \exp(-62|\eta_\phi|). \quad (45)$$

## 5 $f_{NL}$

In this section I will calculate the level of non-gaussianity represented in the parameter  $f_{NL}$  (Komatsu & Spergel, 2001) by taking into account the constraints presented in Section 4 (Cogollo, Rodríguez & Valenzuela-Toledo, 2008a). The level of non-gaussianity, according to the expressions in Eqs. (12), (27), and (30), is in this case given by

$$\begin{aligned} \frac{6}{5} f_{NL} &= \frac{B_\zeta^{1-loop}}{4\pi^4 \frac{\sum_i k_i^3}{\prod_i k_i^4} (\mathcal{P}_\zeta^{tree})^2} \\ &\simeq \frac{\eta_\sigma^3}{\eta_\phi^2 \phi_*^2} \exp[6N(|\eta_\sigma| - |\eta_\phi|)] \left( \frac{H_*}{2\pi} \right)^2 \ln(kL) \\ &= \frac{\eta_\sigma^3}{\eta_\phi^2} \exp[6N(|\eta_\sigma| - |\eta_\phi|)] \left( \frac{m_P}{\phi_*} \right)^2 \frac{r \mathcal{P}_\zeta}{8} \ln(kL) \\ &= \eta_\sigma^3 \exp[6N(|\eta_\sigma| - |\eta_\phi|)] \mathcal{P}_\zeta \ln(kL), \quad (46) \\ \Rightarrow \frac{6}{5} f_{NL} &\approx -2.457 \times 10^{-9} |\eta_\sigma|^3 \exp \left[ 300 \times \right. \\ &\quad \left. \times \ln(5.657 \times 10^{-2} r^{-1/2}) (|\eta_\sigma| - 0.020) \right], \quad (47) \end{aligned}$$

where in the last line I have used expressions in Eqs. (38), (42), and (44).

In figure 4 I show lines of constant  $f_{NL}$  in the plot  $r$  vs  $|\eta_\sigma|$  for the intermediate  $\phi_*$  region in agreement with the constraints in Eqs. (31), (32), and (35). Notice that by implementing the spectral tilt constraint in Eq. (42) to the spectrum normalisation constraint in Eq. (38) and the amount of inflation constraint in Eq. (45) I conclude that the tensor to scalar ratio is bounded from below:  $r \gtrsim 2.680 \times 10^{-4}$ .

## 6 Conclusions

As is evident from the plot, the WMAP (and also PLANCK) observationally allowed  $2\sigma$  range of values for negative  $f_{NL}$ ,

$-9 < f_{NL}$ , is completely inside the intermediate  $\phi_*$  region as required. More negative values for  $f_{NL}$ , up to  $f_{NL} = -20.647$  are consistent within my framework for the intermediate  $\phi_*$  region, but they are ruled out from observation. Nevertheless, it is interesting to see a slow-roll inflationary model with canonical kinetic terms where the observational restriction on  $f_{NL}$  may be violated by an excess and not by a shortfall. So I conclude that if  $B_\zeta$  is dominated by the one-loop correction but  $P_\zeta$  is dominated by the tree-level term, sizeable non-gaussianity is generated even if  $\zeta$  is generated during inflation. I also conclude, from looking at the small values that the tensor to scalar ratio  $r$  takes in figure 4 compared with the present technological bound  $r \gtrsim 10^{-3}$  (Friedman, Cooray, & Melchiorri, 2006), that for non-gaussianity to be observable in this model, primordial gravitational waves must be undetectable.

**Acknowledgments:** This work is supported by COL-CIENCIAS grant No. 1102-333-18674 CT-174-2006, DIF (UIS) grant No. 5134, and by the ECOS-NORD programme grant No. C06P02.

## References

- Ackerman L., Carroll S.M., & Wise M.B., 2007. Imprints of a primordial preferred direction on the microwave background. *Phys. Rev. D* **75**, 083502.
- Ahmad I., Piao Y.-S., & Quiao C.-F., 2008. The spectrum of curvature perturbation for multi-field inflation with a small-field potential. *JCAP* **0802**, 002.
- Alabidi L., 2006. Non-gaussianity for a two component hybrid model of inflation. *JCAP* **0610**, 015.
- Alabidi L. & Lidsey J.E., 2008. Single field inflation after the WMAP five-year data. *Phys. Rev. D* **78**, 103519.
- Alabidi L. & Lyth D.H., 2006. Inflation models after WMAP year three. *JCAP* **0608**, 013.
- Albrecht A. & Steinhardt P.J., 1982. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Phys. Rev. Lett.* **48**, 1220.
- Armendariz-Picon C., Fontanini M., Penco R., & Trodden M., 2008. Where does cosmological perturbation break down?. [arXiv:0805.0114](https://arxiv.org/abs/0805.0114) [hep-th].
- Bardeen J.M., 1980. Gauge invariant cosmological perturbations. *Phys. Rev. D* **22**, 1882.
- Battfeld T. & Easther R., 2007. Non-gaussianities in multi-field inflation. *JCAP* **0703**, 020.
- Boubekeur L. & Lyth D.H., 2005. Hilltop inflation. *JCAP* **0507**, 010.

<sup>11</sup>I hope that the implementation of such a mechanism in my model will keep, or perhaps enhance, the generated non-gaussianity. Nevertheless the opposite behaviour might as well happen. For instance, Ref. (Rigopoulos, Shellard, & van Tent, 2007) studies within a stochastic formalism a quadratic two-component slow-roll model without a dominant constant term in the potential. A momentary violation of the slow-roll conditions around the end of inflation shows to enhance  $f_{NL}$  to observable levels; however, such an enhancement vanishes once inflation ends completely. These results have been confirmed numerically within the  $\delta N$  formalism in Refs. (Vernizzi & Wands, 2006; Yokoyama, Suyama, & Tanaka, 2008a).

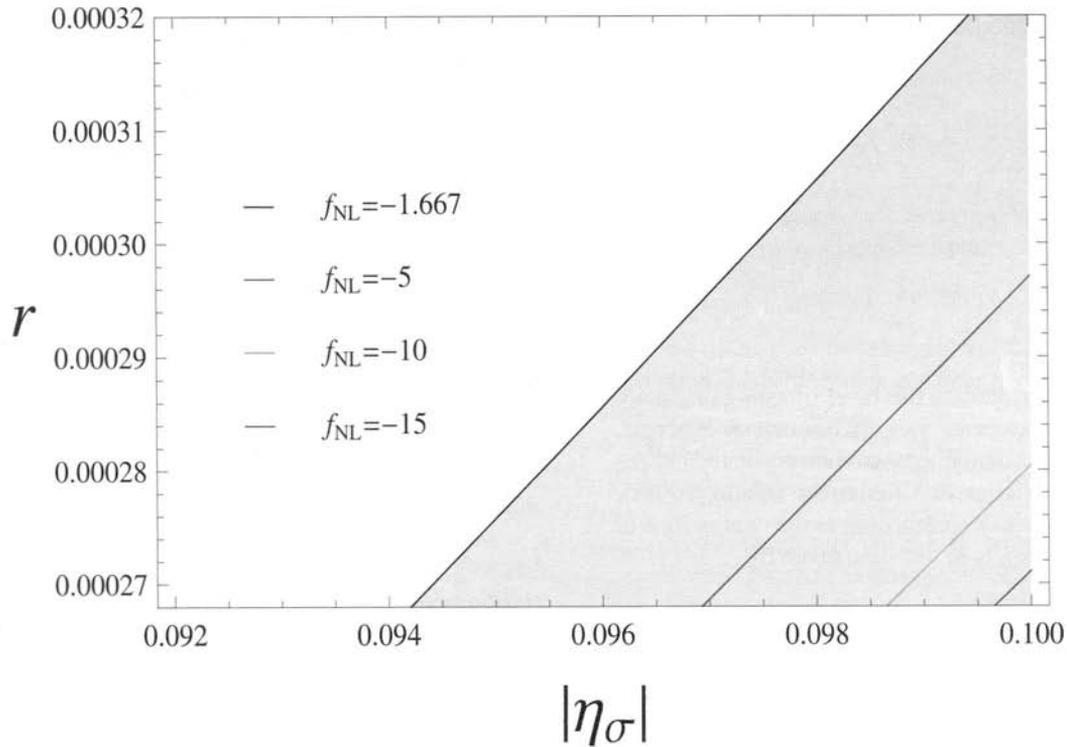


Figure 4: Contours of  $f_{NL}$  in the  $r$  vs  $|\eta\sigma|$  plot. The intermediate (high)  $\phi_*$  region corresponds to the shaded (white) region. The WMAP (and also PLANCK) observationally allowed  $2\sigma$  range of values for negative  $f_{NL}$ ,  $-9 < f_{NL}$ , is completely inside the intermediate  $\phi_*$  region. Notice that the boundary line between the high and the intermediate  $\phi_*$  regions matches almost exactly the  $f_{NL} = -1.667$  line.

**Boubekeur L. & Lyth D.H.**, 2006. Detecting a small perturbation through its non-gaussianity. *Phys. Rev. D* **73**, 021301(R).

**Bunn E.F. & White M.J.**, 1997. The four-year COBE normalization and large-scale structure. *Astrophys. J.* **480**, 6.

**Byrnes C.T., Choi K.-Y., & Hall L.M.H.**, 2008. Conditions for large non-gaussianity in two-field slow-roll inflation. *JCAP* **0810**, 008.

**Byrnes C.T., Koyama K., Sasaki M., & Wands D.**, 2007. Diagrammatic approach to non-gaussianity from inflation. *JCAP* **0711**, 027.

**Byrnes C.T., Sasaki M., & Wands D.**, 2006. The primordial trispectrum from inflation. *Phys. Rev. D* **74**, 123519.

**Carroll S.M., Tseng C.-Y., & Wise M.B.**, 2008. Translational invariance and the anisotropy of the cosmic microwave background. [arXiv:0811.1086](https://arxiv.org/abs/0811.1086) [astro-ph].

**Cogollo H.R.S., Rodríguez Y., & Valenzuela-Toledo C.A.**, 2008a. On the issue of the  $\zeta$  series convergence and loop corrections in the generation of observable primordial non-Gaussianity in slow-roll inflation. Part I: the bispectrum. *JCAP* **0808**, 029.

**Cogollo H.R.S., Rodríguez Y., & Valenzuela-Toledo C.A.**, 2008b. On the issue of the  $\zeta$  series convergence and loop corrections in

the generation of observable primordial non-Gaussianity in slow-roll inflation. Part II: the trispectrum. [arXiv:0811.4092](https://arxiv.org/abs/0811.4092) [astro-ph].

**Cogollo H.R.S., Rodríguez Y., & Valenzuela-Toledo C.A.**, 2008c. Non-gaussianity and loop corrections in a quadratic two-field slow-roll model of inflation. Part II. Submitted to *Rev. Acad. Colomb. Cienc.*

**Cooray A.**, 2006. 21-cm background anisotropies can discern primordial non-gaussianity. *Phys. Rev. Lett.* **97**, 261301.

**Cooray A., Li C., & Melchiorri A.**, 2008. The trispectrum of 21-cm background anisotropies as a probe of primordial non-gaussianity. *Phys. Rev. D* **77**, 103506.

**Dimopoulos K. & Lazarides G.**, 2006. Modular inflation and the orthogonal axion as curvaton. *Phys. Rev. D* **73**, 023525.

**Dimopoulos K., Lyth D.H., & Rodríguez Y.**, 2008. Statistical anisotropy of the curvature perturbation from vector field perturbations. [arXiv:0809.1055](https://arxiv.org/abs/0809.1055) [astro-ph].

**Dodelson S.**, 2003. *Modern cosmology*, Academic Press, San Diego USA.

**Dodelson S., Kinney W.H., & Kolb E.W.**, 1997. Cosmic microwave background measurements can discriminate among infla-

- tion models. *Phys. Rev. D* **56**, 3207.
- Enqvist K. & Vähkönen A.**, 2004. Non-gaussian perturbations in hybrid inflation. *JCAP* **0409**, 006.
- Freese K., Frieman J., & Olinto A.**, 1990. Natural inflation with pseudo-Nambu-Goldstone bosons. *Phys. Rev. Lett.* **65**, 3233.
- Friedman B.C., Cooray A., & Melchiorri A.**, 2006. WMAP-normalized inflationary model predictions and the search for primordial gravitational waves with direct detection experiments. *Phys. Rev. D* **74**, 123509.
- Kogo N. & Komatsu E.**, 2006. Angular trispectrum of CMB temperature anisotropy from primordial non-gaussianity with the full radiation transfer function. *Phys. Rev. D* **73**, 083007.
- Komatsu E.**, 2008. Private communication.
- Komatsu E. & Spergel D.N.**, 2001. Acoustic signatures in the primary microwave background bispectrum. *Phys. Rev. D* **63**, 063002.
- Komatsu E. et al.**, 2008. Five-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological interpretation. [arXiv:0803.0547](https://arxiv.org/abs/0803.0547) [astro-ph].
- Liddle A.R. & Lyth D.H.**, 2000. *Cosmological inflation and large-scale structure*, Cambridge University Press, Cambridge UK.
- Linde A.D.**, 1982. A new inflationary universe scenario: a possible solution to the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Phys. Lett. B* **108**, 389.
- Linde A.D.**, 1994. Hybrid inflation. *Phys. Rev. D* **49**, 748.
- Lyth D.H.**, 2007. The curvature perturbation in a box. *JCAP* **0712**, 016.
- Lyth D.H.**, 2008. Particle physics models of inflation. *Lec. Notes Phys.* **738**, 81.
- Lyth D.H., Malik K.A., & Sasaki M.**, 2005. A general proof of the conservation of the curvature perturbation. *JCAP* **0505**, 004.
- Lyth D.H. & Riotto A.**, 1999. Particle physics models of inflation and the cosmological density perturbation. *Phys. Rep.* **314**, 1.
- Lyth D.H. & Rodríguez Y.**, 2005a. Inflationary prediction for primordial non-gaussianity. *Phys. Rev. Lett.* **95**, 121302.
- Lyth D.H. & Rodríguez Y.**, 2005b. Non-gaussianity from the second-order cosmological perturbation. *Phys. Rev. D* **71**, 123508.
- Maldacena J.**, 2003. Non-gaussian features of primordial fluctuations in single field inflationary models. *JHEP* **0305**, 013.
- Mukhanov V.F.**, 2005. *Physical foundations of cosmology*, Cambridge University Press, Cambridge UK.
- Okamoto T. & Hu W.**, 2002. Angular trispectra of CMB temperature and polarization. *Phys. Rev. D* **66**, 063008.
- The PLANCK Collaboration**, 2006. The scientific programme of Planck. [arXiv:astro-ph/0604069](https://arxiv.org/abs/astro-ph/0604069).
- Rigopoulos G., Shellard E.P.S., & van Tent B.J.W.**, 2007. Quantitative bispectra from multifield inflation. *Phys. Rev. D* **76**, 083512.
- Sasaki M. & Stewart E.D.**, 1996. A general analytic formula for the spectral index of the density perturbations produced during inflation. *Prog. Theor. Phys.* **95**, 71.
- Seery D. & Lidsey J.E.**, 2007. Non-gaussianity from the inflationary trispectrum. *JCAP* **0701**, 008.
- Seery D., Sloth M., & Vernizzi F.**, 2008. Inflationary trispectrum from graviton exchange. [arXiv:0811.3934](https://arxiv.org/abs/0811.3934) [astro-ph].
- Starobinsky A.A.**, 1985. Multicomponent de Sitter (inflationary) stages and the generation of perturbations. *Pisma Zh. Eksp. Teor. Fiz.* **42**, 124. [*JETP Lett.* **42**, 152].
- Vähkönen A.**, 2005. Comment on non-gaussianity in hybrid inflation. [arXiv:astro-ph/0506304](https://arxiv.org/abs/astro-ph/0506304).
- Vernizzi F. & Wands D.**, 2006. Non-gaussianities in two-field inflation. *JCAP* **0605**, 019.
- Weinberg S.**, 2008. *Cosmology*, Oxford University Press, Oxford UK.
- Yadav A.P.S. & Wandelt B.D.**, 2008. Evidence of primordial non-gaussianity ( $f_{NL}$ ) in the Wilkinson Microwave Anisotropy Probe 3-year data at  $2.8\sigma$ . *Phys. Rev. Lett.* **100**, 181301.
- Yokoyama S., Suyama T., & Tanaka T.**, 2007. Primordial non-gaussianity in multi-scalar slow-roll inflation. *JCAP* **0707**, 013.
- Yokoyama S., Suyama T., & Tanaka T.**, 2008a. Primordial non-gaussianity in multi-scalar inflation. *Phys. Rev. D* **77**, 083511.
- Yokoyama S., Suyama T., & Tanaka T.**, 2008b. Efficient diagrammatic computation method for higher order correlation functions of local type primordial curvature perturbations. [arXiv:0810.3053](https://arxiv.org/abs/0810.3053) [astro-ph].
- Zaballa I., Rodríguez Y., & Lyth D.H.**, 2006. Higher order contributions to the primordial non-gaussianity. *JCAP* **0606**, 013.

Recibido: febrero 3 de 2008

Aceptado para publicación: noviembre 13 de 2008

# NON-GAUSSIANITY AND LOOP CORRECTIONS IN A QUADRATIC TWO-FIELD SLOW-ROLL MODEL OF INFLATION. PART II

By

Heiner R. S. Cogollo<sup>1</sup>, Yeinzon Rodríguez<sup>2,3</sup> & César A. Valenzuela-Toledo<sup>4</sup>

## Abstract

**Cogollo H.R.S., Y. Rodríguez, & C.A. Valenzuela-Toledo:** Non-Gaussianity and Loop Corrections in a Quadratic Two-Field Slow-Roll Model of Inflation. Part II. Rev. Acad. Colomb. Cienc. **32** (125): 515-526, 2008. ISSN 0370-3908.

We calculate the trispectrum  $T_\zeta(\vec{k}_1, \vec{k}_2, \vec{k}_3, \vec{k}_4)$  of the primordial curvature perturbation  $\zeta$ , generated during a slow-roll inflationary epoch and considering a quadratic two-component scalar potential and canonical kinetic terms. We consider one-loop and tree level contributions, and show that it is possible to attain *observable* values for the level of non-gaussianity  $\tau_{NL}$  if  $T_\zeta$  is dominated by the one-loop contribution. This work is performed by taking into account that there exists some physical restrictions that constrain the available parameter window. Such conditions are: the existence of some coupling constants that guarantee the calculation in a perturbative regime, the relative weight of the one-loop and tree level contributions, the normalisation of the spectrum, the observed spectral index, and the minimal amount of inflation required to solve the horizon problem.

**Key words:** Primordial curvature perturbation, non-gaussianity, slow-roll inflationary models.

- 
- 1 Escuela de Física, Universidad Industrial de Santander, Ciudad Universitaria, Bucaramanga, Colombia, Correo electrónico: heiner.sarmiento@ciencias.uis.edu.co
  - 2 Centro de Investigaciones, Universidad Antonio Nariño, Cra 3 Este No. 47A-15, Bogotá D.C., Colombia, Correo electrónico: yeinzon.rodriguez@uan.edu.co
  - 3 Escuela de Física, Universidad Industrial de Santander, Ciudad Universitaria, Bucaramanga, Colombia, Correo electrónico: yrodrig@uis.edu.co
  - 4 Escuela de Física, Universidad Industrial de Santander, Ciudad Universitaria, Bucaramanga, Colombia, Correo electrónico: cavalto@ciencias.uis.edu.co

## Resumen

Se calcula el trispectro  $T_\zeta(\vec{k}_1, \vec{k}_2, \vec{k}_3, \vec{k}_4)$  de la perturbación en la curvatura  $\zeta$ , generado durante una época inflacionaria de slow-roll y considerando un potencial escalar cuadrático de dos componentes y términos cinéticos canónicos. En el cálculo se consideran contribuciones a nivel árbol y a un lazo, y se muestra que es posible obtener valores observables para el nivel de no gaussianidad  $\tau_{NL}$  si  $T_\zeta$  es dominado por la contribución a un lazo. El trabajo se desarrolla teniendo en cuenta que existen algunas restricciones físicas que reducen la ventana de parámetros disponible. Estas condiciones son: la existencia de unas constantes de acoplamiento que garantizan la realización del cálculo en un régimen perturbativo, el peso relativo de las contribuciones a nivel árbol y a un lazo, la normalización del espectro, el índice espectral observado, y el monto de inflación mínimo necesario para resolver el problema de horizonte.

**Palabras clave:** Perturbación primordial en la curvatura, no gaussianidad, modelos inflacionarios del tipo slow-roll.

## 1 Introduction

This is the sequel to the paper in Ref. (Rodríguez, 2008). In such a paper one of us discussed the generation of large *and observable* non-gaussianity in the bispectrum  $B_\zeta$  of the primordial curvature perturbation  $\zeta$ , for a particular quadratic two-field slow-roll inflationary model with canonical kinetic terms. The way of doing that is by means of the  $\delta N$  formalism (Lyth, Malik, & Sasaki, 2005; Lyth & Rodríguez, 2005a; Sasaki & Stewart, 1996; Starobinsky, 1985), where  $\zeta$  is expanded in a Taylor-series, much in the way as is done in quantum field theory for the scattering amplitudes. The statistical correlators of the  $\zeta$  series then may be understood as composed of tree-level terms and loop contributions<sup>5</sup>. The key to find large non-gaussianity is by considering the case when the loop corrections dominate over the tree-level terms. For slow-roll inflationary models with canonical kinetic terms (Liddle & Lyth, 2000; Lyth, 2008; Lyth & Riotto, 1999), the level of non-gaussianity  $\tau_{NL}$  associated to  $T_\zeta$  is usually thought to be of order  $\mathcal{O}(\epsilon_i, \eta_i)$  (Seery & Lidsey, 2007; Seery, Sloth, & Vernizzi, 2008), where  $\epsilon_i$  and  $\eta_i$  are the slow-roll parameters with  $\epsilon_i, |\eta_i| \ll 1$  (Lyth & Riotto, 1999). Such values for  $\tau_{NL}$  are far from being detectable (Cooray, Li, & Melchiorri, 2008; Kogo & Komatsu, 2006). However, in order to reach such a conclusion, only the tree-level terms were considered without giving a satisfactory explanation of why the loop contributions are comparatively suppressed. In the present paper we will discuss the generation of large *and observable* non-gaussianity in the trispectrum  $T_\zeta$  of  $\zeta$  for the same inflationary model as in Ref. (Rodríguez, 2008), considering the case when the

loop corrections dominate over the tree-level terms.

The structure of the present paper follows closely that for Ref. (Rodríguez, 2008). Sections 2 and 3 are faithfully reproduced from Ref. (Rodríguez, 2008), just to make the present paper more self-contained. The same is true for Section 4, except for Subsection 4.2. The latter Subsection together with Sections 5 and 6 contain the new material regarding  $\tau_{NL}$ .

## 2 Basic definitions and observation

Given the probability distribution function  $f(\zeta)$ , for the primordial curvature perturbation  $\zeta(\mathbf{x}, t)$ , there are an infinite number of standardized moments that work as statistical descriptors of  $\zeta(\mathbf{x}, t)$ :

$$\text{the mean value : } m_\zeta(1) \equiv \langle \zeta \rangle = \int \zeta f(\zeta) d\zeta, \quad (1)$$

$$\text{the variance : } m_\zeta(2) \equiv \int (\zeta - \langle \zeta \rangle)^2 f(\zeta) d\zeta, \quad (2)$$

$$\text{the skewness : } m_\zeta(3) \equiv \frac{\int (\zeta - \langle \zeta \rangle)^3 f(\zeta) d\zeta}{[m_\zeta(2)]^{3/2}}, \quad (3)$$

$$\text{the kurtosis : } m_\zeta(4) \equiv \frac{\int (\zeta - \langle \zeta \rangle)^4 f(\zeta) d\zeta}{[m_\zeta(2)]^2}, \quad (4)$$

and so on.

Departures from the exact gaussianity come either from

<sup>5</sup>In this paper, we follow the terminology of Ref. (Byrnes, Koyama, Sasaki, & Wands, 2007) to identify the tree-level terms and the loop contributions, within the  $\delta N$  formalism, in a diagrammatic approach. The associated diagrams are called *Feynman-like diagrams*.

non-vanishing odd standardized moments  $m_\zeta(n)$  with  $n \geq 3$ , in which case the probability distribution function is non-symmetric around the mean value, or from higher ( $n \geq 4$ ) even standardized moments different to products of the variance, in which case the probability distribution function continues to be symmetric around the mean value although its “peakedness”<sup>6</sup> is bigger than that for a gaussian function, or from both of them.

Working in momentum space, the standardized moments of the probability distribution function have a direct connection with the correlation functions for the Fourier modes  $\zeta_{\mathbf{k}} \equiv \int d^3k \zeta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}$  defined in flat space. As the  $n$ -point correlators of  $\zeta_{\mathbf{k}}$  are generically defined in terms of spectral functions of the wavevectors involved<sup>7</sup>:

two – point correlator  $\rightarrow$  spectrum  $P_\zeta$  :

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \rangle \equiv (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2) P_\zeta(k), \quad (5)$$

three – point correlator  $\rightarrow$  bispectrum  $B_\zeta$  :

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle \equiv (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times B_\zeta(k_1, k_2, k_3), \quad (6)$$

four – point correlator  $\rightarrow$  trispectrum  $T_\zeta$  :

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \zeta_{\mathbf{k}_4} \rangle \equiv (2\pi)^3 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4) \times T_\zeta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4), \quad (7)$$

and so on,

the standardized moments of the distribution are then written in terms of momentum integrals of the spectral functions for

the modes  $\zeta_{\mathbf{k}}$ :

$$\text{the variance : } m_\zeta(2) = \int \frac{d^3k}{(2\pi)^3} P_\zeta(k), \quad (8)$$

$$\text{the skewness : } m_\zeta(3) = \frac{\int \frac{d^3k_1 d^3k_2}{(2\pi)^6} B_\zeta(k_1, k_2, k_3)}{\left[ \int \frac{d^3k}{(2\pi)^3} P_\zeta(k) \right]^{3/2}}, \quad (9)$$

$$\text{the kurtosis : } m_\zeta(4) = \frac{\int \frac{d^3k_1 d^3k_2 d^3k_3}{(2\pi)^9} T_\zeta(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)}{\left[ \int \frac{d^3k}{(2\pi)^3} P_\zeta(k) \right]^2}, \quad (10)$$

and so on.

Non-gaussianity in  $\zeta$  is, therefore, associated with non-vanishing higher order spectral functions, starting from the bispectrum  $B_\zeta$ .

Now we will parametrize the spectral functions of  $\zeta$  in terms of quantities which are the ones for which observational bounds are given. Because of the direct connection between these quantities and the standardized moments of the probability distribution function  $f(\zeta)$ , we may also call these quantities as the statistical descriptors for  $f(\zeta)$ . The spectrum  $P_\zeta$  is parametrized in terms of an amplitude  $\mathcal{P}_\zeta^{1/2}$  and a spectral index  $n_\zeta$  which measures the deviation from an exactly scale-invariant spectrum (Liddle & Lyth, 2000; Weinberg, 2008):

$$P_\zeta(k) \equiv \frac{2\pi^2}{k^3} \mathcal{P}_\zeta \left( \frac{k}{aH} \right)^{n_\zeta - 1}, \quad (11)$$

where  $a$  is the global expansion parameter and  $H = \dot{a}/a$  is the Hubble parameter, with the dot meaning a derivative with respect to cosmic time. The bispectrum  $B_\zeta$  and trispectrum  $T_\zeta$  are parametrized in terms of products of the spectrum  $P_\zeta$ , and the quantities  $f_{NL}$  and  $\tau_{NL}$  respectively<sup>8</sup> (Boubekeur

<sup>6</sup>Higher even standardized moments different to products of the variance mean more of the variance is due to infrequent extreme deviations, as opposed to frequent modestly-sized deviations.

<sup>7</sup>Statistical homogeneity, i.e., invariance of the correlation functions under translations, requires the presence of the Dirac delta functions (Ackerman, Carroll, & Wise, 2007; Carroll, Tseng, & Wise, 2008; Dimopoulos, Lyth, & Rodríguez, 2008). Statistical isotropy, i.e., invariance of the correlation functions under rotations, requires that the spectrum  $P_\zeta$  and bispectrum  $B_\zeta$  are functions of the wavenumbers only (Ackerman, Carroll, & Wise, 2007; Carroll, Tseng, & Wise, 2008; Dimopoulos, Lyth, & Rodríguez, 2008). For the trispectrum  $T_\zeta$  and the other higher order spectral functions, the momentum dependence also involves the direction of the wavevectors.

<sup>8</sup>There is actually a sign difference between the  $f_{NL}$  defined here and that defined in Ref. (Maldacena, 2003). The origin of the sign difference lies in the way the observed  $f_{NL}$  is defined (Komatsu & Spergel, 2001), through the Bardeen’s curvature perturbation (Bardeen, 1980):  $\Phi^B = \Phi_L^B + f_{NL}(\Phi_L^B)^2$  with  $\Phi^B = (3/5)\zeta$ , and the way  $f_{NL}$  is defined in Ref. (Maldacena, 2003), through the gauge invariant Newtonian potential:  $\Phi^N = \Phi_L^N + f_{NL}(\Phi_L^N)^2$  with  $\Phi^N = -(3/5)\zeta$  (Komatsu, 2008).

& Lyth, 2006; Maldacena, 2003):

$$B_{\zeta}(k_1, k_2, k_3) \equiv \frac{6}{5} f_{NL} \left[ P_{\zeta}(k_1) P_{\zeta}(k_2) + \text{cyclic permutations} \right], \quad (12)$$

$$T_{\zeta}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) \equiv \frac{1}{2} \tau_{NL} \left[ P_{\zeta}(k_1) P_{\zeta}(k_2) P_{\zeta}(|\mathbf{k}_1 + \mathbf{k}_4|) + \text{cyclic permutations} \right]. \quad (13)$$

Higher order spectral functions would be parametrized in an analogous way. Given the present observational state-of-the-art,  $n_{\zeta}$ ,  $f_{NL}$ , and  $\tau_{NL}$  are the statistical descriptors that discriminate among models for the origin of the large-scale structure once  $\mathcal{P}_{\zeta}^{1/2}$  has been fixed to the observed value. Since non-vanishing higher order spectral functions such as  $B_{\zeta}$  and  $T_{\zeta}$  imply non-gaussianity in the primordial curvature perturbation  $\zeta$ , the statistical descriptors  $f_{NL}$  and  $\tau_{NL}$  are usually called the levels of non-gaussianity.

The COBE satellite<sup>9</sup> provided us with a reliable value for the spectral amplitude  $\mathcal{P}_{\zeta}^{1/2}$  (Bunn & White, 1997):  $\mathcal{P}_{\zeta}^{1/2} = (4.957 \pm 0.094) \times 10^{-5}$  which is usually called the COBE normalisation. As regards the spectral index, the latest data release and analysis from the WMAP satellite<sup>10</sup> shows that  $n_{\zeta} = 0.960 \pm 0.014$  (Komatsu et. al., 2008) which rejects exact scale invariance at more than  $2\sigma$ . Such a result has been extensively used to constrain inflation model building (Alabidi & Lidsey, 2008; Alabidi & Lyth, 2006), and although several classes of inflationary models have been ruled out through the spectral index, lots of models are still allowed; that is why it is so important an appropriate knowledge of the statistical descriptors  $f_{NL}$  and  $\tau_{NL}$ . Present observations show that the primordial curvature perturbation  $\zeta$  is almost, but not completely, gaussian. The level of non-gaussianity  $f_{NL}$  in the bispectrum  $B_{\zeta}$ , after five years of data from NASA's WMAP satellite, is in the range  $-9 < f_{NL} < 111$  at  $2\sigma$  (Komatsu et. al., 2008). There is at present no observational bound on the level of non-gaussianity  $\tau_{NL}$  in the trispectrum  $T_{\zeta}$  although it was predicted that COBE should either measure  $\tau_{NL}$  or impose the lower bound  $|\tau_{NL}| \lesssim 10^8$  (Boubekeur & Lyth, 2006; Okamoto & Hu, 2002). It is expected that future WMAP data releases will either detect non-gaussianity or reduce the bounds on  $f_{NL}$  and  $\tau_{NL}$  at the  $2\sigma$  level to  $|f_{NL}| \lesssim 40$  (Komatsu & Spergel, 2001) and  $|\tau_{NL}| \lesssim 2 \times 10^4$  (Kogo & Komatsu, 2006) respectively. The ESA's PLANCK satellite<sup>11</sup> (The Planck Collaboration, 2006), whose launch is currently scheduled in the spring of 2009, promises to re-

duce the bounds to  $|f_{NL}| \lesssim 10$  (Komatsu & Spergel, 2001) and  $|\tau_{NL}| \lesssim 560$  (Kogo & Komatsu, 2006) at the  $2\sigma$  level if non-gaussianity is not detected. In addition, by studying the 21-cm emission spectral line in the cosmic neutral Hydrogen prior to the era of reionization, it is also possible to know about the levels of non-gaussianity  $f_{NL}$  and  $\tau_{NL}$ ; the 21-cm background anisotropies capture information about the primordial non-gaussianity better than any high resolution map of cosmic microwave background radiation: an experiment like this could reduce the bounds on the non-gaussianity levels to  $|f_{NL}| \lesssim 0.2$  (Cooray, 2006; Cooray, Li, & Melchiorri, 2008), and  $|\tau_{NL}| \lesssim 20$  (Cooray, Li, & Melchiorri, 2008) at the  $2\sigma$  confidence. Finally, it is worth stating that there have been recent claims about the detection of non-gaussianity in the bispectrum  $B_{\zeta}$  of  $\zeta$  from the WMAP 3-year data (Yadav & Wandelt, 2008). Such claims, which report a rejection of  $f_{NL} = 0$  at more than  $2\sigma$  ( $26.9 < f_{NL} < 146.7$ ), are based on the estimation of the bispectrum while using some specific foreground masks. The WMAP 5-year analysis (Komatsu et. al., 2008) shows a similar behaviour when using those masks, but reduces the significance of the results when other more conservative masks are included allowing again the possibility of exact gaussianity.

### 3 The model

According to the classification of inflationary models proposed in Ref. (Dodelson, Kinney, & Kolb, 1997), the small-field models are those of the form that would be expected as a result of spontaneous symmetry breaking, with a field initially near an unstable equilibrium point (usually taken to be at the origin) and rolling toward a stable minimum  $\langle \phi \rangle \neq 0$ . Thus, inflation occurs when the field is small relative to its expectation value  $\phi \ll \langle \phi \rangle$ . Some interesting examples are the original models of new inflation (Albrecht & Steinhardt, 1982; Linde, 1982), modular inflation from string theory (Dimopoulos & Lazarides, 2006), natural inflation (Freese, Frieman, & Olinto, 1990), and hilltop inflation (Boubekeur & Lyth, 2005). As a result, the inflationary potential for small-field models may be taken as

$$V = \sum_i \Lambda_i \left[ 1 - \left( \frac{\phi_i}{\mu_i} \right)^p \right], \quad (14)$$

where the subscript  $i$  here denotes the relevant quantities of the  $i$ th field,  $p$  is the same for all fields, and  $\Lambda_i$  and  $\mu_i$  are the

<sup>9</sup>NASA's COBE mission homepage: <http://lambda.gsfc.nasa.gov/product/cobe/>.

<sup>10</sup>NASA's WMAP mission homepage: <http://wmap.gsfc.nasa.gov/>.

<sup>11</sup>ESA's PLANCK mission homepage: <http://planck.esa.int/>.

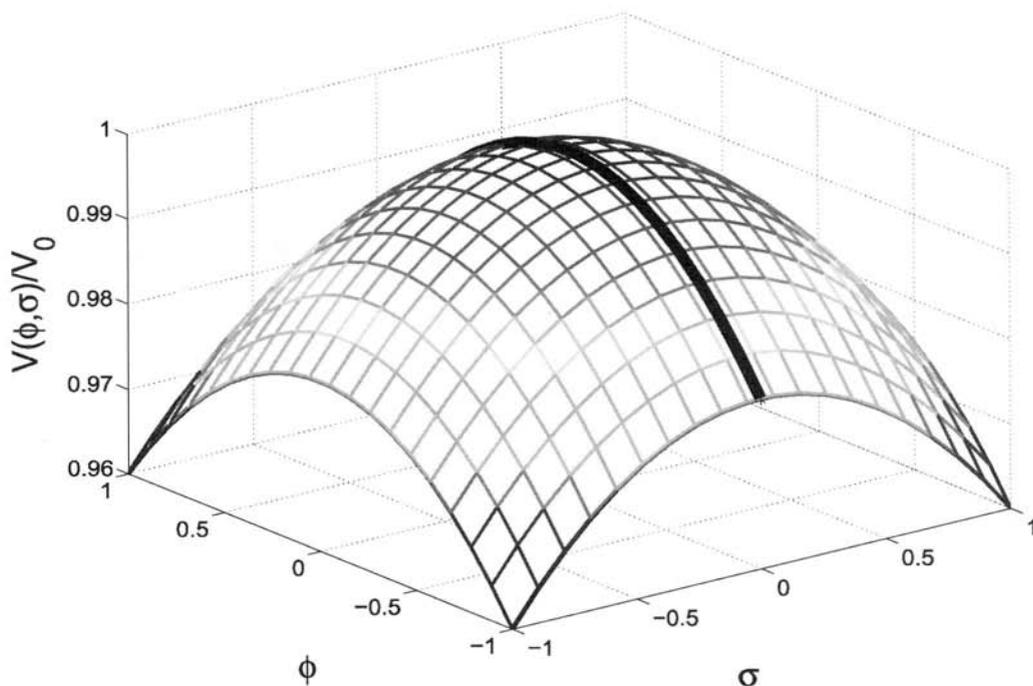


Figure 1: Our small-field slow-roll potential of Eq. (15) with  $\eta_\phi, \eta_\sigma < 0$ . The inflaton starts near the maximum and moves away from the origin following the  $\sigma = 0$  trajectory depicted with the solid black line. (This figure has been taken from Ref. (Alabidi, 2006)).

parameters describing the height and tilt of the potential of the  $i$ th field.

While Ref. (Ahmad, Piao, & Quiao, 2008) studies the spectrum of  $\zeta$  for general values of the parameter  $p$  and an arbitrary number of fields, assuming  $\zeta$  series convergence and tree-level dominance, we will specialize to the  $p = 2$  case for two fields  $\phi$  and  $\sigma$ :

$$V = V_0 \left( 1 + \frac{1}{2} \eta_\phi \frac{\phi^2}{m_P^2} + \frac{1}{2} \eta_\sigma \frac{\sigma^2}{m_P^2} \right), \quad (15)$$

where we have traded the expressions

$$\Lambda_1 + \Lambda_2 \quad \text{for} \quad V_0, \quad (16)$$

$$\frac{\Lambda_1}{\mu_1^2} \quad \text{for} \quad -V_0 \frac{\eta_\phi}{2m_P^2}, \quad (17)$$

and

$$\frac{\Lambda_2}{\mu_2^2} \quad \text{for} \quad -V_0 \frac{\eta_\sigma}{2m_P^2}, \quad (18)$$

and defined  $m_P$  as the reduced Planck mass. On doing this, and assuming that the first term in Eq. (15) dominates,

$\eta_\phi < 0$  and  $\eta_\sigma < 0$  become the usual  $\eta$  slow-roll parameters associated with the fields  $\phi$  and  $\sigma$ .

We have chosen for simplicity the  $\sigma = 0$  trajectory (see Fig. 1) since in that case the potential in Eq. (15) reproduces for some number of e-folds the hybrid inflation scenario (Linde, 1994) where  $\phi$  is the inflaton and  $\sigma$  is the waterfall field. Non-gaussianity in the bispectrum  $B_\zeta$  of  $\zeta$  for such a model has been studied in Refs. (Alabidi, 2006; Byrnes, Choi, & Hall, 2008; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Enqvist & Vähkönen, 2004; Lyth & Rodríguez, 2005a; Lyth & Rodríguez, 2005b; Vähkönen, 2005; Zaballa, Rodríguez, & Lyth, 2006); in particular, Ref. (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a) shows that the one-loop correction dominates over the tree-level terms, generating in this way large values for  $f_{NL}$  even if  $\zeta$  is generated during inflation. Refs. (Alabidi, 2006; Byrnes, Choi, & Hall, 2008), in contrast, work only at tree-level with the same potential as Eq. (15) but relaxing the  $\sigma = 0$  condition, finding that large values for  $f_{NL}$  are possible for a small set of initial conditions. Ref. (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b) studies the

trispectrum  $T_\zeta$  of  $\zeta$  in this model considering dominant loop corrections with  $\sigma = 0$ ; its results show that large values for  $\tau_{NL}$  are generated even if  $\zeta$  is generated during inflation.

The slow-roll conditions for single-field inflationary models with canonical kinetic terms read

$$\dot{\phi}^2 \ll V(\phi), \quad (19)$$

$$|\dot{\phi}| \ll |3H\dot{\phi}|, \quad (20)$$

where  $\phi$  is the inflaton field and  $V(\phi)$  is the scalar field potential. On defining the slow-roll parameters  $\epsilon$  and  $\eta_\phi$  as (Liddle & Lyth, 2000)

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \quad (21)$$

$$\eta_\phi \equiv \epsilon - \frac{\ddot{\phi}}{H\dot{\phi}}, \quad (22)$$

the slow-roll conditions in Eqs. (19) and (20) translate into strong constraints for the slow-roll parameters:  $\epsilon, |\eta_\phi| \ll 1$ . Multifield slow-roll models may also be characterized by a set of slow-roll parameters which generalize those in Eqs. (21) and (22) (Lyth & Riotto, 1999):

$$\epsilon_i \equiv \frac{m_P^2}{2} \left( \frac{V_i}{V} \right)^2, \quad (23)$$

$$\eta_i \equiv m_P^2 \frac{V_{ii}}{V}. \quad (24)$$

By writing the slow-roll parameters in terms of derivatives of the scalar potential, as in the last two expressions, we realize that the slow-roll conditions require very flat potentials to be met.

Since we are considering a slow-roll regime for the model given by the potential in Eq. (15), the evolution of the fields in such a case is given by

$$\phi(N) = \phi_* \exp(-N\eta_\phi), \quad (25)$$

and

$$\sigma(N) = \sigma_* \exp(-N\eta_\sigma), \quad (26)$$

in terms of the amount of inflation  $N$  since horizon exit, and the field values  $\phi_*$  and  $\sigma_*$  at the time when the relevant cosmological scales exit the horizon. Such expressions, together with Eq. (15), seed the  $\delta N$  formalism in order to calculate the spectrum and the trispectrum of the curvature perturbation including the tree-level (see Fig. 2) and the one-loop contributions (see Fig. 3)<sup>12</sup> (see the respective calculational details in Refs. Byrnes, Koyama, Sasaki, & Wands, 2007;

Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b; Lyth & Rodríguez, 2005a; Sasaki & Stewart, 1996):

$$\mathcal{P}_\zeta^{tree} = \frac{1}{\eta_\phi^2 \phi_*^2} \left( \frac{H_*}{2\pi} \right)^2, \quad (27)$$

$$\begin{aligned} \mathcal{P}_\zeta^{1-loop} &\simeq \frac{\eta_\sigma^2}{\eta_\phi^4 \phi_*^4} \exp[4N(|\eta_\sigma| - |\eta_\phi|)] \times \\ &\times \left( \frac{H_*}{2\pi} \right)^4 \ln(kL), \end{aligned} \quad (28)$$

$$\begin{aligned} T_\zeta^{tree} &= \frac{1}{\eta_\phi^4 \phi_*^6} \left( \frac{H_*}{2\pi} \right)^6 \left[ \frac{2\pi^2}{k_2^3} \frac{2\pi^2}{k_4^3} \frac{2\pi^2}{|\mathbf{k}_3 + \mathbf{k}_4|^3} + \right. \\ &\left. + 11 \text{ permutations} \right], \end{aligned} \quad (29)$$

$$\begin{aligned} T_\zeta^{1-loop} &\simeq \frac{\eta_\sigma^4}{\eta_\phi^8 \phi_*^8} \exp[8N(|\eta_\sigma| - |\eta_\phi|)] \left( \frac{H_*}{2\pi} \right)^8 \ln(kL) \times \\ &\times 4 \left[ \frac{2\pi^2}{k_2^3} \frac{2\pi^2}{k_4^3} \frac{2\pi^2}{|\mathbf{k}_3 + \mathbf{k}_4|^3} + \right. \\ &\left. + 11 \text{ permutations} \right], \end{aligned} \quad (30)$$

where  $L$  is the infrared cutoff chosen so that the quantities are calculated in a minimal box (Lyth, 2007), i.e.  $\ln(kL) \sim \mathcal{O}(1)$ , and  $k_1 \sim k_2 \sim k_3$ .

