

Tesis Doctoral

# Aplicación de los pronósticos por ensambles a la predicción del tiempo a corto plazo sobre Sudamérica

Ruiz, Juan José

2008

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in [digital.bl.fcen.uba.ar](http://digital.bl.fcen.uba.ar). It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Ruiz, Juan José. (2008). Aplicación de los pronósticos por ensambles a la predicción del tiempo a corto plazo sobre Sudamérica. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

Ruiz, Juan José. "Aplicación de los pronósticos por ensambles a la predicción del tiempo a corto plazo sobre Sudamérica". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2008.

**EXACTAS** UBA

Facultad de Ciencias Exactas y Naturales



**UBA**

Universidad de Buenos Aires



# Índice

## Volumen I: Texto

	<b>Pag.</b>
<b>Agradecimientos</b> .....	7
<b>Resumen</b> .....	9
<b>Abstract</b> .....	10
 <b>Capítulo 1: Introducción</b>	
<b>1.1 Antecedentes sobre el tema</b> .....	11
<b>1.2 Objetivos</b> .....	15
<b>1.3 Naturaleza del aporte original</b> .....	16
 <b>Capítulo 2: Herramientas y datos utilizados</b>	
<b>2.1 Descripción del modelo WRF</b> .....	18
<b>2.2 Descripción del modelo MRF</b> .....	22
<b>2.3 Datos de precipitación observados</b> .....	24
2.3.1 Fuentes de datos .....	24
2.3.2 Control de calidad adicional .....	25
<b>2.4 Estimaciones de precipitación: CMORPH</b> .....	26
<b>2.5 Datos de superficie</b> .....	27
<b>2.6 Análisis en puntos de retícula</b> .....	28
 <b>Capítulo 3: Evaluación preliminar de las ventajas del pronóstico de precipitación por ensambles en la región de Sudamérica.</b>	
<b>3.1 Introducción</b> .....	30
<b>3.2 Metodologías para la generación de pronósticos probabilísticos</b> .....	31
3.2.1 Sistemas de pronósticos regionales por ensambles utilizados .....	31
3.2.2 Metodologías para la obtención y calibración de los pronósticos probabilísticos de precipitación .....	36
3.2.3 Calibración estática y dinámica .....	41
3.2.4 Obtención de pronósticos probabilísticos a partir de pronósticos determinísticos .....	43
3.2.6 Significancia estadística de los resultados: la técnica de bootstrap .....	45
<b>3.3 Resultados</b> .....	46
3.3.1 Análisis de la relación entre el error y la dispersión para	

	la precipitación .....	46
	3.1.2 Diagramas de confiabilidad y BSS .....	47
	3.1.3 Precisión y resolución .....	53
<b>3.4</b>	<b>Conclusiones</b> .....	57

## **Capítulo 4: El método de los vectores criados.**

<b>4.1</b>	<b>La técnica de los vectores criados</b> .....	61
<b>4.2</b>	<b>Características de los vectores criados y su relación con el error del análisis</b> .....	66
<b>4.3</b>	<b>Aplicación de la metodología en la generación de un ensamble global y otro regional</b> .....	73

## **Capítulo 5: El ensamble regional y global basado en la técnica de los vectores criados: reducción del error y estimación de la incertidumbre**

<b>5.1</b>	<b>Introducción</b> .....	76
<b>5.2</b>	<b>Metodología</b> .....	79
	5.2.1 Relación entre el error en la media del ensamble y el error en el pronóstico control .....	79
	5.2.2 Relación entre la dispersión del ensamble y el error de la media del ensamble .....	80
<b>5.3</b>	<b>Resultados</b> .....	85
	5.3.1 Relación entre el error del control y el error en la media del ensamble .....	85
	5.3.2 Relación entre la dispersión y el error en la media del ensamble .....	90
<b>5.4</b>	<b>Conclusiones</b> .....	99

## **Capítulo 6: Calibración de pronósticos de precipitación durante la estación cálida 2002-2003**

<b>6.1</b>	<b>Introducción</b> .....	103
<b>6.2</b>	<b>Metodología</b> .....	103
<b>6.3</b>	<b>Resultados</b> .....	107
	6.3.1 Impacto de las modificaciones introducidas a la calibración del ensamble en los histogramas de rangos .....	107
	6.3.2 Verificación de los pronósticos probabilísticos generados mediante las diferentes estrategias de calibración .....	109
	6.3.3 Análisis del valor económico del pronóstico .....	113
	6.3.4 Utilización de las estimaciones CMORPH en la calibración .....	115
<b>6.4</b>	<b>Conclusiones</b> .....	118
	6.4.1 Discusión .....	118
	6.4.2 Trabajos futuros .....	120

**Capítulo 7: Conclusiones y trabajos futuros**

<b>7.1</b>	<b>Conclusiones</b> .....	124
<b>7.2</b>	<b>Trabajos futuros</b> .....	131
	<b>Apéndice A</b> .....	134
	<b>Apéndice B</b> .....	145
	<b>Apéndice C</b> .....	154
	<b>Referencias</b> .....	156

**Volumen II: Figuras.**

	<b>Pag.</b>
<b>Figuras capítulo 2</b> .....	2
<b>Figuras capítulo 3</b> .....	4
<b>Figuras capítulo 4</b> .....	14
<b>Figuras capítulo 5</b> .....	18
<b>Figuras capítulo 6</b> .....	30
<b>Apéndice B</b> .....	41

# **Volumen I: Texto.**

## Agradecimientos

Deseo agradecer muy especialmente a Celeste, mi directora de Tesis por toda la dedicación y paciencia con la que me acompañó durante el desarrollo de este trabajo. Gracias también desde lo personal, por toda la ayuda, la comprensión y el afecto recibidos durante estos años.

También deseo agradecer a Eugenia, mi directora de Tesis por haber creído en mi y por introducirme al mundo de los pronósticos por ensambles. Gracias por haberme dado la oportunidad de aprender junto a vos y por el aliento y el afecto constantes que recibí de tu parte.

Deseo agradecer también al Centro de Investigaciones del Mar y la Atmósfera y a su director Dr. Mario Nuñez por haberme proporcionado el lugar de trabajo y todos los medios necesarios para la realización de este trabajo de tesis.

A la Dra. Carolina Vera, mi consejera de estudios y la primera persona con la que tuve oportunidad de hablar acerca de la carrera.

Deseo agradecer a todos aquellos que colaboraron durante la realización de este trabajo: A Alfredo Rolla y María Inés Ortiz de Zárate por haberme ayudado en múltiples oportunidades.

Al Dr. Istvan Szunyogh por haberme provisto los programas necesarios para implementar la técnica de breeding en el modelo MRF y por enseñarme a utilizarlos.

Al Dr. Erick Kostelich por ayudarme con la instalación del modelo MRF en las PCs del CIMA.

A la Dra. Jae Schem por proveerme los reanálisis en el formato requerido por el modelo utilizado.

A Dave Allured por haberme provisto los datos de precipitación de la red ANA.

Al Dr. Pedro Leite Silva Días y a Demerval Soares Moreira por facilitarme los datos del super ensamble de modelos de la Universidad de San Pablo y por atender a múltiples consultas respecto de los mismos.

Al Dr. Dirceu Herdies por facilitarnos los análisis GDAS con los datos SALLJEX asimilados.

A la Dra. Shu-Chi Yang por proveerme los códigos de matlab para realizar los experimentos con el modelo de Lorenz de 3 variables.

Deseo agradecer muy especialmente al Departamento de Ciencias de la Atmósfera y los Océanos en particular por la formación recibida antes y durante el desarrollo del presente trabajo de tesis.

Quiero agradecer muy especialmente a la Dra. Claudia Campetella, por todo el cariño recibido durante estos años y por lo que disfruté trabajando al lado de un ser humano excepcional.

Gracias a Yanina, Paola y Matilde por todo el afecto recibido y por permitirme aprender de ustedes.

Gracias a mis amigos de oficina por aguantarme todo este tiempo!! A Pauli, a Lore y a Sole por tantas charlas y buenos momentos compartidos.

Gracias a mis amigos, por estar siempre cuando los necesitaba: a Luciano, a Matías, a Ramiro, a Lorena, a Romina y a Ricardo, con quienes compartí mi carrera desde el principio e innumerables momentos agradables.

Quiero agradecer también muy especialmente a mi familia por haberme apoyado durante todo este tiempo permitiendo que pudiera llevar adelante este trabajo. Los quiero mucho!! A mi Mamá, mi Papá y a mis hermanos/as: Beatriz, Jorge, María del Carmen, Margarita y Bernardo.

Finalmente (last but not least) quiero hacer una mención especial al grupo M.A.T. sin cuyo apoyo incondicional no podría haber llevado adelante esta tarea. En particular al Dr. Smirnoff cuyo apoyo fue fundamental a lo largo de todo el trabajo y a P\* B\* su socia fundadora.

## Resumen

La técnica de pronósticos por ensambles permite obtener una reducción en el error en los pronósticos a la par que brinda información sobre la confiabilidad de los pronósticos y se basa en la realización de múltiples simulaciones utilizando condiciones iniciales ligeramente perturbadas o modelos diferentes. Esta técnica ha sido utilizada en pronósticos a mediano y corto plazo y funciona operativamente en los principales centros mundiales de pronóstico. En el presente trabajo de tesis, el pronóstico por ensambles es aplicado al pronóstico regional a corto plazo sobre Sudamérica. Tres diferentes metodologías para la generación de ensambles son evaluadas sobre distintas regiones de Sudamérica con diversos regímenes de precipitación, haciendo especial énfasis en las propiedades de los pronósticos probabilísticos de precipitación. Por otra parte se evaluaron diferentes estrategias para calibrar los pronósticos probabilísticos de precipitación de forma tal de aumentar su confiabilidad y por ende el valor económico de los mismos.

Los resultados obtenidos muestran en qué medida la utilización de la técnica de pronósticos por ensambles puede incrementar el valor de la información meteorológica provista sobre Sudamérica.

## **Abstract**

The use of ensembles in weather forecasting has reduced forecasts errors and provides some information about the daily changes in the forecast uncertainty. There are several methods for ensemble generation: some of them introduce perturbations in the initial and / or boundary conditions, others consist on the use of multiple models and some others combine both approaches. Ensembles of forecasts are run operationally at the most important forecast centers around the world to provide short and medium range forecasts and also to generate seasonal forecasts and climate change scenarios.

In this work three ensemble techniques have been applied to obtain short range regional weather forecasts over South America. The performance of these ensembles has been objectively tested using different sources of data. Special attention has been given to development of Probabilistic Quantitative Precipitation Forecasts (PQPF). In most cases the PQPF derived from the ensemble proved to be unreliable, for this reason several calibration techniques have been applied and verified in order to correct systematic biases. The calibration of the PQPF increases the reliability of the forecasts and consistently the economical value of the information provided to the public.

The results of this thesis quantify how ensemble forecasting can improve forecast value over South America.

# Capítulo 1:

## Introducción

### 1.1 Antecedentes sobre el tema.

Los estudios realizados por Lorenz (1963), pusieron de manifiesto el carácter caótico del comportamiento de la atmósfera y por ende, la existencia de incertidumbre en la realización de pronósticos basados en la simulación numérica de la misma. Esta incertidumbre aumenta a medida que se busca realizar pronósticos con mayor antelación y ya desde los primeros estudios se estimó que un plazo de 2 semanas constituía el máximo período “pronosticable” del sistema. A partir de este hecho, se puso de manifiesto que no era posible conocer con exactitud las condiciones futuras del sistema y muchos esfuerzos se orientaron a lograr reducir el error de las predicciones, a entender su comportamiento y eventualmente a estimar cuál es la incertidumbre asociada a los pronósticos del tiempo.

Para poder extender el plazo de utilidad de los pronósticos se han realizado grandes avances en el campo de la asimilación de datos, buscando conocer con mayor precisión las condiciones iniciales y también en el desarrollo de los modelos, para representar más adecuadamente los procesos físicos que determinan el comportamiento de la atmósfera (Kalnay, 2003). Estos avances tienden a reducir la incertidumbre mejorando las herramientas disponibles para llevar a cabo el pronóstico. Por otro lado, Epstein (1969) introdujo la idea del pronóstico estocástico-dinámico, el mismo consiste en intentar determinar la función de distribución de probabilidad de los posibles estados futuros de la atmósfera a partir de un estado inicial. La forma propuesta para determinar esta distribución de probabilidades fue la de realizar múltiples simulaciones de la atmósfera utilizando condiciones iniciales o modelos ligeramente diferentes y de esta manera determinar algunas propiedades (varianza, valor medio) de dicha distribución que no podía ser conocida en detalle dado el gran número de grados de libertad que presenta el problema. Esta metodología se conoce con el nombre de pronóstico por ensambles o conjuntos. Una de las hipótesis fundamentales del pronóstico por ensambles es que si la forma de obtener los diferentes pronósticos es adecuada, las propiedades estadísticas del conjunto se asemejan a las propiedades estadísticas de la

verdadera función de distribución de los estados futuros de la atmósfera.

En las últimas dos décadas, esta metodología que permite reducir la incertidumbre a la par que provee información sobre dicha incertidumbre a lo largo del plazo de pronóstico, ha sido aplicada en modelos globales a mediano plazo (Toth y Kalnay, 1993, Buizza 1997), y también en el pronóstico regional a corto plazo (48-72 horas) (Brooks et al. 1995, Hamill y Colucci 1997, Du et al. 1997, Hou et al. 2001, Chessa et al. 2004) con resultados satisfactorios. El pronóstico por ensambles se ha implementado en forma operativa en los principales centros de pronóstico del mundo como por ejemplo el NCEP (National Centers for Environmental Prediction) (Toth y Kalnay, 1993), el ECMWF (European Centre for Medium Range Weather Forecast) (Molteni et al. 1996, Buizza 1997, Molteni y Buizza 1999) o el CPTEC (Centro de Previsión del Tiempo y del Clima), (Coutinho 1999), entre otros.

Dado el alto costo computacional que trae aparejado el realizar múltiples simulaciones de la atmósfera en el tiempo requerido para proporcionar un pronóstico operativo, el número de miembros que compone un ensamble es limitado (10-50). Por otra parte diversos estudios han mostrado que si bien el espacio de posibles perturbaciones que se pueden agregar en la condición inicial es muy grande (del orden de magnitud de las variables pronosticadas por el modelo), las dimensiones del subespacio conformado por aquellas perturbaciones que experimentan un rápido crecimiento y que dominan los errores del pronóstico en el corto y mediano plazo no son tan grandes (Toth y Kalnay 1993, 1997, Molteni et al. 1996), es por eso que la incertidumbre en la condición inicial y el posterior crecimiento de errores en el pronóstico, puede ser adecuadamente representado con un número relativamente pequeño de perturbaciones en la condición inicial. No obstante, es necesario contar con alguna metodología que detecte cuál es la estructura de aquellas perturbaciones que pueden crecer más rápidamente en el plazo de pronóstico y que además estén relacionadas de alguna manera con los errores presentes en la condición inicial, derivados de los métodos implementados en los ciclos de asimilación de datos. Con el objetivo de encontrar las perturbaciones más adecuadas para capturar las perturbaciones con una tasa de crecimiento mayor se desarrollaron diversos métodos, entre los cuales se cuenta la técnica de vectores criados (Toth y Kalnay 1993, 1997). Esta técnica se utiliza en forma operativa en el NCEP. Otras técnicas para la generación de ensambles son las de los vectores singulares (Molteni et al. 1996, Magnusson et al. 2008 implementada en el ECMWF) y la de perturbaciones basadas en funciones ortogonales empíricas (Zhang y Krishnamurti, 1999 implementada en el CPTEC).

Es importante destacar que la perturbación de las condiciones iniciales y eventualmente de las condiciones de borde, representa únicamente la sensibilidad del sistema a los errores presentes en el análisis. No obstante, los errores presentes en los pronósticos se deben también a los errores del modelo. Existen diversas fuentes de errores vinculados con los modelos numéricos que tienen que ver con los métodos de discretización espacial y temporal de las ecuaciones y errores debido a la incapacidad de los modelos para representar fenómenos de escalas menores a la determinada por su resolución horizontal (i.e. aquellos debidos a la parametrización de efectos de menor escala). Siguiendo la idea de Epstein (1969) y Leith (1974), ambos errores pueden abordarse mediante la generación de conjuntos de pronósticos. Por ejemplo, para considerar los errores presentes en la formulación de los modelos, generalmente se construyen ensambles en donde cada miembro es un modelo distinto (poseen, por ejemplo, diferentes métodos de discretización) o bien dentro de un mismo modelo se utilizan diferentes parametrizaciones (Yang y Arritt, 2002, Krishnamurti et. al. 1999). Cabe destacar que es una práctica común en los centros operativos combinar las dos estrategias de generación de ensambles para obtener las ventajas operativas de ambas (Hou et al. 2001). Un tipo particular de implementación de esta metodología, consiste en combinar los pronósticos de varios centros operativos que utilizan modelos, sistemas de asimilación de datos y parametrizaciones diferentes, para conformar lo que se conoce como super-ensamble de modelos (Ebert 2001, Arribas et al. 2005).

Una de las grandes ventajas de los pronósticos por ensambles es que proveen una estimación más precisa de los estados futuros de la atmósfera a través de la media del ensamble, que actúa como un filtro dinámico sobre las componentes menos previsibles del flujo. Por otro lado, la dispersión (entendida como la desviación estándar entre los miembros del ensamble) provee una medida de la incertidumbre en la previsión. La incertidumbre puede ser considerada como un aspecto invariante (por ejemplo basado en estadísticas de largo plazo sobre el desempeño de un pronóstico sobre una región dada), pero los ensambles podrían aportar información extra sobre cómo se modifica dicha incertidumbre para cada situación en particular (Houtekamer 1993). Este valor agregado en la información, impacta directamente sobre los productos o guías que se pueden desarrollar y que intervienen en el proceso de toma de decisión de los diferentes usuarios del pronóstico. Un ejemplo claro de esto es la posibilidad de obtener pronósticos probabilísticos objetivos, que permiten una utilización más eficiente de la información meteorológica, a la par que

incorporan la incertidumbre asociada a las simulaciones numéricas de la atmósfera.

La incorporación de los pronósticos probabilísticos derivados de los ensambles en el proceso de toma de decisión produce un aumento significativo del valor de la información meteorológica (Richardson 2000, Buizza 2001, Zhu et al. 2002). No obstante, para lograr aprovechar al máximo las ventajas del pronóstico probabilístico, es necesario que el mismo sea confiable en términos estadísticos. La confiabilidad del pronóstico probabilístico entendida como la relación entre la probabilidad pronosticada y la frecuencia observada del evento, puede mejorarse mediante diferentes métodos de calibración estadística de los pronósticos. Se han propuesto diferentes metodologías para obtener pronósticos confiables a partir de un ensamble (Hamill y Colucci 1997, 1998, Stensrud y Yussouf, 2007, Raftery et al., 2005, entre otros), como así también, para obtener pronósticos probabilísticos sin la necesidad de generar un ensamble de modelos (Gallus, y Seagal 2004). Esto último resulta en una alternativa interesante, además de constituir un punto de referencia respecto del valor agregado que puede esperarse de un pronóstico por ensambles, respecto del costo computacional que representa.

Sobre la región de Sudamérica, existen pocos estudios orientados a analizar la potencialidad de las técnicas de ensamble para mejorar el pronóstico regional (Ruiz et al. 2005, Silva Dias et al. 2006, Ruiz et al. 2006, Coutinho 1999, Mendonça y Bonatti 2008). En particular, Silva Dias et al. 2006, analizan la mejora que se obtiene al aplicar la técnica del súper ensamble de modelos al pronóstico de la temperatura de superficie. Los resultados obtenidos muestran que la media del ensamble es un mejor estimador de la temperatura de superficie en términos del error cuadrático medio. Por otra parte, Ruiz et al. 2005, analizan los resultados obtenidos al aplicar la técnica de los vectores criados para generar un ensamble en una situación de baja predecibilidad. En ese trabajo, se observó que la media del ensamble produjo una mejora en los pronósticos de precipitación asociados a un sistema convectivo ocurrido en el norte Argentino. Mondonça y Bonatti 2008, analizaron el impacto sobre el sistema de pronóstico global del CPTEC, modificaciones introducidas al método de generación de perturbaciones basado en funciones ortogonales empíricas. Las mismas generaron un impacto positivo sobre los pronósticos por ensambles.

Sobre la región de Sudamérica, existen pocos estudios orientados a analizar la potencialidad de las técnicas de ensamble para mejorar el pronóstico regional (Ruiz et al. 2005, Silva Dias et al. 2006, Ruiz et al. 2006). En particular, Silva Dias et al. 2006, analizan la mejora que se obtiene al aplicar la técnica del súper ensamble de modelos al pronóstico de la

temperatura de superficie. Los resultados obtenidos muestran que la media del ensamble es un mejor estimador de la temperatura de superficie en términos del error cuadrático medio. Por otra parte, Ruiz et al. 2005, analizan los resultados obtenidos al aplicar la técnica de los vectores criados para generar un ensamble en una situación de baja predecibilidad. En ese trabajo, se observó que la media del ensamble produjo una mejora en los pronósticos de precipitación asociados a un sistema convectivo ocurrido en el norte Argentino.

Estos trabajos sugieren que la aplicación de pronósticos por ensambles puede introducir mejoras significativas en el pronóstico regional a corto plazo en variables de interés para los usuarios finales (i.e. la temperatura en superficie y la precipitación). No obstante, de lo antes expuesto se hace evidente la necesidad de realizar experimentos que permitan indagar más a fondo sobre las propiedades de los pronósticos por ensamble sobre la región. Por ejemplo, es necesario evaluar la potencialidad de pronósticos que incorporen una medida de la incertidumbre en los mismos, la cual es obtenida a partir de la dispersión que presentan los miembros del ensamble. También es importante establecer cuáles métodos de generación de ensambles es más eficiente en términos de costo/beneficio. Por otra parte, resulta de interés conocer la relación que existe entre la dispersión del ensamble y la probabilidad de ocurrencia de errores en el pronóstico, dado que esto permitiría estimar la calidad del pronóstico regional en cada situación particular.

Finalmente resulta de importancia particular estudiar, entre otras, las cualidades del pronóstico probabilístico de precipitación y su potencialidad para incrementar el valor de la información meteorológica.

## **1.2 Objetivos:**

La pregunta central que se formula esta tesis es la siguiente: ¿Cómo puede la utilización de pronósticos por ensambles regionales mejorar la predicción a corto plazo sobre Sudamérica?

La hipótesis básica es que la aplicación de esta metodología permite obtener, por un lado, una reducción de la incertidumbre mediante la utilización de la media de los miembros y por otro lado, una estimación de la incertidumbre asociada a cada una de las situaciones particulares, medida a través de la dispersión entre miembros del ensamble.

Responder la cuestión planteada requiere una serie de pasos que implican la concreción de los siguientes objetivos particulares.

- Explorar las propiedades de un ensamble basado en vectores criados. Y realizar comparaciones con los resultados obtenidos con otras técnicas como el SLAF (Scaled Lagged Averaged Forecast, Ebisuzaki y Kalnay, 1991) y el súper ensamble de modelos.
- Desarrollar pronósticos que incluyan una medida de la incertidumbre: por ejemplo, los pronósticos probabilísticos.
- Comparar las propiedades de pronósticos probabilísticos generados a partir de diferentes ensambles.
- Implementar técnicas de calibración estadística que mejoren la confiabilidad de los pronósticos probabilísticos derivados de diversos tipos de ensamble. Dentro de este marco también se contempla la utilización de diferentes conjuntos de datos para realizar la calibración.
- Definir estrategias para la aplicación y verificación de la calidad de pronósticos probabilísticos, adecuadas para su uso operativo.

### **1.3 Naturaleza del aporte original:**

El aporte original de esta tesis se centra en profundizar en el conocimiento de cuál es la potencialidad de los pronósticos regionales por ensambles sobre Sudamérica, con el fin último de mejorar la calidad de la predicción a corto y mediano plazo. Para ello se aborda el problema desde diferentes aspectos.

En primer lugar es necesario generar un sistema de pronóstico por ensambles que sea viable de implementar en el país. Esta tesis se plantea la evaluación de 3 técnicas diferentes, dos de las cuales (vectores criados y SLAF) no han sido empleadas previamente sobre esta región

Las propiedades de los pronósticos regionales por ensambles para Sudamérica en lo que respecta al pronóstico a corto plazo no han sido abordadas en trabajos previos. El conocimiento de estas propiedades, como por ejemplo la relación entre la dispersión del ensamble y los errores en el pronóstico, es fundamental para sustentar desde el punto de vista teórico la incorporación en los pronósticos operativos de medidas de la incertidumbre.

Otro aspecto importante, íntimamente relacionado con el anterior, es la generación de pronósticos probabilísticos como medio para incorporar medidas de la incertidumbre en

la previsión de una determinada variable. En particular se pondrá énfasis en el pronóstico probabilístico de precipitación. Esta es la primera vez que se realiza un estudio de la potencialidad de los pronósticos probabilísticos de precipitación sobre Sudamérica que además incluye un análisis sobre el impacto que la implementación de estas técnicas podría tener sobre el valor de la información meteorológica. El análisis se plantea de forma tal que se pueda comparar pronósticos probabilísticos generados por diferentes ensambles y considerando regiones con regímenes de precipitación diferentes lo cual aporta al conocimiento general sobre las fortalezas y debilidades de las técnicas implementadas.

Se espera que los resultados obtenidos a partir de esta tesis, puedan contribuir a montar el primer sistema operativo de pronóstico regional por ensambles en Argentina y a realizar un óptimo aprovechamiento de los productos que se puedan derivar del mismo, a la par que permitan continuar profundizando nuestro conocimiento acerca de la predecibilidad y la problemática del pronóstico a corto y mediano plazo sobre la región.

## Capítulo 2:

### Herramientas y datos utilizados.

En este capítulo, se describen las herramientas y los conjuntos de datos utilizados. En particular se presenta una breve descripción de los modelos numéricos empleados y de los controles de calidad aplicados a los datos observados.

#### 2.1 Descripción del modelo WRF

El modelo WRF (Weather Research and Forecasting) es un modelo de mesoescala desarrollado mediante un esfuerzo conjunto entre diferentes instituciones estadounidenses entre las que se cuentan el NCAR (National Centres for Atmospheric Research) y el NCEP (National Centres for Environmental Prediction). Este modelo, está disponible libremente en Internet (<http://www.mmm.ucar.edu/wrf/users/>) y está preparado para funcionar sobre una amplia variedad de plataformas. La distribución del modelo incluye además paquetes que facilitan el pre-procesamiento de los datos, es decir la generación de condiciones iniciales y de borde a partir de diferentes fuentes de datos y el post-procesamiento de los resultados generados. Por otra parte existe un paquete de asimilación de datos que también puede ser descargado y utilizado libremente. La versión utilizada en el presente estudio es la 2.2.0.

Entre los diferentes componentes de un modelo numérico, podemos contar el conjunto de métodos numéricos utilizados para resolver las ecuaciones, lo cual se describe en la bibliografía como núcleo dinámico, y el conjunto de parametrizaciones, que constituyen la física del modelo. En la actualidad el modelo WRF cuenta con 2 núcleos dinámicos, el WRF-ARW (Advanced Research WRF) y el WRF-NMM (Non-hydrostatic Mesoscale Model). En el presente estudio se utilizó el núcleo ARW y es el que se describe a continuación.

De acuerdo con Skamarock (2005) las principales características del WRF-ARW son:

- Ecuaciones compresibles y no hidrostáticas (teniendo la posibilidad de realizar simulaciones en modo hidrostático).
- Coordenada vertical sigma-p con la posibilidad de modificar la cantidad y posición de los niveles en la vertical. El tope del modelo se fija en un nivel de presión constante que puede ser modificado de acuerdo a las necesidades (en los experimentos que se muestran en el presente trabajo se fijó el tope en 50 hPa.).
- Retícula horizontal Arakawa de tipo C. La resolución horizontal del modelo es variable. Por otra parte, existen 3 tipos de proyección que pueden ser utilizadas: Estereográfica polar, Lambert y Mercator. En el presente trabajo se utilizó la proyección de Lambert dado que reduce los problemas vinculados con la variación de la resolución horizontal dentro del dominio.
- Integración temporal a partir del método de Runge Kutta de orden 3. La integración temporal se encuentra subdividida de forma tal que los términos asociados a los modos más rápidos (ondas sonoras y ondas de gravedad) son integrados utilizando un paso de tiempo menor, garantizando así la estabilidad de la solución (Klemp et al. 2007). El paso de tiempo utilizado en los experimentos fue de aproximadamente 150 segundos.
- El orden del método para el cálculo de las derivadas horizontales y verticales puede ser ajustado de acuerdo a las necesidades del usuario.
- El modelo WRF presenta también la posibilidad de incluir retículas anidadas con interacción en una y dos direcciones. En particular también presenta la posibilidad de incluir retículas móviles. Estas capacidades no fueron utilizadas en el presente estudio.
- Las condiciones de borde se introducen en el modelo mediante un término de relajación newtoniana en los 5 puntos de retícula más cercanos al borde. En el punto más externo las variables son iguales a las que indican los datos utilizados como condición de borde.

Dependiendo de las circunstancias, el modelo WRF-ARW no requiere la inclusión de términos difusivos en las ecuaciones para garantizar la estabilidad numérica de la solución. De acuerdo con Klemp et al. (2007) la utilización de métodos de orden impar en la discretización de las derivadas espaciales es equivalente a la introducción de cierta difusión en el modelo. Esta difusión es (siempre y cuando el espaciamiento entre puntos de retícula sea menor que 10 Km) suficiente para mantener la estabilidad de la solución.

Si bien existe gran controversia sobre este punto en particular, algunos resultados parecen indicar que la ausencia de un tratamiento explícito de la difusión le permitiría al modelo WRF-ARW retener una mayor cantidad de energía en las escalas pequeñas dando lugar a una representación más adecuada del espectro de energía (i.e. soluciones más realistas en términos de la variabilidad espacial y temporal de las variables en escalas pequeñas). En todos los experimentos realizados en el presente trabajo de Tesis, no se incluye un tratamiento explícito para la difusión más allá del provisto por el esquema numérico y por la parametrización de la capa límite.

Con respecto a la física del modelo, el WRF-ARW permite la inclusión de varias opciones para el tratamiento de la convección, la interacción con la superficie, la turbulencia y la capa límite planetaria, la radiación y la microfísica. Una descripción somera de estas parametrizaciones puede encontrarse en Skamarock et al. (2005) mientras que Blazquez (2007) presenta experimentos realizados con diferentes configuraciones sobre la región de Sudamérica.

No todos los experimentos descritos en la presente Tesis fueron realizados utilizando las mismas parametrizaciones, es por eso que la configuración del modelo se describe en forma independiente para cada experimento. Las principales características de las parametrizaciones utilizadas son las siguientes:

- Radiación: Se utilizaron las parametrizaciones de Dudhia (1989) que considera una única banda y la interacción con la nubosidad para la onda corta y la parametrización RRTM (Rapid Radiative Transfer Model) que incluye 16 bandas, la interacción con las nubes, el CO<sub>2</sub> y el O<sub>3</sub>, para la onda larga.
- Capa límite: Se utilizaron 2 esquemas, el esquema YSU (Yon-Sei University, Hong y Pan (1996)) que es un esquema no local que incluye un término para considerar los transportes contragradiante y el esquema MYJ (Mellor, Yamada y Janjic, Janjic (2002)) que es un esquema local que utiliza una clausura de orden 2.5.
- Convección: Se utilizaron los esquemas de Grell-Devenyi (Grell y Devenyi, 2002) que estima el efecto de la convección utilizando un ensamble de parametrizaciones utilizando por ejemplo diferentes hipótesis de clausura. También se utilizó el esquema de Kain-Fritsch (Kain, 2004) que se basa en un modelo simplificado de nube que considera mezcla lateral en diferentes niveles y el efecto de las descendentes húmedas. El esquema incluye también un tratamiento de la convección chata.

- Modelo de suelo: Se utilizó el modelo Noah (Chen y Dudhia, 2001). Este es un modelo de 4 capas que posee un tratamiento explícito para la humedad de suelo. Incluye información sobre la cobertura vegetal y el tipo de suelo. Permite simular la presencia de una cubierta de nieve con hasta 2 capas de resolución vertical.
- Microfísica: Se utilizó el esquema Ferrier ETA (Ferrier et al. 2002) que está particularmente diseñado para trabajar en resoluciones intermedias en las cuales no se pueda resolver explícitamente las nubes aumentando la eficiencia computacional. La variable que se advecta entre puntos de retícula vecinos es el condensado total, mientras que un análisis local determina la proporción de las diferentes especies (agua de nube, lluvia, hielo de nube, nieve, graupel y gotas congeladas) que constituyen ese condensado.

Además de las diferentes parametrizaciones y métodos de resolución numérica existentes, la distribución del modelo WRF provee acceso a una base de datos de características de la superficie que se utilizan principalmente para el modelo de suelo, aunque también intervienen en la dinámica general del modelo como en el caso de la altura del terreno.

Los datos que se utilizaron en los experimentos de esta tesis son los siguientes:

- Topografía con una resolución horizontal de 30" provenientes de la United State Geological Survey (USGS).
- Tipo de suelo con una resolución horizontal de 30" (USGS). Estos datos en particular proveen al modelo de suelo información sobre la distribución horizontal de las capacidades caloríficas, conductividad, así como de variables que tienen que ver con la humedad del suelo.
- Uso de suelo con una resolución horizontal de 30" (USGS): Estos datos proveen al modelo de superficie con datos sobre el tipo de cubierta vegetal en el suelo, el albedo y la emisividad de la superficie entre otros. En el caso del albedo, el mismo puede ser especificado a través de la distribución del uso de suelo o bien se puede utilizar un albedo climatológico basado en observaciones satelitales el cual también es provisto en la distribución del WRF. Para el caso del modelo de suelo de 5 capas, la distribución del uso de suelo determina también la distribución de la humedad del suelo ya que en este modelo en particular cada uso de suelo tiene asociado un valor constante de humedad.

Actualmente, este archivo cuenta con 24 categorías de uso de suelo diferentes (USGS Land Use / Land Cover System, Anderson et al., 1976).

- Vegetación con una resolución de 30": Este conjunto de datos provee campos de distribución espacial de la cobertura vegetal en función de la época del año (un campo por mes). Estos datos complementan los datos de uso de suelo. La vegetación interviene principalmente en la modificación del flujo de vapor entre la superficie y la atmósfera y en la determinación del parámetro de rugosidad que interviene en la parametrización de la capa de superficie.

Por lo antes descrito, el modelo WRF constituye una herramienta útil y muy versátil para realizar trabajos de investigación y tareas operativas, destacándose la posibilidad de contar con múltiples parametrizaciones para los diferentes procesos involucrados e incluso 2 núcleos dinámicos y la posibilidad de trabajar con diferentes resoluciones y utilizando retículas anidadas. Estos fueron los principales argumentos que motivaron su elección para el presente trabajo. Por otra parte, su libre distribución hace que un gran número de personas trabajen con este modelo en diferentes aplicaciones en todo el mundo logrando así una más eficiente detección de problemas y mayor facilidad de acceso a la información y ayuda para resolver problemas concretos relacionados con su funcionamiento. No obstante, resulta necesario documentar el desempeño del WRF sobre nuestra región, dado que una configuración puede resultar adecuada sobre determinadas regiones y no sobre otras. Por esa razón se ha considerado que esta evaluación constituirá una parte de los objetivos del presente trabajo de Tesis.

## **2.2 Descripción del modelo MRF**

El modelo MRF (Medium Range Forecast) es global y fue desarrollado por el NCEP en la década del 80' (Sela, 1980). Desde su implementación operativa, se fueron realizando numerosas modificaciones a lo largo de su desarrollo (Kalnay et al. 1990). La versión utilizada en este trabajo data del año 1999.

El MRF es un modelo espectral en ecuaciones primitivas que utiliza la vorticidad y la divergencia como variables de pronóstico. En la vertical, emplea la coordenada sigma. En particular los experimentos realizados en este trabajo, utilizaron 62 componentes espectrales y 28 niveles en la vertical (T62L28) lo cual equivale a una resolución espacial

de 2.5 grados aproximadamente.

El esquema de integración temporal es un "leapfrog" para la advección y un esquema implícito para los términos relacionados con las ondas de gravedad.

En cuanto a las parametrizaciones, el esquema de la radiación incluye un tratamiento para las distintas longitudes de onda (corta y larga) que contempla la interacción con el ozono, el dióxido de carbono, el vapor de agua y las nubes. La convección profunda es parametrizada mediante un esquema del tipo Arakawa y Schubert simplificado (Grell 1993). Dicha simplificación consiste en considerar una única altura posible para las nubes convectivas en cada punto de retícula en lugar de considerar una población de nubes de diferentes alturas como se indica en la versión original del método de Arakawa y Schubert (1974). También se incluye una parametrización de la convección chata.

La precipitación es el resultado de la parametrización de la convección y de la parametrización de la nubosidad de gran escala que tiene en cuenta el proceso de evaporación de la precipitación. La capa límite está parametrizada utilizando una estimación de la altura de la misma en base al número de Richardson. El perfil vertical del coeficiente de transporte turbulento dentro de la capa límite se obtiene como función de la altura de la misma. También se incluye un término de transporte contragradiante (Hong y Pan, 1996). En particular, esta parametrización de la capa límite es una versión previa del esquema YSU implementado en el modelo WRF.

El tratamiento de la superficie incluye un modelo de suelo de 2 capas con ecuaciones de pronóstico para la temperatura y la humedad del suelo. El tratamiento de la humedad del suelo incluye el efecto de la vegetación. La cobertura de nieve también interactúa con la superficie modificando los flujos de calor sensible y latente.

La versión utilizada del modelo cuenta además con un programa de post procesamiento que lleva la solución a un reticulado de tipo gaussiano en formato GRIB.

## **2.3 Datos de precipitación observados:**

### **2.3.1 Fuentes de datos.**

Los datos de precipitación observados utilizados en los diversos experimentos provienen de 3 fuentes:

Datos pluviométricos de la red de precipitación del experimento SALLJEX (South America Low Level Jet Experiment) (Vera et al. 2006). Esta red está conformada por la red operativa del Servicio Meteorológico Nacional y redes de diversas instituciones en el interior del país que colaboraron con el experimento de campo SALLJEX. La Figura 2.1 a, muestra el número total de datos calculado en cajas de 40 km x 40 km, sobre todo el período. En la misma se puede ver que la mayor densidad de datos se da sobre el centro y norte de Argentina.

Los datos SALLJEX cuentan con un control de calidad previo que se describe en Penalba et al. (2004). En dicho trabajo, se analizaron los datos de precipitación utilizando como apoyo información de precipitación estimada por satélite. Los datos catalogados como "dudosos" según ese procedimiento, no fueron utilizados en el presente trabajo. Además, como parte del procesamiento de los datos que se realizó para construir los conjuntos de datos utilizados en el presente estudio, se realizó un control adicional que se describe en la sección 2.3.2.

Otra fuente de datos utilizada, fue la Agencia Nacional de Aguas de Brasil (ANA) que cuenta con mas de 3000 estaciones ubicadas principalmente en el sudeste de Brasil (Figura 2.1 b). Estos datos fueron consistidos al ser incorporados en los análisis diarios de precipitación elaborados por Liebmann y Allured (2005) y fueron generosamente provistos por los autores para la realización del presente trabajo. Estos datos también fueron objeto de los mismos chequeos básicos que los aplicados a los datos SALLJEX.

Durante el período abarcado por el experimento SALLJEX (15 de noviembre de 2002 al 15 de febrero de 2003), se utilizaron los datos SALLJEX en combinación con los de la red ANA. Para aquellos períodos que no coinciden con el experimento de campo SALLJEX, se utilizaron los datos del GTS (Global Telecommunication System). Estos datos son los provenientes de la red de estaciones operativa y tienen una densidad mucho menor (aproximadamente 300 estaciones sobre Sudamérica, Figura 2.2) que las fuentes discutidas previamente.

Todas las fuentes discutidas proveen principalmente datos de precipitación acumulada en 24 horas a las 12 UTC. Si bien algunas estaciones proveen precipitación acumulada en 6 horas, estos datos no fueron utilizados en el presente estudio.

### **2.3.2 Control de calidad adicional.**

El control de calidad adicional que se realizó consiste básicamente en un control sobre los valores de precipitación acumulada en 24 horas. Se removieron del conjunto de datos aquellas estaciones en donde la cantidad de precipitación acumulada en 24 horas superara los 300 mm.

Otro de los controles realizados estuvo orientado a consistir datos de la misma estación que provenían de distintas fuentes. Esto ocurre sobre todo en el caso de los datos SALLJEX en donde dos o más instituciones frecuentemente incorporan información de la misma estación. En estos casos, el algoritmo verifica que el dato de precipitación informado por las distintas fuentes coincida. Cuando no existe coincidencia, el dato puede ser descartado o no dependiendo de lo siguiente:

- Si la diferencia es menor a 2.5 mm (valor fijado arbitrariamente) el dato se conserva y es igual al promedio de las dos observaciones encontradas.
- Si la diferencia es mayor y ambos datos son distintos de 0 entonces el dato se descarta.
- Si uno de los datos es 0: en el caso de las estaciones de la red GTS, muchas veces los datos vienen repetidos y es frecuente encontrar que uno de los informes corresponde a un 0 de precipitación, mientras que el otro valor es mayor que 0. Es por eso que en estos casos el dato 0 se descartó y se contabilizó la precipitación mayor que 0 informada. No obstante, este tipo de error también se encontró asociado a redes que informaban en forma tardía la precipitación ocurrida. Estos casos se analizaron más en profundidad y se llegó a la conclusión de que la fecha de los datos de la red SAGyP (Datos recolectados por la Secretaría de Agricultura, Ganadería y Pesca) en algunas estaciones no estaban indicados correctamente. Como consecuencia de esto, se removió de la base de datos a la red SAGyP.

## 2.4 Estimaciones de precipitación: CMORPH

El algoritmo CMORPH (NOAA Climate Prediction Center Morphing Technique) está basado en el uso de microondas pasivas para estimar la tasa de precipitación asociada a los sistemas nubosos y al uso de las imágenes infrarrojas de los satélites GOES para estimar el desplazamiento de los mismos (ver Joyce et al. (2004) para la descripción detallada del algoritmo). En particular las estimaciones de las tasas de precipitación en base a microondas pasivas son realizadas en base a una constelación de satélites que orbitan la tierra y que están equipados con diferentes sensores. Para cada uno de estos sensores existen algoritmos de calibración diferentes (dependiendo de las bandas disponibles en cada sensor), las diferentes estimaciones son normalizadas utilizando las estimaciones del satélite TRMM (Tropical Rainfall Measurement Mission) que cuenta además con un radar de microondas activas.

La estimación basada en microondas pasivas resulta ser más exacta que las anteriores estimaciones basadas exclusivamente en la temperatura de los toques nubosos (Joyce et al. 2004; Ebert et al., 2007). Estas estimaciones cuentan con una resolución espacial máxima de 0.072 grados y una resolución temporal máxima de media hora. Para el presente estudio, se partió de los productos con una resolución intermedia de 0.125 grados y acumulados en períodos de 3 horas, que luego fueron acumulados sobre 24 horas.

Los datos CMORPH pueden ser descargados libremente en: [http://www.cpc.ncep.noaa.gov/products/janowiak/cmorph\\_description.html](http://www.cpc.ncep.noaa.gov/products/janowiak/cmorph_description.html).

Las estimaciones de precipitación CMORPH han sido verificadas contra datos pluviométricos en diferentes regiones del mundo (Joyce et al. 2004, Ebert et al., 2007). Sobre la región de Sudamérica, Ruiz (2008) realizó un análisis de la calidad del CMORPH comparada contra los mismos datos de la red pluviométrica empleados en esta tesis. Además, Ruiz (2008) utiliza diferentes estrategias tendientes a remover la componente sistemática del error en las estimaciones CMORPH sobre la región. El método que mejores resultados aportó fue el del ajuste de las frecuencias estimadas en base a las frecuencias observadas para diferentes umbrales de precipitación (similar al algoritmo de eliminación de los desvíos sistemáticos de los pronósticos de lluvia utilizado en Hamill y Whitaker, 2006). Básicamente este algoritmo identifica qué valor de precipitación estimada presenta una frecuencia de ocurrencia similar a un determinado valor de precipitación observada,

repetiendo este análisis para diferentes valores de precipitación observados, se obtiene una curva que permite reemplazar los valores estimados, por aquel valor observado que tiene la misma frecuencia de ocurrencia. Este procedimiento permite realizar un ajuste no paramétrico que resulta ser dependiente de la cantidad de lluvia estimada por el CMORPH. En esta tesis se utilizan las estimaciones del CMORPH calibradas como en Ruiz (2008) y las estimaciones no calibradas. Es importante destacar que si bien en este caso existe un abundante caudal de datos pluviométricos que permiten emplearlos tanto en la verificación de pronósticos, como en su posterior calibración, la red de medición operativa de precipitación sobre la región presenta áreas en donde las mediciones son muy escasas (por ejemplo, sobre Bolivia y Paraguay). La inclusión de los datos CMORPH esta orientada a poder aplicar las metodologías propuestas en el presente trabajo a períodos distintos a la estación cálida 2002-2003, en donde la disponibilidad de datos de superficie puede ser sustancialmente menor.

En el presente estudio, la calibración obtenida esta dirigida a la porción sur y sudeste de Sudamérica. Sobre la región norte del dominio (al norte de  $20^{\circ}$  S), los errores sistemáticos de los CMORPH con respecto a los datos pluviométricos de las redes SALLJEX y ANA son mayores que los encontrados al sur de  $20^{\circ}$ S, es por eso que si bien en la porción sur del dominio (al sur de  $20^{\circ}$  S) el comportamiento de los errores sistemáticos es bastante homogéneo, la diferencia entre la región norte y la sur debería ser tomada en cuenta al momento de realizar la calibración de los datos. La corrección aplicada en este caso, no tiene en cuenta la dependencia del error sistemático del CMORPH con la región considerada, cuyo impacto en los experimentos realizados será discutido en el Capítulo 6.