The important factor in the loop corrections is the exponential. This exponential function is directly related to the quadratic form of the potential with a leading constant term. It will give a large contribution if  $|\eta_\sigma| > |\eta_\phi|$ . We have chosen the concave downward potential in order to satisfy the spectral tilt constraint, which makes either  $\eta_\phi < 0$ , if  $\mathcal{P}_\zeta \simeq \mathcal{P}_\zeta^{tree}$ , or  $\eta_\sigma < 0$ , if  $\mathcal{P}_\zeta \simeq \mathcal{P}_\zeta^{1-loop}$ , while keeping  $|\eta_\sigma| > |\eta_\phi|$ .

## 4 Constraints to have a reliable parameter space

### 4.1 Convergence of the $\zeta$ series and existence of a perturbative regime

It has been proved (Cogollo, Rodríguez, & Valenzuela-Toledo, 2008a; Cogollo, Rodríguez, & Valenzuela-Toledo, 2008b) by means of a non-perturbative approach, that there exist what are called two ‘‘coupling constants’’  $x$  and  $y$  for the potential in Eq. (15). Such coupling constants allow us to obtain a necessary condition for the convergence of the  $\zeta$

<sup>12</sup>For an adequate explanation of the Feynman-like diagrams in cosmology, and their application within the  $\delta N$  formalism, see Ref. (Byrnes, Koyama, Sasaki, & Wands, 2007).

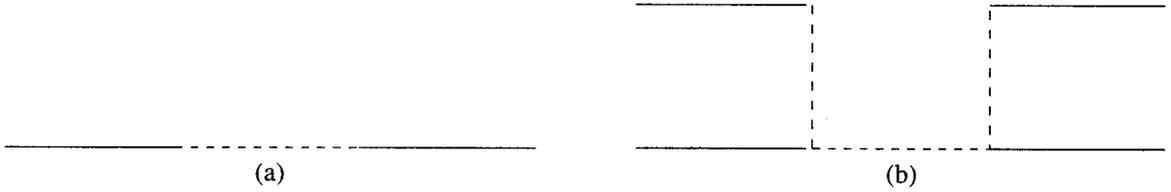


Figure 2: Tree-level Feynman-like diagrams for (a)  $\mathcal{P}_\zeta$ , and (b)  $\mathcal{T}_\zeta$ . The internal dashed lines correspond to two-point correlators of field perturbations.



Figure 3: One-loop Feynman-like diagrams for (a)  $\mathcal{P}_\zeta$ , and (b)  $\mathcal{T}_\zeta$ . The internal dashed lines correspond to two-point correlators of field perturbations.

series and work in a perturbative regime if they are much less than one:

$$|x| \equiv \left| \frac{\delta\phi_\star}{\phi_\star} \right| \approx \left( \frac{H_\star}{2\pi} \right) \frac{1}{\phi_\star} \ll 1, \quad (31)$$

$$|y| \equiv \left\{ \frac{\eta_\sigma^3}{\eta_\phi^3} \frac{\delta\sigma_\star^2}{\phi_\star^2} \exp[2N(\eta_\phi - \eta_\sigma)] \right\}^{1/2} \\ \approx \left\{ \frac{\eta_\sigma^3}{\eta_\phi^3} \left( \frac{H_\star}{2\pi} \right)^2 \frac{1}{\phi_\star^2} \exp[2N(\eta_\phi - \eta_\sigma)] \right\}^{1/2} \ll 1. \quad (32)$$

## 4.2 Tree-level or loop dominance

Because of the exponential factors in Eqs. (28) and (30) it might be possible that the loop corrections dominate over  $\mathcal{P}_\zeta$  and/or  $\mathcal{T}_\zeta$ . There are three possibilities in complete connection with the position of the  $\phi$  field when the relevant scales are exiting the horizon. Here we will consider only the intermediate  $\phi_\star$   $T$ -region<sup>13</sup>, corresponding to the case when  $\mathcal{T}_\zeta$  is dominated by one-loop corrections and  $\mathcal{P}_\zeta$  is dominated by the tree-level term, because this is the only possibility which gives interesting and observationally relevant results.

$\mathcal{T}_\zeta$  dominated by one-loop corrections and  $\mathcal{P}_\zeta$  dominated by the tree-level term: the intermediate  $\phi_\star$   $T$ -region

Looking at Eqs. (27) and (30) we require in this case that

$$\frac{\eta_\sigma^2}{\eta_\phi^2} \exp[4N(|\eta_\sigma| - |\eta_\phi|)] \ll \frac{1}{\phi_\star^2 \left( \frac{H_\star}{2\pi} \right)^2}, \quad (33)$$

$$4 \frac{\eta_\sigma^4}{\eta_\phi^4} \exp[8N(|\eta_\sigma| - |\eta_\phi|)] \gg \frac{1}{\phi_\star^2 \left( \frac{H_\star}{2\pi} \right)^2}, \quad (34)$$

which combine to give

$$\frac{r\mathcal{P}_\zeta}{8} \frac{\eta_\sigma^2}{\eta_\phi^2} \exp[4N(|\eta_\sigma| - |\eta_\phi|)] \ll \left( \frac{\phi_\star}{m_P} \right)^2 \\ \ll \frac{r\mathcal{P}_\zeta}{8} \frac{4\eta_\sigma^4}{\eta_\phi^4} \exp[8N(|\eta_\sigma| - |\eta_\phi|)], \quad (35)$$

where the definition for the tensor to scalar ratio  $r$  (Lyth, 2008) has been employed:

$$r \equiv \frac{\mathcal{P}_T}{\mathcal{P}_\zeta} = \frac{8}{m_P^2} \left( \frac{H_\star}{2\pi} \right)^2. \quad (36)$$

In the latter expression,  $\mathcal{P}_T^{1/2}$  represents the amplitude of the spectrum for primordial gravitational waves.

<sup>13</sup>The  $T$  in  $T$ -region is introduced in this paper in order to differentiate explicitly between these regions and those found in the companion paper (Rodríguez, 2008) for  $B_\zeta$ .

### 4.3 Spectrum normalisation condition

Since we are considering  $\zeta$  being generated during inflation, we must satisfy the appropriate spectrum normalisation condition. According to Eq. (27) if  $\mathcal{P}_\zeta$  is dominated by the tree-level term, we have

$$\mathcal{P}_\zeta^{tree} = \frac{1}{\eta_\phi^2 \phi_*^2} \left( \frac{H_*}{2\pi} \right)^2 = \frac{1}{\eta_\phi^2} \left( \frac{m_P}{\phi_*} \right)^2 \frac{r\mathcal{P}_\zeta}{8} = \mathcal{P}_\zeta, \quad (37)$$

which reduces to

$$\left( \frac{\phi_*}{m_P} \right)^2 = \frac{1}{\eta_\phi^2} \frac{r}{8}. \quad (38)$$

Notice that in such a situation, the value of the  $\phi$  field when the relevant scales are exiting the horizon depends exclusively on the tensor to scalar ratio, once  $\eta_\phi$  has been fixed by the spectral tilt constraint.

### 4.4 Spectral tilt constraint

The current observed value for the spectral tilt is  $n_\zeta - 1 = -0.040 \pm 0.014$  (Komatsu et. al., 2008), and again we will consider only the case when  $\mathcal{P}_\zeta$  is dominated by the tree-level term. That means that the usual spectral index formula (Sasaki & Stewart, 1996) applies:

$$n_\zeta - 1 = -2\epsilon - 2m_P^2 \frac{\sum_{ij} V_i N_j N_{ij}}{V \sum_i N_i^2}, \quad (39)$$

giving the following result once the derivatives of  $N$  with respect to  $\phi_*$  and  $\sigma_*$  have been calculated:

$$n_\zeta - 1 = -2\epsilon + 2\eta_\phi. \quad (40)$$

The effect of the  $\epsilon$  parameter may be discarded in the previous expression since  $\epsilon$  is much less than  $|\eta_\phi|$ :

$$\epsilon = \frac{m_P^2}{2} \frac{V_\phi^2 + V_\sigma^2}{V^2} = |\eta_\phi| \left[ \frac{1}{2} |\eta_\phi| \left( \frac{\phi}{m_P} \right)^2 \right] \ll |\eta_\phi|, \quad (41)$$

according to the prescription that the potential in Eq. (15) is dominated by the constant term. Thus, using the central value for  $n_\zeta - 1$ , we get

$$\eta_\phi = -0.020. \quad (42)$$

### 4.5 Amount of inflation

It is well known that the number of e-folds of expansion from the time the cosmological scales exit the horizon to the end of inflation is presumably around but less than 62 (Dodelson, 2003; Liddle & Lyth, 2000; Weinberg, 2008). The slow-roll evolution of the  $\phi$  field in Eq. (25) tells us that such an amount of inflation is given by

$$N = \frac{1}{|\eta_\phi|} \ln \left( \frac{\phi_{end}}{\phi_*} \right) \lesssim 62, \quad (43)$$

where  $\phi_{end}$  is the value of the  $\phi$  field at the end of inflation. Because of the characteristics of the inflationary potential in Eq. (15), there is no a definite mechanism to end inflation in this model. It could not be by means of the violation of the  $\epsilon < 1$  condition since this would imply extrapolating our results to a region where the potential in Eq. (15) is no longer dominated by the constant term which, in addition, would spoil the large non-gaussianity generated and could send the model to an unknowable quantum gravity regime. Keeping in mind the results of Ref. (Armendariz-Picon, Fontanini, Penco, & Trodden, 2008) which say that the ultraviolet cut-off in cosmological perturbation theory could be a few orders of magnitude bigger than  $m_P$ , we will therefore assume that inflation comes to an end when  $|\eta_\phi| \phi^2 / 2m_P^2 \sim 10^{-2}$ . This allows us to be in a safe side (avoiding large modifications to the potential coming from ultraviolet cutoff-suppressed non-renormalisable terms, and keeping the potential dominated by the constant  $V_0$  term), leaving the implementation of a mechanism to end inflation for a future work<sup>14</sup>. Coming back to Eq. (43), we get then

$$N = \frac{1}{|\eta_\phi|} \ln \left( \frac{m_P}{\phi_*} \right) \lesssim 62, \quad (44)$$

which leads to

$$\frac{\phi_*}{m_P} \gtrsim \exp(-62|\eta_\phi|). \quad (45)$$

## 5 $\tau_{NL}$

In this section we will calculate the level of non-gaussianity represented in the parameter  $\tau_{NL}$  (Boubekeur & Lyth, 2006) by taking into account the constraints presented in Section 4 (Cogollo, Rodríguez & Valenzuela-Toledo, 2008b).

<sup>14</sup>We hope that the implementation of such a mechanism in our model will keep, or perhaps enhance, the generated non-gaussianity. Nevertheless the opposite behaviour might as well happen. For instance, Ref. (Rigopoulos, Shellard, & van Tent, 2007) studies within a stochastic formalism a quadratic two-component slow-roll model without a dominant constant term in the potential. A momentary violation of the slow-roll conditions around the end of inflation shows to enhance  $f_{NL}$  to observable levels; however, such an enhancement vanishes once inflation ends completely. These results have been confirmed numerically within the  $\delta N$  formalism in Refs. (Vernizzi & Wands, 2006; Yokoyama, Suyama, & Tanaka, 2008).

The level of non-gaussianity, according to the expressions in Eqs. (12), (27), and (30), is in this case given by

$$\begin{aligned}
 \frac{1}{2}\tau_{NL} &= \frac{T_{\zeta}^{1-loop}}{8\pi^6 \left[ \frac{1}{k_3^2 k_4^3 |k_3+k_4|^3} + 23 \text{ perm.} \right] (\mathcal{P}_{\zeta}^{tree})^3} \\
 &= \frac{2\eta_{\sigma}^4}{\eta_{\phi}^2 \phi_{\star}^2} \exp[8N(|\eta_{\sigma}| - |\eta_{\phi}|)] \left( \frac{H_{\star}}{2\pi} \right)^2 \ln(kL) \\
 &= \frac{2\eta_{\sigma}^4}{\eta_{\phi}^2} \exp[8N(|\eta_{\sigma}| - |\eta_{\phi}|)] \left( \frac{m_P}{\phi_{\star}} \right)^2 \times \\
 &\quad \times \frac{r\mathcal{P}_{\zeta}}{8} \ln(kL) \\
 &= 2\eta_{\sigma}^4 \exp[8N(|\eta_{\sigma}| - |\eta_{\phi}|)] \mathcal{P}_{\zeta} \ln(kL) \\
 \Rightarrow \frac{1}{2}\tau_{NL} &\simeq 4.91 \times 10^{-9} |\eta_{\sigma}|^4 \times \\
 &\quad \times \exp[400 \ln(5.657 \times 10^{-2} r^{-1/2})] \times \\
 &\quad \times (|\eta_{\sigma}| - 0.020), \tag{46}
 \end{aligned}$$

where in the last line we have used expressions in Eqs. (38), (42), and (44).

In figure 4 we show lines of constant  $\tau_{NL}$  in the plot  $r$  vs  $|\eta_{\sigma}|$  for the intermediate  $\phi_{\star}$   $T$ -region in agreement with the constraints in Eqs. (31), (32), and (35). Notice that by implementing the spectral tilt constraint in Eq. (42) to the spectrum normalisation constraint in Eq. (38) and the amount of inflation constraint in Eq. (45) we conclude that the tensor to scalar ratio is bounded from below:  $r \gtrsim 2.680 \times 10^{-4}$ .

## 6 Conclusions

As is evident from the plot, the observationally expected  $2\sigma$  range of values, for WMAP, PLANCK, and even the 21 cm background anisotropies, and for positive  $\tau_{NL}$ ,  $\tau_{NL} \gtrsim 20$ , are completely inside the intermediate  $\phi_{\star}$   $T$ -region as required. Bigger values for  $\tau_{NL}$ , up to  $\tau_{NL} = 1.7 \times 10^5$  are consistent within our framework for the intermediate  $\phi_{\star}$   $T$ -region.

In the companion paper (Rodríguez, 2008), one of us studied  $f_{NL}$  for the case when  $\zeta$  is generated during inflation,  $B_{\zeta}$  is dominated by the one-loop correction, and  $P_{\zeta}$  is dominated by the tree-level term. Fig. 5 shows the results found. The WMAP (and also PLANCK) observationally allowed  $2\sigma$  range of values for negative  $f_{NL}$ ,  $-9 < f_{NL}$ , is completely inside the intermediate  $\phi_{\star}$  region<sup>15</sup>. More negative values for  $f_{NL}$ , up to  $f_{NL} = -20.647$ , are consistent within our framework for the intermediate  $\phi_{\star}$  region, but they

are ruled out from observation. Fig. 6 shows both Figs. 4 and 5 in the same plot. Incidentally, for the available parameter window, lines for constant  $\tau_{NL}$  almost exactly matches lines for constant  $f_{NL}$ . Thus, it is possible to see that, according to the observational state-of-the-art presented in Section 2, *non-gaussianity is more likely to be detected through the trispectrum than through the bispectrum*, for the particular inflationary model studied in this paper, and from the WMAP, PLANCK, and even the 21 cm background anisotropies observations. Fig. 6 also shows some *consistency relations between the values of  $f_{NL}$  and  $\tau_{NL}$*  that will be useful at testing the particular model considered against observations. For instance, if WMAP detected non-gaussianity through the trispectrum with  $\tau_{NL} \geq 8 \times 10^4$  at the  $2\sigma$  level, the quadratic two-field slow-roll inflationary model considered in this paper would be ruled out since the predicted  $f_{NL}$  would be outside the current observational interval.

Similarly to the  $f_{NL}$  case studied in the companion paper (Rodríguez, 2008), it is interesting to see a slow-roll inflationary model with canonical kinetic terms where large, and observable, values for  $\tau_{NL}$  may be obtained (in contrast to the expected  $\tau_{NL} \sim \mathcal{O}(\epsilon, \eta_i)$  from the tree-level calculation (Seery & Lidsey, 2007; Seery, Sloth, & Vernizzi, 2008)). So we conclude that *if  $T_{\zeta}$  is dominated by the one-loop correction but  $P_{\zeta}$  is dominated by the tree-level term, sizeable non-gaussianity is generated even if  $\zeta$  is generated during inflation*. We also conclude, from looking at the small values that the tensor to scalar ratio  $r$  takes in figure 6 compared with the present technological bound  $r \gtrsim 10^{-3}$  (Friedman, Cooray, & Melchiorri, 2006), that *for non-gaussianity to be observable in this model, primordial gravitational waves must be undetectable*.

**Acknowledgments:** This work is supported by COLCIENCIAS grant No. 1102-333-18674 CT-174-2006, DIEF (UIS) grant No. 5134, and by the ECOS-NORD programme grant No. C07P02.

## References

- Ackerman L., Carroll S.M., & Wise M.B., 2007. Imprints of a primordial preferred direction on the microwave background. *Phys. Rev. D* **75**, 083502.
- Ahmad I., Piao Y.-S., & Quiao C.-F., 2008. The spectrum of curvature perturbation for multi-field inflation with a small-field potential. *JCAP* **0802**, 002.
- Alabidi L., 2006. Non-gaussianity for a two component hybrid model of inflation. *JCAP* **0610**, 015.
- Alabidi L. & Lidsey J.E., 2008. Single field inflation after the WMAP five-year data. *Phys. Rev. D* **78**, 103519.

<sup>15</sup>The intermediate  $\phi_{\star}$   $T$ -region (where  $T_{\zeta}$  is dominated by the one-loop correction and  $P_{\zeta}$  is dominated by the tree-level term) encloses the intermediate  $\phi_{\star}$  region (where  $B_{\zeta}$  is dominated by the one-loop correction and  $P_{\zeta}$  is dominated by the tree-level term).

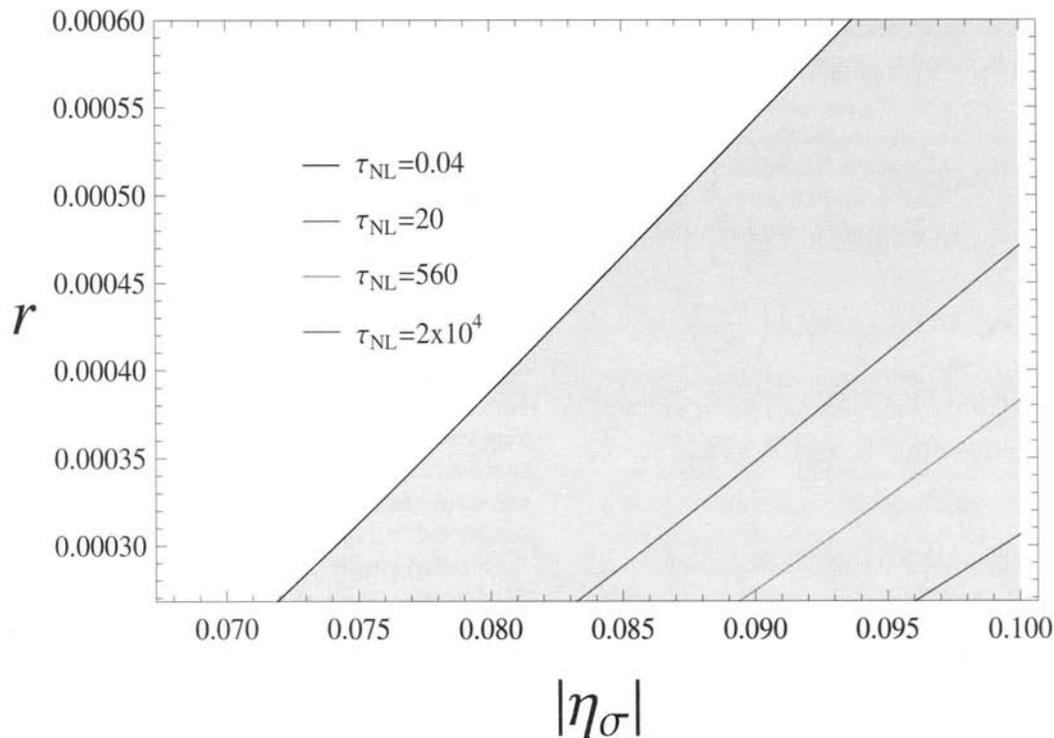


Figure 4: Contours of  $\tau_{NL}$  in the  $r$  vs  $|\eta_\sigma|$  plot. The intermediate (high)  $\phi_*$   $T$ -region corresponds to the shaded (white) region. The observationally expected  $2\sigma$  range of values, for WMAP, PLANCK, and even the 21 cm background anisotropies, and for positive  $\tau_{NL}$ ,  $\tau_{NL} > 20$  are completely inside the intermediate  $\phi_*$   $T$ -region. Notice that the boundary line between the high and the intermediate  $\phi_*$   $T$ -regions matches almost exactly the  $\tau_{NL} = 0.04$  line.

**Alabidi L. & Lyth D.H.**, 2006. Inflation models after WMAP year three. *JCAP* **0608**, 013.

**Albrecht A. & Steinhardt P.J.**, 1982. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Phys. Rev. Lett.* **48**, 1220.

**Armendariz-Picon C., Fontanini M., Penco R., & Trodden M.**, 2008. Where does cosmological perturbation break down?. *arXiv:0805.0114* [hep-th].

**Bardeen J.M.**, 1980. Gauge invariant cosmological perturbations. *Phys. Rev. D* **22**, 1882.

**Boubekeur L. & Lyth D.H.**, 2005. Hilltop inflation. *JCAP* **0507**, 010.

**Boubekeur L. & Lyth D.H.**, 2006. Detecting a small perturbation through its non-gaussianity. *Phys. Rev. D* **73**, 021301(R).

**Bunn E.F. & White M.J.**, 1997. The four-year COBE normalization and large-scale structure. *Astrophys. J.* **480**, 6.

**Byrnes C.T., Choi K.-Y., & Hall L.M.H.**, 2008. Conditions for large non-gaussianity in two-field slow-roll inflation. *JCAP* **0810**, 008.

**Byrnes C.T., Koyama K., Sasaki M., & Wands D.**, 2007. Diagrammatic approach to non-gaussianity from inflation. *JCAP* **0711**,

027.

**Carroll S.M., Tseng C.-Y., & Wise M.B.**, 2008. Translational invariance and the anisotropy of the cosmic microwave background. *arXiv:0811.1086* [astro-ph].

**Cogollo H.R.S., Rodríguez Y., & Valenzuela-Toledo C.A.**, 2008a. On the issue of the  $\zeta$  series convergence and loop corrections in the generation of observable primordial non-Gaussianity in slow-roll inflation. Part I: the bispectrum. *JCAP* **0808**, 029.

**Cogollo H.R.S., Rodríguez Y., & Valenzuela-Toledo C.A.**, 2008b. On the issue of the  $\zeta$  series convergence and loop corrections in the generation of observable primordial non-Gaussianity in slow-roll inflation. Part II: the trispectrum. *arXiv:0811.4092* [astro-ph].

**Cooray A.**, 2006. 21-cm background anisotropies can discern primordial non-gaussianity. *Phys. Rev. Lett.* **97**, 261301.

**Cooray A., Li C., & Melchiorri A.**, 2008. The trispectrum of 21-cm background anisotropies as a probe of primordial non-gaussianity. *Phys. Rev. D* **77**, 103506.

**Dimopoulos K. & Lazarides G.**, 2006. Modular inflation and the orthogonal axion as curvaton. *Phys. Rev. D* **73**, 023525.

**Dimopoulos K., Lyth D.H., & Rodríguez Y.**, 2008. Statistical

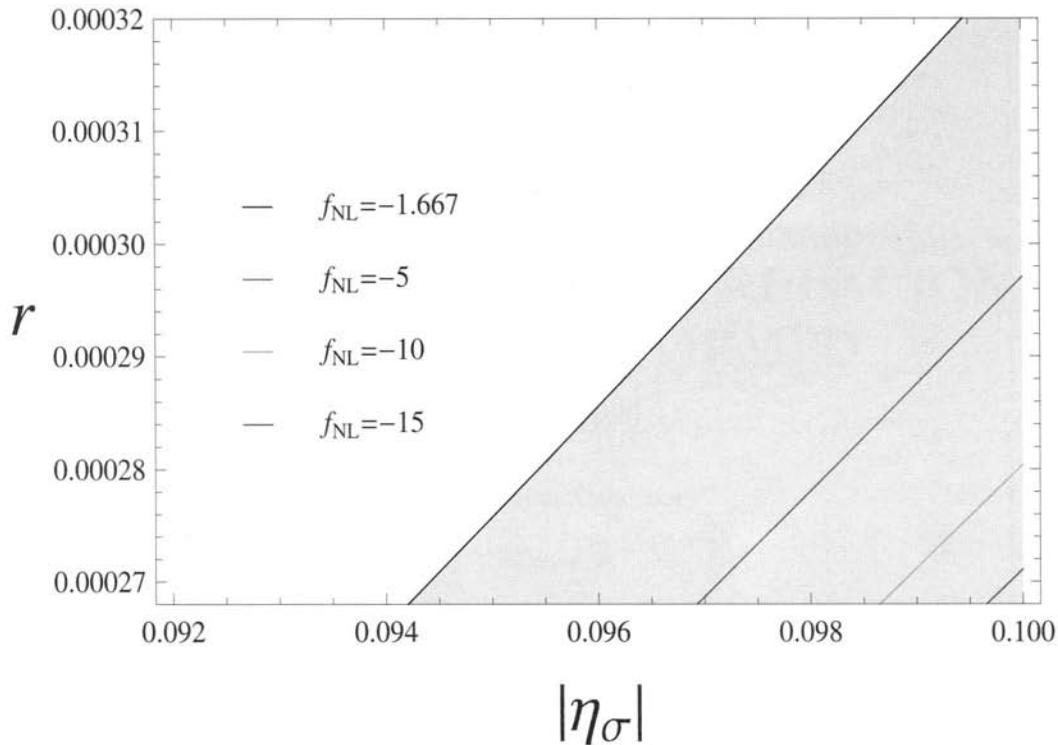


Figure 5: Contours of  $f_{NL}$  in the  $r$  vs  $|\eta\sigma|$  plot. The intermediate (high)  $\phi_*$  region corresponds to the shaded (white) region. The WMAP (and also PLANCK) observationally allowed  $2\sigma$  range of values for negative  $f_{NL}$ ,  $-9 < f_{NL}$ , is completely inside the intermediate  $\phi_*$  region. Notice that the boundary line between the high and the intermediate  $\phi_*$  regions matches almost exactly the  $f_{NL} = -1.667$  line. (This figure has been taken from Ref. (Rodríguez, 2008)).

anisotropy of the curvature perturbation from vector field perturbations. arXiv:0809.1055 [astro-ph].

Dodelson S., 2003. Modern cosmology, Academic Press, San Diego USA.

Dodelson S., Kinney W.H., & Kolb E.W., 1997. Cosmic microwave background measurements can discriminate among inflation models. Phys. Rev. D **56**, 3207.

Enqvist K. & Väihkönen A., 2004. Non-gaussian perturbations in hybrid inflation. JCAP **0409**, 006.

Freese K., Frieman J., & Olinto A., 1990. Natural inflation with pseudo-Nambu-Goldstone bosons. Phys. Rev. Lett. **65**, 3233.

Friedman B.C., Cooray A., & Melchiorri A., 2006. WMAP-normalized inflationary model predictions and the search for primordial gravitational waves with direct detection experiments. Phys. Rev. D **74**, 123509.

Kogo N. & Komatsu E., 2006. Angular trispectrum of CMB temperature anisotropy from primordial non-gaussianity with the full radiation transfer function. Phys. Rev. D **73**, 083007.

Komatsu E., 2008. Private communication.

Komatsu E. & Spergel D.N., 2001. Acoustic signatures in the primary microwave background bispectrum. Phys. Rev. D **63**, 063002.

Komatsu E. et. al., 2008. Five-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological interpretation. arXiv:0803.0547 [astro-ph].

Little A.R. & Lyth D.H., 2000. Cosmological inflation and large-scale structure, Cambridge University Press, Cambridge UK.

Linde A.D., 1982. A new inflationary universe scenario: a possible solution to the horizon, flatness, homogeneity, isotropy and primordial monopole problems. Phys. Lett. B **108**, 389.

Linde A.D., 1994. Hybrid inflation. Phys. Rev. D **49**, 748.

Lyth D.H., 2007. The curvature perturbation in a box. JCAP **0712**, 016.

Lyth D.H., 2008. Particle physics models of inflation. Lec. Notes Phys. **738**, 81.

Lyth D.H., Malik K.A., & Sasaki M., 2005. A general proof of the conservation of the curvature perturbation. JCAP **0505**, 004.

Lyth D.H. & Riotto A., 1999. Particle physics models of inflation and the cosmological density perturbation. Phys. Rep. **314**, 1.

Lyth D.H. & Rodríguez Y., 2005a. Inflationary prediction for primordial non-gaussianity. Phys. Rev. Lett. **95**, 121302.

Lyth D.H. & Rodríguez Y., 2005b. Non-gaussianity from the second-order cosmological perturbation. Phys. Rev. D **71**, 123508.

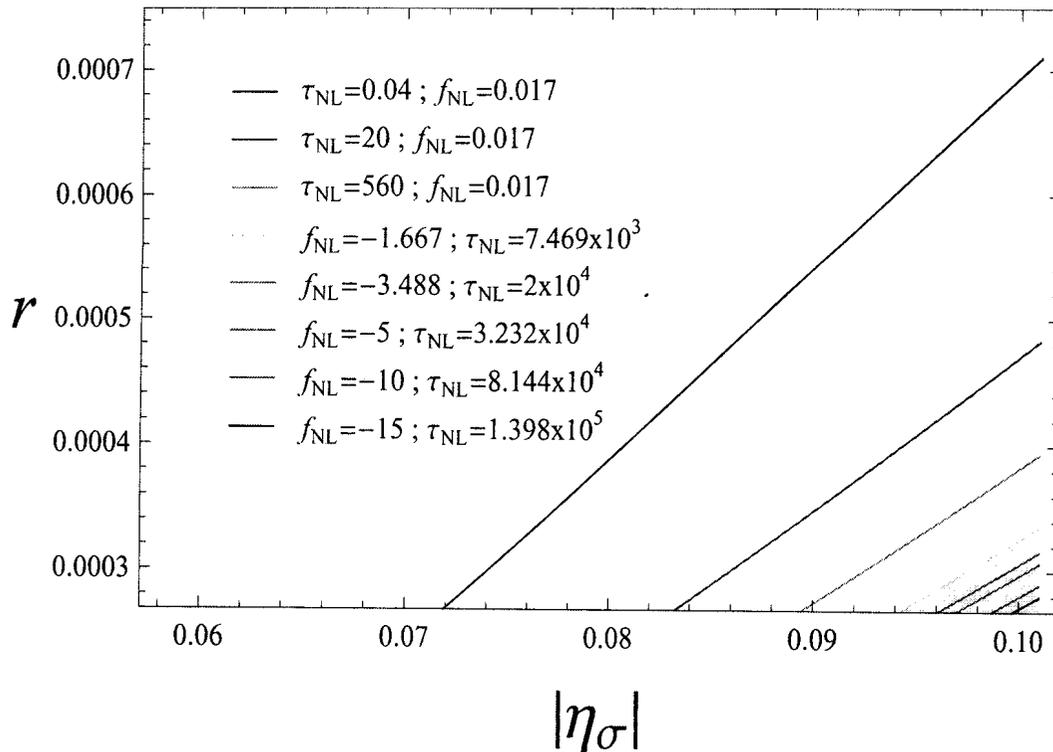


Figure 6: Contours of both  $f_{NL}$  and  $\tau_{NL}$  in the  $r$  vs  $|\eta_\sigma|$  plot. The intermediate (high)  $\phi_*$  region corresponds to the shaded (white) region. Lines for constant  $\tau_{NL}$  almost exactly matches lines for constant  $f_{NL}$ . According to this figure, and to the observational state-of-the-art, *non-gaussianity is more likely to be detected through the trispectrum than through the bispectrum*, for the particular inflationary model studied in this paper, and from the WMAP, PLANCK, and even the 21 cm background anisotropies observations. These lines also show some *consistency relations between the values of  $f_{NL}$  and  $\tau_{NL}$*  that will be useful at testing the particular model considered against observations.

**Maldacena J.**, 2003. Non-gaussian features of primordial fluctuations in single field inflationary models. *JHEP* **0305**, 013.

**Okamoto T. & Hu W.**, 2002. Angular trispectra of CMB temperature and polarization. *Phys. Rev. D* **66**, 063008.

**The PLANCK Collaboration**, 2006. The scientific programme of Planck. [arXiv:astro-ph/0604069](http://arxiv.org/abs/astro-ph/0604069).

**Rigopoulos G., Shellard E.P.S., & van Tent B.J.W.**, 2007. Quantitative bispectra from multifield inflation. *Phys. Rev. D* **76**, 083512.

**Rodríguez Y.**, 2008. Non-gaussianity and loop corrections in a quadratic two-field slow-roll model of inflation. Part I. Submitted to *Rev. Acad. Colomb. Cienc.*

**Sasaki M. & Stewart E.D.**, 1996. A general analytic formula for the spectral index of the density perturbations produced during inflation. *Prog. Theor. Phys.* **95**, 71.

**Seery D. & Lidsey J.E.**, 2007. Non-gaussianity from the inflationary trispectrum. *JCAP* **0701**, 008.

**Seery D., Sloth M., & Vernizzi F.**, 2008. Inflationary trispectrum from graviton exchange. [arXiv:0811.3934](http://arxiv.org/abs/0811.3934) [astro-ph].

**Starobinsky A.A.**, 1985. Multicomponent de Sitter (inflationary)

stages and the generation of perturbations. *Pisma Zh. Eksp. Teor. Fiz.* **42**, 124. [*JETP Lett.* **42**, 152].

**Väihkönen A.**, 2005. Comment on non-gaussianity in hybrid inflation. [arXiv:astro-ph/0506304](http://arxiv.org/abs/astro-ph/0506304).

**Vernizzi F. & Wands D.**, 2006. Non-gaussianities in two-field inflation. *JCAP* **0605**, 019.

**Weinberg S.**, 2008. *Cosmology*, Oxford University Press, Oxford UK.

**Yadav A.P.S. & Wandelt B.D.**, 2008. Evidence of primordial non-gaussianity ( $f_{NL}$ ) in the Wilkinson Microwave Anisotropy Probe 3-year data at  $2.8\sigma$ . *Phys. Rev. Lett.* **100**, 181301.

**Yokoyama S., Suyama T., & Tanaka T.**, 2008. Primordial non-gaussianity in multi-scalar inflation. *Phys. Rev. D* **77**, 083511.

**Zaballa I., Rodríguez Y., & Lyth D.H.**, 2006. Higher order contributions to the primordial non-gaussianity. *JCAP* **0606**, 013.

Recibido: febrero 3 de 2008

Aceptado para publicación: noviembre 13 de 2008

# HACIA UNA ONTOLOGÍA RELACIONAL DEL ESPACIO-TIEMPO

Por

Favio Ernesto CalaVitery<sup>1</sup>

## Resumen

Cala Vitery F. E.: Hacia una ontología relacional del espacio-tiempo. Rev. Acad. Colomb. Cienc. 32 (125): 527-543, 2008. ISSN 0370-3908.

En este artículo argumento –en contra de la opinión más generalizada– que la Teoría General de la Relatividad recibe su interpretación natural como una teoría relacional sobre el espacio-tiempo. Para ello me apoyo en tres argumentos: i) La ausencia de identidad primitiva para los puntos del espacio-tiempo en teorías covariantes. ii) El hecho de que el tensor métrico, al tener una estructura genuinamente dinámica, se entiende mejor como un objeto material en toda regla. iii) La posibilidad tanto hipotética como concreta de formular alternativamente –en el contexto de la dinámica clásica– la geometría y la inercia como estructuras relacionales y no como estructuras intrínsecamente sustantivistas.

**Palabras clave:** Ontología, Espacio-Tiempo, Relatividad General, Relacionismo, Substantivalismo.

## Abstract

I argue that –contrary to the widespread opinion– General Relativity should be interpreted as a relational spacetime theory. To do so, I rely on three arguments: i) The absence of primitive identity for spacetime points within generally covariant theories. ii) The fact that the metric tensor field, having a genuinely dynamical structure, should be better understood as a material object field in its own right. iii) The hypothetical and concrete possibility of formulating geometry and inertia –in the context of classical mechanics– as alternative relational structures rather than as intrinsically substantial structures.

**Key words:** Ontology, Spacetime, General Relativity, Relationism, Substantivalism.

<sup>1</sup> Universidad de Bogotá “Jorge Tadeo Lozano”. Colombia. Correo electrónico: favio.cala@utadeo.edu.co

## 1. Introducción

En la arena de la filosofía de la ciencia, ciertamente es común cuestionar la interpretación de la Mecánica Cuántica (MQ). Esto porque, a pesar del éxito generado por el uso, aplicación y desarrollo de su propio aparato formal, no estamos del todo persuadidos por la pintura que revela del mundo. Hay un problema de inteligibilidad. Su ontología, es decir, la interpretación de su formalismo, sigue siendo largamente discutida. Por otra parte, la interpretación de la Teoría General de la Relatividad (TGR), suele ser tomada por menos problemática a pesar de que importantes cuestiones estructurales no han estado exentas de un alto grado de controversia. Sin ir lejos, el cúmulo de dificultades para esclarecer el significado físico de la covariancia general, que seguramente es el rasgo más distintivo de la TGR, es buena muestra de esto.

La revolución científica operada conjuntamente por la teoría de gravitación de Einstein y la Mecánica Cuántica supuso una profunda transformación de las concepciones fundamentales de la física. Newton había construido la teoría científica más exitosa que ha conocido la humanidad. Con su dinámica logró asentar bien las nociones de espacio, tiempo, gravitación e inercia. Nociones todas estas indispensables para tratar el problema del movimiento y nociones todas estas que culminan su transformación posterior con la Teoría General de la Relatividad. En parte por esto, con el tiempo, la Teoría General de la Relatividad ha llegado a ser vista como la legataria natural de la dinámica newtoniana, mientras que la mecánica cuántica, al ocuparse fundamentalmente de cuestiones diferentes como la estructura de la materia y la naturaleza de la radiación, parece ajena a esta misma tradición.

Como resultado ha circulado la falsa impresión de que, al igual que la dinámica de Newton en su momento, la Teoría General de la Relatividad no tiene problemas de interpretación significativos.

El error es doble. Primero, porque procede de la falsa opinión de que las nociones de espacio, tiempo, inercia y movimiento 'bien establecidas' por Newton fueron, o maduraron hasta llegar a ser, inobjetables. Segundo, porque, una vez reconocidas algunas de las objeciones a la ontología de la dinámica newtoniana, procede de la no menos infundada suposición de que éstas han sido superadas del todo por su heredera natural, la Teoría General de la Relatividad. No es este el caso.

La dinámica de Newton, más allá de su eficacia predictiva, resultó desde un comienzo controvertible. Entre otras razones porque ésta -contra **Descartes**- parecía conferir realidad física independiente al espacio y al tiempo. **Huygens**, encontró absurda la idea de un movimiento 'verdadero' sin relación a cuerpos materiales, y por esta misma vía, **Leibniz** intentó rebatir, desde un inicio, la existencia física independiente del espacio y el tiempo absolutos argumentando que el soporte físico del movimiento debería estar anclado en la materia y sus relaciones. Así nació el debate sobre la existencia física del espacio. Debate este que quedó documentado en la serie epistolar que sostuvo el propio **Leibniz** durante sus dos últimos años de vida (1715-1716) con **Samuel Clarke**, fiel portavoz de **Newton**. Por entonces se estaban sentando los fundamentos de la física. Y es que es común afirmar que en buena medida la física moderna debe su origen al esfuerzo consistente por encontrar la unidad formal de un problema singular. Se trata del que la filosofía natural, desde **Aristóteles**, planteaba como el problema del movimiento local. Todo esto porque, antes de **Newton**, quienes se ocuparon de estudiar el movimiento local, que comprendía la caída de los graves, las órbitas planetarias y las trayectorias de proyectiles, buscaron expresar de una forma más o menos inteligible lo que significa que un cuerpo se mueva de un lugar a otro. En el fondo esta cuestión siempre estuvo latente la discusión sobre la naturaleza física independiente del espacio y el tiempo.

Discutiendo el asunto, **Newton** y **Leibniz** ya anticiparon buena parte de las dificultades interpretativas que nos siguen ocupando. Dejaron documentada su discusión al respecto en la polémica epistolar ya mencionada. La tradición recogió del enfrentamiento dos posiciones encontradas que pretendían responder a la siguiente cuestión: ¿Son el espacio y el tiempo entidades físicas reales en toda regla o son simplemente un conjunto de relaciones entre cuerpos materiales?

La herencia reciente de este debate germinal ha denominado **substantivalista** a la posición según la cual, siguiendo a **Newton**, el espacio es una entidad física que no debe su existencia a la presencia de objetos materiales y **relacionista** a la posición según la cual, siguiendo a **Leibniz**, el espacio no es otra cosa que el conjunto de relaciones entre objetos materiales coexistentes. Consideraciones similares sobre la naturaleza del tiempo son defendidas por los correspondientes costados del debate. Entretanto, mientras que para el substantivalista el espacio puede ser considerado como una especie

de contenedor o receptáculo para los objetos materiales y, en consecuencia, como el soporte universal de los fenómenos físicos, para el relacionista afirmar la sustancialidad (existencia) del espacio parece una concesión metafísica ociosa o un desafortunado desliz semántico. Para él, existe la materia y sus relaciones.

De este debate inicial entre relacionistas y substantivalistas (R-S) se acostumbra afirmar que hemos aprendido que el relacionista puede objetar, amparado en consideraciones epistemológicas, la invisibilidad del espacio. Ésta lleva al substantivalista a afirmar la existencia de situaciones físicas ontológicamente diferentes pero físicamente (empíricamente) indistinguibles. El espacio newtoniano permite este tipo de cosas y, discutiblemente, esto raya en el absurdo metafísico (Cala, 2006).

Por otra parte, **Newton** y la tradición substantivalista pudieron desechar esta objeción cargando a los relacionistas con la pesada loza de la inercia. Y esta no es una loza meramente epistémica. El cambio relacional, es decir, el cambio en la configuración relativa entre objetos materiales coexistentes resultaba insuficiente para proveer una justificación dinámica de los efectos inerciales. Baste con recordar los argumentos que **Newton** formuló en el esolío de sus Principia apoyado en el experimento del balde de agua en rotación y el de los dos globos atados por una cuerda, en un universo vacío, que giran en torno a su centro de gravedad común generando una tensión en la cuerda. En ambos casos parece insuficiente una justificación relacional de la inercia (Laymon 1978, Cala 2006).

El relacionismo de **Leibniz** fue recogido por el obispo **Berkeley** (1712), aunque sus objeciones al espacio absoluto newtoniano tomaron mejor forma, más tarde, con la crítica efectiva de **Mach** a la dinámica newtoniana. **Mach** no dudó en tildar de atavismo medieval a cualquier intento de conceder realidad física al espacio absoluto (Mach 1883, p.272). Y aunque el espacio invisible le pareciera una monstruosidad metafísica, hizo bien en reconocer que para proporcionar una ontología relacional alternativa se debía enfrentar el problema de la inercia. Esbozó un programa relacional que permitiría soportar la loza de la inercia. A sus críticas se sumo toda la tensión provocada por el desarrollo de la electrodinámica sobre los conceptos fundacionales de la mecánica. **Poincaré** y **Einstein** también se subieron al carro relacional, pero fue este último quien entendió, en la línea de **Mach**, que para erradicar la existencia física del espacio absoluto había que amarrar la estructura inercial de la dinámica a la distribución de materia estelar.

A la idea, **Einstein** terminó bautizándola como el Principio de Mach (**Einstein**, 1918). De hecho en 1916 ya había afirmado su satisfacción por haber estructurado una teoría relacional machiana en la forma definitiva de la Teoría General de la Relatividad (TGR). Hoy esta afirmación sigue siendo discutible. Y su discusión es el objeto central de este artículo.

Pero, ¿reivindica, finalmente, la Teoría General de la Relatividad el relacionismo à la *Leibniz-Mach* como **Einstein** pretendía? O, por el contrario, ¿en la tradición newtoniana, afirma la existencia independiente del espacio-tiempo, en este caso?

La cuestión parece estar lejos de ser zanjada. Y de esta cuestión me ocuparé en este artículo. Para ello, antes de entrar propiamente en los pormenores de la discusión actual, continúo con una breve historia de las tensiones que han caracterizado el debate en la era post-relativista.

**Einstein** había pretendido materializar el relacionismo inspirado en las ideas de Mach, al tiempo que buscaba una nueva teoría de gravitación. De paso trasladó la arena de discusión desde la sustancialidad del espacio y el tiempo hacia la sustancialidad del espaciotiempo. Durante el que quizá sea uno de los capítulos más interesantes de la historia de la gestación de la TGR, se sabe que **Einstein** logró sobreponerse a la limitada covarianza de sus ecuaciones de campo preliminares superando las objeciones que él mismo había impuesto mediante las restricciones debidas a su argumento del agujero (**Einstein** 1914).

Al rescatar la covarianza general en sus ecuaciones de campo definitivas **Einstein** había salido de su propio agujero argumentando que 'todas nuestras verificaciones espaciotemporales invariablemente corresponden a la determinación de coincidencias espacio-temporales' (**Einstein** 1918). Detalles sobre esta historia pueden hallarse en **Hoefler** 1994 y **Cala et al.** 2007.

La covarianza general es uno de los aspectos más distintivos de la teoría de **Einstein** y este tipo de aclaración vaga de su significado fue recogido con entusiasmo en tempranas interpretaciones de la teoría. En particular quienes se agrupaban bajo el positivismo lógico vieron en ésta una confirmación de la filosofía positiva de Mach y, presuntamente, una reivindicación del relacionismo à la Leibniz-Mach.

Al fin y al cabo, **Einstein** había superado las viejas nociones absolutas de espacio y tiempo newtoniano.

Este tipo de argumentación un tanto difusa fue defendida, entre otros, por **Hans Reichenbach**. Su libro sobre filosofía del espacio y del tiempo (**Reichenbach** 1928) enfatizaba el carácter empírico de la geometría física. En las notas introductorias a la edición inglesa, Rudolf Carnap, juzga al libro como la mejor contribución al tema hasta la fecha (1956)<sup>2</sup>. Así que en las tempranas interpretaciones positivistas de la TGR, la vertiente relacional fue felizmente patrocinada. Esta clima intelectual permaneció estable hasta que el desmoronamiento del positivismo lógico, el renacimiento del interés general por la cosmología relativista y el surgimiento del realismo científico obligaron a un replanteamiento de las cosas hacia finales de los años sesentas del siglo anterior.

En su lectura más inocente, el realismo científico interpreta las estructuras y objetos de nuestras mejores teorías literalmente. Es decir, le concede realidad física a las estructuras formales de la teoría. Los modelos de la TGR consisten en una variedad (manifold) y unos campos distribuidos sobre los puntos (eventos) de la variedad espaciotemporal. Desde la perspectiva realista pareció natural interpretar la variedad espaciotemporal como un espaciotiempo substantivista que aloja campos físicos. No es de sorprender que por entonces la variedad se interpretara como una entidad estructurada cuya existencia es independiente de la presencia de campos materiales. Ésta se había convertido en el sucedáneo del contenedor absoluto espacial de la física de Newton (**Earman** 1970, **Stein** 1970, **Friedman** 1983). La variedad es parte integral de nuestras mejores teorías y las ecuaciones de campo hacían ver a los campos como propiedades de los puntos -las partes- de la variedad espaciotemporal.

Este consenso sobre la existencia independiente del espaciotiempo en la forma de la variedad se rompió a finales de los años ochentas del siglo pasado. La filosofía del espaciotiempo se encontró entonces redescubriendo el viejo argumento del agujero de Einstein.

En su momento **Einstein** pensaba que la covarianza general implicada en el argumento violaba la ley de la causalidad. Igualmente, en su versión actualizada (**Earman y Norton** 1987), el argumento del agujero señalaba una dificultad interpretativa, a saber: si se suscribe una interpretación substantivista de la TGR,

ésta teoría debe ser juzgada como una teoría indeterminista. Esta conclusión es, desde luego, contraria a la concepción general de la teoría implicada en su práctica científica. Las reacciones no se hicieron esperar, bien fuera impugnando la forma de indeterminismo que involucra el argumento por resultar físicamente irrelevante, bien fuera suscribiendo directamente el relacionismo o, alternativamente, modificando la forma de entender el substantivalismo en la TGR.

Recuperados de las secuelas del agujero reabierto por **Earman y Norton** (1987) la mayoría de filósofos del espaciotiempo han optado por la última alternativa. A esta se le conoce con el nombre de substantivalismo sofisticado (SS) (**Mundy** 1992, **Brighouse** 1994, **Di Salle** 1994, **Hofer** 1996, **Bartels** 1996, **Pooley** 2002). En este tipo de doctrina ontológica se argumenta que la mejor forma de entender el espaciotiempo es suscribiendo una interpretación realista del mismo, pero que es erróneo juzgar a la variedad despojada de estructuras adicionales como si se tratase de un espaciotiempo en toda regla. Estas estructuras adicionales están codificadas en la métrica del espaciotiempo (el campo gravitacional). Es por esto que el substantivista sofisticado considera a la dupla variedad+métrica  $(M, g)$  como la forma correcta de entender el espaciotiempo. Para superar las dificultades generadas por el argumento del agujero el sustantivista sofisticado suscribe el principio relacional de la identidad de los indiscernibles en su versión moderna, conocida como equivalencia de Leibniz (EL). De este substantivalismo han quedado importantes lecciones relacionales que serán argumentadas más adelante.

En todo caso, la tensión del debate R-S sigue latente a pesar de que alguna de las reacciones más escépticas considera que el debate en el contexto actual está gastado (**Rynasiewicz** 1996). La razón para ello es que, presumiblemente, las categorías que le dieron vida en los tiempos de Newton y Leibniz se proyectan indistintamente en la física actual. Mi opinión es contraria. La cuestión sobre la mejor -la correcta- interpretación de la TGR es filosóficamente relevante, por una parte y, por otra, conduce a importantes consideraciones físicas (**Belot** 1996, **Hofer** 1998, **Belot e Earman** 2001, **Butterfeld e Isham** 1999, **Callender y Hugget** 2001).

En los tiempos de Descartes, Newton, Leibniz y Huygens se discutían los problemas fundacionales de la

<sup>2</sup>Hoy buena parte de los argumentos generales de **Reichenbach**, a pesar de su incuestionable estatura intelectual, son considerados como desacertados. Véase **Earman** 1989 p. 6. En este sentido el trabajo de **Friedman** (1983) y el del propio **Earman** con la herramienta de la geometría diferencial ha contribuido enormemente a una mejor comprensión de aspectos interpretativos y fundacionales de la TGR y de la filosofía del espaciotiempo.

mecánica. Se estaba pintando la imagen unitaria del mundo mediante la estructuración de una posible teoría física. Cuando esta teoría se estableció en la forma de la dinámica de Newton, algunos de sus problemas interpretativos no desaparecieron, pero la confianza en su poder pudo disipar la discusión filosófica por parecer físicamente intrascendente. Hoy, tres siglos después, el regreso a este tipo de filosofía de la naturaleza viene por cuenta de esa nueva búsqueda de la imagen unitaria del mundo, esto es, de la búsqueda de una teoría de unificación de todas las fuerzas (TU) o, alternativamente, de una teoría de gravitación cuántica (GQ). En este contexto muchos de los principales investigadores ven el debate como un instrumento importante para su trabajo (Smolin 1991, Rovelli 1996, Ashtekar 1998, Baez 2001, Barbour 1993, 2002). Pensar que el debate ha caducado conduce al error de suponer que la TGR es una teoría definitiva sin problemas interpretativos, pero también es el resultado de ignorar algunos adelantos hacia una mejor comprensión de la estructuras espaciotemporales de la dinámica pre-relativista (e.g. Barbour y Bertotti 1982).

Se ha mencionado insistentemente el argumento del agujero y aunque este es conocido empezaré por ofrecer una sucinta reconstrucción del mismo. Para esto utilizaré el lenguaje más económico de la geometría diferencial, en detrimento de la enunciación original hecha por Einstein en 1913 en lenguaje de coordenadas. Entre otras razones porque la enunciación del argumento en la presentación original del lenguaje de coordenadas se convirtió en una de las mayores fuentes de confusión para el propio Einstein. Tras presentar el argumento del agujero podré perfilar con mayor claridad mi posición en todo este asunto.

## 2. El argumento del agujero

En 1913, con ayuda la ayuda del matemático Marcel Grossman, Einstein publicó su primera teoría de gravitación tensorial. Esta teoría -conocida como la teoría *Entwurf*- postulaba ya un tensor métrico asociado mediante ecuaciones de campo a la distribución de materia-energía del universo. Pero a diferencia de la siguiente TGR ésta no era completamente covariante. Y aunque Einstein, en principio, estaba buscando una teoría de campo completamente covariante ya que estaba convencido de que la covarianza general era garantía teórica de un principio formal de relatividad general, llegó a convencerse durante un par de años (1913-1915)

de que la limitada covarianza de su teoría preliminar era una necesidad de orden relacional. En otras palabras, Einstein estaba convencido de que una teoría de gravitación covariante violaba la ley de la causalidad. Esto porque, según su interpretación de este tipo de ecuaciones de campo, una distribución de materia dada no definía unívocamente las componentes del tensor métrico y por tanto, la estructura inercial del espaciotiempo no quedaba formalmente determinada por la materia. Ello implicaba una violación de lo que después se conocería como el principio de Mach (Einstein 1918). La historia del calvario de Einstein para recuperar la covarianza general de sus ecuaciones de campo es compleja y muy sutil, pero no hay duda de que la principal objeción de Einstein a la covarianza general de la teoría estaba motivada por su determinación de edificar una teoría relacional al estilo anticipado por Mach. Tampoco hay duda de que esta objeción estaba soportada por el ya célebre argumento del agujero. El argumento del agujero fue concebido por Einstein en 1913, pero -como ya había mencionado- fue redescubierto por Earman y Norton en 1987. A continuación ofrezco una breve reconstrucción del agujero en una versión más cercana a la de Earman y Norton por considerar que su presentación en el lenguaje de la geometría diferencial resulta más limpia que la enunciación original en lenguaje de coordenadas. (¡Tomó casi dos años a Einstein superar la dificultades interpretativas para salir de su propio agujero!). Veamos<sup>3</sup>:

Cualquier modelo de la TGR (i.e., una solución a las ecuaciones de campo), que representa un universo posible, es usualmente representado por la triplete  $\langle M, g, T \rangle$ . Donde  $M$  es una variedad de puntos diferenciable con una cierta estructura topológica,  $g$  es el tensor métrico que codifica gravedad, estructura inercial y geometría, y  $T$  es el tensor de energía-momento que representa la distribución de materia-energía del universo posible.

La covarianza general de la TGR -y de cualquier teoría similar- implica, por definición, que si cualquier tensor  $X$  de la variedad es una solución a las ecuaciones de campo, también lo es el tensor  $\phi * X$  que resulta de empujar  $X$  mediante la acción de un difeomorfismo activo. En este caso  $X$  es cualquier tensor métrico o material que pueda definirse sobre la variedad. Un difeomorfismo activo es esencialmente una transformación de coordenadas que lleva tensores de sus puntos de origen a otros alterando de paso la forma del tensor. Es importante distinguirlos de los difeomorfismos pasivos

<sup>3</sup>Salvo algunas modificaciones menores, reproduzco *verbatim* el argumento del agujero del apéndice de Cala et al. 2007

que actúan localmente como transformaciones de coordenadas que simplemente rebautizan los puntos de la variedad sin generar *movimientos* ni *deformaciones*. Este tipo de transformaciones pasivas resultan triviales para nuestra discusión ya que cualquier teoría puede hacerse covariante general pasiva simplemente requiriendo que sus ecuaciones de campo sean escritas en forma tensorial (Kretschmann 1917, Friedman 1983). A cambio, un difeomorfismo activo  $\phi : M \rightarrow M$  lleva un punto  $p$  de  $M$  a otro punto  $q := \phi(p)$  en  $M$ , moviendo o deformando los ‘contenidos’ de la variedad.

El argumento del agujero es posible debido a la libertad para generar modelos de la TGR mediante la acción de este tipo de difeomorfismos activos. El conjunto de estos difeomorfismos forma un grupo denotado habitualmente como  $Diff(M)$ . Con esto en mente enuncio el argumento del agujero de Earman y Norton en forma compacta.

Sea  $U = \langle M, g, T \rangle$  un modelo de la TGR. Por definición, es posible escoger cualquier difeomorfismo  $\phi \in Diff(M)$  para generar un nuevo modelo  $U^* = \langle M, \phi * g, \phi * T \rangle$  de la TGR. Se escoje  $\phi^* = id$  (transformación identidad) en toda la variedad  $M$  exceptuando una región  $H \subset M$  dentro de la cual  $\phi$  difiere suavemente de la identidad. Esta región  $H$  es el *agujero* de Einstein. Dentro de éste  $\phi^* \neq id$ . Por conveniencia suponemos que el espaciotiempo  $M$  admite una foliación (3+1) mediante rebanadas (hipersuperficies) globales de tipo espacial (spacelike). Es decir, suponemos que podemos partir el espaciotiempo en tajadas espaciales tabuladas por un parámetro temporal  $t$  que aumenta en la dirección futura a medida que nos desplazamos a lo largo de una curva tipo-tiempo (timelike). Ahora escogemos  $H \subset M$  para  $t > 0$ .

De todo esto resulta que, como  $\phi^* = id$  para  $t \leq 0$ , pero difiere suavemente para  $t > 0$  (i.e dentro de  $H \subset M$ ), tenemos dos modelos de la teoría que difieren a partir de  $t = 0$ . Esto es,  $U = \langle M, g, T \rangle$  y  $U^* = \langle M, \phi * g, \phi * T \rangle$  son idénticos hasta  $t = 0$  pero difieren a partir de entonces<sup>4</sup>. Y esto debería entenderse como una violación del determinismo en la TGR ya que la especificación completa del espaciotiempo y sus contenidos materiales ( $g, T$ ) fuera del agujero no determina unívocamente la forma en que éstos evolucionan (o se

distribuyen) dentro de los puntos del agujero. En nuestro caso, con un agujero puesto en el futuro, según la teoría, el futuro no vendría determinado unívocamente por esta especificación completa del pasado. Falla el determinismo.

Pero aquí es importante advertir la forma en que falla el determinismo. Por ejemplo, si el centro de una galaxia  $\chi$  se encuentra ubicado hoy en el punto  $p$ , las ecuaciones de campo más el pasado no podrían determinar si mañana éste pasa por el punto  $q$  de  $M$  o por el punto  $r$  de  $M$ . Lo chocante del argumento no es que el determinismo falle, al fin y al cabo, el determinismo no tiene necesariamente que ser una camisa de fuerza de nuestras teorías. Ya tenemos a la mecánica cuántica en su interpretación estándar. Lo chocante proviene de la manera en que falla. El indeterminismo del agujero proviene de la imposibilidad para saber qué punto específico de la variedad subyace a qué proceso material concreto.

Las ecuaciones de campo no pueden escoger entre diferentes evoluciones de los campos dentro del agujero. En otras palabras, no pueden seleccionar a  $U = \langle M, g, T \rangle$  sobre  $U^* = \langle M, \phi * g, \phi * T \rangle$  para determinar si, por ejemplo, el centro de la galaxia  $\chi$  pasará por  $q$  de  $M$  o por  $r$  de  $M$ . Pero, el asunto importante es que los modelos  $U$  y  $U^*$  de la TGR son empíricamente indistinguibles. Comparten la misma variedad espaciotemporal de puntos subyacentes pero discrepan, dentro del *hueco*, sobre la forma en que los campos y la materia son distribuidos sin que esto suponga ningún tipo de diferencia observacional.

Todos los invariantes físicos de la teoría son preservados ante el tipo de transformaciones que generan el indeterminismo en el agujero. En tales circunstancias un realista ingenuo, al hacer una lectura literal de la variedad como una entidad física real cuyos puntos existen objetivamente y forman el tejido sustancial del espaciotiempo, enfrenta la siguiente traba interpretativa: Si la variedad  $M$  es considerada como un espacio físico real en toda regla, esto es, como una entidad capaz de existencia concreta con independencia de la ocurrencia de procesos materiales, entonces él debería considerar a  $U$  y  $U^*$  como modelos físicamente, ontológicamente, distintos aunque estos sean empíricamente equivalentes. Esto claramente recuerda el argumento de los mundos desplazados de Leibniz, donde un universo probable  $U_I$

<sup>4</sup>En la versión original de Einstein  $H$  no estaba, necesariamente, en el futuro. Pero fuera de  $H$  se suponía que el universo estaba lleno de materia ( $T \neq 0$ ), mientras que en su interior no había materia ( $T = 0$ ). Por esto recibió la denominación original de ‘agujero’. Era un hueco vacío. El difeomorfismo activo generado por Einstein parecía violar ‘la ley de la causalidad’ ya que la materia fuera del hueco no determinaba unívocamente la estructura métrica ni inercial dentro de éste. Esto, de paso, parecía una violación flagrante del principio Mach o de la *condicionalidad material del espacio*, como Einstein le llamaba por entonces.

resultaba empíricamente indistinguible de un universo probable  $U_s$ , al conservar el conjunto de relaciones entre objetos materiales coexistentes, pero difería de éste por su ubicación en el espacio absoluto. El Newtoniano estaba obligado a conceder, dada la identidad presupuesta para los puntos del espacio, que los dos casos correspondían a situaciones ontológicamente diferentes. En su lugar Leibniz apelaba a la identidad de los indiscernibles (PII) para concluir que se trataba de dos representaciones de un mismo universo o dicho de otro modo de dos maneras de nombrar la misma cosa (Cala 2006.). En la terminología y contexto actuales la identidad de los indiscernibles recibe el nombre de *Equivalencia de Leibniz*. Este es un principio de inspiración relacional y puede enunciarse así:

Equivalencia de Leibniz (EL): Dos distribuciones de campos relacionadas por un difeomorfismo activo representan la misma situación física, esto es,  $U = \langle M, g, T \rangle$  y  $U^* = \langle M, \phi * g, \phi * T \rangle$  son equivalentes para cualquier  $\phi \in Diff(M)$ .

En resumen, según Earman y Norton, el substantivista (realista ingenuo) debe rechazar la EL y concluir extrañamente que la TGR es una teoría indeterminista. Por supuesto que las leyes de la física pueden admitir el indeterminismo, bien sea porque la teoría sea intrínsecamente probabilista, bien sea por la presencia de singularidades, o por la incursión de *invasores espaciales* (Earman 1989 Cáp. 3). Pero todas estas parecen aceptables, o buenas, razones físicas. La cuestión es que el determinismo debería fallar por buenas razones de la física y no por compromisos con alguna doctrina ontológica que de entrada lo proscriba. En palabras de Earman, 'al determinismo debería dársele una oportunidad' (Earman 1989 p.180).

Los filósofos han visto diversas reacciones al argumento del agujero, pero la mayoría concuerdan con Earman y Norton en que el determinismo merece una oportunidad en la interpretación de la TGR. Al fin y al cabo, esta es la forma habitual de concebir la teoría en la práctica científica. Estoy de acuerdo. Aquí se trata de un indeterminismo que no afecta la determinación predictiva empírica de la teoría. Esto es raro. Por esto la mayoría suscribe la EL y rechaza el substantivalismo directo que sostiene el argumento del agujero. No es de extrañar que la conclusión es la misma a la que Einstein llegó escapando del agujero, esto es, que los puntos de la variedad no tienen significado ontológico independiente, es decir, que presuponer la identidad primitiva de los puntos de la variedad lleva al tipo de indeterminación

'ingenua' que a él le hizo caer en el enredo del agujero desde el principio. Cuando Einstein se dió cuenta de esto pudo, por fin, formular sus ecuaciones de campo gravitacional completamente covariantes entendiéndolas como ecuaciones de campo deterministas.

En cualquier caso, esta es ciertamente la concepción ampliamente mayoritaria entre filósofos del espacio-tiempo y cosmólogos activos. En su famoso tratado sobre la estructura del universo a gran escala, Hawking y Ellis (1973) simplemente se refieren a todos los modelos isomorfos de la TGR relacionados mediante  $Diff(M)$  como una clase equivalente que representa la misma solución de las ecuaciones de campo (i.e. un único universo posible) sin ninguna mención al argumento del agujero. La presentación de Wald (1984) es similar.

La presunción tácita de la EL, común a prácticamente todas las representaciones de la TGR, prohíbe pues la individuación apriorística de los puntos de la variedad. Para eleverlos a la categoría de eventos se requiere, como Einstein declaraba, la ocurrencia de coincidencias espaciotemporales y estas no pueden señalarse sin la presencia de algún campo físico definido sobre la misma. Pero aunque el rechazo generalizado a la identidad primitiva de los puntos del espaciotiempo -vía EL- pueda sonar muy relacional, la posición más consensuada se ha decantado por una interpretación substantivista de la TGR en la forma ya mencionada del substantivalismo sofisticado. La evidente tensión proviene, primero, de las diversas formas de definir el substantivalismo, y luego de los detalles interpretativos de estas definiciones. Por ahora, del argumento del agujero retengo el siguiente corolario: *En la TGR los puntos de la variedad no tienen identidad primitiva.*

A continuación argumentaré que esto, en parte, conduce mejor a una interpretación relacional de la TGR.

### 3. Sobre la mejor forma de plantear el debate cuando los puntos del espaciotiempo son inidentificables *per se*

El argumento del agujero está montado, se ha visto, sobre la suposición de que el substantivalismo es una doctrina realista sobre los puntos del espaciotiempo y que éstos constituyen el tejido del soporte universal de los procesos materiales y en general de los fenómenos. Por el contrario, el relacionismo debería refutar este estatus ontológico para el soporte universal de todos los fenómenos. Recientemente John Earman y Gordon

**Belot** han insistido en esta forma de plantear el debate R-S. Ellos escriben (**Belot y Earman** 2001, p 227):

*Substantivalists understand the existence of spacetime in terms of the existence of its pointlike parts, and gloss spatiotemporal relations between material contents in terms of the spatiotemporal relations between points at which they occur. Relationists will deny that spacetime points enjoy this robust sort of existence, and will accept spatiotemporal relations between events as primitive.*

En discusiones recientes es bien cierto que por momentos una distinción clara entre los participantes de cada costado del debate R-S puede resultar cuando menos desconcertante. La distinción que hacen **Earman y Belot** resulta interesante por cuanto parece estar, en principio, atada estrechamente a las premisas del debate original, cuando **Newton y Leibniz**, presumiblemente, podían 'darse el lujo de saber de que estaban hablando'<sup>5</sup>. Puntos o Materia. Contenedor o contenido. Esta forma de hablar hace parecer el debate una mera discusión sobre la aceptación de la EL (relacionismo) o su rechazo (substantivalismo).

Mi impresión es que la EL, o el rechazo a la identidad robusta o primitiva de los puntos del espaciotiempo, es una premisa necesaria aunque no suficiente para cualquier forma de relacionismo. Por su parte el substantivalismo es neutro en este asunto. Esto debe quedar claro al pasar de la representación convencional newtoniana de la dinámica clásica a su representación en el espaciotiempo neo-newtoniano. En esta última los puntos son despojados de su identidad, al remover el espacio absoluto, y el énfasis del sustancialismo está puesto en la existencia robusta de las estructuras espaciotemporales. Estas estructuras son independientes de la presencia de

materia y hacen inteligible la idea de un cuerpo único moviéndose en el espacio sin relación a nada (material)<sup>6</sup>.

De todas formas **Earman y Belot** trazan su distinción argumentando que en el contexto de las probables teorías de campo unificado (TU) o, alternativamente, de gravitación cuántica (GQ) debe aparecer esta separación limpia en el debate. Conuerdo en que la eventual formulación de una teoría en la que los puntos del espaciotiempo estuviesen cargados con el tipo de existencia robusta de la que hablan **Earman y Belot**, haría del substantivalismo una doctrina mucho más irrefutable, pero que yo sepa no existen teorías de este tipo asomando por el horizonte teórico actual. Por esto considero que es más saludable presentar al substantivalismo como una doctrina realista sobre la existencia independiente de estructuras espaciotemporales.

En esta línea, esta es la definición enunciada por **Hoefer** (1996, p. 5):

*A modern-day substantivalist thinks that spacetime is a kind of thing which can, in consistency with the laws of nature, exist independently of material things (ordinary matter, light, and so on) and which is properly described as having its own properties, over and above the properties of any material things that may occupy parts of it.*

Estoy de acuerdo en la definición general. Pero discrepo en los detalles importantes. Especialmente en la clasificación de las 'cosas materiales'<sup>7</sup>. **Hoefer** hace su definición hablando como un substantivalista sofisticado. Ya he dicho que el substantivalismo sofisticado (SS) ha sido considerado mayoritariamente como la mejor forma de entender el espaciotiempo y como la reacción más natural al argumento del agujero (**Mundy** 1992, **Brighouse** 1994, **Di Salle** 1994, **Hoefer** 1996, **Bartels**

<sup>5</sup>La frase citada es original de **Earman** (1989, p 18). También es citada por **Rynasiewicz** (1996, p. 286) y **Hoefer** (1998, p. 452).

<sup>6</sup>En la representación neo-newtoniana (El término es de **Sklar** 1976) de la mecánica clásica se remueve el espacio absoluto, es decir, se suprime la idea de que exista un marco de referencia universal privilegiado. Dicho de otro modo, se suprime la noción de reposo absoluto y retenemos la clase o familia equivalente de sistemas inerciales conectados por las simetrías del grupo de Galileo. Al espaciotiempo neo-newtoniano, por esta razón, también se le conoce como el espaciotiempo galileano

<sup>7</sup>No obstante, existen voces discordantes que vuelven a darle la vuelta al asunto. Por ejemplo **Saunders** (2002), sostiene que el relacionismo (no reductivo) es neutro en la distinción entre materia y espacio. Él está interesado en los principios relacionales de Leibniz (Principio de Razón Suficiente y PII) y su proyección en la forma de principios de simetría e invarianza en las teorías de campo (**Saunders** 2003). Considerando que la covarianza general activa de la TGR funciona como un principio de invariancia ante cualquier transformación del grupo  $Diff(M)$ , la EL parece -según el- suficiente para hacer de la TGR una teoría relacional, sin que esto implique una referencia explícita a la distinción entre materia y espacio. En este caso lo importante es que los invariantes de la teoría corresponden a especificaciones físicas relacionales, es decir, que no precisan puntos de referencia predeterminados. **Teller** (1991) argumenta en esta misma línea. Considero, al contrario, que la clave de la discusión está en la correcta distinción entre materia y espacio.

1996, Pooley 2002). Pero, ¿Qué es el substantivalismo sofisticado?

El substantivalista sofisticado es alguien que suscribe la EL, es realista sobre la existencia independiente del espaciotiempo pero no asigna el estatus ontológico de espaciotiempo completo a la variedad desnuda. Una de las razones, aparte del argumento del agujero, es que la variedad funciona como una colección de puntos con estructura diferenciable y topológica pero carece prácticamente de todas las propiedades *paradigmáticas* que permiten referirse a ésta como si se tratara de un espaciotiempo en toda regla.

Por supuesto, para encontrar las propiedades paradigmáticas del espaciotiempo habría que volver la mirada hacia la dinámica de Newton. Y es que consideramos el espacio absoluto newtoniano como una entidad real en toda regla porque éste tenía todas las estructuras que hacían teóricamente inteligible la idea de un cuerpo único en movimiento sin relación a nada (material). Y la variedad por sí sola no permite ni siquiera definir distancias espaciales ni temporales. No contiene la estructura inercial o conexión afín y tampoco permite distinguir pasado de futuro. En definitiva, la variedad dista mucho de cumplir el papel paradigmático que cumplían el espacio y el tiempo newtonianos o, en su defecto, el espaciotiempo neo-newtoniano.

Creo que esta razón es bastante fuerte. La variedad es una entidad abstracta no exclusiva de la TGR y sin cargarla con estructuras adicionales está lejos de cumplir el rol explicativo que cumplía el espacio en la dinámica de Newton. Estas estructuras adicionales son básicamente estructuras métricas. La variedad carece de la estructura métrica necesaria para hablar, de alguna forma, de un espaciotiempo inteligible. Resultaría más que paradójico considerarla como un espaciotiempo sin estructuras espaciotemporales. Por esto, en esencia, para el substantivalista sofisticado es más natural hablar de la dupla métrica+variedad  $(M, g)$  para referirse a un espaciotiempo físico real en toda regla<sup>8</sup>. El substantivalista sofisticado está convencido de que en la TGR, el rol explicativo que cumplía el espacio newtoniano es proporcionado por el tensor métrico  $(g)$ .

A mi juicio, el SS hace bien al señalar las debilidades de la variedad para ser considerada como un espaciotiempo en toda regla. Empero, resulta poco convincente al asignarle a la métrica el estatus de estructura

substantivalista. Al hacerlo, el substantivalista sofisticado está tomando todas las propiedades de la métrica como propiedades intrínsecas de un espaciotiempo real. La fuerza de esta línea argumental, vimos, depende de la clasificación de las estructuras espaciotemporales heredadas de la dinámica de Newton. Considero que esto no es correcto. Pero incluso si juzgamos a la métrica desde la perspectiva newtoniana las cosas no quedan muy claras. Veamos por qué.

El campo métrico de la TGR tiene energía y momento. Resulta más natural pensar que la energía y el momento son cargados por campos físicos materiales como el campo electromagnético o cualquier campo de distribuciones de densidades en un fluido, por ejemplo. ¿Por qué elevar, entonces, la métrica a la categoría de espaciotiempo real independiente sin más? ¿Dónde ponemos el campo gravitacional? ¿En la materia como todos los demás campos físicos o en las propiedades del espacio? ¿Qué pasa con las ondas gravitacionales y la energía contenida en campos gravitacionales sin fuentes? **Stachel** se expresa así en todo este asunto (Stachel 1993, p.144):

*Several philosophers of science have argued that the general theory of relativity actually supports spacetime substantivalism (if not separate spatial and temporal substantivalisms) since it allows solutions consisting of nothing but a differentiable manifold with a metric tensor field and no other fields present (empty spacetimes). This claim, however ignores the second role of the metric tensor field; if it is there chronogeometrically, it inescapably generates all the gravitational field structures. Perhaps the culprit here is the words 'empty spacetime'. An empty spacetime could also be called a pure gravitational field, and it seems to me that the gravitational field is just as real a physical field as any other. To ignore its reality in the philosophy of spacetime is just as perilous as to ignore it in everyday life.*

Estoy de acuerdo con **Stachel** en que es mejor considerar al campo métrico como un campo físico material pero no sólo porque integre las estructuras del campo gravitacional. En breve volveré sobre esto.

<sup>8</sup>Dado que el SS suscribe la EL sería más preciso referirse a la dupla  $(M, \{g\})$  como un espaciotiempo probable en toda regla. Aquí  $\{g\}$  es una clase equivalente de métricas relacionadas entre sí mediante cualquier  $\phi \in Diff(M)$ .

La TGR incorpora esencialmente, geometría, gravedad e inercia en la forma del campo métrico. En la física de Newton, geometría e inercia son cualidades estructurales del espacio, mientras que el campo gravitacional es un campo físico. El SS escoge geometría e inercia -o las estructuras cronogeométricas- mientras que los pocos relacionistas como **Stachel**, **Rovelli** y otros escogen el campo gravitacional para enfatizar sus respectivas posiciones. Esto es lo que dice **Rovelli** al respecto (**Rovelli** 1997, p. 193):

*Einstein's identification between gravitational field and geometry can be read in two alternative ways:*

*i. as the discovery that the gravitational field is nothing but a local distortion of spacetime geometry; or*

*ii. as the discovery that spacetime geometry is nothing but a manifestation of a particular field, the gravitational field.*

*The choice between these two points of view is a matter of taste, at least as long as we remain within the realm of nonquantistic and nonthermal general relativity. I believe, however, that the first view, which is perhaps more traditional, tends to obscure, rather than enlighten, the profound shift in the view of spacetime produced by general relativity*

...

Encuentro convincente la interpretación relacional (ii), pero a estas alturas uno puede imaginarse por qué **Rynasiewicz** (1996) argumenta que el debate ha caducado, o por qué **Saunders** (2002) intenta definir un relacionismo neutro en cuanto a la distinción entre materia y espacio. Creo que a pesar de las dificultades rendirse no es un camino acertado y tampoco lo es ignorar la distinción entre materia y espacio. Este es el vértice de la cuestión.

Recordemos que la Teoría Especial de la Relatividad (TER) tiene básicamente las mismas estructuras -geometría e inercia- que la dinámica de Newton incorporadas a las cualidades del espacio<sup>9</sup>. En este sentido perpetúa la tradición newtoniana. En la TGR las cosas son distintas. El desencanto de **Rynasiewicz** y **Saunders** es comprensible porque, al ser usualmente juzgada desde la tradición newtoniana, la estructura teórica y el papel explicativo del tensor métrico en la TGR no

parece permitir una distinción clara para decantar el debate. 'Parece una cuestión de gusto'. La razón, creo, para haber convertido a la interpretación substantivalista en la más tradicional es, tautológicamente, la fuerza de la tradición.

Quiero insistir en lo siguiente: La dinámica Newtoniana, paradigmática de nuestras ideas sobre la materia y el espacio, es entendida como una teoría substantivalista sobre el espacio (y el tiempo). No es de extrañar que mayoritariamente físicos, cosmólogos y astrofísicos, hayan tomado a la TGR, siguiendo la usanza newtoniana, como una teoría substantivalista sobre el espacio-tiempo. La mayoría de filósofos de la ciencia también han suscrito esta interpretación pero, ante las dificultades presentadas por la ambigüedad en la lectura de la TGR, deben existir mejores argumentos que la tradición.

Es bien cierto que **Newton** tenía una teoría completa y **Leibniz** no. Pero liquidar al relacionismo amparándose en la victoria de **Newton** es, a mi juicio, puro atavismo gratuito. Podemos superar esto. Yo propongo el siguiente ejercicio: ¿Qué tal si tomamos por buenas algunas de las premisas relacionales de **Leibniz** (o **Mach**)? En este caso, la métrica -el nudo del asunto- ¿qué papel explicativo cumpliría? ¿Estaría bien planteada la elección entre geometría (i.) y campo gravitacional (ii.) planteada por **Rovelli** (y compartida prácticamente por todos los participantes del debate R-S en la TGR)? ¿Sería la geometría entendida como una cualidad intrínseca de un espacio real independiente de la materia?

Veamos: Para **Leibniz** el espacio no es otra cosa que el conjunto de relaciones entre objetos materiales coexistentes. En este caso, la geometría debe funcionar como una estructura que codifica el conjunto de relaciones métricas entre objetos materiales coexistentes. Pensemos que **Leibniz** tuvo una teoría buena. ¿Sería sano concluir que la geometría *per se* es una entidad relacional pura y por tanto la TGR es mejor entendida como una teoría relacional? Estoy casi seguro de que todo el mundo respondería negativamente a esto. Sin embargo, muchos filósofos encuentran muy natural hacer lo propio con la geometría para argumentar una interpretación substantivalista de la TGR. En realidad la geometría, sin más, parece neutra en este asunto. Forma parte de las estructuras que sirven al relacionista para codificar relaciones entre objetos materiales o al substantivalista para

<sup>9</sup>Igualmente, es posible pensar que el campo gravitacional funciona como un campo físico independiente. Aunque por razones estructurales de fondo (localidad) la dificultad para incorporarlo a la teoría motivó la gestación de la TGR

describir al espacio o al espaciotiempo. En este grupo también cabe la estructura diferencial y topológica de la variedad o inclusive toda la estructura inercial. Utilizar *per se* el papel geométrico, o métrico, o cronogeométrico, del campo gravitacional de la TGR como un argumento decisivo en favor de la interpretación substantivalista se cae, desde esta perspectiva, por su propio peso.

La fragilidad de la interpretación substantivalista queda en evidencia. Por su parte el relacionalista tiene razones importantes para considerar que el campo métrico es un campo material (energía, momento, campo gravitacional). Parece que la interpretación substantivalista se hubiera olvidado del asunto de fondo: la naturaleza *independiente* de las estructuras espaciotemporales que cumplen un cierto papel explicativo en la dinámica. Para ilustrarlo, propongo el siguiente ejercicio: Imaginemos que N y L discuten sobre si el rojo es una propiedad del espacio o el resultado de relaciones materiales. N y L logran construir buenas teorías en las cuales, por una parte, el rojo aparece como una cualidad del espacio y, por otra, como una entidad relacional (i.e. dependiente de la materia). N y L están de acuerdo sobre las propiedades esenciales de la materia y su papel dinámico y, en consecuencia, entienden el campo gravitacional como un objeto material igual que los demás campos. Por alguna razón la teoría de L se pierde durante 200 años. Nos acostumbramos a pensar, con N, que el rojo es una propiedad del espacio; de hecho que caracteriza un espacio independiente de la materia. Pero la teoría de N es mejorada por E. En la teoría de E, el campo gravitacional, un objeto dinámico como todos los demás campos físicos, también resulta siendo rojo. Los seguidores de N, la gran mayoría ya que L cayó en el olvido, se encontrarán ante un dilema espeluznante. Interpretan al campo gravitacional como un espaciotiempo real en toda regla únicamente porque es rojo o lo interpretan como un campo material por las razones físicas conocidas. Parece una elección arbitraria.

De todas formas la mayoría se decanta por el rojo porque al parecer este ha calado como la cualidad esencial para definir un espacio independiente de los objetos materiales. Pero su estatus de cualidad esencial (del espacio o la materia) era precisamente lo que discutían N y L desde el principio. Si rescatamos la teoría de L y la ponemos a la par con la de N, el asunto cambia. En este caso si los seguidores de N no ignoran a la teoría de L, concluirán que el rojo en realidad es neutro en este asunto, que no es una razón convincente para dejar de

considerar al campo gravitacional como un campo físico de materia al igual que los demás campos. Ya sabíamos con L, desde el principio, que la materia podía codificar al rojo dentro de sus cualidades estructurales. El rojo *per se* no prueba nada contra la materia. Éste era el objeto de discusión. No puede funcionar en sí mismo como una justificación valedera. La cuestión aparecerá, en cambio, ante los ojos de los seguidores de L como un asunto resuelto. L tiene buenas razones para entender el campo gravitacional como un campo de materia; son las mismas razones que compartía con N desde el principio (energía, momento, estructura de campo físico, etc...) y, superada la teoría de N, el rojo ha dejado de lucir como una cualidad intrínseca de un espacio físico real independiente de la presencia de materia. El relacionismo no afirma que el espaciotiempo no exista, tan sólo supedita su estatus ontológico a las propiedades estructurales de la materia. No hay espaciotiempo sin materia según esta doctrina ontológica. Con este ejemplo inocente sólo quiero enfatizar que una vez que estructuras espaciotemporales como la inercia y la geometría también puedan recibir un tratamiento relacional, parecerá mucho más natural considerar un campo que integra dichas estructuras como un campo físico ya que dichas estructuras no pasarían a definir necesaria o intrínsecamente a un espaciotiempo independiente.

El relacionismo ha sido proscrito por la evidencia histórica de que ni **Leibniz**, ni **Mach** construyeron una dinámica relacional capaz de rivalizar con la dinámica de Newton. Pero aún así, superada la dinámica newtoniana, la doctrina merece una oportunidad. Que se haya descartado en el contexto de la TGR por razones que en el fondo están dictaminadas por la contingencia histórica resulta ahora entendible pero no justificable.

He argumentado que, en buena medida, la interpretación substantivalista (SS) de la TGR debe su popularidad a la carga atávica de la dinámica de Newton. Con el fin de dar una oportunidad equitativa al relacionismo y limpiar parte de la inercia histórica de toda esta discusión, propuse el ejercicio de imaginar que Leibniz, o un seguidor suyo o cualquier Leibniziano, tenía una teoría relacional buena, a la par con la de Newton. Pues bien, este no tiene que ser tan sólo un ejercicio de pura imaginación ya que, a mi juicio las ideas de **Leibniz** y **Mach**, en el contexto de la dinámica clásica, han encontrado buena forma en la teoría relacional desarrollada inicialmente por **Barbour** y **Bertotti** (BB2,1982)<sup>10</sup>. En consecuencia, la condena histórica del relacionismo

<sup>10</sup>La teoría de Barbour se conoce como la dinámica intrínseca. Los elementos principales de la formulación intrínseca de la dinámica son las configuraciones materiales instantáneas *tres-dimensionales* y la *acción* que permite escoger -vía principio de Jacobi- la historia

puede ser subsanada. No es este el lugar para mostrar que **Barbour** y **Bertotti**, efectivamente, han logrado construir una dinámica clásica genuinamente relacional y que además ésta es mejor, por ser más explicativa, que la dinámica convencional substantivalista newtoniana (para esto véase directamente **Barbour** y **Bertotti** 1982 (BB2) o alternativamente **Brown** y **Pooley** 2001). De todas formas, de ser así creo que, como he intentado argumentar, la interpretación de la TGR y en particular el papel desempeñado por el campo gravitacional reciben su lectura más natural en la línea relacional.

Mi impresión es que en este caso, toda vez que sabemos que las estructuras espaciotemporales -como la geometría y la estructura inercial- reciben una interpretación relacional consistente, la obligación explicativa se invierte en el debate R-L en la TGR. No respondería al relacionista mostrar por qué el campo gravitacional no funciona como un espaciotiempo independiente ya que éste cumpliría funciones, o tendría propiedades, representativas de estructuras genuinamente relacionales; así que la deuda explicativa quedaría pendiente, más bien, para el costado substantivalista que necesitaría argumentos más fuertes para elevar el campo gravitacional a la categoría de entidad espaciotemporal independiente y mostrar por qué no puede ser entendido como una entidad relacional material.

En la TGR las variables físicas, bien sea la ubicación de una partícula material, la intensidad de un campo en algún lugar, o la trayectoria de un rayo de luz, tienen que especificarse relacionamente asumiendo la EL, es decir, tomando los invariantes físicos de la clase equivalente  $\{g\}$ . En los tiempos de Leibniz la identidad de los indiscernibles (PII), precursora natural de la EL, aparecía como un principio relacional fuera de toda duda. Similarmente, la geometría aparece en la física de Newton como una propiedad substantivalista difícil de objetar. He admitido que la EL *per se* es neutra en el debate, e insisto, la geometría y las estructuras cronogeométricas *per se* también lo son. Lo que está en juego es su naturaleza independiente.

De la misma forma en que los substantivalistas han respondido al argumento del agujero reconociendo hábilmente que la EL no es una razón fuerte contra el substantivalismo, debían reconocer que por sí solas las estructuras espaciotemporales de la métrica tampoco pueden decantar en forma alguna el debate hacia su

costado. Éstas (y su naturaleza independiente) son el objeto de discusión. No pueden funcionar como un argumento decisivo. Mientras el relacionismo tiene buenas razones para juzgar a la métrica como un campo de materia, el substantivalismo sofisticado, a mi juicio, no tiene argumentos importantes para rechazar la naturaleza material del campo gravitacional. Este fue bien entendido como un campo físico antes del argumento del agujero. Esto debió tomarse como una buena nota de advertencia, ya que por entonces a casi nadie pareció incomodarle su estructura espaciotemporal para discriminar su naturaleza material. Si estamos preparados para vencer la inercia histórica y dar una oportunidad a **Leibniz** y **Mach** en todo esto, veremos que la TGR recibe su interpretación más natural como una teoría relacional limpia.

#### 4. La naturaleza dinámica de la materia

A continuación quiero argumentar en favor del papel dinámico de la materia. Como veremos, este aspecto importante suele ser ignorado por el substantivalismo sofisticado en su interpretación del papel de la métrica en la TGR. Las variables dinámicas de nuestras teorías están ligadas a propiedades estructurales de la materia. Esta es notablemente una característica distintiva de nuestras teorías físicas. Sin embargo, este consenso es roto por la interpretación substantivalista (SS) de la TGR.

Recordemos que un modelo de la TGR  $\langle M, g, T \rangle$  está dado por la variedad  $M$ , el campo métrico  $g$ , que funciona como un objeto dinámico que debe representar las propiedades del espaciotiempo, y  $T$  que, naturalmente, es también un objeto dinámico que representa los contenidos materiales del mismo. Hemos visto que no hay nada contradictorio en hablar de propiedades espaciotemporales relacionadas con la materia. El substantivalista quiere proscribir cualquier estructura espaciotemporal de la materia.

En este sentido el substantivalista sofisticado quiere que en su teoría los objetos materiales aparezcan separados, limpiamente, de las estructuras espaciotemporales.

Supongamos que puede hacerse tal tipo de distinción limpia entre objetos espaciotemporales  $S_i$  y objetos materiales  $m_i$ . En este caso los modelos  $\aleph$  de una buena teoría substantivalista estarán representados de la siguiente forma  $\aleph = \langle M, S_1, S_2, \dots, m_1, m_2, \dots \rangle$ .

conformada por la sucesión de tales configuraciones materiales relativas prescindiendo *ab initio* de los sistemas inerciales de la dinámica clásica en la versión newtoniana. El enfoque desarrollado por Barbour y Bertotti puede hallarse en forma germinal en **Zanstra** 1928.

Afortunadamente, como es bien sabido, esta suposición funciona bien para modelos newtonianos, neo-newtonianos, y tipo-Minkowski (TER). La separación entre materia y espacio-tiempo está bien definida en estos casos. Funciona bien si tomamos cualquier ensamble de campos físicos (electromagnéticos, gravitacionales, fluidos de polvo, etc.) inmersos en los respectivos espacios newtoniano, neo-newtoniano o tipo-Minkowski (TER). En estos casos un modelo de la teoría puede representarse así:  $\aleph = \langle M, h, \Gamma, \phi, \rho, E, \dots \rangle$ <sup>11</sup>. Donde  $M$  es la variedad,  $h$  es la estructura métrica,  $\Gamma$  es la conexión afín o estructura inercial,  $\phi$  es el campo gravitacional,  $\rho$  es una densidad de fluido que puede representar la distribución de materia ponderable convencional,  $E$  es el campo electromagnético y los puntos suspensivos corresponden a cualquier otra forma de materia que queramos incluir en el modelo.