## **2.5 Datos de superficie.**

Los datos de superficie utilizados para la verificación del modelo WRF en los diferentes experimentos, provienen de la red GTS que coincide con la posición de las estaciones de medición de precipitación que se muestra en la Figura 2.2. Esta red provee datos en forma tri-horaria de temperatura, temperatura de rocío, presión reducida a nivel del mar, presión de la estación y las componentes del viento a 10 metros. Sobre estos datos no se realizó un control de calidad sistemático.

Si bien en principio las estaciones envían información cada 3 horas, existe una sensible reducción de la cantidad de datos disponibles, sobre todo a las 6 UTC (3 AM hora

local), siendo 00 y las 12 UTC (21 y 9 hora local respectivamente) las horas en que se cuenta con mayor cantidad de información.

## **2.6 Análisis en puntos de retícula**

Los análisis en punto de retícula fueron utilizados para la inicialización de los experimentos con los modelos WRF y MRF. Para la inicialización del modelo global (MRF) se utilizaron los reanálisis del NCEP (Kalnay et al., 1996) en coordenadas espectrales, que fueron gentilmente preparados por la Dra. Jae Schemm para poder ser utilizados con la versión del modelo utilizada en el presente trabajo. Los reanálisis del NCEP proveen información sobre las variables de atmósfera e información de las variables de superficie en 2 niveles de suelo.

Para la inicialización del modelo regional, se utilizaron los análisis operativos del NCEP GDAS (Global Data Assimilation System). Los mismos fueron almacenados durante el experimento SALLJEX para su posterior utilización en estudios de modelado numérico sobre la región. Estos análisis proveen información del estado de la atmósfera 4 veces al día (00, 06, 12 y 18 UTC) en 26 niveles isobáricos y con una resolución espacial de 1°. También incluyen información sobre las variables de superficie en 2 capas. Actualmente el modelo de suelo cuenta con 4 capas, pero en el momento del experimento SALLJEX solo había información de suelo en 2 niveles.

Durante la estación cálida 2002-2003, se generó un conjunto de análisis basado en el mismo tipo de asimilación utilizado por el GDAS, pero incorporando los datos SALLJEX (Herdies et al 2007) que fueron cedidos gentilmente por Dirceu Herdies. Estos análisis fueron utilizados en experimentos orientados a determinar la relación existente entre los vectores criados y los errores en el análisis como se detalla en el Capítulo 4.

También se utilizaron los reanálisis del proyecto JRA-25 (Japanese 25-year Reanalysis project). Estos análisis están disponibles desde 1979 hasta el 2004 y fueron generados con un modelo global de resolución T106L40 con un esquema 3DVAR para la asimilación de datos. Los datos provistos fueron archivados en puntos de retícula con una resolución uniforme de 2.5°x2.5°. Una descripción de este conjunto de datos se puede encontrar en Onogi et al. (2007).



## **Capítulo 3:**

# **Evaluación preliminar de las ventajas del pronóstico de precipitación por ensambles en la región de Sudamérica**

### **3.1 Introducción**

En este capítulo se describe un experimento orientado a explorar el potencial del pronóstico por conjuntos sobre la región de Sudamérica, en particular en su aplicación a la obtención de pronósticos probabilísticos de precipitación. Uno de los principales objetivos, es determinar las ventajas que el pronóstico probabilístico puede introducir frente a productos similares que pueden ser derivados de pronósticos determinísticos. Para esto se utilizan diferentes estrategias para generar pronósticos probabilísticos, por un lado dos sistemas de pronósticos por conjuntos, uno basado en la técnica de Scaled Lagged Averaged Forecast (SLAF) y otro basado en la técnica de utilización de diferentes configuraciones (Super Model Ensemble System, SMES) y por otro lado, pronósticos probabilísticos derivados a partir de pronósticos determinísticos siguiendo la metodología propuesta por Gallus y Seagal (2004).

Además se ensayan diferentes estrategias que permiten mejorar la utilidad de la información derivada del ensamble. En particular estas estrategias están orientadas a mejorar la relación entre la probabilidad pronosticada y la probabilidad real de que el fenómeno ocurra.

Los pronósticos probabilísticos derivados de las diferentes metodologías propuestas, son comparados entre sí utilizando diferentes medidas que permiten evaluar la confiabilidad, la resolución (o capacidad de discriminar) y la habilidad de cada uno.

## 3.2 Metodologías para la generación de pronósticos probabilísticos

### 3.2.1 Sistemas de pronóstico regional por ensambles utilizados:

#### Conjunto SLAF-WRF

Este sistema de pronósticos por conjuntos está basado en un solo modelo, en donde todos los miembros comparten la misma física (no hay cambios en las parametrizaciones utilizadas). En este caso particular, se utilizó el modelo WRF-ARW versión 2.0 con la siguiente configuración:

- Proyección mercator con 50 km de resolución horizontal y 31 niveles verticales.
- Parametrización Ferrier para la microfísica (Ferrier et al. 2002).
- Parametrización de Grell para la convección (Grell y Devenyi, 2002).
- Esquema RRTM (Mlawer et al., 1997) para la radiación de onda larga y Dudhia (1989) para la radiación de onda corta.
- Capa límite de acuerdo con la formulación Mellor-Yamada-Janjic (Janjic, 2002).
- Modelo de superficie NOAH (Chen y Dudhia, 2001).

En este experimento los diferentes miembros del conjunto se obtienen perturbando las condiciones iniciales y de borde utilizando la técnica de Scaled Lagged Averaged Forecast (SLAF) (Ebisuzaki y Kalnay 1991). Esta técnica fue utilizada por primera vez para el pronóstico regional en Hou et al. (2001). La Figura 3.1, describe la implementación de la técnica de SLAF para generar perturbaciones a las condiciones iniciales y de borde. Supongamos que nos encontramos en un tiempo  $t$ , en el cual contamos con el análisis de un modelo global y una serie de pronósticos previos inicializados en los tiempos  $t-1$ ,  $t-2$ , etc. obtenidos con dicho modelo global. Los mismos serán utilizados para generar perturbaciones a las condiciones iniciales y de borde del dominio regional. Por ejemplo, para perturbar el pronóstico en el día  $t$ , se utiliza el pronóstico global inicializado en el tiempo  $t-1$  de la siguiente manera.

Se computa la diferencia entre el pronóstico para el tiempo  $t$  (inicializado en  $t-1$ ) y el análisis en el tiempo  $t$ . La diferencia se computa en todas las variables de atmósfera que se utilizan para inicializar el modelo regional. En la Figura 3.1, las curvas negras a trazos, muestran la evolución de diferentes pronósticos inicializados en los tiempos  $t-1$  y  $t-$

2, y la línea negra continua representa el análisis. La distancia A-B, representa la diferencia entre el pronóstico inicializado en  $t-1$  en el tiempo  $t$  y el análisis en el tiempo  $t$ . En este caso no se tuvieron en cuenta los errores sistemáticos en el modelo, es decir que la resta entre los pronósticos y el análisis incluye los errores del día y la componente sistemática del error. En una implementación futura de este algoritmo sería mejor utilizar solo la componente no sistemática de las diferencias entre pronóstico y análisis para calcular las perturbaciones.

Como es de esperar y se representa en la Figura 3.1, el pronóstico inicializado en  $t-2$ , muestra mayores diferencias con el análisis en el tiempo  $t$  que aquel inicializado en  $t-1$ . Para contrarrestar este efecto y que todas las perturbaciones tengan una amplitud similar se modifica la amplitud de estas diferencias por un factor (este se aplica a todas las diferencias obtenidas). En este caso particular se eligió un factor que equipara la amplitud media de las perturbaciones en el geopotencial de 200 hPa fijando dicha amplitud en un valor cercano al valor de la amplitud de perturbaciones generadas por el método de los vectores criados (ver Capítulo 4 de esta tesis). Dicha amplitud como se discutirá más adelante, es similar al nivel de incertidumbre presente en la condición inicial que está representado por las líneas grises a trazos en la Figura 3.1.

Las diferencias entre el pronóstico inicializado en  $t-1$  y el análisis en el tiempo  $t$ , con la amplitud modificada se suman y se restan al análisis en el tiempo  $t$  de esta manera se obtiene un par de análisis perturbados. El mecanismo se repite con el pronóstico inicializado en el tiempo  $t-2$ ,  $t-3$ , etc. Es importante notar que en este caso el número de perturbaciones que se pueden obtener está limitado por la longitud máxima de los pronósticos globales y además por el hecho de que si se utilizan diferencias entre el análisis y pronósticos inicializados en un tiempo previo demasiado lejano, estas perturbaciones podrían no tener valor para su implementación en el pronóstico por ensambles. A medida que utilizamos pronósticos más alejados de nuestra condición inicial, la perturbación generada se parece más a una perturbación al azar las cuales no tienen la capacidad de crecimiento que tienen las perturbaciones generadas por SLAF cuando los pronósticos son cercanos al tiempo de inicialización, porque no representan adecuadamente las inestabilidades del flujo en el momento de la inicialización del ensamble. En el presente estudio, se utilizaron pronósticos inicializados hasta 48 horas previas al momento de inicialización del ensamble, espaciados en un período de 12 horas

con lo cual se generaron 4 pares de perturbaciones en la condición inicial que sumados al pronóstico control dan un total de 9 miembros en el ensamble. El intervalo de 12 horas utilizado obedece a que solo estaban disponibles los pronósticos inicializados a las 00 y 12 UTC, pero los pronósticos globales del NCEP se inicializan operativamente 4 veces por día (a las 00, 06, 12 y 18 UTC) lo cual permitiría generar un número mayor de perturbaciones.

Para las condiciones de borde, se procede en forma análoga, solo que en lugar de hacer la resta entre el análisis (que no está disponible para los tiempos futuros) y pronósticos inicializados previamente, se hace la resta entre el pronóstico más reciente y los pronósticos previos que verifican a la misma hora (La Figura 3.1, muestra un ejemplo de esta diferencia indicada por la recta D-E). En este caso particular, el control de amplitud que se hace sobre las perturbaciones es el mismo para la condición inicial o las condiciones de borde. Este procedimiento no es óptimo en el sentido de que lo ideal es permitir que las perturbaciones en las condiciones de borde tengan mayor amplitud que las introducidas en las condiciones iniciales, sin embargo la dispersión en el ensamble obtenido tuvo un comportamiento adecuado (Ruiz et al. 2006). Los pronósticos del ensamble regional se integraron por un plazo de 48 horas. Teniendo en cuenta el procedimiento especificado, se requirieron pronósticos globales de una longitud de 96 horas cada uno para poder perturbar consistentemente las condiciones iniciales y de borde de un pronóstico a 48 horas.

Una de las ventajas fundamentales del método de SLAF es que no requiere el acceso a un pronóstico por conjuntos del modelo global. En la actualidad si bien existen pronósticos por ensambles globales disponibles en tiempo real, la transmisión de un volumen de datos tan grande a través de Internet dificulta una implementación operativa. Por otra parte, las perturbaciones del método SLAF son dinámicamente consistentes (porque provienen de la resta de campos que originalmente lo eran). Y además tienen información de los "errores del día", es decir que las perturbaciones son dependientes del flujo (Kalnay, 2003). Las mayores perturbaciones se dan en donde la discrepancia entre los pronósticos anteriores y el análisis son mayores, es decir en donde los errores del análisis podrían ser importantes.

En el presente trabajo, la técnica de SLAF se aplicó a los pronósticos del Global Forecasting System (GFS). Los mismos están disponibles a través del FTP de la National Oceanic and Atmospheric Administration (NOAA) (<ftp://ftpprd.ncep.noaa.gov>). Se

realizaron 2 experimentos utilizando la misma configuración para el modelo WRF, uno desde el 22 de octubre al 27 de noviembre de 2005 y otro desde el 3 de octubre al 31 de diciembre de 2006. Los dos períodos generados serán utilizados uno como período de entrenamiento y otro como período de verificación como se describe en las secciones subsiguientes.

### **Super model ensemble system (SMES):**

El SMES es un proyecto desarrollado por el laboratorio de Meteorología Aplicada a Sistemas de Tiempo Regional (MASTER) de la Universidad de San Pablo (USP) cuyas características principales están descritas en Silva Dias et al. (2006) y cuyos productos pueden ser consultados a través de Internet en ([http://www.master.iag.usp.br/graf\\_phplot/index.php](http://www.master.iag.usp.br/graf_phplot/index.php)). Este pronóstico por conjuntos es un esfuerzo coordinado de varios centros operativos y de investigación sobre la región de Sudamérica. El SMES estaba compuesto al momento de realizar este experimento por más de 10 modelos diferentes y más de 34 configuraciones, situación que se ha ido modificando a lo largo del tiempo con la incorporación de nuevos miembros. También cuenta con modelos que se inicializan con diferentes condiciones iniciales, obtenidas a partir de análisis desarrollados en diferentes centros operativos. Algunos de los modelos cubren una región relativamente pequeña, es por eso que en el presente trabajo no se utilizan todos los modelos que participan en el SMES, sino que se seleccionan aquellos en donde la cobertura espacial es suficientemente grande, permitiendo así un mayor número de datos para la verificación y para la calibración de los pronósticos probabilísticos de precipitación. El dominio abarcado por los modelos utilizados en el presente trabajo puede verse en la Figura 3.2.

La lista completa de modelos utilizados en el presente experimento, se muestra en la Tabla 3.I. Los miembros individuales del conjunto global del CPTEC, no fueron incluidos en este conjunto dado que se buscaba evitar que el conjunto estuviera sesgado hacia la dinámica de un modelo particular. La mayoría de los miembros del conjunto SMES son inicializados a las 00 y a las 12 UTC. En estos casos, ambas corridas son incluidas en el conjunto como 2 miembros independientes. De esta manera, el SMES no solo cuenta con cambios en la física y en los métodos de resolución numéricos entre los diferentes

miembros sino que además cuenta con perturbaciones en las condiciones iniciales y de borde en algunos miembros. Esto también es cierto para la media del conjunto global de CPTEC que se incluye como un miembro más del conjunto y que contiene información sobre los miembros perturbados del conjunto global.

Modelo	Res. (km)	Descripción adicional
Eta 40 CPTEC	40	Mesinger et al. (1988)
Eta 20 CPTEC	20	Mesinger et al. (1988)
Global T213 CPTEC	63*	Kinter et al. (1997)
BRAMS UBA	80-20	<a href="http://www.bramsuba.com.ar/">http://www.bramsuba.com.ar/</a>
HRM	30	<a href="http://www.mar.mil.br/dhn/chm/meteo/">http://www.mar.mil.br/dhn/chm/meteo/</a>
ETA RPSAS CPTEC	40	<a href="http://www.cptec.inpe.br/~ioweb/index.shtml">http://www.cptec.inpe.br/~ioweb/index.shtml</a>
BRAMS CATT CPTEC	30	Freitas et al. (2006)
WRF CIMA	60-20	Skamarock et al. (2005) ( <a href="http://wrf.cima.fcen.uba.ar">http://wrf.cima.fcen.uba.ar</a> )
ETA UMD	80-22	Mesinger et al. (1988)
MRF NCEP	200*	Kalnay et al. (1990)
Ens. CPTEC	100*	Mendonça et al. (2006)

**Tabla 3.I: Lista de modelos que forman parte del SMES utilizado en el presente trabajo. El (\*) indica resolución aproximada para los modelos espectrales.**

Los pronósticos de precipitación de los miembros del SMES son archivados en la USP por el laboratorio MASTER en puntos correspondientes a la ubicación de estaciones de superficie. El método de interpolación utilizado es bi-lineal (Demerval Soares Moreira, comunicación personal). Es por eso que los resultados que se muestran a continuación están basados en puntos de estaciones y no en campos reticulados. Los pronósticos del SMES utilizados en este estudio van del 1 de Septiembre al 31 de diciembre de 2006.

Uno de los principales desafíos vinculados con la utilización de este conjunto, es que frecuentemente por problemas técnicos de origen o de transmisión de los datos, algunos de los miembros no se encuentra disponible. Si se hubieran eliminado aquellos días en los que alguno de los miembros no estuvo disponible, se hubiera reducido demasiado el conjunto de datos. Es por eso que se optó por utilizar cada día los miembros disponibles (siempre y cuando el número de miembros no fuera menor a 10). La calibración se realizó en estos casos utilizando la verificación de los miembros disponibles. Es por eso que para este conjunto en particular el número de miembros es variable, es decir que el conjunto puede tener menos miembros algunos días en particular. El algoritmo así diseñado, tiene aplicaciones directas a la calibración operativa de los

pronósticos probabilísticos derivados del SMES, lo cual es muy ventajoso.

A diferencia del conjunto SLAF, el conjunto SMES no cuenta con un pronóstico control o no perturbado, dado que a priori todas las configuraciones y modelos utilizados son igualmente válidos. No obstante, existen algunas diferencias en la precisión de los pronósticos de precipitación de los diferentes miembros del conjunto. Con el fin de obtener un miembro control o un pronóstico determinístico que pueda ser comparado con el resto del conjunto, se buscó el miembro que mostrara mejores valores de ETS (Equitable Threat Score) y Bias (ver Apéndice A para una definición de ambos parámetros) en el pronóstico de precipitación sobre la región para todo el período considerado. El miembro que mostró los mejores resultados fue el ETA-CPTEC con una resolución espacial de 20 km en la horizontal. Este miembro será referido como el mejor miembro (MM) de ahora en más. En este caso, la selección del MM fue realizada conociendo de ante mano la verificación de los pronósticos de lluvia sobre el período de estudio. Pero de contarse con pronósticos de lluvia sobre un período mayor, el MM puede determinarse de acuerdo con una verificación en base a datos previos.

### **3.2.2 Metodologías para la obtención y calibración de pronósticos probabilísticos de precipitación.**

El método más simple para obtener un pronóstico probabilístico de lluvia a partir de un conjunto consiste en, dado un umbral de precipitación para el cual se quiere pronosticar la probabilidad de ocurrencia, se cuenta la cantidad de miembros que pronostican precipitaciones por encima de dicho umbral y se divide dicho número por el total de miembros. La principal desventaja de este método es que tiene baja fiabilidad ligada a la sobreestimación de la probabilidad (sobre todo para valores altos de probabilidad) como se verá en las secciones siguientes. Estos apartamientos del comportamiento esperado tienen su origen en los errores sistemáticos del modelo y en las deficiencias del ensamble para pronosticar adecuadamente la incertidumbre asociada a las diferentes situaciones. Para corregir parcialmente algunos aspectos que producen los errores sistemáticos en el pronóstico probabilístico, existen diferentes métodos de calibración de las probabilidades obtenidas a partir del ensamble. Entre los más utilizados se encuentran los desarrollados por Hamill y Colucci (1998, HC98 de aquí en adelante),

Raftery et al. (2005), Stensrud y Yussouf (2007), entre otros. En el presente trabajo se utiliza el método de calibración basado en el histograma de rangos desarrollado por Hamill y Colucci (HC98).

El histograma de rangos o diagrama de Talagrand (Talagrand et al. 1997, Hamill 2001), se construye de la siguiente manera: primero, para cada punto o estación con pronósticos, se ordenan de menor a mayor los valores de precipitación pronosticados por los diferentes miembros del conjunto. Luego se toma la observación correspondiente a dicho punto y se identifica qué lugar ocupa entre los diferentes pronósticos del conjunto, es decir, cuántos miembros del conjunto pronosticaron precipitaciones inferiores y cuántos pronosticaron precipitaciones por encima de lo observado. Si ningún miembro pronostica precipitaciones por debajo de lo observado diremos que la observación cae en el intervalo 1 (rango 1), si un solo miembro pronostica precipitaciones por debajo a lo observado el rango será 2 y así siguiendo. El número total de rangos es igual al número de miembros en el conjunto + 1. Al repetir este procedimiento sobre varios puntos (estaciones) y días, se puede construir un histograma en donde se muestra la frecuencia relativa con la que la observación ocupa cada uno de los rangos definidos anteriormente quedando conformado así el histograma de rangos. Si el conjunto fuera perfecto en el sentido de que pudiera representar adecuadamente la incertidumbre asociada con los pronósticos, el histograma de rangos debería ser chato, es decir que la probabilidad de la observación de situarse en cualquiera de los rangos debiera ser la misma. En este supuesto caso, evidentemente, no sería necesario aplicar una calibración. Existen algunos casos particulares, como por ejemplo el caso en que los pronósticos de todos los miembros del conjunto son 0 y la verificación también es 0. En este caso es imposible asignar un intervalo a la verificación y por ende el intervalo se asigna en forma aleatoria como se describe en HC98. Ejemplos de histogramas de rangos calculados para la región norte y sur del dominio en base a los pronósticos del SMES se pueden ver en la Figura 3.3, en donde además se construyeron histogramas de rangos para subconjuntos del ensamble con diferentes valores de dispersión. Por ejemplo para el caso en el que la dispersión es baja (Figuras 3.3 a y e) se puede apreciar claramente que la distribución dista bastante de ser uniforme. Los histogramas de rango también nos dan información sobre los errores sistemáticos. Por ejemplo si la verificación frecuentemente ocupa rangos bajos indica que por lo general la mayoría de los miembros del conjunto suele pronosticar valores de precipitación más altos

que los observados y por ende indica un bias positivo en la precipitación (ver por ejemplo Figura 3.3 d). Para una discusión profunda sobre la interpretación de los histogramas de rangos, ver Hamill (2001).

A continuación se describe el método de calibración utilizando la información provista en el histograma de rangos siguiendo a HC98. Supongamos que se tiene un vector  $X$  con  $N$  componentes ( $X_i$ ) que representan los  $N$  pronósticos de un ensamble de  $N$  miembros para una localidad o punto de retícula determinados.  $V$  representa la observación correspondiente a dicho pronóstico que no se conoce al momento de la calibración. Por otra parte tenemos el histograma de rangos construido a partir de un conjunto previo de verificaciones y pronósticos ( $R$ ) que tiene  $N+1$  componentes.

Supongamos que queremos conocer la probabilidad de que la precipitación  $V$  esté por debajo de un determinado umbral  $U$ . Previamente notamos la siguiente relación:

$$P(V < X_i) = \sum_{j=1}^i R_j \quad (3.1)$$

La Ecuación 3.1 nos dice que la probabilidad de que la verificación esté por debajo de la precipitación pronosticada por el miembro  $i$ -ésimo del conjunto es la suma de las frecuencias relativas del histograma de rangos ( $R_j$ ) desde el rango 1 hasta el  $i$ -ésimo (en este punto debemos recordar que  $X$  está ordenado de menor a mayor).

Una suposición extra que se introduce es que la probabilidad de que  $V$  sea menor que un umbral  $U$  entre  $X_i$  y  $X_{i+1}$  ó entre 0 y  $X_1$  es lineal. (por ejemplo si 10 y 20 son 2 pronósticos consecutivos del ensamble y la probabilidad de que  $V$  sea menor que 10 es 0.5 y que sea menor que 20 es 1 entonces la probabilidad de que  $V$  sea menor que 15 es 0.75).

De esa manera se llega a que para todo umbral  $U$ , se puede calcular la probabilidad de ocurrencia en función de  $X$  y  $R$  de la siguiente manera:

$$P(V < U) = \left(\frac{U}{X_1}\right)R_1 \quad \text{si } 0 < U < X_1 \quad (3.2)$$

$$P(V < U) = P(V < X_i) + \left(\frac{U - X_i}{X_{i+1} - X_i}\right)R_{i+1} \quad \text{si } X_i < U < X_{i+1} \quad (3.3)$$

El único caso que no está contemplado en las ecuaciones 3.1-3.3 es cuando se quiere evaluar la probabilidad para un umbral que está por encima del máximo de  $X$  ( $X_n$ ). Como destacan HC98 no se puede aplicar un razonamiento de tipo lineal en este caso porque eso implicaría asumir que existe un valor de precipitación para el cual la probabilidad de ocurrencia es 0 %. En este caso, en lugar de recurrir a una proporción lineal se recurre a la proporción mostrada en la Ecuación 3.4; cuánto aumenta la probabilidad es proporcional a cuánto más probable es un valor menor a  $U$  con respecto a un valor menor a  $X_n$ . Esta proporcionalidad cumple con los requisitos, ya que cuando  $U$  es igual a  $X_n$  la probabilidad de que  $V$  esté entre  $U$  y  $X_n$  resulta 0 y cuando  $U$  tiende a infinito la probabilidad resulta igual a  $R_{N+1}$ .

$$P(X_n \leq V < U) = \frac{P(V > U) - P(V > X_n)}{1 - P(V > X_n)} R_{N+1} \quad (3.4)$$

Y análogamente al caso 3.3.

$$P(V < U) = P(V < X_n) + P(X_n \leq V < U) \quad \text{Si } V > X_n \quad (3.5)$$

La ecuación 3.5 completa las ecuaciones que nos permiten calcular la probabilidad calibrada en base al diagrama de Talagrand, un ejemplo numérico de la implementación de este algoritmo se puede encontrar en el apéndice de HC98. Sin embargo todavía persiste un problema a resolver y es la forma en la que se calcula la probabilidad  $P(V > Y)$ , donde  $Y$  es un valor dado de precipitación. HC98 proponen un ajuste de una distribución Gumbel para la estimación de dichas probabilidades. En la presente implementación del algoritmo, las probabilidades de que la verificación esté por encima de un determinado umbral, independientemente de otros condicionamientos, se estima directamente a partir de los datos observados (es decir que no se utilizan ajustes paramétricos para el cálculo de dichas probabilidades sino que se utiliza la frecuencia empírica de ocurrencia del evento).

Una de las características fundamentales de esta calibración es que depende de la forma del histograma de rangos. Por lo tanto, es natural pensar que todos aquellos aspectos que produzcan un cambio en la forma del histograma producirán un cambio en la

calibración. Dicho de otra forma, es necesario identificar los aspectos que influyen sobre el histograma de rangos y realizar una calibración selectiva basada en dichos aspectos. El efecto más estudiado es la dependencia del histograma de rangos con la dispersión del conjunto. Para ver esto se calcula en forma independiente el histograma de rangos para puntos de retícula en los que la dispersión del conjunto es alta, media y baja por separado. La Figura 3.3 muestra un ejemplo de como son los resultados para cada caso. Se puede observar que existe una fuerte dependencia de la forma del histograma con la dispersión. Es por eso que al momento de realizar la calibración, los puntos en donde la dispersión del conjunto es alta deben calibrarse con un histograma distinto al utilizado en los casos en los que la dispersión es baja, porque, si se utilizara el mismo histograma para todos los casos, se estaría introduciendo un error sistemático en el cálculo de las probabilidades. Este aspecto de la calibración está discutido en HC98 quienes introducen una distinción en 3 rangos de dispersión y utilizan un histograma para cada uno de dichos rangos y en Eckel y Walters (1998) en donde se incrementa el número de categorías de dispersión utilizadas. También existe una dependencia del histograma de rangos con el régimen de precipitación y con los errores sistemáticos del modelo de acuerdo con lo discutido en párrafos anteriores. El efecto de los errores del modelo se vuelve más difícil de tratar en el caso del SMES ya que cada uno de los miembros tiene diferentes errores sistemáticos. En este caso es probable que una corrección previa de los diferentes miembros utilizando información sobre sus errores sistemáticos antes de realizar la calibración pueda mejorar los resultados obtenidos (Silva Días, comunicación personal, Hamill y Whitaker 2006), por otro lado existen técnicas de calibración que se basan en el comportamiento de cada miembro por separado y que podrían ser más adecuadas para el tratamiento de este tipo de ensambles (Raftery et al. 2005).

En este capítulo se considera la dependencia del histograma de rangos con la dispersión del conjunto (tomada como la desviación estándar de los pronósticos de los miembros del conjunto) y se definen 3 categorías (dispersión baja, media y alta) cuyos límites se fijan a partir de los terciles de la distribución de la dispersión (es decir que la dispersión baja representa el 30 % de los casos de menor dispersión excluyendo los casos en los que la dispersión es nula). En este caso en particular como solo se utilizan los datos del GTS, no se puede subdividir el rango de dispersiones en más categorías porque en cada categoría el número de datos sería insuficiente para construir un histograma de

rangos estable. También se tiene en cuenta la dependencia del histograma de rangos con el régimen de precipitación ya que se subdivide el dominio en dos regiones, una posicionada al norte de los 20°S y otra al sur de dicha latitud (ver Figura 3.1). La justificación para esta subdivisión se verá más claramente al analizar los resultados de las verificaciones de los pronósticos de lluvia sobre ambas, pero inicialmente sabemos que la región norte está caracterizada por un régimen más tropical en donde la convección ocurre en forma generalizada y con un marcado ciclo diurno. Mientras que en la región sur, si bien durante el verano los fenómenos convectivos explican la mayor parte de la precipitación, los eventos están más asociados con forzantes de escala sinóptica. Las diferencias entre la bondad de los pronósticos para regiones tropicales y de latitudes medias fue documentada por (Ebert et al. 2003) y varios estudios previos muestran que el comportamiento de los errores del pronóstico son muy diferentes en las regiones tropicales respecto de lo que sucede en las latitudes medias (Kalnay 2003). Esto se pudo verificar para el ensamble SLAF al realizar un análisis del comportamiento de los errores de la media del ensamble y de la dispersión del mismo para varias variables (Ruiz et al. 2006).

Para la implementación de las calibraciones y la obtención de otros pronósticos probabilísticos de precipitación, se utilizan los datos de la red GTS cuya distribución puede encontrarse en la Figura 3.2 y cuyas características generales se describieron en el Capítulo 2.

Por otra parte existen diferentes estrategias que pueden adoptarse en cuanto a la selección del período de datos utilizado para calibrar los pronósticos probabilísticos. Las estrategias evaluadas en el presente trabajo se describen a continuación.

### **3.2.3 Calibración estática y dinámica:**

La diferencia entre una calibración estática y una dinámica radica fundamentalmente en los datos que se van a utilizar para la construcción del histograma de rangos. En el caso de una calibración estática, se utilizan datos provenientes de un período equivalente: por ejemplo de la misma estación pero de años anteriores. De esta manera se construye el histograma con el comportamiento típico o climatológico de la atmósfera para una determinada época del año. En este caso particular, este tipo de

calibración solo puede aplicarse al conjunto SLAF debido a que no se contó con datos del conjunto SMES para el año previo a este experimento. Para el conjunto SLAF el experimento previo es el que va desde el 22 de octubre 2005 al 27 de noviembre de 2005, y se utiliza, entonces, como período de entrenamiento. El largo del período utilizado como entrenamiento en este caso puede resultar escaso, pero está limitado por la cantidad de pronósticos disponibles.

Por otra parte, en la calibración dinámica, el histograma de rangos se calcula utilizando un número fijo de pronósticos y verificaciones correspondientes a los días previos al momento de calibrar el pronóstico. En este caso, el número de pronósticos (pares pronóstico – observación) elegidos para calibrar el conjunto fue de 1000. Este número de pronósticos equivale aproximadamente a 20 días previos cuando se utilizan los datos del GTS como fuente de calibración. Este período es similar al propuesto por Silva Dias et al. (2006) para la calibración de los pronósticos de diferentes variables, incluyendo precipitación y por Stensrud y Yussouf (2007). La selección de la longitud del período de entrenamiento, se basó en la necesidad de contar con un número de estaciones suficiente para obtener un histograma de rangos estable y por otra parte que no requiriera extender el período de entrenamiento a un tiempo mucho mayor que pudiera acortar el período disponible para la verificación de los pronósticos. En este caso, la calibración toma en cuenta el estado de la atmósfera en los días previos, que puede ser muy ventajoso cuando el sistema se encuentra dentro de un mismo régimen y puede resultar adverso cuando hay un cambio abrupto (por ejemplo un paso de período seco a húmedo marcado). El efecto adverso se minimiza en la medida que se acorta el período de entrenamiento. Para poder utilizar un intervalo de tiempo más corto en el entrenamiento es necesario contar con un gran número de datos para poder construir los histogramas de rango requeridos por el método propuesto por HC98 con los pronósticos de los días previos. Este requisito se hace aún más complicado cuando la región donde se realiza la calibración está caracterizada por una red de precipitación que en algunos casos resulta insuficiente. En el capítulo 6 se discuten alternativas para poder incrementar el número de observaciones disponibles utilizando estimaciones de precipitación.

### **3.2.4 Obtención de pronósticos probabilísticos a partir de pronósticos determinísticos**

Como se destaca en HC98, la media del conjunto, tratada como un pronóstico determinístico, tiene un potencial para generar pronósticos probabilísticos con una confiabilidad similar a los pronósticos probabilísticos derivados con la técnica basada en el histograma de rangos. En HC98, los autores encuentran que la frecuencia de ocurrencia de precipitación por encima de un determinado umbral está relacionada con el valor de la precipitación media pronosticada por el ensamble. Utilizando esta relación, obtienen un algoritmo que permite evaluar cuál es la probabilidad de que la precipitación esté por encima de un determinado umbral para distintos valores de la media del conjunto.

Gallus y Seagal (2004), mostraron que también puede utilizarse este algoritmo para obtener pronósticos probabilísticos a partir de un pronóstico determinístico. De acuerdo con estos autores, existe una relación entre la cantidad de precipitación pronosticada por un modelo determinístico y la probabilidad de ocurrencia de precipitación por encima de un umbral determinado.

Para obtener dicha relación, Gallus y Seagal (2004) a diferencia de lo realizado por HC98 utilizan un método no paramétrico. Primero se separa el rango de la variable lluvia en intervalos (la longitud de los intervalos no es uniforme, los mismos se hacen más grandes a medida que aumenta la cantidad de precipitación). Para cada intervalo se buscan todos los días en que la cantidad de precipitación pronosticada esté dentro del intervalo especificado. A cada intervalo se le asocia la verificación (observaciones) que se corresponde con dichos pronósticos. Es decir que para cada intervalo tenemos una colección de observaciones que son las que se produjeron cuando el pronóstico estuvo dentro de dicho intervalo. Una vez hecho esto, en cada intervalo se puede calcular a partir de las observaciones la probabilidad de que la precipitación esté por encima de un umbral dado, por ejemplo qué probabilidad hay de que llueva más de 10 mm dado que el valor pronosticado está entre 50 y 60 mm. El valor de probabilidad obtenido se puede asignar al valor medio del intervalo de lluvia pronosticada y a partir de este punto realizar un ajuste que nos permita obtener una curva continua. Esta curva será la probabilidad de ocurrencia de precipitación por encima del umbral  $U$ , como función de la cantidad de lluvia pronosticada por el modelo. La Figura 3.4, muestra un ejemplo en donde se muestra la

probabilidad de tener lluvia por encima del umbral de 2.5 mm como función de la cantidad de lluvia pronosticada por el MM del SMES y la media del ensamble SMES. La línea verde en esta figura representa la probabilidad como función de la lluvia pronosticada para un pronóstico perfecto; en este caso, si el pronóstico es mayor a 2.5 mm entonces la verificación es siempre mayor a 2.5 mm y la probabilidad asociada debería ser del 100%, mientras que si el pronóstico está por debajo de 2.5 mm entonces la verificación también lo estaría y la probabilidad de ocurrencia de lluvia por encima de 2.5 mm sería 0. Otro caso posible sería un pronóstico en donde la probabilidad de ocurrencia de lluvia por encima del umbral es independiente de la cantidad de lluvia pronosticada e igual a la probabilidad climatológica, en este caso se trataría de un pronóstico que no tendría ninguna utilidad. Los pronósticos reales se encuentran entre estos dos extremos, las probabilidades de ocurrencia son función de la cantidad de lluvia pronosticada, pero no llegan a ser 0% cuando la cantidad de lluvia está por debajo del umbral ni alcanzan el 100% cuando están por encima debido a los errores en el pronóstico de lluvia. La diferencia entre la mínima probabilidad y la máxima probabilidad observada en cada una de las curvas puede ser tomada como una medida de la resolución. Cuanto mayor es la distancia a la curva verde o menor la pendiente de la curva entonces menor es la resolución que tiene el pronóstico si entendemos por resolución la capacidad de discriminar entre la ocurrencia o no ocurrencia del evento (por ejemplo un pronóstico totalmente incapaz de discriminar mostraría una curva totalmente plana en donde la probabilidad de ocurrencia del evento es totalmente independiente del valor de precipitación pronosticado). La forma de las curvas roja y azul en la Figura 3.4, muestra que el poder de discriminar entre la ocurrencia y no ocurrencia de un evento no es perfecta ya sea para la media del ensamble o para el MM.

La Figura 3.5 muestra la relación entre la frecuencia observada de lluvia por encima de diferentes umbrales y la cantidad de lluvia pronosticada por la media y los controles (MM en el caso del SMES) de los ensambles considerados. En las Figuras 3.4 y 3.5 se puede apreciar que la media del conjunto tiene una mayor resolución, es decir que permite discriminar más eficientemente la ocurrencia y no ocurrencia del fenómeno. También se puede apreciar que la resolución disminuye en todos los casos a medida que aumenta el umbral seleccionado. Esto muestra que el pronóstico probabilístico se degrada a medida que el fenómeno a pronosticar se vuelve más extremo o infrecuente, aunque

también hay que tener en cuenta que al aumentar el umbral de precipitación disminuye el número de casos y por ende el tamaño de la muestra para la verificación.

### **3.2.5 Significancia estadística de los resultados a través de la técnica de "bootstrap"**

En la verificación de los pronósticos probabilísticos, se utilizaron medidas tanto de la resolución como de la confiabilidad del pronóstico (ver Apéndice A), la mayoría de estas medidas, tienen el problema de que no se conoce su distribución teórica de probabilidades con lo cual no se pueden aplicar los tests de hipótesis tradicionales. Esto presenta una limitación cuando se quiere evaluar la significancia estadística de las diferencias expresadas a través de estos índices. Es por eso que para tal efecto, se utilizan técnicas que permiten evaluar significancia en casos en los que se desconoce la distribución teórica de un determinado parámetro basados en técnicas de remuestreo, como por ejemplo la técnica de "bootstrap". La aplicación de esta técnica es como sigue: Supongamos que contamos con una muestra de ND datos que nos permiten calcular el índice I. La técnica requiere que se generen NM muestras con ND datos cada una a partir de la muestra original. Para generar cada muestra utilizamos una selección aleatoria de elementos de la muestra original con reposición. El número de muestras NM es generalmente grande (en este caso se utilizaron 1000 muestras al azar). Luego se calcula el índice para cada una de las muestras y se obtiene así una estimación de la función de distribución del parámetro a partir del valor obtenido para cada muestra. A partir de esto se pueden calcular probabilidades de que 2 valores del índice difieran en más de una cierta cantidad.

Esta metodología fue aplicada por primera vez a los índices de verificación de pronóstico por Hamill (1999) y luego incluida por otros autores como Accadia et al. 2003, 2005, Tencer 2006, entre otros.

En este trabajo se utilizó un nivel de confianza del 90 % para obtener los límites de significancia para los valores de los índices calculados. Por otra parte se utilizaron diferentes tamaños de muestras y se observó que el valor de 1000 utilizado, produce resultados estables en cuanto a la posición de los límites de los intervalos de confianza.

### **3.3. Resultados.**

#### **3.3.1 Análisis de la relación entre el error y la dispersión en la precipitación.**

De acuerdo con el análisis de los histogramas de rangos en las secciones precedentes (Figura 3.3), se puede observar que para el caso del SMES, existe una relación entre la forma del histograma y la dispersión del conjunto. En particular, cuando la dispersión del conjunto es pequeña, las mayores frecuencias se observan en los rangos bajos y centrales. Esto está indicando que cuando la dispersión del conjunto es pequeña, el error es pequeño ya que la verificación se encuentra contenida frecuentemente entre los miembros del conjunto que a su vez están próximos entre si (Figura 3.3 a y e). Por otra parte, para valores altos de dispersión, se observa el fenómeno opuesto, la verificación frecuentemente se ubica en los extremos del intervalo de rangos (es decir que frecuentemente el valor de la verificación está por debajo del mínimo valor pronosticado por el ensamble o por encima del máximo valor pronosticado). En este caso, el error en la precipitación frecuentemente es mayor y más aún si tenemos en cuenta que en este caso los miembros del conjunto están más separados unos de otros.

En principio este comportamiento sugiere una relación estrecha entre la dispersión del conjunto y el error en el pronóstico de la precipitación, lo cual sería una característica deseable de todo sistema de pronóstico por conjuntos. La existencia de dicha relación en el presente caso se muestra en las Figuras 3.6 (a y b). No obstante como notaron HC98, esta relación puede no ser un resultado de la calidad del conjunto o una consecuencia de su capacidad para pronosticar la incertidumbre en el pronóstico.

De acuerdo con HC98, existe una relación muy estrecha entre el valor medio de la precipitación pronosticada por el conjunto y el error (como también se da entre la magnitud de la lluvia pronosticada por un pronóstico determinístico). A su vez, existe una relación estrecha entre el valor de la dispersión y la media del conjunto (a mayor valor de la media del conjunto mayor dispersión), esto podría resumirse en que a medida que aumenta el valor de la precipitación pronosticada por los miembros del conjunto, aumenta la discrepancia entre los mismos y por ende la dispersión del conjunto (Figura 3.6 a y b).

Es decir que en el caso de la precipitación, la incertidumbre es tan grande que el error crece al aumentar la magnitud de la precipitación pronosticada casi en la misma

proporción como se puede ver por ejemplo en la Figura 3.7, al comparar la magnitud del error como función del valor de precipitación pronosticado por la media del ensamble. (Las pendientes de las curvas que vinculan el error del pronóstico con la cantidad de lluvia pronosticada son cercanas a uno). Es decir que para evaluar si efectivamente el conjunto es capaz de distinguir entre eventos con mayor error o menor error, primero debemos remover la dependencia del error con el valor medio de la precipitación. Para lograr esto se procedió de la siguiente manera: se dividió el rango de la media de los pronósticos de lluvia del ensamble en intervalos y en cada intervalo se calculó el valor medio de la dispersión y el valor medio del error y sus respectivas desviaciones estándar. Luego en cada uno de estos intervalos, se estandarizaron los valores de error y dispersión con respecto a los valores medios y los desvíos propios de cada intervalo. Luego se construyó un diagrama de dispersión con los valores estandarizados utilizando los valores medios y la desviación de cada intervalo (Figura 3.8). En dicha figura, puede observarse que cuando se remueve la dependencia de dispersión y error con la media del conjunto, no existe relación entre ambas variables. Las correlaciones lineales confirman este resultado dando para el SLAF y el SMES correlaciones cercanas a 0. Estos resultados coinciden con los encontrados por HC98 que no pudo demostrar la existencia de una relación entre la dispersión del conjunto y el error en el pronóstico más allá de la que existe a través de la relación entre precipitación media y error. Es preciso destacar que este comportamiento no se extiende a otras variables con un comportamiento espacial y temporalmente más homogéneo como temperatura, humedad, geopotencial como se discutirá en el Capítulo 5.

### **3.3.2. Diagramas de confiabilidad e índice de acierto de Brier.**

La confiabilidad del pronóstico mide el grado de relación entre la probabilidad pronosticada y la probabilidad observada. En la presente tesis, dicha relación se estudia a través de los diagramas de confiabilidad, que se construyeron para los pronósticos a 24 y 48 horas y para todos los umbrales utilizados en los cálculos de probabilidades.

La Figura 3.9, muestra el diagrama de confiabilidad para el umbral de 2.54 mm para los pronósticos probabilísticos no calibrados y las diferentes calibraciones ensayadas en este trabajo. La línea verde diagonal marca la relación esperable para un conjunto ideal y la línea verde horizontal indica el valor de la probabilidad climatológica de

ocurrencia de precipitación por encima del umbral para cada región. Un pronóstico carente de confiabilidad mostraría una curva horizontal similar a esta línea, indicando que independientemente de la probabilidad pronosticada por el conjunto, la probabilidad de ocurrencia del fenómeno es siempre la misma. Como se puede ver en la figura, en general el pronóstico no calibrado muestra una relación que se aparta bastante de la ideal (con excepción del SMES sobre la región sur que se muestra en la Figura 3.9). Esto indica que el pronóstico probabilístico calculado según el método de los votos democráticos no es confiable. La introducción de la calibración por el método del histograma de rangos (curvas azules) mejora en todos los casos la confiabilidad de los pronósticos. Como se puede observar de estos diagramas, no existe una diferencia significativa entre la calibración estática y la dinámica en cuanto a la confiabilidad del pronóstico.

El método para generar un pronóstico probabilístico a partir de un pronóstico determinístico se aplicó a la media del conjunto, al pronóstico control del SLAF y al MM del SMES. Estos métodos también mejoran significativamente la confiabilidad. HC98 habían ya mostrado que la media del conjunto tenía un importante potencial para generar un pronóstico probabilístico confiable. Por otra parte, es importante destacar que el pronóstico control (que no cuenta con ninguna información de los miembros del conjunto) también muestra una cierta habilidad para generar un pronóstico probabilístico confiable. Estos resultados están de acuerdo con los de Gallus y Seagal (2004) que exploraron esta potencialidad. El algoritmo propuesto por Gallus y Seagal (2004), utiliza información de cómo se distribuyen los errores del pronóstico a partir de eventos pasados, mientras que los pronósticos derivados de los ensambles agregan la información provista por los diferentes miembros.

Por otra parte, la calibración produce cambios importantes en la distribución de la frecuencia con la que se pronostican los diferentes valores de probabilidad. Los pronósticos no calibrados en la región sur por ejemplo muestran dos máximos bien marcados en la frecuencia, uno en valores muy bajos de probabilidad y otro en valores muy altos de probabilidad. No obstante, la curva del diagrama de confiabilidad muestra que en este último caso se está sobreestimando el valor de la probabilidad. La calibración reduce el número de veces que se pronostica un valor de probabilidad alto.

Cuando se compara la región norte y la región sur en términos de las frecuencias con las que se pronostica los diferentes valores de probabilidad, se observa que en la

región norte existe una tendencia a pronosticar frecuentemente valores de probabilidad alrededor del 50%. Esto reduce la capacidad del pronóstico de distinguir entre ocurrencia o no ocurrencia de los eventos, ya que si bien cuando se pronostican probabilidades más altas o más bajas estas son confiables, la mayor parte del tiempo, las probabilidades pronosticadas no muestran esos valores con lo cual parecería indicar que el pronóstico calibrado tiene menor resolución, no obstante como se verá más adelante esta aparente merma en la resolución es compensada por el aumento de la pendiente de la curva en los diagramas de confiabilidad.