En todos estos modelos pre-TGR,  $h$  y  $\Gamma$  son entendidos como propiedades del espacio-tiempo, mientras que  $\phi, \rho, E$  y el resto son objetos materiales. Todo el mundo desearía que las cosas fueran tan limpias cuando pasamos a los modelos de la TGR y que una simple traducción directa de estas estructuras y campos bastara para la identificación, pero allí el tensor  $g$  integra simultáneamente los roles de  $\phi$  -un campo físico de materia- y de  $h$  y  $\Gamma$  -estructuras espaciotemporales-. Así que necesitamos un criterio más fuerte que la simple identificación de estructuras. Aparte del transporte de energía-momento, común a todos los campos materiales, un criterio que encuentre concluyente viene dado por la naturaleza dinámica de estos objetos.

En los modelos pre-TGR las estructuras espaciotemporales ( $h, \Gamma$ ) son entidades absolutas que describen las propiedades del espaciotiempo. Por su parte, toda la materia está representada por objetos dinámicos ( $\phi, \rho, E, \dots$ )<sup>12</sup>. Es pertinente aclarar que un objeto absoluto -en el sentido de la geometría diferencial- es un objeto que retiene su estructura en todos los modelos de la teoría afectando la evolución dinámica de la materia sin ser recíprocamente afectado. En general los objetos absolutos son caracterizados por su grupo de simetría y el efecto de las correspondientes transformaciones de simetría es producir movimientos o arrastre de los campos dinámicos mientras el fondo absoluto permanece invariante<sup>13</sup>.

En todos los modelos (pre-TGR) del tipo  $\aleph = \langle M, S_1, S_2, \dots, m_1, m_2, \dots \rangle$ , los  $S_i$  son objetos absolutos, mientras que los  $m_i$  son objetos dinámicos que cambian en correspondencia con sus respectivas ecuaciones de campo<sup>14</sup>.

En los modelos  $\langle M, g, T \rangle$  de la TGR,  $g$  y  $T$  son objetos dinámicos. Encuentro algo obstinado desconocer la naturaleza material de los objetos dinámicos, común a los modelos de todas nuestras teorías, simplemente para salvar una doctrina ontológica. Sería algo obstinado desconocer la naturaleza material del campo gravitacional  $g$ . El campo gravitacional es un objeto dinámico material como todos los demás campos de fuerza. La TGR recibe, por tanto, su interpretación natural como una teoría relacional en la que las estructuras espaciotemporales son absorbidas por un campo material  $g$ .

<sup>11</sup>Es claro que las estructuras espaciotemporales cambian detalles formales en el paso del espacio newtoniano, a los espaciotiempos neo-newtoniano y tipo-Minkowski de la relatividad especial, pero en todos los casos se está de acuerdo sobre cuáles son esencialmente las estructuras espaciotemporales. Éstas son, básicamente, estructura métrica e inercial. De todas formas no sobra recordar que incluso en el contexto de la dinámica newtoniana es posible integrar el potencial gravitacional en la conexión afín  $\Gamma$  -como curvatura variable- pero en este caso la conexión afín se dinamiza de la mano del campo gravitacional y parece más natural considerarla entonces como un campo de materia. A este respecto véase especialmente **Friedman** (1983), Cap 4: *Gravitation Theory: Curved Space-Time*, pp 108-114.

<sup>12</sup>Nótese que la distinción también aplica a cualquier modelo mecanocuántico ya que en estos casos los campos físicos se definen sobre fondos espaciotemporales absolutos. Incluso en los casos en que se intenta formular modelos mecanocuánticos sobre espaciotiempos curvos, éstos espaciotiempos son representados como estructuras fijas -no dinámicas-.

<sup>13</sup>La distinción entre los objetos absolutos y los objetos dinámicos de las teorías espaciotemporales es originaria de **Anderson** (1967 p. 83-84) y es bien tratada por **Friedman** (1984 p. 47-60)

<sup>14</sup>En general las ecuaciones de campo tienen la estructura  $D\varphi = \sigma$  donde el lado izquierdo de la ecuación representa la estructura diferencial del campo (i.e. un operador  $D$  actuando sobre el campo  $\varphi$ ) y el lado derecho ( $\sigma$ ) representa las fuentes de éste. En todos los casos ambos lados de la ecuación -el campo y las fuentes- son entendidos como objetos materiales ya que los campos son portadores de energía. Esta equivalencia material izquierda-derecha es rota por la interpretación substantivalista (SS) de la TGR.

<sup>15</sup>Nótese que esta distinción entre objetos absolutos que caracterizan la estructura espaciotemporal -el fondo fijo- y objetos dinámicos -que representan la materia y su evolución- se aplica únicamente a los objetos (campos) que se cuelgan sobre la variedad. La variedad proporciona la estructura topológica y diferencial requerida por estos objetos. Pero no es en sí misma cobijada por la distinción. Su falta de estructura no permite entenderla ni como un espaciotiempo de fondo, en el caso de teorías con variedad absoluta (fija), ni como un objeto material -no es portadora de ningún tipo de interacción, ni tiene la estructura de un campo- en caso de teorías con variedad dinámica (TGR).

La TGR es una teoría completamente dinámica. Este es quizá su aspecto más revolucionario. Todos sus modelos son de la forma  $\mathcal{N} = (M, m_1, m_2, \dots)^{15}$ .

En todo caso la interpretación del campo gravitacional como un campo físico de fuerzas, descansa, en alguna medida en la viabilidad para entenderlo como una entidad portadora de energía-momento, al igual que todos los demás campos. Y aunque casi nadie duda que la energía del campo gravitacional se intercambie libremente con la energía de los demás campos -de hecho se supone que es la fuente de las grandes cantidades de energía liberada en la forma de radiación y calor en la descripción del colapso estelar-, en la TGR, como consecuencia del principio de equivalencia, no se tiene una buena definición local de la densidad de energía del campo gravitacional. Es una asignatura pendiente de resolver en la TGR, estrechamente vinculada con los problemas para cuantizar el campo gravitacional. De todas formas, que la energía no sea 'localizable' en regiones infinitesimales no significa que no exista, al menos, una forma aproximada para hacer estimaciones físicas empíricamente contrastables. Lo que se tiene es una definición cuasi-local (i.e. localizable en regiones finitas del espaciotiempo) de la energía del campo gravitacional para una clase restringida de modelos de la TGR (Szabados 2004) o un pseudotensor que representa la densidad de energía en ciertos sistemas de coordenadas<sup>16</sup>.

Aunque los pseudotensores no den una buena definición local de la energía del campo gravitacional sus integrales son útiles para calcular la energía gravitatoria de sistemas relativamente aislados. En la práctica se han utilizado para calcular la radiación gravitacional generada por sistemas autogravitantes. El primero en hacerlo fue **Einstein**. En su momento se llegó a discutir si las ondas gravitacionales eran reales o meros artificios de los sistemas de coordenadas pero **Bondi** y otros (1957, 1963) mostraron que se trataba de efectos intrínsecos y el análisis específico de ondas cortas de **Isaacson**

(1968) pareció confirmarlo. Las predicciones sobre la energía radiada gravitacional han sido empíricamente contrastadas al menos para el pulsar binario PSR 1913+16 descubierto por **Hulse** y **Taylor** en 1974. Y a pesar de los problemas para tener una definición (local) de la energía gravitacional y su conservación, éstas consideraciones inclinan hoy a la inmensa mayoría a aceptar la realidad de la energía de las ondas gravitacionales y en general de la energía del campo gravitacional en la TGR. Estoy de acuerdo.

## 5. Algo más sobre la variedad: ¿Hay espaciotiempo?

Debido a las tensiones que siguen existiendo al interior del debate R-S, con respecto al rol desempeñado por la variedad quisiera despejar cualquier duda sobre el estatus ontológico de la misma. Ésta dejó de ser considerada como un espaciotiempo que alojaba campos físicos ya que, hemos visto, la falta de identidad primitiva de los puntos que la componen y la carencia de estructuras que le permitan alcanzar algún rol explicativo en los fenómenos físicos son razones fuertes para dejar de considerarla como el contenedor espaciotemporal de todos los fenómenos físicos.

La variedad proporciona la estructura diferencial y topológica para levantar cualquier teoría de campos. Es incluso menos estructurada que la posible métrica utilizada por algún Leibniziano. Algunos substantialistas han equivocado este asunto al objetar el uso de la variedad y otras estructuras *espaciotemporales* en la concepción relacional de una teoría física<sup>17</sup>. El asunto importante en realidad pasa por determinar si el espaciotiempo físico puede -o no- funcionar como una entidad independiente de la presencia de materia conservando su poder explicativo, soportando la dinámica de la teoría.

En la TGR los puntos de la variedad no tienen identidad, ni siquiera la estructura topológica está determinada en forma independiente por la colección de puntos. Éstos sólo pueden elevarse a la categoría de posibles

<sup>16</sup>En Cala 2008 se presentan en detalle las dificultades para localizar la energía del campo gravitacional y para formular una correcta ley de la conservación de energía-momento en la TGR. Allí se concluye que estas dificultades están directamente relacionadas con la estructura dinámica-relacional de la teoría. Esto porque, en resumen, la Teoría General de la Relatividad carece de un fondo espaciotemporal independiente y, como consecuencia de ello la TGR carece también de objetos absolutos cuyas simetrías globales permitirían una formulación adecuada de la ley de conservación de energía-momento.

<sup>17</sup>**Earman**, por ejemplo, no ha dejado de manifestar su malestar porque, en principio, la teoría de campos requiere en el fondo algún tipo de sustanciación ya que éstos especifican sus cantidades en función de las coordenadas adaptadas a un fondo espacial sustancial al estilo de la variedad de la TGR. **Earman** no duda que el campo gravitacional haga parte de los contenidos materiales del espaciotiempo, pero el uso de este tipo de puntos (no individuados) de la variedad le llevó a sugerir que el debate debía dirimirse fuera de las formas tradicionales de entender el substantialismo y el relacionismo (**Earman** 1989, p.208).

eventos espaciotemporales cuando un tensor métrico específico -un campo material- es definido sobre la variedad. Incluso la estructura topológica global requiere la especificación del tensor métrico. En cierto sentido, la variedad es también una estructura dinámica ya que ésta cambia de modelo a modelo de la teoría de la mano del campo gravitacional<sup>18</sup>. Su estructura dinámica es heredada del campo gravitacional vía ecuaciones de campo (**Stachel** 1994, **Isham** 1994, **Sorkin** 1997). De modo que al remover el campo gravitacional no puede hablarse, en sentido alguno, del significado físico o espaciotemporal de la variedad.

El espacio (y el tiempo) newtoniano, por el contrario, puede pensarse sin la presencia de ningún objeto material. Removida la materia éste conserva toda la riqueza de su estructura espaciotemporal. Este no es el caso en la TGR. Ni siquiera la variedad -insisto- retiene algún tipo de significado espaciotemporal sin la presencia del campo gravitacional. Todo es materia en la TGR. Al remover la materia desaparece con ella el espaciotiempo. No existe, ni siquiera como hipótesis teórica, la posibilidad de definir un espaciotiempo vacío. La Teoría General de la Relatividad es, en definitiva, una teoría relacional sobre el espaciotiempo. Esta es mi conclusión general, como también esta fue la conclusión madura de **Einstein**.

En la edición de 1954 de su *Relativity, The Special and the General Theory* (la última antes de su muerte), **Einstein** incluyó un apéndice especial. Éste se titulaba *Relativity and the Problem of Space*. Para cerrar, le cito en cierta extensión porque considero que este no deja dudas al respecto (**Einstein** 1954 Appendix V, P.155):

*We are now in a position to see how far the transition to the general theory of relativity modifies the concept of space. In accordance with classical mechanics and according to the special theory of relativity, space (space-time) has an existence independent of matter or field. [...] On the basis of the general theory of relativity, on the other hand, space as opposed to 'what fills space', which is dependent on the co-ordinates, has no separate existence. Thus a pure*

*gravitational field might have been described in terms of the  $g_{ik}$  (as functions of the co-ordinates), by solution of the gravitational equations. If we imagine the gravitational field, i.e. the functions  $g_{ik}$ , to be removed, there does not remain a space of type (I)<sup>19</sup>, but absolutely nothing, and also no 'topological space'. For the functions  $g_{ik}$  describe not only the field but, at the same time also the topological and metrical structural properties of the manifold. A space of type (I), judged from the standpoint of the general theory of relativity, is not a space without field, but a special case of the  $g_{ik}$  field, for which -for the co-ordinate system used, which itself has no objective significance- the functions  $g_{ik}$  have values that do not depend on the co-ordinates. There is no such thing as empty space, i.e. a space without a field. Space-time does not claim existence on its own, but only as a structural quality of the field<sup>20</sup>.*

## 6. Conclusión

A pesar de que circula la versión según la cuál la TGR instancia la existencia de un espacio-tiempo físico independiente de la presencia de objetos materiales o de la ocurrencia de fenómenos, concluyo que resulta más conveniente e inteligible la interpretación relacional del espaciotiempo en la TGR. Los puntos de la variedad espaciotemporal no tienen identidad primitiva y ello obliga a trasladar la discusión sobre la existencia física del espaciotiempo al poder explicativo de las estructuras espaciotemporales que se cuelgan sobre esta entidad, es decir, al rol desempeñado por el tensor métrico y los campos materiales (portadores de energía-momento) en la dinámica relativista. Y aunque el tensor métrico codifica, por así decirlo, la estructura cronogeométrica típicamente adjudicada a un espacio substantivalista al estilo del espacio newtoniano, puede mostrarse que estas estructuras cronogeométricas pueden recibir un tratamiento genuinamente relacional (**Barbour** y

<sup>18</sup>La variedad no tiene forma local intrínseca. Ésta es heredada de los tensores definidos sobre ella y éstos a su vez la cargan de estructura.

<sup>19</sup>Aquí por (I), Einstein se está refiriendo al espaciotiempo de Minkowski. Éste hubiera sido inaceptable en la época en que el *problema del espacio* se había convertido para Einstein en el problema de satisfacer el principio de Mach. Pero, incorporada consistentemente la noción de campo a la de objeto material, no existe, ni siquiera teóricamente, la posibilidad de un espacio vacío sin la presencia de materia. Difícilmente puede uno imaginarse algo más relacional.

<sup>20</sup>El énfasis es mío.

Bertotti 1982, Pooley y Brown 2001), con lo cual se necesitan razones más fuertes para elevar la métrica al estatus ontológico de un espaciotiempo independiente. Los campos materiales -y la materia en general- son portadores de energía-momento y son representados en *todas* las teorías espaciotemporales por objetos dinámicos en el sentido de la geometría diferencial (Anderson 1967, Friedman 1983). No existe, por tanto, una razón fuerte para suponer que el tensor métrico, objeto dinámico portador de energía-momento, no caiga en esta clasificación. Se trata de un objeto material en toda regla. Las estructuras cronogeométricas son subsidiarias de un objeto material extendido: el campo métrico o gravitacional. Ello obliga a concluir, con Einstein, que si se remueve la materia también se remueve el espaciotiempo; la tesis relacional.

**Agradecimientos.** Deseo agradecer a un anónimo revisor científico quien contribuyó con pertinentes observaciones a mejorar la presentación y el contenido de este artículo.

#### Referencias

- [1] Alexander, H. G. (ed) (1984) *The Leibniz Clarke Correspondence (1717)*. Barnes and Noble. New York.
- [2] Anderson, J. L. (1967) *Principles of Relativity Theory*. Academic Press. New York
- [3] Ashtekar, A. 'Geometric Issues in Quantum Gravity', en *The Geometric Universe: Science Geometry and The Work of Roger Penrose*
- [4] Baez, J. C. (2001) 'Higher-dimensional algebra and Planck scale physics', en *Physics Meets Philosophy at the Planck Scale*, C. Callender y N. Hugget (eds.) Cambridge University Press. Cambridge, pp. 177-198.
- [5] Barbour, J. B. (1982) 'Relational Concepts of Space and Time' *British Journal for the Philosophy of Science*, **33**, pp 251-274.
- [6] Barbour, J. B. (1994) 'The timelessness of quantum gravity: I The evidence from the classical theory, II The appearance of dynamics in static configurations'. *Class. and Quant. Grav* **11**, pp. 2853.
- [7] Barbour, J. B. & Bertotti, B. (1982) 'Mach's principle and the Structure of Dynamical Theories' *Proceedings of the Royal Society*, **382**, pp 295-306.
- [8] Barbour, J. B., Foster, B.Z., & Murchadha, N. (2002) 'Relativity without Relativity' *Class. Quantum Grav.* **19**, pp. 3217. Disponible en *gr-qc/0012089*.
- [9] Bartels, A. (1996) 'Modern essentialism and the problem of individuation of spacetime points', *Erkenntnis* **45**, pp 25-43.
- [10] Belot, G. (1998) 'Why General Relativity Does Need an Interpretation' *Philosophy of Science (Proceedings)*, **63**, pp. 80-88.
- [11] Belot, G. & Earman, J. (2001) 'Pre-Socratic Quantum Gravity', en *Physics Meets Philosophy at the Planck Scale*, C. Callender y N. Hugget (eds.) Cambridge University Press. Cambridge, pp. 213-255.
- [12] Brighouse, C. (1994) 'Spacetime and Holes', en *PSA 1994, Vol 1*, D. Hull, M. Forbes, y R. Burian (eds.). East Lansing Michigan. Michigan, pp. 117-125.
- [13] Bondi, H. (1957) *Nature*, **179**, p. 1072.
- [14] Bondi, H., J. van der Burg, M. G. & Metzner, A. W. K. (1962) *Proceedings for The Royal Society*, **A 269**, p.21-51.
- [15] Butterfield, J. & Isham, C. J. (1999) 'On the Emergence of Time in Quantum Gravity', en *The Arguments of Time*, J. Butterfeld (ed.). Oxford University Press, Oxford. pp. 111-168.
- [16] Cala, F. E. (2006) 'La Identidad de las Partes del Espacio y el Problema de la Inercia' *Praxis Filosófica*, **22**, pp. 153-169.
- [17] Cala, F. E., Gómez, A. R. & Ramos, J. F. (2007) 'De la Relatividad de la Inercia al Universo Cerrado' *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, **31**(118), pp. 97-108.
- [18] Cala, F.E. (2008) 'Sobre la Dinámica Relacional del Espacio-Tiempo y la Conservación de la Energía en la Teoría General de la Relatividad' *TEORÍA: Revista de Teoría, Fundamentos e Historia de la Ciencia*, en prensa.
- [19] Callender, C. & Hugget, N. (eds.) (2001) *Physics Meets Philosophy at the Planck Scale*, Cambridge University Press. Cambridge.
- [20] DiSalle, R. (1994) 'On dynamics, indiscernibility, and spacetime ontology', *British Journal for the Philosophy of Science*, **45**, pp. 265-287.
- [21] Earman, J. (1970) 'Who's Afraid of Absolute Space', *Australian Journal of Philosophy*, **48**, pp 287-317.
- [22] Earman, J. (1986) *A Primer on Determinism*. D. Reidel. Dordrecht, The Netherlands.
- [23] Earman, J. & Norton, J. (1987) 'What Price Space-Time Substantivalism? The Hole Story', *British Journal for The Philosophy of Science*, **38**, pp 515-525.
- [24] Earman, J. (1989) *World Enough and Space-Time*. MIT Press. Cambridge, Massachusetts.
- [25] Einstein, A. (1914) 'Prinzipielles zur verallgemeinerten Relativitätstheorie', *Physikalische Zeitschrift*, **15**, pp 176-180.
- [26] Einstein, A. (1918) 'Prinzipielles zur allgemeinen Relativitätstheorie', *Annalen der Physik*, **55**, pp 241-244.
- [27] Einstein, A. (1954) 'Relativity and the Problem of Space', Appendix 5, *Relativity: The Special and General Theory*. 15 edición. Methuen.
- [28] Friedman, M. (1983) *Foundations of Space-Time Theories*. Princeton University Press. Princeton.

- [29] **Hawking, S. W. & Ellis, G. F. R.** (1973) *The Large Scale Structure of Space-Time*. Cambridge University Press. Cambridge.
- [30] **Hofer, C.** (1994) 'Einstein's Struggle for a Machian Gravitation Theory', *Studies for the History and Philosophy of Science*, Vol. 25, 3, pp 287-335.
- [31] **Hofer, C.** (1996) 'The Metaphysics of Space-Time Substantivalism', *The Journal of Philosophy*, 93, pp. 5-27.
- [32] **Hofer, C.** (1998) 'Absolute Versus Relational Spacetime: For Better or Worse, the Debate Goes On', *British Journal for the Philosophy of Science*, 49, pp. 451-467.
- [33] **Isaacson, R. A.** (1968) *Physical Review*, 166, p.1263-1272.
- [34] **Isham, C.** (1994) 'Prima Facie Questions in Quantum Gravity', en *Canonical Quantum Gravity: From Classical to Quantum (Lecture Notes in Physics 434)*. J Ehlers y H. Friedrich (eds.). Springer-Verlag. Berlín, pp.1-21.
- [35] **Kretschmann, E.** (1917) *Annalen der Physik* 53, pp. 575-614.
- [36] **Laymon, R.** (1978) 'Newton's Bucket Experiment', *Journal of the History of Philosophy*, 16, pp 399-413.
- [37] **Mach, E.** (1883) *The Science of Mechanics*. Trad. de T.J Mc Cormack. Open Court. La Salle, Illinois. 1960.
- [38] **Maudlin, T.** (1993) 'Buckets of Water and Waves of Space: Why Spacetime is Probably a Substance', *Philosophy of Science*, 60, pp. 183-203.
- [39] **Mundy, R.** (1992) 'Space-Time and Isomorphism', en *PSA 1992, Vol 1*, D. Hull, M. Forbes, y K. Okruhlik (eds.). East Lansing Michigan. Michigan ,pp. 515-527.
- [40] **Nagel, E.** (1961) *The Structure of Science*. Harcourt, Brace and World. New York.
- [41] **Pooley, O. & Brown, H. R.** (2001) 'Relationism Rehabilitated? I: Classic Mechanics'. *British Journal for the Philosophy of Science*, 53, pp 183-204. E Print en <http://philsci-archive.pitt.edu/archive/00000220>.
- [42] **Pooley, O.** (2002) 'Relationism Rehabilitated? II: Relativity'. E Print en <http://philsci-archive.pitt.edu/archive/00000221>.
- [43] **Reichenbach, H.** (1957) *Space and Time*. Dover. New York.
- [44] **Rovelli, C.** (1996) 'Relational Quantum Mechanics', *International Journal of the Theoretical Physics*, 35, pp 1637.
- [45] **Rovelli, C.** (1997) 'Halfway Through the Woods: Contemporary Research on Space and Time', en *The Cosmos of Science: Essays of Exploration*, J. Earman y J. Norton (eds.). University of Pittsburgh Press. Pittsburgh, pp 180-223.
- [46] **Rynasiewicz, R.** (1996) 'Absolute Versus Relational Space-Time: An Otomodded Debate?' , *Journal of Philosophy*, 93, pp.279-306.
- [47] **Saunders, S.** (2002) 'Indiscernibles, General Covariance, and Other Symmetries: The Case for Non-Reductive Relationism', en *Revisiting the Foundations of Relativistic Physics*, J. Renn (ed.). Kluwer Press.
- [48] **Saunders, S.** (2003) 'Physics and Leibniz Principles', en *Symmetries in Physics: Philosophical Reflections*, K. Brading and E. Castellani (eds.), Cambridge University Press.
- [49] **Sklar, L.** (1976) *Space, Time, and Space-Time*. University of California Press. Berkeley.
- [50] **Smolin, L.** (1991) 'Space and Time in the Quantum Universe', en *Conceptual Problems of Quantum Gravity*, A. Ashtekar, y J. Stachel (eds.). Birkhäuser. Boston, pp 228-291.
- [51] **Sorkin, R. D.** (1997) 'Forks in the Road, on the Way to Quantum Gravity', *International Journal of Theoretical Physics*, 36, pp. 2759-81.
- [52] **Stachel, J.** (1993) 'The Meaning of General Covariance', en *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, J. Earman, A. Janis, G. Massey, y N. Rescher, (eds.), Pittsburgh University Press. Pittsburgh, pp. 129-160.
- [53] **Stachel, J.** (1994) 'Changes in the Concepts of Space and Time Brought About by Relativity', en *Artifacts, Representation and Social Practice*, C.C. Gould y R.S. Cohen (eds.). Kluwer Academic. Dordrecht, pp. 141-162.
- [54] **Stein, H.** (1970), 'Newtonian space-time', en *The Annus Mirabilis of Sir Isaac Newton 1666-1966*. The M. I.T. press. Cambridge, Mass., pp. 258-284.
- [55] **Szabados, L. B.** (2004) 'Quasi-Local Energy-Momentum and Angular Momentum in GR: A Review Article.' , *Living Rev. Relativity*, 7, 4.
- [56] **Teller, P.** (1991) 'Substances, Relations and Arguments About the Nature of Spacetime', *The Philosophical Review*, C3, pp. 363-97.
- [57] **Wald, R. M.** (1984) *General Relativity*. University of Chicago Press. Chicago.
- [58] **Zanstra, H.** (1928) 'A study of relative motion in connection with classical mechanics', *Physical Review*, 23, nn 528-545.

Recibido: noviembre 8 de 2007

Aceptado para publicación: diciembre 3 de 2008

# ON THE BEHAVIOUR OF THE $\text{Ca-SO}_4\text{-H}_2\text{O}$ SYSTEM

By

Iván Berdugo<sup>1</sup>, Enrique Romero<sup>1</sup>, Maarten Saaltink<sup>1</sup> & María Albis<sup>2</sup>

## Abstract

**Berdugo, I., E. Romero, M. Saaltink & M. Albis:** On the behaviour of the  $\text{Ca-SO}_4\text{-H}_2\text{O}$  system. *Rev. Acad. Colomb. Cienc.* **32**(125): 545-557, 2008. ISSN 0370-3908.

Thermotropia, barotropia and hygotropia of the  $\text{Ca-SO}_4\text{-H}_2\text{O}$  system is reviewed and discussed and a three-dimensional representation of the phase diagram for system at atmospheric pressure is derived. It is concluded that the behaviour of the system is a coupled thermo-hydro-chemical phenomenon in which water activity is the key factor for existence of high concentration of  $\text{Ca}^{2+}$  and  $\text{SO}_4^{2-}$  and confinement pressure plays a secondary role.

**Key words:** anhydrite, gypsum, water, solubility, temperature, pressure, vapour, relative humidity.

## Resumen

Se revisan y discuten la termotropía, barotropía e higriscopía del sistema  $\text{Ca-SO}_4\text{-H}_2\text{O}$  y se deriva una representación tridimensional del diagrama de fases del sistema a presión atmosférica. Se concluye que el comportamiento del sistema es un fenómeno termo-hidro-químico acoplado en el cual la actividad del agua es el factor determinante para la existencia de altas concentraciones de  $\text{Ca}^{2+}$  y  $\text{SO}_4^{2-}$  y la presión de confinamiento juega un papel secundario.

**Palabras clave:** anhidrita, yeso, agua, solubilidad, temperatura, presión, vapor, humedad relativa.

## Introduction

Interacting with water calcium sulphate can exist in six different solid phases, presented hierarchically here in the

order of increasing solubility: one dihydrate (gypsum,  $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ ); two hemihydrates (bassanite:  $\alpha\text{-CaSO}_4 \cdot \frac{1}{2}\text{H}_2\text{O}$  and  $\beta\text{-CaSO}_4 \cdot \frac{1}{2}\text{H}_2\text{O}$ ); and three anhydrous (anhydrite: AI-unstable- or  $\alpha\text{-CaSO}_4$ , AII –moderately soluble– or  $\beta\text{-CaSO}_4$ ,

<sup>1</sup> Departament d'Enginyeria del Terreny, Cartogràfica i Geofísica – Universitat Politècnica de Catalunya, Barcelona, Spain. Correos electrònics: ivan.berdugo@upc.edu, enrique.romero-morales@upc.edu, maarten.saaltink@upc.edu

<sup>2</sup> Programa de Ingeniería Civil – Universidad de La Salle, Bogotá, D.C. Colombia. Correo electrónico: malbis@unisalle.edu.co

and AIII-totally soluble- or  $\gamma$ -CaSO<sub>4</sub>). However, it is pointed out by **Sattler & Brückner** (2001) that in addition to the hemihydrate other “sub-hydrates” with crystal water contents between 1/2 and 4/5 have recently been produced and investigated crystallographically.

The behavior of calcium sulphate-rich water is an important aspect in several environmental, geotechnical and industrial processes involving either dissolution or crystallization of calcium sulphate-based minerals (Table 1). The doline collapse in anhydritic-gypsiferous soils and rocks is a typical case of sulphate-based mineral dissolution conducing to ground loss and instabilities of existing structures. On the other hand, solvent-way gypsum growth is typical case of crystallization that often causes –among other–, scale induced duct obstructions as well as swelling in soils, rocks and stabilized materials. In general, excluding production of building materials, calcium sulphate minerals –gypsum in particular–, often appear as undesirable crystallizations.

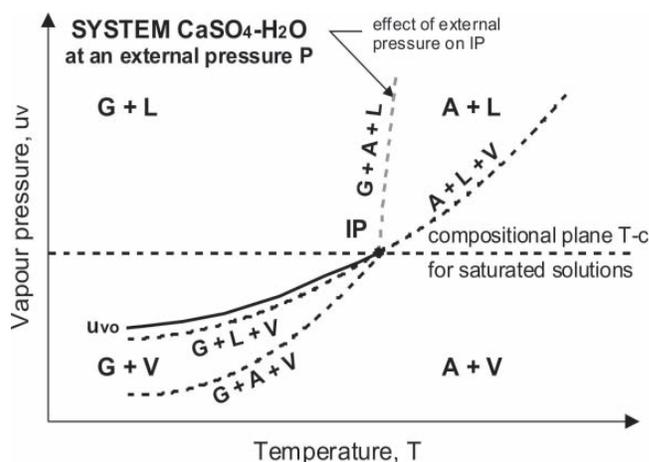
An accurate analysis of mechanisms underlying dissolution and crystallization of calcium sulphate-based minerals requires a previous knowledge on factors controlling the behaviour of calcium sulphate-rich water. This paper deals with the effects of temperature, confinement (external) pressure and water activity on the solubility and the characteristic vapour pressure of the Ca-SO<sub>4</sub>-H<sub>2</sub>O system. A comprehensive review of contributions on the theme is presented and a three-dimensional representation of the phase diagram for the Ca-SO<sub>4</sub>-H<sub>2</sub>O system at atmospheric pressure is derived.

### Univariants of the Ca-SO<sub>4</sub>-H<sub>2</sub>O system

Due to the hemihydrate is an intermediate phase between the dihydrate and the anhydrous, which is metastable at all temperatures, in standard analyses of the Ca-SO<sub>4</sub>-H<sub>2</sub>O system only the dihydrate and the anhydrous phases are considered. Solubility of gypsum (G) and anhydrite (A) depends on three basic variables: (i)

temperature, T, (ii), confinement (external or reference) pressure P, and (iii) water activity, *a<sub>w</sub>*. Then, the occurrence of stable phases of both minerals in the presence of saturated solutions (L) and vapour (V) (namely, binary states G + V, G + L, A + L and A + V) is a multi-dependent phenomenon associated with specific values of vapour pressure, *u<sub>v</sub>*.

Boundaries for these binary states constitutes the so-called univariant three-phase equilibria (G + L + V, A + L + V, G + A + L and G + A + V), which intersect at IP point representing the invariant four-phase equilibrium (G + A + L + V); where gypsum, anhydrite, saturated solution and vapour are all stable and phases exist at a unique condition temperature-solubility-vapour pressure (figure 1).



**Figure 1.** Schematic representation of the CaSO<sub>4</sub>-H<sub>2</sub>O system phase diagram on the barometric plane.

Solubility of calcium sulphate has been extensively studied during decades since the pioneer contributions by van't Hoff and co-workers at the early twenty century (**van't Hoff et al.**, 1903). Criteria on their values as a function of temperature, pressure and water activity have a long

**Table 1.** Some enviromental, geotechnical and industrial processes involving calcium sulphate-rich water.

Environmental processes	Geotechnical processes	Industrial processes
<ul style="list-style-type: none"> <li>• Marine intrusion</li> <li>• Water desalination</li> <li>• Nanofiltration of saline drainage</li> <li>• Geothermal energy generation</li> <li>• Hydrothermal energy generation</li> <li>• Flue gas desulfurization</li> <li>• Waste storage in evaporite rocks</li> </ul>	<ul style="list-style-type: none"> <li>• Doline collapse</li> <li>• Gypsum growth induced swelling</li> <li>• Ground improvement using calcium-based stabilizers</li> <li>• Sulphate attack to mortars and concrete</li> </ul>	<ul style="list-style-type: none"> <li>• Production plaster of Paris</li> <li>• Production of phosphoric acid</li> <li>• Production of hydrogen fluoride</li> <li>• Production of cooper</li> <li>• Refining of zinc</li> <li>• Secondary oil recovery</li> </ul>

history with successive changing opinions. However, a review of some recent contributions reveals that within the ranges of temperature and pressure relevant to environmental and geotechnical processes (0°C to 50°C and 0 to 10 MPa) there is a consensus regarding the characteristic values (Freyer, 2000; Kontrec *et al.*, 2002; Freyer & Voigt, 2003, 2004; Vanko & Bach, 2005). Most important experimental and theoretical contributions to the study of effects of temperature, pressure and water activity on solubility of gypsum and anhydrite are summarized and discussed below.

### Thermotropia

Univariants G + L + V and A + L + V, representing the equilibrium of gypsum and anhydrite with saturated solutions and vapour, constitute the solubility diagram -or the compositional plane temperature-concentration (T-c)-, for the dihydrate and the anhydrous phases of the CaSO<sub>4</sub>-H<sub>2</sub>O system. Figure 2 is the solubility diagram for gypsum and anhydrite (AII) in pure water within the

temperature range 10°-80°C in an ideal thermodynamically closed system at atmospheric pressure. Curves from empirical correlations between temperature and calcium sulphate molality or concentration proposed by Blount & Dickson (1973) (Eq. 1 and Eq. 2), and Innorta *et al.* (1980) (Eq. 3 and Eq. 4) are also presented. Finally, results obtained by thermodynamic calculations by Møller (1988) are also included as theoretical references.

Figure 2 indicates that under atmospheric conditions gypsum has direct solubility behaviour until 49,5°C; it is that solubility increases with increasing temperature. In excess of this temperature threshold inverse solubility behaviour holds. On the contrary, anhydrite has always inverse solubility behaviour. At a given temperature the phase with the lowest solubility represents the stable phase. Below the temperature threshold gypsum is the stable phase and at high temperatures it is anhydrite. Thus, the transition temperature (IP) in the compositional plane temperature-concentration (T-c) is required to define the right occurrence of hydrated or anhydrous phases.

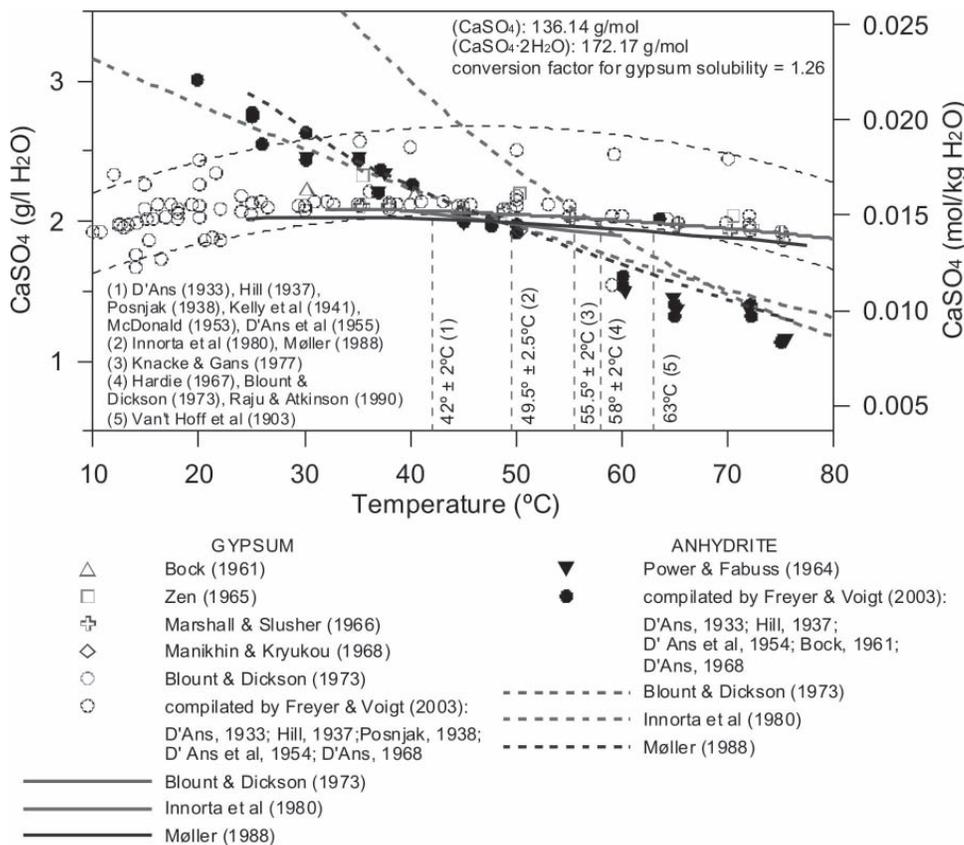


Figure 2. Solubility diagram for the CaSO<sub>4</sub>-H<sub>2</sub>O system at 0,10 MPa (1 bar).

Blount & Dickson (1973)

Gypsum:

$$\ln m(G)_{(CaSO_4)} = -4.355 + 8.4 \times 10^{-4} P + 1.05 \times 10^{-2} T - 1.7 \times 10^{-4} T^2 + 5.8417 \times 10^{-7} T^3 \quad (\text{Eq. 1})$$

Anhydrite:

$$\ln m(A)_{(CaSO_4)} = -2.87 + 1.22 \times 10^{-3} P - 2.37 \times 10^{-2} T - 2.8 \times 10^{-6} PT + 1.0 \times 10^{-6} T^2 + 1.4 \times 10^{-8} PT^2 \quad (\text{Eq. 2})$$

$$m: \left[ \frac{\text{mol}}{\text{kg } H_2O} \right]; P: [\text{bar}]; T: [^\circ\text{C}]$$

Innorta et al (1980)

Gypsum:

$$c(G)_{(CaSO_4)} = 8.69 \times 10^{-5} T^2 - 1.73 \times 10^{-2} T + 2.624 \quad (\text{Eq. 3})$$

Anhydrite:

$$c(A)_{(CaSO_4)} = 1.14 \times 10^{-4} T^2 - 3.667 \times 10^{-2} T + 3.517 \quad (\text{Eq. 4})$$

$$c(G, A)_{(CaSO_4)}: \left[ \frac{\text{g}}{\text{l } H_2O} \right]; T: [^\circ\text{C}]$$

The invariant IP (four-phase equilibrium point G + A + L + V or transition temperature gypsum-anhydrite) is defined as the intersection of solubility curves for gypsum and anhydrite on the compositional plane (T-c). However, two shortcomings are associated with this criterion:

(i) Anhydrite does not crystallize in water with measurable rate at temperatures below 70°C -even in the presence of anhydrite seed crystals (**Hardie**, 1967). Then, the solubility equilibrium of this phase cannot be proved experimentally from both sides, that is, from under and supersaturation. In fact, most solubility measurements approached equilibrium only from the undersaturated side (**Freyer & Voigt**, 2003, 2004). Only experimental values reported by **Innorta et al.** (1980) have been obtained from the supersaturated side.

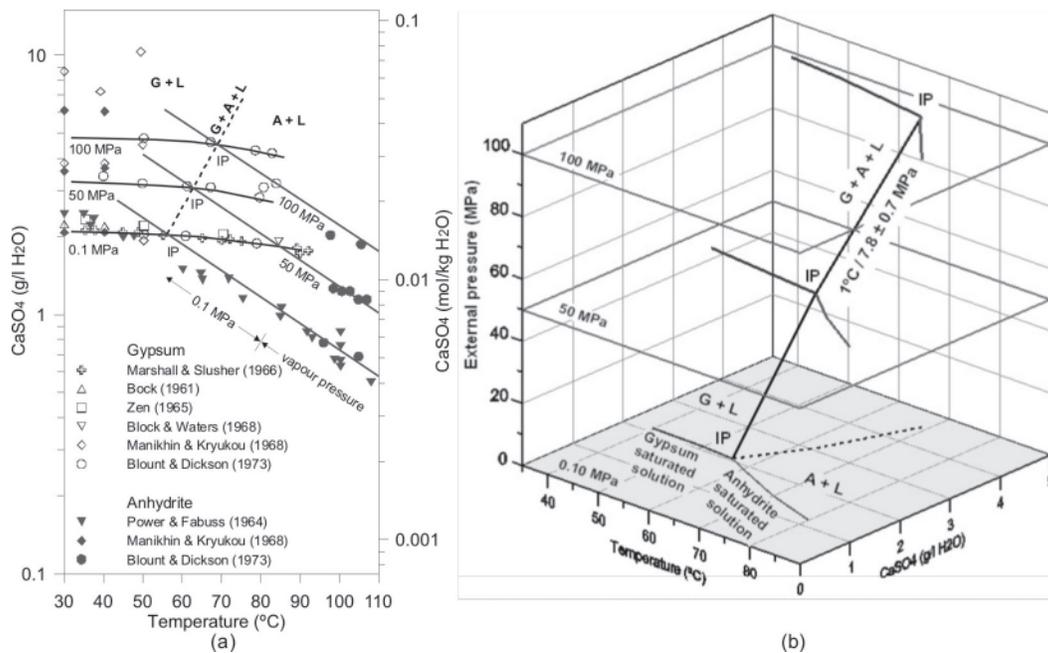
(ii) Experimental data are often considerably scattered; then, IP vary depending on selected parameters.

Due to these shortcomings, some authors have postulated that temperature transition predicted using this criterion would be considered only as a minimum value (**Hardie**, 1967; **Blount & Dickson**, 1973; **Knacke & Gans**, 1977; **Raju & Atkinson**, 1990). It can be seen in figure 2 that borderlines for experimental data on solubility of gypsum and anhydrite yield transition temperatures from about 25°C to 58°C; although the usual proposed values varies between 42°C and 63°C.

## Barotropia

The univariant G + A + L -the equilibrium of gypsum, anhydrite and saturated solutions in pure water-, represents the effect of the confinement pressure on the behaviour of the CaSO<sub>4</sub>-H<sub>2</sub>O system. It is really the envelope of invariants IP respect to P and T, in which invariants IP are deduced from solubility curves for gypsum and anhydrite (G + L + V and A + L + V) obtained at different confinement pressures, in excess of the atmospheric pressure. Unfortunately measurements of the vapour pressure imposed by the CaSO<sub>4</sub>-H<sub>2</sub>O system at high external pressures are not available. Then, even nowadays only projections of this univariant on the compositional plane T-c are really known, and projections on the barometric plane T-uv remain unknown.

In general, the solubility of gypsum and anhydrite increases with pressure; a fact that has been verified repeatedly in laboratory tests (figure 3a); however, the increase of solubility with pressure is higher for anhydrite than for gypsum. Consequently, an increase in transition temperature gypsum-anhydrite is generated when the CaSO<sub>4</sub>-H<sub>2</sub>O system is exposed to pressures in excess of the atmospheric pressure. A series of values representing such a dependence has been postulated, for example: 1°C/8,3 MPa (**Marsal**, 1952), 1°C/8,54 MPa (**McDonald**, 1953), 1°C/7,1 ± 0,19 MPa (**Zen**, 1965) and 1°C/7,8 ± 0,7 MPa



**Figure 3.** Effect of the confinement pressure on the solubility of the CaSO<sub>4</sub>-H<sub>2</sub>O system. (a) Experimental data. (b) Three-dimensional representation of the univariant G + A + L –equilibrium of gypsum, anhydrite and saturated solutions–, following Blount & Dickson (1973).

(Blount & Dickson, 1973). Subsequent researches seem to be in agreement with the latest of these values (Monnin, 1990; Freyer, 2000). A three-dimensional representation of the univariant G + A + L following Blount & Dickson (1973) is presented in figure 3b.

### Hygrotopia

Information on the vapour pressure ( $u_v$ ) or the relative humidity ( $u_v/u_{v0}$ ) imposed by the CaSO<sub>4</sub>-H<sub>2</sub>O system at atmospheric pressure in data-bases on properties of saturated solutions is poor, and only isolate data are reported by some authors (Schneider, 1960; Lide & Frederikse, 1997; Delage *et al.*, 1998; Romero, 2001; Tang & Cui, 2005). The reason is, of course, that gypsum and anhydrite are not effective dehydrators and it lacks interest for industrial applications. Analyses presented here concern to gypsum-saturated solutions (univariant G + L + V) due to information on the relative humidity imposed by anhydrite-saturated solutions (univariant A + L + V) is inconclusive.