En general en la Figura 3.9, se puede observar que la confiabilidad del SMES es superior a la del SLAF tanto antes como después de la calibración. Esto se observa particularmente sobre la región norte del dominio (Figuras 3.9 a y b). En dichas figuras puede observarse que el SMES muestra mayor resolución ya que la pendiente de la curva indica que es mayor su capacidad de discriminación. La razón para que ocurra esto no es del todo clara, pero está de acuerdo con la experiencia que el uso de multimodelos es superior al de un modelo con distintas condiciones iniciales (Hou et al 2001, Krishnamurti et al 1999). Puede ser que la inclusión de perturbaciones en la física incluidas en el SMES sea más eficiente sobre la región norte donde los procesos convectivos no resueltos juegan un rol fundamental en la dinámica de la atmósfera. También es posible que la configuración del modelo WRF-ARW utilizada en el presente experimento podría no ser la adecuada para la región norte en cuyo caso esto estaría afectando negativamente el pronóstico probabilístico derivado del conjunto SLAF.

La Figura 3.10, muestra diagramas de confianza pero para el umbral de 25.4 mm. La diferencia fundamental que puede observarse con respecto a los pronósticos de umbrales más bajos, es que hay una drástica reducción en la resolución del pronóstico. La máxima probabilidad que se puede derivar del ensamble con respecto a la ocurrencia de precipitaciones por encima de 25.4 es inferior a 50% para la región sur, tanto para el SLAF como para el SMES. Esto podría interpretarse como un incremento de la incertidumbre cuanto mas extremo es el fenómeno que se quiere pronosticar e implica una limitación a la aplicación de los pronósticos probabilísticos a la ocurrencia de eventos extremos. Por otra parte en el presente estudio existe una limitación adicional que es la falta de un volumen grande de datos, que se hace más evidente cuanto menos frecuente es la ocurrencia del fenómeno que se quiere pronosticar/verificar. La degradación de la

resolución conforme se pronostican umbrales cada vez mayores se observa siempre en la implementación de pronósticos probabilísticos (HC98, Eckels y Walters 1998, Buizza et al. 1999).

Es importante destacar además que si bien las Figuras 3.9 y 3.10 muestran los resultados para el pronóstico a 24 horas, los pronósticos a 48 horas no presentan un comportamiento distinto respecto a las conclusiones que se derivaron de los pronósticos a 24 horas, si bien se observa que conforme avanza el plazo de pronóstico existe una degradación en la resolución de los mismos. Por ejemplo para el SMES y para el umbral de 2.54 mm, la máxima probabilidad calibrada se reduce un 10 % sobre ambas regiones (no mostrado).

Para resumir la información contenida en los diagramas de confiabilidad para distintos umbrales se utiliza el índice de acierto de Brier (de aquí en más BSS -ver Apéndice A ecuación A.2-) que constituye una medida de la confiabilidad y la resolución del pronóstico probabilístico comparado contra la probabilidad climatológica de ocurrencia del evento. La Figura 3.11 muestra los valores de BSS obtenidos como función del umbral de precipitación considerado para el pronóstico sin calibrar y para los diferentes pronósticos calibrados analizados en el presente capítulo. Como se puede ver en dicha figura, la calibración mejora la relación entre la probabilidad observada y la pronosticada para los dos sistemas en la región norte (Figura 3.11 a y b) y para el SLAF sobre la región sur. Para el caso del SLAF, las mejoras son significativas sobre ambas regiones para casi todos los umbrales considerados, no obstante para la región norte, aún el pronóstico calibrado por diferentes metodologías exhibe valores de BSS por debajo de 0, indicando que no pudo agregar información más allá del conocimiento de la frecuencia climatológica del evento. En todos los casos se observa una disminución del BSS a medida que aumenta el umbral que queremos pronosticar, de acuerdo con lo observado previamente en los diagramas de confiabilidad. En la región sur por ejemplo (Figuras 3.11 c y d), los pronósticos probabilísticos calibrados a partir del SLAF, logran superar a la climatología hasta el umbral de 25.4 mm, mientras que para el SMES, el umbral hasta el cual el pronóstico logra superar a la climatología se extiende hasta los 38.2 mm.

El pronóstico no calibrado para ambos conjuntos muestra valores de BSS superiores a 0 sobre la región sur, no obstante para la región norte, sobre todo en el caso del SLAF, el pronóstico no calibrado muestra un desempeño inferior incluso al de la climatología.

Sobre la región sur y para los pronósticos derivados de la media del conjunto y del pronóstico control se observa una mejora significativa en el BSS, únicamente para los primeros umbrales. Para los umbrales más altos no se observa una mejora en el BSS con respecto al pronóstico no calibrado.

En el caso del SMES en particular sobre la región sur, se observa poca mejora con la calibración, las diferencias solo son significativas para el primer umbral. En este caso, el pronóstico probabilístico derivado del MM muestra valores de BSS tan buenos como los demás pronósticos probabilísticos. Esto podría estar indicando que existe poco incremento del valor de la información por parte de este conjunto sobre la región sur, ya que en términos del BSS pueden esperarse los mismos resultados utilizando un único pronóstico control. Es posible que la inclusión en este conjunto de modelos cuya exactitud sea inferior al resto de los miembros del conjunto pueda estar degradando el pronóstico probabilístico sobre esta región. Esto sugiere que la aplicación de una técnica como el SLAF utilizando un modelo como el "mejor modelo" (MM) podría dar mejores resultados sobre la región.

Para intentar probar esta hipótesis se utilizó la técnica LAF (Lagged Averaged Forecast, Dalcher et al. 1988) sobre el modelo MM utilizando pronósticos de hasta 120 horas, lo cual resultó en peores resultados que los obtenidos con el pronóstico control (no se muestra la figura). No obstante esto puede deberse a que el LAF no es una técnica adecuada para la generación de pronósticos probabilísticos de precipitación.

La Figura 3.11 confirma los resultados obtenidos a partir de los diagramas de confiabilidad, en el sentido de que no existen grandes diferencias entre la calibración estática y la dinámica siendo la primera la que muestra mejores resultados. En el presente caso, la ventaja de la calibración estática sobre la dinámica puede deberse a que esta última usa un número limitado de observaciones para el proceso de calibración. Para explorar este punto, se realizaron experimentos en donde la calibración dinámica utilizó el doble de observaciones (2000 datos). Estos experimentos mostraron que la diferencia entre ambas estrategias de calibración continúan siendo pequeñas (es decir que no se puede establecer una ventaja en término de la confiabilidad en el presente caso). También se realizaron pruebas aplicando las estrategias de calibración dinámicas y estáticas al pronóstico probabilístico derivado de la media del conjunto y del control del conjunto SLAF. En este caso las diferencias encontradas entre ambas estrategias también fueron

pequeñas (no se muestra la figura).

Para indagar un poco más sobre el origen de las diferencias entre los diversos esquemas de pronósticos probabilísticos estudiados, se recurre a la descomposición del índice de Brier (BR de aquí en más) como se describe en el Apéndice A (Ecuación A.3), en una componente asociada a la confiabilidad y otra componente asociada a la resolución. En este caso, el sentido de las variaciones se interpreta de la misma manera que con el BSS, es decir que cuanto mayor es el valor, mejor es el desempeño del pronóstico probabilístico comparado contra la climatología.

La Figura 3.12, muestra la componente de confiabilidad. De acuerdo con la descomposición propuesta un valor de 0, indica una confiabilidad óptima. Como se puede apreciar en la figura, la calibración tiene un efecto importante sobre la confiabilidad en todos los casos con excepción del SMES sobre la región sur en donde, como se había discutido previamente, la calibración tuvo poco impacto. Esto se debe principalmente a que el SMES mostraba un pronóstico más confiable que el del SLAF antes de la calibración, por lo que era de esperarse que el impacto de la misma fuera menor. Por otra parte, los pronósticos derivados del control del SLAF y del MM, como así también los derivados de la media del conjunto muestran una mejora importante de esta componente con respecto al pronóstico no calibrado. Es decir que estos métodos son tan o más eficientes en incrementar la relación entre probabilidad observada y pronosticada como la implementación de la calibración por el histograma de rangos. Al igual que lo observado en los diagramas de confiabilidad, el SMES no calibrado muestra mejores valores de confiabilidad que el SLAF no calibrado, asimismo la confiabilidad es menor en la región norte con respecto a la sur para ambos conjuntos. Es decir que podemos indicar que el aumento del BR sobre la región norte no está únicamente ligado a un aumento de la incertidumbre (debido a que en la región norte la probabilidad de ocurrencia de lluvia más cercana a 0.5 para los primeros umbrales y por ende la incertidumbre es mayor que en la región sur) sino que además existe un aporte extra que tiene que ver directamente con la calidad del pronóstico probabilístico sobre la región.

La Figura 3.13 muestra la componente de resolución del BSS, donde se puede apreciar que la misma disminuye (empeora) a medida que aumenta el umbral indicando que para umbrales más grandes el pronóstico tiene menor capacidad para discriminar entre la ocurrencia y no ocurrencia de los eventos. Algo importante de destacar, es que la

resolución no se ve prácticamente afectada por el proceso de calibración debido a que los algoritmos de calibración ensayados están orientados a corregir los errores sistemáticos que afectan a la confiabilidad del pronóstico. Con respecto a los pronósticos probabilísticos derivados del control en el caso del SLAF, se puede apreciar, que existe una diferencia importante en la resolución con respecto a los pronósticos derivados del ensamble (Figura 3.13 c). En esta figura aparecen diferencias notorias entre ambos ensambles. Por un lado en el caso del SLAF, la resolución del pronóstico derivado del ensamble no logra superar a la del pronóstico control sobre la región norte, mientras que en la región sur, existe una clara superioridad del ensamble con respecto al pronóstico control. Esto sugiere que la técnica aplicada es más efectiva sobre la porción sur del dominio, probablemente debido a que esta técnica se basa en buscar perturbaciones óptimas que respondan a las inestabilidades del flujo y responda mejor en regiones donde la baroclinicidad sea el mecanismo de inestabilidad dominante. Por otra parte como se mencionó anteriormente, es posible que la configuración utilizada del modelo WRF no sea adecuada para la región tropical. El SMES, muestra un comportamiento diferente, ya que en la región norte, muestra que el ensamble logra superar en resolución al MM, mientras que en la región sur el MM y el ensamble muestran valores de resolución similares. Esto sugiere que la técnica del super ensamble podría ser más adecuada en la región tropical a diferencia de la técnica del SLAF que mostraba mejores resultados sobre la región sur.

### **3.3.3 Precisión y resolución:**

Una medida de la precisión de un pronóstico probabilístico es el diagrama ROC (Relative Operating Characteristics). Como se discute en el Apéndice A, este diagrama permite graficar como es el coeficiente de aciertos como función del coeficiente de falsas alarmas para un umbral de precipitación dado. El diagrama ROC ofrece un punto de vista diferente y complementario al que se utiliza al analizar el diagrama de confiabilidad. En este caso, analizamos los pronósticos que surgen de utilizar diferentes umbrales de probabilidad en la toma de decisión (es decir diferentes valores de probabilidad para decidir si se pronostica o no la ocurrencia del evento) para establecer cuál es la probabilidad de que el fenómeno haya sido pronosticado si se presenta (tasa de acierto) y cual es la probabilidad de que haya sido pronosticado si no se presenta (falsa alarma). Es

decir que analizamos las probabilidades de que el fenómeno haya sido pronosticado en función de la ocurrencia o no del mismo. Si el pronóstico carece de exactitud la probabilidad de que el pronóstico sea positivo será la misma independientemente de si el fenómeno ocurre o no (esto indica también una falta de resolución).

La Figura 3.14, muestra el diagrama ROC para el umbral de 2.5 mm. Los umbrales de probabilidad utilizados comienzan en 10% y terminan en 90% a intervalos de 10 %. Cuanto más cerca está la curva a la esquina superior izquierda mayor es la exactitud del pronóstico. Como puede observarse en este gráfico, el ensamble muestra una exactitud mayor que el control del SLAF para la región sur principalmente, de acuerdo con los resultados analizados previamente. En cuanto al SMES, el ensamble muestra mejores resultados sobre la región norte con respecto al MM, pero en la región sur la exactitud de ambos pronósticos es comparable. La región norte, muestra curvas que se acercan más a la diagonal que en el caso de la región sur; esto indica que la exactitud de los pronósticos es menor sobre dicha región. Como en otras medidas que no están directamente relacionadas con la confiabilidad del pronóstico, las curvas del diagrama ROC no se ven significativamente afectadas por la calibración ya que de hecho el diagrama ROC no se ve directamente afectado por los bias condicionales en la probabilidad.

En los diagramas ROC, un pronóstico determinístico se representa con un punto, mientras que los pronósticos probabilísticos constituyen una curva o familia de valores. Esto está vinculado con una de las ventajas fundamentales del pronóstico probabilístico que es la de poder "ajustar" los valores de falsa alarma y acierto a las necesidades de diferentes usuarios, posibilidad que no existe cuando el pronóstico se presenta en forma de pronóstico determinístico.

La información contenida en los diagramas ROC para diferentes umbrales, se puede resumir calculando el área de la región limitada por la curva del diagrama ROC y la diagonal y computando el índice ROCS (ecuación A.8). La Figura 3.15, muestra los resultados únicamente para el pronóstico no calibrado, para el pronóstico control del SLAF y para el MM del SMES. Esta Figura permite confirmar las conclusiones antes extraídas y generalizarlas a los diferentes umbrales. El conjunto SLAF muestra una menor degradación de la exactitud del pronóstico al incrementarse el valor umbral a la par que se observa un mayor aporte de información del conjunto con respecto al pronóstico control, mientras que en la región sur para algunos umbrales el MM del SMES muestra una

exactitud mayor que el conjunto. Este resultado en particular, es distinto de resultados obtenidos en trabajos previos como por ejemplo Zhu et al. (2002). En este trabajo en particular, se compararon las áreas bajo la curva del diagrama ROC de pronósticos determinísticos y probabilísticos. El cálculo del área del pronóstico determinístico se realizó tomando el valor de acierto y falsa alarma correspondiente a dicho pronóstico y uniéndolo con las esquinas inferior izquierda y superior derecha mediante una interpolación lineal (ver su Figura 6). En este caso, no se está teniendo en cuenta el potencial del pronóstico determinístico para generar un pronóstico probabilístico. Si tenemos en cuenta ese potencial (al comparar por ejemplo con el pronóstico probabilístico derivado del pronóstico control o del MM), el área bajo la curva aumenta. Es por eso que a la hora de comparar las ventajas de los pronósticos derivados de un conjunto con las de un pronóstico determinístico es aconsejable tener en cuenta que se pueden obtener pronósticos probabilísticos a partir del pronóstico determinístico y que en algunos casos pueden mostrar valores de área tan buenos o mejores que el conjunto al cual pertenecen.

Otra forma de evaluar la exactitud del pronóstico es a través del ETS y el bias (ver Apéndice A). Estos índices al igual que los aciertos y las falsas alarmas fueron extensamente utilizados en la evaluación de pronósticos determinísticos. El ETS al igual que los aciertos y las falsas alarmas, puede ser calculado para un pronóstico probabilístico utilizando diferentes valores de probabilidad como umbral para la toma de la decisión. De esta manera, al igual que en el caso del diagrama ROC, podemos obtener varios valores de ETS cada uno correspondiente a un valor de probabilidad dado. De manera análoga se puede proceder con el bias. El bias en particular, tenderá a ser mayor cuanto menor sea el umbral de probabilidad seleccionado (es decir que si elegimos un umbral de probabilidad bajo para pronosticar la ocurrencia del evento es de esperar que aumente la cantidad de veces que el evento es pronosticado y por lo tanto aumente el bias). Hamill (1999) mostró que los pronósticos que tienen un bias muy diferente entre sí no pueden ser comparados directamente ya que el ETS de un pronóstico varía cuando sólo se le suma una constante al campo pronosticado. Para solucionar este problema es que se utiliza el diagrama ETS-bias que es similar al ROC, sólo que muestra el ETS en función del bias. De esta forma, se puede comparar la exactitud de los pronósticos probabilísticos considerados a igual valor de bias.

Es importante tener en cuenta que este problema es particularmente notorio

cuando se evalúa la exactitud de pronósticos determinísticos obtenidos a partir de pronósticos probabilísticos utilizando diferentes umbrales de probabilidad (como en el caso del ROC) ya que por ejemplo los umbrales de probabilidad más bajos producirán pronósticos con valores relativamente altos de bias y viceversa.

La Figura 3.16 (a) y (b), muestra los diagramas ETS-bias sólo para la región sur, el pronóstico a 24 horas y el umbral de 2.5 mm. En este caso los pronósticos (puntos de la curva) con menor bias (lado izquierdo) corresponden a los umbrales mayores de probabilidad. Los máximos de ETS se observan en general para bias cercanos a 1 pero ligeramente superiores. En este diagrama se puede ver claramente como la exactitud del pronóstico probabilístico derivado a partir de la media del ensamble es superior en exactitud al pronóstico control para el caso del conjunto SLAF. En ocasiones el ETS de la media del conjunto, calculado como si fuera un pronóstico determinístico arroja resultados similares o incluso inferiores a los de un pronóstico control. Esto sin embargo no contradice lo que muestra el diagrama ETS-bias, dado que por lo general la media del conjunto para umbrales bajos posee un bias muy alto (mientras que lo contrario ocurre con los valores de precipitación más altos). Este bias elevado degrada el ETS asociado a la media del conjunto cuando se lo calcula de esta manera. Sin embargo el diagrama ETS-bias muestra que la media tiene un potencial de exactitud mucho mayor que el pronóstico control.

Con respecto a los experimentos analizados en la presente sección, es importante mencionar que el período de calibración / verificación utilizado para comparar ambos ensambles no coincide exactamente (ver Tabla 3.II), dado que existen distintos días faltantes en uno y otro conjunto de datos. Dado que la muestra sobre la cual se realizó la verificación no es muy grande, las diferencias en los días considerados para uno u otro ensamble podrían producir diferencias significativas en los índices previamente calculados. Es por eso que se realizó un experimento en donde se utilizaron únicamente los días en los que ambos conjuntos estaban disponibles. Los resultados obtenidos con dicho experimento condujeron a exactamente las mismas conclusiones que en el presente caso.

### 3.4 Conclusiones:

Los experimentos realizados en el presente capítulo, muestran que existe un gran potencial de mejora en la previsión de la precipitación a corto plazo (24-48 horas) sobre la región mediante la implementación de pronósticos por ensamble a escala regional.

Esta ventaja se basa principalmente en dos factores, por un lado los pronósticos de la media de los ensambles utilizados como pronósticos determinísticos mostraron niveles de exactitud mayores que los pronósticos determinísticos obtenidos con una única simulación del modelo. Esta ventaja fue más evidente para el modelo WRF sobre la región sur del dominio. Sobre la región norte, la mejora en la exactitud puede encontrarse en ambos sistemas de pronósticos por ensambles, pero las diferencias no son tan importantes como para el SLAF sobre la región sur. Esto indica que la implementación del pronóstico por ensambles en este aspecto podría tener ventajas más limitadas sobre la región norte del dominio. Como se discutió previamente, esta limitación tiene que ver con una menor exactitud del pronóstico de lluvia sobre la región tropical donde los procesos convectivos son más importantes y las precipitaciones no están tan ligadas a los forzantes de escala sinóptica.

Pronóstico	Calibración	Entrenamiento (*)	Verificación
Ensamble SLAF	Estática y dinámica	22 de octubre al 27 de noviembre de 2005	28 de octubre al 31 de diciembre de 2006
Media ensamble SLAF	Estática y dinámica	22 de octubre al 27 de noviembre de 2005	28 de octubre al 31 de diciembre de 2006
Control SLAF	Estática y dinámica	22 de octubre al 27 de noviembre de 2005	28 de octubre al 31 de diciembre de 2006
Ensamble SMES	Dinámica	-	18 de octubre al 31 de diciembre de 2006
Media ensamble SMES	Dinámica	-	18 de octubre al 31 de diciembre de 2006
SMES MM	Dinámica	-	18 de octubre al 31 de diciembre de 2006

**Tabla 3.II: Resumen de los pronósticos utilizados en el presente trabajo consignando tipo de calibración, período de entrenamiento (\*) (sólo para la calibración estática) y período de verificación**

Otro aspecto importante del aporte de los pronósticos derivados de conjuntos sobre la región es la posibilidad de implementar un pronóstico probabilístico confiable. Como se discutió a lo largo del capítulo, un pronóstico probabilístico no es patrimonio exclusivo de los sistemas de pronósticos por ensamble. Los mismos se pueden derivar de un pronóstico determinístico y obtener un grado de confiabilidad relativamente alto. Por otra parte, al igual que como se mostró en trabajos anteriores, el pronóstico probabilístico derivado del ensamble requiere de una calibración basada en una estadística del comportamiento del mismo para poder incrementar su nivel de confianza. Utilizando diferentes estrategias de calibración, se pudo ver que el pronóstico por conjuntos tiene un potencial mayor para generar pronósticos probabilísticos confiables que los pronósticos derivados de una única simulación. Por otra parte también se mostró que la resolución del pronóstico es mayor para el pronóstico por conjuntos. No obstante se podría decir que los pronósticos probabilísticos derivados del pronóstico control o del "mejor modelo" (MM) fueron competitivos. Esto implica que la capacidad de los pronósticos determinísticos para generar pronósticos probabilísticos a través de algoritmos como el utilizado en este estudio, debería ser tenida en cuenta cuando se compara la exactitud de dichos pronósticos contra los pronósticos derivados de un conjunto.

Con respecto específicamente a las diferentes estrategias de calibración examinadas, en todos los casos, la calibración mejoró la calidad de los pronósticos conduciendo principalmente a un aumento de la confiabilidad. Esto se observó tanto para la región norte como para la sur y en ambos sistemas de pronóstico, aunque en el caso del SMES el impacto de la calibración fue menor sobre la región sur. El algoritmo propuesto basado en HC98, no incluye información sobre el desempeño individual de los diferentes modelos que componen el SMES. Este tratamiento es, por tanto, más adecuado a ensambles como el basado en SLAF, en donde los errores del modelo son comunes a todos los miembros. Para el caso de ensambles multi-modelo como el SMES, existen algoritmos que se basan en promedios Bayesianos (Raftery et al. 2005) y que pueden ser aplicados a los pronósticos probabilísticos de precipitación (Mc Lean et al. 2007). La aplicación de este tipo de técnicas a los productos derivados del SMES podrían aportar mejoras significativas con respecto a los resultados obtenidos en el presente trabajo.

Con respecto a las estrategias dinámicas y estáticas en la implementación de la calibración, la primera presenta ventajas en cuanto a que puede adaptarse rápidamente al

régimen de tiempo imperante. También en un marco operativo, es indiferente a los posibles cambios en el modelo ya que requiere un período de entrenamiento relativamente corto. En el presente capítulo, no se encontraron diferencias significativas entre ambas implementaciones de los algoritmos, no obstante la calibración estática mostró casi siempre resultados ligeramente superiores. Esto puede deberse en el presente caso a que los histogramas de rango contaban con mayor número de datos en el caso de la calibración estática. Este punto en particular está vinculado con una de las grandes desventajas que presenta la calibración dinámica sobre nuestra región que está caracterizada por una red de estaciones que es insuficiente.

Con respecto a las diferentes estrategias para generar pronósticos por ensambles analizadas en el presente estudio, ambos sistemas mostraron un comportamiento similar en términos de exactitud y confiabilidad sobre la porción sur del dominio, no obstante sobre la región norte, el comportamiento del SMES fue claramente superior al del SLAF. El SMES incluye cambios en la física de los modelos, y dado que la convección y otros fenómenos de escala no resuelta son muy importantes en la dinámica de la atmósfera en la región tropical, esta estrategia puede tener un mayor potencial para generar soluciones diferentes y lograr un pronóstico probabilístico más confiable que en el caso del SLAF.

Por otra parte, es posible que la configuración utilizada en el modelo ARW-WRF en estos experimentos no sea la más apropiada para la región norte del dominio. Vinculados a esta configuración se observaron importantes bias fríos en la capa límite que podrían estar asociados con una reducción de la energía potencial disponible para la convección y por ende en una degradación del pronóstico de lluvia sobre la región.

Los resultados de este capítulo muestran claramente que la generación de pronósticos por conjuntos presenta un valor agregado comparado con pronósticos de una única realización, aún en pronósticos a corto plazo, como los valorados en esta tesis. Este resultado justifica invertir esfuerzos en esta línea e indagar otras metodologías, con vistas a tener un ensamble operativo sobre la región de Sudamérica.

Para estudiar con mayor profundidad el impacto de las calibraciones, otras estrategias de generación de pronósticos por conjunto y la posibilidad de implementar estrategias de calibración en base a estimaciones de precipitación, se realizaron nuevos experimentos utilizando la estación cálida 2002-2003 que coincide con la realización del experimento de campo SALLJEX. Los experimentos realizados durante dicho período se

Evaluación preliminar ...

describen en los capítulos subsiguientes.

## Capítulo 4:

### El método de los vectores criados

En el presente capítulo se hace una breve revisión del método de los vectores criados para la generación de perturbaciones iniciales que permitan la construcción de un ensamble. Además se mostrarán algunas de sus características a partir de los experimentos aquí realizados.

#### 4.1 La técnica de los vectores criados

Los pronósticos por ensambles buscan representar la incertidumbre asociada a la predicción numérica de la atmósfera que surge como consecuencia de la naturaleza caótica de la misma y de esta manera poder obtener información adicional que sea de utilidad para los usuarios de la información meteorológica. Los pronósticos numéricos poseen dos grandes fuentes de incertidumbre: la primera viene dada por el inexacto conocimiento de las condiciones iniciales y de borde. Y la segunda por la falta de precisión de los modelos numéricos para reproducir diversos aspectos de los fenómenos que se dan en la atmósfera. Para generar ensambles que reproduzcan la incertidumbre en el conocimiento de las condiciones iniciales se han desarrollado diversas técnicas, entre las cuales se destacan los vectores singulares (Molteni et al. 1996) que se utilizan en el Centro Europeo, los vectores criados (Toth y Kalnay 1993, 1997) que se utilizaba en forma operativa en el NCEP, los métodos que se basan en la perturbación de los ciclos de asimilación y los métodos de Montecarlo. Los métodos de Montecarlo son limitados en el sentido de que no contemplan los errores asociados a las inestabilidades del flujo ("errores del día"). En cuanto a las técnicas de los vectores singulares y de los vectores criados, ambas buscan generar perturbaciones que tengan un rápido crecimiento durante el pronóstico, lo cual puede generar una dispersión de los miembros del ensamble que sea similar en magnitud al apartamiento que sufre el pronóstico de la atmósfera real, a la par de que tienen en cuenta las inestabilidades del flujo.

La técnica de los vectores singulares, busca aquellas perturbaciones que tendrán un rápido crecimiento a lo largo de un cierto plazo (por ejemplo 3 días). Estas perturbaciones se obtienen utilizando argumentos de crecimiento lineal y usualmente utilizando una versión simplificada del modelo: es un método complejo y muy costoso computacionalmente. Los vectores criados, por otra parte, constituyen un método que, como se detallará a continuación, se basa en la utilización del modelo no lineal (es decir, el modelo completo) y es muy económico computacionalmente, a la vez que provee resultados similares a los obtenidos con la técnica de los vectores singulares. Esta característica de ser barato computacionalmente fue uno de los factores principales para determinar su elección, debido a que al momento de realizar los experimentos de este trabajo de tesis, no se contaba con recursos computacionales suficientes, ni el software adecuado para implementar un ensamble basado en los vectores singulares.

El objetivo general del método de los vectores criados es generar (criar) perturbaciones que puedan crecer rápidamente durante el pronóstico. De acuerdo con Kalnay (2003), diversos experimentos mostraron que las perturbaciones que contenían información sobre las inestabilidades presentes en el flujo (que variaban día a día) podían crecer más rápidamente que lo que crecían las perturbaciones generadas al azar. Por otra parte, de acuerdo con Toth y Kalnay (1997), el subespacio de perturbaciones que son capaces de crecer a lo largo del plazo de pronóstico es relativamente pequeño comparado con las dimensiones del espacio donde se encuentra la solución de un modelo numérico. Es por eso que al seleccionar perturbaciones de este subespacio, un número menor de perturbaciones puede muestrear más eficientemente las posibilidades futuras de evolución del sistema.

El método de los vectores criados consiste en un ciclo que se puede inicializar tomando una perturbación generada al azar (o como en el presente caso, calculada como la diferencia entre dos análisis cualesquiera). Esta perturbación inicial se suma y se resta al análisis al comienzo del ciclo para obtener un par de análisis perturbados. A continuación se realizan 2 pronósticos utilizando el modelo global partiendo de los análisis perturbados. El modelo se integra por un período de tiempo (en este estudio es de 6 horas), al cabo del cual, se toman los pronósticos del modelo global y se restan obteniéndose así una nueva perturbación cuyo tamaño, es por lo general mayor al de la perturbación inicial, debido a que ha crecido como consecuencia de la presencia de

inestabilidades en el flujo. Es por eso que para controlar la amplitud de la perturbación al cabo de la integración, esta se reduce de forma tal de que su magnitud sea similar a la que tenía la perturbación inicial. Esta reducción de la amplitud se realiza multiplicando a la perturbación por un factor que se define según la Ecuación 4.1. De acuerdo con la ecuación 4.1, la modificación de la amplitud de las perturbaciones se realiza en forma local (es decir que el factor de reducción  $f_a$  depende de la posición y del tiempo), mediante el cálculo de un coeficiente que depende de un parámetro global ( $f_g$ , independiente de la posición), una máscara ( $M_g$ ) que es función de la latitud, la longitud y el tiempo pero que no depende del flujo y la amplitud de la perturbación medida según la energía cinética ( $EC_s$ ) en 500 hPa., suavizada. Una particularidad de los vectores criados es que su evolución es independiente de la forma en la que se mida su amplitud (es decir la norma utilizada para calcular su amplitud), esta característica no es compartida por los vectores singulares, que son dependientes de la norma utilizada para medir las perturbaciones (por ejemplo la perturbación inicial que produce un máximo crecimiento de la temperatura en 850 hPa no es la misma que produce un máximo crecimiento de la energía cinética en 500 hPa.).

$$f_a = \frac{f_g M_g}{EC_s} \quad (4.1)$$

La máscara global utilizada en el cálculo del coeficiente que reduce la amplitud de la perturbación depende de la posición y de la época del año (ver Figura 4.1 a) según lo establecido en Toth y Kalnay (1997). La función de la máscara es representar la distribución espacial y temporal de la incertidumbre en el análisis, independientemente de los errores del día, las regiones con mayor incertidumbre corresponden a regiones donde la red de observación es menos densa y la asimilación de datos es menos eficiente en reducir los errores presentes en el análisis. En estas regiones, el valor de la máscara es mayor lo que, de acuerdo con la ecuación 4.1, permite un mayor crecimiento local de la perturbación. De esta manera, la máscara representa la inhomogeneidad en la distribución de las observaciones, donde estas son más densas, se asume que las perturbaciones en la condición inicial deberían ser menores, ya que esta condición debería estar mejor representada por el análisis. La variación temporal de la máscara obedece a

las variaciones estacionales de la incertidumbre en los análisis, las mismas fueron estudiadas a partir de la comparación entre dos ciclos de análisis independientes (I. Szunyogh, 2002, comunicación personal).

El parámetro global ( $f_g$ ) que interviene en la reducción de amplitud de las perturbaciones constituye un valor fijo que es independiente del tiempo y la posición. Si este factor se reduce, las perturbaciones resultantes serán más pequeñas, como se discutirá más adelante, las inestabilidades que juegan un rol preponderante en la generación de las perturbaciones por el método de los vectores criados, varían de acuerdo a la amplitud de las mismas, es decir que son sensibles al valor de  $f_g$  utilizado.

Finalmente, la energía cinética suavizada en 500 hPa., realiza una reducción de amplitud selectiva. En las regiones donde el crecimiento fue más rápido, la reducción es mayor, mientras que las regiones donde el crecimiento fue más lento, la reducción de la amplitud es menor. La Figura 4.1 b, muestra la energía cinética suavizada en 500 hPa., de uno de los pares perturbados utilizados, correspondiente al día 17 de diciembre. Los aspectos de menor escala han sido eliminados mediante la aplicación de un filtro gaussiano. La Figura 4.1 c, muestra el factor de escala resultante para este día particular, que se usará para reducir la amplitud de la perturbación. Es importante notar que la amplitud de la perturbación siempre se reduce siendo 1 el máximo valor posible para el factor de escala. Finalmente la Figura 4.1 d, muestra la energía cinética de la misma perturbación en 500 hPa., luego de que dicha perturbación fuera sometida a la reducción de la amplitud. En esta figura puede verse cómo los máximos que se observan en la perturbación original al sudoeste de Australia y sobre el sur de Sudamérica, han sido reducidos debido a la amplitud de la perturbación en dichas regiones. La máscara modula este proceso: por ejemplo, puede verse que la mayor reducción de la amplitud en el extremo sur de Sudamérica se da sobre el continente donde el valor de la máscara global es menor. Otro caso similar ocurre entre el Pacífico nororiental y la región oeste de Estados Unidos.

Una vez que se generó la nueva perturbación, esta perturbación se suma y se resta al análisis correspondiente y el proceso se repite nuevamente. Este ciclo así ejecutado da lugar a perturbaciones que van evolucionando con el tiempo. Durante el período de integración del modelo global, las perturbaciones crecen más en las regiones en donde existen inestabilidades en el flujo. Estas regiones son además los lugares más propicios

para la presencia de errores en el análisis. Este ciclo está resumido en la Figura 4.2.

El procedimiento descrito, indica cómo se puede llevar a cabo un ciclo de generación de perturbaciones que da origen a un par de análisis perturbados. Este ciclo, puede repetirse utilizando una perturbación inicial diferente que originará un par de perturbaciones distinto. Debido a que las perturbaciones generadas están estrechamente vinculadas con las inestabilidades del flujo, localmente puede existir similitud en las estructuras generadas por dos ciclos diferentes, en otras ocasiones distintos pares pueden mostrar estructuras diversas.

El ciclo de vectores criados guarda relación con los ciclos de asimilación de datos a partir del cual se obtiene el análisis. En el caso de los análisis, los pronósticos a corto plazo (6 horas a escala global) son utilizados como punto de partida (campo preliminar), los errores presentes en estos pronósticos son reducidos parcialmente, mediante la asimilación de datos. Esta reducción del error es más importante en aquellas áreas donde la disponibilidad de datos es mayor. Por otra parte, los errores presentes en el campo preliminar son más importantes en las regiones en donde la presencia de inestabilidades conducen a una mayor divergencia entre el pronóstico y la evolución de la atmósfera real. Esta reducción en la magnitud del error que se produce durante el proceso de asimilación guarda una estrecha similitud con la reducción en la amplitud de las perturbaciones que tiene lugar en el ciclo de generación de vectores criados (Toth y Kalnay, 1997). Es por eso que es esperable que exista una cierta correlación entre los errores del análisis y las perturbaciones generadas por el método de los vectores criados. De hecho, Corazza et al. (2003) demostraron que existe una importante similitud entre las estructuras de los errores presentes en el campo preliminar y los vectores criados, para un experimento realizado con un modelo cuasigeostrófico.

De acuerdo con Toth y Kalnay (1997), el subespacio de perturbaciones que son capaces de crecer a lo largo del plazo de pronóstico es relativamente pequeño comparado con las dimensiones del espacio donde se encuentra la solución de un modelo numérico. Es por eso que al seleccionar perturbaciones de este subespacio, un número menor de perturbaciones puede muestrear más eficientemente las posibilidades futuras de evolución del sistema.

La amplitud y estructura horizontal de las perturbaciones generadas a partir de esta metodología es, como se mencionó previamente, función de la amplitud que se le impone

a la perturbación en el momento en que la misma es reducida. Experimentos realizados con menores amplitudes de la perturbación, dan lugar a perturbaciones que responden más a inestabilidades de tipo convectivas, caracterizadas por una menor extensión horizontal (Kalnay, 2003). Mientras que las perturbaciones obtenidas permitiendo una mayor amplitud de las perturbaciones muestran una mayor extensión horizontal, comparable con las longitudes de onda del flujo de mayor escala (asociado principalmente a la inestabilidad baroclínica). De acuerdo con Kalnay 2003, este comportamiento representa una ventaja, ya que al seleccionar una amplitud conveniente, el método permite filtrar aquellas perturbaciones que pueden tener una tasa de crecimiento muy rápida inicialmente (por ejemplo aquellas asociadas a la inestabilidad convectiva), pero que saturan a las pocas horas debido a las interacciones no lineales y que no son relevantes en el problema del pronóstico a corto y mediano plazo.

El método de los vectores criados presenta además una importante ventaja en términos computacionales, ya que no requiere prácticamente de la realización de cálculos adicionales para la obtención de las perturbaciones. Por otra parte el algoritmo requerido por el método es de fácil implementación. Esto contrasta con otros métodos computacionalmente más costosos y más sofisticados como el método de los vectores singulares (Molteni et al. 1996).

En el presente trabajo, se utilizó la técnica de los vectores criados aplicada al modelo global MRF descrito en el Capítulo 2. Los programas necesarios para la implementación del método en este modelo fueron provistos por Itsvan Szunyogh (Comunicación personal). La frecuencia de reducción de amplitud en las perturbaciones se fijó en 6 horas que es coincidente con los valores utilizados operativamente en el ensamble global del NCEP y que producen perturbaciones que responden a las inestabilidades de mayor escala.

## **4.2 Características de los vectores criados y su relación con el error del análisis.**

Utilizando el modelo global, se inició un ciclo de generación de perturbaciones mediante vectores criados el día 1 de diciembre de 2002, finalizando el experimento el día 15 de febrero de 2003. Se generaron 5 pares de perturbaciones de acuerdo con el

esquema de la Figura 4.2. De estos 5 pares, se seleccionó un par para la realización de diagnósticos que brindaran información sobre las características básicas de las perturbaciones sobre la región.

Uno de los parámetros calculados sobre el par seleccionado es la diferencia cuadrática media (DCM) entre la perturbación positiva y la perturbación negativa para diferentes variables como se muestra en la ecuación 4.2, en donde  $P_+$  representa el valor de una determinada variable para la perturbación positiva y  $P_-$  es como en el caso anterior pero para la perturbación negativa. Estas diferencias son elevadas al cuadrado y promediadas en el tiempo en cada punto de retícula del dominio.

$$DCM = \sqrt{\frac{\sum (P_+ - P_-)^2}{N}} \quad (4.2)$$

En primer lugar, se buscó comparar la amplitud de las perturbaciones con la variabilidad de las variables en el período considerado. Para esto se parte de la base que el máximo error posible en un pronóstico es función de la variabilidad de las variables y que el error del análisis se estima en aproximadamente un 10% de la variabilidad. Este análisis, mostró que en latitudes medias la amplitud de las perturbaciones medida a través del DCM es un 10-20 % de la varianza de las variables temperatura y geopotencial en niveles medios y altos (la Figura 4.3, muestra un ejemplo para el campo de geopotencial en 500 hPa). Estos valores están cerca de los valores óptimos que mostraron una mejor relación entre la velocidad de crecimiento de las perturbaciones y el crecimiento del error en el pronóstico (Toth y Kalnay, 1997).

Por otra parte, la amplitud de las perturbaciones, debería ser similar a la amplitud de los errores presentes en el análisis. La estimación de dichos errores es difícil, no obstante existen algunas alternativas que pueden explorarse para determinar como podrían ser las características estadísticas de dichos errores y su distribución espacial. Una posible aproximación al problema, es utilizar ciclos de análisis independientes que trabajen con modelos y esquemas de asimilación independientes (aunque los datos asimilados sean los mismos) y utilizar las diferencias entre ambos como un estimador de las incertezas presente en el análisis, de ahora en adelante se utilizará la expresión incertezas para designar la diferencia entre los ciclos de análisis independientes y así distinguirlos de los

errores del análisis que no pueden ser conocidos. Si bien este procedimiento no nos permite ver cuales son los verdaderos errores del análisis, nos permite estudiar sus propiedades estadísticas y las magnitudes de los mismos. Para llevar este procedimiento a la práctica, en la presente tesis, se utilizaron los reanálisis JRA-25 en una resolución de 2.5x2.5 grados (algo menor que la resolución nativa de este conjunto de datos), los cuales fueron comparados con los reanálisis del NCEP durante todo el período en el que se desarrolló el experimento de los vectores criados. La comparación consistió fundamentalmente en el cálculo de las diferencias día a día y para diferentes variables entre ambos conjuntos de datos. A dicha diferencia se le restó la diferencia media entre ambos análisis calculada sobre el período de interés. La misma se sustrajo para evitar incluir en el análisis posibles diferencias sistemáticas vinculadas con errores sistemáticos en los modelos utilizados en ambos ciclos de asimilación. La diferencia cuadrática media, a partir de las diferencias diarias a las que se les sustrajo la diferencia media se muestra en la Figura 4.4 a. Dicha diferencia cuadrática media constituye una estimación de la magnitud de la incertidumbre presente en el análisis y como tal es comparada con la diferencia cuadrática media de una perturbación generada por el método de los vectores criados (Figura 4.4 b). En la Figura 4.4, la variable seleccionada es la altura geopotencial de 500 hPa. Lo que muestra el cociente entre la diferencia cuadrática media entre los dos análisis y la diferencia cuadrática media de la perturbación generada por el método de los vectores criados, es que la relación entre ambas oscila entre 0.5 y 1.6 indicando una buena relación entre las perturbaciones y la incertidumbre en el análisis en términos de amplitud. Es importante notar que las perturbaciones generadas por el método de los vectores criados reproducen adecuadamente los cambios norte sur en la amplitud de los errores presentes. Tanto las perturbaciones generadas como la incertidumbre presente en el análisis, muestran mayor amplitud en la porción sur del dominio donde la actividad de ondas baroclínicas es mayor. Por otra parte, la incertidumbre del análisis muestra un mínimo de incertidumbre sobre el continente extendiéndose corriente abajo sobre el Océano Atlántico. La amplitud de las perturbaciones reproduce un patrón similar, pero algo más débil. Este patrón responde probablemente a la presencia de mayor cantidad de observaciones sobre el continente, hecho que es tenido en cuenta en la generación de las perturbaciones a través de la utilización de una máscara (Figura 4.1 a).

$$\begin{aligned}
 BIAS &= \frac{1}{N} \sum_i (REA2 - JRA) \\
 DCM &= \sqrt{\frac{1}{N} \sum_i (REA2 - NCEP - BIAS)^2}
 \end{aligned}
 \tag{4.3}$$

Estos resultados sugieren que a grandes rasgos los vectores criados tienen estructuras similares a las de las incertezas presentes en el análisis tal cual sugiere la teoría. El interrogante que surge a continuación, es si existe alguna vinculación entre la estructura espacial de los vectores criados y de las incertezas presentes en el análisis, y también si existe una relación entre los cambios de amplitud de las mismas.

Para evaluar la existencia de la relación entre la estructura espacial y amplitud de los vectores criados y la incerteza en el análisis, se calcularon correlaciones espaciales para cada uno de los días entre los 5 pares de vectores criados y las diferencias entre los reanálisis del NCEP y los reanálisis JRA-25, considerando una subregión ubicada entre 60 y 45° de latitud sur y entre 70 y 45° de longitud oeste. El tomar una región más acotada, mejora las chances de tener una mejor correlación entre las incertezas del análisis y los vectores criados, dado que si la región estudiada es demasiado grande, los vectores criados y las incertezas del análisis responden en diferentes regiones a diferentes fuentes de inestabilidad, los signos podrían ser diferentes en cada caso lo cual tendería a disminuir la correlación. El coeficiente de correlación entre cada uno de los vectores criados y nuestra estimación de la incertidumbre presente en el análisis, tiene una gran variabilidad temporal. De todas maneras, lo importante es que al menos alguna de las perturbaciones utilizadas tenga una gran similitud con las incertezas del análisis. Por eso para resumir la información vinculada con las 5 perturbaciones obtenidas, para cada día se tomó el valor del máximo coeficiente de correlación elevado al cuadrado, es decir la máxima varianza explicada por el vector criado que explica un mayor porcentaje de la incertidumbre en el análisis. Por otra parte, sobre la subregión antes mencionada, se calculó la diferencia cuadrática media entre ambos ciclos de análisis sobre la región (para cada día) y la diferencia cuadrática media de la perturbación (también para cada día), de forma de tener 2 series temporales, una que representa la amplitud media de la perturbación en la región y otra que representa la amplitud del error del análisis estimado.

Para determinar la existencia de una relación entre la amplitud cuadrática media de

la perturbación y la de las incertezas presentes en el análisis se calcularon los coeficientes de correlación lineal entre las series temporales de ambos para las 5 perturbaciones. En todos los casos el resultado arrojó valores de coeficiente de correlación que no eran significativamente distintos de 0. Por otra parte se calcularon los valores medios de los coeficientes de correlación espacial y su sensibilidad a la magnitud de la incertidumbre presente en el análisis. Para esto se calculó la mediana de la distribución de errores cuadráticos medios para la incerteza del análisis y se calculó el porcentaje de varianza explicada máximo medio para los casos en los que la incerteza del análisis estaba por debajo de la mediana y para aquellos casos en los que estaba por encima de la mediana. Lo mismo se hizo con la amplitud cuadrática media de las perturbaciones. Este cálculo se llevó a cabo para varias variables. Los resultados se resumen en la Tabla 4.I.