Blount & Dickson (1973) postulate a value of vapour pressure of 124 torr (16.53 kPa) at the transition temperature gypsum-anhydrite of 56°C (239,15°K). Using Eq. 5 to calculate the vapour pressure imposed by pure water at 56°C and atmospheric pressure, a value of relative humidity ( $u_v/u_{v0}$ ) =

99,5% is obtained following these authors. On other hand, in measurements using a saturated solution of pure gypsum Panreac® (purity > 98%) in distilled water the dependence of the relative humidity on the temperature was obtained for the range 5°C £ T £ 56°C, and a value of relative humidity ( $u_v/u_{v0}$ ) = 97.8% was obtained at 56°C (Berdugo, 2007).

$$u_{v0} = 136075 e^{\left(\frac{-5239.7}{T+273.15}\right)} \text{ [MPa]} \quad (\text{Eq. 5})$$

$$T: (\text{°C})$$

$$1 \text{ MPa} = 7500.617 \text{ torr}$$

An alternative to evaluate the relative humidity imposed by gypsum-saturated solutions is the use of thermodynamical considerations in the analysis of the stoichiometric equation for the formation of this mineral (Eq. 6). It is possible to write the Gibbs free energy of the reaction in terms of the water activity ( $a_w$ ) as a simplified expression for the solubility constant (K) (Eq. 7). The validity of this approach –used since the early contribution by Kelly *et al.* (1941) until recent works by Freyer & Voigt (2003) and Vanko & Bach (2005), among others–, was confirmed in a recent publication by Coussy (2006) on crystallization of salts in porous media.

At equilibrium,  $\Delta G_{(T)}$  is zero. Therefore, at equilibrium the relative humidity will be related to the standard Gibbs free energy,  $\Delta G_{(T)}^{\phi}$  (Eq. 8).  $\Delta G_{(T)}^{\phi}$  corresponds strictly to the difference between the total standard Gibbs free energy of formation of the products and the reactants.

This exercise was performed in this research using expressions for  $\Delta G_{(T)}^{\phi}$  (gypsum) proposed by **McDonald** (1953) (Eq. 9), and **Hardie** (1967) (Eq. 10) in order to calculate the dependence of the relative humidity imposed by gypsum-saturated solutions at atmospheric pressure on the temperature. Results of this exercise are presented in figure 4, with relative humidity data from other sulphate-saturated solutions.

It is clear from this figure that relative humidity obtained by means of thermodynamical calculations depends on the transition temperature gypsum-anhydrite selected as thermodynamical reference: 40°C (**McDonald**, 1953) and 58°C (**Hardie**, 1967). In spite of the direct or inverse dependence of relative humidity on temperature, sulphate-saturated solutions in general and gypsum-saturated solutions in particular impose high relative humidities; usually above 80%.



$$\Delta G_{(T)} = \Delta G_{(T)}^{\phi} + RT \ln a_w \quad (\text{Eq. 7})$$

$\Delta G_{(T)}$ : Gibbs free energy of the reaction ( $\text{J mol}^{-1}$ )

$\Delta G_{(T)}^{\phi}$ : standard Gibbs free energy ( $\text{J mol}^{-1}$ )

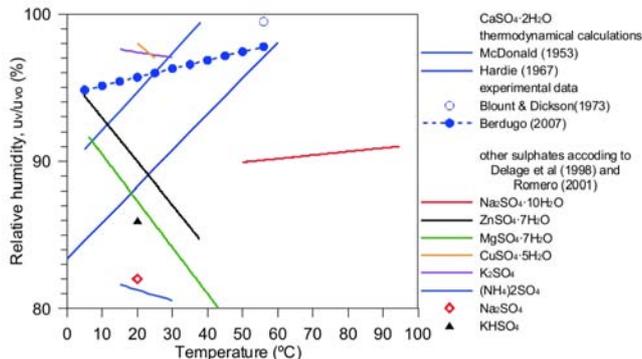
$R$ : universal gas constant ( $8.314 \text{ J mol}^{-1} \text{ K}^{-1}$ )

$T$ : absolute temperature in °K

$a_w$ : water activity or relative humidity,  $\frac{u_v}{u_{vo}}$

$u_v$ : vapour pressure imposed by the solution at atmospheric pressure

$u_{vo}$ : vapour pressure imposed by pure water at atmospheric pressure



**Figure 4.** Dependence of the relative humidity on the temperature for different saturated sulphate solutions at atmospheric pressure.

$$0 = \Delta G_{(T)}^{\phi} (\text{gypsum}) + 2RT \ln \frac{u_v}{u_{vo}} \quad (\text{Eq. 8})$$

$$-(-2495 + 163.89T + 0.0215T^2 - 65.17T \log T) =$$

$$2RT \ln \frac{u_v}{u_{vo}} (\text{McDonald, 1953}) \quad (\text{Eq. 9})$$

$$-(-2870 + 180.43T + 0.0262T^2 - 71.44T \log T) =$$

$$2RT \ln \frac{u_v}{u_{vo}} (\text{Hardie, 1967}) \quad (\text{Eq. 10})$$

The univariant  $G + A + V$ —representing the equilibrium of gypsum, anhydrite and vapour—, only can be estimated by means of second-order empirical correlations between calcium sulphate concentration–vapour pressure, once the latter has been obtained from data on temperature–vapour pressure–salt concentration measured in electrolytic solutions in simultaneous equilibrium with anhydrite and gypsum. The reason is that, even nowadays, temperatures and vapour pressures at which gypsum and anhydrite coexist in the presence of vapour cannot be determined directly due to kinetic hindrances to attain the equilibrium (**Freyer & Voigt**, 2004, **Vanko & Bach**, 2005).

Due to interactions between electrolytic solutions and calcium sulphate do not generate new solid phases, the univariant  $G + A + V$  has been usually studied using the  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  system as geochemical reference. Then, the influence of electrolytes on gypsum and anhydrite solubilities—in terms of either NaCl concentration or solution activity—, can be studied and transition temperature values for the  $\text{CaSO}_4\text{-H}_2\text{O}$  system can be determined by means of two different approaches:

(i) The experimental approach: a procedure based on the equilibrium of gypsum and anhydrite in calibrated solutions of  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  (**Toriumi & Hara**, 1938; **Posnjak**, 1940; **D’Ans et al.**, 1955; **Bock**, 1961; **Power & Satterfield**, 1966; **Hardie**, 1967; **Bount & Dickson**, 1973; **Freyer**, 2000; **Freyer & Voigt**, 2003).

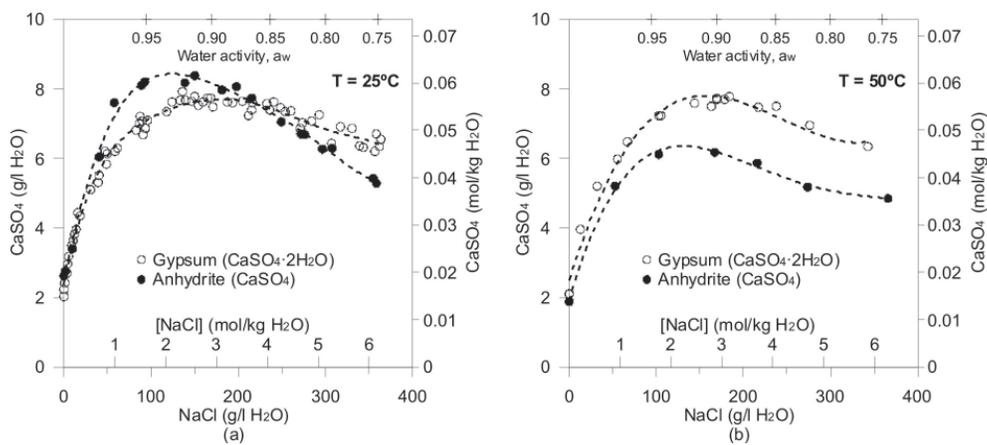
(ii) The transformation energy approach: a theoretical procedure based on the Gibbs free energy dependence on solution activity (**D’Ans**, 1933; **Hill**, 1937; **Posnjak**, 1938; **D’Ans et al.**, 1955; **McDonald**, 1953; **Harvie & Weare**, 1980; **Møller**, 1988; **Raju & Atkinson**, 1990; **Freyer & Voigt**, 2003; **Vanko & Bach**, 2005).

Results of experimental approaches indicate that the addition of non-common ion electrolytes under isothermal conditions increases the solubility of both gypsum and

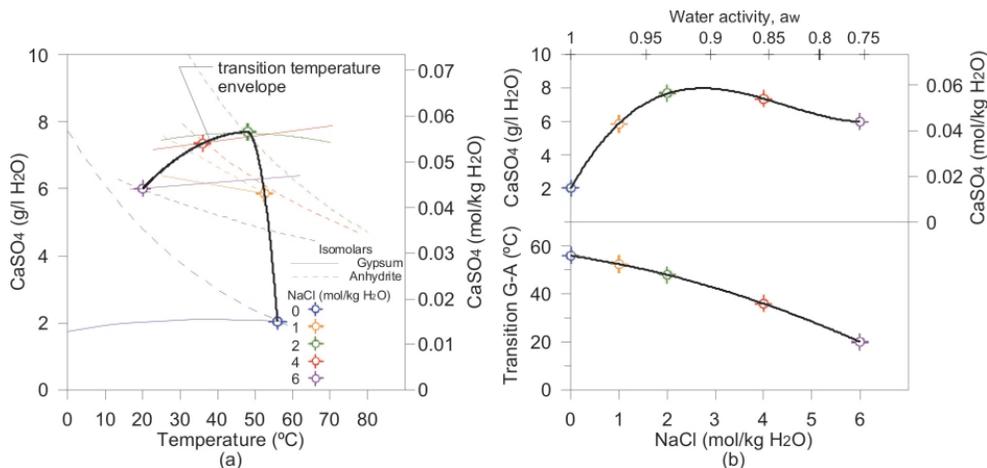
anhydrite until a certain optimum solution activity. Then, a decrease in solubility is related to the hydration ability of the electrolyte; and at very low values of solution activity –a very high salinity–, solution could be undersaturated (or subsaturated) respect to both gypsum and anhydrite. These features are illustrated in figure 5. At 25°C, a classic reference temperature in geochemistry, solubility curves intersect at a NaCl concentration of approximately 4 mol/kg H<sub>2</sub>O (233.8 g/l). Below this value gypsum represents the stable phase, and above it is anhydrite (figure 5a). On other hand, at 50°C, just above the transition temperature gypsum-anhydrite proposed by **Innorta et al.** (1980) and **Møller** (1988), anhydrite is always the stable phase (figure 5b).

When solubility diagrams for gypsum and anhydrite are obtained for several systems CaSO<sub>4</sub>-NaCl-H<sub>2</sub>O following the experimental approach –that is, isomolal solutions with respect to NaCl–, the variation of the transition temperature gypsum-anhydrite with the NaCl molality (or the water activity) can be obtained. This exercise was performed in this research applying the geometrical criterion for IP determination (intersection between solubility curves for anhydrite and gypsum) to isolated experimental data reported by **Blount & Dickson** (1973), and their results are presented in figure 6.

From this figure it seems that the apparent super-saturation associated whit the presence of non-common



**Figure 5.** Solubility of gypsum and anhydrite and its dependence on sodium chloride concentration at atmospheric pressure: (a) at 25°C (Freyer & Voigt, 2003), (b) at 50°C (Møller, 1988).



**Figure 6.** CaSO<sub>4</sub>-NaCl-H<sub>2</sub>O system: gypsum-anhydrite-sodium chloride equilibrium at 0.10 MPa (1 bar). (a) isomolal curves (solubility diagrams) for gypsum and anhydrite, (b) transition conditions in terms of temperature and solubility.

ion electrolytes has an upper limit at a water activity value near 0.92 (figure 6b). However, an inverse relationship between the temperature transition gypsum-anhydrite and the electrolyte concentration remains in spite of the limit value for IP imposed by the water activity.

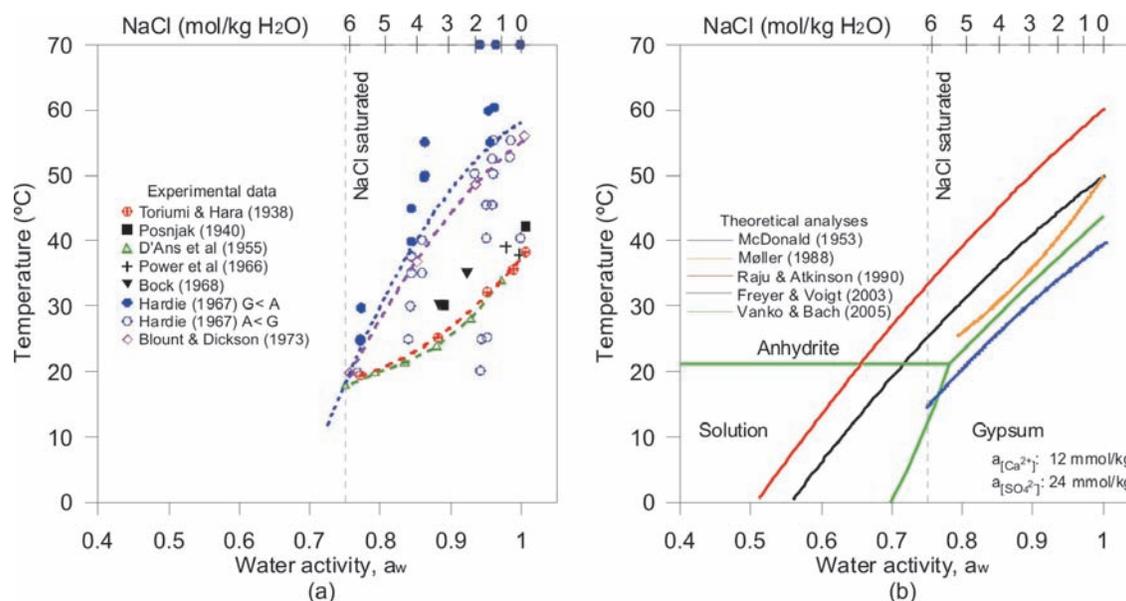
A comparison between experimental data and thermodynamical calculations on the dependence of the transition temperature gypsum-anhydrite on the water activity, using the  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  system at 0,10 MPa as barometric reference, is presented in figure 7. Transition temperatures in pure water according to these results are also indicated in figure 2. From figures 6 and 7 the fact that the transition temperature decreases with decreasing water activity is unquestionable, but two characteristic values seem exist for pure water: 40°C and 56°C.

Discrepancies between these values are often attributed to partial anhydrite-gypsum or gypsum-anhydrite transition in experimental procedures (Hardie, 1967 and Blount & Dickson, 1973); as well as to differences in standard Gibbs free energies selected in order to fit the transition temperature in theoretical procedures (Berdugo, 2007). The latter shortcoming has been avoided correcting the standard entropy of anhydrite in by  $1.6 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ .

Univariant  $G + A + V$  can be obtained from information obtained in the  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  system, but it requires some preliminary considerations.

(i) According to a basic principle of physical chemistry, if one of the phases conforming a multi-phase system in equilibrium at given temperature and pressure is removed the system remains at equilibrium, whenever other circumstances are being held constant. Then, gypsum and anhydrite equilibrated with vapour and NaCl (aqueous) remain in equilibrium with the vapour phase even if the  $\text{NaCl}\cdot\text{H}_2\text{O}$  solution is removed. Consequently, the vapour pressure along univariant  $G + A + V$  can be estimated from relationships between NaCl concentrations and transition temperatures gypsum-anhydrite obtained in  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  systems. It requires empirical correlations between the NaCl concentration (or the water activity) and the vapour pressure of  $\text{NaCl}\cdot\text{H}_2\text{O}$  at different temperatures.

(ii) A series of equations for thermodynamic properties of the  $\text{NaCl-H}_2\text{O}$  system has been proposed by Sparrow (2003). They are polynomials that depend only on composition and temperature, which were adjusted using relationships for the thermodynamical properties of aqueous sodium chloride proposed by other authors (Pitzer *et al.*, 1984 and Archer, 1992). For the case of vapour pressure ( $u_v$ ) between 0°C and 150°C, a expression given by Eq. 11 is proposed for molalities  $> 0$ . Concentration of NaCl is expressed in terms of the salt mass fraction,  $M_f$ , a parameter related to the solution molality,  $m$  (mol/kg  $\text{H}_2\text{O}$ ), and the molar mass of the salt (58.443 g/mol). The vapour pressure associated with pure water ( $u_{v0}$ ) can be calculated using Eq. 5.



**Figure 7.**  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  system: gypsum-anhydrite-saturated solution equilibria at 0.10 MPa (1 bar). Summary of approaches to define the transition temperature gypsum-anhydrite following the activity criterion. (a) experimental approach, (b) transformation energy approach.

$$u_v = A + BT + CT^2 + DT^3 + ET^4 \quad [\text{MPa}] \quad (\text{Eq. 11})$$

$$A = (0.9083 - 0.569M_i + 0.1945M_i^2 - 3.736M_i^3 + 2.82M_i^4) \times 10^{-3}$$

$$B = (-0.0669 + 0.0582M_i - 0.1668M_i^2 + 0.6761M_i^3 - 2.09M_i^4) \times 10^{-3}$$

$$C = (7.541 - 5.143M_i + 6.482M_i^2 - 52.62M_i^3 + 115.7M_i^4) \times 10^{-6}$$

$$D = (-0.0922 + 0.0649M_i - 0.1313M_i^2 + 0.80241M_i^3 - 1.986M_i^4) \times 10^{-6}$$

$$E = (1.237 - 0.753M_i + 0.1448M_i^2 - 6.964M_i^3 + 14.61M_i^4) \times 10^{-9}$$

$$M_i = \frac{mM_{(\text{NaCl})}}{1000 + mM_{(\text{NaCl})}}$$

$m$ : Molality, mol NaCl / kg H<sub>2</sub>O

$M_{(\text{NaCl})}$ : molar mass of NaCl = 58.443 g / mol

A comparison between Sparrow's formulation and experimental data by **Romero** (2001) for the relative humidity ( $u_v/u_{v0}$ ) in the NaCl-H<sub>2</sub>O system at 20°C (Eq. 12) is illustrated in figure 8a. The percent error (figure 8b) is expressed using experimental values as reference.

$$\frac{u_v}{u_{v0}} = 1 - 0.035m - \chi m(m-3) \quad (\text{Eq. 12})$$

$$\chi = 1.142 \times 10^{-3} \text{ for } m < 3 \text{ mol/kg}$$

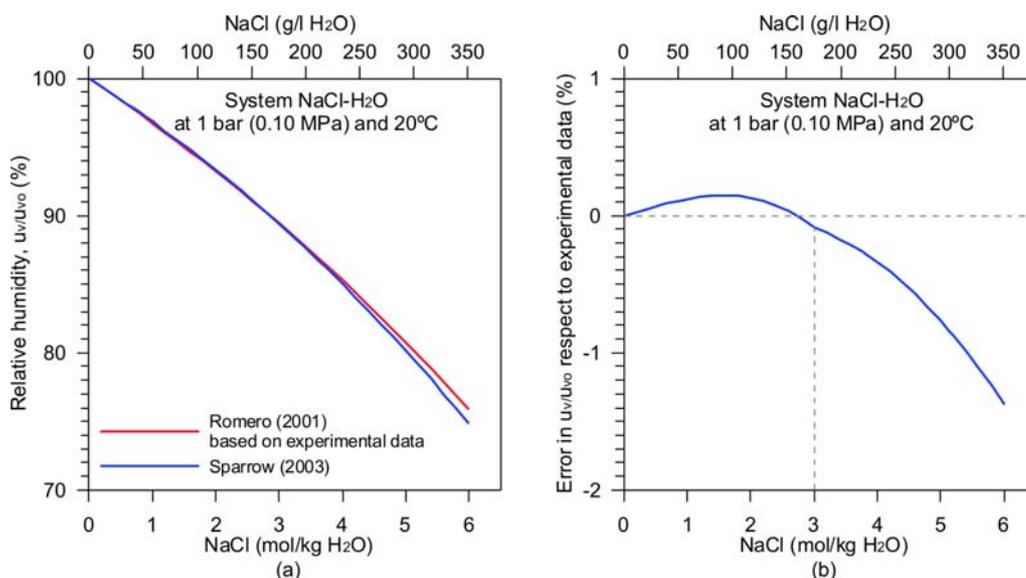
$$\chi = 1.739 \times 10^{-3} \text{ for } m \geq 3 \text{ mol/kg}$$

From this exercise it can be postulated that Eq. 11 predicts the vapour pressure of a sodium chloride solution with reasonable accuracy up to  $m = 3$  mol/kg H<sub>2</sub>O, just the critical value recognized by **Romero** (2001). Beyond 3 mol/kg H<sub>2</sub>O

reliability diverge significantly and may be used with reduced confidence to estimate  $u_v$  at high concentrations.

The formulation by **Sparrow** (2003) was applied to experimental data reported by **Hardie** (1967) and **Blount & Dickson** (1973) for the CaSO<sub>4</sub>-NaCl-H<sub>2</sub>O system at atmospheric pressure –which converge at a transition temperature 56°C in pure water–. The result was the univariant G + A + V shown in figure 9a. At 56°C this univariant give a saturation vapour pressure  $u_v = 121,8$  torr (16,24 kPa); associated with  $u_{v0} = 124,6$  torr (16,61 kPa). Then, a relative humidity ( $u_v/u_{v0}$ ) = 97,8% is obtained at the invariant four-phase equilibrium G + A + L + V. This value for  $u_v$  at IP is within the experimental error recognized by **Hardie** (1967) and **Blount & Dickson** (1973):  $124 \pm 9$  torr. However, in those works the invariant IP was fixed to  $a_w = 1$  in order to extrapolate the vapour pressure from low activity to the saturation condition in pure water. This implies that the invariant point was estimated neglecting the effect that dissolved Ca<sup>2+</sup> and SO<sub>4</sub><sup>2-</sup> exert on the vapour pressure. This simplification is founded on the idea that the tiny lowering of the vapour pressure resulting from dissolved Ca<sup>2+</sup> and SO<sub>4</sub><sup>2-</sup> is compensated by the positive effect of the confining pressure (0,10 MPa), as was pointed out by **Blount & Dickson** (1969, 1973). In reality, when saturated in calcium sulphate at atmospheric pressure water is characterized by a truly high relative humidity (near 98%, as was calculated here), but not by an ideal value of 100%.

This consideration invalidates the theoretical relative humidity value reported by **Blount & Dickson** (1973)



**Figure 8.** NaCl-H<sub>2</sub>O system at 20°C: (a) relative humidity according to Romero (2001) and Sparrow (2003), (b) error of the formulation proposed by Sparrow (2003) respect to experimental data reported by Romero (2001).

(figure 4), based on the saturation vapour pressure at 56°C: 99,5% for  $u_v = 124$  torr. It is proposed that a value of  $u_v/u_{vo} = 97,8\%$  for  $u_v = 121.8$  torr at 56°C under atmospheric pressure –considering the solubility at equilibrium–, is more consistent with experimental data. From figure 9b it is clear that at high water activity hydration ability of electrolyte governs the imposed relative humidity and, as it was mentioned above, calcium sulphate solutions could be subsaturated respect to both gypsum and anhydrite.

### Three-dimensional representation of the phase diagram for the Ca-SO<sub>4</sub>-H<sub>2</sub>O system

Considerations presented above suggest that for a proper analysis of phase relationships in the CaSO<sub>4</sub>-H<sub>2</sub>O system the incorporation of the compositional plane T-c is necessary and the three-dimensional representation of univariants in terms of temperature-concentration-vapour pressure is unavoidable.

In this research the CaSO<sub>4</sub>-H<sub>2</sub>O system at atmospheric pressure as barometric reference is associated with a temperature transition gypsum-anhydrite of 56°C (**Blount and Dickson, 1973**), and the solubility curves for gypsum and anhydrite proposed by these authors are used as reference (Eq. 1 and Eq. 2, respectively). In these conditions, at IP the equivalent (CaSO<sub>4</sub>) concentration is 2.05 g/l, which in terms of gypsum (CaSO<sub>4</sub>·2H<sub>2</sub>O) solubility corresponds to 2.58 g/l.

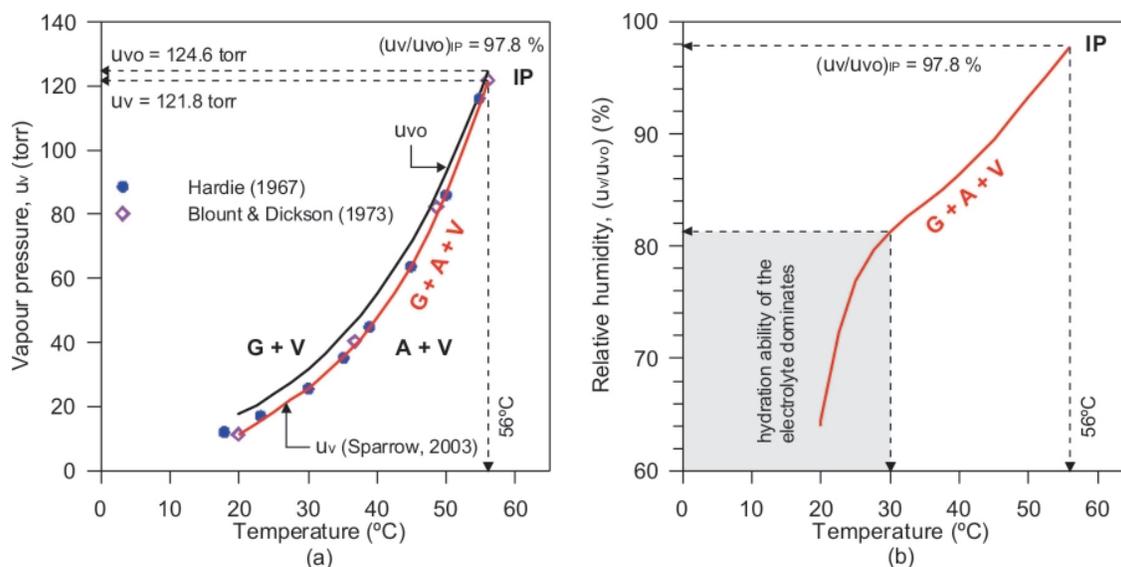
In the absence of additional information on the relative humidity imposed by gypsum-saturated solutions at atmospheric pressure, the single linear relationship T-( $u_v/u_{vo}$ ) based on experimental data by **Berdugo (2007)** is adopted (Eq. 13). Then, the combination of Eq. 1, Eq. 5 and Eq. 13 give the three-dimensional form of univariant G + L + V. Univariant A + L + V is only represented as a projection on the compositional plane due to information on relative humidity imposed by anhydrite-saturated solutions is only available for very high temperatures ( $T > 90^\circ\text{C}$ ) –the case of desalinization processes or industrial applications (**Freyer & Voigt, 2004**)–.

For univariant G + A + L, the criterion proposed by **Blount & Dickson (1973)** ( $1^\circ\text{C}/7,8 \pm 0,7$  MPa) is adopted. Finally, univariant G + A + V is formulated using parameters presented in table 2, which were obtained from figures 6 and 9a. The result of this exercise is presented in figure 10.

A important feature presented in figures 3 and 6 can be now properly visualized in figure 10: water activity (salinity or the associated relative humidity) is the key factor for existence of high concentration of Ca<sup>2+</sup> and SO<sub>4</sub><sup>2-</sup> in the aqueous system, and confinement pressure plays a secondary role, even at very high values (for example  $> 50$  MPa). On the other hand, the increases in water activity generates true metastable conditions for both gypsum and anhydrite occurring at vapour pressures below the vapour pressure imposed in pure water.

$$\left(\frac{u_v}{u_{vo}}\right)_{\text{gypsum}(\%)} = 5.81 \times 10^{-2} T + 94.548 \quad (25^\circ\text{C} \leq T \leq 56^\circ\text{C}) \quad (\text{Eq. 13})$$

$P = 1 \text{ atm}, T$



**Figure 9.** CaSO<sub>4</sub>-H<sub>2</sub>O system at 0.10 MPa (1 bar) deduced from the CaSO<sub>4</sub>-NaCl-H<sub>2</sub>O system: (a) univariant G + A + V in terms of vapour pressure, (b) univariant G + A + V in terms of relative humidity.

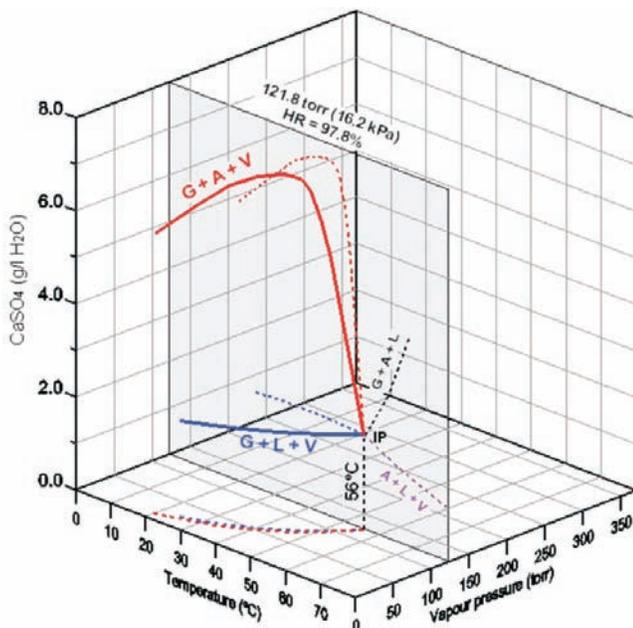
**Table 2.** Representative parameters for univariant  
G + A + V, P = 0,10 MPa (1 bar).

NaCl (mol/kg H <sub>2</sub> O)	T (°C)	CaSO <sub>4</sub> (g/l H <sub>2</sub> O)	uv (torr)
0	56	2.05	121.8
1	52	5.86	97.0
2	48	7.70	76.8
4	36	7.34	37.5
6	20	6.00	11.4

### Conclusive remarks

From the review presented above it is clear that only univariants G + L + V, A + L + V and G + A + L have been directly obtained by means of experimental methods using pure water as solvent, and the univariant G + A + V is often indirectly obtained using the CaSO<sub>4</sub>-NaCl-H<sub>2</sub>O system as referent.

The precise value for the transition temperature gypsum-anhydrite is an open discussion –even in the simplest case of calcium sulphate saturated solutions in pure water at atmospheric pressure–, and an accurate knowledge on the vapour pressure associated with univariants at different confinement pressures is not available. Therefore, the three-dimensional representation of the CaSO<sub>4</sub>-H<sub>2</sub>O system presented in figure 10 illustrates only partially the behaviour of the system.

**Figure 10.** Three-dimensional representation of the CaSO<sub>4</sub>-H<sub>2</sub>O system at atmospheric pressure as barometric reference

Experimental and thermodynamical analyses indicate that the transition temperature gypsum-anhydrite in pure water at atmospheric pressure varies between 42°C and 63°C. In this case a direct dependence of the relative humidity on the temperature can be adopted for gypsum-saturated solutions within the range 5°C=T=56°C, and the characteristic relative humidity is about 96%. On the other hand, the solubility, the transition temperature and the vapour pressure are strongly affected by presence of other ions. The addition of non-common ion electrolytes increases the solubility of gypsum and anhydrite and decreases both the transition temperature and the vapour pressure; so, a reduction in relative humidity occurs. However, thermotopia of both phases solubility is not affected.

Metastable states for gypsum and anhydrite without dissolution or precipitation can exist below the transition temperature as a result of water salinity. Alterations of this metastable equilibrium are undoubtedly related with temperature changes, but mainly with variations in the relative humidity imposed by the surrounding environment. The cause is that in spite of the direct or inverse nature of both solubility and vapour pressure, sulphate-rich solutions impose high values of relative humidity, often above 80%. Then, exposed to moderate dry environments these solutions are capable to transfer vapour towards the environment increasing the concentration in Ca<sup>2+</sup> and SO<sub>4</sub><sup>2-</sup> and generating true supersaturated conditions; the basic requirement for solvent-way precipitation and gypsum crystal growth. Nevertheless, above the transition temperature anhydrite is the stable phase at any relative humidity.

The behaviour of the CaSO<sub>4</sub>-NaCl-H<sub>2</sub>O system is a coupled thermo-hydro-chemical (THC) phenomenon in which water activity is the key factor for existence of high concentration of Ca<sup>2+</sup> and SO<sub>4</sub><sup>2-</sup>, and confinement pressure plays a secondary role. Consequently, the selection of the most representative values for solubility, transition temperature gypsum-anhydrite and relative humidity deepens on the boundary conditions imposed by environmental, geotechnical or industrial concerning processes.

### Acknowledgements

The Spanish Ministry for Infrastructures was the financial supporter for this research. The authors wish to thank the advising provided by their colleagues Prof. Dr. Ing. Lucila Candela-Lledó, Prof. Dr. Ing. Marcos Arroyo-Álvarez de Toledo and Dr. Ing. Juliana Knobeldorff.

## References

- Archer, D.G.** 1992. Thermodynamic properties of the NaCl-H<sub>2</sub>O system II. J. Physical and Chemical Reference Data 21, Issue 4: 793-829.
- Berdugo, I.R.** 2007. Tunnelling in sulphate-bearing rocks – expansive phenomena. Dr. Thesis, Universitat Politècnica de Catalunya.
- Block, J & Waters, O.B.** 1968. The CaSO<sub>4</sub>-Na<sub>2</sub>SO<sub>4</sub>-NaCl-H<sub>2</sub>O system at 25° to 100°C. J. Chem. Eng. Data 13: 336-344.
- Blount C.W. & Dickson F.W.** 1969. The solubility of anhydrite CaSO<sub>4</sub> in NaCl-H<sub>2</sub>O from 100 to 450°C and 1 to 1000 bars. Geochim. Cosmochim. Acta 33: 227-245.
- Blount C.W. & Dickson F.W.** 1973. Gypsum anhydrite equilibria in systems CaSO<sub>4</sub> and CaCO<sub>3</sub>-NaCl-H<sub>2</sub>O. The American Mineralogist, 58: 323-331.
- Bock, E.** 1961. On the solubility of anhydrous calcium sulfate and of gypsum in concentrated solutions of sodium chloride at 25°C, 30°C, 40°C and 50°C. Can. J. Chem. 39: 1746-1751.
- Coussy, O.** 2006. Deformation and stress from in-pore drying-induced crystallization of salt. Journal of the Mechanics and Physics of Solids, v 54, n 8: 1517-1547.
- D'Ans, J.** 1933. Die Lösungsgleichgewichte der System der Salz ozeanischer Salzablagerungen. Berlin: 118-123.
- D'Ans, J., Bredtscheider, D., Eick, H. & Freund, H.E.** 1955. Untersuchungen über die Calciumsulfate. Kali Steinsalz, 9: 17-38.
- D'Ans, J.** 1968. Der Übergangspunkt Gips – Anhydrite. Kali U. Steinsalz, 5: 109-111.
- Delage, P., Howat, M.D. & Cui, Y.J.** 1998. The relationship between suction and swelling properties in a heavily compacted unsaturated clay. Engineering Geology 50: 31-48.
- Freyer, D.** 2000. Zur Phasenbildung und -stabilität im System Na<sub>2</sub>SO<sub>4</sub>-CaSO<sub>4</sub>-H<sub>2</sub>O. Dissertation, TU Bergakademie Freiberg.
- Freyer, D. & Voigt, W.** 2003. Crystallization and Phase Stability of CaSO<sub>4</sub> and CaSO<sub>4</sub> – Based Salts. Monatshefte für Chemie 134: 693-719.
- Freyer, D. & Voigt, W.** 2004. The measurement of sulfate mineral solubilities in the Na-K-Ca-Cl-SO<sub>4</sub>-H<sub>2</sub>O system at temperatures of 100, 150 and 200°C. Geochimica et Cosmochimica Acta, Vol. 68, No. 2: 307-318.
- Hardie, L.A.** 1967. The gypsum-anhydrite equilibrium at one atmosphere. Amer. Mineral. 52: 171-200.
- Harvie, C.E. & Weare, J.H.** 1980. The prediction of mineral solubilities in natural waters: Na-K-Mg-Ca-Cl-SO<sub>4</sub>-H<sub>2</sub>O system from zero to high concentration at 25°C. Geochim. Cosmochim. Acta 44: 981-997.
- Hill, A.** 1937. The transition temperature of gypsum to anhydrite. J. Am. Chem. Soc. 59 (11): 2242-2244.
- Innorta, G., Rabbi, E. & Tomadin, L.** 1980. The gypsum-anhydrite equilibrium by solubility measurements. Geochim. Cosmochim. Acta 44: 1931-1936.
- Kelly, K.K., Southard, J.C. & Anderson, C.T.** 1941. Thermodynamic properties of gypsum and its dehydration products. U.S. Bur. Mines, Tech. Papers. 625.
- Knacke O, Gans W.** 1977. The thermodynamics of the system CaSO<sub>4</sub>-H<sub>2</sub>O. Z. Phys. Chem. 104: 41-48.
- Kontrec, J., Kralj, D. & Brecevic, L.** 2002. Transformation of anhydrous calcium sulphate into calcium sulphate dihydrate in aqueous solutions. Journal of Crystal Growth, 240, n 1-2: 203-211.
- Lide, D.R. & Frederikse, H.P.R.** 1997. CRC Handbook of chemistry and physics. A ready reference book of chemical and physical data. CRC Press, New York.
- Manikhin, V.I. & Kryukou, P.A.** 1968. Effect of pressure on the solubility of sodium and calcium sulfates. Porovye Rastvory Metody Ikh. Izuch., 133-44.
- Marsal, D.** 1952. Der Einfluss des Druckes auf das System CaSO<sub>4</sub>-H<sub>2</sub>O. Heidelberger Beit. Mineral Petrol 3: 289-296.
- Marshall, W.L. & Slusher R.** 1966. Thermodynamics of calcium sulphate dihydrate in aqueous sodium chloride solutions, 0-110°C. J. Phys. Chem. 70: 4015-4027.
- McDonald, G.J.F.** 1953. Anhydrite-gypsum equilibrium relations. American Journal of Science, vol. 251: 884-898.
- Monnin, C.** 1990. The influence of pressure on the activity coefficients of the solutes and on the solubility of minerals in the system Na-Ca-Cl-SO<sub>4</sub>-H<sub>2</sub>O to 200°C and 1 kbar and to high NaCl concentration. Geochimica et Cosmochimica Acta, Volume 54, Issue 12: 3265-3282.
- Møller N.** 1988. The prediction of mineral solubilities in natural waters: A chemical equilibrium model for the Na-K-Ca-Cl-SO<sub>4</sub>-H<sub>2</sub>O system, to high temperature and concentration. Geochim. Cosmochim. Acta 52: 821-837.
- Pitzer, K.S., Peiper, J. C., & Busey, R.H.** 1984. Thermodynamic properties of aqueous sodium chloride solutions. J. Phys. Chem. Ref. Data 13: 1-102.
- Posnjak E.** 1938. The system CaSO<sub>4</sub>-H<sub>2</sub>O. Am. Jour. Sci., 5th ser., 35A: 247-272.
- Posnjak E.** 1940. Deposition of calcium sulfate from sea water. Am. J. Sci. 238: 559-568.
- Power, W.H. & Fabuss, B.M.** 1964. Thermodynamic properties of saline water. Office of Saline Water Research and Development Progress Rept. 104.
- Power, W.H. & Satterfield, C.N.** 1966. Transient solute concentrations and phase changes of calcium sulfate in aqueous sodium chloride. J. Chem. Eng. Data 11: 149-154.
- Raju, K. & Atkinson, G.** 1990. Thermodynamics of 'scale' mineral solubilities. 3. Calcium sulfate in aqueous NaCl. Journal of Chemical and Engineering Data, v 35, n 3: 361-367.
- Romero, E.** 2001. Controlled-suction techniques. 4° Simpósio Brasileiro de Solos Não Saturados. W.Y.Y. Gehling & F. Schnaid (eds.): 535-542.
- Sattler, H. & Brückner, H.** 2001. Volumen- und Dichteänderungen bei der Hydratation von Gipsbindemitteln in Abhängigkeit vom

Wasserangebot. Zement, Kalk, Gips international, Vol. 54 No. 9: 522-529.

**Schneider, A.** 1960. Neue diagramme zur bestimmung der relativen luftfeuchtigkeit uber gesattigten wasserigen salzlosungen und wasserigen schwefelsaurelosungen bei verschiedenen temperaturen. Holz als Rohund Werkstoff 18: 269-272.

**Sparrow, B.** 2003. Empirical equations for the thermodynamic properties of aqueous sodium chloride. Desalination 159: 161-170.

**Tang, A.M. & Cui, Y.J.** 2005. Controlling suction by the vapour equilibrium technique at different temperatures and its application in determining the water retention properties of MX80 clay. Can. Geotech. J. 42: 287-296.

**Toriumi, T. & Hara, R.** 1938. On the calcium sulfate in sea water II. Solubilities of calcium sulphate hemihydrate in sea water of

various concentrations at 65°-1 50°. Technol. Rep. Tohoku Imp. Univ. 12: 560-571.

**Vanko, D.A. & Bach, W.** 2005. Heating and freezing experiments on aqueous fluid inclusions in anhydrite: recognition and effects of stretching and the low-temperature formation of gypsum. Chemical Geology 223: 35-45.

**van't Hoff, J.H., Armstrong, E.F., Hinrichsen, W., Weigert, F. & Just, G.** 1903. Gips und Anhydrit. Z. phys. Chem. 45: 257-306.

**Zen, E.-AN.** 1965. Solubility measurements in the system  $\text{CaSO}_4\text{-NaCl-H}_2\text{O}$  at 35°C, 50°C and 70°C and one atmosphere pressure. J. Petrol. 6: 124-164.

Recibido: noviembre 5 de 2008

Aceptado para su publicación: noviembre 18 de 2008

# ESQUEMAS DE REACCIÓN PARA LICUEFACCIÓN Y GASIFICACIÓN HIDROTÉRMICA DE MACROMOLÉCULAS ORGÁNICAS. UNA REVISIÓN

Por

Alejandro Amadeus Castro V.<sup>1\*</sup>, Luis Ignacio Rodríguez V.<sup>2</sup>, José de Jesús Díaz V.<sup>1</sup>

## Resumen

Castro V., A.A., L.I. Rodríguez & J. de J. Díaz V.: Esquemas de reacción para licuefacción y gasificación hidrotérmica de macromoléculas orgánicas. Una revisión. Rev. Acad. Colomb. Cienc. 32(125): 559-570, 2008. ISSN 0370-3908.

Las macromoléculas orgánicas son las principales componentes de la biomasa, entre ellas: carbohidratos, lignina y proteínas. La abundancia de enlaces tipo éter y éster en sus mesómeros los hacen susceptibles de ser despolimerizados y la aplicación de un ambiente hidrotérmico permite escoger la ruta predominante: hidrólisis o deshidratación, descarboxilación y escisión de grupos funcionales. En esta revisión se compilan esquemas de reacción hidrotérmicos, especialmente para polisacáridos y lignanos, los cuales son los principales contribuyentes a la diversidad química de los productos de conversión hidrotérmica de biomasa lignocelulósica y sus propiedades combustibles, debido a que la reducción en el contenido de oxígeno y el consecuente aumento del poder calorífico está ligado a las reacciones de estas moléculas. La fracción hidrofóbica de los productos obtenidos por licuefacción y gasificación hidrotérmica parcial de los constituyentes de la biomasa puede utilizarse como bio-combustible.

**Palabras clave:** macromoléculas orgánicas, procesos hidrotérmicos, biomasa, bio-crudo, mecanismos de reacción.

## Abstract

The main components of biomass are organic macromolecules: carbohydrates, lignin and proteins. The abundance of ether-like and ester bonds in their mesomeric units does them susceptible to be depolymerized and the usage of a hydrothermal atmosphere allows us to choose the predominant

<sup>1</sup> Universidad Nacional de Colombia, Sede Bogotá, Facultad de Ciencias, Departamento de Química, Laboratorio de Investigación en Combustibles y Energía.

\* Autor correspondiente, teléfono: (+57+1+316) 4131309, correos electrónicos: aacastrov@unal.edu.co, aacastrov@gmail.com

<sup>2</sup> Universidad Nacional de Colombia, Sede Bogotá, Facultad de Ingeniería, Departamento de Ingeniería Química.

route: hydrolysis or dehydration, decarboxylation and scission of functional groups. In this review are compiled hydrothermal reaction pathways, specially polysaccharides and lignanes, these ones are the main molecules that contribute to chemical diversity and fuel properties in lignocellulosic biomass hydrothermal conversion products, due to the oxygen content reduction in these macromolecules and the inherent raise of heating value. The hydrophobic products from liquefaction and partial hydrothermal gasification of biomass components can be used as bio-fuel.

**Key words:** organic macromolecules, hydrothermal processes, biomass, bio-crude, reaction pathways.

## 1. Introducción

La transformación de fisicoquímica de los constituyentes de la biomasa (macromoléculas orgánicas, grasas, agua y minerales) se realiza por medio de diversos procesos tecnológicos, entre los que se tienen: combustión, carbonización, pirólisis, gasificación, fermentación, digestión, licuefacción, trans-esterificación (para lípidos) y conversión hidrotérmica (Veringa, 2004). Cada proceso expande la capacidad de procesamiento y permite transformar un universo más amplio de sustancias, las cuales pueden convertirse en combustibles (Fröling *et al.*, 2005), dado su valor agregado superior. El más importante de los combustibles, el petróleo, proviene de la conversión hidrotérmica de macromoléculas orgánicas (biomasa), especialmente de lípidos (Siskin *et al.*, 1990b), gomas y ceras (Schobert, 1990) y moléculas isoprenoides (Curry, 1988). Estas moléculas se protegieron de la descomposición natural al ser cubiertas por agua y sedimentos (Schobert, 1990) bajo condiciones anaerobias (Suárez, 2005) y su descomposición originó el kerógeno por reacciones de descarboxilación y deshidratación (Rouxhet, 1978; Schobert, 1990). La presencia de marcadores biológicos como clorofila y sus porfirinas en la composición de pizarras bituminosas, asfaltos y carbón, evidencian y soportan el postulado de la rápida protección de la biomasa en contra de la oxidación y su transformación en combustibles (gas, petróleo y carbones) (Landes, 1977; Siskin *et al.*, 1990b).

La combinación de agua en estado líquido y temperaturas elevadas constituye un ambiente hidrotérmico (Mandoki, 1986; Shaw, 1991; Luijkx, 1994) y en tales condiciones actúa simultáneamente como solvente, catalizador y reactante (Kuhlmann *et al.*, 1994; Katritsky *et al.*, 1996). La condición hidrotérmica se favorece por el desprendimiento de dióxido de carbono en estado supercrítico, que incrementa la presión del sistema aunque frecuentemente es suficiente la presión endógena para mantener el agua en fase líquida (Luijkx, 1994; Saleh, 1994). La evolución hidrotérmica de macromoléculas or-

gánicas (diagénesis) hacia kerógeno se demuestra con suficiencia al someter crudos inmaduros a pirólisis en medio acuoso (Hoering, 1984; Bakr, 1990; Koopmans, 1998). Un moderado aumento en la temperatura durante largos tiempos geológicos produce los mismos resultados que los conseguidos por una temperatura alta en pocos minutos cuando están presentes catalizadores (Landes, 1977; Siskin y Katrytzky, 2000) y por esta razón, la conversión hidrotérmica es un proceso con posibilidades industriales. En la naturaleza, estas reacciones pueden catalizarse por minerales ácidos, básicos, sales o metales alcalinos (Siskin y Katrytzky, 2000). Los ambientes geológicos que pudieron dar origen al petróleo se han simulado, efectuando reacciones por tiempos muy prolongados (semanas), en presencia de cloruro de sodio y Montmorillonitas (Siskin *et al.*, 1990b).

El uso eficiente de los combustibles es una prioridad frente al inevitable incremento de los precios del petróleo y al agotamiento de los campos productores. Desde la década de 1970 se ha mostrado la factibilidad de la despolimerización hidrotérmica (Campbell, 1976; Mandoki, 1986) para la obtención de hidrocarburos y otras sustancias combustibles a partir de la biomasa y moléculas artificiales tóxicas (Modell, 1978). Investigaciones posteriores muestran el 'reformado' acelerado de los constituyentes de la biomasa y los productos mismos de despolimerización (Annee, 1986; Goudriaan, 1990; Mok, 1992; Luijkx, 1994; Minowa, 1997; Szabo, 1998; Yan, 1999; Antal *et al.*, 2000; Saisu, 2003; Kruse, 2003; Kruse, 2004; Bicker, 2005; Kamio, 2006).

El Laboratorio de Investigación en Combustibles y Energía inició en 2005 actividades de tipo teórico y experimental en los procesos de conversión hidrotérmica. En este artículo se presenta una revisión de los mecanismos de reacción de la biomasa durante la conversión hidrotérmica (licuefacción y gasificación parcial) y se consolidan esquemas globales de conversión de las principales macromoléculas que la constituyen (polisacáridos, lignanos, y polipéptidos) y algunos productos de degradación de éstas mismas.

## 2. Mecanismos de conversión hidrotérmica (licuefacción y gasificación) de macromoléculas orgánicas (biomasa)

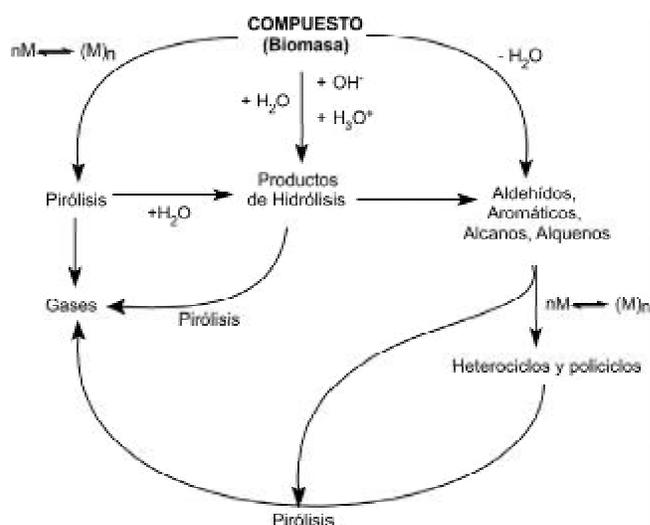
### 2.1. Consideraciones previas: de la pirólisis hacia la despolimerización hidrotérmica

El levoglucosano (1-6 deshidroglucosa) es el principal producto de despolimerización en la pirólisis de la celulosa (poli  $\beta$ -1,4 glucano). Los oligosacáridos con unidades terminales de anhidroglucosa se degradan en levoglucosano por escisión aleatoria de moléculas. Adicionalmente a estos mecanismos de ruptura, existen especies en competencia, las cuales presentan rutas más ‘complicadas’ como deshidratación y desproporciónación (**Schwarzinger**, 2005). En general, la pirólisis de macromoléculas orgánicas produce anhídroligómeros (**Piskorz**, 2000), que tras sucesivas degradaciones, pueden generar gases y líquidos con contenidos de oxígeno altos, carbonizados y cenizas. En el procedimiento clásico de pirólisis de biomasa, cuando el material orgánico se procesa húmedo (contenido de humedad cercano a 50%) (**Demirbas**, 2004), se observa el incremento de la fracción líquida de biocrudo. Este fenómeno constituye el primer acercamiento y transición entre la pirólisis y la conversión hidrotérmica, removiéndose oxígeno de las macromoléculas de la biomasa (**Goudriaan et al.**, 2003) y los productos presentan un poder calorífico superior, asemejándose al de los hidrocarburos (**Anne**, 1987). La remoción del oxígeno sucede principalmente por eliminación de dióxido de carbono y agua (**Demirbas**, 2000).

La licuefacción y gasificación hidrotérmica de biomasa inician con mecanismos de ruptura. Las moléculas que contienen enlaces tipo éter, fenol o amina se hidrolizan y pirolizan simultáneamente. Existe una condición de equilibrio entre los monómeros y el polímero (**Sawada**, 1976), y los radicales orgánicos formados se recombinan a una velocidad aproximada de  $10^9$  a  $10^{10} \text{ mol}\cdot\text{s}^{-1}$  (**Stein**, 1985). Los grupos funcionales pirolizados generan moléculas pequeñas y gases.

La solubilización de los monómeros de biomasa es apreciable a partir de 603 K (**Bobleter**, 1986; **Mok**, 1992). En la celulosa ocurre despolimerización hacia glucosa, isómeros y oligómeros. Los radicales pueden estabilizarse transfiriendo un átomo de hidrógeno a otros radicales (**Stein**, 1985), repolimerizando, adicionándose a otros radicales o participando en reacciones de abstracción o escisión (**Savage**, 2005). Sin dicha estabilización aparecen alquitranes y carbonizados entre los productos.

La información anterior permite sugerir el esquema global de conversión que se muestra en la Figura 1 para las macromoléculas orgánicas despolimerizadas en condición



**Figura 1.** Esquema global de rutas y productos de conversión hidrotérmica de macromoléculas orgánicas de la biomasa.

hidrotérmica presentes en la biomasa. Este esquema es válido solamente si la conversión se realiza en ausencia de especies oxidantes, condición última necesaria para que el procedimiento se asemeje a la fosilización natural de macromoléculas que dieron origen al petróleo. La biomasa puede despolimerizar por tres vías: la primera es una escisión aleatoria en la que la macromolécula se encuentra en equilibrio con sus mesómeros ( $nM \rightleftharpoons \chi(M)_n$ ) para dar productos de pirólisis que pueden estabilizarse por hidratación o generar productos gaseosos. La segunda vía es la hidrólisis ocasionada por la acción del agua y sus especies dissociadas; los monómeros y oligómeros obtenidos pueden presentar pirólisis de sustituyentes para producir gases o presentar reacciones de condensación aldólica. La tercera vía es una deshidratación que conduce directamente a los productos de condensación aldólica, tales como aldehídos, aromáticos y si existen lípidos, pueden tenerse alcanos y alquenos. Estos productos pueden repolimerizar en hetero y policiclos, fuentes de alquitranes y carbonizados, o pirolizar para dar gases.

### 2.2. Conversión hidrotérmica de polisacáridos

Los polisacáridos son los principales componentes de la biomasa (**Brucker**, 1912; **Braverman**, 1980). Su conversión hidrotérmica se asemeja a una disolución por hidrólisis (**Sasaki et al.**, 2002a) que es evidente a partir de 463 K (**Bobleter**, 1986; **Deguchi et al.**, 2006). La velocidad de descomposición incrementa conjuntamente con la temperatura. Los productos de hidrólisis son oligosacáridos, monómeros e isómeros (**Kabyemela**, 1999; **Sasaki et al.**,

2002a). Los productos de degradación se componen principalmente de furfurales, aldehídos de cadena corta, ácidos y fenoles polihidroxilados, cetonas, otros aldehídos, fenoles y ácidos carboxílicos (Luijckx, 1994; Srokol *et al.*, 2004; Ponder, 1992; Siskin y Katrytzky, 2000). Estos últimos predominan cuando la reacción se desarrolla en condición subcrítica (condensación retro-aldol, tautomerismo ceto-enol, deshidratación) (Sasak *et al.*, 2002), y descarboxilación (Luijckx, 1994). Es posible que ocurra deshidrogenación por desprendimiento de moléculas de agua durante reacciones de deshidroxilación (Barth, 2000).

Polisacáridos con funcionalidades como las pectinas, los queratanos, glicolípidos y la quitina pueden presentarse en la biomasa (Braverman, 1980; Dumitriu, 2005); siendo éstas moléculas las que le siguen en importancia a la celulosa (Kurita, 2003). Las pectinas y quitinas se hidrolizan hacia poli-(ácido galacturónico) y quitosano, respectivamente y oligómeros de grado de polimerización entre 2 y 10, los cuales también producen monómeros (Miyazawa, 2004; Sakai, 1998). Posteriormente se presenta degradación (Miyazawa, 2004) y procesos simultáneos de descarboxilación, desaminación, y desacetilación; originando oligómeros y monómeros de tipo celulósico (Quitain, 2001) cuya estructura degrada bajo los mecanismos para polisacáridos.

### 2.3. Conversión hidrotérmica de lignanos

La estructura mesomérica básica de los lignanos es el fenilpropano (Demirbas, 2001), pero son el guayacol, catecol, fenol y ácido vanílico las moléculas que se han tomado como compuestos modelo de lignina debido a que son éstas las que se producen en experiencias de conversión hidrotérmica. El mecanismo de descomposición hidrotérmica de la lignina es compartido entre hidrólisis y pirólisis (Serikawa, 1992) ~~del enlace~~ ?-O-4. Derivados del metoxifenol, fenoles, catecoles y trazas de benceno se producen bajo una cinética de primer orden (Demirbas, 2000). Las reacciones de pirólisis ocasionan desmetilación, deshidroxilación y deshidratación (Barth, 2000). El ácido vanílico empleado como compuesto modelo evidencia conversiones hacia guayacol (2-metoxifenol) y degradación hacia catecol (González, 2004) y generación de dióxido de carbono (Demirbas, 2001). Los enlaces éter existente en la lignina también presentan hidrólisis y pirólisis (Katritzky, 1990c); extendiéndose dicho comportamiento a ácidos y a aldehídos fenólicos (Demirbas, 2001) que constituyan algún polímero.

### 2.4. Conversión hidrotérmica de polipéptidos

En las proteínas, los aminoácidos pueden aislarse al causar hidrólisis de enlaces peptídicos (Braverman, 1980). Un ambiente hidrotérmico es, por supuesto apto para ocasio-

nar esta despolimerización. En la práctica se ha encontrado un punto óptimo de presión y temperatura de 543 K y 5,583 kPa (Quitain, 2001) para realizarlo y favorecer la degradación (Sato, 2004; Kang, 2001). Los polipéptidos y proteínas resultan licuados y gasificados como aminas, ácidos orgánicos, amoníaco y dióxido de carbono (Sato, 2004).

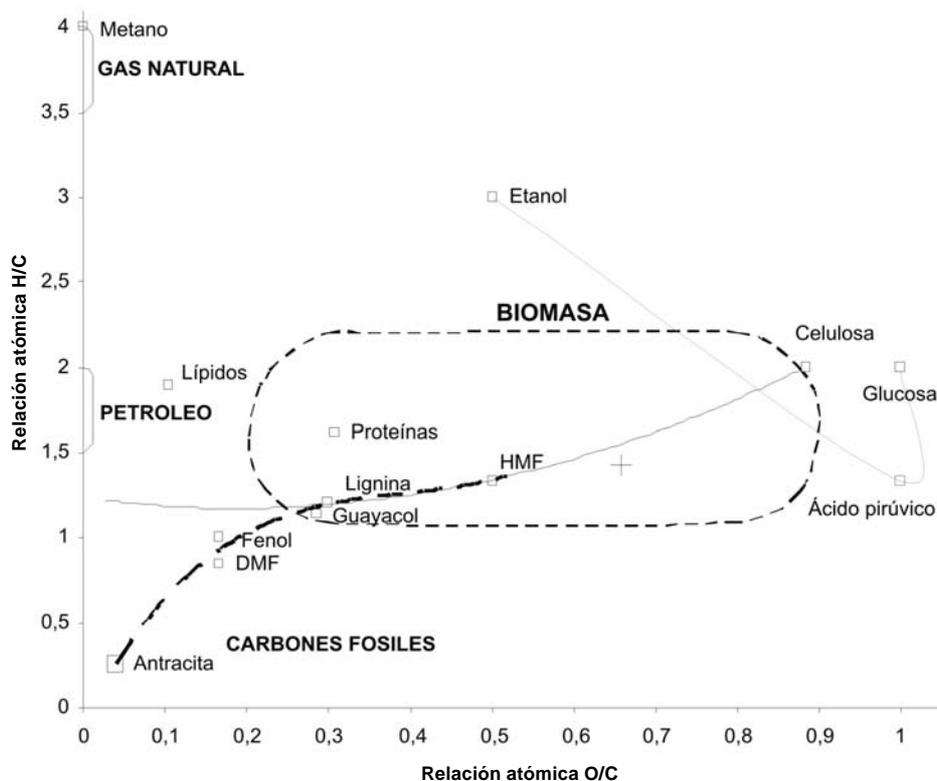
La glicina se ha propuesto como compuesto modelo de las proteínas (Peterson, 2006), ya que posee la estructura básica de partida para otros aminoácidos. Para tales sustancias, la primera reacción es la hidrólisis (Shaw, 1991). Posteriormente, cada aminoácido u oligopéptido puede descarboxilarse para producir dióxido de carbono y aminas; o presentar desaminación para generar amoníaco y ácidos orgánicos (Sato, 2004). El amoníaco formado durante la hidrólisis y desaminación cataliza la reacciones de descarboxilación (Katritzky *et al.*, 1990). Los aminoácidos con estructuras piridínicas pueden degradarse si contienen sustituyentes aldehído, hidroxilo y carboxilo (Katritzky *et al.*, 1990a. Katritzky *et al.*, 1990b).

## 3. Esquemas generalizados de reacciones

Las moléculas, producto de conversión hidrotérmica, siguen un patrón de evolución regido por la intensidad de la descarboxilación como se muestra en la Figura 2, en la cual se recopilan las relaciones atómicas oxígeno/carbono (O/C) e hidrógeno/carbono (H/C) de dichas moléculas y de los polímeros que les dieron origen. La evolución de los productos de conversión hidrotérmica desde celulosa hacia biocrudos se muestra en línea continua y los carbonizados en línea punteada. La celulosa parte de una relación atómica (O/C) cercana a 0.9 (línea gris continua); si se somete a fermentación, esta relación disminuye (etanol). Durante la fermentación se tienen intermedios ricos en oxígeno como glucosa y ácido pirúvico. Por el contrario, la conversión hidrotérmica de celulosa reduce la relación (O/C) desde el comienzo de la reacción y en cada una de las etapas intermedias (HMF y DMF). La lignina se comporta de la misma manera, pero su punto de partida es una relación (O/C) más baja.

Cuando los productos presentan una relación (H/C) igual a uno se tienen biocrudos. Las reacciones de pirólisis conducen a la producción de gases (metano) y a reacciones de repolimerización que, aunque reducen la relación (O/C), también disminuye la relación (H/C) a valores por debajo de 1 y los productos muestran semejanzas con la composición de los carbones.

Estas trayectorias coinciden con las que presentaría un conjunto de macromoléculas orgánicas (biomasa) sujeta a una evolución geológica en condiciones anaerobias. Puede inferirse que la conversión hidrotérmica (licuefac-



+ Relación O/C-H/C promedio típica de la biomasa calculada a partir de Obenberger (2005)

**Figura 2.** Representación (tipo Van Krevelen) de los principales productos de conversión hidrotérmica de biomasa.

ción y gasificación) de biomasa se considera como un proceso semejante pero acelerado para producción de combustibles fósiles. Sin embargo, a pesar de la similitud de las propiedades fisicoquímicas de un biocruído respecto a las del petróleo, este producto requiere estabilización y redistribución de enlaces hidrógeno y carbono-carbono mediante una etapa de hidroxigenación.

Para el caso de la conversión hidrotérmica de polisacáridos, la reducción del valor O/C es más pronunciada que la de la relación H/C; situación favorable para el propósito de crear un conjunto de moléculas similares a los hidrocarburos, pero el descenso en la relación atómica H/C hace que parte de los productos de licuefacción tengan también afinidad elemental con los alquitranes, pizarras bituminosas y carbones. En el evento que una fuente de biomasa posea macromoléculas con una relación atómica O/C inferior a 0,2, las reacciones que ocasionan gasificación hidrotérmica tienen el potencial de generar biocruídos con relaciones atómicas H/C entre 1 y 2, siempre y cuando se cumpla la similitud de trayectoria que presenta la conversión hidrotérmica de celulosa. En esta área se

encuentran los lípidos, pero su costo de adquisición es elevado. Un análisis rápido permite calcular que los aceites vegetales adquirirían competitividad como insumo de procesos de conversión hidrotérmica para producción de biocruídos si el petróleo alcanza un valor de US\$100/barril. Sin embargo esta cota no está muy lejana y, día a día, la economía conduce a esas posibilidades.

Para la biomasa de tipo lignocelulósico, la conversión hidrotérmica es rentable. El inicio de actividades exploratorias en este campo requiere un estudio preliminar de los mecanismos de reacción. Si los polisacáridos fueran las únicas moléculas en la biomasa, su ruta de conversión hidrotérmica general puede ser la esquematizada en la Figura 3. El principal producto de licuefacción lo constituyen furfural, siendo predominante el hidroximetil-furfural (HMF) (Srokol, 2004). El único producto de gasificación es dióxido de carbono. Tal gas no procede de un proceso de combustión sino de una transformación similar a la respiración o el metabolismo; por lo tanto, las emisiones de dióxido de carbono generadas durante la conversión hidrotérmica de polisacáridos no puede considerarse como

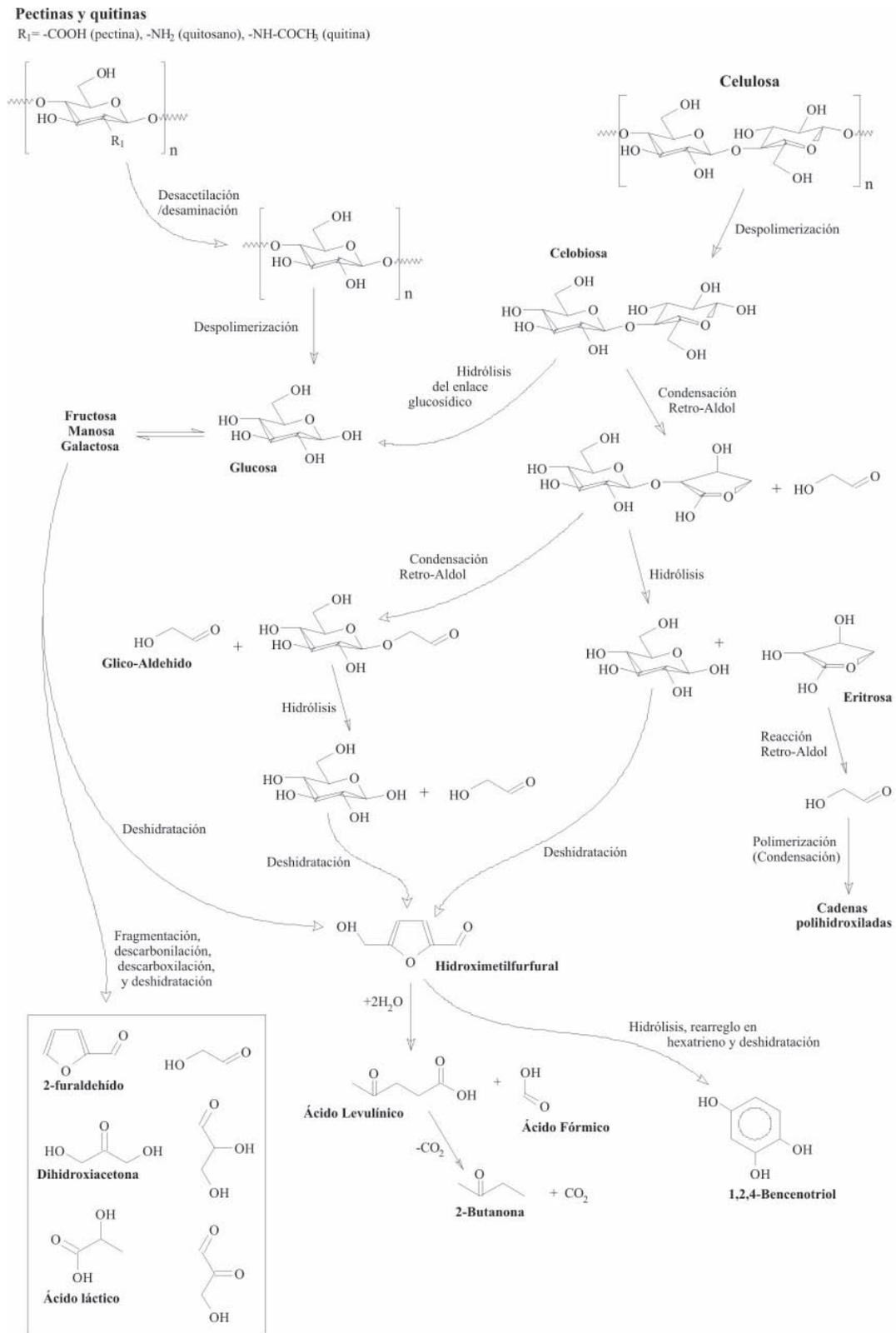
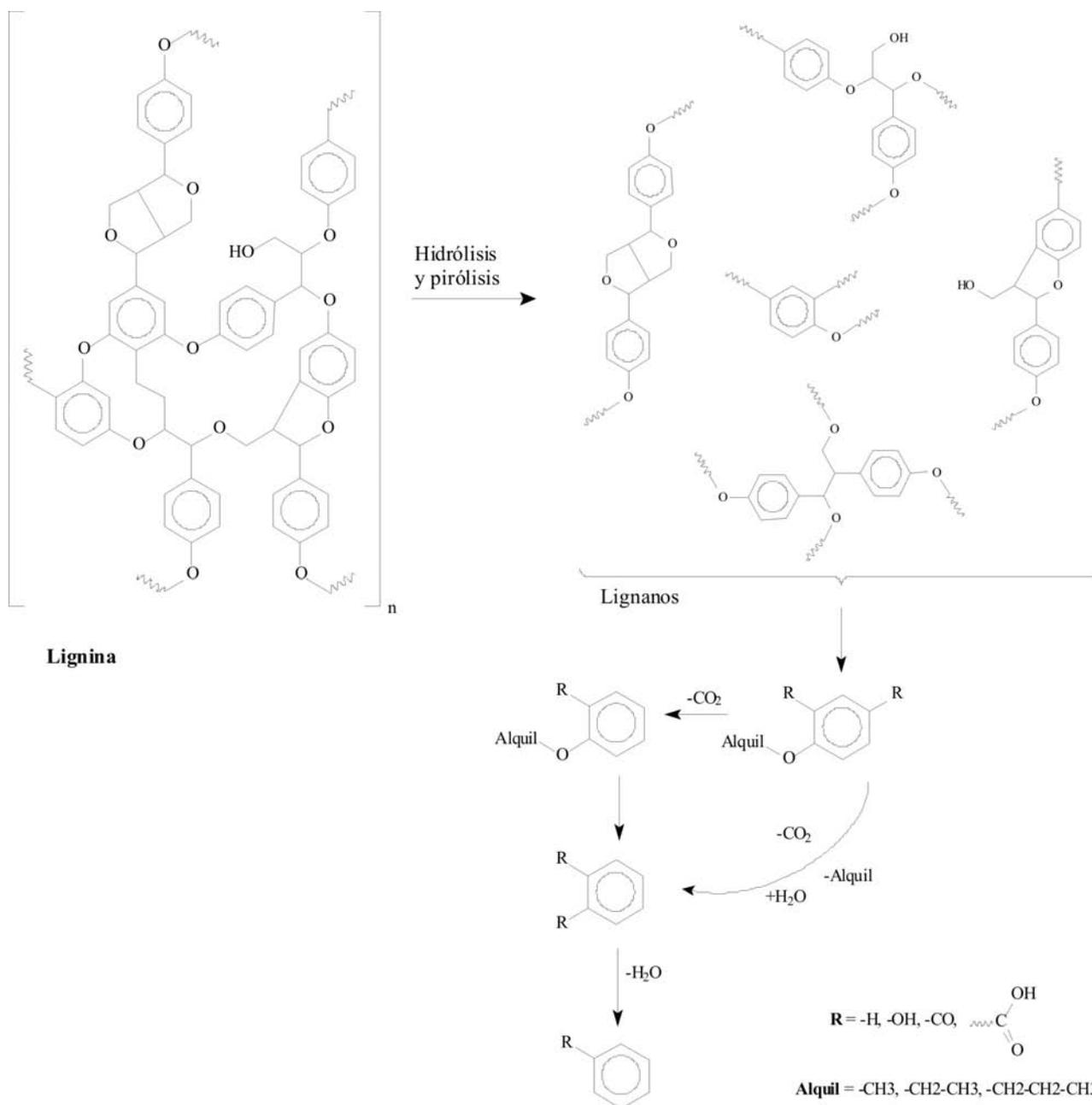


Figura 3. Esquema global de reacciones de la celulosa al ser sometida a conversión (licuefacción-gasificación) hidrotérmica.

contribuyente al efecto invernadero y adicionalmente la cantidad emitida siempre será menor que la que debió fijarse en la biomasa cuando se formó.

El estadio inicial de reacción es la pérdida de los grupos funcionales por hidrólisis y pirólisis cuando las sustancias o tejidos involucrados en conversión hidrotérmica son polisacáridos con funcionalidades. De acuerdo con la

literatura (Quitain, 2001), los principales productos serán moléculas pequeñas, y en la mayoría de los casos el resultado es gasificación de los grupos removidos (metano, dióxido de carbono, amoníaco). El desarrollo de las reacciones de conversión hidrotérmica posteriores suceden tal como ocurre con los polisacáridos sin funcionalidades (Srokol, 2004). En la Figura 4 se esquematiza este comportamiento para las pectinas y la quitina. El mismo comporta-



**Figura 4.** Esquema de reacción generalizado para monómeros de lignina y derivados fenólicos sometidos a conversión hidrotérmica subcrítica (Estructuras básicas de lignanos adaptadas de Vivas et al., 2006).

miento puede postularse para sustituyentes sulfatados como queratanos o el polisacárido funcionalizado que sea objeto de ensayo.

La lignina y los lignanos presentan todos los tipos de reacción pertinentes a especies con grupos éter oxigenados. Debido a su naturaleza, los principales productos a esperar son lignanos y derivados fenólicos. El esquema generalizado de reacción hidrotérmica de sustancias con estructura de tipo lignina se muestra en la Figura 4. Las rutas de reacción planteadas exigen un ambiente carente de oxígeno, ya que en el caso contrario, la conversión hidrotérmica de lignina producirá exclusivamente fenoles, los cuales se oxidan hacia ácidos carboxílicos y productos carbonizados (Boukis, 2004) y la conversión puede desviarse a gasificación exclusivamente, con generación de dióxido y monóxido de carbono.

En la Figura 5 se propone un esquema generalizado para las reacciones de conversión hidrotérmica de un polipéptido. El sustituyente  $-R$  puede ser un grupo fenilo, piridina, isopropilo, hidroxilo, carboxilo, etc. Únicamente se han considerado las rutas de desaminación y descarboxilación. Algunas proteínas con enlaces sulfuro y sulfhidrilo presentarán desprendimiento de ácido sulfhídrico al instante de ocurrir la hidrólisis. Sin embargo, se encuentran casos reportados en los que el ácido se reduce hasta azufre elemental (Baskis, 1993), y la gasificación de amoníaco puede ser mitigada por adición de agentes acomplejantes/mineralizantes, de forma que éste se puede recuperar como sales de amonio (Fröling, 2004; 2005). Debido a la gran cantidad de grupos carboxilo contenidos en una proteína (al menos uno por péptido) la gasificación hidrotérmica se manifiesta con desprendimiento de dióxido y monóxido de carbono.

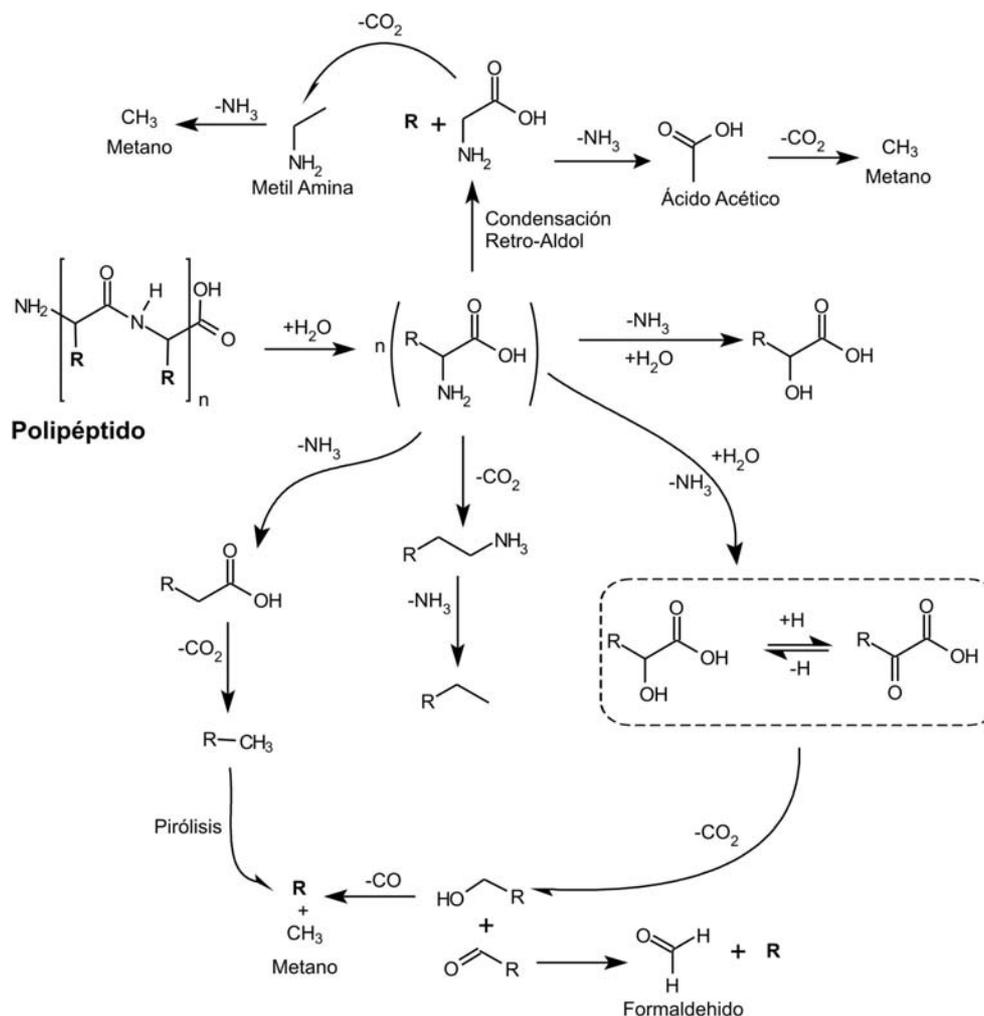


Figura 5. Esquema generalizado de rutas de conversión hidrotérmica de un polipéptido.

no. Nuevamente, las consideraciones hechas respecto de la naturaleza de su emisión para los polisacáridos son aplicables a los polipéptidos. Estos gases no deben considerarse como emisiones que contribuyan al aumento del efecto invernadero.

En todos los tipos de biomasa hay confluencia de las principales macromoléculas discutidas en este artículo; si bien en la biomasa vegetal predominan los polisacáridos (Dumitriu, 2005), y los lípidos solo son considerables en pocas plantas especializadas, la existencia e inclusión de trazas o bajos porcentajes de lípidos en sistemas de conversión hidrotérmica es benéfica, pues su intervención en las demás reacciones de conversión ocasiona un aumento en el rendimiento de productos líquidos, lo cual supone que los lípidos inhiben la gasificación sin afectar la conversión o rutas de reacción (Kruse, 2005). De confirmarse esta proposición, podría considerarse la aplicación de conversión hidrotérmica a biomasa con contenido de lípidos, en mezcla con sales de potasio para desfavorecer la carbonización (Schmieder, 2000) y un medio de reacción acidulado (Siskin, 1995) para producir bio-crudos con alto rendimiento y poca gasificación.

La posibilidad de que se formen macromoléculas con entrecruzamientos internos (carbonizados) y heterociclo y policiclos voluminosos (alquitranes) puede atacarse por medio del empleo de las tres condiciones mencionadas en el párrafo anterior. Si bien es cierto que las sales de potasio favorecen la gasificación hidrotérmica, este inconveniente es útil si este tipo de catalizadores se emplea en proporciones identificadas, para evitar la condensación de radicales no estabilizados y contribuir a la despolimerización de policiclos creados por repolimerización. Se debe recordar que en estas condiciones la recombinación de los radicales orgánicos se hace reversible y las macromoléculas, sus monómeros e incluso algunas moléculas degradadas se encuentran en equilibrio (Sawada, 1976).

#### 4. Conclusiones

Los principales productos de conversión hidrotérmica subcrítica de la biomasa y residuos de biomasa son los furfurales y los ácidos carboxílicos. Esta situación viene precedida por la proporción en que se presenta la celulosa y sustancias de similar estructura en la composición misma de los materiales de partida. Se sugiere que moléculas como el hidroximetil furfural (HMF) y sus productos de reducción y oxidación, tienen el potencial de ser insumos industriales, ya sea como precursores químicos o como aditivos para nuevos combustibles.

Fuentes de biomasa y de residuos son abundantes en el país, tales como las maderas, el bagazo de caña, los pastos, o los desechos domésticos y estos materiales como cualquier material orgánico son susceptibles de someterse a conversión hidrotérmica para la obtención de biocrudos. Surgen modelos de mecanismos de reacción para cada material, pero la existencia de combinaciones de tipos de biomasa origina modificaciones en dichos mecanismos.

La energía contenida en la biomasa presente en residuos y vertimientos puede contribuir a cubrir la creciente demanda energética de los procesos productivos. Su conversión hidrotérmica hacia combustibles líquidos, además de reducir las tasas retributivas o costos de disposición, tiene el potencial de agregar valor a los residuos, proporcionando una forma de energía alternativa y un mecanismo progresivo de destoxificación del hábitat humano.

Los compuestos obtenidos por conversión hidrotérmica de biomasa tienen un valor agregado que supera al valor de la materia prima en, al menos, 4 a 6 veces su valor de adquisición, pero la idea de obtener estos productos a partir de residuos orgánicos y de residuos de biomasa es sin duda seductora, dada su ubicuidad y bajo o prácticamente nulo costo de adquisición. Adicionalmente, la transformación hidrotérmica (subcrítica o supercrítica) de cualquiera de las familias de macromoléculas abordadas en este documento contribuyen a la reducción de emisión de gases de invernadero (CO<sub>2</sub>) pues nunca se podrá emitir una cantidad mayor de estos gases que la que han fijado por fotosíntesis la biomasa y residuos de biomasa que alimentan el proceso.

#### 5. Agradecimientos

Los autores agradecen a la Universidad Nacional de Colombia por el apoyo a esta investigación a través del proyecto "Combustibles líquidos por procesos de conversión hidrotérmica de biomasa y residuos de biomasa" código DIB 8003064 y a Colciencias por el apoyo a través del proyecto "Combustibles líquidos por procesos de conversión hidrotérmica supercrítica de biomasa y sus residuos" código 1101-405-20173, contrato 202-2007.

#### Bibliografía

- Annee, J., Ruyter, H. "Process for producing hydrocarbon-containing liquids from biomass". Patente europea EP 0204 354. Oficina Europea de Patentes. Diciembre 10 de 1986.
- Antal, M., Carlsson, M., Xu, X. 1998. "Mechanism and kinetics of the acid-catalyzed dehydration of 1- and 2-propanol in hot compressed liquid water". *Industrial and Engineering Chemistry Research*. **37**(10): 3820-3829.