	<b>950</b>		<b>850</b>		<b>500</b>		<b>250</b>	
	$r^2$	amplitud	$r^2$	amplitud	$r^2$	Amplitud	$r^2$	amplitud
<b>Geopotencial</b>								
Error bajo	0.23	11.96	0.25	11.73	0.27	13.65	0.35	15.40
Error alto	0.32	12.83	0.37	11.95	0.38	14.3	0.44	14.89
<b>Temperatura</b>								
Error bajo	0.19	0.63	0.24	0.71	0.23	0.73	0.26	1.03
Error alto	0.24	0.65	0.25	0.70	0.26	0.76	0.29	0.98
<b>Viento U</b>								
Error bajo	0.18	1.73	0.19	1.55	0.17	1.89	0.24	1.52
Error alto	0.20	1.98	0.21	1.73	0.24	1.88	0.35	1.56

**Tabla 4.I: Amplitud media de las perturbaciones y varianza máxima media explicada por las perturbaciones para los casos en los que la incerteza del análisis es alta y los casos en los que la incerteza del análisis es baja. La amplitud de los perturbaciones/incertezas está medida en mgp, K y  $m s^{-1}$ .**

Los resultados que se muestran en la Tabla 4.I, indican que el comportamiento es dependiente de la variable y el nivel considerados. El geopotencial, muestra poca sensibilidad en la amplitud de la perturbación a la amplitud estimada de la incerteza presente en el análisis, mientras que sí muestra cierta sensibilidad en la correlación espacial entre ambas. En particular, muestra que la correlación es mayor cuanto mayor es la magnitud de la incerteza en el análisis. Esto es importante, ya que muestra que los errores más grandes son más eficientemente capturados por el método de los vectores criados, pudiendo llegar a un 44 % de varianza explicada en promedio por la perturbación que más explica en los 250 hPa. En niveles bajos estos porcentajes resultan menores, no

obstante se mantiene la dependencia con la magnitud de la incerteza presente en el análisis. La temperatura muestra una menor cantidad de varianza explicada por las perturbaciones y una menor sensibilidad de la misma a la amplitud de la perturbación. Por otra parte tampoco se observa sensibilidad en la magnitud de las perturbaciones con la de la incerteza del análisis. En el caso de la componente U del viento, tampoco se observa mucha sensibilidad a la magnitud de la incerteza en cuanto a la de la perturbación, si existe algún indicio de sensibilidad en cuanto a la correlación espacial sobre todo en niveles altos. Para ejemplificar este comportamiento la Figura 4.5 muestra 2 ejemplos de la perturbación en el geopotencial de 500 hPa, contra la estimación de la incerteza del análisis utilizada en estos experimentos. En la parte a de la figura, se puede ver un caso en el que la correlación espacial resultó alta, como se puede apreciar sobre el extremo sur de Sudamérica y sobre el Pacífico, existe una diferencia entre ambos sistemas de análisis que tiene una estructura de dipolo. La estructura de la perturbación generada por los vectores criados tiene una estructura muy similar e incluso una magnitud que es acorde a las diferencias observadas entre ambos análisis. Notar que si bien la correlación es elevada, el signo de la perturbación es opuesto a las diferencias entre ambos análisis, esto es totalmente arbitrario, dado que depende de cómo se tomen las diferencias entre los pronósticos utilizados para calcular la perturbación o las diferencias entre los análisis. Los vectores criados de alguna manera en este caso están mostrando la posible forma que podrían tener los errores del análisis y su amplitud, pero no pueden determinar el signo que tendrán dichos errores. Estos errores estuvieron asociados a un sistema de baja presión intenso (no se muestra la figura). Por otra parte, la Figura 4.5 b, muestra una situación en que la estimación de las incertezas del análisis no parece tener mucha relación con la distribución espacial de las perturbaciones (en este caso al igual que en el anterior, la perturbación que se muestra es aquella que obtuvo el mayor coeficiente de correlación). En este caso, la metodología falló en representar los errores importantes que se observan en el Atlántico sobre el sector sudeste del dominio mostrado en la figura. Sobre el extremo sur de Sudamérica y sobre el Atlántico sur hay dos centros que tienen el mismo signo en el error del análisis pero diferente signo en la perturbación, indicando que la metodología falló en este caso en capturar esa estructura. Esta situación corresponde a un flujo más zonal en donde no se apreciaba la presencia de un sistema importante de baja o alta presión sobre el dominio analizado.

Un aspecto que puede oscurecer la relación entre los vectores criados y las incertezas presentes en el análisis tal cual son estimadas en el presente capítulo, es que las diferencias entre los dos sistemas de análisis independientes no sólo responden a las inestabilidades del flujo, sino que también responden a los errores presentes en el modelo y en los errores vinculados al sistema de asimilación en si. Por otra parte los vectores criados, responden únicamente a las inestabilidades del flujo y no a los errores presentes en el modelo. Esta limitación intentó ser tenida en cuenta al restar la componente sistemática de la diferencia entre ambos ciclos de análisis, no obstante esto no garantiza que todo el error debido a errores en el modelo haya sido removido. Por otro lado, los errores del análisis no pueden ser conocidos con exactitud en experimentos aplicados a la atmósfera real como en el presente estudio.

Otra alternativa disponible para estimar la magnitud y distribución espacial de los errores presentes en el análisis, es la de utilizar análisis obtenidos utilizando el mismo sistema de asimilación, pero habiendo incorporado distintos conjuntos de datos en el proceso de asimilación. Para llevar a cabo este experimento, se utilizaron los análisis GDAS para el período Diciembre 2002-Febrero 2003 con los datos SALLJEX asimilados (Herdies et al. 2007). La diferencia entre análisis (el que tiene los datos SALLJEX asimilados y el que no) es tomada como una medida del error presente en el análisis. En estos experimentos no se pudo encontrar una relación entre la estructura de las diferencias entre ambos análisis y la estructura de la perturbación generada a partir del método de los vectores criados. Por otra parte y a diferencia del experimento anterior, tampoco se pudieron encontrar similitudes entre la amplitud y distribución espacial en términos estadísticos tal como lo muestra la Figura 4.6, que ilustra la amplitud media de las diferencias del análisis y la amplitud de los vectores criados para la altura geopotencial de 500 hPa. En dicha figura se puede ver que los vectores criados alcanzan su máxima amplitud en la zona de los oestes donde las inestabilidades del flujo favorecen su crecimiento, mientras que las diferencias entre los análisis se concentra en la región en donde se asimilaron las observaciones del experimento SALLJEX (es decir en la región donde los conjuntos de datos incorporados durante el proceso de asimilación son diferentes). Esto muestra que las diferencias entre ambos análisis responden principalmente a la distribución de las observaciones asimiladas y maximizan en esa región que no coincide necesariamente con la región donde se producen los máximos de los

vectores criados. Por otra parte experimentos de sensibilidad realizados con estos análisis, utilizando un modelo regional, muestran que las diferencias experimentan poco crecimiento durante el pronóstico, a diferencia de los vectores criados que como se discutirá más adelante muestran un mayor crecimiento (Paegle et al. 2007). En la Figura 4.7 se compara el crecimiento de la amplitud de la perturbación como función del plazo de pronóstico, para pronósticos regionales inicializados utilizando un vector criado como perturbación en la condición inicial y pronósticos inicializados a partir de los dos conjuntos de análisis (SALLJEX y NO-SALLJEX).

En resumen, los experimentos orientados a encontrar una relación entre los vectores criados y las incertezas presentes en el análisis, permitieron encontrar una similitud en la distribución espacial de las zonas donde la amplitud de las perturbaciones es mayor y de las zonas donde el error del análisis es mayor, este último estimado como las diferencias entre los análisis generados por dos ciclos de asimilación independientes (JRA-25 y Reanálisis 2). También se encontró que la correlación entre las incertezas del análisis y los vectores criados en algunas variables como el geopotencial, son sensibles a la magnitud de la incertidumbre presente en el análisis, es decir que cuanto más grande es la incertidumbre, mayor es la correlación entre el vector criado y las incertezas. No obstante, la amplitud de las perturbaciones, no mostró sensibilidad, indicando que una amplitud mayor en las perturbaciones no estaría indicando que las incertezas en el análisis son potencialmente mayores. Por otra parte se identificaron algunas limitaciones para estimar los errores del análisis en un caso real. Para poder profundizar un poco en estas cuestiones sin las limitaciones que presenta la estimación de los errores del análisis en casos reales, el problema de la relación entre los errores del análisis y los vectores criados se re-examinó utilizando el modelo de Lorenz de 3 variables (Lorenz 1969). Los resultados se muestran en el Apéndice B.

### **4.3 Aplicación de la metodología en la generación de un ensamble global y otro regional.**

Una vez que se generaron las perturbaciones mediante el método de los vectores criados, para el período correspondiente a la estación cálida 2002-2003, se inicializaron ensambles globales utilizando las condiciones iniciales perturbadas (11 miembros en total,

5 pares de perturbaciones y un pronóstico control). Los pronósticos se inicializaron 2 veces por día a las 00 y 12 UTC y a un plazo máximo de 48 horas. Las integraciones del modelo global necesarias para completar esta tarea equivalen a 7.5 años de simulación.

De acuerdo con los objetivos planteados en el presente trabajo de tesis, a partir del ensamble global, se procedió a generar un ensamble regional utilizando el modelo WRF. En el ensamble regional, cada uno de los miembros obtiene condiciones iniciales y de borde de uno de los miembros del conjunto global y es integrado por un plazo total de 48 horas.

El modelo WRF fue configurado de una manera similar a la utilizada en el Capítulo 3:

- Proyección Lambert conformal con 40 km de resolución horizontal y 31 niveles verticales.
- Parametrización Ferrier para la microfísica (Ferrier et al. 2002).
- Parametrización de Kain-Fritsch, (Kain, 2004).
- Esquema RRTM para la radiación de onda larga (Mlawer et al, 1997) y el tratamiento descrito por Dudhia (1989) para la radiación de onda corta.
- Capa límite YSU (Hong y Pang, 1996).
- Modelo de superficie NOAH (Chen y Dudhia, 2001)

El archivo de configuración ("namelist") utilizado para correr el modelo se puede encontrar en el Apéndice C.

A diferencia del ensamble global que utiliza los reanálisis del NCEP como condición inicial, en el ensamble regional se utilizan los análisis GDAS como condición inicial, los mismos son más adecuados para inicializar los pronósticos en el modelo regional ya que tienen una mayor resolución. Las perturbaciones iniciales (que fueron generadas en una resolución de  $2.5^{\circ} \times 2.5^{\circ}$ ) fueron interpoladas utilizando interpolación bilineal a la resolución de  $1^{\circ} \times 1^{\circ}$  y sumadas y restadas a dichos análisis. Las condiciones de borde fueron provistas directamente por el modelo MRF en su resolución original, dado que cada miembro del ensamble original está anidado en un miembro del ensamble global, las condiciones de borde difieren entre los diferentes miembros del modelo regional permitiendo un mayor crecimiento de las perturbaciones en el dominio regional.

En general las diferencias sistemáticas que presentan los análisis GDAS y los

reanálisis (que se utilizan como condición inicial para el pronóstico control del conjunto global) son pequeñas (no se muestra la figura) indicando que pocas discrepancias podrían surgir de inicializar ambos ensambles con condiciones iniciales diferentes.

El conjunto regional fue inicializado 2 veces al día a las 00 y 12 UTC al igual que el conjunto global e integrado por un período total de 48 horas. Las salidas de este conjunto se archivaron cada 6 horas para reducir la cantidad de espacio en disco necesarias para su almacenamiento. El período cubierto por los pronósticos del conjunto regional es el mismo que el abarcado por el conjunto global y va desde el 15 de diciembre de 2002 hasta el 15 de febrero de 2003. La evaluación de la calidad de los pronósticos con estos nuevos pronósticos por conjuntos se trata en los Capítulos 5 y 6.

## **Capítulo 5:**

# **El ensamble regional y global basado en la técnica de los vectores criados: reducción del error y estimación de la incertidumbre.**

### **5.1 Introducción:**

De acuerdo con Kalnay (2003), los objetivos principales del pronóstico por ensambles son reducir el error mediante la utilización de la media del ensamble, determinar las regiones donde la incertidumbre de la previsión es mayor utilizando la dispersión de los miembros del ensamble y producir pronósticos probabilísticos que sean confiables para diferentes variables. En esta sección se presentarán algunos resultados orientados a mostrar en qué medida los ensambles basados en la técnica de los vectores criados, generados en el experimento realizado para la estación cálida 2002-2003, permiten avanzar en los primeros 2 objetivos sobre nuestra región. El análisis de la capacidad del ensamble para producir pronósticos probabilísticos confiables sobre la región será analizado en el Capítulo 6.

Varios trabajos han demostrado que la media del ensamble efectivamente reduce el error de las previsiones tanto en los pronósticos a largo plazo (Kalnay 2003 y referencias allí citadas) como en la predicción a corto plazo (Hamill y Colucci 1997, Ruiz et al. 2006, Silva Dias et al. 2006). Esta reducción se debe principalmente a que el promedio sobre los miembros del ensamble actúa como un filtro dinámico que reduce la amplitud de aquellos patrones del flujo que tienen asociada una mayor incertidumbre en el pronóstico. En este sentido la media del ensamble ofrece una herramienta muy útil a aquellos usuarios que requieren una estimación del valor futuro de alguna variable en particular (por ejemplo la temperatura máxima). No obstante, para otras aplicaciones como la predicción de fenómenos extremos, la media del ensamble puede no ser la herramienta más adecuada

dado que produce un suavizado de los campos pronosticados.

El segundo aspecto resaltado por Kalnay (2003) tiene que ver con la capacidad del ensamble de detectar períodos y regiones donde la incertidumbre en el pronóstico es mayor. Desde el punto de vista de la aplicación, si existe una relación importante entre la dispersión del ensamble y el error de la media, entonces es posible tomar decisiones basadas en la incertidumbre del pronóstico. Una aplicación directa podría ser la asignación de rangos de error para la previsión que varíen día a día como función de la dispersión del ensamble. La discusión y los experimentos que se detallan a continuación tienen como principal objetivo, describir la relación entre la dispersión del ensamble y los errores en el pronóstico para los ensambles regionales y globales. Dicha relación será evaluada en términos probabilísticos, pudiendo eventualmente constituir la base para la elaboración de un pronóstico probabilístico del error en el pronóstico.

En un ensamble ideal, la verificación debería ser indistinguible (desde el punto de vista estadístico) de los miembros individuales del ensamble (Kalnay 2003, Hou et al. 2001, Whitaker y Lough 1998). Esto implica que donde la dispersión entre los miembros es mayor, la probabilidad de que la verificación y la media del ensamble se alejen aumenta (es decir que existe una adecuada relación entre la dispersión del ensamble y la función de distribución de los errores del modelo). Para que un ensamble se comporte como un ensamble ideal, es necesario que las diferentes fuentes de incertidumbre estén consideradas dentro del ensamble de forma tal que la dispersión del ensamble considere a los errores del modelo y a la incertidumbre en las condiciones iniciales como posibles fuentes de error en el pronóstico.

Houtekamer (1993) desarrolló un modelo teórico sencillo que ejemplifica la relación que debería esperarse entre la dispersión del ensamble y el error en la media del ensamble para un ensamble ideal. Básicamente este modelo propone que el error en la media proviene de una distribución normal con media 0 y un desvío estándar igual a la dispersión del ensamble. La dispersión del ensamble ( $S$ ) se calcula de acuerdo a la ecuación 5.1, donde  $M$  es el número total de miembros del ensamble,  $f_i$  es el pronóstico del miembro  $i$ -ésimo y  $\bar{f}$  es el valor de la media del ensamble. De acuerdo con los resultados de Houtekamer (1993), en el marco de este modelo simplificado, la correlación lineal entre el error y la dispersión depende de la variabilidad temporal de la dispersión en el ensamble (a mayor variabilidad temporal, mayor coeficiente de correlación) y tiene un

valor límite de 0.79. Es decir que aún para un ensamble perfecto, la relación entre error y dispersión medida a través del coeficiente de correlación lineal, no puede superar dicho valor. Whitaker y Loughe (1998), exploraron la correlación lineal existente entre la dispersión del ensamble y los errores en la media para el hemisferio norte y encontraron valores relativamente bajos (menores a 0.6) y que decaían hacia el ecuador.

$$S = \sqrt{\frac{1}{M-1} \sum_M (f_i - \bar{f})^2} \quad (5.1)$$

Los resultados obtenidos por estos autores, ponen de manifiesto las limitaciones de estudiar la vinculación entre dispersión y error desde el punto de vista de la correlación lineal. Por otra parte Whitaker y Loughe (1998) y en mayor medida Gritit y Mass (2007), muestran otras formas de medir la relación que se basan en análisis de la distribución conjunta de dispersión y error que se independizan de la hipótesis de la existencia de una relación lineal entre ambas. Gritit y Mass (2007) retoman el modelo teórico de Houtekamer (1993) y lo utilizan para analizar la relación existente entre la dispersión y el error medido en diferentes formas, sugiriendo que para demostrar la relación existente entre ambas variables, se debe discutir la consistencia estadística entre la distribución del error y la distribución de la dispersión del ensamble. Por otra parte los autores confirman un resultado sugerido en varios trabajos anteriores en el sentido de que la relación entre la dispersión y el error es más intensa cuando la magnitud de la dispersión es mayor.

En el presente trabajo se busca encontrar medidas de la relación entre la dispersión y el error que no asuman una relación lineal entre ambas variables y que representen en alguna medida el potencial de utilidad práctica de dicha relación. A continuación se describen los aspectos metodológicos de cómo se llevaron a cabo dichos objetivos en la presente tesis.

## 5.2 Metodología

### 5.2.1 Relación entre el error de la media del ensamble y el error en el pronóstico control:

Para evaluar la reducción en el error de la media del ensamble con respecto al error del pronóstico control primero se calculó el bias del pronóstico como la media temporal ( [ ] ) en cada punto de retícula de la diferencia entre el pronóstico y la observación (Ecuación 5.2) donde  $f_i$  es el pronóstico para un determinado punto de retícula en el tiempo  $i$ ,  $o_i$  es el valor de la verificación en dicho tiempo. Esta definición corresponde a la componente constante del error del modelo la cual se computa en forma independiente sobre cada punto de retícula. El error absoluto (EA) se define para cada punto de retícula y para cada tiempo como el módulo de la diferencia entre el pronóstico y la verificación, menos el bias (Ecuación 5.2). Finalmente, el error absoluto medio (EAM) se calcula como el promedio del EA sobre todo  $i$ , donde  $N$  es el número total de pares pronóstico-verificación utilizados en el cálculo de dicho promedio. En el presente trabajo, el promedio de los errores absolutos puede ser tomado como un promedio temporal (en cuyo caso  $N$  sería el número total de tiempos considerados para el promedio), espacial (en cuyo caso  $N$  sería el número total de puntos de retícula promediados) o ambos al mismo tiempo, aclarándose en cada caso particular de qué promedio se trata.

$$BIAS = [f_i - o_i]$$

$$EA_i = |f_i - o_i - BIAS| \quad (5.2)$$

$$EAM = \frac{1}{N} \sum_n EA_i$$

Para la verificación se utilizaron diferentes conjuntos de datos. Por un lado se utilizaron los reanálisis JRA-25. Estos datos proveen una fuente de verificación diferente a las condiciones iniciales (y en el caso del modelo regional de las de borde) utilizadas por los ensambles analizados. También se utilizaron datos de superficie de la red GTS que

fueron interpolados a la retícula de los reanálisis para verificar que los resultados obtenidos fueran consistentes con las observaciones.

En el caso del ensamble regional, si bien los experimentos fueron realizados con una resolución de 40 km, en la mayoría de los resultados presentados en este capítulo, las salidas del ensamble regional fueron llevadas a una resolución de  $2.5^{\circ} \times 2.5^{\circ}$  utilizando el método de promedio por cajas. Este cambio en la resolución tiene como objetivo hacer los resultados del modelo regional comparables con el JRA-25. Si la interpolación se hiciera en sentido inverso, es decir si se llevaran los reanálisis JRA-25 a la retícula de 40 km podrían generarse errores espurios, porque patrones de menor escala representados en el modelo regional serían comparados contra un análisis que no tiene dicha información. La forma en la que se llevó a cabo la interpolación permite una comparación más equitativa entre los errores del pronóstico global y los errores del pronóstico regional, ya que en ambos casos se abarcan las mismas escalas espaciales.

### **5.2.2 Relación entre la dispersión del ensamble y el error de la media del ensamble.**

De acuerdo con Gritit y Mass (2007) el análisis de la relación entre la dispersión y el error se basará en medidas que no supongan una relación lineal entre ambas y analizando diferentes aspectos de la distribución conjunta.

En el presente trabajo de tesis, el análisis se lleva a cabo de forma tal que permita estudiar cómo es la dependencia espacial y con el plazo de pronóstico entre error y dispersión. No obstante, dado que la región abarcada es muy extensa y comprende regiones con climas totalmente distintos es necesario primero transformar los valores de dispersión y error de forma tal que sean comparables entre las diferentes regiones. Para eso se calculó el valor medio y la desviación estándar de las series temporales del EA y de dispersión (S) para cada uno de los puntos de retícula. Utilizando estos valores se estandarizó el valor de error y dispersión en todo el dominio utilizando el valor medio y la desviación estándar de las series temporales en cada punto de retícula. Esto permite una mejor comparación de los resultados obtenidos en las diferentes regiones.

Las variaciones en los valores medios y la desviación estándar de los errores y la dispersión del ensamble no solo dependen de la región considerada sino que también

dependen de la hora del día. En este caso, los errores son calculados a las 12 y a las 00 y ambos conjuntos de datos son combinados de manera tal que permita incrementar el tamaño de las muestras con el que se trabaja. No obstante, existe un importante ciclo diurno en los errores y en la dispersión del ensamble, sobre todo en los niveles cercanos a la superficie (variables de superficie, 925 y 850 hPa). Para que estas variaciones diurnas no contaminen la relación existente entre dispersión y error la estandarización se realiza en forma independiente para el ciclo de las 00 y el de las 12 UTC, es decir que se calculan valores de media y desviación estándar para los pronósticos de ambos ciclos en forma independiente. Los valores estandarizados serán identificados con un subíndice e, por ejemplo  $EA_e$  identifica el valor del error absoluto estandarizado.

El análisis realizado se basa fundamentalmente en cuantificar los cambios en la función de distribución de probabilidades del error de acuerdo con los valores de dispersión obtenidos a partir del ensamble. Para eso se divide el rango de valores de  $S_e$  en  $k$  categorías, las mismas están definidas a partir de los límites de los  $q$ -percentiles de la distribución de la  $S_e$  (donde  $q=1/k$ ) de forma tal de que todas contengan el mismo número de elementos. Para cada categoría de dispersión, vamos a tener una muestra de errores y sus respectivos  $EA_e$ .

Debido a que el análisis se lleva a cabo en cada punto de retícula uno de los principales limitantes es el tamaño de las muestras utilizadas. Si tenemos en cuenta que el período comprendido va entre el 15 de diciembre de 2002 y el 15 de febrero de 2003, tenemos 120 valores de dispersión y de error. Si por ejemplo dividimos estos valores en 4 categorías de dispersión, obtenemos muestras de 30 datos para cada categoría lo cual es insuficiente para realizar un análisis de la función de distribución de probabilidades para los errores. Es por eso que para aumentar el tamaño de las muestras, se utiliza un procedimiento como el sugerido por Cusack y Arribas (2008) en donde a la muestra total en cada punto de retícula, se la compone por los valores en dicho punto y las series de dispersión y error en los 8 puntos vecinos con lo cual el tamaño de la muestra total asciende a 1080 elementos y las muestras para cada categoría de dispersión a un total de 270 (asumiendo una división en 4 categorías). En esta tesis vamos a utilizar 4 categorías para no reducir más el tamaño de las muestras de error para cada categoría de dispersión. Es importante destacar que con esta metodología los resultados obtenidos para puntos de retícula adyacentes no son independientes entre sí, ya que las muestras utilizadas en los

cálculos en puntos de retícula adyacentes comparten un 66% de la muestra. De todas maneras, los valores de dispersión y error en puntos de retícula adyacentes no son independientes entre si. Es importante destacar que uno de los objetivos de realizar la estandarización previa de las variables en cada punto de retícula es también remover cualquier correlación espuria que pueda surgir al combinar puntos de retícula adyacentes dentro de una misma muestra. Este efecto podría ser importante sobre todo en zonas donde existe una fuerte variación espacial de la magnitud media de los errores o de la dispersión. Estos aspectos serán discutidos más adelante cuando se analice la significancia estadística de los resultados.

De acuerdo con Gritmit y Mass (2007) y Houtekamer (1993), debería existir una relación lineal 1 a 1 entre la dispersión del ensamble y la desviación estándar del error dentro de cada categoría. En el presente caso, esta relación se estudia analizando cómo es la desviación estándar del error en cada una de las categorías. A partir de estos valores se puede estudiar si existe una relación lineal entre la desviación estándar del error y el valor central de dispersión correspondiente a cada una de las categorías analizadas.

Para estudiar la relación existente entre las variaciones temporales y espaciales de la dispersión y del error, se estudian los cambios en la distribución de probabilidades del  $EA_e$  dentro de cada una de las categorías de dispersión definidas. Al igual que como se hizo para definir las categorías de dispersión, se definen umbrales de error a partir de los límites de los intervalos cuartiles de la distribución de errores en cada punto. Se eligen los valores de  $EA_e$  tal que la probabilidad de ocurrencia de errores por encima de dicho valor sea 25, 50 y 75%. Estos valores son independientes de la dispersión y la probabilidad asociada a cada umbral se la llamará  $P_c$ , no obstante, dichos valores tienen una importante variación espacial. La idea es determinar cómo se modifican las probabilidades de ocurrencia de errores por encima de dichos umbrales para las diferentes categorías de dispersión definidas. A la probabilidad de ocurrencia de  $EA_e$  por encima de un determinado umbral asociado a cada categoría de dispersión lo vamos a denominar con  $P_{cs}$  (Es decir la probabilidad condicional de tener un error por encima del valor del umbral dado una determinada categoría de dispersión). De no existir relación entre la dispersión y el error  $P_{cs}$  debería ser igual a  $P_c$  para todas las categorías de dispersión utilizadas.

La Figura 5.1 ilustra esquemáticamente la metodología propuesta. En este esquema la línea roja identifica el valor de  $P_c$  para un umbral de error con probabilidad de

ocurrencia climatológica de 50%. Si la relación entre error y dispersión fuera determinística y perfecta, entonces Pcs sería 1 cuando se supera determinado umbral de dispersión (aquel cuyo valor de frecuencia acumulada sea igual a Pc) y 0 si el valor de la dispersión está por debajo de dicho valor, a esta curva la denominaremos Pd. Esta situación está representada por la curva azul en el diagrama 5.1. Por otra parte, como la relación entre error y dispersión no es determinística y además existen diversos factores que limitan la capacidad de los ensambles reales de poder representar adecuadamente la incertidumbre presente en los pronósticos la relación que se observa en algunos casos se parece más a la que muestra la curva roja discontinua que es una representación esquemática de la curva Pcs para un ensamble real. En este ejemplo la probabilidad de tener errores que superen el umbral, aumenta a medida que el valor de la dispersión en el ensamble es mayor. Cuanto mayor es la pendiente de la curva Pcs como función de la categoría, mayor es la utilidad potencial de la relación existente entre dispersión y error ya que indica que la resolución o la capacidad de discriminar entre eventos con errores grandes o errores chicos es mayor.

Para resumir la información contenida en un diagrama del tipo esquematizado en la Figura 5.1, se define un índice de resolución (IR) como lo expresa la ecuación 5.3, donde k es un subíndice que identifica las diferentes categorías de dispersión. Este índice mide la distancia entre Pcs y Pd. La ecuación 5.4, muestra como sería el valor del índice si no existiera ninguna relación entre ambas variables (IRc), es decir si solo tuviéramos la información climatológica del comportamiento de los errores en el pronóstico. En el contexto de la generación de un pronóstico probabilístico, la resolución es uno de los atributos más importantes, porque como se ha discutido en el Capítulo 3, los pronósticos probabilísticos pueden ser calibrados de forma tal de mejorar la confiabilidad pero estos algoritmos no modifican la resolución del pronóstico.

$$IR = \sum_k (p_{cs}(k) - p_d(k))^2 \quad (5.3)$$

$$IRc = \sum_k (p_c - p_d(k))^2 \quad (5.4)$$

A partir de estas cantidades se puede definir un parámetro que tenga un valor máximo de 1 (Pcs coincida con Pd) y un valor de 0 que indique la ausencia total de

relación entre ambas variables, que se lo denominará IRSS (Ecuación 5.5). Los índices propuestos cumplen con el objetivo de no asumir una relación lineal entre dispersión y error y están orientados a medir la utilidad práctica potencial de dicha relación.

$$IRSS = 1 - \frac{IR}{IRc} \quad (5.5)$$

En la presente tesis, se analiza la relación entre EA y la dispersión, otros trabajos como Gritit y Mass (2007) exploraron la relación para otras formas de definir el error del pronóstico. En este caso se utiliza el EA por considerar que su estimación tiene una aplicación práctica directa, mientras que otras medidas como la dispersión del error de los miembros del ensamble si bien muestran una relación más directa con la dispersión (por motivos de definición) no resultan herramientas tan útiles en la evaluación del nivel de error que un usuario final podría esperar.

La metodología propuesta es aplicable también a la verificación de pronósticos determinísticos, dado que evalúa la resolución de los mismos. Además, si se calcula utilizando como umbrales diferentes percentiles de la distribución de las variables, la metodología permite discriminar variaciones en la resolución del pronóstico como función del valor de la variable (por ejemplo permitiría determinar si el pronóstico en condiciones extremas es más o menos preciso que cuando las condiciones están cercanas a los valores medios.

Para analizar la significancia estadística de los resultados y ante la inexistencia de un test que permita evaluarla en forma teórica y bajo hipótesis que se verifiquen en las muestras utilizadas, se recurrió a la generación de un gran número de muestras al azar. En dichas muestras se generaron series de error y dispersión sin ninguna correlación entre si a partir de una distribución normal. A dichas series se les calculó el valor absoluto para simular las asimetrías de las series de EA y S en los casos reales y para cada serie se calculó el valor del IRSS. Se generaron en total 10000 muestras aleatorias de 1080 elementos cada una para representar muestras con el mismo tamaño que las utilizadas en la presente tesis. Si asumimos que los elementos de estas muestras son independientes, entonces 99% de los casos tenían valores de IRSS inferiores a 0.15 en valor absoluto y 95% de los casos están por debajo de 0.11 constituyendo estos los umbrales de significancia medidos empíricamente bajo el supuesto de independencia de los elementos

de la muestra. No obstante, los elementos que componen la muestra, no son independientes dada la metodología utilizada para la construcción de las mismas. Los errores en los puntos de retícula adyacentes utilizados para incrementar el tamaño de las muestras no son independientes de los errores en el punto central (Parrish y Derber, 1992), de hecho en algunos casos la relación se extiende por varios cientos de kilómetros, sobre todo para variables como la presión y la temperatura. Por otra parte también existe una persistencia en los valores de error y dispersión (la correlación temporal es función de la posición, pero es menor que la correlación espacial con valores de  $r$  del orden de 0.7) Para tener estos elementos en cuenta, se generó una serie con un  $r$  en lag 1 de 0.95. Las series que representan los puntos vecinos fueron obtenidas a partir de esta misma serie con una correlación lineal también de 0.95. De esta manera, las muestras fueron construidas de una forma similar a como sucede en el caso real. Utilizando este procedimiento se generaron 10000 series más bajo estas suposiciones y se obtuvo que 99% de los valores de IRSS estuvieron por debajo de 0.26 en valor absoluto y 95% estuvo por debajo de 0.21. Estos valores fueron obtenidos con coeficientes de correlación más altos que los que se observan en los casos reales por lo cual podrían tomarse como una cota superior para el nivel de significancia para las variables discutidas en el presente trabajo.

## 5.3 Resultados

### 5.3.1 Relación entre el error del control y el error de la media del ensamble.

A continuación, se muestran los resultados respecto de la reducción de los errores en el modelo utilizando la media del ensamble como predictor con respecto al error del pronóstico control. Los resultados se expresan en cambio porcentual en el error de acuerdo con la Ecuación 5.6, donde  $\Delta\epsilon$  es el cambio porcentual en el error, valores negativos indican una mejora en la media del ensamble con respecto al control. Para este cálculo en particular se utilizó el EAM tomado como promedio temporal sobre cada punto de retícula y calculado a partir del EA (no estandarizado). Los subíndices  $m$  y  $c$  representan a la media del ensamble y al control respectivamente.

$$\Delta\epsilon = \frac{EAM_m - EAM_c}{EAM_c} \quad (5.6)$$

El valor de diferencia porcentual se calculó para la presión a nivel del mar (PNM), la temperatura a 2 metros (T2M), la humedad específica a 2 metros (Q2M) y para las variables altura geopotencial, temperatura (T), humedad específica (Q), componente zonal (U) y meridional (V) del viento para los niveles de 925, 850, 500 y 250 hPa. Por una cuestión de brevedad, se hace hincapié en las variables de superficie que son las que presentan un mayor interés para diversas aplicaciones.

La Figura 5.2, muestra un ejemplo de los resultados obtenidos para la humedad específica a 2 metros para el ensamble global. El error asociado a esta variable, es máximo en la región centro y norte de Argentina tanto a 24 como a 48 horas de pronóstico. En esta región se observa también una zona de fuerte gradiente en el campo medio de distribución de la variable sugiriendo que en parte las advecciones asociadas a las ondas baroclínicas en esta región pueden producir cambios importantes en la humedad (no se muestra). Por otra parte, los errores son más importantes sobre el continente que sobre el océano. Dos factores podrían estar contribuyendo a estas diferencias, por un lado la determinación de los flujos de superficie es mucho más compleja sobre el continente que sobre el océano lo cual podría conducir a un mayor error, y por otra parte, en los océanos no hay mediciones directas de la humedad cerca de la superficie.

Con respecto a la diferencia porcentual entre el error de la media del ensamble y el error del control, se puede observar que la reducción es de entre un 2.5 y un 10% para la mayor porción del dominio con una distribución espacial bastante uniforme. La reducción del error a 48 horas es mayor en términos porcentuales que lo que se observa a 24 horas.

La Figura 5.3, muestra estos mismos campos, pero contruidos a partir de observaciones de la red GTS. Es importante destacar que en esta figura, el patrón del error en Q a 2 metros, es muy similar al encontrado utilizando los reanálisis JRA-25 como verificación. Por otra parte la distribución espacial y los órdenes de magnitud encontrados para la reducción del error en el pronóstico al utilizar la media del ensamble también son similares a los encontrados con los JRA-25, no solo para esta variable sino también para PNM y T2M. Esto sugiere que la utilización de los JRA-25 permite estudiar las propiedades básicas de los errores del pronóstico y del ensamble.

La Figura 5.4, muestra el error, comparando contra los JRA-25 para los pronósticos de Q2M obtenidos a partir del ensamble regional. En dicha figura se puede apreciar que el error en el pronóstico regional es menor al observado para el ensamble global, sobre todo para la región centro y norte de Argentina que era una de las regiones donde el ensamble global tenía más problemas para reproducir la evolución de esta variable. Con respecto a la reducción de la magnitud de los errores por parte de la media del ensamble, la reducción asociada al ensamble regional es similar a la observada para el ensamble global con excepción de lo que ocurre sobre la región amazónica a 24 horas de pronóstico donde la reducción del error para el ensamble regional es menor que para el ensamble global.

La Figura 5.5, muestra los errores y la mejora porcentual en la media del ensamble para la presión a nivel medio del mar. En este caso, las diferencias entre el control y la media del ensamble son menores. De hecho el cambio porcentual es menor a 2.5 % sobre la mayor parte de Argentina. Para el ensamble regional, los errores son similares y resultados similares se observan con respecto a la reducción del error por parte de la media del ensamble (no se muestra la figura).

La Figura 5.6 resume la evolución a distintos plazos de los errores de la media del ensamble global y del ensamble regional promediados temporalmente y espacialmente sobre las regiones indicadas en la Figura 5.2. Estas regiones fueron seleccionadas para ilustrar diferentes comportamientos en los errores y como se discutirá más adelante, en la relación entre el error y la dispersión. La Región 1, se ubica sobre el norte y centro de Argentina, la Región 2, abarca gran parte de la región tropical de Sudamérica y la Región 3 se encuentra sobre el océano cerca del extremo sur del dominio en una región dominada por las perturbaciones baroclínicas. Las mejoras más notorias en cuanto a la reducción del error por parte del ensamble, se producen para Q2M y sobre todo para la región norte de Argentina (Región 1), para esta misma variable también son notorias las diferencias que existen entre el pronóstico global y el regional sobre la Región 2. Para PNM y T2M, la reducción relativa del error es menor. Es posible que esto se deba en parte a la mayor variabilidad espacial del campo de humedad específica en donde el promedio pueda filtrar más eficientemente variaciones de pequeña escala poco pronosticables. Los campos de temperatura y presión a nivel del mar son más homogéneos. La eficiencia del ensamble regional en reducir el error es comparable a la del ensamble global, con la ventaja de que el ensamble regional tiene de por sí un error menor debido a la realización

de simulaciones con una mayor resolución, la excepción a esta situación lo constituye el pronóstico de temperatura a 2 metros sobre la región 3 en donde se ve que el ensamble regional tiene un error algo mayor al del ensamble global. Por otra parte, la variación de la magnitud de los errores entre las diferentes regiones analizadas está estrechamente vinculada a la variabilidad temporal de las variables dentro de cada una de las regiones. Dado que el valor de saturación de los errores es función de la variabilidad temporal de la variable, es de esperar que en las regiones con mayor variabilidad los errores alcancen valores más importantes (Kalnay, 2003).

Para variables como la temperatura y para el ensamble global, los mayores errores en niveles bajos se observan sobre el centro y norte de Argentina, mientras que para los niveles altos los mayores errores se encuentran cerca del extremo sur del dominio. Este comportamiento en niveles bajos probablemente obedezca a la mayor complejidad de los flujos de superficie como en el caso de la humedad específica a 2 metros, mientras que en los niveles altos los errores sean mayores en las regiones donde las advecciones son mayores debido a la mayor baroclinicidad del flujo. Este comportamiento se observa para otras variables como por ejemplo para el viento U y V. En particular, la distribución para V en 850 hPa. en el ensamble global, se muestra en la Figura 5.7. En esta Figura se encuentra que los máximos errores se observan sobre la región norte y centro de Argentina, extendiéndose también hacia Paraguay y el sur de Bolivia. Esta región coincide con la porción sur de la corriente en chorro en capas bajas durante la época estival (Salio et al. 2002) y está asociado a un máximo relativo de variabilidad en la circulación de niveles bajos (Liebmann et al. 2004). No obstante, este error, responde también a un error sistemático que tiende a subestimar con el tiempo la intensidad de la corriente en chorro en niveles bajos (por ejemplo los pronósticos a 48 horas sistemáticamente pronostican vientos en 850 menos intensos que los pronósticos a 24 horas). Como el error es mayor en las situaciones en las que se encuentra presente la corriente en chorro en dicha región, esta componente sistemática del error no puede ser removida totalmente restándole la media del error y deja una señal que se encuentra aún luego de haber sustraído dicha media. Este error también está presente aunque en menor magnitud en el ensamble regional.

De los resultados expuestos podemos ver que la media del ensamble permite reducir los errores en el pronóstico, (principalmente a 48 horas); tanto en la región

tropical como en la zona más baroclínica. Las mayores reducciones se encontraron para U, V, T y Q, mientras que el campo de geopotencial y presión a nivel del mar fueron los menos afectados. Este resultado es particularmente interesante para el pronóstico a corto plazo de las variables de superficie, donde se pudo demostrar utilizando datos de estaciones, que la media del ensamble produce una mejora en los pronósticos. También es importante destacar que los resultados obtenidos con los datos de superficie fueron muy similares a los obtenidos con los JRA-25 lo cual refuerza los resultados obtenidos a partir de este conjunto de datos.

La reducción del error en el ensamble regional es del mismo orden que la encontrada para el modelo global. Esto implica de alguna manera que en lo que respecta a la reducción del error, el ensamble regional no es más eficiente que el global, pero la utilización del modelo regional permite mejorar algunos aspectos del pronóstico reduciendo el error en algunas variables. El ensamble regional, por otra parte, tiene la ventaja de aplicar el mismo filtro dinámico a aquellas componentes de escalas más pequeñas que el modelo global no puede resolver.

La Figura 5.8 muestra la mejora en el error en el ensamble regional calculados a partir de los datos GTS teniendo en cuenta la resolución original de 40 km. Los errores en la región 3 no pudieron ser computados, ya que la misma se encuentra mayormente sobre el océano donde no hay disponibilidad de datos de superficie. En la Figura 5.8, se puede apreciar que la reducción del error por parte de la media del ensamble se mantiene cuando consideramos los pronósticos en su resolución original. Un aspecto importante que se destaca en esta figura es que los errores en las diferentes variables comparados con los datos de estaciones en alta resolución, muestran claramente errores más altos que los obtenidos cuando la verificación se hacía en baja resolución (es decir interpolando las salidas a la retícula de los JRA-25). Esto muestra que una verificación en alta resolución no es comparable a una verificación en baja resolución y por eso para comparar los ensambles globales y regionales se decidió llevar ambos conjuntos de datos a un dominio común y de preferencia al dominio que incluyera aquellas escalas que ambos conjuntos de datos estaban en condiciones de resolver.

### 5.3.2 Relación entre la dispersión del ensamble y el error en la media del ensamble.

Una primera aproximación al problema consiste en evaluar que dentro del marco teórico propuesto por Houtekamer (1993), exista una relación entre las distribuciones del error y la dispersión que se asemeje a la distribución de un ensamble ideal. Para esto, resulta útil comparar cómo son las tasas de crecimiento del error y de la dispersión del ensamble. Dentro del marco teórico planteado se puede demostrar que el EAM (definido en 5.2) es una función de dispersión (ecuación 5.7). En este caso, el EAM puede ser obtenido al promediar los errores espacial o temporalmente. Esto implica que para un ensamble perfecto, la tasa de crecimiento de la dispersión del ensamble debería estar estrechamente vinculada con la tasa de crecimiento del error en el pronóstico

$$EAM = \frac{2}{\sqrt{2\pi}} S \quad (5.7)$$

En la Figura 5.9, se muestra la evolución media de la dispersión y del error en el pronóstico, promediados sobre diferentes regiones y sobre todo el período abarcado por el experimento, como función del plazo del pronóstico, tanto para el ensamble regional como para el ensamble global. La figura muestra esta evolución para las variables de superficie (T2M, PNM y Q2M). Lo que se puede apreciar es que en general la magnitud del error es mayor que la dispersión del ensamble y tiene una tasa de crecimiento mayor lo cual se verifica para ambos ensamble. En general la tasa de crecimiento de la dispersión es menor a la óptima que es una característica observada en otros ensambles regionales y globales (Hou et al. 2001, Hamill y Colucci 1997 entre otros). Por otra parte es importante destacar el comportamiento de diferentes regiones: en la región 2 que está dentro de la zona tropical, el crecimiento de los errores en el plazo de pronóstico es muy escaso y la dispersión se comporta de una manera acorde. En esta región, los errores crecen rápidamente durante las primeras horas del pronóstico, y luego muestran un comportamiento estacionario o de lento crecimiento. Este comportamiento es consistente con una región en donde dominan inestabilidades como la convectiva que producen un rápido crecimiento del error que alcanza la saturación en plazos de tiempo relativamente cortos (Kalnay 2003). En general la calidad de los pronósticos en esta región se deteriora

rápidamente. Esto puede explicarse por qué el crecimiento de la dispersión y el error dentro del ensamble es tan pequeño a lo largo del pronóstico. En la región 3, que está en una zona donde la actividad de ondas baroclínicas es mayor (y por ende hay mayor inestabilidad en el flujo) la tasa de crecimiento de los errores y de la dispersión es sensiblemente mayor para ambos ensambles. No obstante la dispersión crece más lentamente que el error. Existen diversos motivos que pueden explicar el menor crecimiento de la dispersión frente al error del pronóstico (característica que es común a todas las regiones) por un lado el modelo no es perfecto y tiene errores debido a las parametrizaciones, la discretización, etc. La dispersión del ensamble en este caso, no está capturando los errores presentes en el modelo. Si bien el error medio fue removido durante el cálculo, existen errores en el modelo que no son sistemáticos y errores sistemáticos que no son necesariamente constantes con el tiempo (como por ejemplo el error en la componente V del viento en la región del jet de capas bajas que es más intenso en los eventos de corriente en chorro intensa, este tipo de error claramente no puede ser corregido mediante la resta del error medio). Estos errores no van a estar bien representados por un ensamble que sólo considera perturbaciones en las condiciones iniciales. Esto podría explicarse por qué, por ejemplo, la región 1 que se ubica en la región de salida de la corriente en chorro en capas bajas y es en donde los errores en diferentes variables mostraban un máximo de amplitud, los ensambles muestran valores de dispersión muy por debajo del valor del error y con un crecimiento menor al del error. El estudio de los errores sistemáticos demanda un gran volumen de información entre pronósticos y observaciones y un estudio muy pormenorizado de la forma que adoptan los errores sistemáticos en las diferentes variables. Por eso, una alternativa más eficiente para estudiar esta limitación es la utilización de un super ensamble de modelos donde se utilicen diferentes condiciones iniciales y diferentes modelos que tengan asociados errores sistemáticos diferentes. De esta forma, el error sistemático, siempre y cuando no sea compartido por todos los miembros del ensamble, pasa a formar parte de la incertidumbre asociada al pronóstico. (Hou et al. 2001, Kalnay 2003)

Existen otros factores a tener en cuenta que podrían debilitar el crecimiento de la dispersión del ensamble contra el crecimiento de los errores. El método de los vectores criados está diseñado para obtener perturbaciones eficientes que sean sensibles a las inestabilidades del flujo y que muestren un rápido crecimiento durante el pronóstico. En

particular los parámetros elegidos, permiten capturar el crecimiento debido a las inestabilidades de la escala sinóptica. Es por eso que es esperable encontrar un mejor desempeño de este método en las regiones donde la inestabilidad del flujo es importante en la escala sinóptica. Esto se puede ver en la Figura 5.9, comparando la región 3 que se ubica donde la actividad de ondas baroclínicas es mayor y la región 2 que se ubica en la zona tropical. La medida en la que los vectores criados permitan capturar las inestabilidades del flujo y crear perturbaciones que crezcan de acuerdo a las mismas va a determinar también la relación entre el crecimiento de la dispersión y el crecimiento del error. En este punto vale la pena retomar la Figura 4.7 del Capítulo 4, en donde claramente se ve que el efecto de utilizar perturbaciones que no estén optimizadas desde el punto de vista de las inestabilidades del flujo puede afectar el crecimiento de la dispersión a lo largo del pronóstico.