- Antal, M., Allen, S., Schulman, D., Xu X., Divilio, R.** 2000. "Biomass gasification in supercritical water". *Industrial and Engineering Chemical Research*. **39**: 4040-4053.
- Bakr, M., Akiyama, M., Sanada, Y.** 1990. "ESR assessment of kerogen maturation and its relation with petroleum genesis". *Organic Geochemistry*, **15** (6): 595-599. ISSN: 0146-6380.
- Barth, T., Abbott, G., Pedersen, K.** "Aqueous thermal conversion of biomass to hydrocarbons in the fluid fuel range". Reporte final publicable del contrato JOR3-CT97-0176. Diciembre 1 de 1997 – noviembre 31 de 2000. Programa de energía no nuclear Joule III, Comisión Europea-Statoil.
- Bicker, M., Kaiser, D., Ott, L. et al.** 2005. "Dehydration of d-fructose to hydroxymethylfurfural in sub- and supercritical fluids". *Journal of supercritical fluids*. **36**: 118-126.
- Bobleter, O.** 1986. "Hydrolysis of Cellulose in Dilute Sulfuric Acid and Under Hydrothermal Conditions". *Journal of Carbohydrate Chemistry*. **5**: 387-399.
- Boukis, N. et al.** 2004. "Advances with the process of biomass gasification in supercritical water". Documento internmo Institut für Technische Chemie CPV, Forschungszentrum Karlsruhe. Alemania.
- Braverman, J.** "Introducción a la bioquímica de los alimentos". 3ª Ed. Ediciones Omega. 1980 Barcelona, España. 354 p.
- Brucker, E.** "Iniciación Botánica. Colección de iniciaciones científicas". París, (1912). Editorial E. Brucker.
- Campbell, G., Meluch, W.** 1976. "Polyurethane foam recycling superheated steam hydrolysis". *Environmental Science & Technology*. **10** (2): 182-185.
- Curry, D., Simpler, T.** 1988. "Isoprenoid constituents in kerogens as a function of depositional environment and catagenesis". *Organic Chemistry*, **13** (4-6): 995-1001.
- Deguchi S., Tsujii K., Horikoshi K.** 2006. Cooking cellulose in hot and compressed water, *Chem. Commun.* pp. 3293-3295.
- Demirbas, A.** 2000. "Mechanism of liquefaction and pyrolysis reaction of biomass". *Energy conversion & Management*. **41**: 633-646.
- \_\_\_\_\_. 2001. "Biomass to methanol via pyrolysis process". *Energy conversion & Management*. **42**: 1349-1356.
- \_\_\_\_\_. 2004. "Effect of initial moisture content on the yields of oily products from pyrolysis of biomass". *Journal of Analytical and Applied Pyrolysis*, **71** (2): 803-815.
- Dumitriu, S.** (Editor) "Polysaccharides: Structural Diversity and Functional Versatility" Marcel Dekker, New York, U.S.A. 2005.
- Franck, E.** 1987. "Fluids at high pressures and temperatures. *Journal of Chemical Thermodynamic*. **19**: 225-242.
- Fröling, M., Peterson, A., Tester, J.W.** "Turkey processing wastes to diesel oil- Investigating the environmental performance of a hydrothermal process for transferring food industry waste into useful products." Proceeding of the 10th European Biosolids and Biowastes Conference, West Yorkshire, Reino Unido, Noviembre de 2004.
- Fröling, M., Peterson, A., Svanström, M., Tester, J.W.** "Choosing between green innovative technologies: Hydrothermal processing of biowastes". Proceedings of the International Society for Industrial Ecology: 2005 Conference, Estocolmo, Suecia (junio de 2005).
- González, G.; Salvadó, J.; Montané, D.** 2004. "Reactions of vanillic acid in sub- and supercritical water". *Journal of Supercritical Fluids*. **31**: 57-66.
- Goudriaan, F.; Peferoen, D.** 1990. "Liquid fuels from biomass via a hydrothermal process". *Chemical Engineering Science*. **45** (8): 2729-2734.
- Höering T.** 1984. "Thermal reactions of kerogen with added water, heavy water and pure organic substances". *Organic Geochemistry*. **5** (4): 267-278.
- Huang, D.** 1999. "Advances in hydrocarbon generation theory (1): Generation and evolution model for immature oils and hydrocarbon". *Journal of petroleum Science and Engineering*. **22** (1-3): 121-130.
- Jin, F.M.** 2001. "Kinetics of oxidation of food wastes with H<sub>2</sub>O<sub>2</sub> in supercritical water". *The Journal of Supercritical Fluids*. **19**: 251-262.
- Kabyemela, B., et al.** 1999. "Glucose and Fructose Decomposition in Subcritical and Supercritical Water: Detailed Reaction Pathway, Mechanisms, and Kinetics". *Industrial and Engineering Chemical Research*. **38** (8): 2888-2895.
- Kamio, E.; Takahashi, S.; Noda, H. et al.** 2006. "Liquefaction of cellulose in hot compressed water under variable temperatures". *Industrial and Engineering Chemistry Research*. **45**: 4944-4953.
- Kang, Kilyoon et al.** 2001. Optimization of aminoacids production from waste fish entrail by hydrolysis in sub and supercritical water. *Canadian Journal of Chemical Engineering*. **79**: 65-70.
- Katritzky, A., Lapucha, R., Siskin, M.** 1990. "Aqueous High-Temperature Chemistry of Carbo- and Heterocycles. 12. Benzonitriles and Pyridinecarbonitriles, Benzamides and Pyridinecarboxamides, and Benzylamines and Pyridylamines". *Energy & Fuels*. **4**: 555-561.
- \_\_\_\_\_. 1990a. "Aqueous High-Temperature Chemistry of Carbo- and Heterocycles. 3. 2-Substituted Pyridines". *Energy & Fuels*. **4**: 506-510.
- \_\_\_\_\_. 1990b. "Aqueous High-Temperature Chemistry of Carbo- and Heterocycles. 4. 4-Substituted Pyridines". *Energy & Fuels*. **4**: 510-514.
- \_\_\_\_\_. 1990c. "Aqueous High-Temperature Chemistry of Carbo- and Heterocycles. 2. Monosubstituted Benzenes: Benzyl Alcohol, Benzaldehyde, and Benzoic Acid". *Energy & Fuels*. **4**: 499-505.
- Kiennemann, A.** *Comunicación personal (Entrevista)*. Profesor-Investigador, Laboratorio de Materiales, Superficies y Procedimientos para Catálisis- ECPM. Universidad Louis Pasteur. Strasbourg, Francia. Septiembre de 2004. (kiennemann@chimie.u-strasbg.fr)
- Koch B., Christiansen, F.** 1993. "Maturation of lower palaeozoic kerogens from North Greenland". *Organic Geochemistry*. **20** (3): 405-413.
- Koopmans, M., Rijpstra, IM., Leeuwa, J. et al.** 1998. "Artificial maturation of an immature sulfur- and organic matter-rich

- limestone from the Ghareb Formation, Jordan". *Organic Geochemistry*. **28** (7-8): 503-521.
- Kruse, A., Gawlick, A.** 2003. "Biomass Conversion in Water at 330-410°C and 30-50 MPa. Identification of Key Compounds for Indicating Different Chemical Reaction Pathways". *Industrial and Engineering Chemical Research*. **42**: 267-279.
- Kruse, A., Gawlick, A., Henningsen, T.** 2004. "Biomass liquefaction and gasification in near and supercritical water: key compounds as a tool to understand chemistry". Documento. Institut für Technische Chemie CPV, Forschungszentrum Karlsruhe. Alemania.
- Kuhlmann, B., Arnett, E., Siskin, M.** 1994. "Classical Organic Reactions in Pure Superheated Water". *Journal of Organic Chemistry*. **59**: 3098-3101.
- Kurita, K., Akao, H., Yang, J., Shimojoh, M.** 2003. "Nonnatural branched polysaccharides: synthesis and properties of chitin and chitosan having disaccharide maltose branches". *Biomacromolecules*. **4**: 1264-1268.
- Landes, K.** "Geología del petróleo". 3ª Ed. Ediciones Omega. (1977) Barcelona, España. 463 p.
- Lorant, F., Behar, F.** 2002. "Late Generation of Methane from Mature Kerogens". *Energy & Fuels*. **16** (2): 412-427.
- Luijckx, G.** 1993. "Hydrothermal conversion of carbohydrates and related compounds". Tesis para aspirar al título de Doctor en Ingeniería Química, Universidad Tecnológica de Delft. Países Bajos (Holanda).
- Mandoki, J.** "Depolymerization of condensation polymers." Patente norteamericana US 4,605,762. Oficina norteamericana de patentes, Agosto 12 de 1986.
- Minowa, T., Zhen, F., Ogi, T. et al.** 1997. "Liquefaction of cellulose in hot compressed water using sodium carbonate: products distribution at different reaction temperatures. *Journal of Chemical Engineering (Japón)*. **30**: 186.
- Miyazawa, T., Funazukuri, T.** 2004. "Hydrothermal production of mono (galacturonic acid) and oligomers from poly (galacturonic acid) with water under pressures". *Industrial and Chemical Engineering Research*. **43**: 2310-2314.
- Modell, M., Reid, R., Amin, S.** "Gasification process". Patente norteamericana US 4.113.446. Oficina norteamericana de Patentes y registros. Septiembre 12 de 1978.
- Mok, W., Antal, M.** 1992. "Uncatalyzed Solvolysis of Whole Biomass Hemicellulose by Hot Compressed Liquid Water". *Industrial and Engineering Chemistry Research*. **31**: 1157-1161.
- Obenberger, P.** "Proximal analysis of biomass". Version 2,4, © BIOS Bioenergiesysteme GmbH. Graz, Austria. Disponible en Inea en [www.bios-bioenergy.at](http://www.bios-bioenergy.at)
- Peterson, A.** "Studying the detailed chemistry of the hydrothermal reaction of glucose and glycine, which are model compounds for biomass, in a high-pressure and high-temperature environment". Citado en "Curriculum vitae". Disponible en [web.mit.edu/aap/www/CV](http://web.mit.edu/aap/www/CV).
- Piskorz, J. et al.** 2000. "Flash pyrolysis of cellulose for production of anhydro-oligomers". *Journal of Analytical and Applied Pyrolysis*. **56**: 145-166.
- Ponder, G., Richards, G., Stevenson, T.** 1992. "Influence of linkage position and orientation in pyrolysis of polysaccharides: A study of several glucans". *Journal of Analytical and Applied Pyrolysis*. **22** (3): 217-229.
- Quitain, A. et al.** 2001. "Production of valuable materials by hydrothermal treatment of shrimp shells". *Industrial and Engineering Chemistry*. **40**: 5885-5888.
- Rouxhet, P., Robin, P.** 1978. "Infrared study of the evolution of kerogens of different origins during catagenesis and pyrolysis". *Fuel*. **57** (9): 533-540.
- Saisu, M., Sato, T., Watanabe, M. et al.** 2003. "Conversion of lignin with supercritical water-phenol mixtures". *Energy Fuels*. **17**: 922.
- Sakai, K.** 1998. "Function and Application of Chitin-Chitosan Oligomers" (Artículo en japonés). *New Food Industries*. **40**: pp. 17. (Referenciado por Quitain, 2001).
- Saleh, R., Siskin, M., Knudsen, G.** "Process for depolymerization of butyl rubbers and halobutyl rubbers using hot liquid water". Patente norteamericana US 5, 283, 318. Febrero 1 de 1994.
- Sasaki, M. et al.** 2002. "Kinetics and Mechanism of Cellobiose Hydrolysis and Retro-Aldol condensation in Subcritical and Supercritical Water", *Industrial and Engineering Chemistry Research*, **41**: 6642-6649.
- Sasaki, T. et al.** 2002. "Saccharification of Cellulose Using a Hot-Compressed Water-Flow Reactor". *Industrial Engineering Chemical Research*. **41**: 661-665.
- Sato, N. et al.** 2004. "Reaction kinetics of amino acid decomposition in high-temperature and high-pressure water". *Industrial and Engineering Chemical Research*. **43**: 3217-3222.
- Savage, P.** 1999. "Organic Chemical Reactions in Supercritical Water". *Chemical reviews*. **99**: 603-621.
- \_\_\_\_\_. "Comunicación personal". Ingeniero Químico, investigador del Departamento de Ingeniería Química de la Universidad de Michigan. Michigan, Estados Unidos de Norteamérica. ([psavage@umich.edu](mailto:psavage@umich.edu)), agosto de 2005.
- Sawada, H.** 1976. "Thermodynamics of polymerization". Marcel Dekker. New York, Estados Unidos de América.
- Schmieder, H. et al.** 2000. "Hydrothermal gasification of biomass and organic wastes". *The Journal of Supercritical Fluids*. **17**: 145-153.
- Schobert, H.** "The chemistry of hydrocarbon fuels". Butterworth & Co (actualmente Elsevier Science Ltd). Londres, Reino Unido. 1990. ISBN 0-408-03825-X
- Schwarzinger, C.** "Comunicación Personal". Investigador del Instituto para Tecnología Química de Materiales Orgánicos, Universidad Johannes Kepler. Enero de 2005. Linz, Austria. ([clemens.schwarzinger@jku.a](mailto:clemens.schwarzinger@jku.a))
- Serikawa, R., Funazukuri, T., Wakao, N.** 1992. "Oil conversion of vinasse with high density water". *Fuel*. **71**: 283-287.
- Shaw, R. et al.** 1991. *Supercritical Water, a medium for chemistry*. Chemical & Engineering News. **69** (51): 26-39.
- Siskin, M., Katritzky, A., Balasubramanian R.** 1995. "Aqueous organic chemistry. 9. Reactivity of 1,5-, 1,6-, 1,7- and 2,6-

dihydroxynaphthalenes, dibenzofuran, 2-hydroxydibenzofuran, carbazole, 2-hydroxycarbazole, acridine and 4-hydroxyacridine". *Fuel*. **74**: 1509-1511.

**Siskin, M., Katritzky, A.** 2000. "A review of the reactivity of organic compounds with oxygen-containing functionality in superheated water". *Journal of analytical and applied pyrolysis*. **54**: 193-214.

**Siskin, M. et al.** "Aqueous Organic Chemistry. 3. Aquathermolysis: Reactivity of Ethers and Esters". *Energy & Fuels*. **5**: 1990, pp. 488-492.

\_\_\_\_\_. 1990 a. "Aqueous Organic Chemistry. 2. Cross-Linked Cyclohexyl Phenyl Compounds". *Energy Fuels*. **4**: 482-488.

\_\_\_\_\_. 1990 b. "Aqueous organic chemistry. 1. Aquathermolysis: Comparison with thermolysis in the reactivity of aliphatic compounds". *Energy & Fuels*. **4**: 475-482.

**Srokol, Z.** "Hydrothermal liquefaction of biomass: reaction paths and kinetics". Tesis de doctorado en Ingeniería Química, Universidad Tecnológica de Delft. 2004. Países Bajos.

**Stein, S.** "Free radicals in coal conversion, Chemistry of coal conversion", editado por Richard H. Schlosberg. Plenum Press. Nueva York, Estados Unidos de América. 1985.

**Suárez, M.** "Comunicación personal". Bióloga, especialista en microbiología de organismos patógenos. Instituto de Ciencia y

Tecnología de Alimentos. Universidad Nacional de Colombia. abril de 2005. Bogotá, Colombia.

**Szabo, P., Minowa, T., Ogi, T.** 1998. "Liquefaction of cellulose in hot compressed water using sodium carbonate. Part 2: Thermogravimetric mass spectrometric study of the solid residues". *Journal of Chemical Engineering (Japón)*. **31**: 571.

**Veringa, H.** "Advanced Techniques for generation of energy from biomass and waste". Document ECN-Biomass. Países Bajos (Holanda). Octubre de 2004.

**Vivas N., Nonier M-F., Pianet I., Vivas de Gaulejac N., Fouquet È.** 2006. Structure of extracted lignins from oak heartwood (*Quercus petraea* Liebl., *Q. Robur* L.); *Comptes Rendus Chimie*; **9**: 1221-1233.

**Yan, Y., Xu, J., Li, T. et al.** 1999. "Liquefaction of sawdust for liquid fuel". *Fuel Processing Technology*. **60**: 135-143.

**Zhong, C., Peters, C., De Swaan Arons, J.** 2002. "Thermodynamic modeling of biomass conversion processes". *Fluid Phase Equilibria*. **194-197**: 805-815.

Recibido: agosto 31 de 2006

Aceptado para su publicación: noviembre 16 de 2008

# HISTOLOGÍA COMPARADA DEL SISTEMA REPRODUCTIVO DE CUATRO ESPECIES DE RAYAS CARIBEÑAS (CHONDRICHTHYES: BATOIDEA)

Por

Arturo Acero P.<sup>1</sup>, Marcela Grijalba-Bendeck<sup>2</sup>, Fabián Moreno<sup>2</sup>, Kelly Acevedo<sup>2</sup> & Esperanza González<sup>3</sup>

## Resumen

**Acero, A., M. Grijalba-Bendeck, F. Moreno, K. Acevedo & E. González:** Histología comparada del sistema reproductivo de cuatro especies de rayas caribeñas (Chondrichthyes: Batoidea). Rev. Acad. Colomb. Cienc. **32**(125): 571-586, 2008. ISSN 0370-3908.

Con el fin de comparar histológicamente los sistemas reproductivos de diferentes linajes de batoideos se examinaron 93 ejemplares (62 hembras y 31 machos) de cuatro especies: *Narcine bancroftii* (familia Narcinidae), *Rhinobatos percellens* (familia Rhinobatidae), *Urotrygon venezuelae* (familia Urotrygonidae) y *Dasyatis guttata* (familia Dasyatidae). Se diferenciaron tres tipos de sistemas femeninos: *R. percellens* ejemplifica un sistema generalizado; *U. venezuelae* y *D. guttata* muestran un ovario que alberga ovocitos con estructuras internas únicas; *N. bancroftii* posee el ovario más derivado, en cuanto a la presencia y forma de estructuras exclusivas. Sin embargo, al ser las características claramente autapomórficas no es posible plantear hipótesis de relaciones a partir de los sistemas femeninos. Las diferencias son relativamente menos profundas en los sistemas masculinos, con *N. bancroftii* mostrando un testículo generalizado, mientras *R. percellens* presenta cierto grado de lobulación externa; *D. guttata* presenta un testículo de tipo Radial, que en *U. venezuelae* aparece modificado.

**Palabras clave:** Sistema reproductor, *Narcine bancroftii*, *Rhinobatos percellens*, *Urotrygon venezuelae*, *Dasyatis guttata*.

## Abstract

In order to compare histologically the reproductive systems of four different batoid lineages 93 specimens (62 females and 31 males) of four species were examined: *Narcine bancroftii* (family

- 
- 1 Universidad Nacional de Colombia sede Caribe, CECIMAR/INVEMAR, Cerro Punta de Betín, Santa Marta, Colombia. Correo electrónico: aacero@invemar.org.co
  - 2 Programa de Biología Marina, Universidad Jorge Tadeo Lozano, Mundo Marino, El Rodadero, Santa Marta, Colombia. Correos electrónicos: marcela.grijalba@utadeo.edu.co, fabian.morenorodriguez@gmail.com, kelly.acevedo@utadeo.edu.co
  - 3 Facultad de Ciencias Naturales, Universidad Jorge Tadeo Lozano, Bogotá D.C., Colombia.

Narcinidae), *Rhinobatos percellens* (family Rhinobatidae), *Urotrygon venezuelae* (family Urotrygonidae) and *Dasyatis guttata* (family Dasyatidae). Three types of female systems were differentiated: a generalized one, represented by *R. percellens*, one with ovaries including unique internal structures, such as those of *U. venezuelae* and *D. guttata*, and a highly derived one, with uniquely shaped internal structures, in *N. bancroftii*. However, those features are autapomorphic precluding the establishment of any hypothesis of relationships. Differences between male systems are less marked; testicles in *N. bancroftii* are generalized, while they show some degree of external lobulation in *R. percellens*; Radial testicles appear in *U. venezuelae* and *D. guttata*, which are modified in the former species.

**Key words:** reproductive system, *Narcine bancroftii*, *Rhinobatos percellens*, *Urotrygon venezuelae*, *Dasyatis guttata*.

## Introducción

El taxón Batoidea [considerado cohorte por **McEachran & Aschliman** (2004) y subdivisión por **Nelson** (2006)], un grupo natural de peces cartilaginosos elasmobránquios, incluye casi 550 especies, agrupadas en algo más de 70 géneros (**Nelson**, 2006) y conocidas casi universalmente en español como rayas. No existe, sin embargo, unanimidad en cuanto al número de taxones familiares y supra-familiares incluidos en Batoidea ni mucho menos acerca de las relaciones entre ellos. Mientras **Compagno** (1999) presenta una clasificación prolífica en número de órdenes (seis) y de familias (21), **McEachran & Aschliman** (2004) y **Nelson** (2006) coinciden en la aceptación de sólo cuatro órdenes (Torpediniformes, Pristiformes, Rajiformes y Myliobatiformes), pero desacuerdan acerca de cuantas familias de rayas deben ser reconocidas; **McEachran & Aschliman** (2004) aceptan sólo 14 familias mientras **Nelson** (2006) reconoce 17.

En las plataformas continentales caribeñas existen diez familias de rayas, representativas de cuatro órdenes (**Nelson**, 2006; **Mejía-Falla et al.**, 2007). Sin embargo, el precario estatus de conservación de los peces sierra (familia Pristidae), únicos representantes del orden Pristiformes (**Acero et al.**, 2002), impide completamente cualquier estudio biológico de esos peces, por lo cual no es posible realizar estudios comparativos de todos los taxones involucrados. Recientemente se ha desarrollado un esfuerzo importante para conocer la biología reproductiva y otros aspectos biológicos y ecológicos de los batoideos comunes en el Caribe colombiano (**Téllez et al.**, 2006; **Acevedo et al.**, 2007; **Grijalba-Bendeck et al.**, 2007a; 2007b; 2008; **Moreno et al.**, en prensa). Dentro de ese gran objetivo, se presenta aquí una visión comparada a nivel histológico de los procesos de maduración gonadal de las rayas de aguas someras del Caribe, cotejando cuatro especies comunes en la costa norte de Colombia: un Torpediniformes, *Narcine*

*bancroftii* (Griffith & Smith, 1834) (familia Narcinidae), un Rajiformes *Rhinobatos percellens* (Walbaum, 1792) (familia Rhinobatidae) y los Myliobatiformes *Urotrygon venezuelae* Schultz, 1949 (familia Urotrygonidae) y *Dasyatis guttata* (Bloch & Schneider, 1801) (familia Dasyatidae).

## Metodología

Se colectaron ejemplares de las cuatro especies de rayas en los sectores de playa Salguero (11°10'-11°N, 74°13'-14°W) y Don Jaca (11°06'-07°N, 74°13'-14°W), Santa Marta, entre agosto 2005 y octubre 2007, con muestreos mensuales. Los especímenes se disectaron ventralmente para extraer los principales órganos que conforman el sistema reproductor, los cuales fueron fijados en paraformaldehído al 10% (**Snelson et al.**, 1988) y luego conservados en una solución de alcohol isopropílico al 40 % (**Natanson & Cailliet**, 1986).

Para el proceso histológico los órganos fijados (hembras: ovario, glándula nidamental, oviducto, útero y ovocito; machos: testículo, epidídimo y vesícula seminal) se lavaron con agua corriente para eliminar el exceso de fijador. De cada órgano se realizaron cortes en la parte anterior, media y posterior, longitudinales y transversales, abarcando parcial o totalmente la estructura. El tejido se sometió a las técnicas de deshidratación, impregnación, penetración, corte y coloración en sucesivas soluciones de hematoxilina-eosina de Harris propuestas por **Martoja & Martoja-Pierson** (1970), estandarizadas para peces cartilaginosos por el Grupo de Investigación en Peces del Caribe GIPECA. Se realizaron descripciones macroscópicas (Tabla 1) e histológicas (Tabla 2) de las estructuras del sistema reproductor en machos y hembras durante cada una de las etapas del desarrollo gonadal (**Grijalba-Bendeck et al.**, 2007a; 2007b).

**Tabla 1.** Escala de desarrollo gonadal macroscópico adaptada para peces cartilagosos.

<b>HEMBRAS</b>	
<b>I</b>	Inmadura: Sin desarrollo ovárico, ovarios pequeños de textura acuosa, oviductos delgados poco diferenciados.
<b>II</b>	En maduración: Ovarios delgados, consistentes y de apariencia uniforme, sin ovocitos visibles, oviductos delgados y diferenciables.
<b>IIIa</b>	Madura sin embriones: Ovario con ovocitos visibles, oviductos gruesos y útero parcialmente ensanchado.
<b>IIIb</b>	Madura con embriones: Ovarios con ovocitos, embriones ubicados dentro de los úteros, que se encuentran ensanchados, ocupando gran parte de la cavidad visceral.
<b>IV</b>	Maduras Postparto: Útero ensanchado con paredes sanguinolentas y flácidas. Abertura cloacal expandida.
<b>MACHOS</b>	
<b>I</b>	Inmaduro: Testículos acuosos, ductos poco diferenciados, cláspers pequeños, flácidos, rifiodón cerrado, ausencia de semen y rotación menor a 360°.
<b>II</b>	En maduración: Testículos consistentes, ductos bien desarrollados, epidídimos enrollados en la porción anterior y media, del ducto eferente, cláspers aumentan de tamaño y están parcialmente calcificados.
<b>III</b>	Maduro: Testículos con lóbulos prominentes indicativos de activa producción espermática, conducto eferente se extiende adyacente a la columna vertebral, epidídimo enrollado a lo largo de todo el conducto hasta llegar a la vesícula seminal, cláspers totalmente calcificados, con rifiodón abierto y presencia de semen. Rotación 360°.

**Tabla 2.** Escala de desarrollo gonadal microscópico generalizado para rayas marinas.

<b>HEMBRAS</b>	
<b>Estado 1</b>	Ovocito con el núcleo y nucléolo visibles, este último se puede encontrar hacia uno de los bordes de la membrana nuclear y pueden ser entre uno y seis. El ovocito está rodeado de epitelio plano simple, que se transforma en cúbico simple y capas de tejido conjuntivo.
<b>Estado 2</b>	Los ovocitos están rodeados de la zona pelúcida (membrana acelular), y epitelio cúbico simple. Se reconocen granulaciones muy parecidas a vitelo. Aumenta el número de ovocitos de todos los tamaños. En el interior del ovocito se diferencia el núcleo con un nucléolo.
<b>Estado 3</b>	Ovocitos bordeados de la zona pelúcida y células cúbicas alrededor, que se van transformando a columnar simple, externamente se observa un tejido conjuntivo laxo con fibroblastos y fibras de colágeno.
<b>Estado 4</b>	Ovocitos conteniendo gránulos de vitelo en variado estado de desarrollo, rodeados de la zona pelúcida, epitelio columnar y tejido conjuntivo laxo.
<b>MACHOS</b>	
<b>Estado 1</b>	Cistos de forma circular, bordeados de tejido conjuntivo y fibroblastos, alrededor células espermáticas “espermatogonias” también redondeadas, conteniendo células germinativas y cromatina granulosa basófila. También células de Sertoli y tejido conjuntivo, los nucléolos son visibles y de coloración rojiza.
<b>Estado 2</b>	Aumenta la abundancia de espermatogonias, cromatina menos granulosa y adquiere mayor densidad, las células de Sertoli se diferencian por su coloración blanquecina, todavía son visibles los nucléolos de color rojo oscuro.
<b>Estado 3</b>	Los cistos aumentan su diámetro, la cromatina adquiere un aspecto alargado y espiralado, son visibles los núcleos y las células de Sertoli, que son menos numerosas, más claras y se ubican basalmente.
<b>Estado 4</b>	La célula espermática “espermatocono”, reduce su tamaño, la cromatina se hace muy condensada y reticulada, los cistos son grandes y en los bordes siguen ubicadas las células de Sertoli
<b>Estado 5</b>	Las células espermáticas se hacen más pequeñas y levemente alargadas; son visibles las células de Sertoli en los márgenes del cisto, donde se ven separadas y desorganizadas.
<b>Estado 6</b>	Comienza la fase de espermátide, con fuerte reducción celular, la cromatina se alarga y comienza a formar la cabeza y la cola de los espermatozoides, en <i>R. percellens</i> no se diferencia el cuello, las células de Sertoli reducen su material citoplasmático.
<b>Estado 7</b>	La célula espermática continúa haciéndose más pequeña, la cabeza es marcadamente reducida y es evidente el alargamiento y adelgazamiento de la cola; comienza el alineamiento de las células al interior del cisto. Algunas células de Sertoli siguen siendo conspicuas, otras se han perdido.
<b>Estado 8</b>	La célula espermática adquiere la denominación de “espermatozoide”, las cabezas se observan muy pequeñas y contienen la cromatina muy condensada a manera de puntos oscuros, la cola es larga y casi recta, el esperma continúa la organización por paquetes y alrededor se diferencian bien las células de Sertoli.
<b>Estado 9</b>	Los espermatozoides están completamente desarrollados, en <i>R. percellens</i> las cabezas son diminutos puntos oscuros de las que se desprenden colas rizadas. Los espermatozoides forman paquetes denominados espermatozuogmata. Las cabezas espermáticas se orientan a través de la periferia del espermatocisto y se embeben en la región apical de las células de Sertoli; cada una se asocia a un paquete de esperma cuya ruptura inicia la liberación del esperma (espermación).

## Resultados

Se examinó un total de 93 ejemplares de las cuatro especies de batoideos. Dicha suma es discriminada así: 16 ejemplares de *N. bancroftii* entre 196 y 328 mm de ancho del disco (AD) (13 hembras, tres machos), 24 *U. venezuelae* (14 hembras, diez machos) entre 66 y 167 mm AD, 29 *R. percellens* (17 hembras, doce machos) entre 66 y 245 de longitud total y 24 *D. guttata* (18 hembras, seis machos) entre 413 y 1710 mm AD. En cada especie los ejemplares analizados estuvieron distribuidos en todo el espectro de tallas.

El resumen de los principales resultados encontrados para las 62 hembras de las cuatro especies estudiadas se presenta en la Tabla 3. Se determinó la estructura histológica de los ovarios de todas las especies. La pared ovárica fue observada en tres de las especies estudiadas, con presencia de epitelio cúbico simple y tejido conjuntivo laxo; sin embargo, en *U. venezuelae* no se pudo detallar esa estructura. Se encontró presencia de gránulos en las cuatro especies objeto (Figura 5), siendo más abundantes en los ejemplares inmaduros. En todas las especies se estudió en detalle al menos el estado IIIa de madurez de los ovarios, caracterizado por presentar ovocitos con estados medios a finales de desarrollo y vitelo abundante (Figura 6), llegándose en el caso de *N. bancroftii* a detallar la situación en cuatro estados de madurez. Así mismo, se halló en las especies estudiadas, excepto *D. guttata*, el estado de desarrollo IIIb (Figura 7), caracterizado por la presencia de embriones uterinos. En *N. bancroftii* los ovocitos del estado IIIb atraviesan un proceso de reabsorción. En *U. venezuelae* no se observó reabsorción, sino que se hallaron ovocitos en estados medios a finales de desarrollo, lo cual podría indicar que ellos funcionan como un lote de reserva. Además, para el caso de *N. bancroftii*, se detectó el estado ovárico IV (postparto) (Figura 8), caracterizado por su elevada irrigación y carencia de ovocitos.

Se detectaron entre cuatro (condición usual) y siete estados de desarrollo de los ovocitos con diámetros entre 0,10 y 5,86  $\mu\text{m}$ , según la especie; se destaca que en *N. bancroftii* se produce una elongación secundaria de los ovocitos, que llegan a medir hasta 19,81  $\mu\text{m}$ . En cuanto a los oviductos, sus paredes están, en general, conformadas por tejido conjuntivo y muscular liso, mientras que los canales en todas las especies, excepto en *R. percellens*, están recubiertos de epitelio columnar pseudoestratificado ciliado. La glándula nidamental, por su parte, sólo fue detectada y estudiada en dos especies, *R. percellens* y *U. venezuelae*. Finalmente, el útero de las cuatro rayas fue analizado en detalle, encontrándose que en general la pared es de tejido muscular liso y conjuntivo, irrigada por venas, arterias, vasos y capilares.

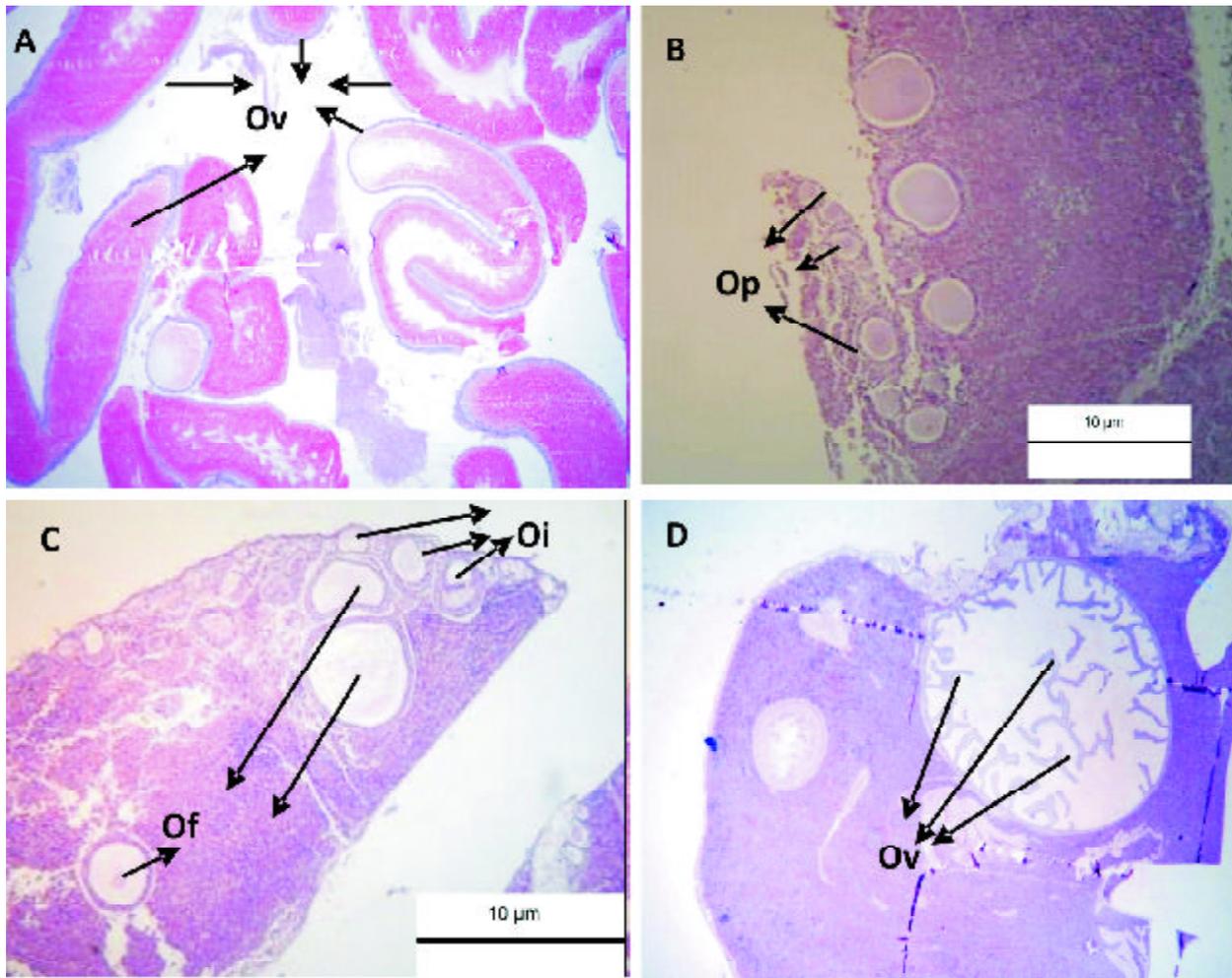
La información recopilada tras estudiar los cortes efectuados en las gónadas de los 31 machos detallados se presenta en la Tabla 4. Los testículos de las cuatro especies son del tipo compuesto y se realizaron cortes representativos de condiciones inmaduras, en maduración y maduras. Así mismo, para todas las especies se detectaron siete estados de desarrollo espermático y fue posible el estudio del órgano epigonal. El epidídimo fue localizado en sólo tres especies, no en *N. bancroftii*, mientras que únicamente se describió la vesícula seminal de *R. percellens* y *U. venezuelae*.

## Discusión

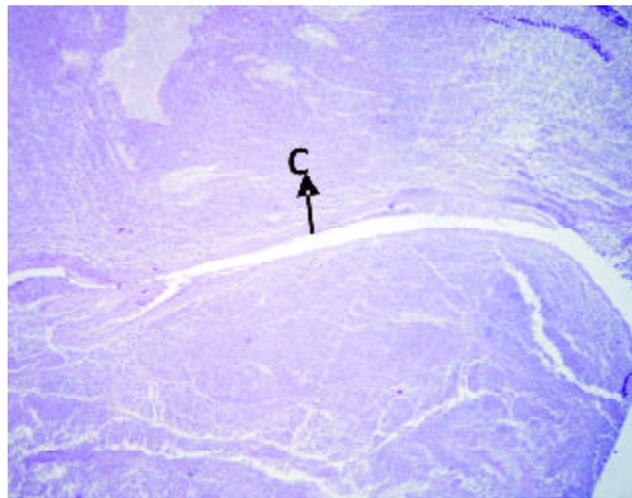
Los ovarios de dos de las especies estudiadas (*R. percellens* y *U. venezuelae*) coinciden en su carácter mixto (en este trabajo llamado Compuesto), pues poseen atributos que definen los tipos descritos por **Pratt** (1988), es decir presencia de ovocitos externos (Figura 1) y de un canal (Figura 2) para su salida. El ovario de *D. guttata* encajaría aparentemente con el tipo Externo de **Pratt** (1988), pues no se detectó un canal. Sin embargo, considerando el tamaño relativamente mayor de esta especie, es posible que los cortes no hayan caracterizado en su totalidad la gónada; así mismo, es factible que **Pratt** (1988) hubiese tenido problemas para ubicar el canal central en el ovario de *Prionace glauca*, su especie modelo, debido al gran tamaño de ese carcharhínido.

Por otro lado, el ovario de *N. bancroftii* no se asemeja a ninguno de los tipos descritos por **Pratt** (1988) ni con el Compuesto. El ovario de ese Torpediniformes, aquí denominado Atípico, se caracteriza por la presencia de varias estructuras tubulares, aguzadas en sus extremos, plegadas dentro del ovario (**Moreno et al.**, en prensa). Además, el ovario de esta especie presenta un conducto interno para la salida de los ovocitos. Este tipo de ovario no ha sido descrito anteriormente ni por **Pratt** (1988), en su descripción general de las gónadas de los elasmobranquios, ni por **Villavicencio** (1995), en su trabajo sobre la especie mexicana *N. entemedor*. Se considera que este tipo de estructura puede constituir un estado derivado, autapomórfico, cuya presencia debe ser corroborada en otras rayas eléctricas.

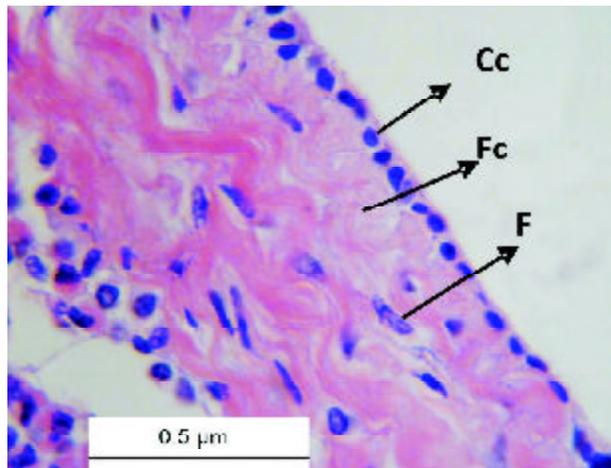
Se encontraron tres tipos distintos de la composición estructural histológica de los ovocitos. La estructura relativamente más sencilla parece ser la de *R. percellens*, cuyo ovario se caracteriza por el bajo número de estados, con cambios en la pared y aparición de células secretoras y de gránulos de vitelo en el estado más avanzado (Figura 8). Las dos especies de Myliobatiformes estudiadas coinciden



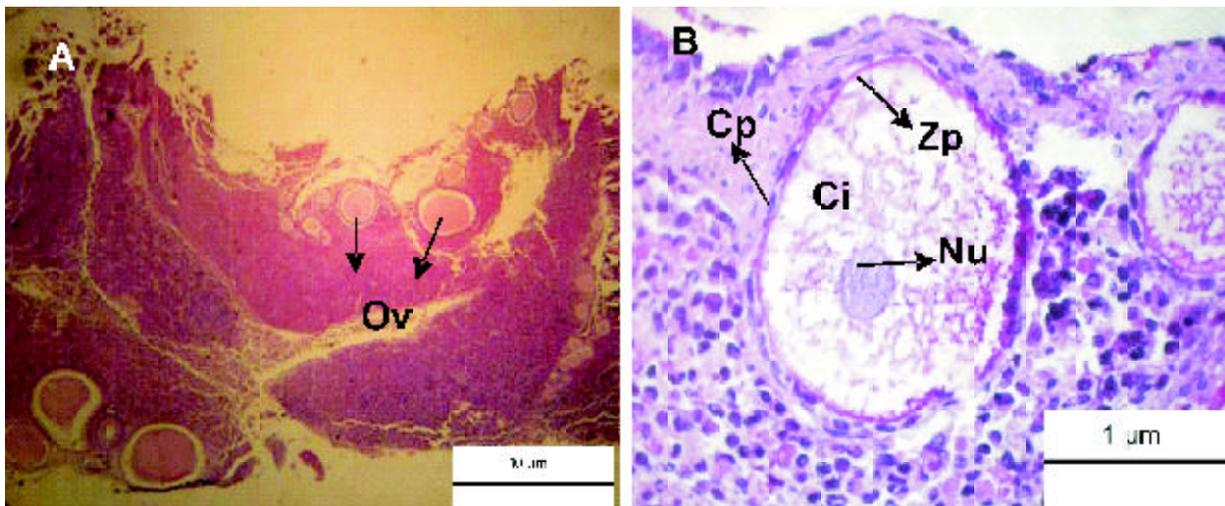
**Figura 1.** A. Ovario de *N. bancroftii* estado IIIa con ovocitos (Ov) alargados llenos de vitelo 16x, B. Corte de ovario con ovocitos en la periferia (Op) en hembra inmadura de *R. percellens* 40x, C. Ovario inmaduro de *U. venezuelae* con ovocitos inmaduros (Oi) y en formación (Of) ubicados cerca a la pared del ovario 4x y D. Ovario IIIa de *D. guttata* con ovocitos (Ov) en estado seis 10x.



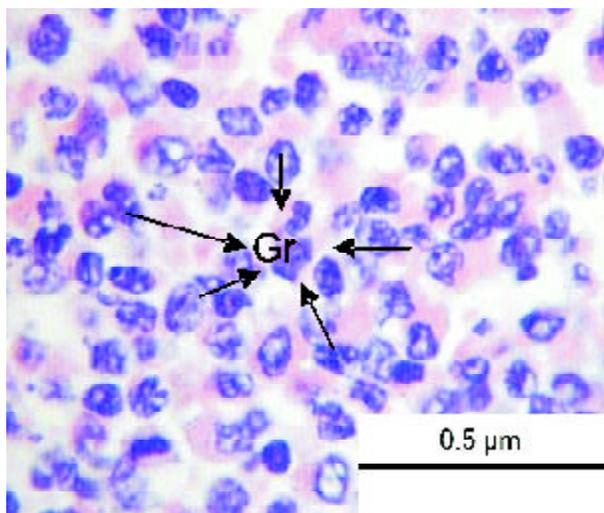
**Figura 2.** Conducto (C) presente en el ovario de *R. percellens* 25x.



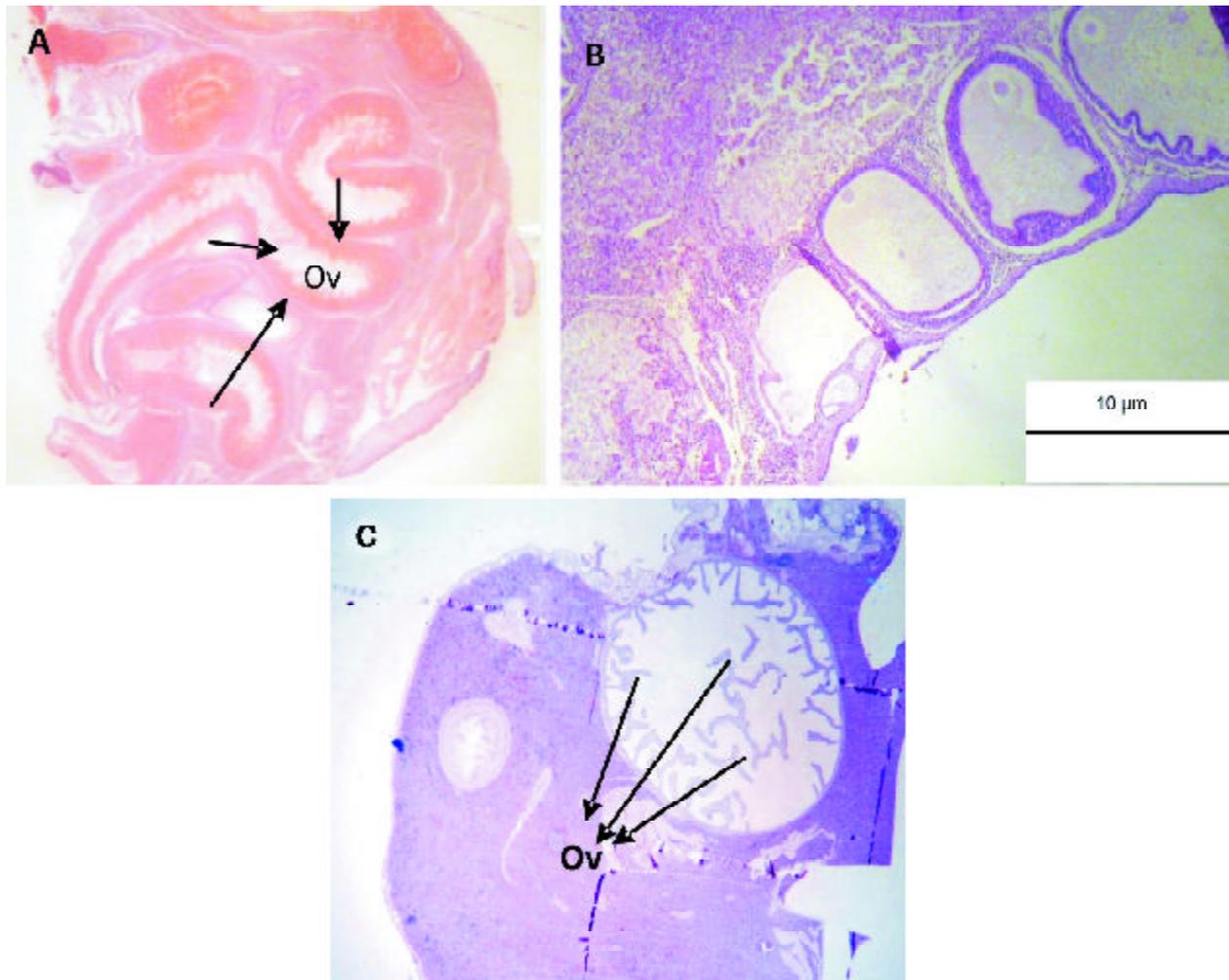
**Figura 3.** Pared ovárica de *N. bancroftii* con epitelio cúbico simple (Cc) tejido conjuntivo laxo con fibras de colágeno (Fc) y fibroblastos (F) 100x.



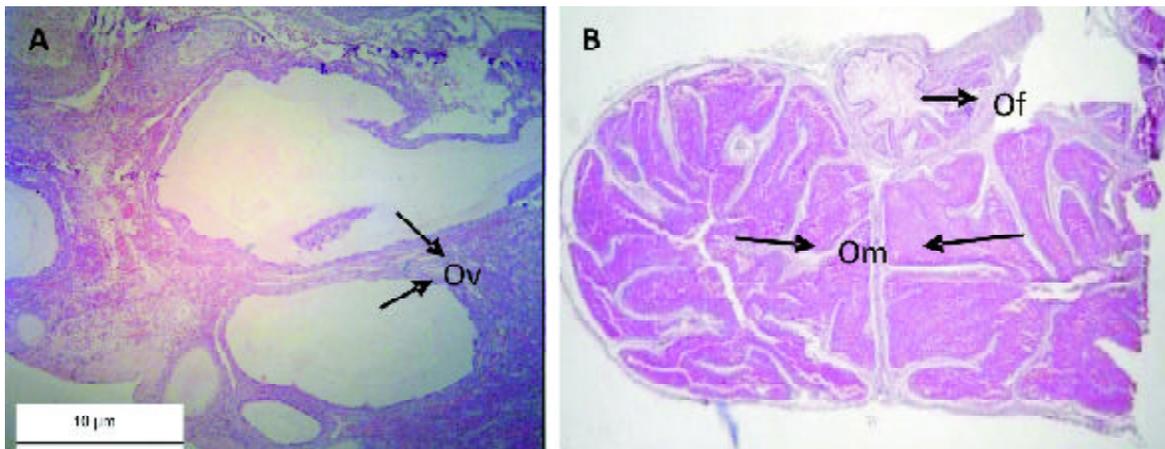
**Figura 4.** A. Ovario inmaduro de *R. percellens* con ovocitos (Ov) 40x y B. Primer estado de desarrollo para *U. venezuelae* con células planas (Cp) en la pared folicular, núcleo (Nu), Citoplasma (Ci) y zona pelúcida (Zp) 40x.



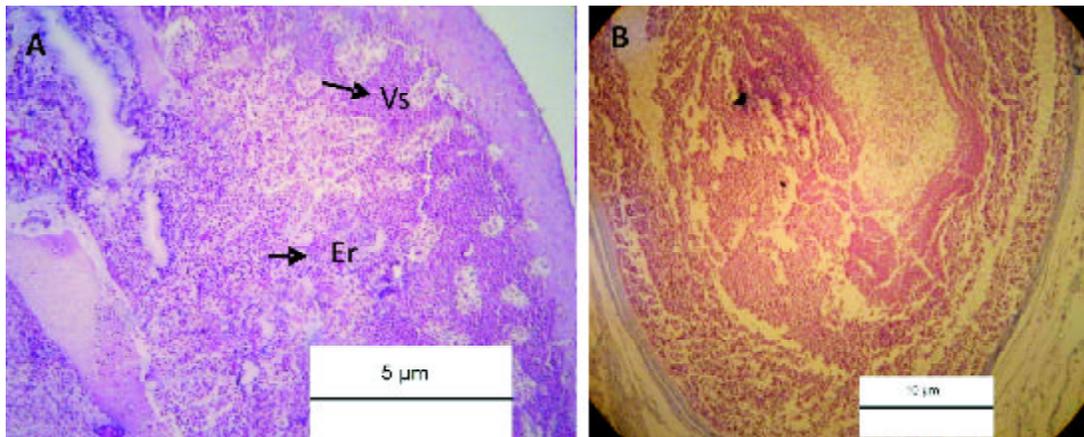
**Figura 5.** Gránulos en *U. venezuelae* (Gr) 100x.