Con respecto a la relación entre el crecimiento de la dispersión en el ensamble regional y en el global, a priori, no es del todo claro que el modelo regional pueda reaccionar a las perturbaciones en la condición inicial de la misma forma que lo hace un modelo global en donde el crecimiento del error puede darse en escalas de movimiento mayores a las permitidas por un modelo de área limitada. En particular, el comportamiento de la dispersión en un ensamble regional, va a estar condicionado también por la forma en la que se perturben las condiciones de borde. En el presente estudio, las condiciones de borde están perturbadas de manera consistente con las perturbaciones iniciales, ya que los pronósticos globales inicializados a partir de las condiciones perturbadas son los que además dan condiciones de borde a cada uno de los miembros del ensamble regional. Esto hace que la divergencia entre los miembros del ensamble esté forzada en los bordes por la divergencia de las soluciones del modelo global. En la Figura 5.9 se observa que si bien el error en el modelo regional es menor, también lo es su dispersión.

La Figura 5.10, muestra los resultados para todo el dominio considerado, comparando la tasa de crecimiento del error entre el pronóstico a 12 horas y el pronóstico a 48 horas con la tasa de crecimiento de la dispersión para el ensamble global. Las tasas de crecimiento se calculan de acuerdo con la Ecuación 5.8, donde el incremento en  $S$  está normalizado por el error a 12 horas de pronóstico para que los resultados puedan ser comparados con la situación teórica planteada por la ecuación 5.7.

$$\Delta EAM = \frac{AEM_{48} - AEM_{12}}{AEM_{12}} \quad (5.8)$$

$$\Delta S = \frac{S_{48} - S_{12}}{AEM_{12}}$$

En la Figura 5.10 se puede apreciar que el crecimiento del error es máximo sobre el océano en la región sudeste del dominio. Este máximo probablemente esté vinculado estrechamente con una mayor actividad baroclínica en la zona y con la menor disponibilidad de observaciones (ver por ejemplo Figura 4.3 a), por otra parte sobre el continente existen máximos sobre el norte/centro de Argentina y en el centro de Brasil alrededor de 15°S y 50°O. En cuanto a la dispersión, también muestra un máximo en la porción sudeste del dominio, acompañando el mayor crecimiento del error en esa zona aunque no muestra un máximo tan marcado sobre el norte de Argentina o sobre Brasil. En general podemos concluir que existe un acuerdo importante entre las regiones donde el crecimiento de la dispersión es mayor y las regiones en donde el crecimiento del error es mayor indicando que el ensamble global puede capturar correctamente la distribución espacial del crecimiento de los errores en el pronóstico en la escala sinóptica. Esto sugiere además que gran parte de esa distribución del crecimiento de los errores obedece a la distribución de las regiones que son más sensibles al crecimiento de la incertidumbre en la condición inicial que es el tipo de incertidumbre que estamos considerando al utilizar los vectores criados para generar el ensamble. El acuerdo es menor para aquellas regiones donde la fuente de incertidumbre podría estar dominada por los errores en el modelo como por ejemplo sucede en la región norte y centro de Argentina. Como en los casos anteriores, este mismo análisis se repitió para otras variables y otros niveles verticales pero los resultados no se muestran dado que no se considera que agreguen información a la discusión.

La Figura 5.11, muestra los resultados obtenidos para el ensamble regional. En este caso como se había discutido al analizar el comportamiento de las diferentes regiones, se ve que el crecimiento del error es menor que en el modelo global, con excepción de lo que se observa sobre Paraguay. Por otra parte el crecimiento de la dispersión también se redujo con respecto al modelo global, pero en este caso, al estar normalizado por el error del modelo regional, aparece como si la tasa de crecimiento hubiera aumentado en

algunas regiones respecto a lo que se observa en el modelo global, lo cual no contradice los resultados discutidos previamente.

Lo presentado hasta ahora no demuestra la existencia de una relación entre la dispersión y el error. Pero dado que nuestro ensamble no se comporta como un ensamble ideal es necesario determinar en qué medida los cambios en la dispersión están vinculados con los errores en el pronóstico. Para verificar la existencia de la relación entre dispersión y error se consideraron las 4 categorías de dispersión definidas previamente y para cada categoría se computó la desviación estándar de los errores y la pendiente de la recta de regresión entre el valor central del intervalo de dispersión y el valor de dicha desviación de los errores para cada intervalo. Para el presente caso, esta correlación se calcula utilizando 4 puntos que es el número de categorías disponibles. La Figura 5.12, muestra los coeficientes de correlación obtenidos y la pendiente de las rectas de regresión para el ensamble global y para Q2M. En la misma se puede observar que hay cierta consistencia en la mayor parte del dominio, sin embargo, regiones como el sur de Brasil y el norte de Argentina, muestran una clara falta de consistencia, indicando poca sensibilidad de la distribución del error a los cambios en la dispersión. Por otro lado, se indican los valores de la pendiente de la recta de regresión. Estos valores son en todos los casos muy inferiores a 1 indicando que la sensibilidad de la desviación estándar del error a los cambios en la dispersión está por debajo de lo que se espera para un ensamble ideal (por ejemplo cuando la dispersión es baja se producen más errores grandes que lo esperado y cuando la dispersión es alta hay más ocurrencia de errores chicos que lo esperado, con lo cual la pendiente de la recta disminuye). Otro aspecto destacable en la Figura 5.12, son los cambios con el plazo de pronóstico. La consistencia estadística aumenta a medida que se extiende el plazo de pronóstico. Esto confirma resultados obtenidos previamente que muestran que la relación entre la dispersión y el error, aumenta con el plazo de pronóstico en los primeros 7 días de pronóstico (Whitaker y Loughé 1998). Los resultados obtenidos para el ensamble regional, son bastante similares: también se observa una notoria falta de relación entre dispersión y error sobre la región norte de Argentina (no se muestra la figura)

La Figura 5.13, muestra los resultados obtenidos para la presión a nivel del mar; estos difieren del caso anterior principalmente en que para la región tropical de Sudamérica, no existe sensibilidad de la distribución del error al valor de la dispersión.

Esto ilustra el hecho de que diferentes variables muestran comportamientos distintos y que en general las conclusiones obtenidas para una variable no son extensibles directamente a otras variables. Para las variables de superficie en particular, se trabajó tanto con los JRA-25 como con los datos de superficie. En las regiones donde la red de estaciones hizo posible la comparación entre ambos resultados, los mismos fueron muy similares. Es importante destacar que no parece existir una relación demasiado estrecha entre las regiones donde la consistencia estadística es mayor y las regiones donde la reducción del error por parte de la media del ensamble es mayor (comparar por ejemplo las Figuras 5.2 y 5.5).

A continuación se muestran los resultados orientados a mostrar la relación existente entre dispersión y EA desde un punto de vista probabilístico como el discutido en la sección 5.2. La Figura 5.14, muestra ejemplos de cómo varía  $P_{cs}$  como función de la categoría de la dispersión para las regiones 1, 2 y 3 discutidas previamente y para el ensamble global y regional. Todos los puntos de retícula presentes en cada una de las regiones fueron utilizados para generar las muestras a partir de las cuales se calcularon las probabilidades que se muestran en las figuras. Como se puede ver en la figura, las regiones 1 y 3, muestran un cierto grado de relación entre  $P_{cs}$  y la dispersión del ensamble. Esta relación es más fuerte en la región 3 que es la que está ubicada más al sur. La relación al igual que como se observó al analizar la consistencia estadística entre ambas distribuciones muestra un incremento a medida que aumenta el plazo de pronóstico (en particular para la región 1 recién a las 48 horas de plazo parece observarse una relación más estrecha entre ambas variables). En la región 2, por otra parte con un comportamiento más tropical, no se puede apreciar relación alguna entre la dispersión y la función de distribución de probabilidades del error. Como bien señalan Gritit y Mass (2007), aún para un ensamble ideal, la relación entre ambas variables es débil, aunque podría ser de gran utilidad para un amplio rango de usuarios dependiendo de los modelos de toma de decisión utilizados. Los resultados observados para el ensamble regional son similares a los que se encuentran para el ensamble global demostrando que la efectividad de determinar la incertidumbre en la previsión en las escalas más grandes es la misma para ambos ensambles.

En el caso de la Figura 5.14, (panel inferior derecho) se puede apreciar además que existe una relación más fuerte para los umbrales de error más altos –menos frecuentes- lo

cual coincide con lo que encontraron diversos autores en estudios previos. Por ejemplo, para el umbral más alto de error considerado, la diferencia de  $P_{cs}$  entre la categoría de dispersión 1 y la categoría 4 es de aproximadamente 20 %, es decir que es 20% más probable que ocurra un error grande si la dispersión es grande que si la dispersión es pequeña. Mientras que para el umbral más bajo la diferencia observada es menor.

Para obtener una distribución espacial de la intensidad de la relación entre dispersión y error, se calculó el IR y el IRSS para diferentes variables de acuerdo con la metodología propuesta en la sección 5.2. La Figura 5.15, muestra la distribución del IRSS para la presión a nivel del mar, para diferentes umbrales de error y para diferentes plazos de pronóstico. Se puede apreciar que en general, para esta variable, la mayor relación se encuentra sobre la porción sur del dominio, siendo muy bajo el valor del IRSS para la región tropical. Esta distribución no obedece a los umbrales de error seleccionados, ya que en cada punto de retícula, los umbrales fueron tomados a partir de los límites de los intervalos cuartiles de la distribución local del error de forma tal que los resultados sean comparables. La Figura 5.5, muestra claramente que la magnitud de los errores en la presión a nivel del mar aumenta a medida que nos acercamos al polo, por ende es necesario aplicar una estandarización para poder obtener umbrales que permitan comparar los resultados de diferentes regiones. Confirmando lo que se había visto en la Figura 5.14, para diferentes regiones, la mayor relación se observa para el umbral de error menos probable (es decir el umbral más alto). En la Figura 5.15 se incluyen también contornos que indican la variabilidad temporal de la dispersión en cada punto. Varios autores (Houtekamer 1993, entre otros) muestran que la relación entre el error y la dispersión, en el ensamble ideal, aumenta cuando aumenta la variabilidad temporal de la dispersión, para un ensamble ideal. En el presente caso se observa cierta correspondencia entre las regiones donde la variabilidad de la dispersión es mayor y las zonas donde la relación entre ambas variables es más fuerte. En general, la distribución espacial del índice para la PNM es bastante similar a la distribución que se observa en el geopotencial en niveles bajos (925 y 850 hPa., no se muestra la figura). De acuerdo con la determinación de los valores de significancia para el IRSS realizada en la sección 5.2, valores superiores a 0.21 son significativos al 95 %, lo cual nos da para este caso, pocas regiones con valores de relación significativos a 12 y 24 horas de pronóstico, siendo significativa la relación a 48 horas sobre una extensa área sobre el centro y sudeste del dominio. Al

analizar los valores de IRSS para la PNM para el ensamble regional (no se muestra la figura), se obtuvieron básicamente la misma distribución espacial, y la misma dependencia del IRSS con el tiempo de pronóstico y los umbrales de error considerados, la única diferencia fue que los valores de IRSS fueron algo menores a los obtenidos en el ensamble global, tal como lo sugiere la Figura 5.14.

La Figura 5.16, muestra la distribución espacial del IRSS para la humedad a 2 metros para el ensamble global. Como se veía para el caso de la consistencia estadística, las distribuciones asociadas a distintas variables pueden ser muy diferentes. En el caso de la humedad a dos metros, se observa que la región tropical muestra un relativamente alto grado de relación entre la dispersión y el error, sobre todo para los errores grandes, acompañado por un máximo de variabilidad temporal de la dispersión en dicha región. Por otra parte se observa otro máximo en la porción sur de la región considerada, a la par de que la relación es casi nula para la región norte y centro de Argentina. Esto refuerza la idea de que la utilización de esta herramienta para el pronóstico del error debe ser evaluada en forma individual por variables y regiones de acuerdo con las necesidades de los potenciales usuarios de la información.

La Figura 5.17, muestra la distribución del IRSS calculado para el ensamble regional en la humedad específica a 2 metros. Al igual que sucede con la presión, en este caso, los valores de IRSS son más bajos que los obtenidos en el modelo global. En particular sobre la región norte del dominio y sobre el Atlántico sur, donde se observa la existencia de alguna relación entre la dispersión y el error, el ensamble regional, muestra una menor variabilidad temporal de la dispersión que de acuerdo con el modelo teórico propuesto por Houthekamer (1993) podría explicar el debilitamiento de la relación.

Con respecto a otras variables, la temperatura presenta un patrón de distribución del índice que se asemeja más al de la humedad específica (no se muestra la figura), con un máximo en la región tropical y otro en la porción sur del dominio. Sin embargo el nivel de relación entre la dispersión y el error es algo menor para la humedad específica. Esto también es válido para los niveles más bajos considerados en este trabajo (925 y 850 hPa).

En general los niveles bajos muestran una distribución espacial del IRSS que no difiere mucho de la encontrada para las variables de superficie pero observándose valores de IRSS superiores a los que se encuentran para dichas variables en superficie, un

ejemplo de esto se puede observar para la humedad específica en el nivel de 925 hPa (Figura 5.18), un comportamiento similar se encuentra en la temperatura. En niveles medios y altos esto no sigue siendo válido. En algunos casos como en la temperatura, los niveles altos muestran una menor concordancia entre la dispersión y el error, mientras que para otras variables como la componente V del viento, la relación entre ambas variables se incrementa, sobre todo en la región de mayor actividad baroclínica. Un caso extremo lo constituye la humedad específica en niveles altos que muestra una relación muy estrecha entre dispersión y error. La relación es incluso notoria en los campos diarios de ambas variables. En este caso, el hecho de que la humedad específica sea tan baja en 500 hPa, puede llevar a una relación estrecha entre las regiones donde el valor es relativamente alto (y por ende el error puede serlo) y las regiones donde el valor de  $q$  es tan pequeño que difícilmente errores en la advección horizontal podrían generar errores importantes.

Para comparar los resultados obtenidos con la forma tradicional de evaluar la relación entre dispersión y error utilizando el coeficiente de correlación lineal, se calculó para todas las variables y los niveles, la relación entre dispersión y error utilizando esta metodología. Las muestras a partir de las cuales se calculó el valor del coeficiente de correlación fueron exactamente las mismas que se utilizaron para calcular IRSS. La Figura 5.19, muestra un ejemplo de los resultados obtenidos para la humedad específica a 2 metros. En general los valores de correlación lineal son bajos (inferiores a 0.3) similares a los encontrados por (Whitaker y Loughé 1998) para latitudes semejantes en el hemisferio norte. No obstante, en este caso es notoria la aparición de valores de correlación casi tan grandes, en zonas subtropicales, como los encontrados en la porción sur del dominio.

Si bien los resultados obtenidos por ambas metodologías son similares en cuanto a los patrones espaciales obtenidos y en cuanto a la dependencia de la relación con el plazo de pronóstico, la utilización del IRSS permitió discriminar los cambios en la resolución del pronóstico dependiendo de la magnitud del error que se quiere pronosticar, mostrando que existe una mayor resolución en la relación entre los errores más grandes (con menos de 25% de probabilidad de ocurrencia) y la dispersión del pronóstico. Por otra parte, el IRSS no asume que exista una relación de tipo lineal entre ambas variables.

## 5.4 Conclusiones

Los resultados presentados en este capítulo, muestran que ambos ensambles permiten reducir el error de los pronósticos de diferentes variables (en particular de las variables de superficie) a través de la utilización de la media del ensamble como predictor. En particular el ensamble regional, no mostró ser más eficiente en la reducción porcentual del error (comparando media con su control), pero sí mostró tener menores errores debido a la mayor resolución con la que trabaja. En esta tesis, tanto el ensamble de alta resolución como el de baja resolución, fueron verificados en baja resolución para garantizar que pudieran ser evaluados contra una fuente de datos que contenga información de las mismas escalas espaciales. En este marco se pudo verificar, que el ensamble regional pudo reproducir los aspectos de mayor escala de la circulación con un error menor al presente en el ensamble global.

Tanto el ensamble global como el ensamble regional, presentaron errores sistemáticos en la circulación en niveles bajos. En particular se observó una subestimación de la intensidad de la corriente en chorro en capas bajas que es más notoria en el ensamble global, aunque también esta presente en el ensamble regional. Dado que dicha circulación esta vinculada al transporte de aire cálido y húmedo desde la región Amazónica hacia el centro y norte de Argentina, su subestimación puede estar vinculada con errores en otras variables como por ejemplo en la humedad. De hecho, en el presente caso, se observó un máximo de error en la humedad específica sobre la región norte y centro de Argentina que coincide con la región de salida de la corriente en chorro en capas bajas.

Es importante destacar que en general, dado que se están considerando pronósticos a corto plazo, la diferencia que se observa visualmente entre la media del ensamble y el pronóstico control es escasa, en el sentido de que en la mayoría de las situaciones estas diferencias podrían no tener un impacto significativo en el análisis de los sistemas de tiempo predominantes (por ejemplo en determinar posición e intensidad de las cuñas y vaguadas), pero sí presenta ventajas para aquellas tareas que requieran un pronóstico cuantitativo del valor de alguna variable en particular (por ejemplo pronóstico de temperaturas diarias). Es importante destacar que este análisis no incluye a la precipitación, variable cuyo comportamiento es analizado en los Capítulos 3 y 6 del presente trabajo de tesis.

Por otro lado se evaluó también la capacidad de ambos ensambles para proveer información sobre el grado de incertidumbre de los pronósticos en el corto plazo. La metodología aplicada permitió evaluar el potencial del uso de la dispersión para el pronóstico probabilístico de errores en un marco que permite la comparación de diferentes variables y regiones. Las relaciones encontradas entre la dispersión y el error fueron similares en ambos ensambles, en particular el crecimiento de la dispersión en el ensamble regional, resultó ser similar a la del ensamble global cuando se consideran escalas espaciales similares. Este resultado es importante ya que indica que el forzado en las condiciones de borde es adecuado para garantizar un crecimiento de las perturbaciones dentro del dominio regional. No obstante, los valores de IRSS resultaron algo menores en el dominio regional para las variables estudiadas.

La distribución espacial del IRSS para diferentes variables, revela que en algunas como la PNM, la mayor relación entre el error y la dispersión del ensamble se encuentra para latitudes medias y altas, mientras que en latitudes bajas es prácticamente nula. Esto en principio parecería sugerir que la técnica utilizada no es adecuada para representar la incertidumbre de los pronósticos en la región tropical de Sudamérica. No obstante, al analizar otras variables como Q2M o T2M, una de las zonas con mayor relación entre dispersión y error se encuentra justamente en la porción norte del dominio, es decir que la técnica en principio es sensible a los errores que se producen en algunas variables aún en la región tropical. Más bien, la falta de relación entre error y dispersión para la PNM en la región tropical tiene más que ver con la falta de variabilidad temporal de la dispersión sobre dicha región para la PNM (Houtekamer, 1993).

Un aspecto importante a tener en cuenta en el análisis de estos resultados es que trabajos anteriores han reportado una importante variabilidad interanual en las regiones que muestran una mayor relación entre dispersión y error para el hemisferio norte (Whitaker y Lough 1998) por tanto los patrones obtenidos no necesariamente son representativos de todos los veranos en nuestra región, con lo cual es necesario extender en el futuro este tipo de estudios a un período más largo de tiempo.

Los resultados obtenidos mostraron que en general la relación entre la dispersión del ensamble y el error es débil dentro de las primeras 48 horas de pronóstico para las variables de superficie. Una de los posibles factores que podrían estar debilitando esta relación son los errores del modelo. Dado que tanto el ensamble global como el regional

están basados en la perturbación de sus condiciones iniciales pero no incluyen los errores del modelo, la dispersión del ensamble no estaría respondiendo a los mismos. Es por eso que en todas aquellas situaciones donde los errores del modelo dominan el error del pronóstico, la dispersión del ensamble no resultará un buen predictor de la incertidumbre. Esta limitación se puede resolver o bien introduciendo un ensamble de varios modelos (como el SMES descrito en el Capítulo 3) o bien incluyendo un estudio pormenorizado de los errores del pronóstico. Este estudio podría basarse por ejemplo en el cálculo del IRSS para medir la relación existente entre los errores y el valor de diferentes variables. Esto revelaría la distribución espacial de los errores sistemáticos asumiendo que los mismos podrían tener una relación no-lineal con las variables del modelo. El análisis de la relación entre los errores del pronóstico y las variables del modelo es parte de un trabajo en preparación.

No obstante la debilidad de la relación entre la dispersión y el error en las variables de superficie, existe una mayor relación entre el error y la dispersión del ensamble cuando se analiza la distribución de las variables en niveles bajos, medios y altos. En algunos casos como por ejemplo para la temperatura y la humedad específica se obtuvieron mejores resultados en niveles bajos que lo que se observó con las variables de superficie, obteniéndose valores de IRSS que marcan la existencia de una relación más estrecha entre el error y la dispersión. En particular esto es interesante para el pronóstico de precipitación que puede verse afectado por las fluctuaciones de T y Q en niveles bajos. La aplicación de este sistema de ensamble al pronóstico probabilístico será evaluada en el Capítulo 6. Por otra parte también podría afectar directamente a medidas de la inestabilidad de la atmósfera como el CAPE, siendo un importante factor a tener en cuenta en el posible desarrollo de pronósticos probabilísticos de ocurrencia de convección y eventualmente condiciones favorables para la ocurrencia de tiempo severo. Para eso también es fundamental trabajar para corregir los errores sistemáticos del modelo dado que en regiones claves como el centro y norte de Argentina, la subestimación de los valores de Q y de la componente meridional del viento en niveles bajos limitan sensiblemente la información que se puede obtener a partir del ensamble.

Por otra parte, a la hora de estimar la probabilidad de error para un día determinado de forma tal que pueda ser útil para los usuarios finales del pronóstico, se debe tener en cuenta aquellos errores vinculados a los errores del modelo y aquellos que

tengan que ver con la inestabilidad del flujo. En algunos casos los dos factores estarán muy relacionados entre si, pero en otros, sobre todo en aquellos vinculados a los errores sistemáticos, la dependencia con el valor de las variables del modelo puede aportar más información que la dispersión del ensamble. Las Figuras 5.20 y 5.21 proveen un ejemplo del tipo de pronóstico operativo de incertidumbre que se podría generar a partir de un ensamble regional. En este caso, se presentan los valores del error estandarizado de la media del ensamble, la dispersión estandarizada del ensamble, y la media estandarizada del ensamble para el día 24 de enero de 2003 a las 00 UTC, en este caso, hubo gran incertidumbre y errores en el campo de PNM (Figura 5.20) vinculados a dos sistemas de baja presión sobre el Atlántico y el Pacífico Sur. Mientras que en el campo de humedad, los errores más importantes sobre la región continental están vinculados a un frente frío ubicado sobre el norte de Argentina, en esta región se observa una buena correspondencia entre el error y la dispersión del ensamble. En este caso, tanto los errores como la dispersión y el valor medio del ensamble están estandarizados con respecto al punto de retícula donde fueron calculados, y al error se le sustrajo el valor medio.

La mayor parte de los resultados del presente capítulo fueron obtenidos en una resolución menor a la provista por el ensamble regional, para facilitar la comparación directa con el ensamble global. No obstante es necesaria una comparación utilizando datos con mayor resolución para poder determinar el desempeño del ensamble regional en su resolución original. En este sentido una de las principales limitaciones es que no se cuenta con una red de observaciones lo suficientemente densa. No obstante recientemente se generaron para la región análisis de alta resolución que asimilaron los datos de la red operativa y los provenientes de la red del experimento SALLJEX (García Skabar, 2007) que permitirían repetir algunos experimentos utilizando la máxima resolución del ensamble regional. Esto también incrementaría el tamaño de las muestras permitiendo la utilización de un mayor número de categorías de dispersión.

## Capítulo 6

# Calibración de pronósticos probabilísticos de precipitación durante la estación cálida 2002-2003.

### 6.1 Introducción.

En el presente capítulo, se exploran los resultados obtenidos con diferentes formas de generar pronósticos probabilísticos de precipitación para el ensamble regional basado en la técnica de los vectores criados para la estación cálida 2002-2003. Los pronósticos probabilísticos son calibrados retomando la metodología descrita en el Capítulo 3, pero ensayando modificaciones que permitan mejorar los resultados obtenidos. Por otra parte también se utilizan los datos CMORPH en reemplazo de los datos pluviométricos como alternativa para la calibración.

### 6.2 Metodología.

Los pronósticos de precipitación obtenidos para los diferentes miembros del ensamble y para el pronóstico control, son calibrados de diferentes maneras para obtener pronósticos probabilísticos de precipitación para la región de Sudamérica. Dichos pronósticos se generan para 3 áreas por separado: región norte de Argentina, ubicada entre 35° y 20° de latitud sur y 68° y 58° de longitud oeste (Región 1), región sudeste de Brasil ubicada entre 35° y 20° de latitud sur y 58° y 40° de longitud oeste (Región 2) y región norte ubicada entre 20° y 8° de latitud sur y 68° y 45° de longitud oeste (Región

3). El objetivo de esta subdivisión es evaluar en forma independiente las propiedades del pronóstico probabilístico sobre diferentes regímenes de precipitación. Las regiones del norte de Argentina y sudeste de Brasil, tienen una mayor influencia de las perturbaciones baroclínicas, si bien en esta época del año el ingreso de frentes no es tan frecuente como durante el invierno. Por otra parte, en el sudeste de Brasil hay una importante influencia de la zona de convergencia del Atlántico Sur (ZCAS).

Para obtener diferentes variantes del pronóstico probabilístico de lluvia, se ensayaron diversas calibraciones que fueron obtenidas como se detalla a continuación.

- Pronóstico no calibrado: La probabilidad obtenida del ensamble utilizando la metodología de voto democrático (Capítulo 3).
- Pronóstico probabilístico calibrado según HC98 (**Calibración 1**): Calibración en función del histograma de rangos como se describió en el Capítulo 3.
- Pronóstico probabilístico calibrado en base al histograma de rangos con tratamiento diferencial de los ceros (**Calibración 2**): En este caso se introduce una modificación en el algoritmo propuesto por HC98. Dichos autores proponen que cuando la verificación sea 0 y todos los miembros del ensamble estén por encima de dicho valor se incorpore dicho caso en el rango 1 del histograma de rangos. Por otra parte si algunos miembros del ensamble pronostican 0 y la precipitación observada es 0, entonces el rango se determina aleatoriamente. El número de veces que la verificación es 0 es importante y dado que el valor 0 en precipitación es muy frecuente en los regímenes analizados, se ha considerado aquí que este valor requiere un tratamiento especial. Como se describirá en los resultados, la inclusión de los 0 de precipitación como en HC98, junto con la hipótesis de que la probabilidad se distribuye linealmente en el primer rango del histograma de rangos, puede conducir a la sobreestimación de la probabilidad sobre todo para valores de probabilidad elevados. Para corregir esto se propone una sencilla modificación a la metodología: cada vez que la verificación es 0, esta no se asigna a ningún intervalo del histograma de rangos, sino que se utiliza para calcular la frecuencia observada del valor 0 en cada categoría de dispersión utilizada. Luego, al dividir cada intervalo del histograma por el número total de pronósticos, se tienen en cuenta aquellos en

- los que la precipitación observada fue 0, esto produce que la suma de los rangos del histograma no sea 1, sino que ahora sea  $1-P_0$ , donde  $P_0$  es la probabilidad de que la verificación sea 0 para cada una de las categorías de dispersión. La idea es que esto asigne probabilidades más realistas en cada caso reduciendo la sobreestimación para los valores altos de probabilidad.
- Pronóstico probabilístico calibrado en base al histograma de rangos incorporando categorías basadas en los errores sistemáticos del modelo (**Calibración 3**): En este experimento se utiliza una metodología basada en el histograma de rangos (sin tratamiento diferencial de los ceros) pero además de dividir los pronósticos en 3 categorías según el valor de dispersión del ensamble como se hizo en el Capítulo 3 y como lo proponen HC98, se agregó una nueva categorización. Para cada categoría de dispersión, se definieron 3 categorías de error sistemático de la precipitación. El error sistemático de precipitación fue definido como la diferencia entre la precipitación media pronosticada y la precipitación media observada en el período de entrenamiento utilizado. Los puntos del ensamble dentro de cada una de las regiones son divididos en puntos con error sistemático bajo, medio y alto. Los umbrales para la definición de estas categorías se definen a partir de los límites de los terciles de la distribución de errores sistemáticos, es decir que como ocurre con las categorías de dispersión, se adaptan a cada región y a cada fecha en particular. Dado que como se describe en Hamill (2001), los errores sistemáticos del modelo afectan la forma del histograma de rangos, es natural pensar que obtener un histograma de rangos único a partir de pronósticos y observaciones que provienen de regiones con diferentes errores sistemáticos puede conducir a una degradación del pronóstico calibrado. Al introducir esta categorización adicional en base al error sistemático del modelo, se intenta hacer una subregionalización de la calibración agrupando puntos que tengan un comportamiento similar en cuanto a los errores sistemáticos. Es importante destacar que esta alternativa requiere un mayor número de datos de precipitación disponibles para poder ser implementada, dado que requiere obtener 9 histogramas de rangos en lugar de los 3 requeridos por la implementación de la metodología que se hizo en el Capítulo 3. Las observaciones pluviométricas del experimento SALLJEX y las de la red ANA en Brasil permitieron evaluar esta metodología durante la estación cálida 2002-2003.

- Pronóstico probabilístico derivado de la media del ensamble como se describe en el Capítulo 3 (**Calibración 4**).
- Pronóstico probabilístico derivado del control (**Calibración 5**): Como en el caso anterior, pero computando la probabilidad a partir del pronóstico control (no se utilizan los miembros del ensamble) (Ver Capítulo 3).
- Calibración regionalizada (**calibración 6**): Este caso fue diseñado para explorar las ventajas que podrían surgir de la utilización de un conjunto de datos como los CMORPH sobre la región. La idea consiste en utilizar un algoritmo de calibración (en este caso se utiliza la calibración 4 basada en la media del ensamble) y calibrar los pronósticos de cada punto de retícula utilizando un entorno del punto. El entorno es variable y puede ser ajustado para obtener mejores resultados siempre y cuando se cuente con un número suficiente de observaciones para calcular las estadísticas necesarias para el proceso de calibración. En este caso se utilizaron entornos rectangulares con un tamaño de 5 grados alrededor de cada punto de retícula, lo cual permite contar con un número suficiente de pares observación-pronóstico para realizar la calibración.

En base al trabajo de Eckel y Walters (1998) se verificó también el impacto de utilizar el doble de categorías de dispersión (6 categorías en total) que las utilizadas en el Capítulo 3, esta implementación no mostró mejoras significativas y por lo tanto no es discutida en los resultados.

Todas estas calibraciones fueron implementadas de manera dinámica como se describe en el Capítulo 3, utilizando un tiempo de entrenamiento de 15 días. Debido a esto, el plazo de verificación se extiende desde 1 de enero al 15 de febrero de 2003, dado que los pronósticos de precipitación utilizados están disponibles desde el 15 de diciembre de 2002. Todas las calibraciones fueron aplicadas a los pronósticos a 24, 36 y 48 horas (los pronósticos a 24 y 48 horas corresponden al ciclo de las 12 UTC y el de 36 horas al ciclo de las 00 UTC)

Como se describió previamente, algunas calibraciones en particular requieren un mayor volumen de datos para obtener histogramas de rangos estables. Si bien durante la estación cálida 2002-2003 había un gran número de observaciones pluviométricas disponibles, la red operativa de precipitación no es suficiente para poder implementar

dichas alternativas con períodos de entrenamiento cortos como el propuesto en este trabajo de tesis. Es por eso que además de utilizar datos pluviométricos se analiza el impacto que tiene la utilización de estimaciones de precipitación CMORPH en el proceso de calibración. Si bien dichos datos presentan errores sistemáticos sobre la región, estos no son significativos para umbrales medios y bajos de precipitación (Ruiz, 2008 enviado), con la ventaja de que los mismos presentan una excelente cobertura espacial y temporal que podría resultar muy útil para la calibración de pronósticos de precipitación.

## **6.3 Resultados.**

### **6.3.1 Impacto de las modificaciones introducidas a la calibración del ensamble en los histogramas de rangos.**

En primer lugar, resulta interesante analizar las diferencias que surgen en los histogramas de rangos a partir de las modificaciones introducidas a la metodología de calibración. La Figura 6.1, muestra los histogramas de rangos correspondientes al experimento en donde se introducen categorías en base al error sistemático del modelo. Con respecto a los errores sistemáticos del modelo, el comportamiento de los histogramas de rangos está de acuerdo con lo esperado dado que en los casos en los que el error sistemático es bajo (alto) aumenta (disminuye) la frecuencia de casos en los que la verificación está por encima del máximo miembro del ensamble y disminuye (aumenta) la frecuencia de casos en los que la verificación está por debajo del mínimo miembro del ensamble. Los demás rangos se modifican de manera acorde, pero el impacto en sus respectivas frecuencias es menor. El impacto crece a medida que aumenta la dispersión del ensamble. Para los casos de dispersión alta, las diferencias en las frecuencias de los intervalos extremos del intervalo de rangos pueden alcanzar el 20%, es decir que la probabilidad asociada al caso en el que el máximo miembro del ensamble coincide con el umbral podría ser de 70% o de 90% dependiendo de si el punto de retícula en cuestión tiene un error sistemático alto o bajo respectivamente. Si bien todas las regiones muestran un comportamiento similar con respecto a lo antes descrito, para la región norte de Argentina, la mayor sensibilidad a los errores del modelo se encuentra en la frecuencia del máximo intervalo del histograma. Mientras que en las otras 2 regiones la mayor

sensibilidad se observa en el primer intervalo del histograma. Esto puede obedecer a que el norte de Argentina está caracterizado por un déficit de precipitación pronosticada lo cual incrementa la frecuencia de veces en las que la verificación supera al valor pronosticado por todos los miembros del ensamble en relación a la frecuencia con que la verificación está por debajo de los valores pronosticados por todos los miembros del ensamble. Por el contrario, en las restantes 2 regiones, los errores sistemáticos de precipitación muestran una sobreestimación de los valores de precipitación, lo cual hace que la frecuencia más sensible a los cambios en el error sistemático sea la frecuencia de veces en las que la precipitación está por debajo de los valores pronosticados por todos los miembros del ensamble (Figura 6.2).

En la Figura 6.2, se puede apreciar además el patrón espacial de los errores de la media del ensamble para los pronósticos de precipitación a 48 horas. La precipitación es fuertemente subestimada sobre el centro y norte de Argentina, mientras que en la región de la ZCAS se observa una clara sobreestimación de los valores de precipitación. Es interesante destacar que este patrón es muy similar al observado a 24 y 36 horas de pronóstico, con la diferencia de que la amplitud de los errores se acentúa a medida que aumenta el plazo de pronóstico. Vinculando los errores de precipitación con los errores en la circulación en niveles bajos descritos en el Capítulo 4, se podría decir que, independientemente del régimen de precipitación en el cual se inicie la simulación, el pronóstico tiende hacia una configuración tipo ZCAS activa como la descrita en Nogués-Paegle y Mo, 1997, en donde la subsidencia compensatoria de la actividad convectiva en la zona de la ZCAS inhibe la convección sobre el centro y norte de Argentina. Este error sistemático es consistente con los errores encontrados por Pessacg, 2008, en simulaciones de más largo plazo y con los encontrados en la versión operativa del modelo WRF sobre Sudamérica (S. Cardazzo, comunicación personal). Para estudiar cuan estable es la determinación de los errores sistemáticos utilizando un período de entrenamiento de 15 días, en la Figura 6.2 se incluye la desviación estándar del bias sobre el período de estudio. Las regiones donde el error sistemático es mayor, también muestran una importante variación de los errores sistemáticos dependiendo del período considerado.

Otra de las modificaciones propuestas en la metodología presentada por HC98, es el tratamiento diferencial de los ceros en la verificación (calibración 2). La Figura 6.3, muestra la forma del histograma de rangos obtenido con el algoritmo propuesto por HC98

y el obtenido considerando el tratamiento diferencial de los ceros. Como se puede apreciar en dicha figura, la probabilidad de tener un cero en la precipitación, es muy grande cuando la dispersión del ensamble es baja (es decir cuando la precipitación pronosticada es baja), alcanzando valores de entre 60 y 80 %. A medida que aumenta la precipitación media pronosticada por el ensamble también lo hace la dispersión del mismo y la probabilidad de que la verificación sea 0 se reduce. El no incluir los ceros en el histograma de rangos no sólo modifica la frecuencia del intervalo más bajo del histograma, sino que modifica también los valores más bajos. El único rango que no se ve modificado es el más alto. En este caso, como solo estamos considerando aquellos casos en los que al menos uno de los miembros del ensamble pronosticó lluvia por encima de 0, el valor de verificación 0 nunca puede ocupar el último rango del histograma. Los demás rangos sí pueden ser ocupados por ceros en la verificación, porque en el caso de que más de un miembro del ensamble pronostique 0, el rango se define aleatoriamente.

Las diferencias observadas en esta sección, servirán para poder entender mejor los cambios que cada una de estas modificaciones introducen a los pronósticos probabilísticos y a evaluar los resultados obtenidos a partir de la verificación de los mismos.

### **6.3.2 Verificación de los pronósticos probabilísticos generados mediante las diferentes estrategias de calibración.**

Los diagramas de confiabilidad que se muestran en la Figura 6.4 muestran que, como era de esperarse, el pronóstico no calibrado sobreestima la precipitación para casi todos los umbrales de probabilidad. Las calibraciones 1 y 3, muestran una mejora significativa con respecto al pronóstico no calibrado tal cual se había mostrado para otro período del año en el Capítulo 3, pero no logran solucionar totalmente el problema, ya que continúan exhibiendo una sobreestimación de la probabilidad sobre todo para valores altos de probabilidad pronosticada. Los valores más altos de probabilidad pronosticada, se alcanzan cuando el umbral se encuentra en el primer intervalo del histograma de rangos. HC98, en estos casos proponen que la probabilidad de que la verificación se encuentre entre el mínimo miembro del ensamble y el umbral se calcule en forma proporcional a la distancia entre el umbral y el mínimo miembro del ensamble. Esta hipótesis de que la distribución de probabilidades dentro de este intervalo es lineal no está convenientemente

fundamentada (como sí lo está en el caso de los restantes intervalos a excepción del último). Muchos de los casos en donde la precipitación observada es cero, caen dentro del primer intervalo del histograma de rangos, pero dichos valores nunca estarán por encima de ningún umbral lo cual puede conducir a una sobreestimación de la probabilidad que es lo que se observa en la Figura 6.4. La calibración 2, al remover los ceros en la verificación del histograma de rangos utilizado para la calibración, mejora sensiblemente este problema, reduciendo la sobreestimación de la probabilidad para los valores altos de probabilidad. Esto también se puede apreciar en los gráficos que muestran la frecuencia con la cual se pronostican los diferentes valores de probabilidad que se incluyen en la Figura 6.4. Un comportamiento muy diferente puede observarse entre los pronósticos calibrados y el no calibrado en la frecuencia de pronósticos con 0% de probabilidad. En el umbral de 2.5 mm se puede apreciar que esta frecuencia es mayor para el ensamble no calibrado que para el calibrado, esto se debe a que en el ensamble no calibrado para que la probabilidad pronosticada sea 0% basta con que todos los miembros del ensamble estén por debajo del umbral indicado, mientras que para el caso del pronóstico calibrado, la probabilidad es 0% solo cuando todos los miembros del ensamble pronostican 0 mm. En el caso del pronóstico control, la probabilidad es 0% cuando el pronóstico control pronostica 0 mm, y esto ocurre en este caso aún con mayor frecuencia que para el ensamble no calibrado. Esto puede conducir a un aumento de las sorpresas como se ve sobre todo en la región 3 en donde el aumento de la frecuencia de los 0 de probabilidad está acompañado por un aumento en la probabilidad de ocurrencia de precipitaciones cuando el pronóstico indica 0% de probabilidad de lluvia por encima del umbral lo cual tiende a degradar la confiabilidad del pronóstico.

Por otra parte, se incluyen además en la Figura 6.4, los pronósticos obtenidos a partir de las calibraciones 4 y 5. Desde el punto de vista de la confiabilidad, en ambos casos se obtiene un pronóstico confiable en concordancia con los resultados obtenidos en el Capítulo 3.

En cuanto al comportamiento de las diferentes regiones analizadas, las regiones 1 y 2, muestran mayor resolución que la región 3 como se desprende de la distribución de la confiabilidad y la frecuencia de ocurrencia de los diferentes valores de probabilidad pronosticada. En la región 1, el menor número de datos produce mayor ruido en los resultados obtenidos a partir de la verificación que también se puede apreciar como una

mayor amplitud del intervalo de confianza para la confiabilidad. No obstante, el comportamiento de las calibraciones es similar en todas las regiones analizadas.

Para ejemplificar lo que sucede en términos del diagrama de confiabilidad para umbrales de precipitación mayores, la Figura 6.5 muestra el diagrama de confiabilidad para el umbral de 38.1 mm. En este caso a diferencia del anterior podemos apreciar que la curva de confiabilidad del pronóstico no calibrado muestra poca o ninguna resolución. No obstante existe una diferencia significativa entre la probabilidad de ocurrencia del fenómeno cuando la probabilidad pronosticada es 0 y aquella que se observa cuando la probabilidad es mayor que 0. La calibración en este caso, reduce la probabilidad de ocurrencia del fenómeno cuando la probabilidad pronosticada es 0% y mejora la confiabilidad permitiendo pronósticos confiables en el rango 0-30% en las regiones 1 y 2 y 0-20% en la región 3. Como se puede observar en esta figura, la calibración permite mejorar levemente el pronóstico para umbrales extremos, aunque lo reducido del período de entrenamiento y el de verificación hacen difíciles obtener estadísticas confiables para estos umbrales de precipitación confirmando que el pronóstico de valores extremos continúa siendo un problema.

Para resumir el comportamiento de las diferentes calibraciones ensayadas para los distintos umbrales de precipitación utilizados, se calculó el BSS y sus componentes como en el Capítulo 3. Los resultados se muestran en la Figura 6.6. A partir del BSS se puede apreciar que los mejores resultados se obtuvieron para la región SE de Brasil, seguidos bastante de cerca por la región norte de Argentina y que las diferentes calibraciones ensayadas muestran una mejora ampliamente significativa en el BSS. Dicha mejora está explicada principalmente por el aumento de la confiabilidad de los pronósticos. La descomposición del BSS permite ver que los pronósticos calibrados a partir de la media del ensamble y del pronóstico control son los que muestran mejores resultados en cuanto a la confiabilidad del pronóstico para todos los umbrales considerados, con excepción de lo que sucede para el pronóstico control en la región 3. Las calibraciones 1 y 3, por otra parte, muestran resultados sensiblemente inferiores sobre todo para los umbrales bajos, esto es así porque cuanto más bajo es el umbral mayor es la probabilidad de que el mismo esté por debajo del mínimo miembro del ensamble y que por lo tanto se aplique la hipótesis de que la probabilidad se distribuye linealmente en dicho intervalo. El tratamiento diferencial de los ceros introducido en la calibración 2, permite corregir

adecuadamente este problema, ya que se puede apreciar que en esta calibración la confiabilidad en los umbrales bajos se incrementa. No obstante no llega a superar el desempeño de las calibraciones 4 y 5 en términos de la confiabilidad de los pronósticos. Al analizar la componente de resolución, llama la atención el hecho de que si bien el pronóstico derivado del control muestra resoluciones más cercanas a 0 en las 3 regiones consideradas -como era de esperar-, las diferencias no llegan a ser significativas al 95 %. No obstante, a medida que se extiende el plazo de pronóstico, las diferencias entre los pronósticos probabilísticos derivados del ensamble y los del pronóstico control sí se hacen significativos indicando una mayor degradación con el plazo de pronóstico para el control (Figura 6.7), esto también se cumple para otros índices utilizados en la verificación. Estos resultados concuerdan con el comportamiento observado en el Capítulo 5, en donde se veía que las ventajas de la utilización del ensamble comenzaban a ser evidentes a partir de las 36-48 horas de pronóstico (aunque este plazo puede depender del método utilizado para perturbar las condiciones iniciales). Tal como se discutió en el Capítulo 3, las calibraciones no impactan significativamente sobre la resolución del pronóstico, con excepción de la calibración 3. Esta calibración muestra valores de resolución más negativos (mejores) respecto de las restantes calibraciones y del pronóstico no calibrado y además las diferencias resultan ser significativas. Este resultado es esperable, porque se está agregando información extra a la calibración. Es decir que la misma no se hace solo en función de la dispersión y los valores de precipitación pronosticados por el ensamble, sino que además se introduce la información referente a los errores sistemáticos lo cual permite detectar mejor los eventos de lluvia y no lluvia. Las ventajas de la calibración 3, se destacan sobre todo en la región 1 que está caracterizada por un déficit de precipitación pronosticada como se ve en la Figura 6.2. Esta región es la que sufre una degradación más rápida de la calidad de los pronósticos probabilísticos de precipitación debido a que muchos sistemas precipitantes son totalmente ignorados a plazos iguales o mayores de 48 horas.

En la Figura 6.8, se muestran diagramas ETS-bias para las calibraciones 1 a 5 y para el pronóstico no calibrado sobre las 3 regiones. Sobre la región 1 en particular, se puede apreciar que la habilidad de los pronósticos en detectar la ocurrencia de los diferentes umbrales sufre una degradación durante el proceso de calibración. Esta degradación no se observa en la región 2, donde los pronósticos calibrados muestran

valores de ETS similares o mejores al del pronóstico sin calibrar, pero vuelve a insinuarse en la región 3. La razón de este comportamiento no es del todo clara, pero los resultados obtenidos en la región dos parecen sugerir que no se trata de una limitación inherente al proceso de calibración. Algunos experimentos posteriores, indicaron que el bajo número de estaciones disponibles para la calibración podría ser una de las razones: las regiones 1 y 3 tienen un menor número de estaciones que la 2.

### **6.3.3 Análisis del valor económico del pronóstico.**

Una herramienta útil para la verificación de pronósticos es el cálculo del valor relativo del pronóstico. Este parámetro se calcula asumiendo un esquema sencillo de costos y beneficios que dependen básicamente de la información meteorológica provista (el esquema más sencillo es aquel en donde el usuario debe tomar algún tipo de acción preventiva frente a la ocurrencia de un fenómeno). El detalle de la definición y la forma en la que se calcula el valor del pronóstico se describe en el Apéndice A. Dado un pronóstico probabilístico de ocurrencia de un determinado fenómeno, podemos decir que el valor relativo es el valor adicional que aporta el pronóstico por sobre el conocimiento del comportamiento climatológico del fenómeno en cuestión. La confiabilidad es un factor importante que incide directamente sobre el valor del pronóstico, dado que si  $L$  es la pérdida que produce la ocurrencia no pronosticada del fenómeno y  $C$  es el costo de protección, el umbral de probabilidad óptimo para tomar la decisión de protegerse o no contra el fenómeno está dado por  $C/L$  (Zhu et. al. 2002). Si el pronóstico no es confiable, entonces el umbral óptimo podría ser otro, pero a priori no sabríamos tampoco cual es dicho umbral.

La Figura 6.9, muestra el valor relativo de los diferentes pronósticos analizados en el presente capítulo para los umbrales de 1, 2.5 y 6.35 milímetros. Las regiones 1 y 2, muestran valores similares, mientras que la región 3, muestra valores relativos significativamente más bajos, lo cual es consistente con las limitaciones observadas en otros aspectos previamente analizados. Además el rango de valores de  $C/L$  para los cuales el valor relativo del pronóstico es positivo (es decir que aporta información útil) es menor que en las regiones 1 y 2. Esto es importante porque indica que el rango de usuarios a los cuales la información meteorológica podría resultar de utilidad es menor dado que cada

usuario en principio no puede modificar la relación entre el costo de protección y la pérdida ocasionada por el fenómeno.

Cuanto mayor es el valor de  $C/L$ , quiere decir que el costo de protección es más caro frente a la pérdida y por ende, el umbral de probabilidad a partir del cual conviene protegerse del fenómeno es mayor. Es decir que valores altos de  $C/L$  van a estar asociados con valores de probabilidad altos en la toma de decisión. Como se puede ver en la Figura 6.9, todas las calibraciones muestran un mejor desempeño en términos del valor del pronóstico (con respecto al pronóstico no calibrado) cuando la relación  $C/L$  es relativamente alta. Esto está vinculado con la falta de confiabilidad que el pronóstico no calibrado tiene para el presente sistema de ensambles y pone de manifiesto la importancia del proceso de calibración en la generación de un producto útil para los usuarios finales del pronóstico. Con respecto al comportamiento de las diferentes calibraciones, en general se puede observar que las calibraciones 1 y 3, muestran los peores resultados. Esto es porque como se discutió previamente, estas calibraciones son las que tienen mayores problemas para reducir la sobrestimación de la probabilidad para valores relativamente altos de probabilidad. Por otra parte, en la región 1 y 3, la calibración 3, muestra resultados levemente superiores respecto a la calibración 1, mostrando el impacto positivo que tiene incluir los errores sistemáticos del modelo como variable adicional en el proceso de calibración.

Existe otro rango de valores de  $C/L$  que varía de acuerdo a la región, donde las calibraciones degradan el valor del pronóstico (por ejemplo para la región 1 se extiende desde 0.1 hasta 0.4 sin depender demasiado del umbral considerado). Este comportamiento es similar al encontrado por Ruiz et al., 2008. Si bien la Figura 6.9 parece sugerir que estas diferencias solo son importantes para la región 1, el origen de las mismas podría estar relacionado con los máximos de frecuencia de ocurrencia de algunos umbrales de probabilidad que se observan en la Figura 6.4, en particular para algunas calibraciones como la calibración 2 que es la que muestra una degradación más marcada con respecto al pronóstico no calibrado en la región 1. Por otro lado como se explorará más adelante esta degradación podría estar relacionada con una cantidad de datos insuficiente, dado que la región 1 es la que menor disponibilidad de datos de lluvia tiene para aplicar las diferentes calibraciones.

#### **6.3.4 Utilización de las estimaciones CMORPH en la calibración.**

A continuación se resumen los resultados obtenidos al haber utilizado las estimaciones CMORPH para la calibración de los pronósticos probabilísticos. Es importante destacar que las regiones utilizadas para la verificación así como los datos utilizados en la verificación son siempre los mismos.

Cuando se utilizan las estimaciones CMORPH, surge la ventaja de la mayor disponibilidad de datos dada la mayor cobertura espacial de los mismos. No obstante estas estimaciones pueden presentar errores sistemáticos y no sistemáticos respecto de los datos pluviométricos. Los errores sistemáticos tienen que ver con bias de las estimaciones respecto a las mediciones pluviométricas, que ya fueron analizados sobre la región en Ruiz 2008 (enviado) y que fueron parcialmente corregidos para la estación cálida 2002-2003. Debe destacarse que la forma en la que fueron removidos los errores sistemáticos podría no ser tan eficiente para otros períodos, dada la menor cantidad de pluviómetros disponibles para estimar el signo y magnitud de dichos errores. Por otra parte, los errores no sistemáticos incluyen posibles fallas en la detección de precipitaciones asociadas a tormentas que pudieran no ser detectadas por los satélites de la constelación que componen los CMORPH o bien en las suposiciones realizadas por el algoritmo de propagación de los sistemas (Joyce et. al. 2004).

Para poder analizar el impacto global que estos errores tienen sobre la calibración de los pronósticos, se propone la utilización del algoritmo de calibración 4, basado en la media del ensamble, utilizado con los datos pluviométricos (como en las secciones anteriores), con los datos CMORPH sin la remoción de los errores sistemáticos y con los datos CMORPH con el error sistemático removido como se propone en Ruiz 2008 (enviado). Para extraer la mayor ventaja posible de este conjunto de datos se utilizaron todas las estimaciones disponibles dentro de las regiones analizadas (tanto aquellas donde había datos pluviométricos como aquellas en donde los mismos no existían).

La Figura 6.10, muestra el impacto de utilizar los diferentes conjuntos de datos sobre la forma de la curva de probabilidad para el umbral de 2.5 mm, es decir como varía la probabilidad de ocurrencia de lluvia por encima de 2.5 mm en función de la lluvia pronosticada por la media del ensamble. Como se puede apreciar en la figura, las

diferencias son pequeñas. Sobre todo para los valores más bajos de probabilidad. Para los valores más altos, las diferencias son un poco más importantes, probablemente debido a que la frecuencia con la que ocurren precipitaciones asociadas a valores altos de probabilidad es menor y por ende el impacto de tener una mayor cantidad de información disponible utilizando los datos CMORPH es más evidente en esta región de la curva. Por otro lado las diferencias son más evidentes aún sobre la región 3 y 1 que son las que tienen una cobertura espacial más pobre en cuanto a la disponibilidad de datos pluviométricos. Es importante destacar que no existen grandes diferencias entre el CMORPH calibrado y el CMORPH sin calibrar. Esto probablemente se deba a que para el umbral considerado, los errores sistemáticos en las estimaciones CMORPH no son tan importantes.

La Figura 6.11, muestra los BSS obtenidos a partir de los diferentes conjuntos de datos. Para las regiones 1 y 2, se puede apreciar que los resultados obtenidos con datos pluviométricos y con CMORPH son muy similares a iguales algoritmos de calibración, hasta el umbral de 25 mm. A partir de dicho valor, los datos CMORPH sin la remoción de sus errores sistemáticos, muestran una significativa degradación de los pronósticos que está principalmente ligada a una menor confiabilidad de los pronósticos probabilísticos calibrados utilizando dicho conjunto de datos. Esta degradación probablemente se deba a la sobreestimación que realizan los CMORPH principalmente de los valores altos de precipitación, lo cual conduce a un bias positivo en la probabilidad de los umbrales más altos (no se muestra la figura). Cuando se corrigen los errores sistemáticos en los datos CMORPH, mejora la confiabilidad, sobre todo para los umbrales de precipitación altos en las regiones 1 y 2. En la región 3, la corrección de los errores sistemáticos produce un deterioro en la confiabilidad de los pronósticos, esto puede deberse a que los errores sistemáticos fueron estudiados a partir de datos que corresponden a la región 1 y 2 principalmente y como se discute en Ruiz 2008 (enviado) los resultados obtenidos en la región tropical de Sudamérica no fueron tan satisfactorios como en otras regiones.

La Figura 6.11, muestra además el desempeño de la calibración 6 que fue diseñada para sacar más provecho de la información provista por los CMORPH. Sobre la región 1, el desempeño de esta calibración, es similar al de las demás. No obstante, la Figura 6.11 indica que esta misma calibración produce una sensible degradación del pronóstico sobre la región 2 particularmente para umbrales mayores a 2.5 mm. Es probable que la

reducción de la cantidad de datos disponibles para la calibración que resultan de tomar un dominio local para realizar la calibración sea el principal responsable de esta degradación. Por otra parte la figura indica que las mayores diferencias se producen en la componente de resolución del BSS que se ve afectada negativamente sobre la región 2 al introducir el algoritmo de calibración local.

La Figura 6.12, muestra los diagramas ETS-BIAS para la calibración 4 utilizando pluviómetros, CMORPH calibrado y CMORPH sin calibrar. Se observa que en la región 1, la utilización de las estimaciones CMORPH tiene un impacto positivo sobre la habilidad de los pronósticos. Es probable que esto sea consecuencia principalmente de la mayor cantidad de datos disponibles para construir las curvas que relacionan la precipitación pronosticada por el ensamble con la probabilidad, como se discutió previamente. De hecho, al utilizar los CMORPH, el ETS de los pronósticos calibrados es similar al del pronóstico no calibrado, es decir que no se observa en el caso de los CMORPH la degradación de la habilidad de los pronósticos calibrados que se encontró cuando se utilizaron pluviómetros para la calibración sobre la región 1 (ver por ejemplo la Figura 6.8). Sobre la región 2 en donde la disponibilidad de datos pluviométricos es mayor, el impacto de utilizar los datos CMORPH es casi imperceptible.

A partir de la combinación de las Figuras 6.12 y 6.13 se puede sintetizar el efecto de las estimaciones CMORPH sobre la habilidad del pronóstico y sobre su valor relativo respectivamente. La utilización de los CMORPH en el proceso de calibración produjo un impacto positivo en la región 1, mientras que en la región 2 y 3 producen resultados muy similares a los obtenidos con los datos pluviométricos. Por otra parte, la inclusión de una calibración de tipo local, produjo resultados similares en cuanto a la habilidad y el valor relativo del pronóstico sobre la región 1, pero no produjo mejoras sobre las regiones 2 y 3, en donde en algunos casos trajo aparejado una degradación de la calidad del pronóstico. Es importante destacar que si bien en la región 2 y 3 el impacto de utilizar los CMORPH en la calibración no es significativo, la cantidad de datos pluviométricos utilizados en este estudio es muy superior a la cantidad de datos disponible en forma operativa. El impacto de la utilización de las estimaciones CMORPH comparada con la utilización de las estaciones de la red operativa podría resultar superior a las diferencias obtenidas en el presente trabajo.

## **6.4 Conclusiones.**

### **6.4.1 Discusión.**

En este capítulo se utilizaron los pronósticos de precipitación derivados de un ensamble regional basado en la técnica de los vectores criados para evaluar diferentes estrategias de calibración para la estación cálida 2002-2003. La red pluviométrica del experimento SALLJEX en conjunto con los datos de ANA, permitieron evaluar la calidad de los pronósticos probabilísticos generados y compararlos con las estimaciones CMORPH que presentan una mayor densidad espacial.

Al igual que lo encontrado en el Capítulo 3 del presente trabajo de tesis, los pronósticos probabilísticos obtenidos a partir del ensamble fueron superiores a los derivados a partir de un pronóstico determinístico, diferencia que se va acentuando a medida que se incrementa el plazo de pronóstico. No obstante, queda claro que para que la información proveniente del ensamble sea aprovechada en forma adecuada, los pronósticos probabilísticos de precipitación deben ser calibrados. Esto es fundamental a la hora de generar productos para los usuarios finales, dado que para una determinada actividad con un cierto costo de protección y una cierta pérdida que se produce cada vez que el fenómeno no fue pronosticado, existe un valor óptimo de probabilidad en el cual hay que tomar la decisión entre proteger o no proteger la actividad. Si la decisión se toma a partir de un pronóstico probabilístico que no es confiable se estará utilizando un umbral que no es el óptimo para la toma de decisiones y por ende el valor asociado a ese pronóstico no será el mejor que se pueda obtener, de hecho puede ser que el valor de dicho pronóstico sea equiparable al obtenido a partir de una simple estadística de los errores y un pronóstico determinístico basado en un único modelo.

Las diversas calibraciones propuestas y evaluadas en el presente capítulo dieron resultados satisfactorios en cuanto a que permitieron mejorar la confiabilidad del pronóstico y por ende incrementar su valor. No obstante existen algunas diferencias entre ellas:

- La calibración 1 basada en el histograma de rangos (HC98) presenta deficiencias en cuanto a la estimación de probabilidades cuando el umbral de lluvia se encuentra en el primer rango del histograma. Estas limitaciones tienen que ver principalmente

con la hipótesis de que la probabilidad se distribuye linealmente y con la forma en la que se construye el histograma de rangos.

- La calibración 2 logra corregir parcialmente el problema presente en la calibración 1, mediante la reformulación del método utilizado para construir el histograma de rangos. Los resultados obtenidos mediante esta modificación mejoran sensiblemente la sobreestimación de los valores altos de probabilidad presentes en la formulación anterior.
- La calibración 3 incorpora información sobre la distribución espacial de los errores del modelo, intentando agrupar los puntos con valores similares de bias sistemático. Como los errores sistemáticos del modelo modifican la forma del histograma de rangos, es natural pensar que la agrupación por regiones con valores similares de error sistemático puede producir mejoras en la calibración. De hecho, este algoritmo produjo mejoras en la resolución del pronóstico siendo los resultados más notorios sobre la región 1. En cuanto a la confiabilidad, la incorporación de los errores sistemáticos del modelo también produjo algunas mejoras que fueron más notorias para los umbrales altos de precipitación.
- La calibración 4, basada puramente en el valor medio de precipitación pronosticado por el ensamble, tuvo los mejores resultados en cuanto a la confiabilidad y con respecto a la resolución resultados similares a los del pronóstico no calibrado (lo cual era de esperarse dado que no estábamos incorporando información adicional). Esta metodología es además de aplicación más sencilla y requiere un menor costo computacional, aunque es preciso aclarar que en ninguno de los casos el costo computacional representa un problema a la hora de pensar en una implementación operativa. Es por eso que este tipo de calibración es una de las alternativas más interesantes al considerar un algoritmo para desarrollar un pronóstico probabilístico de precipitación calibrado que funcione en forma operativa y además como se discutió en los resultados, puede ser aplicada también a pronósticos determinísticos obteniéndose interesantes resultados.
- Con respecto a la utilización de los datos CMORPH en el proceso de calibración, se pudo observar que la mayor disponibilidad de datos tuvo un impacto positivo en la calibración que se notó principalmente en las regiones 1 y 3 donde la disponibilidad de datos pluviométricos es menor. Es importante destacar que si bien durante la

estación cálida 2002-2003 había una gran disponibilidad de datos sobre el centro y norte de Argentina debido a los datos aportados por el experimento SALLJEX, la red operativa accesible en tiempo real que podría utilizarse en la calibración de pronósticos en la región es sensiblemente más escasa. Es por eso que los resultados obtenidos a partir del uso de las estimaciones CMORPH sobre la región cobran mayor relevancia.

Las estimaciones CMORPH presentan errores sistemáticos que son más importantes en los umbrales de precipitación superiores (de 12 mm en adelante). Esto produce una degradación de las probabilidades calibradas en base a las estimaciones CMORPH para los umbrales de precipitación más elevados, pero no afecta prácticamente a los umbrales más bajos. Por otra parte, los errores sistemáticos presentes en las estimaciones CMORPH pueden ser parcialmente corregidos como en Ruiz (2008) para obtener un conjunto de datos que no produzca una degradación importante de los pronósticos para los umbrales de precipitación más altos.

Finalmente se buscó aprovechar la mayor disponibilidad de datos presentes en las estimaciones CMORPH, implementando la calibración 6 de manera local. En esta implementación preliminar del algoritmo, no se logró obtener una mejora significativa mediante la aplicación local del algoritmo con lo cual es necesario seguir indagando en el tema. No obstante en la región 1 se observó una leve mejora en términos de la resolución de los pronósticos.

#### **6.4.2 Trabajos futuros.**

Con respecto a los trabajos futuros, los experimentos realizados en el presente capítulo, abren posibilidades para seguir mejorando la calidad de los pronósticos probabilísticos de precipitación a corto plazo.

- En cuanto a los pronósticos derivados de un pronóstico determinístico (no de la media del ensamble), una de las posibilidades es utilizar en lugar del valor de precipitación pronosticado por el modelo, un valor medio espacial entre cada uno de los puntos y los puntos adyacentes. Esta sería una manera de considerar la

- incertidumbre en cuanto a la posición de los sistemas precipitantes y eventualmente podría ser extendida para considerar el tiempo de ocurrencia de la precipitación. En cualquiera de estos casos, se puede evaluar la sensibilidad de los resultados obtenidos al tamaño de la región promediada, es decir cuantos puntos adyacentes se van a considerar en el promedio, y eventualmente la inclusión de pesos dependientes de la distancia.
- Con respecto a los algoritmos basados en el pronóstico control o la media del ensamble es interesante evaluar si existe alguna sensibilidad de la calidad del pronóstico probabilístico obtenido a los intervalos utilizados para construir la curva de calibración. Si la hubiera, sería necesario encontrar un algoritmo objetivo que permitiera determinar un valor óptimo o al menos razonable de los intervalos dependiendo de las características climatológicas de la precipitación en cada región.
  - Las calibraciones basadas en el pronóstico control o la media del ensamble, también pueden ser combinadas con información sobre los errores sistemáticos del modelo análogamente a como se hizo con los histogramas de rangos para incrementar la resolución del pronóstico. No obstante, los errores sistemáticos pueden no ser la variable con la cual se obtengan los mejores resultados. Es necesario experimentar con otras variables relacionadas que puedan influir directamente en los procesos asociados a la ocurrencia de precipitación como por ejemplo el agua precipitable, el CAPE o alguna medida de la inestabilidad que sea más continua como el Lifted Index.
  - Las estimaciones CMORPH, no son las únicas estimaciones de precipitación disponibles. Eventualmente se debería evaluar el impacto de utilizar otras estimaciones de precipitación o una combinación de diferentes estimaciones de precipitación que permita mejorar aún más los resultados.
  - En el caso de la implementación local del algoritmo de calibración (calibración 6) que mostró resultados dispares en las diferentes regiones consideradas, es necesario evaluar en este caso la sensibilidad de los resultados al tamaño de la región utilizada para realizar la calibración local. Una región demasiado pequeña, podría generar una curva con demasiado ruido debido a la falta de datos, mientras que por otro lado una región demasiado grande podría incorporar datos de regiones que tengan características climatológicas muy distintas o en donde los

errores del modelo se comporten de manera muy diferente lo cual impactaría negativamente sobre la calibración.

- El algoritmo de calibración local puede ser mejorado utilizando áreas en donde la forma no esté determinada de ante mano (por ejemplo regiones cuadradas como las que se utilizaron en el presente trabajo de Tesis) sino que sean regiones que busquen maximizar la homogeneidad dentro del conjunto utilizado para la calibración. Por ejemplo, en las laderas orientales de la cordillera de los Andes, las zonas más homogéneas desde el punto de vista climatológico se extienden con orientación norte sur, dado que el mayor gradiente de las propiedades está en la dirección oeste-este. En este caso, la región óptima que podríamos utilizar tendría una mayor extensión norte-sur que este-oeste garantizando una mayor homogeneidad en el comportamiento climatológico dentro de la región utilizada en la calibración. El diseño e implementación de un algoritmo para definir regiones afines es parte de un trabajo en curso.
- Por otra parte es necesario considerar además otros algoritmos como el de Stensrud y Yussouf (2007) para comparar su desempeño en la calibración de los pronósticos probabilísticos de precipitación sobre la región y evaluar la posibilidad de incorporar alguno de los aspectos discutidos dentro de dicho algoritmo. En particular, el algoritmo planteado por Stensrud y Yussouf (2007) propone reemplazar los miembros del ensamble por elementos al azar tomados de una muestra de entrenamiento, en donde la muestra depende de la precipitación pronosticada por cada miembro del ensamble. Este algoritmo tiene la ventaja de que permite incluir fácilmente diferentes modelos dentro del ensamble, ya que cada uno de ellos podría tener sus propias muestras de entrenamiento lo cual permitiría incluir la información de los errores sistemáticos de cada miembro del ensamble dentro de la calibración.
- Finalmente una de las limitaciones más notorias de los resultados obtenidos en el presente trabajo es el acotado período de tiempo durante el cual se lleva a cabo la verificación de los resultados (2 meses). Esto se debe principalmente al costo computacional asociado a la generación del ensamble, a la disponibilidad de condiciones iniciales para la inicialización del mismo y de datos de pluviómetros más allá del período del experimento SALLJEX. Es por eso que es importante

explorar cómo es la variación a lo largo del año de las propiedades de los pronósticos probabilísticos sobre la región, en particular con relación a la ventaja de la utilización de los datos CMORPH en períodos en donde la precipitación convectiva no es tan frecuente.

## Capítulo 7:

### Conclusiones y trabajos futuros.

#### 7.1 Conclusiones:

Una de las principales preguntas planteadas en el presente trabajo de tesis fue en qué medida la implementación del pronóstico por ensambles permitiría mejorar el pronóstico a corto plazo sobre Sudamérica. Desde el punto de vista de la implementación operativa, la técnica de pronóstico por ensambles representa un importante costo computacional que puede ser mayor o menor dependiendo de la técnica implementada, pero que en todos los casos es muy superior a realizar pronósticos puramente determinísticos. Es por eso que se buscó abordar esa pregunta a partir no sólo de la valoración de la calidad de pronósticos por conjuntos sino también comparándolos con aquélla resultante de los pronósticos determinísticos. Esta comparación es fundamental al momento de encarar el esfuerzo requerido para la implementación operativa de un ensamble. Por otra parte, estudios como el desarrollado en el presente trabajo de Tesis conducen, además, a un aprovechamiento más eficiente de la información provista por un ensamble sobre nuestra región, identificando aquellas regiones en donde la ventaja de utilizar ensambles es mayor y a través de la evaluación de diferentes estrategias que permitan incrementar el valor de la información provista a los usuarios finales del pronóstico.

Más aún, la comparación de los productos derivados del ensamble con productos similares (por ejemplo pronósticos probabilísticos) pero derivados de pronósticos determinísticos a través de un estudio estadístico de sus errores, permitió ir más a fondo en la valoración de los productos del ensamble.

Dentro de los productos del ensamble que fueron evaluados, se encuentra la media del ensamble, utilizada como predictor para diferentes variables. Los resultados obtenidos en este sentido, muestran que la reducción del error que se puede lograr mediante la

utilización de la media del ensamble en los pronósticos a corto plazo (hasta 48 hs) no supera el 10% en la mayoría de las variables analizadas. Por otra parte, esta reducción del error sería útil en la estimación de valores numéricos como temperaturas extremas diarias, pero por lo general no presenta diferencias importantes que aporten a la intensidad o ubicación de los sistemas de escala sinóptica en la región considerada en los pronósticos a corto plazo. Es importante mencionar, además, que la utilización de la media del ensamble introduce un suavizado de los campos que puede conducir a la reducción de la habilidad del pronóstico para detectar eventos extremos sobre todo en variables con gran variabilidad espacial como la precipitación.

La media del ensamble también se evaluó para el pronóstico cuantitativo de precipitación a corto plazo. En este caso, también se encontró que conduce a una sistemática sobreestimación de los valores de precipitación bajos y a una sub-estimación de los máximos, por lo cual no se recomienda la utilización de este campo directamente como herramienta para la elaboración de pronósticos, si bien puede resultar de utilidad en otras aplicaciones.

Una de las maneras más eficientes de aprovechar y resumir la información provista por el ensamble es la generación de pronósticos probabilísticos. La ventaja de este producto, es que los extremos no se suavizan y que además constituye una herramienta útil para la toma de decisiones por parte de los usuarios finales del producto. Como principal limitación figura el hecho de que debemos definir con antelación el fenómeno para el cual queremos estimar la probabilidad de ocurrencia (por ejemplo lluvia por encima de 20 mm), mientras que el producto nos dice cuan probable es la ocurrencia de ese fenómeno pero no nos dice nada sobre otros posibles fenómenos (por ejemplo la probabilidad de que llueva más que 20 mm no da mucha información sobre si puede llover más que 100 mm). En el presente trabajo de tesis, se desarrollaron pronósticos probabilísticos para la precipitación, aunque la misma metodología utilizada se puede aplicar a otras variables. De acuerdo a lo documentado en la bibliografía la precipitación es una de las variables que presenta mayores errores tanto en los pronósticos determinísticos como en los pronósticos probabilísticos y por lo general las técnicas utilizadas para mejorar los pronósticos de precipitación pueden ser aplicadas con igual o mejores resultados a otras variables del modelo.

Las comparaciones realizadas permitieron evaluar las ventajas de los pronósticos

probabilísticos derivados de un ensamble comparándolos con aquellos generados a partir de un pronóstico determinístico. La comparación se realizó en forma independiente para diferentes sub regiones que poseían regímenes climáticos diferentes. En general se pudo observar que los productos derivados del ensamble fueron superiores en cuanto a la capacidad de detectar la ocurrencia o no ocurrencia de diferentes valores de precipitación para todos los regímenes analizados. Sin embargo, las ventajas de utilizar pronósticos por ensambles fueron más evidentes en regiones donde la actividad baroclínica era mayor. Por otro lado, se mostró que la utilización de los pronósticos por ensambles produce un incremento en el valor económico de la información, lo cual es fundamental para el desarrollo de aplicaciones.

Las ventajas de la utilización de los ensambles aumentan a medida que aumenta el plazo de pronóstico: en el caso de los pronósticos de precipitación, los pronósticos derivados de una corrida individual del modelo, resultaron competitivos a 24 horas de plazo, pero a 36 y 48 horas los pronósticos derivados del ensamble mostraron mejores resultados.

Es por eso que a partir de los resultados obtenidos en este trabajo de tesis, se recomienda la implementación de un sistema de pronóstico por ensambles a corto plazo para la generación de pronósticos probabilísticos de precipitación y de todas aquellas variables que puedan resultar de interés para distintas aplicaciones y para la reducción del error en aquellas aplicaciones que requieran la estimación de algún parámetro en particular (como por ejemplo las temperaturas extremas diarias).

Por otra parte, en la generación de pronósticos probabilísticos y en su aplicación a actividades específicas, es fundamental contar con un pronóstico confiable en el sentido de que la probabilidad pronosticada se corresponda lo más estrechamente posible con la probabilidad observada del fenómeno. Para explorar este punto, en este trabajo de tesis se ensayaron diferentes estrategias de calibración de los pronósticos probabilísticos derivados del ensamble, basadas en las metodologías propuestas por Hamill y Colucci (1998) y Gallus y Seagal (2004). En particular también se propusieron mejoras a la metodología propuesta por los primeros, que impactaron favorablemente sobre la confiabilidad de los pronósticos. Todos los algoritmos implementados mejoraron significativamente la confiabilidad, incrementando el valor de la información obtenida a partir del ensamble. Por lo tanto, a partir de estos resultados, se recomienda que la

generación de pronósticos probabilísticos a partir de un ensamble en nuestra región esté acompañada de un algoritmo de calibración que incremente el valor de la información.

Como lo sugieren varios de los resultados encontrados en esta Tesis, uno de los aspectos claves del proceso de calibración es la disponibilidad de observaciones que permitan obtener una estadística robusta acerca de los errores del modelo pero sin la necesidad de recurrir a períodos de tiempo excesivamente largos (la utilización de tales períodos en pronósticos de alta resolución representa un enorme esfuerzo computacional). Para ello es fundamental contar con una red de observación que tenga una mayor densidad espacial. Esto permitirá contar con mayores volúmenes de datos para garantizar la estabilidad estadística de todos aquellos parámetros estimados durante el proceso de calibración y también permitirá obtener información adecuada sobre los errores sistemáticos del modelo y sobre los diferentes regímenes de precipitación que podrá ser utilizada para mejorar la implementación de la calibración sobre la región. Si la red de observaciones es lo suficientemente densa, se puede implementar una calibración de tipo dinámica, en la que la información de algunos días previos es utilizada para la calibración; esta alternativa tiene la desventaja de que las estadísticas de los errores obtenidos podrían estar demasiado sesgadas al comportamiento durante dicho período y podrían experimentar problemas durante los cambios de regímenes. No obstante, los algoritmos de calibración propuestos se adaptan rápidamente a los cambios en la configuración del / de los modelo/s utilizados.

Una alternativa para salvar la escasa cantidad de información pluviométrica disponible en la actualidad es la utilización de estimaciones de precipitación en base a sensores remotos, como por ejemplo las estimaciones CMORPH que fueron utilizadas aquí en el proceso de calibración de los pronósticos. Los resultados mostraron que estos datos pueden convertirse en una excelente alternativa que permitiría aplicar la calibración sobre regiones más pequeñas y homogéneas. De hecho en los resultados analizados se observó un impacto positivo de los datos CMORPH sobre la calidad de los pronósticos probabilísticos de lluvia que, presumiblemente, se debe a la mayor disponibilidad de información, ya que fue más notorio en las regiones donde la disponibilidad de datos era menor, mientras que en las regiones donde la disponibilidad de datos era similar, los resultados fueron comparables a los obtenidos con datos pluviométricos. Es importante destacar que los datos CMORPH, presentan errores sistemáticos que impactan

negativamente sobre la calibración afectando sobre todo a los umbrales más altos (por encima de 20 mm). Este aspecto puede ser corregido parcialmente como se describe en Ruiz (2008 enviado); dicha corrección impacta positivamente sobre la calidad de los pronósticos probabilísticos de lluvia para umbrales altos.

Otro aspecto fundamental a tener en cuenta a la hora de implementar operativamente un pronóstico por ensambles es la técnica utilizada para la generación del ensamble. En el presente trabajo de tesis se utilizaron 3 técnicas diferentes, SLAF (Scalled Lagged Averaged Forecast), vectores criados y el super ensamble de modelos. Las 3 técnicas utilizadas tienen en común el bajo costo computacional requerido para su implementación. Dentro de las 3, la técnica más costosa computacionalmente es la de los vectores criados, dado que en la presente implementación requirió la generación de un ensamble global de baja resolución para proveer condiciones iniciales perturbadas y condiciones de borde al ensamble regional. No obstante esta técnica requiere mucha menor capacidad computacional que otras como, por ejemplo, la técnica de los vectores singulares.

Dado que no todos los experimentos realizados con las diferentes técnicas fueron simultáneos no se pudo realizar una intercomparación de los resultados obtenidos con cada una de las mismas. No obstante se encontró que la técnica de SLAF presenta un comportamiento adecuado y permitió obtener pronósticos probabilísticos de precipitación con valores de habilidad (Skill Score de Brier) similares a los reportados por ejemplo por Hamill y Whitaker, 2006, para la misma época del año. Por otra parte, las comparaciones realizadas con los resultados del super-ensamble de modelos para el mismo período mostraron que sobre todo en la región de mayor actividad baroclínica, los resultados del SLAF fueron iguales o ligeramente superiores. Esto sugiere que el SLAF es una técnica sencilla y computacionalmente económica (no requiere generar un ensamble global) para la obtención de pronósticos por ensamble a escala regional. Considerando que actualmente en el país no se cuenta con un modelo global funcionando operativamente, la implementación del SLAF como técnica para generar perturbaciones en las condiciones iniciales y de borde resulta una buena alternativa. Por otra parte, la implementación del SLAF se puede complementar con la generación de vectores criados dentro del dominio regional para generar la perturbación en la condición inicial, manteniendo el SLAF para la perturbación exclusivamente de las condiciones de borde. La técnica de SLAF no obstante

no parece ser adecuada para la generación de perturbaciones adecuadas en la región tropical de Sudamérica como lo ponen de manifiesto el pobre desempeño de los pronósticos probabilísticos de lluvia derivados de este ensamble y el escaso crecimiento de la dispersión del ensamble en esta región como fue mostrado por Ruiz y otros (2006).

La técnica de super ensamble de modelos, mostró sobre todo un mejor desempeño en la región tropical de Sudamérica, aunque los resultados en el resto de la región fueron comparables a la técnica de SLAF. La ventaja de la técnica de super-ensamble es que incorpora la incertidumbre en la formulación del modelo dentro del ensamble. Esto puede introducir mejoras sobre todo en aquellas regiones donde el error del pronóstico esté dominado por los errores del modelo. Esto sugiere que la implementación de los pronósticos por ensamble en Sudamérica, deben incorporar de alguna manera la incertidumbre en los modelos. Este hecho se refuerza dado los errores sistemáticos presentes en los modelos analizados sobre nuestra región y que son discutidos en el presente trabajo de tesis. La incorporación de dichos errores puede realizarse mediante la técnica del super-ensamble o bien utilizando diferentes configuraciones dentro de un mismo modelo como el WRF. La introducción de los errores en el modelo como fuente de incertidumbre ciertamente trae aparejado algunas complicaciones posteriores, sobre todo en la calibración de los pronósticos probabilísticos. Las metodologías utilizadas en el presente trabajo asumen que los errores sistemáticos de los diferentes miembros del ensamble son iguales para todos los miembros, lo cual no se cumple estrictamente para un ensamble compuesto por modelos diferentes. Es por eso que para la calibración de los pronósticos probabilísticos en estos casos es necesaria la introducción de otros algoritmos, como el sugerido por McLean et al., 2007.

Con respecto a la técnica de los vectores criados, la misma permitió obtener buenos resultados en términos del pronóstico probabilístico de precipitación para la época estival. Estos resultados no pudieron ser comparados directamente con los del SLAF o los del super-ensamble de modelos. No obstante, para este ensamble los resultados mostraron que permite obtener ventajas sobre el pronóstico determinístico que son similares a las encontradas para las otras metodologías aplicadas. Para este ensamble en particular se encontró además que existe una cierta relación entre la incertidumbre presente en el análisis y las perturbaciones introducidas en la condición inicial, en concordancia con resultados anteriores y con resultados obtenidos para un sistema con 3 grados de libertad.

Las ventajas antes mencionadas se presentan en la región donde la actividad de ondas baroclínicas es mayor, en la región tropical de Sudamérica, se observó en general que existe poca relación entre la dispersión del ensamble y los errores del pronóstico utilizando esta metodología, lo cual sugiere que podría no ser la forma más apropiada para generar perturbaciones en la condición inicial en dicha región.

Para este ensamble en particular también se estudió la relación entre la dispersión del ensamble y el error en el pronóstico. Los resultados mostraron que si bien dicha relación no es óptima, en gran parte del dominio y para algunas variables, podría contribuir a generar productos de estimación de la incertidumbre en el pronóstico. Por otra parte, la existencia de una relación entre la dispersión y el error en el pronóstico es una de las bases que sustenta la generación de pronósticos probabilísticos a partir del ensamble, en lugar de utilizar estadísticas climatológicas del error. En este sentido los resultados obtenidos sugieren que en aquellas variables donde se encontró la existencia de cierta relación entre la dispersión del ensamble y los errores del pronóstico, es razonable pensar que la utilización del ensamble puede proveer pronósticos probabilísticos con mayor valor agregado respecto de una estadística de largo plazo. Para el caso particular de la precipitación, en concordancia con estudios previos, la relación entre dispersión y error, fue más difícil de analizar debido a la existencia de una fuerte relación entre el error y el valor de la variable. Cuanto mayor es la precipitación acumulada pronosticada por el modelo mayor es el error en dicha variable.

Finalmente, la implementación del modelo WRF para la generación de pronósticos a corto plazo sobre la región de Sudamérica pudo ser evaluada dentro del marco de los experimentos presentados en este trabajo de Tesis. Los valores de verificación de los diferentes parámetros obtenidos, muestran un buen desempeño general del modelo sobre la región. Sin embargo hay varios aspectos que es necesario mejorar en vistas a una implementación operativa. Uno de los aspectos negativos que más se destaca es una inadecuada representación de la intensidad de la corriente en chorro en capas bajas, que consiste en una subestimación de la misma, la cual se acentúa a medida que se extiende el plazo de pronóstico, a la par que la precipitación sobre la SACZ es fuertemente sobreestimada. Esto también está estrechamente vinculado con un error sistemático en la humedad sobre el norte de Argentina en donde se pronostican valores inferiores a los observados (justamente porque falla el transporte de aire cálido y húmedo desde la región

amazónica) impactando negativamente en el pronóstico de lluvia en dicha región, generando una subestimación en la cantidad de precipitación que se acentúa también al aumentar el plazo de pronóstico. En altura se observan apartamientos en la circulación que probablemente estén vinculados con las diferencias introducidas en el calentamiento de la atmósfera debido a los errores en la distribución de la precipitación. Si bien el WRF cuenta con muchas ventajas en cuanto a la disponibilidad de información, la sencillez de su instalación y manejo, es necesario adecuar su configuración a la región de Sudamérica para garantizar la reducción de los errores antes descritos, que producen una importante degradación en el pronóstico a plazos de 48 horas en adelante. En este sentido es importante notar que modelos como el ETA que funciona operativamente en el CPTEC, mostraron valores de ETS y bias mejores a los obtenidos con el modelo WRF durante el período de los experimentos presentados en este trabajo de Tesis, reforzando los argumentos antes expuestos. La utilización de un buen modelo en la generación del ensamble es fundamental, dado que el ensamble (salvo que esté compuesto por varios modelos) no va a poder remover las deficiencias del modelo y las ventajas que se puedan obtener mediante la aplicación de esta técnica se pueden perder rápidamente debido a la utilización de un modelo con errores sistemáticos excesivos.

## **7.2 Trabajos futuros:**

Con respecto a las futuras investigaciones, existen numerosos aspectos que pueden ser abordados en la región. En primer lugar es necesaria una comparación más exhaustiva de las diferentes metodologías para la generación de ensambles que permita comparar directamente el SLAF con los vectores criados y con una técnica de vectores criados regional que utilice SLAF para las condiciones de borde únicamente y que no requiera la generación de un ensamble global reduciendo significativamente el costo computacional. A estas técnicas debería sumarse alguna técnica de super ensamble o bien alguna técnica mixta que permita combinar perturbaciones en la condición inicial y en las parametrizaciones del modelo permitiendo medir los impactos que esto trae en las diferentes regiones. La relación entre el error del pronóstico y la dispersión debería ser estudiada en las diferentes alternativas como elemento para determinar aquella que tiene

la mayor potencialidad para introducir mejoras en el pronóstico a corto plazo. Existe una técnica en particular que permite la combinación de los pronósticos por ensamble y la asimilación de datos: el filtro de Kalman local basado en ensambles (LETKF) (Hunt, et al., 2006). Esta metodología no sólo permitiría generar condiciones iniciales para un ensamble regional, sino que provee un método eficiente para la asimilación de datos. En este contexto generar perturbaciones adecuadas para generar el ensamble cobra una renovada importancia, y las diferencias entre el desempeño de las diferentes metodologías sobre la región tropical y extra-tropical se convierte en un importante desafío. Para avanzar en determinar cuales pueden ser las perturbaciones óptimas que sirvan tanto al pronóstico por ensambles como a la asimilación de datos sobre ambas regiones es necesario llevar a cabo experimentos con otros métodos como los llevados adelante por Mendonça y Bonatti (2008). En particular sería interesante evaluar si los vectores criados generados con una menor amplitud pueden capturar las inestabilidades asociadas a los procesos convectivos en la región tropical que son una de las principales fuentes de incertidumbre en la región.

Es importante destacar que los experimentos realizados se limitaron a las primeras 48 horas de pronóstico, principalmente por motivos computacionales y de disponibilidad de almacenamiento. No obstante, dado que los resultados en este período muestran que las ventajas de utilizar pronósticos probabilísticos calibrados obtenidos a partir del ensamble, se incrementan a medida que el plazo de pronóstico aumenta, sería interesante evaluar estas ventajas a plazos mayores. También es importante explorar la capacidad de los ensambles y de los pronósticos individuales para pronosticar la lluvia acumulada en períodos más largos y más cortos de tiempo. En el presente trabajo se utilizaron acumulados a 24 horas de plazo para poder hacer una comparación directa con los datos pluviométricos, pero mediante la utilización de los CMORPH, se pueden estudiar también los acumulados en plazos de tiempo más cortos. Por otro lado el estudio de la precipitación sobre plazos de tiempo mayores (48, 72, 96 horas) podría constituiría una herramienta importante para el pronóstico de inundaciones.

Por otra parte, dado que el super ensamble de modelos funciona operativamente, es necesario avanzar en el desarrollo e implementación de técnicas de calibración más sofisticadas para la generación de pronósticos probabilísticos de precipitación confiables sobre la región, que incrementen el valor económico de la información provista a los usuarios finales. En un primer paso, la utilización del CMORPH para la regionalización de la

calibración podría ofrecer una ventaja importante con respecto a la utilización de los datos de la red pluviométrica operacional, sobre todo en aquellas regiones donde la densidad de estaciones no es adecuada para la calibración de los pronósticos. Las estimaciones CMORPH presentan errores sistemáticos que pueden ser corregidos. Sin embargo es necesario profundizar en el conocimiento de dichos errores y en las metodologías utilizadas para su remoción de forma tal de reducir los aspectos negativos de la utilización de dichas estimaciones en la calibración de los pronósticos probabilísticos. Por otra parte también es necesario realizar experimentos utilizando otras estimaciones satelitales como las mencionadas en Ebert et. al. (2007).

Finalmente y en virtud de los errores sistemáticos encontrados en el modelo WRF para la región de Sudamérica es importante estudiar dichos errores realizando experimentos que permitan aislar las causas de los mismos y eventualmente encontrar mecanismos para su corrección.

## **Apéndice A:**

# **Índices utilizados en la verificación de los pronósticos probabilísticos de precipitación.**

### **Medidas de confiabilidad:**

En el presente trabajo, la confiabilidad asociada a un pronóstico se mide utilizando los diagramas de confiabilidad. Estos diagramas muestran la frecuencia observada de un fenómeno (por ejemplo precipitación por encima de un umbral  $U$ ) como función de la probabilidad pronosticada por el conjunto. Para el cálculo de esta relación se divide el rango de probabilidades pronosticadas en intervalos de 10%; en cada uno de los intervalos, se reúne el conjunto de observaciones que corresponden a los pronósticos que caen dentro de dicho intervalo. Para cada conjunto de observaciones se computa la frecuencia observada de ocurrencia del fenómeno. Luego se grafica la frecuencia observada en función de la probabilidad pronosticada. De esta manera cuanto más se acerque la curva asociada al pronóstico a una relación 1 a 1 entre la frecuencia real y la pronosticada, mayor será la confiabilidad de nuestro pronóstico.

En la mayoría de los casos, los diagramas de confiabilidad están acompañados de un diagrama auxiliar que muestra la cantidad de veces que el pronóstico por conjuntos pronostica cada uno de los intervalos de probabilidad antes definidos (i.e el número de pronósticos como función de la probabilidad pronosticada). Este gráfico es una medida de la resolución del pronóstico, en la medida que permite detectar si las diferentes probabilidades son efectivamente pronosticadas por el conjunto o si existe una tendencia a pronosticar siempre el mismo valor de probabilidad independientemente de lo que ocurra. En particular, un pronóstico perfecto debería tener solo pronósticos de 0% y 100 % de probabilidad.

Por otra parte, una buena herramienta que permite resumir la información contenida en el diagrama de confiabilidad es el índice de Brier (BR) (Brier 1950). Este

parámetro da información sobre todas las características antes mencionadas. La definición se muestra en la Ecuación A.1, en donde vemos que el parámetro se calcula como el promedio de los cuadrados de la resta entre la probabilidad pronosticada y la observada sobre todos los pares pronóstico/observación disponibles. La probabilidad observada se considera 1 cuando el evento tuvo lugar y 0 cuando el evento no tuvo lugar. Estas diferencias se elevan al cuadrado y se promedian en forma análoga a como se realiza en el cálculo del error cuadrático medio que se utiliza frecuentemente en la verificación de los pronósticos de diversas variables. El BR puede interpretarse como la distancia del pronóstico probabilístico del conjunto a un pronóstico probabilístico perfecto que pronostica una probabilidad de 100 % cuando el fenómeno ocurre y de 0 % cuando el fenómeno no ocurre. Usualmente se utiliza una forma del BR en el cual se lo compara con el pronóstico climatológico. Para eso se calcula el BR asociado a la climatología (BR<sub>clim</sub>) y se procede como muestra la Ecuación A.1 y A.2 para obtener lo que se denomina el índice de acierto de Brier (BSS). Este último parámetro muestra cuanta más información aporta el pronóstico del conjunto con respecto al pronóstico de la climatología.

$$BR = \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2 \quad (A.1)$$

$$BSS = 1 - \frac{BR}{BR_{climato\log y}} \quad (A.2)$$

Como describen Hamill y Juras (2006), el BR de la climatología, debe ser calculado cuidadosamente para evitar subestimar su valor. En trabajos previos, el BR de la climatología se calculaba directamente utilizando la probabilidad climatológica de ocurrencia de un evento sobre una región que podía ser muy amplia y sobre un período de tiempo que eventualmente podía abarcar cambios en la probabilidad de ocurrencia del evento. Esto conduce sistemáticamente a una subestimación del BR de la climatología y de acuerdo con la ecuación A.2 a una sobreestimación del BSS. Para evitar este problema, Hamill y Juras (2006) proponen que el BR de la climatología se calcule en forma independiente sobre regiones lo más homogéneas posibles y sobre períodos de tiempo lo más homogéneos posibles y, dependiendo de los datos disponibles, utilizando un período mayor al de verificación. Dado que algunas de las regiones utilizadas en el presente trabajo de tesis son muy amplias y que muestran regímenes diferentes en términos de

probabilidad de ocurrencia de precipitación (Robledo 2007) se separó todo el dominio en cajas de 2.5°x2.5° para lograr un régimen más homogéneo dentro de cada región. El BR de la climatología fue calculado en forma independiente sobre cada una de las cajas utilizando datos de 3 años consecutivos incluyendo el año correspondiente a los pronósticos que fueron verificados. Por otra parte, el BR fue calculado en forma independiente sobre los 3 meses que comprenden el período de verificación (octubre-noviembre-diciembre). Los resultados obtenidos para cada caja y cada mes fueron promediados sobre las regiones utilizadas para la verificación de los pronósticos pesando cada valor por la cantidad de datos utilizados en cada caso de acuerdo con lo sugerido por Hamill y Juras (2006) (dado que la red de observaciones de precipitación sobre la región no es homogénea cada caja contaba con un número diferente de observaciones). De esta manera se garantiza una comparación más justa entre los resultados obtenidos para diferentes regiones.