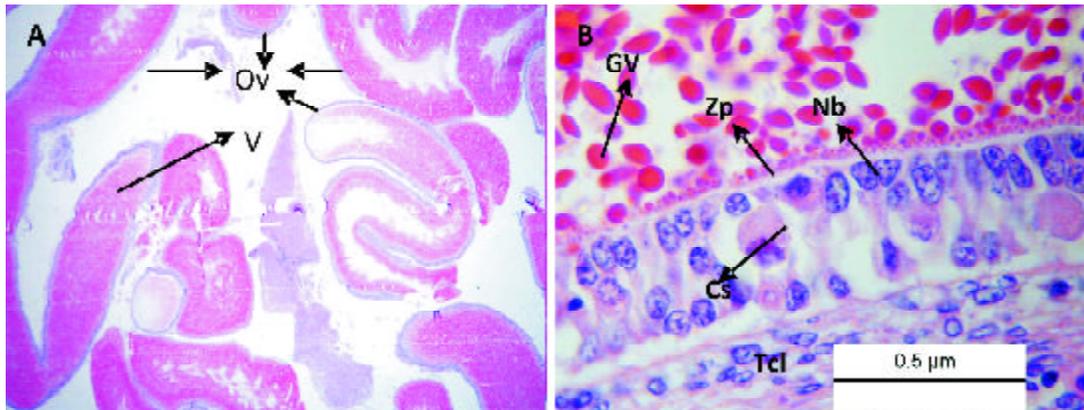
**Figura 6.** Cambios de estados de madurez **A.** Ovario IIIa con ovocitos (Ov) en *N. bancroftii* 8x, **B.** Ovario maduro IIIa de *U. venezuelae* 4x y **C.** Ovario IIIa de *D. guttata* con ovocitos (Ov) en estado seis 10x.



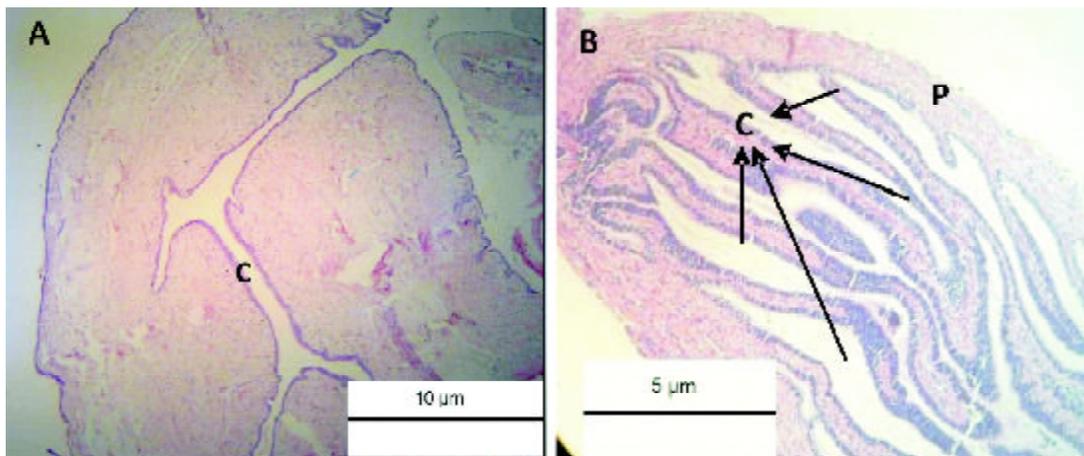
**Figura 7.** Cambios de estados de madurez en hembras **A.** Ovario IIIb con ovocitos (Ov) 4x de *N. bancroftii* y **B.** Ovario IIIb 20x con ovocitos maduros llenos de vitelo (Om) y ovocitos con inicios de formación de vitelo (Of) de *U. venezuelae*.



**Figura 8.** A. Ovario postparto IV con vasos sanguíneos (Vs) y eritrocitos (Er) 10x en *N. bancroftii* y B. Ovocito en cuarto estado de *R. percellens* conteniendo vitelo y rodeado del folículo (4X).



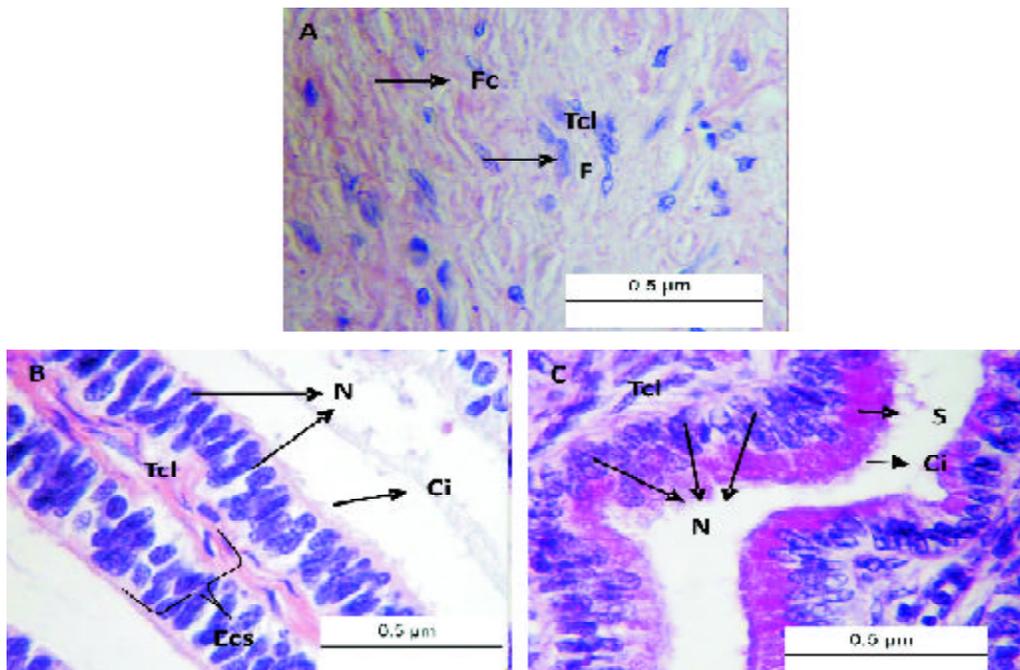
**Figura 9.** A. Ovocitos en estado seis (Ov) y vitelo (V) en *N. bancroftii* 16x. B. Pared de ovocito en estado seis con células secretoras (CS) núcleos basófilos (Nb), tejido conjuntivo laxo (Tcl), zona pelúcida (Zp) y gránulos de vitelo maduros (GV) 100X.



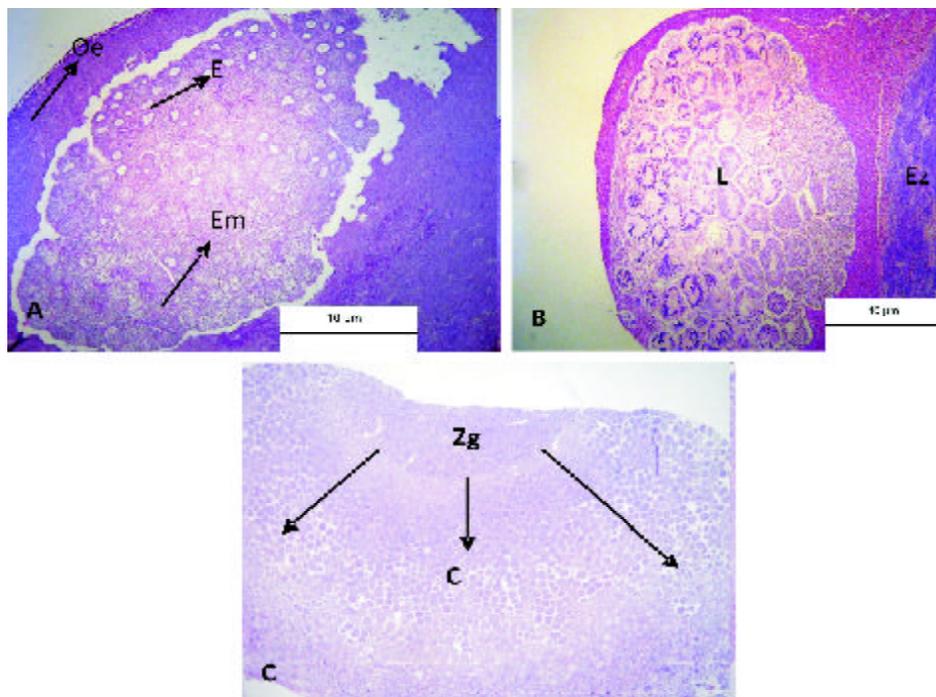
**Figura 10.** A. Oviducto de *N. bancroftii* con conducto central (C) 4x. B. Oviducto de *U. venezuelae*, se observa la pared (P) y los canales (C) 10x.

**Tabla 3.** Sumario de las principales características histológicas de los órganos reproductivos femeninos para cuatro especies de rayas de aguas someras del Caribe sur.

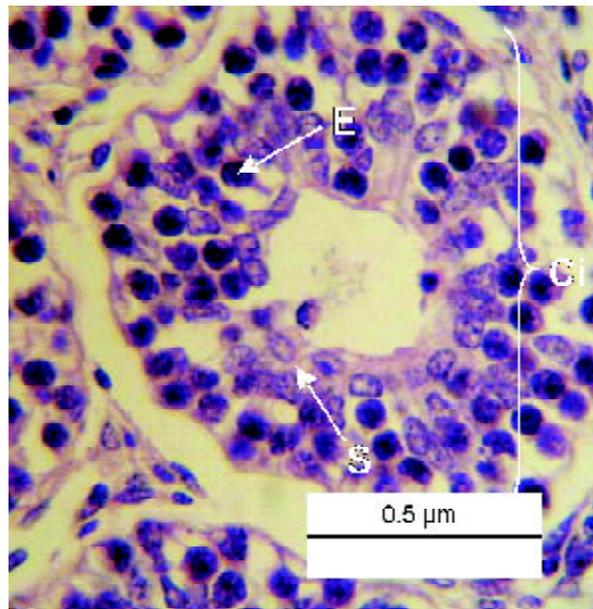
	<i>N. bancroftii</i>	<i>R. percellens</i>	<i>U. venezuelae</i>	<i>D. guttata</i>
<b>OVARIO</b>				
<b>Tipo</b>	Atípico, apariencia de hilos; presenta canal para Salida de ovocitos (Interno).	Compuesto, con ovocitos externos (Externo) y conducto central (Interno)	Compuesto, con ovocitos externos (Externo) y conducto central (Interno).	Externo, con ovocitos externos y ausencia de canal.
<b>Pared</b>	Epitelio cúbico simple y tejido conjuntivo laxo.	Capa externa de tejido epitelial cúbico simple y tejido conjuntivo laxo.	?	Epitelio cúbico simple y tejido conjuntivo laxo.
<b>Presencia de gránulos</b>	Presentes, inmersos en los espacios restantes de los ovocitos, cambian su proporción de acuerdo a la madurez del ejemplar.	Presentes, en mayor proporción en inmaduros que en maduros.	Presentes, en mayor proporción en inmaduros, que en maduros.	Presentes, no se pudo establecer diferencias entre inmaduros y maduros.
<b>Diferencias en estados de madurez</b>	Se establecieron diferencias en IIIa, IIIb y IV.	Las diferencias entre inmaduro, IIIa y IIIb radican en la proporción de cistos y en sus estados de madurez.	Descripciones para inmadura, IIIa y IIIb.	Sólo hay una descripción de IIIa.
<b>Estados de desarrollo ovocitario</b>	Seis estados, debidos a cambios en la pared folicular de células planas en inmaduros a cúbicas, rodeadas de tejido conjuntivo laxo, aparecen en la pared células secretoras, con incrementos en diámetro; cambios en la forma ya que se alargan y poseen vitelo en formación.	Cuatro estados, cambios en la pared que en inmaduros posee epitelio plano simple a columnar rodeado por tejido conjuntivo laxo, con dos a tres nucléolos, en maduros, inmerso en el epitelio columnar, aparecen células secretoras y gránulos de vitelo, incrementos en diámetros.	Siete estados, cambios en la pared, en inmaduros células planas a maduros con epitelio cúbico con tejido conjuntivo, sin células secretoras inmersas, aparecen vellosidades, con igual conformación de la pared, que ocupan el citoplasma, cambios de coloración en el vitelo, no se identificaron nucléolos.	Siete estados, diferencias principalmente en la composición de la pared que cambió de células planas en inmaduros (núcleo con un nucléolo) a cúbicas rodeadas de tejido conjuntivo laxo en maduras, aparición de prolongaciones derivadas que invaden el citoplasma hasta cubrirlo completamente y desarrollo de vitelo, aparecen células secretoras en las vellosidades.
<b>Diámetros máximo y mínimo</b>	0,10 $\mu\text{m}$ a 4,04 mm, con largo de 19,81mm, cambian de forma circular a alargados.	0,21 $\mu\text{m}$ a 13 mm, manteniendo la forma circular.	0,29 $\mu\text{m}$ a 1,48 mm, la forma se mantiene circular.	1,42 $\mu\text{m}$ a 5,86 mm, se mantiene la forma circular.
<b>OVIDUCTO</b>				
<b>Pared</b>	Tejido conjuntivo laxo y muscular liso.	Tejido conjuntivo y muscular liso.	Tejido muscular liso.	Tejido muscular liso.
<b>Canal</b>	Con un canal central recubierto de epitelio columnar pseudoestratificado ciliado.	?	Con varios canales recubiertos por tejido epitelial columnar pseudoestratificado y de tejido conjuntivo.	Tejido epitelial columnar pseudoestratificado ciliado, que presentó células secretoras, con tejido conjuntivo laxo inmerso entre los epitelios.
<b>Glándula nidamental</b>	?	Forma ovalada y dos porciones glandulares una circular y otra alargada, con secreción de mucus en su interior, con un ducto central y otros secundarios, canal recubierto de microvellosidades, ubicada anterior al útero	Un canal central con una pared de tejido conjuntivo laxo y muscular liso, recubierto de epitelio columnar pseudoestratificado ciliado, con dos porciones glandulares de la misma forma (circular) diferenciadas por presentar o carecer de granulaciones, ubica da en la parte anterior al útero.	?
<b>ÚTERO</b>				
	Capa externa de tejido epitelial columnar simple. Pared con tejido muscular liso y conjuntivo, irrigada por venas, arterias, vasos y capilares, con vellosidades tapizadas por epitelio columnar simple, sin células granulosas.	Con dos regiones, una glandular y otra con conductos rodeados de epitelio pseudoestratificado ciliado. En la luz del útero se observaron pliegues o prolongaciones uterinas, conformadas por células glandulares de tipo caliciforme, además de tejido conjuntivo y túbulos glandulares, irrigado por vasos sanguíneos y epitelio columnar provisto de microvellosidades.	Pared con tejido muscular liso, con tejido conjuntivo laxo, poco irrigada, con algunos capilares, prolongaciones no irrigadas, constituidas por tejido conjuntivo laxo y recubiertas con tejido columnar simple, no se encontró tapizada por cilios o microvellosidades.	Pared con tejido muscular liso irrigado con arteriolas y capilares grandes, vellosidades, irrigada con arteriolas, vasos y capilares sanguíneos, recubiertas por epitelio cúbico y tejido conjuntivo en el interior.



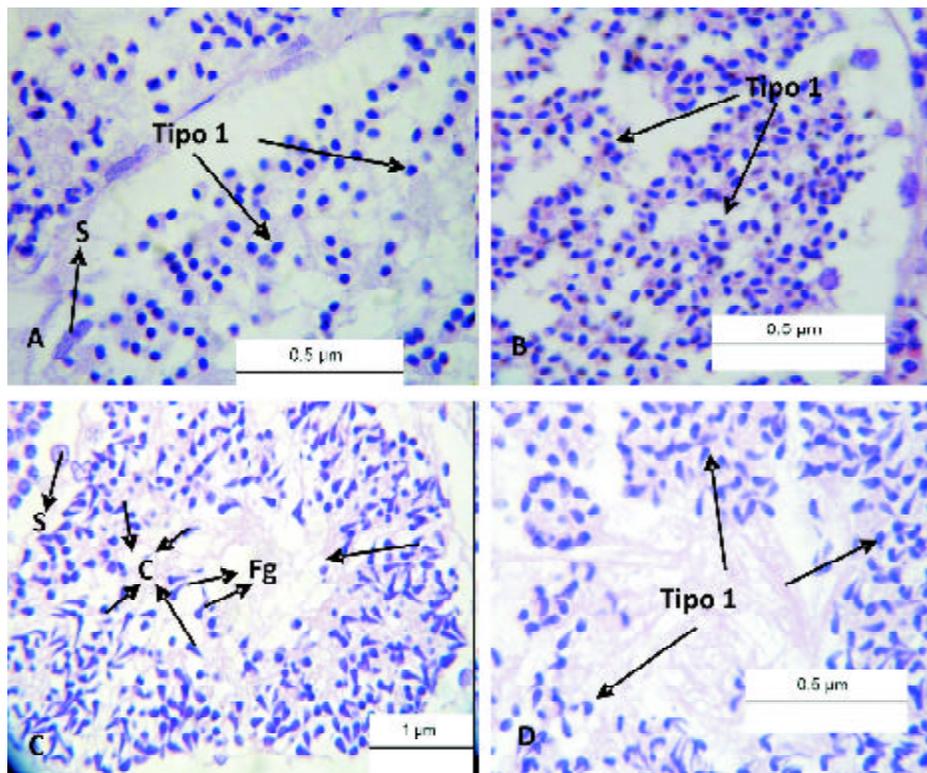
**Figura 11.** A. Pared del oviducto de *D. guttata* formada por tejido conjuntivo laxo (Tcl) con fibroblastos (F) y fibras de colágeno (Fc) 100x. B. Canales de *D. guttata* recubiertos de epitelio columnar pseudoestratificado (Ecs) con núcleos (N) en diferente posición, cilios (Ci) y en la parte central tejido conjuntivo (S) 100x. C. Superficie de canales en *D. guttata* con células columnares con secreciones (S) 100x.



**Figura 12.** A. Testículo de *N. bancroftii* en maduración con espermatogonias (E), órgano epigonal (Oe) y espermatozoides maduros (Em) 4x. B Testículo maduro de *U. venezuelae* con lóbulos (L), se observan diferentes estados de maduración de los espermatozoides (Ez) 4x. C. Testículo radial en maduración de *D. guttata* con quistes (C), células espermáticas en diferentes estados de desarrollo y zona germinal (Zg) 16x.



**Figura 13.** Primera etapa de desarrollo testicular en *R. percellens*, se observan los cistos (Ci), células de Sertoli (S) y espermatogonias (E) 100x.



**Figura 14.** A. Proceso de espermiogénesis en *N. bancroftii*, se observan espermatídes tipo I acompañadas de células de Sertoli (S) 100x. B. Espermatídes tipo I en *R. percellens* 100x.

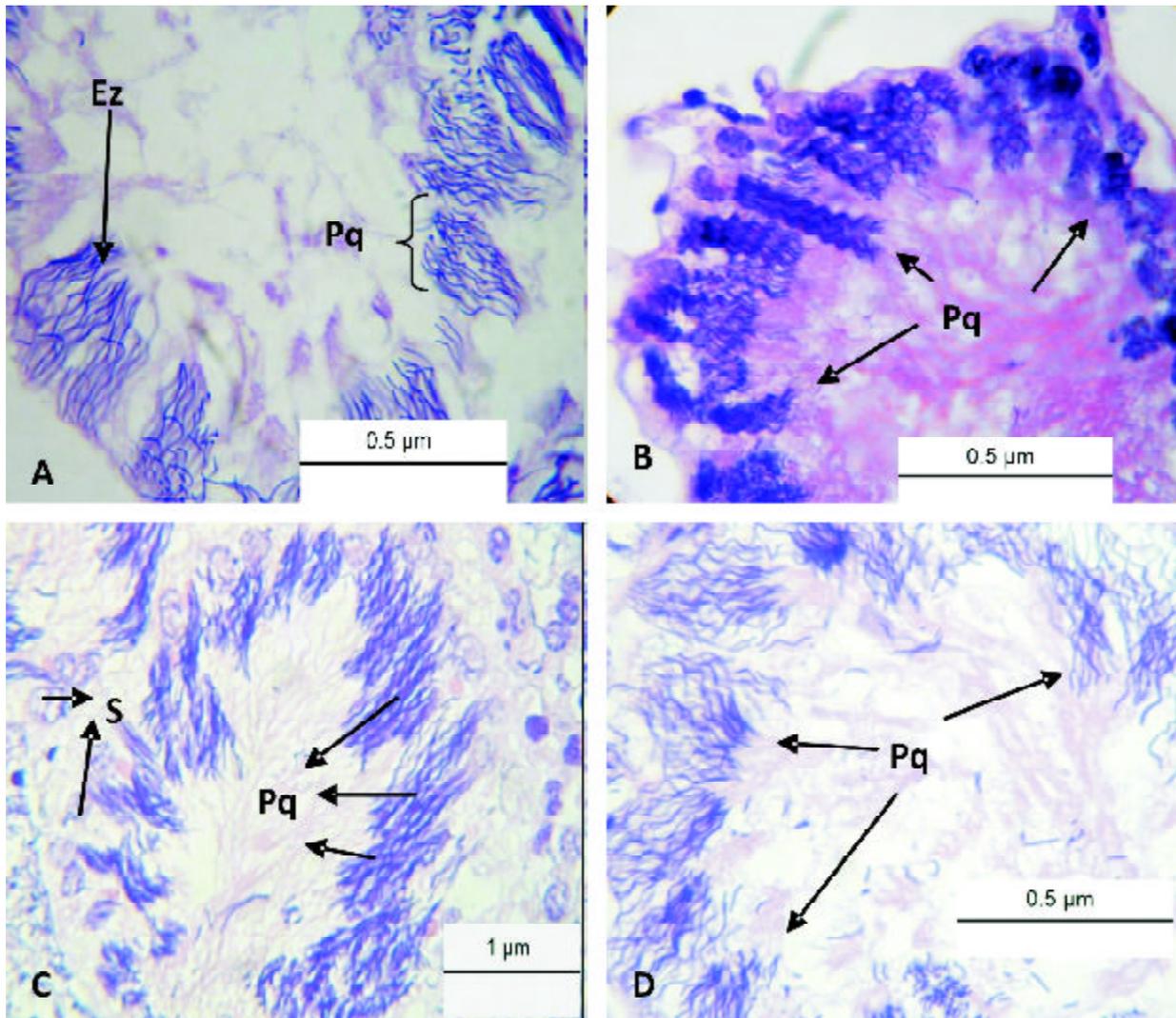
C. Desarrollo de espermatídes tipo I en *U. venezuelae*, se aprecian las cabezas (C) y flagelos (Fg) y las células de Sertoli (S) 40X. D. Espermatídes tipo I ovaladas con un triángulo que será el flagelo en *D. guttata* 100X.

**Tabla 4.** Sumario de las principales características histológicas de los órganos reproductivos masculinos para cuatro especies de rayas de aguas someras del Caribe sur.

	<i>N. bancroftii</i>	<i>R. percellens</i>	<i>U. venezuelae</i>	<i>D. guttata</i>
<b>TESTÍCULO</b>				
<b>Tipo</b>	Compuesto: alargados y comprimidos dorsoventralmente, área germinal en región lateral, grupos de desarrollo folicular con áreas más desarrolladas cercanas o en la misma dirección.	Compuesto: aspecto lobular externo, zona germinal en superficie lateral, cistos desarrollándose radialmente y hacia el interior del testículo.	Radial: cilíndrico aplanado dorsoventralmente constituido por unidades formadas por folículos seminíferos que funcionan independientemente, rodeadas de órgano epigonal.	Radial: cilíndrico, aplanado dorsoventralmente, superficie lobular sin unidades funcionales formadas por cistos, con canal de salida de espermatozoides.
<b>Estados de maduración</b>	Inmaduros: sólo se observan cistos de espermatogonias. En maduración: presentaron espermatogonias y espermatoцитos primarios y secundarios. Maduros: todos los estados de desarrollo.	Inmaduros: bajo número de cistos y espermatogonias. En maduración: mayor cantidad de cistos con espermátides y espermatoцитos primarios y secundarios. Maduros: todos los estados de desarrollo.	Inmaduros: pocos cistos con espermatogonias y espermatoцитos primarios. En maduración: aumento del número de cistos, presentando espermátides. Maduros: todos los estados, con predominio de espermatozoides.	Inmaduros: algunos cistos con espermatogonias. En maduración: solo una parte con espermatogonias a espermatoцитos secundarios. Maduros: todos los estados desde espermatogonias hasta espermatozoides.
<b>Espermatogénesis</b>	Siete estados de desarrollo espermático en las cuatro especies (espermatogonias, espermatoцитos primarios, secundarios, espermátides tipo I, tipo II, tipo III y espermatozoide), se registraron variaciones principalmente en forma de células espermáticas avanzadas (espermátides), durante las etapas cambiada forma y tamaño de cabeza (redonda u ovalada) y flagelo ondulado en las especies <i>N. bancroftii</i> , <i>U. venezuelae</i> y <i>D. guttata</i> a rizado en <i>R. percellens</i> .			
<b>EPIDÍDIMO</b>	No se colectó	Inmaduros, conductos pequeños y estrechos. Maduros, conductos abiertos llenos de espermatozoides con flagelos bien rizados, paredes de tejido epitelial acompañados de musculatura lisa y tejido conjuntivo.	Inmaduros túbulos cerrados. Maduros, dos sectores uno de túbulos cerrados y otro de abiertos llenos de espermatozoides, con paredes de tejido cúbico.	Inmaduros, cerrados con dos sectores. Maduros, abiertos con espermatozoides, con paredes de tejido cúbico y cerrados con paredes de tejido columnar.
<b>VESÍCULA SEMINAL</b>	No se colectó	Tapizada por musculatura lisa. Inmaduros, constituida por tejido muscular y conjuntivo. Maduros, formado además por ductos tapizados internamente por una capa de epitelio columnar.	Ductos rodeados de tejido conjuntivo laxo, sin células columnares en interior.	No se colectó

en el alto número de estados, presentando prolongaciones derivadas de la pared, denominadas pliegues foliculares (Babel, 1967), las cuales ocupan el citoplasma. En *D. guttata* se observaron células secretoras en los pliegues, las cuales no se vieron en *U. venezuelae*. En *N. bancroftii* ocurre una elongación de los ovocitos, paralela

a la de las masas tubulares descritas anteriormente; de ese modo, los ovocitos de *N. bancroftii* llegan a alcanzar longitudes de casi dos centímetros, mientras que su anchura es de menos de medio centímetro (Figura 9). Además, las paredes de los ovocitos de esta especie también poseen las células secretoras ya mencionadas.

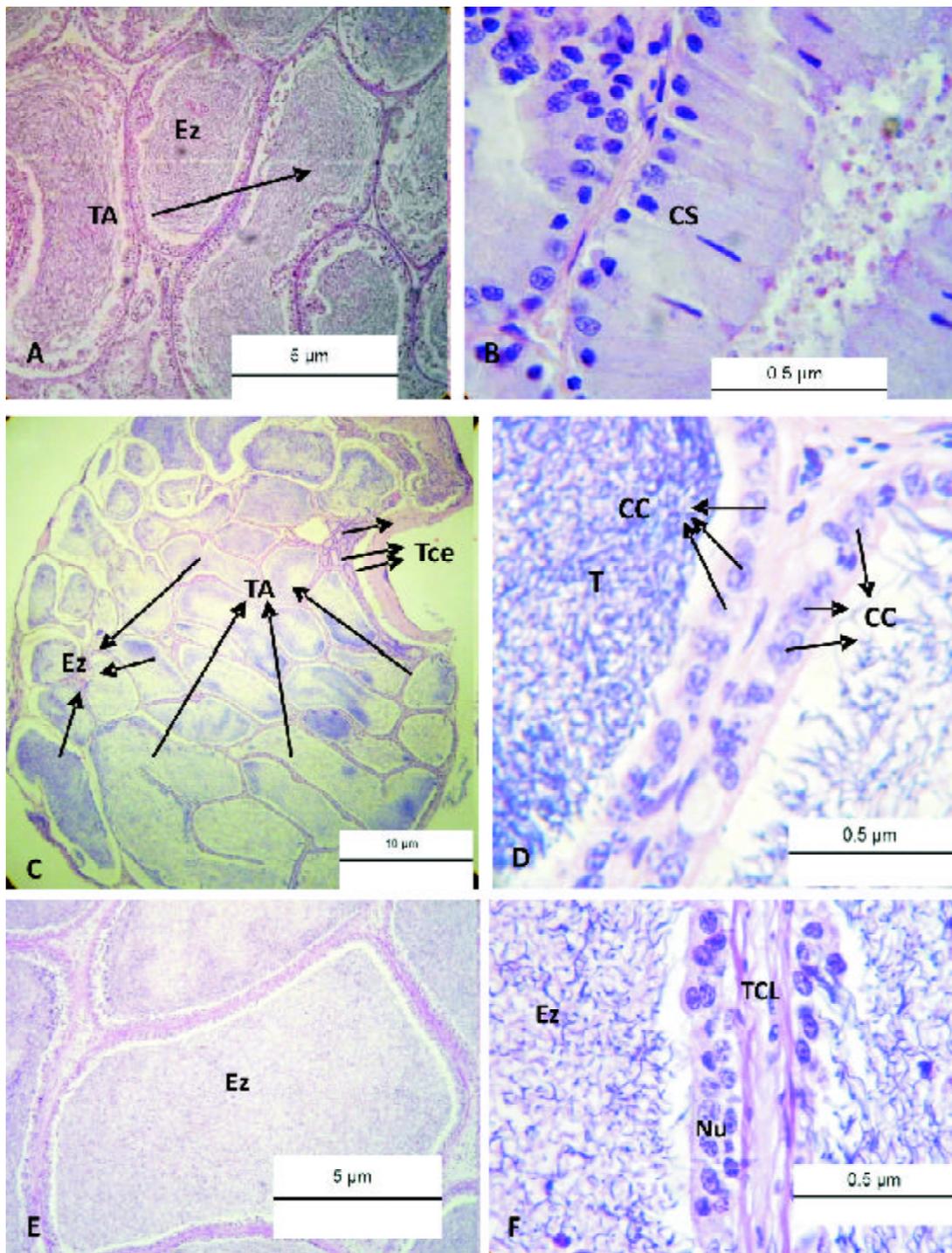


**Figura 15.** A. Espermatozoides (Ez) agrupándose en paquetes (Pq) en *N. bancroftii*, los flagelos están poco ondulados 100x. B. Espermatozoides de *R. percellens* completamente maduros, en paquetes y muy rizados 100x. C. Espermatozoides largos y ondulados en *U. venezuelae*, formando paquetes, acompañados de células de Sertoli (S) 40x. D. Espermatozoides que se están empaquetando en *D. guttata* 100X.

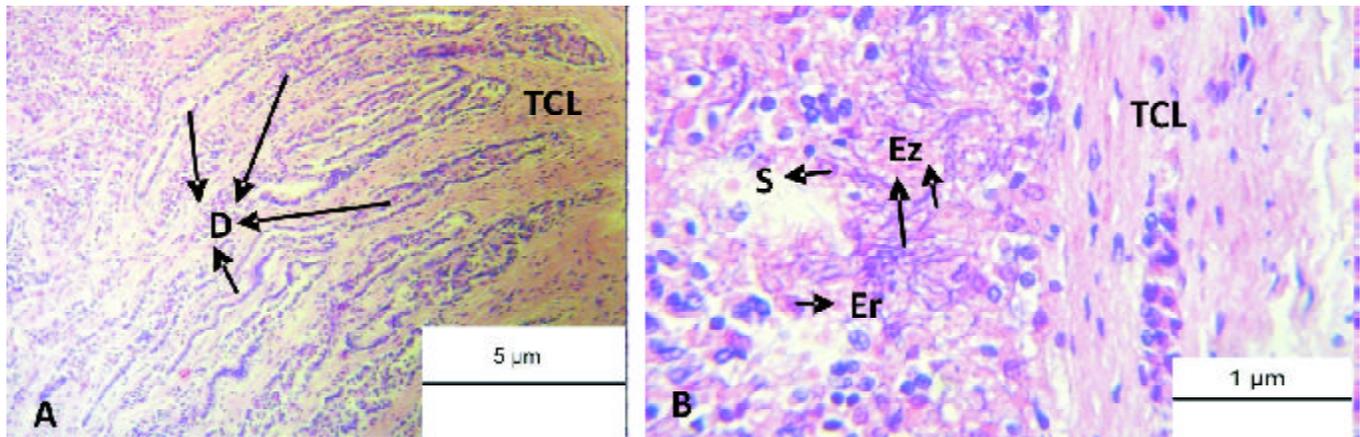
Los oviductos de *U. venezuelae* y *N. bancroftii* (Figura 10) difieren en que en la última sólo hay un canal, mientras que en la primera se detectaron varios canales; no fue posible dilucidar si se trata, en este caso, de varios canales o de uno solo, ramificado. La otra especie de Myliobatiformes estudiada (*D. guttata*) parece presentar también varios canales del oviducto y se observaron células secretoras en su superficie (Figura 11). Sólo fue posible estudiar las glándulas nidamentales de *R. percellens* y *U. venezuelae*, las que difieren en que una de las porciones glandulares en *R. percellens* es alargada. Además, esta especie tiene células caliciformes, secretoras de mucus.

El útero de *R. percellens* difiere del de las otras especies pues posee dos regiones, una de ellas con células glandulares caliciformes; en este caso se podría tratar de una autapomorfía de los rhinobátidos. No se pudieron detectar diferencias ligadas a la estrategia de nutrición embrionaria de los Myliobatiformes, conocida como trofonemata (Hamlett *et al.*, 1999).

Los testículos de las cuatro especies estudiadas son aplanados dorsoventralmente, acorde con la morfología general de las rayas. *Narcine bancroftii* y *R. percellens* son similares al tipo Compuesto (Pratt, 1988), caracterizado



**Figura 16.** A. Corte transversal epidídimo en *R. percellens*, se observan espermatozoides (Ez) y túbulos abiertos (TA) 10x. B. Pared celular en *R. percellens* constituida por tejido columnar simple (CS) 100x. C. Corte transversal de epidídimo en *U. venezuelae*, con túbulos abiertos (TA) llenos de espermatozoides (Ez) y túbulos cerrados (Tce) sin células espermáticas 4x. D. Membrana que divide los túbulos (T) conformada por células cúbicas (CC) en *U. venezuelae* 100x. E. Ampliación de ductos abiertos con espermatozoides (Ez) en *D. guttata* 10x. F. Pared de ductos abiertos con espermatozoides (Ez) en *U. venezuelae*, formado por células cúbicas con núcleos en diferentes posiciones (Nu) y tejido conjuntivo (TCL) 100x.



**Figura 17.** A. Vesícula seminal en *U. venezuelae* con los ductos (D) donde se almacena el semen rodeados de tejido conjuntivo laxo (TCL) 10x. B. Acercamiento del interior del ducto con semen, constituido por espermatozoides (Ez), partes de células de Sertoli (S) y eritrocitos (Er), además de tejido conjuntivo laxo (TCL) 40x.

porque el tejido epigonal está reducido a una capa delgada ~~alrededor del tejido conectivo~~ (Figura 12). Pratt (1988) registró que el testículo de *Dasyatis* sp. es del tipo Radial, definiéndolo como una estructura lobulada con desarrollo compuesto a través de cada lóbulo; los cortes estudiados de los testículos de *D. guttata* parecen cumplir con esa descripción (Figura 12). *Urotrygon venezuelae* difiere de ese patrón en que el testículo está formado por lóbulos, cada uno rodeado por órgano epigonal y albergando los diferentes estados de espermatogénesis (Figura 12); este tipo sería parcialmente similar al Radial, difiriendo en la relación entre los lóbulos y el tejido epigonal, lo cual podría constituir un estado derivado, autapomórfico, de los urotrygónidos. Pratt (1988) había advertido que, al aumentar el número de batoides examinados, se incrementarían, así mismo, los tipos estructurales de sus gónadas.

El desarrollo de la madurez testicular incluye, para las cuatro especies, tres estados generales: inmaduros, en maduración y maduros. La espermatogénesis, por su parte, es similar para las especies estudiadas, con siete estados (Figuras 13-15). Se destaca que las cabezas de los espermatozoides, que se empiezan a formar a partir del cuarto estado, cambian ontogenéticamente de redondeadas a ovaladas. Los flagelos de los espermatozoides se desarrollan similarmente en las cuatro especies, pero los de *R. percellens* en el séptimo estado difieren claramente de los de las otras especies en que son altamente rizados (Figura 15). Esto ya había sido registrado por Wenbin y Shuyuan, (1993) para la especie asiática *R. hynnicephalus*, por lo cual es aparentemente una carácter derivado compartido, sinapomorfía, de estas dos especies de rhinobátidos.

El epidídimo fue detectado para tres de las especies, pues no se encontró en *N. bancroftii*. Se halló que en *R. percellens* los ductos están formados por paredes de tejido epitelial con musculatura lisa y tejido conjuntivo (Figura 16), mientras que en *U. venezuelae* y *D. guttata* las paredes están formadas por tejido cúbico (Figura 16). Además, en *D. guttata* se presentan también células columnares. La vesícula seminal sólo fue vista en *R. percellens* y *U. venezuelae*, difiriendo en que en *R. percellens* los ductos se hallan tapiados internamente por una capa de epitelio columnar, mientras que en *U. venezuelae* ellos solo poseen tejido conjuntivo laxo (Figura 17).

Resumiendo, se detectaron tres tipos diferentes de sistemas reproductivos femeninos en las rayas estudiadas. *Rhinobatos percellens* ejemplifica un sistema generalizado, que incluye características ampliamente representadas en los peces cartilagosos. Los Myliobatiformes, por su parte, muestran un ovario relativamente generalizado, pero que alberga ovocitos con estructuras internas únicas. *Narcine bancroftii* posee el ovario más derivado, en cuanto a la presencia y forma de estructuras exclusivas. Las diferencias parecen ser menos marcadas en los sistemas reproductivos masculinos, con *N. bancroftii* mostrando un testículo generalizado, mientras *R. percellens* presenta cierto grado de lobulación externa. El testículo de *D. guttata* es de tipo Radial; *U. venezuelae* podría poseer un testículo Radial modificado.

Desde el punto de vista filogenético, los resultados no aportan información que permita dilucidar las relaciones entre los tres órdenes de rayas estudiados. Se detectaron estados claramente derivados, pero autapomórficos, para

*N. bancroftii*, *R. percellens* y *U. venezuelae*. Al no encontrarse estados que constituyan sinapomorfías, salvo uno probable para las dos especies de *Rhinobatos*, se hace imposible aun utilizar esos datos para el esclarecimiento de las afinidades dentro de los batoideos.

### Agradecimientos

La Fundación para la Promoción de la Investigación y la Tecnología del Banco de la República, Bogotá, financió parcialmente este trabajo a través del proyecto 2174 "Biología reproductiva de cuatro especies de batoideos: una contribución histológica". La Universidad Nacional de Colombia, por intermedio del Centro de Investigaciones en Ciencias del Mar y el Instituto de Investigaciones Marinas y Costeras financiaron en parte y colaboraron con infraestructura para el trabajo del primer autor. El programa de Biología Marina de la Facultad de Ciencias Naturales de la Universidad de Bogotá Jorge Tadeo Lozano, en Bogotá y Santa Marta, contribuyó con el uso de sus laboratorios para el análisis de las muestras y el procesamiento de la información. A los pescadores de playa Salguero y Don Jaca en Santa Marta, Magdalena y a sus familias. Contribución 325 CECIMAR, Universidad Nacional de Colombia y Contribución 008 del Grupo de Investigación de Peces Cartilagosos (GIPECA-UJTL).

### Bibliografía

- Acero P., A., L.S. Mejía & M. Santos-Acevedo.** 2002. *Pristis pectinata*, *Pristis perotteti*: 47-52 En **Mejía, L.S. & A. Acero P.** (Eds.) Libro rojo de peces marinos de Colombia. Inst. Cienc. Nat., Univ. Nal. Colombia/INVEMAR, Bogotá, 174 p.
- Acevedo, K., J. Bohórquez-Herrera, F. Moreno, C. Moreno, E. Molina, M. Grijalba-Bendeck & P. Gómez-Canchong.** 2007. Tiburones y rayas (subclase Elasmobranchii) descartados por la flota de arrastre camaronero en el Caribe de Colombia. *Acta Biol. Colomb.*, **12** (2): 71-81
- Babel, J.** 1967. Reproduction, life history and ecology of the round stingray *Urolophus halleri* Cooper. *Calif. Fish Game Bull.*, **137**: 1-104.
- Compagno, L.J.V.** 1999. Checklist of living elasmobranchs: 471-498. En **Hamlett, W.C.** (Ed.) *Sharks, skates, and rays. The biology of elasmobranch fishes*, John Hopkins University, Baltimore.
- Grijalba-Bendeck, M., C. Polo-Silva & A. Acero P.** 2007a. Una aproximación a la abundancia de los batoideos capturados artesanalmente en Santa Marta (Colombia). *Bol. Invest. Mar. Cost.*, **36**: 251-268.
- Grijalba-Bendeck, M., Polo C.J., K. Acevedo, F. Moreno & D. Mojica.** 2007b. Aspectos biológicos de algunos Batoideos en Santa Marta (Caribe colombiano). Centro de Investigaciones Científicas. Facultad de Ciencias Naturales. Programa de Biología Marina. Universidad Jorge Tadeo Lozano. Santa Marta. 180 p.
- Grijalba-Bendeck, M., K. Acevedo, F. Moreno, D. Mojica & A. Acero P.** 2008. Condición reproductiva y descripción anatómica del sistema reproductor de algunos batoideos capturados artesanalmente en Santa Marta (Caribe colombiano). *Resultados de Investigación* **1**(3): 1-20.
- Hamlett, W., M. Jezior & R. Spieler.** 1999. Ultrastructural analysis of folliculogenesis in the ovary of the yellow spotted stingray *Urolophus jamaicensis*. *An. Anat.*, **181**: 159-172.
- Martoja, R & M. Martoja-Pierson.** 1970. *Técnicas de Histología Animal*. Toray-Masson, Barcelona, 350 p.
- McEachran, J.D. & N. Aschliman.** 2004. Phylogeny of Batoidea: 79-113. En J.C. Carrier, J.A. Musick y M.R. Heithaus (Eds.) *Biology of sharks and their relatives*, CRC, Boca Raton, EE.UU.
- Mejía-Falla, P.A., A.F. Navia, L.M. Mejía-Ladino, A. Acero P. & E. Rubio.** 2007. Tiburones y rayas de Colombia (Pisces: Elasmobranchii): lista actualizada, revisada y comentada. *Bol. Invest. Mar. Cost.*, **36**: 111-149.
- Moreno, F., K. Acevedo, M. Grijalba-Bendeck, A. Acero P. & J. Páramo.** En prensa. Reproducción de la raya eléctrica *Narcine bancroftii* (Griffith & Smith) en Santa Marta, Caribe de Colombia. *Latin Am. J. Aquat. Res.*
- Natanson, L. & G. Cailliet.** 1986. Reproduction and development of the Pacific angel shark, *Squatina californica*, off Santa Barbara, California. *Copeia*, (4): 987-994.
- Nelson, J.S.** 2006. *Fishes of the world*. Cuarta Ed. Wiley, Hoboken, EE.UU., 601 p.
- Pratt, H.L.** 1988. Elasmobranch gonad structure: a description and survey. *Copeia*. (3): 719-729.
- Snelson, F., S. Williams-Hooper & T. Schmidt.** 1988. Reproduction of the Atlantic stingray, *Dasyatis sabina*, in Florida coastal lagoons. *Copeia*. (3): 729-739.
- Téllez, L., C. Vargas & M. Grijalba-Bendeck.** 2006. Algunos aspectos biológicos de *Urotrygon venezuelae* Schultz, 1949 (Elasmobranchii, Rajiformes, Urolophidae), capturada en playa Salguero, Santa Marta, Caribe de Colombia. *Rev. U.D.C.A. Actual. Divulg. Cient.*, **9** (2): 75-87.
- Villavicencio, C.** 1995. Observaciones sobre la biología reproductiva de *Narcine brasiliensis* (Olfers) (Pisces: Narcinidae), en la bahía Almejas, B.C.S. México. *Rev. Inv. Cient.*, **4** (1): 95-99.
- Wenbin, Z & Q. Shuyuan.** 1993. Reproductive biology of the guitarfish, *Rhinobatos hynnicephalus*. *Env. Biol. Fish.*, **38**: 81-93.

Recibido: abril 5 de 2008

Aceptado para su publicación: noviembre 13 de 2008