El pronóstico climatológico por definición tiene una confiabilidad muy alta (buena relación entre la probabilidad pronosticada y la frecuencia observada) pero ninguna resolución (no puede distinguir entre la ocurrencia o no del evento). Valores positivos de BSS indican una mejora del pronóstico del conjunto por sobre la climatología (siendo 1 el valor máximo posible), mientras que los valores negativos indican que no hay aporte extra de información. El BSS puede interpretarse como el incremento porcentual relativo del BR respecto a la climatología. El BR por su parte vale 0 cuando tenemos un pronóstico perfecto.

A su vez, de acuerdo con Murphy (1973), el BR puede ser descompuesto en 3 componentes: confiabilidad, resolución e incertidumbre. Para lograr esta descomposición, primero es necesario particionar en intervalos el rango de probabilidad pronosticada (por ejemplo en 10 intervalos con un largo de 10 % cada uno). En este caso el BR y el BSS pueden escribirse como muestra la ecuación A.3:

$$\begin{aligned}
 BR &= \frac{1}{N} \sum_{k=1}^T n_k (f_k - \bar{o}_k)^2 - \frac{1}{N} \sum_{k=1}^T n_k (\bar{o}_k - \bar{o})^2 + \frac{1}{N} \sum_{k=1}^T \sum_{i=1}^{n_k} (\bar{o} - o_{ki})^2 \\
 BSS &= \frac{-\frac{1}{N} \sum_{k=1}^T n_k (f_k - \bar{o}_k)^2 + \frac{1}{N} \sum_{k=1}^T n_k (\bar{o}_k - \bar{o})^2}{BR_{climatología}} \tag{A.3}
 \end{aligned}$$

Donde  $n_k$  es el número de pronósticos en el k-ésimo intervalo (por ejemplo probabilidad pronosticada entre 10% y 20 %).  $f_k$  es el valor medio del intervalo (por ejemplo 15 %). El valor de  $o_{ki}$  representa la probabilidad asociada a la ocurrencia del fenómeno para cada evento  $i$  dentro de la categoría  $k$  y se toma según la convención del cálculo del BR.  $\bar{o}_k$  es la frecuencia observada dentro de la categoría k-ésima y  $\bar{o}$  es la probabilidad de ocurrencia del fenómeno independientemente de la categoría que si calculáramos sobre un período lo suficientemente largo de tiempo tendería a la probabilidad climatológica de ocurrencia del fenómeno. El primer término del lado derecho de la descomposición del BR es una medida de la confiabilidad, ya que es la suma del cuadrado de la distancia de la curva del diagrama de confiabilidad a la curva del conjunto ideal (diagonal), pesada por la cantidad de veces que se pronosticó cada una de las categorías. El segundo término es una medida de la resolución ya que mide la diferencia entre la frecuencia observada asociada a cada categoría y la frecuencia observada del fenómeno independientemente del valor pronosticado por el conjunto, pesada a su vez por el número de casos dentro de cada categoría. Cuanto mayor es en módulo este término mayor es la resolución del pronóstico. Finalmente el tercer término del lado derecho de la ecuación es el término de incertidumbre, se puede demostrar que este término se puede escribir como  $\bar{o}(1 - \bar{o})$  es decir que depende exclusivamente de la probabilidad de ocurrencia del fenómeno que se quiere pronosticar y además que es igual al BR de la climatología que aparece en la formulación del BSS. Este término es máximo (mayor incertidumbre) cuando la probabilidad de ocurrencia del fenómeno es 50% es decir que existe la misma probabilidad de que el fenómeno ocurra o que no ocurra.

Es interesante notar que si incrementamos el número de categorías o intervalos de forma tal de tener un solo pronóstico en cada categoría entonces los términos 2 y 3 de la descomposición del BR, se cancelan y el término 1 pasa a tomar la forma de la definición del BR. En particular si analizamos el pronóstico basado en la "climatología" entonces para el primer término debemos notar que solo vamos a poblar aquel intervalo que contenga a la probabilidad climatológica de ocurrencia del fenómeno. El primer término será cercano a 0 debido a que en dicho intervalo la probabilidad observada será cercana al valor medio del intervalo, la resolución será 0 debido a que el pronóstico climatológico es incapaz de distinguir entre la ocurrencia o no ocurrencia de un evento y la incertidumbre no se

modificará porque no depende del pronóstico considerado. Es decir que el BR de un pronóstico supera al de la climatología cuando el segundo término es mayor en módulo al primero. Es decir que la frecuencia media observada en cada categoría está más cerca de la probabilidad central del intervalo que de la probabilidad de ocurrencia del fenómeno (Wilks, 1995).

En la Ecuación A.3 se muestra además la descomposición aplicada al BSS, en este caso, la componente asociada a la incertidumbre desaparece y solo quedan las componentes de confiabilidad y resolución.

La descomposición propuesta en la ecuación A.3 está basada en analizar, dado un valor de probabilidad pronosticada, cuáles son los valores de probabilidad observados. Existe una aproximación alternativa que es, dados los valores de probabilidad observada, cuáles son los valores de probabilidad pronosticados, que permite una descomposición diferente del BR (Murphy y Winkler 1987, Roulin y Vannitsem 2005).

### **Medidas de resolución y habilidad:**

Para cuantificar la exactitud del pronóstico probabilístico, primero es necesario utilizar el mismo como un pronóstico determinístico. Para eso, es necesario convertir un valor de probabilidad  $P$  en un pronóstico que determine la SI / NO ocurrencia de un fenómeno. Esto se logra utilizando un umbral de probabilidad  $U_p$  arbitrario (que puede ser modificado para la evaluación del pronóstico como se discutirá más adelante). Si la probabilidad pronosticada es superior al valor  $U_p$  entonces se considerará un pronóstico positivo, de lo contrario se considera un pronóstico negativo. Dado que las probabilidades están formuladas como la probabilidad de que ocurra precipitación por encima de un determinado umbral, el pronóstico determinístico nos pronosticará la SI/NO ocurrencia de la precipitación por encima de dicho umbral de lluvia. Es por eso que en este caso se trata a la precipitación como una variable dicotómica y la misma se verifica utilizando herramientas que se apliquen a este tipo de variables. Dado un pronóstico por si o por no, pueden ocurrir 4 cosas:

- Acierto positivo (A): El fenómeno es pronosticado y ocurre.
- Falsa alarma (FA): El fenómeno es pronosticado pero no ocurre.
- Sorpresa (S): El fenómeno no es pronosticado pero ocurre.

- Acierto negativo (NA): El fenómeno no es pronosticado y no ocurre.

La habilidad del pronóstico se evalúa utilizando el índice Equitable Threat Score (ETS) (Schaefer, 1990, Hamill, 1999) que se define con las Ecuaciones A.5 a A.7. Este índice, tiene en cuenta los aciertos, pero también incorpora información sobre las falsas alarmas, las sorpresas e identifica qué porcentaje de los aciertos podrían haberse dado por azar. La cantidad de aciertos que se dan por azar depende de la climatología de la región considerada y de la cantidad de veces que el modelo pronostica la ocurrencia del evento. El índice ETS varía de  $-1/3$  a 1 siendo 1 el valor ideal.

Como se describe en Hamill (1999) el ETS es dependiente del bias de los pronósticos. El bias se calcula en este caso como la relación entre la frecuencia pronosticada para un determinado umbral y la frecuencia observada de dicho umbral. Es decir que varía entre más infinito a 0 siendo 1 el valor óptimo. La dependencia del ETS con el bias es tal que el ETS de un pronóstico puede ser incrementado o reducido simplemente modificando el bias sistemático del mismo. Esto hace que la comparación utilizando el ETS de 2 pronósticos cuyos bias sean muy diferentes no sea válida. Es por eso que en el presente trabajo para analizar la habilidad de los diferentes pronósticos se utiliza un diagrama ETS-Bias en donde los ETS de los diferentes pronósticos se grafican como función del bias de los mismos. La aplicación de este diagrama al presente trabajo se discutirá más adelante.

$$ETS = \frac{aciertos - aciertos_r}{aciertos + sorpresas + falsa\ alarma - aciertos_r} \quad (A.5)$$

$$aciertos_r = \frac{(aciertos + sorpresas)(aciertos + falsa\ alarma)}{total}$$

$$tasa\ de\ acierto = \frac{aciertos}{aciertos + falsas\ alarmas} \quad (A.6)$$

$$coeficiente\ de\ falsa\ alarma = \frac{falsa\ alarma}{pronosticos\ positivos} \quad (A.7)$$

También se define la tasa de acierto como el cociente entre los aciertos y el total

de pronósticos positivos (Ecuación A.6) (es decir aquellos en los que se pronosticó la ocurrencia del fenómeno que se obtiene sumando los aciertos y las falsas alarmas). Este coeficiente no tiene en cuenta los aciertos negativos (NA). En el tipo de régimen de lluvias preponderante sobre la región, el número de días con lluvia al año es menor que el número de días con no lluvia. Esto tiende a producir un valor elevado de aciertos negativos que no necesariamente están vinculados con un pronóstico eficiente.

Para evaluar la incidencia de las falsas alarmas dentro del pronóstico se puede calcular el coeficiente de falsa alarma (far) definido en la Ecuación A.7 como el cociente entre las falsas alarmas y la totalidad de los pronósticos positivos. Este índice representa la probabilidad de que el fenómeno no se produjera dado que el pronóstico indicaba que se iba a producir. Esta definición es la misma que la utilizada por Hou et al. (2001).

La tasa de acierto y el far se utilizan para la construcción del diagrama ROC (Relative Operating Characteristics). Como se discutió previamente, la tasa de acierto y el far están definidos para un pronóstico determinístico de una variable dicotómica. Para obtener un pronóstico determinístico a partir de un pronóstico probabilístico, podemos elegir un valor de probabilidad ( $U_p$ ) tal que si la probabilidad está por encima de dicho valor entonces se está pronosticando la ocurrencia del evento y si es menor no se está pronosticando la ocurrencia del evento. Una vez hecho esto podemos calcular el valor de HR y far para el pronóstico determinístico así obtenido. En este caso, podemos tomar diferentes valores de  $U_p$  y para cada valor de  $U_p$  vamos a obtener un valor de tasa de acierto y un valor de far. El diagrama ROC consiste justamente en visualizar la tasa de acierto como función de la far para los diferentes valores de  $U_p$ . La curva obtenida se denomina curva ROC y es una medida de la resolución y de la exactitud del pronóstico probabilístico.

Usualmente se resume la información contenida en el diagrama ROC mediante el cálculo del área bajo la curva del diagrama. Cuanto mayor es el área, mayores son los coeficientes de acierto asociados al pronóstico y por tanto mayor la exactitud para pronosticar el fenómeno de interés. Sin embargo para el cálculo del área se debe asumir que la curva del ROC comienza en la esquina inferior izquierda del gráfico (far = 0 tasa de acierto = 0 es decir el caso extremo en donde el pronóstico no indica la ocurrencia del fenómeno) y termina en la esquina superior derecha (far = 1, tasa de acierto = 1, el pronóstico siempre indica la ocurrencia del fenómeno y por ende el acierto es 1, pero las

falsas alarmas también). En función del valor del área bajo la curva se puede definir el índice ROCS (ROC area based skill score) según Richardson (2000) (Ecuación A.8), donde ROCA denota el valor del área bajo la curva del diagrama ROC.

La resta  $ROCA - 0.5$  es una medida de la precisión del pronóstico, ya que se puede demostrar que un pronóstico en donde la tasa de acierto es igual a la falsa alarma, equivale a un pronóstico que no tiene ninguna habilidad para distinguir entre la ocurrencia o no de un fenómeno (por ejemplo, si construimos varios pronósticos por azar en donde vamos modificando la cantidad de veces en que el pronóstico es positivo, la curva ROC asociada a dichos pronósticos sería la diagonal del diagrama que va desde la esquina inferior izquierda a la superior derecha). Es por eso que este índice le resta al área del ROC el área del pronóstico por azar.

$$ROCS = 2(ROCA - 0.5) \quad (A.8)$$

### **Medidas del valor económico del pronóstico.**

Otro aspecto a tener en cuenta para la evaluación de pronósticos, son las medidas del valor del pronóstico. Estas medidas están orientadas a cuantificar en términos relativos, el valor de la información meteorológica cuando es utilizada por diferentes usuarios en el proceso de toma de decisiones. El algoritmo por el cual el usuario de la información meteorológica toma una determinada decisión, depende de cada aplicación en particular y puede involucrar diferentes grados de complejidad (por ejemplo puede depender de una o de varias variables meteorológicas, de forma continua o discreta), algunos ejemplos de aplicaciones directas se pueden encontrar en Roebber y Bosart (1996).

En el presente trabajo, vamos a utilizar uno de los modelos más sencillos de toma de decisión y el más frecuentemente utilizado en la literatura para la evaluación de los pronósticos, es aquel en el cual la decisión depende de una sola variable meteorológica, y se toma cuando el pronóstico indica que la misma estará por encima de un determinado umbral. Esta situación, describe el caso de un tipo particular de usuario, aquel que frente a la ocurrencia de un fenómeno que es perjudicial para su actividad, puede tomar alguna precaución que minimice los efectos negativos de dicho fenómeno. Este sería el caso por

ejemplo de la ocurrencia de heladas en las que un productor ante un pronóstico que indica la ocurrencia de heladas, toma una acción preventiva para evitar la pérdida que la helada podría ocasionarle. Dentro de este esquema de toma de decisiones, se pueden dar las 4 posibilidades descritas previamente (acierto, falsa alarma, sorpresa y acierto negativo). Para cuantificar el valor relativo de la información meteorológica, a cada una de estas posibilidades se les asigna un valor económico. En el caso del acierto, si el pronóstico indica la ocurrencia del fenómeno y el mismo se produce, hay un costo asociado que es el costo de protección (C), en algunos casos a este costo se le agrega un factor de pérdida no evitable (U) que representa aquella parte de los efectos negativos que no pueden ser contrarrestados. La falsa alarma, tiene únicamente asociado el costo de protección (C). Los aciertos negativos, representan aquellos casos en los que el fenómeno no fue pronosticado y no se presenta, y por ende la actividad se desarrolla normalmente, en este caso se supone que dependiendo de la actividad existe un cierto rédito, el cual no es considerado en la evaluación del valor ya que es independiente de la información meteorológica provista al usuario, con lo cual normalmente se lo considera como 0. Finalmente la sorpresa tiene asociado una pérdida (L) porque el fenómeno no fue pronosticado y no se tomaron las acciones preventivas correspondientes.

La pérdida vinculada con la ocurrencia del fenómeno meteorológico de interés puede ser calculada de esta forma en base a una tabla de contingencia como la utilizada para definir el ETS. Simplemente multiplicando el costo asociado a los aciertos, sorpresas y falsas alarmas por su ocurrencia empírica en el pronóstico analizado. La Ecuación A.9, muestra por ejemplo cual es el costo medio ( $E_f$ ) asociado a un pronóstico donde N es el número total de pronósticos evaluados. El valor obtenido es siempre negativo, porque como se mencionó anteriormente el rédito obtenido por la actividad no es tenido en cuenta y solo se analiza la pérdida ocasionada por la ocurrencia de un fenómeno y su reducción por parte de la utilización de un pronóstico y algún método de prevención. Otra simplificación que frecuentemente se hace es suponer que el método de prevención es perfecto y que por ende U es 0.

$$E_f = \frac{1}{N} [\text{acierto}(C + U) + \text{falsa\_alarma}(C) + \text{sorpresa}(L)] \quad (\text{A.9})$$

Por lo general la calidad de los pronósticos en diversos aspectos es comparada

contra la climatología. En el caso de la estimación del pronóstico, de acuerdo con Zhu et al. 2002, el valor del pronóstico por climatología se expresa en la Ecuación A.10 donde  $o$  es la probabilidad climatológica de ocurrencia del evento en cuestión.

$$E_c = oU + \min[oL, C] \quad (A.10)$$

De acuerdo con la Ecuación A.10, en el caso del valor del pronóstico por climatología, pueden existir dos situaciones, una en la cual la probabilidad de ocurrencia del fenómeno sea tan baja que resulte más conveniente no tomar ninguna protección frente al evento, y otra situación donde el evento ocurre tan frecuentemente que lo más conveniente es protegerse siempre. El umbral que separa estos dos comportamientos es función de la relación entre  $L$  y  $C$  y es  $o_u = C/L$ . Por otra parte el valor económico asociado a un pronóstico perfecto ( $E_p$ ) se puede expresar como lo indica la Ecuación A.11.

$$E_p = o(C + U) \quad (A.11)$$

Usualmente se utiliza una expresión que combina las ecuaciones A.9-A.11 y que se denomina valor relativo del pronóstico que asumiendo que el método de protección es perfecto se puede expresar como en la Ecuación A.12 (Zhu et al. 2002), donde  $r$  es el cociente entre  $C$  y  $L$ .

$$E = \frac{E_c - E_f}{E_c - E_p} = \frac{\min(o, r) - \frac{1}{N}(\text{acierto} + \text{falsa\_alarma})r - \frac{1}{N}\text{sorpresa}}{\min(o, r) - or} \quad (A.12)$$

Como se desprende de la Ecuación A.12, el valor relativo del pronóstico vale 1 cuando el pronóstico tiene un valor igual al del pronóstico perfecto, y vale 0 si su valor es igual al de la climatología. Por otra parte es importante destacar que en este marco sencillo, la sensibilidad a los costos de diferentes usuarios se resume en el parámetro  $r$  que mide la relación entre el costo de protección y la pérdida que ocasiona el fenómeno. Es por eso que este parámetro ( $E$ ) generalmente se grafica como función de  $r$  para describir como es el valor de un pronóstico determinado para diferentes usuarios con

diferentes relaciones entre C y L.

El valor del pronóstico depende directamente de la habilidad del mismo para poder distinguir entre la ocurrencia y no ocurrencia de un evento. Por ende es una medida de la resolución del pronóstico. El presente marco conceptual tiene importantes limitaciones para estimar el verdadero valor del pronóstico, en primer lugar porque todo el análisis se aplica a una categoría de usuarios muy particular que claramente no agota todas las posibles aplicaciones de la información, por otro lado aún en el caso de los usuarios que supone este análisis, se simplifican diversos factores como la posible dependencia de la pérdida con la magnitud del error cometido en el pronóstico o la intensidad del fenómeno. En algunos casos esto puede ser un factor importante, para un productor sujeto al riesgo de heladas no es lo mismo el daño producido por una helada corta y débil que el producido por una helada prolongada e intensa. En este ejemplo y en otros casos la pérdida es función de la magnitud del error cometido en el pronóstico, y este planteo falla en considerar dicha dependencia (que por otra parte es distinta para cada ejemplo considerado y depende también de la actividad realizada).

## Apéndice B:

### Experimentos con un modelo de 3 grados de libertad.

Los experimentos discutidos en el Capítulo 4, mostraron que si bien las estimaciones del error del análisis tenían una amplitud y una distribución espacial promedio similar a la amplitud y distribución espacial de los vectores criados, en cada caso particular, no había mucha relación entre la amplitud de las perturbaciones y el error del análisis, aunque si se observaba alguna similitud en los patrones espaciales, sobre todo cuando el error en el análisis era grande. Una de las principales limitaciones para poder estudiar la relación entre el error del análisis y los vectores criados es que los primeros sólo pueden ser estimados - en nuestro caso a través de calcular las diferencias entre ciclos de análisis independientes-; es por eso que resulta de interés investigar la relación existente entre los vectores criados (VC) y los errores del análisis en un sistema en donde podamos conocer con exactitud los errores del análisis y que, además, tenga menos grados de libertad haciendo más sencilla la interpretación y visualización de los resultados. Para eso se recurrió al modelo de 3 variables de Lorenz (1969), más específicamente se utilizó una implementación en Matlab que fue empleada por Yang et al. (2006), Evans et al. (2004) entre otros y que integra el modelo con un paso de tiempo de 0.01 segundos utilizando un método de Runge Kutta de cuarto orden. El modelo se describe en la Ecuación B1, donde  $X=(X_1, X_2, X_3)$  es el vector estado del sistema y  $r$ ,  $\sigma$  y  $b$  son parámetros constantes ( $r=28$ ,  $\sigma=10$  y  $b=8/3$ ).

$$\begin{aligned}\frac{dX_1}{dt} &= \sigma(X_2 - X_1) \\ \frac{dX_2}{dt} &= rX_1 - X_2 - X_1X_3 \\ \frac{dX_3}{dt} &= X_1X_2 - bX_3\end{aligned}\tag{B1}$$

Como ha sido extensamente discutido en la bibliografía este modelo presenta un comportamiento caótico para los valores seleccionados de los parámetros (Kalnay, 2003 y referencias allí citadas).

El procedimiento utilizado fue el siguiente: se integró el modelo de Lorenz partiendo de una condición inicial arbitraria hasta que la solución converge al atractor. Una vez en esta región, se realizó una simulación por un período de 1600 segundos (equivalente a  $16 \cdot 10^4$  pasos de integración) la cual es utilizada como la evolución real del sistema ( $X_v$ ), es decir que se utiliza la hipótesis del modelo perfecto. Paralelamente, se agrega una pequeña perturbación a la condición inicial y se inicia un ciclo de asimilación de datos que opera cada 8 pasos de tiempo (en total se realizaron 20000 ciclos de asimilación) incorporando pseudo-observaciones que se generan como perturbaciones aleatorias de la simulación que se utiliza como verdad y que dan lugar a un conjunto de análisis ( $X_a$ ). El método de asimilación utilizado es el de interpolación óptima y su implementación para el modelo de Lorenz fue realizada por Yang et al.(2006). La matriz de covarianza de los errores del modelo es constante a lo largo de los ciclos de asimilación realizados y su valor se obtuvo en base a estadísticas de los errores en el modelo de Lorenz. En cada asimilación, se generaron y asimilaron observaciones para las 3 variables del modelo.

A la par de que se generó este ciclo de asimilación, se puso en marcha un ciclo de generación óptima de perturbaciones utilizando el método de los vectores criados (VC). Una única perturbación fue generada a lo largo de todo el período abarcado por el experimento y aplicando una reducción en la amplitud de las perturbaciones cada 8 pasos de tiempo, es decir que coincide con la frecuencia de la asimilación de datos. Esta reducción se llevó a cabo multiplicando a la perturbación por el cociente entre la amplitud media del error del análisis (que fue calculado previamente) y la amplitud de la perturbación luego de la integración. En el caso en que dicho cociente fuera mayor que uno se fija el valor del cociente en uno. De esta manera, la amplitud de la perturbación resultante es similar a la magnitud de los errores en el análisis (que en este caso sí pueden ser calculados dado que asumiendo la hipótesis del modelo perfecto conocemos la verdadera evolución del sistema), pero la amplitud de la perturbación nunca es incrementada, sino que por el contrario se reduce para que no supere a la incertidumbre media de la condición inicial (Toth y Kalnay, 1997).

Por otra parte, se realizó un tercer ciclo en donde se perturbó el análisis agregando

una perturbación al azar, cuya magnitud es similar a los errores presentes en el análisis, pero a diferencia de lo que sucede con el vector criado, las perturbaciones de dos análisis sucesivos son totalmente independientes. Esta perturbación al azar se construye generando un número aleatorio a partir de una distribución normal con media 0 y desviación estándar igual a la estimación del error del análisis. En cada ciclo se generan 3 números al azar (uno para cada variable del modelo).

Para simplificar el análisis, los resultados se muestran en función de la posición en el plano  $X_1$ - $X_2$  a pesar de que las trayectorias del modelo son tridimensionales. La mayoría de los resultados se muestran como función de la posición en el atractor, dado que es conocido que en el mismo existen regiones más inestables, es decir donde los errores tienen mayor capacidad de crecimiento (Evans et al. 2004).

Para visualizar los resultados, el plano  $X_1$ - $X_2$  se dividió en cajas de  $0.5 \times 0.5$  y dentro de cada caja se calcularon los errores en el campo preliminar, la amplitud de los vectores criados y sus respectivas tasas de crecimiento en el tiempo transcurrido entre dos sucesivas asimilaciones. La forma en la que se definieron las diferentes cantidades se muestra en las Ecuaciones agrupadas en B2, donde  $\varepsilon_b(t)$  es el error en el campo preliminar que se computa como la diferencia entre el campo preliminar y la verdad en el tiempo  $t$ ,  $\varepsilon_a(t)$  es el error en el análisis, que se computa como la diferencia entre el análisis y la verdad en el tiempo  $t$ ,  $p_{vc}(t)$  es la diferencia entre el pronóstico a 8 pasos de tiempo ( $X_f(t)$ ) inicializado a partir del análisis en  $(t-1)$  y el pronóstico inicializado a partir del análisis perturbado con el vector criado en el tiempo  $t-1$  ( $X_{fvc}(t)$ ), ambos verificando en el tiempo  $t$ . La tasa de crecimiento del error en el pronóstico ( $A_b$ ) se define como el cociente entre el error en el campo preliminar sobre el error del análisis en el tiempo anterior. La tasa de crecimiento del error del análisis ( $A_a$ ), se define como el cociente entre el error del análisis en el tiempo  $t$  y el error del análisis en el tiempo  $t-1$ . Finalmente, la tasa de crecimiento de los vectores criados ( $A_{vc}$ ), se mide como la diferencia entre el pronóstico control y el pronóstico perturbado con el vector criado, sobre la amplitud del vector criado en el tiempo anterior. Una vez que se obtuvieron las tasas de crecimiento de los vectores criados, de los errores del análisis y de los errores en el campo preliminar dentro de cada caja, se definieron 3 categorías de crecimiento: alto, medio y bajo. Los umbrales se determinaron a partir de considerar los terciles de la distribución de tasas de crecimiento de los vectores criados, los errores del análisis y los errores en el campo preliminar para

todo el experimento (independientemente de la posición). Es importante aclarar que los umbrales se calcularon en forma independiente para cada una de estas 3 cantidades. En los resultados que se muestran a continuación, se hizo hincapié en los casos de mayor crecimiento de los errores del pronóstico y del análisis que son los que representan mayor interés en el presente trabajo de tesis.

$$\begin{aligned}
 \varepsilon_b(t) &= X_b(t) - X_v(t) \\
 \varepsilon_a(t) &= X_a(t) - X_v(t) \\
 p_{vc}(t) &= X_f(t) - X_{fvc}(t) \\
 A_b(t) &= \frac{|\varepsilon_b(t)|}{|\varepsilon_a(t-1)|} \\
 A_a(t) &= \frac{|\varepsilon_a(t)|}{|\varepsilon_a(t-1)|} \\
 A_{vc}(t) &= \frac{|p_{vc}(t)|}{|VC(t-1)|}
 \end{aligned} \tag{B2}$$

La Figura B.1 (a y b), muestra la probabilidad de que la tasa de crecimiento en los errores del campo preliminar y el crecimiento de los vectores criados sea alta. Como se puede apreciar y de acuerdo con lo mencionado anteriormente, existen regiones preferenciales para el crecimiento del error en  $X_b$ , las cuales se corresponden estrechamente con el crecimiento en la magnitud de los VC. Existen por otra parte amplias zonas del atractor donde la probabilidad de que el crecimiento del error en  $X_b$  o de los VC sea grande, es muy pequeña. Esta similitud entre el crecimiento de los VC y de los errores de los pronósticos es consistente con los resultados encontrados por Corazza et al. (2002) que mediante un experimento utilizando un modelo cuasigeostrófico, mostraron que existía una estrecha relación entre la amplitud y estructura de los vectores criados y los errores en los pronósticos a muy corto plazo. Esto tiene una importancia muy grande para la mejora de la asimilación de datos y es uno de los pilares teóricos en los que se sustenta la utilización de ensambles para obtener una estimación de la matriz de covarianza de los errores del modelo que tenga información de las inestabilidades del flujo. El período entre dos asimilaciones es comparable con los experimentos mostrados en este capítulo, dado que si bien los modelos describen sistemas diferentes, el plazo entre dos sucesivas asimilaciones se eligió en forma tal que el crecimiento del error fuera

similar al que se observa en un período de 6 horas en los modelos que simulan el comportamiento de la atmósfera.

La Figura B.2, muestra la probabilidad de tener tasas de crecimiento altas como función de la posición en el atractor para las perturbaciones al azar. En este caso, si bien se sigue observando un máximo de probabilidad similar al observado para los errores en los pronósticos a corto plazo y para los VC, los valores de probabilidad no se asemejan tanto a la distribución observada en la Figura B1a. Esto sugiere que los VC reproducen mejor el comportamiento de los errores en el pronóstico que resultan de las inestabilidades del flujo, confirmando resultados previos (Corazza et al. 2003).

Por otra parte, es interesante ver como es el comportamiento de los errores del análisis ( $\epsilon_a$ ) y de la tasa de crecimiento de los mismos ( $A_a$ ) que se definen en la Ecuación B.2. Dado que el análisis es una combinación óptima de pronósticos a corto plazo y observaciones, es natural pensar que los errores presentes en el análisis van a estar en parte determinados por los errores en el campo preliminar. La Figura B.3, muestra la tasa de crecimiento del error del análisis como función de la posición del sistema en el plano  $X_1$ - $X_2$ . A diferencia de lo que se observa en las Figuras B.1 a y b, la Figura B.3 no muestra valores tan altos de probabilidad de tener un crecimiento alto de los errores, si bien los máximos de probabilidad tienen cierta similitud con las regiones donde la probabilidad de crecimiento alto en el error del pronóstico es máxima. Por otra parte, en regiones del atractor donde la probabilidad de observar un crecimiento alto en el error del pronóstico es muy baja, el error del análisis muestra una mayor probabilidad de tener errores altos. Este comportamiento sugiere que en estas regiones, el crecimiento en los errores del análisis esté gobernado principalmente por el error en las observaciones.

A continuación, se busca describir la relación existente entre las regiones del plano  $X_1$ - $X_2$  en donde se observa el máximo crecimiento de los errores y las perturbaciones y aquellas donde ocurren los máximos valores de error en  $X_b$  y en  $X_a$ . La Figura B.4 (a y b), muestra la distribución dentro del atractor de la probabilidad de tener un error alto en el análisis y en el pronóstico (respectivamente). Las regiones donde la probabilidad de tener un error alto en el análisis son mayores (Figura B.4a), no coinciden exactamente con la región en donde se observan las mayores probabilidades de que la tasa de crecimiento del error en el pronóstico sea alta, sino que se encuentran desplazadas "corriente abajo" de las mismas, esto se puede ver analizando la dirección media de desplazamiento del

sistema en el plano  $X_1$ - $X_2$  que está dada por los vectores en la Figura B.4 a. Esta figura ejemplifica cómo un error en el análisis se amplifica debido a las inestabilidades del flujo en una determinada región del atractor, pero luego se propaga a otras regiones del atractor debido a que un error elevado en el análisis es más probable que genere un error alto en el pronóstico a corto plazo, aunque la región del atractor no sea tan favorable para el crecimiento del error. Dicho error en el pronóstico a corto plazo impactará negativamente en el análisis del ciclo siguiente. Esta situación se va corrigiendo paulatinamente a medida que la incorporación de observaciones reduce el error del análisis. Como se puede ver en la Figura B.4b, la probabilidad de encontrar errores altos en el campo preliminar sigue una distribución similar a la de los errores del análisis encontrándose desplazada corriente abajo de las regiones de máximo crecimiento de los errores. Por otra parte, la Figura B.5, muestra la probabilidad de tener un VC con una amplitud alta. Las regiones donde la probabilidad es mayor coinciden con las regiones donde los errores en el análisis y en el campo preliminar son elevados.

De acuerdo con los resultados obtenidos en el Capítulo 4, es interesante ver si en los experimentos con menos grados de libertad y bajo la hipótesis de modelo perfecto se puede encontrar una correspondencia entre la estructura (orientación) de los VC y la de los errores en el campo preliminar como así también entre la estructura de los VC y los errores en el análisis. Por otra parte resulta interesante verificar si dicha relación depende de la amplitud del error como lo sugieren los resultados obtenidos en el Capítulo 4. Para eso se calculó el coseno del ángulo comprendido entre el VC y el error del campo preliminar y entre el VC y el error del análisis. La Ecuación B3, muestra la expresión matemática utilizada, donde  $\alpha$  es el ángulo comprendido entre el VC y el error del campo preliminar; para calcular el ángulo correspondiente entre el error del análisis y el VC se utiliza la misma expresión, pero reemplazando  $\varepsilon_b$  por  $\varepsilon_a$ . Un valor de 0 para el coseno del ángulo comprendido entre ambos vectores estaría indicando que los vectores son perpendiculares con lo cual no existe relación entre ambos, mientras que un valor de 1 indica que ambos vectores están en la misma dirección (aunque podrían tener módulos y sentido diferentes). Los valores negativos cercanos a 1 también indican que la dirección de los vectores es similar, pero su sentido es opuesto. Como el sentido de los vectores es arbitrario, vamos a considerar únicamente el valor del módulo del coseno del ángulo, ya que tanto valores cercanos a -1 como valores cercanos a 1 indican una buena

correspondencia en la dirección de ambos vectores que es lo que nos interesa evaluar.

$$\cos(\alpha) = \frac{VC \cdot \varepsilon_b}{|VC| |\varepsilon_b|} \quad (B.3)$$

El producto interno a diferencia de los cálculos anteriores, mide una correspondencia caso por caso, es decir que no basta que los VC crezcan en las mismas regiones donde crecen los errores del análisis, sino que además en cada caso particular se mide cuánto se parecen entre sí ambos vectores. La Figura B.6, muestra el valor medio del módulo de la proyección como función de la posición del sistema en el plano  $X_1$ - $X_2$ . En este caso, se toma el módulo del producto interno porque no estamos interesados en determinar el sentido de los errores, sino la dirección en la que se producen. Comparando con la Figura B.4, se puede apreciar que los valores más altos de correspondencia entre la dirección del VC y de los errores tanto de  $X_b$  como de  $X_a$ , coinciden a grandes rasgos con las regiones donde la probabilidad de tener un error alto es mayor, esto significa que a mayor error, los vectores criados resultan más eficientes en capturar los mecanismos que dieron origen a dicho error. Este comportamiento es similar al observado en el experimento realizado en el Capítulo 4, dado que en dicho experimento se observó que la correlación lineal espacial máxima día a día entre los vectores criados y la estimación de los errores del análisis era mayor para los días en los que la magnitud de los errores en el análisis era grande. La proyección entre los errores en el análisis y los VC resulta menor que la observada entre el campo preliminar y los VC, lo cual era esperable dado que en los errores del análisis intervienen otros factores (como por ejemplo el error en las observaciones) que no dependen de las inestabilidades del flujo. El comportamiento de la proyección como función de la amplitud del error en el análisis se sintetiza en la Tabla B.I, en donde se muestra la variación de la media del módulo del coseno del ángulo comprendido entre el VC y el error en  $X_a$  para 3 categorías diferentes de error en  $X_a$ . Los resultados muestran que cuanto mayor es el error en  $X_a$ , más similares son las direcciones en las que apunta VC y el error en  $X_a$ . Por otra parte, los resultados obtenidos se comparan con el coseno del ángulo comprendido entre el error del análisis y la perturbación obtenida a partir de introducir un error al azar en el análisis. En este caso,

los valores de proyección son menores que las obtenidas con el VC, aunque persiste cierta sensibilidad a la magnitud del error en el análisis. Esto probablemente se debe a que dado que los grados de libertad del modelo utilizado son pocos, aún una perturbación al azar puede proyectarse con cierta eficiencia sobre las direcciones de máximo crecimiento, pero las probabilidades de que esto ocurra deberían ser menores a medida que aumentan los grados de libertad del sistema.

Para complementar este análisis, se dividió a las proyecciones obtenidas en 3 categorías (baja, media y alta) de manera similar a como se procedió con los errores y con las tasas de crecimiento, y se graficó la probabilidad de tener una proyección alta entre el VC y los diferentes errores como función de la posición en el plano  $X_1$ - $X_2$ . En la Figura B.7 se puede apreciar más claramente que la probabilidad de que la proyección sea alta es mayor en las regiones donde los errores en  $X_b$  son mayores. También, es importante notar que donde los errores no muestran un crecimiento elevado, la probabilidad de que ambos vectores estén en una dirección similar es casi nula, es decir que cuando el error en el análisis no está dominado por las inestabilidades del flujo, la estructura de los VC no tiene por qué parecerse a los errores en el campo preliminar.

La Figura B.8, muestra, por otra parte, la probabilidad de tener un valor de proyección alto entre la perturbación al azar y el error en el análisis. Si bien todavía pueden observarse probabilidades mayores en las regiones del plano  $X_1$ - $X_2$  donde los errores de  $X_a$  resultaron mayores, la probabilidad de que la proyección sea elevada es menor que en el caso de los VC. Además hay que tener en cuenta que la proyección media entre estas perturbaciones y el error en  $X_a$  es menor que en el caso del VC, es decir que una proyección "alta" en este caso puede significar un valor de proyección menor que lo que se considera una proyección "alta" para los VC.

En conclusión, los experimentos realizados permitieron encontrar que los VC guardan una estrecha relación con los errores presentes en el análisis o "errores del día", ya que representan muy adecuadamente el comportamiento de los errores en el campo preliminar. No obstante un factor que puede degradar la relación entre los VC y los errores del análisis, son los errores del modelo. Este aspecto no fue considerado en los experimentos idealizados, dado que se ha abordado desde diversas perspectivas en los casos reales estudiados en el Capítulo 4 de la presente tesis. Los VC no pueden capturar los errores del modelo, y cualquier error en el análisis que derive de dichos errores no va a

estar vinculado con los VC, con lo cual este factor contribuye a debilitar la relación entre ambos.

En experimentos futuros, se indagará sobre la relación existente entre el error del análisis y los VC utilizando métodos de asimilación que tengan en cuenta las variaciones de la matriz de covarianza de los errores como función de su posición en el atractor. Es de esperar que en este caso, mejore la relación entre los vectores criados y los errores del análisis en la medida en que estos respondan más estrechamente a las inestabilidades del flujo.

Error en Xa	Azar	VC
Bajo	0.56	0.59
Medio	0.59	0.62
Alto	0.64	0.71

**Tabla B.I: Proyección media entre la perturbación al azar y los errores del análisis y entre el VC y los errores del análisis.**

## Apéndice C:

### Configuración utilizada en el modelo numérico WRF en los experimentos utilizando la técnica de vectores criados.

#### Archivo de opciones del modelo (namelist):

A continuación se muestran las diferentes variables que se definen en el archivo de configuración del modelo WRF, tal cual fueron definidas para los experimentos utilizando la técnica de los vectores criados.

&domains	
Time_step	= 150,
Time_step_fract_num	= 0,
Time_step_fract_den	= 1,
max_dom	= 1,
s_we	= 1,
e_we	= 150,
s_sn	= 1,
e_sn	= 150,
s_vert	= 1,
e_vert	= 36,
num_metgrid_levels	= 32
dx	= 40000,
dy	= 40000,
Grid_id	= 1,
parent_id	= 0,
num_metgrid_levels	= 24
interp_type	= 2
zap_close_levels	= 500
lowest_lev_from_sfc	= .false.
force_sfc_in_vinterp	= 1
p_top_requested	= 5000
eta_levels	= 1.000, 0.983, 0.965, 0.947, 0.928, 0.910, 0.891,

	0.871, 0.851, 0.830, 0.809, 0.788, 0.766, 0.743, 0.719, 0.695, 0.670, 0.643, 0.616, 0.587, 0.557, 0.525, 0.491, 0.455, 0.415, 0.371, 0.322, 0.276, 0.241, 0.207, 0.172, 0.138, 0.103, 0.069, 0.034, 0.000,
&physics	
mp_physics	= 5,
ra_lw_physics	= 1,
ra_sw_physics	= 1,
Radt	= 30,
sf_sfclay_physics	= 1,
sf_surface_physics	= 2,
bl_pbl_physics	= 1,
Bldt	= 0,
cu_physics	= 1,
Cudt	= 5,
isfflx	= 1,
ifsnow	= 1,
icloud	= 1,
surface_input_source	= 1,
num_soil_layers	= 4,
ucmcall	= 0,
mp_zero_out	= 1,
surface_input_source	= 1,
&bdy_control	
Spec_bdy_width	= 5,
Spec_zone	= 1,
relax_zone	= 4,
specified	= .true.,
nested	= .false.,
&dynamics	
w_damping	= 0,
rk_ord	= 3,
Diff_opt	= 0,
km_opt	= 4,
Diff_6th_opt	= 0,
Diff_6th_factor	= 0.12,
Damp_opt	= 0,
non_hydrostatic	= .true.,
h_mom_adv_order	= 5,
v_mom_adv_order	= 3,
h_sca_adv_order	= 5,
v_sca_adv_order	= 3,
Time_step_sound	= 4,

## Referencias:

- Accadia, C., S. Mariani, M. Casaioli, A. Lavagnini, and A. Speranza, 2003 : Sensitivity of Precipitation Forecast Skill Scores to Bilinear Interpolation and a Simple Nearest-Neighbor Average Method on High-Resolution Verification Grids. *Weather and Forecasting*, **18**, 918-932.
- Accadia, C., S. Mariani, M. Casaioli, A. Lavagnini, and A. Speranza, 2005 : Verification of Precipitation Forecasts from Two Limited-Area Models over Italy and Comparison with ECMWF Forecasts Using a Resampling Technique. *Weather and Forecasting*, **20**, 276-300.
- Anderson, J. R., Hardy E. E., Roach J. T. and Witmer R. E., 1976: A land use and land cover classification system for use with remote sensor data: *U.S. Geological Survey Professional Paper* **964**, 28 p.
- Arakawa A. and W. H. Schubert, 1974: Interaction of a Cumulus Cloud Ensemble with the Large-Scale Environment, Part I. *Journal of the Atmospheric Sciences*, **31**, 674-701.
- Arribas, A., K. B. Robertson, y K. R. Mylne, 2005: Test of a Poor Man's Ensemble Prediction System for Short-Range Probability Forecasting. *Monthly Weather Review*. **133**, 1825-1839.
- Blazquez, J. 2007: Validación regional del modelo WRF bajo condiciones meteorológicas persistentes. Tesis de Licenciatura del Departamento de Ciencias de la Atmósfera y los Océanos. Director: Mario Nuñez.
- Brier, G. W., 1950: Verification of Forecasts Expressed in Terms of Probability. *Mon. Wea. Rev.* **78**, 1-3.
- Brooks, H. E., M. S. Tracton, D. J. Stensrud, G. DiMego, and Z. Toth, 1995: Short-range ensemble forecasting: Report from a workshop, 25-27 July 1994. *Bull. Amer. Meteor. Soc.*, **76**, 1617-1624.
- Buizza R., 1997: Potential Forecast Skill of Ensemble Prediction and Spread and Skill Distributions of the ECMWF Ensemble Prediction System. *Monthly Weather Review* **125**, 99-119.
- Buizza R., A. Hollingsworth, F. Lalauette, and A. Ghelli, 1999: Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System. *Weather and Forecasting*, **14**, 168-189.
- Buizza R., 2001: Accuracy and Potential Economic Value of Categorical and Probabilistic Forecasts of Discrete Events. *Monthly Weather Review*. **129**, 2329-2345.
- Chen, F., and J. Dudhia, 2001: Coupling an advanced land-surface / hydrology model with the Penn State / NCAR MM5 modeling system. Part I: Model description and implementation. *Monthly Weather Review*, **129**, 569-585.

- Chessa, P. A., G. Ficca, M. Marrocu, y R. Buizza, 2004: Application of a Limited-Artea Short-Range Ensemble Forecast System to a Case of Heavy Rainfall in the Mediterranean Region. *Weather and Forecasting*, **19**, 566-581.
- Corazza M., E. Kalnay, D. J. Patil, S.-C. Yang, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt, and J. A. Yorke, 2003: Use of the breeding technique to estimate the structure of the analysis "errors of the day". *Nonlinear Processes in Geophysics*, **10**, 1-11.
- Coutinho, M. M. 1999 Previsão por conjuntos utilizando perturbações baseadas em componentes principais. São José dos Campos. 136p. Dissertação (Mestrado em Meteorologia) - Instituto Nacional de Pesquisas Espaciais
- Cusack, S. y A. Arribas, 2008: Assessing the Usefulness of Probabilistic Forecasts. *Monthly Weather Review*, **136**, 1492-1504.
- Dalcher, A., E. Kalnay, and R. N. Hoffman, 1988: Medium Range Lagged Average Forecasts. *Mon. Wea. Rev.*, **116**, 402-416.
- Du J., S. L. Mullen, and F. Sanders, 1997: Short-Range Ensemble Forecasting of Quantitative Precipitation. *Mon. Wea. Rev.*, **125**, 2427-2459.
- Dudhia J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, *J. Atmos. Sci.*, **46**, 3077-3107.
- Ebert E. E., 2001: Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation. *Mon. Wea. Rev.*, **129**, 2461-2480.
- Ebert E. E., U. Damrath, W. Wergen, and M. E. Baldwin, 2003: The WGNE Assessment of Short-term Quantitative Precipitation Forecasts. *Bulletin of the American Meteorological Society*, **84**, 481-492.
- Ebert E. E., J. E. Janowiak, and C. Kidd, 2007: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *Bulletin of the American Meteorological Society*, **88**, 47-64.
- Ebisuzaki, W., and E. Kalnay, 1991: Ensemble experiments with a new lagged average forecasting scheme. *WMO Report* #15.
- Eckel, F. A. and M. K. Walters, 1998: Calibrated Probabilistic Quantitative Precipitation Forecasts Based on the MRF Ensemble. *Weather and Forecasting*, **13**, 1132-1147.
- Epstein E. S., 1969: Stochastic-dynamic prediction. *Tellus*, **21**, 739-759.
- Evans E., N. Bhatti, J. Kinney, L. Pann, M. Peña, S.-C. Yang, E. Kalnay, and J. Hansen, 2004: RISE: Undergraduates Find That Regime Changes in Lorenz's Model are Predictable. *Bulletin of the American Meteorological Society*, **85**, 520-524.

Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. Preprints, 15th Conf. on Numerical Weather Prediction, San Antonio, TX, *Amer. Meteor. Soc.*, 280-283.

Freitas, S. R., K. Longo, M. Silva Dias, P. Silva Dias, R. Chatfield, A. Fazenda, L. F. Rodrigues, 2006: The coupled aerosol and tracer transport model to the Brazilian developments on the Regional Atmospheric Modeling System: validation using direct and remote sensing observations. Proceedings, *8<sup>th</sup> International Conf. on Southern Hemisphere Meteorology and Oceanography*, Foz do Iguazu, Brazil, American Meteorological Society, 101-107.

García Skabar, Y., 2007: Análisis enriquecidos, evaluación de su impacto en el pronóstico y su aplicación para describir circulaciones de mesoescala en el centro y norte de Argentina. Tesis doctoral de la Universidad de Buenos Aires, 177 pp. Director: Matilde Nicolini.

Gallus, W. A., and M. Segal, 2004: Does Increased Predicted Warm-Season Rainfall Indicate Enhanced Likelihood of Rain Occurrence? *Weather and Forecasting*, **19**, 1127-1135.

Grell G. A., 1993: Prognostic Evaluation of Assumptions used by Cumulus Parameterizations. *Monthly Weather Review*, **121**, 764-787.

Grell G. A. and D. Devenyi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29(14)**, Article 1639.

Grimit y C. Mass, 2007: Measuring the Ensemble Spread-Error Relationship with a Probabilistic Approach: Stochastic Ensemble Results. *Monthly Weather Review*, **135**, 203-221.

Hamill T., 1999: Hypothesis Tests for Evaluating Numerical Precipitation Forecasts. *Weather and Forecasting*, **14**, 155-167.

Hamill T. and S. J. Colucci, 1997: Verification of Eta-RSM Short-Range Ensemble Forecast, *Monthly Weather Review.*, **125**, 1312-1327.

Hamill T. and S. J. Colucci, 1998: Evaluation of Eta-RSME Ensemble Probabilistic Precipitation Forecast. *Monthly Weather Review.*, **126**, 711-724.

Hamill T. M., 2001: Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Monthly Weather Review*, **129**, 550-560.

Hamill T. and J. S. Whitaker, 2006: Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application. *Monthly Weather Review.* **134**, 3209-3229.

Hamill T. M. and J. Juras, 2006: Measuring forecast skill: is it real skill or is the varying climatology?. *Quarterly Journal of the Royal Meteorological Society*, **132**, 2905.

- Herdies D. L., V. E. Kousky, and Wesley Ebisuzaki, 2007: The Impact of High-Resolution SALLJEX Data on Global NCEP Analyses. *Journal of Climate*, **20**, 5765-5783.
- Hong S. and H. Pan, 1996: Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model. *Monthly Weather Review*, **10**, 2322-2339.
- Hou, D., E. Kalnay, and K. K. Droegemeier, 2001: Objective Verification of the SAMEX`98 Ensemble Forecast. *Monthly Weather Review*, **129**, 73-91.
- Houtekamer, P., 1993: Global and Local Skill Forecasts. *Monthly Weather Review*, **121**, 1834-1846.
- Hunt, B. R., E. J. Kostelich and I. Szunyogh, 2006: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter. *Physica D: Nonlinear phenomena*, **1-2**, 112-126.
- Janjic, Z. I., 2002: Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso model, NCEP Office Note, No. 437, 61 pp.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, **5**, 487-503.
- Kain J. S., 2004: The Kain-Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*. **43**, 170-181.
- Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press.
- Kalnay, M. Kanamitsu, and W.E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. *Bulletin of the American Meteorological Society*, **71**, 1410-1428.
- Kalnay E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Roplewski, J. Wang, Roy Jenne, and Dennis Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*.
- Kinter, J. L. III, D. Dewitt, P. A. Dirmeyer, M. J. Fennessy, B. P. Kirtman, L. Marx, E. K. Schneider, J. Shukla and D. M. Straus, 1997: The COLA-Atmosphere-Biosphere General Circulation Model Volume 1: Formulation. *COLA Technical Report # 51*, Center for Ocean Land Atmosphere Studies, Calverton Md.
- Klemp, J. B., W. C. Skamarock, and J. Dudhia, 2007: Conservative split-explicit time integration methods for the compressible nonhydrostatic equations. *Monthly Weather Review*, **8**, 2897-2913.

- Krishnamurti, T. N., C. M. Kshtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Willford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecast from multi model superensemble. *Science*, **285**, 1548-1550.
- Leith C. E., 1974: Theoretical Skill of Monte Carlo Forecasts. *Monthly Weather Review*, **102**, 409-418.
- Liebmann B. and D. Allured: Daily Precipitation Grids for South America. *Bulletin of the American Meteorological Society*. **86**, 1576-1570.
- Liebmann B., G. N. Kiladis, C. S. Vera, A. C. Saulo and L. M. V. Carvalho, 2004: Subseasonal Variations of Rainfall in South America in the Vicinity of the Low-Level Jet East of the Andes and Comparison to Those in the South Atlantic Convergence Zone. *Journal of Climate*, **17**, 3829-3842.
- Lorenz E. N., 1963: Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, **20**, 130-141.
- Magnusson, L., M. Leutbecher y E. Kallén, 2008: Comparison between Singular Vectors and Breeding Vectors as Initial Perturbations for the ECMWF Ensemble Prediction System. *Monthly Weather Review*, **136**, 4092-4104.
- McLean, J., A. Raftery, T. Gneiting, and C. Fraley, 2007: Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging. *Monthly Weather Review*, **135**, 3209-3220.
- Mendonça A. M. Desempenho do modelo global CPTEC/COLA durante episódios de ZCAS, utilizando os esquemas de convecção profunda tipo Kuo e Arakawa-Schubert relaxada. Dissertação de mestrado em meteorologia, Instituto Nacional de Pesquisas Espaciais, Brasil.
- Mendonça A. M. and J. P. Bonatti, 2008: Experiments with EOF-Based Perturbation Methods and their Impact on the CPTEC/INPE Ensemble Prediction System. *Monthly Weather Review*, en prensa.
- Mesinger, F., Z. I. Janjic, S. Nickovic, D. Gavrilov and D. G. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of Appalachian redevelopment. *Monthly Weather Review*, **116**, 1493-1518.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long wave. *J. Geophys. Res.*, **102** (D14), 16663-16682.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73-119.

- Molteni F. y R. Buizza, 1999: Validation of the ECMWF Ensemble Prediction System Using Empirical Orthogonal Functions. *Monthly Weather Review*, **127**, 2346-2358.
- Murphy A. H., 1973: A New Vector Partition of the Probability Score. *Journal of Applied Meteorology*, **12**, 595-600.
- Murphy A. H. and R. L. Winkler, 1987: A General Framework for Forecast Verification. *Monthly Weather Review*, **115**, 1330-1338.
- Nogués-Paegle J. N. and K. C. Mo, 1997 : Alternating Wet and Dry Conditions over South America during Summer. *Monthly Weather Review*, **125**, 279-291.
- Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, S. Kadokura, K. Wada, K. Kato, R. Oyama, T. Ose, N. Mannoji and R. Taira, 2007: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan*, **85**, 369-432.
- Paegle J, Lee A. Byerle, Celeste Saulo and Juan Ruiz: Global, Synoptic, and Mesoscale Predictability with Barotropic, Primitive Equation and Euler Models. AMS Numerical Weather Prediction Meeting. Utah, Junio 2007.
- Penalba O., C. Vera, B. Cerne, M. Rusticucci, P. Salio, L. Ferreira, B. Liebmann, D. Allured y A. Diaz, 2004: Daily Rainfall Data over Argentina and Uruguay during SALLJEX. CLIVAR Exchanges. Vol 9, No.1. PP 29-31.
- Pessacg, N., 2008: Experimentos de Sensibilidad Utilizando un Modelo Regional para la Simulación de la Precipitación en el Sur de Sudamérica. Tesis de Licenciatura del Departamento de Ciencias de la Atmósfera y los Océanos (FCEyN) Universidad de Buenos Aires, 80 pp.
- Parrish D. F. and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review*, **120**, 1747-1763.
- Raftery, A. E., T. Gneiting, F. Balabdaoui, y M. Polakowski: Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review*, **133**, 1155-1174.
- Richardson, D. S., 2000: Skill and economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **126**, 649-668.
- Robledo F. A., 2007: Régimen estacional de la lluvia diaria en el Sudeste de Sudamérica. Cuantificación del cambio en la segunda mitad del siglo 20. Tesis de Licenciatura del Departamento de Ciencias de la Atmósfera y los Océanos (FCEyN) Universidad de Buenos Aires, 76 pp.
- Roebber P. J. and L. F. Bosart, 1996: The Complex Relationship between Forecast Skill and Forecast Value: A Real-World Analysis. *Weather and Forecasting*, **11**, 544-559.

- Roulin E. and S. Vannitsem, 2005: Skill of Medium-Range Hydrological Ensemble Predictions. *Journal of Hydrometeorology*, **6**, 729-744.
- Ruiz J. J., E. Kalnay y C. Saulo, 2005: Aplicación de la técnica de vectores criados a una situación de baja predictabilidad. Actas de IX Congreso Argentino de Meteorología. Buenos Aires, Argentina. 10 pp.
- Ruiz J. J., 2008: Verificación y calibración de las estimaciones de precipitación CMORPH durante el SALLJEX. Enviado a la Revista Brasileira de Meteorología.
- Ruiz, J. J., A. C. Saulo and E. Kalnay, 2008: Comparison of methods to generate probabilistic quantitative precipitation forecasts over South America. *Weather and Forecasting*. In press.
- Ruiz J. J., A. C. Saulo and E. Kalnay, 2006: A regional ensemble forecast system for southeastern south America: preliminary assessment. Proceedings, *8<sup>th</sup> International Conf. on Southern Hemisphere Meteorology and Oceanography*, Foz do Iguazu, Brazil, American Meteorological Society, 1977-1984.
- Salio, P., M. Nicolini, y A. C. Saulo, 2002: Chaco Low-Level Jet Events Characterization During the Austral Summer Season by ERA Reanalysis. *Journal of Geophysical Research – Atmosphere*, **107**, D.24.10.1019/2001JD001315.
- Schaefer, J. T., 1990: The Critical Success Index as an Indicator of Warning Skill. *Weather and Forecasting*, **5**, 570-575.
- Sela, J. G., 1980: Spectral modeling at NMC, *Monthly Weather Review*, **108**, 1279-1292.
- Silva Dias, P. L., D. Soares Moreira, and G. D. Neto, 2006: The MASTER Model Ensemble System (MSMES). Proceedings of 8 ICSHMO, Foz do Iguazu, Brazil, April 24-28, INPE, p. 1751-1757.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF Version 2. NCAR Tech Notes-468+STR
- Stensrud, D. and N. Yussouf, 2007: Reliable Probabilistic Quantitative Precipitation Forecast from a Short-Range Ensemble Forecasting System. *Weather and Forecasting*, **22**, 3-17.
- Talagrand O., R. Vautard and B. Strauss, 1997: Evaluation of probabilistic prediction systems. Proc. ECMWF Workshop on Predictability.
- Tencer B., 2006: Períodos de retorno de valores extremos de temperatura en Argentina: El cambio observado desde comienzos del siglo XX. Tesis de Licenciatura del Departamento de Ciencias de la Atmósfera y los Océanos (FCEyN Universidad de Buenos Aires), 72 pp.
- Toth, Z. y E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations.

*Bulletin of the American Meteorological Society*, **74**, 2317-2330.

Toth, Z., and E. Kalnay, 1997: Ensemble Forecasting at NCEP and the Breeding Method. *Monthly Weather Review*, **125**, 3297-3319.

Vera, C., J. Baez, M. Douglas, C. B. Emmanuel, J. Marengo, J. Meitin, M. Nicolini, J. Noguez-Paegle, J. Paegle, O. Penalba, P. Salio, C. Saulo, M. A. Silva Dias, P. Silva Dias, and E. Zipser, 2006: The South American Low-Level Jet Experiment. *Bulletin of the American Meteorological Society*, **87**, 63-77.

Whitaker J. and A. F. Lough, 1998: The Relationship between Ensemble Spread and Ensemble Mean Skill. *Monthly Weather Review*, **126**, 3292-3302.

Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences: An Introduction. International Geophysics Series, Vol. 59, Academic Press, 467 pp.

Yang S.-C., D. Baker, H. Li, K. Cordes, M. Huff, G. Nagpal, E. Okereke, J. Villafañe, E. Kalnay, and G. S. Duane, 2006: Data Assimilation as Synchronization of Truth and Model: Experiments with the Three-Variable Lorenz System. *Journal of the Atmospheric Sciences*, **63**, 2340-2354.

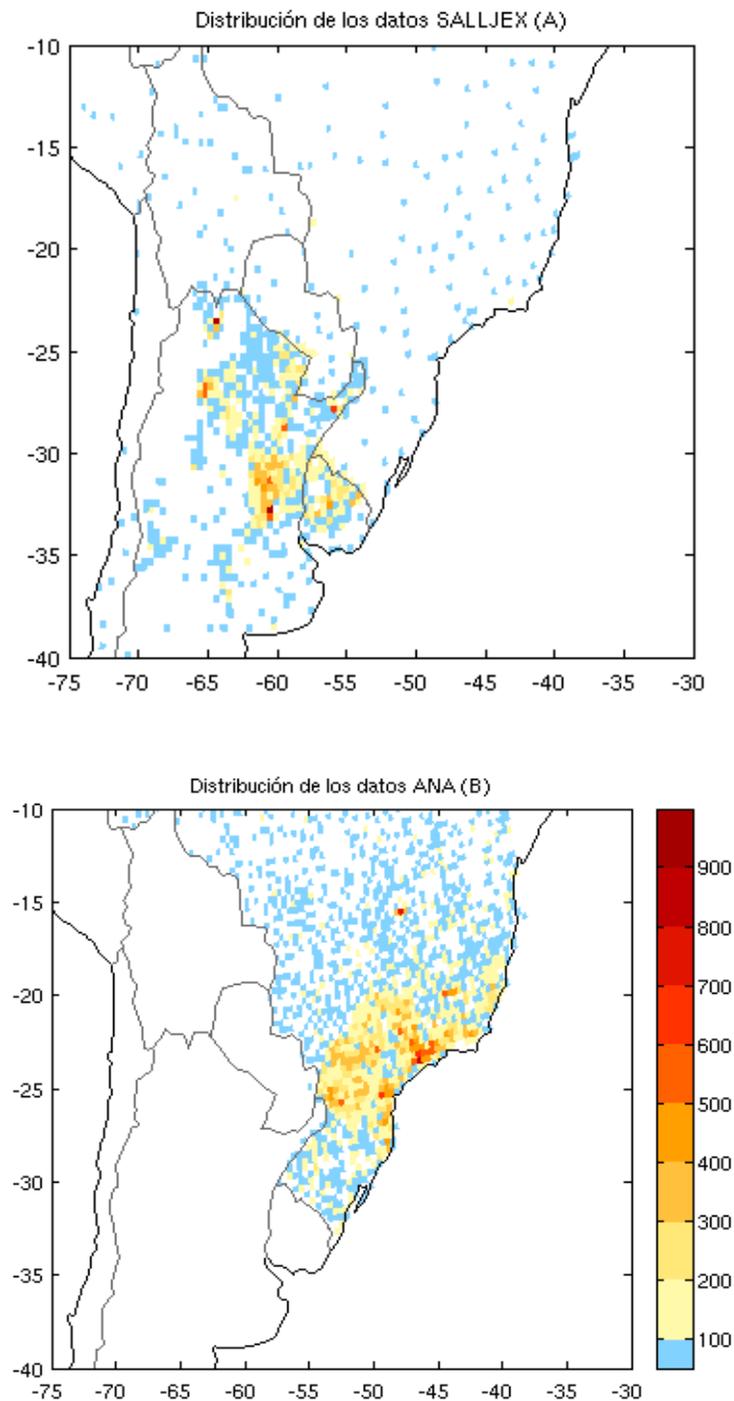
Yang, Z. and R. W. Arritt, 2002: Test of a perturbed physics ensemble approach for regional climate modeling. *J. Climate*, **15**, 2881-2896.

Zhang, Z., Krishnamurti, T. N., 1999: A perturbation method for hurricane ensemble predictions. *Mon. Wea. Rev.*, **127**, 447-469.

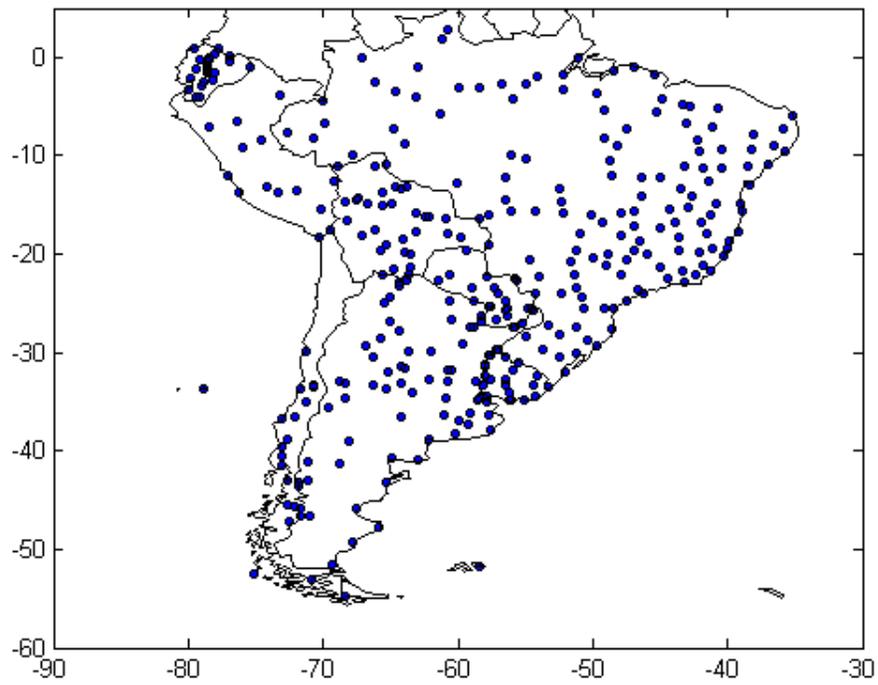
Zhu Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The Economic Value of Ensemble-Based Weather Forecast. *Bulletin of the American Meteorological Society*, **83**, 73-83.

# **Volumen II: Figuras.**

## Figuras Capítulo 2

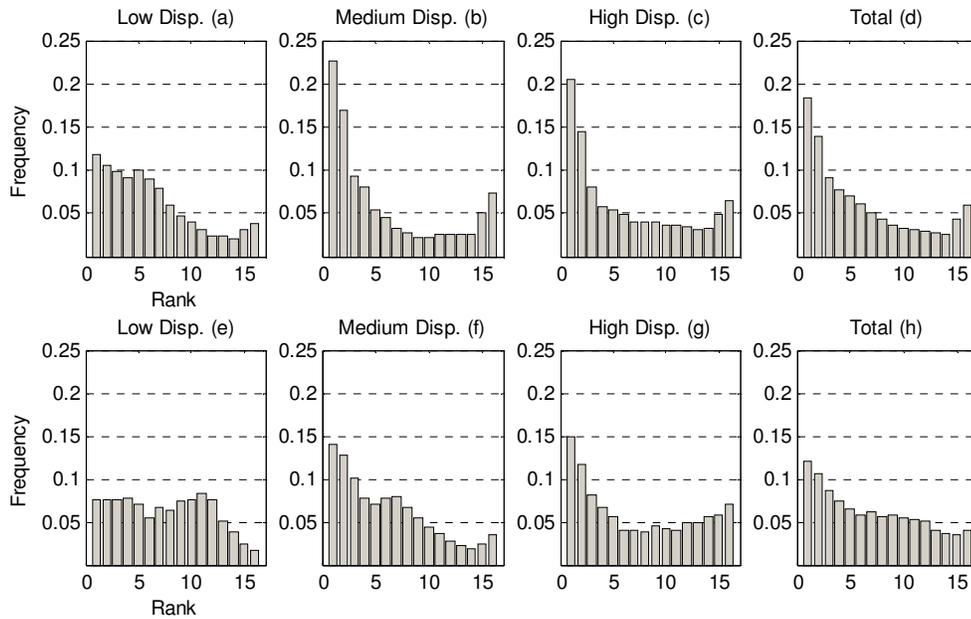


**Figura 2.1: Distribución del número total de datos de precipitación en cajas de 40km por 40km, en el período comprendido entre el 15 de diciembre de 2002 y el 15 de febrero de 2003 para: (a) la red de precipitación SALLJEX y (b) la red de precipitación ANA.**

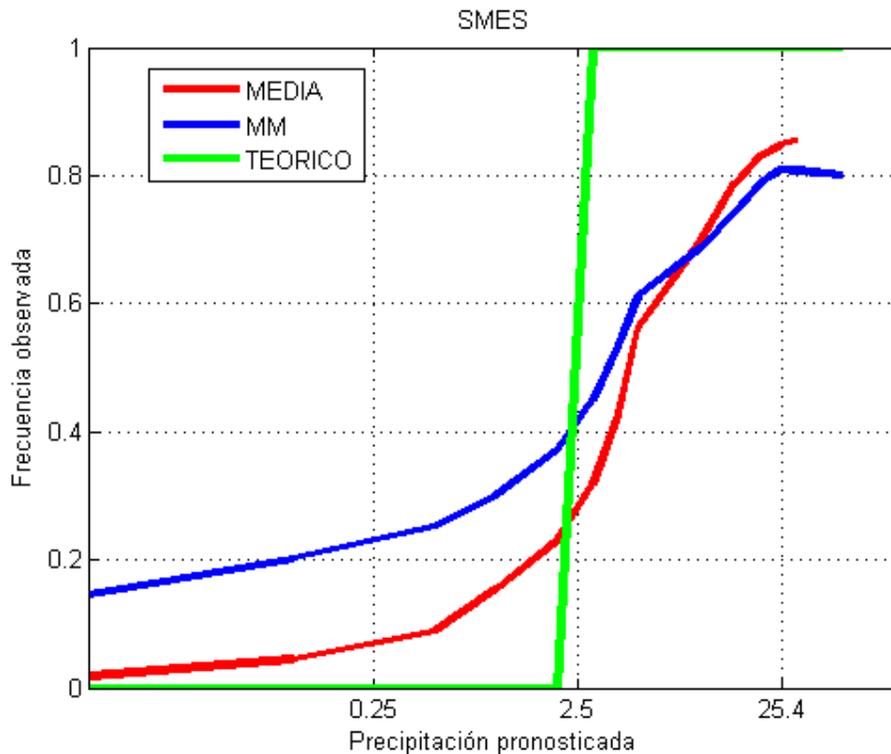


**Figura 2.2: Ubicación de las estaciones pluviométricas y de superficie de la red GTS.**

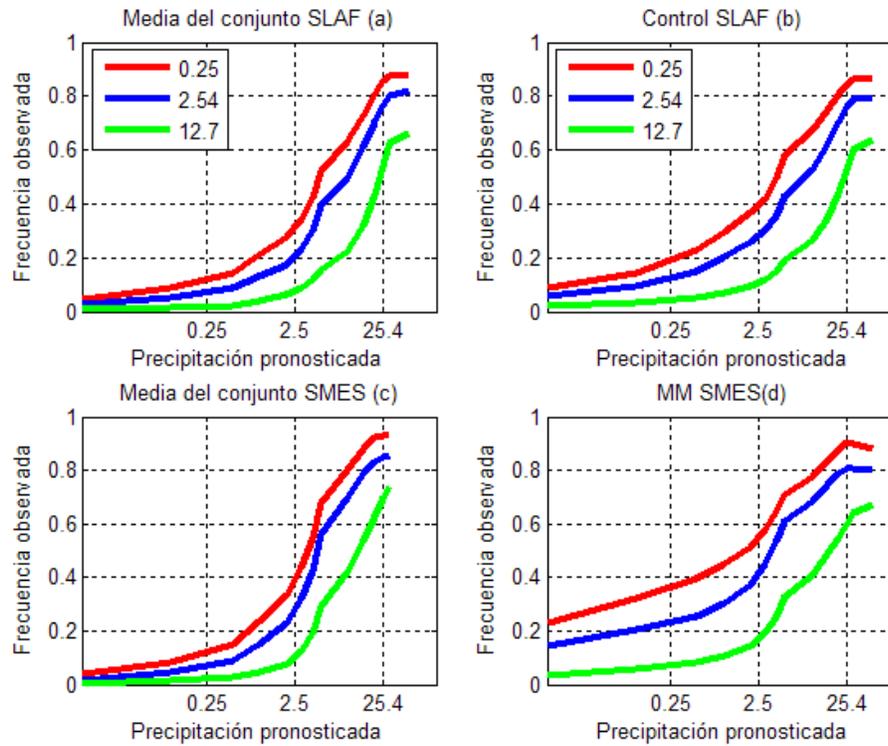




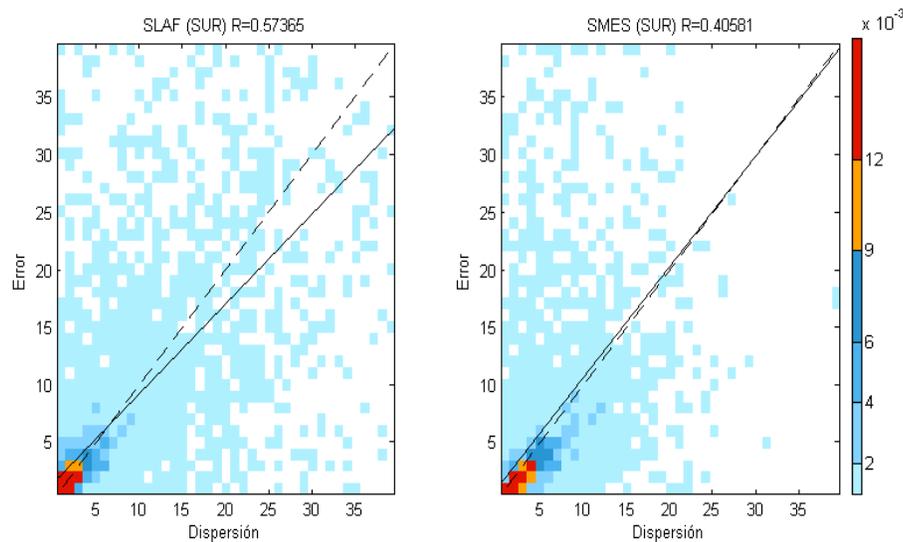
**Figura 3.3: Histograma de rangos para el SMES para distintos valores de dispersión del conjunto y para el total de casos. Los paneles superiores corresponden a la región norte, mientras que los inferiores corresponden a la región sur.**



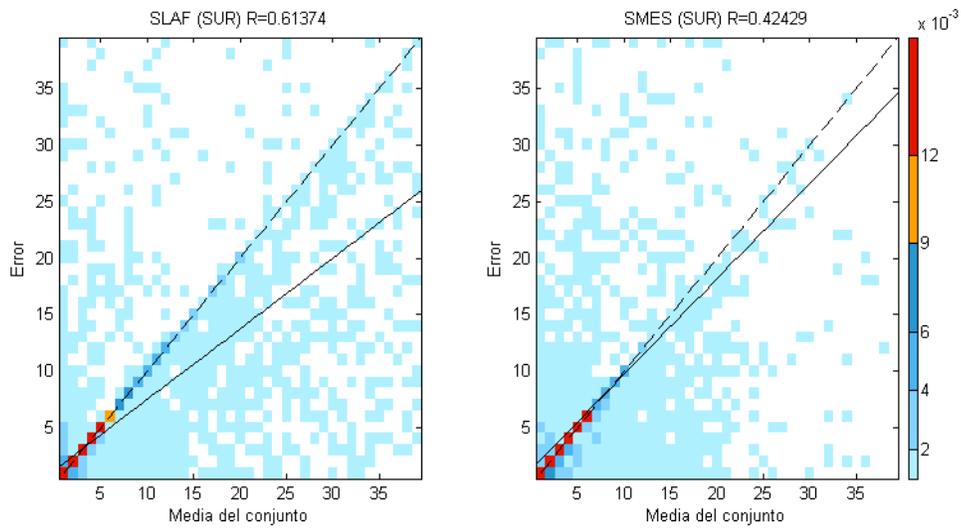
**Figura 3.4: Relación entre la precipitación pronosticada y la probabilidad de precipitación por encima de 2.5 mm para la media del SMES (línea roja) el MM (línea azul) y un pronóstico perfecto (línea verde).**



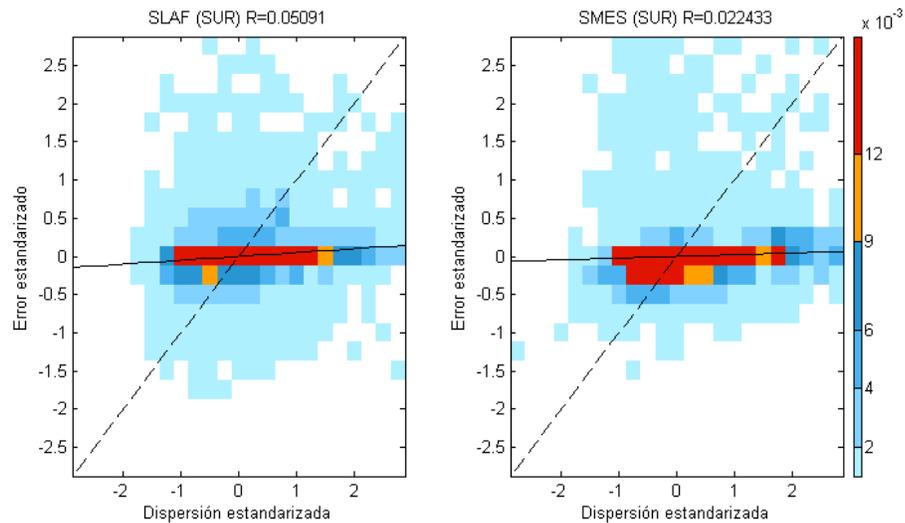
**Figura 3.5: Relación entre probabilidad de ocurrencia de precipitación por encima de un determinado umbral y la precipitación pronosticada por el modelo. Umbral de 0.25 mm (línea roja), umbral de 2.5 mm (línea azul) y umbral de 12.7 mm (línea verde). Para la media del conjunto SLAF (a), el control del conjunto SLAF (b), la media del conjunto SMES (c) y el MM del conjunto SMES (d).**



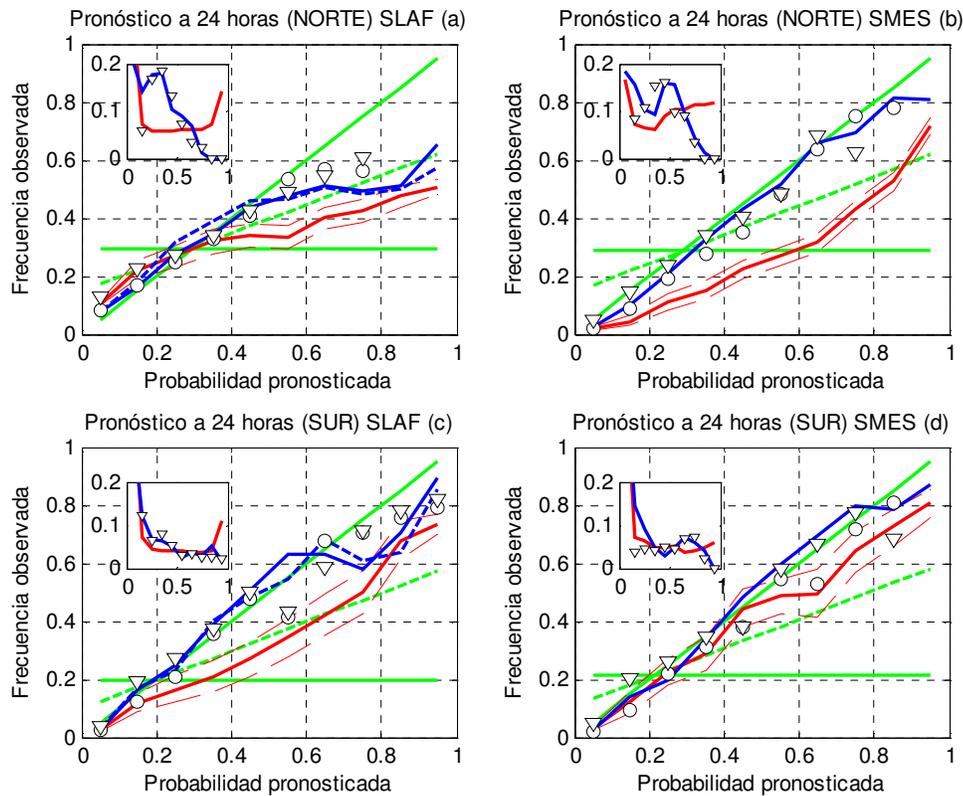
**Figura 3.6: Gráfico de dispersión de error de la media del conjunto como función de la dispersión del conjunto. La línea continua negra indica la recta de regresión lineal entre ambas variables y la línea a trazos la relación 1 a 1 entre las mismas.**



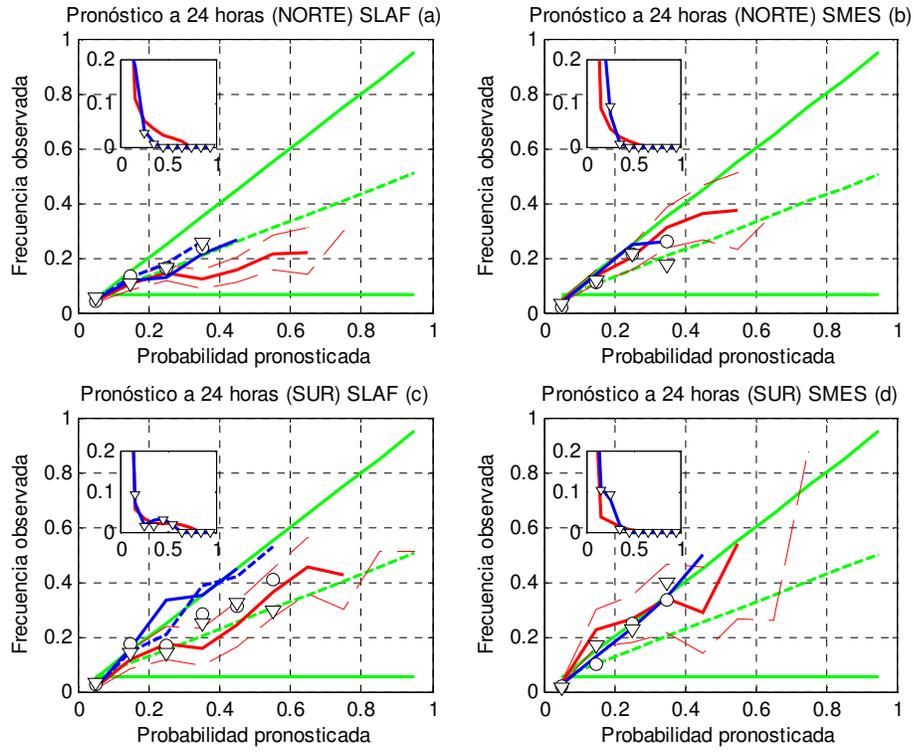
**Figura 3.7:** Ídem Figura 3.6 pero para la relación entre el error de la media del conjunto y el valor de la media del conjunto.



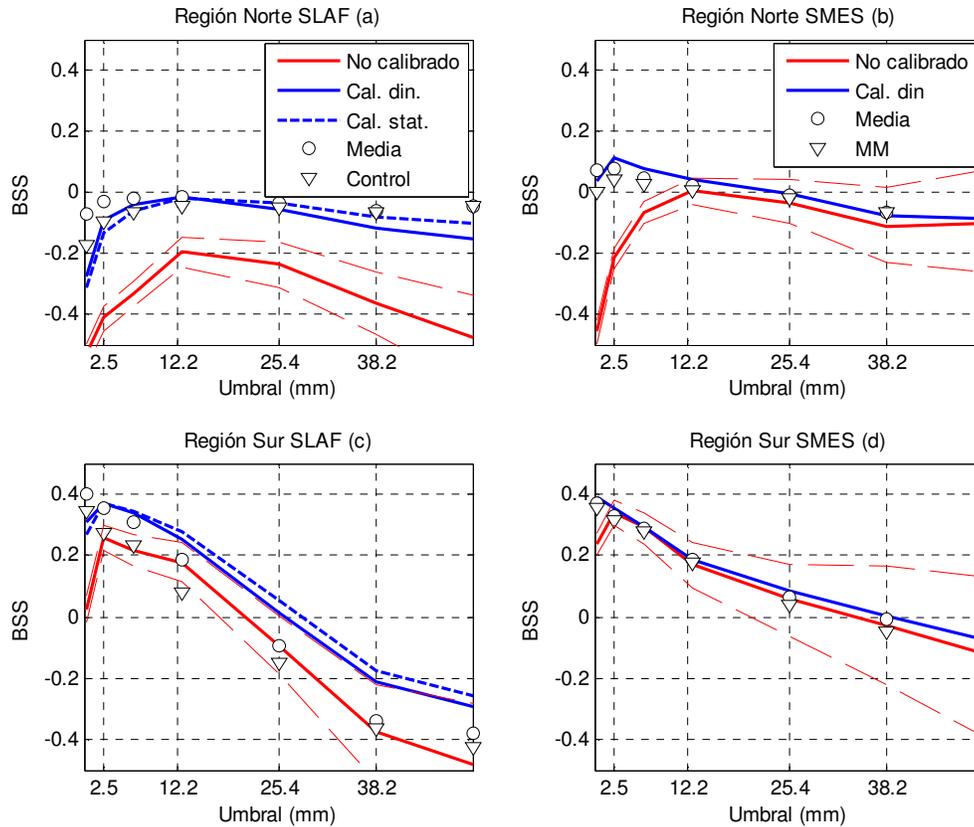
**Figura 3.8:** Como en la Figura 3.6 pero para la relación entre la dispersión estandarizada y el error estandarizado.



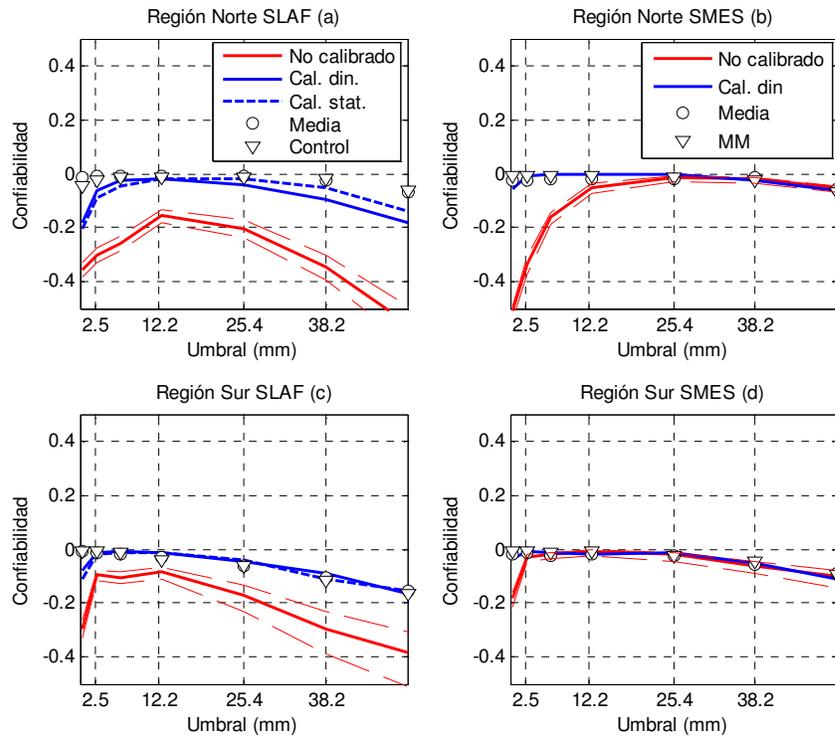
**Figura 3.9: Diagrama de confiabilidad del pronóstico para el umbral de 2.54 mm. La línea roja continua corresponde al pronóstico no calibrado y las líneas discontinuas son sus límites de confianza al 90%. La línea azul continua es el pronóstico calibrado dinámicamente. La línea azul discontinua corresponde a la calibración estática. Los triángulos corresponden al pronóstico determinístico (control SLAF y MM SMES) y los círculos a la media del ensamble. Los recuadros muestran la distribución de frecuencia con la que se pronostica cada uno de los valores de probabilidad respetando los mismos colores que en el gráfico de confiabilidad. (a) Ensamble SLAF, región norte, (b) ensamble SMES, región norte, (c) ensamble SLAF, región sur y (d) ensamble SMES, región sur. La línea diagonal verde continua indica la relación esperada para un ensamble ideal, la línea verde continua horizontal marca la confiabilidad esperada para un pronóstico sin resolución. Los puntos que se ubican entre la línea verde discontinua y la línea verde diagonal continua son los que aportan positivamente al BSS (ver discusión Apéndice A)**



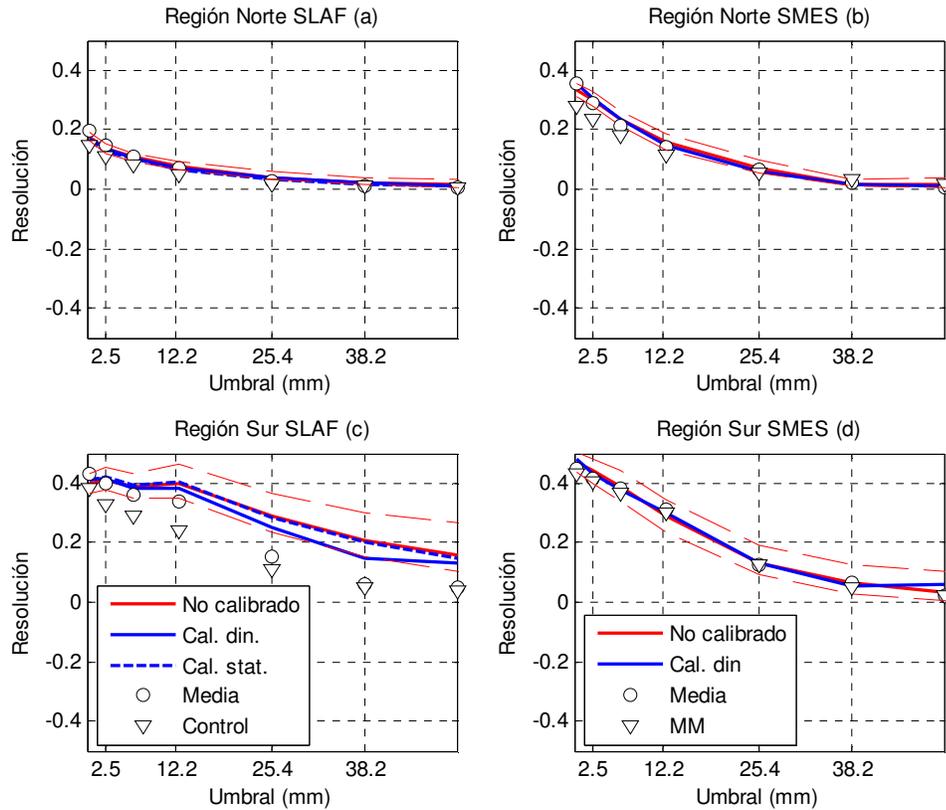
**Figura 3.10: Como en la Figura 3.9 pero para el umbral de 25.4 mm.**



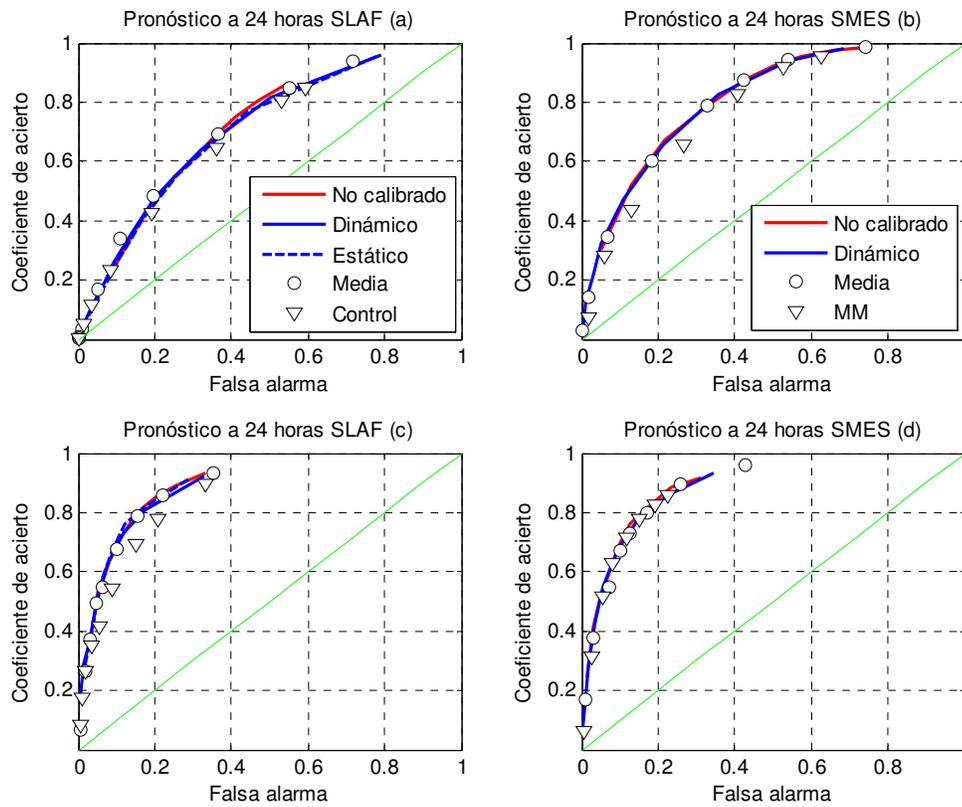
**Figura 3.11: Índice de acierto de Brier (BSS) como función del umbral pronosticado. Línea roja continua, pronóstico no calibrado y en línea roja discontinua los límites del intervalo de confianza. La línea azul continua es la calibración dinámica, la línea azul discontinua la calibración estática, los triángulos representan el pronóstico probabilístico derivado del control y del MM según corresponda y los círculos blancos representan el pronóstico probabilístico derivado de la media del ensemble. (a) Ensemble SLAF región norte, (b) ensemble SMES región norte, (c) ensemble SLAF región sur y (d) ensemble SMES región sur.**



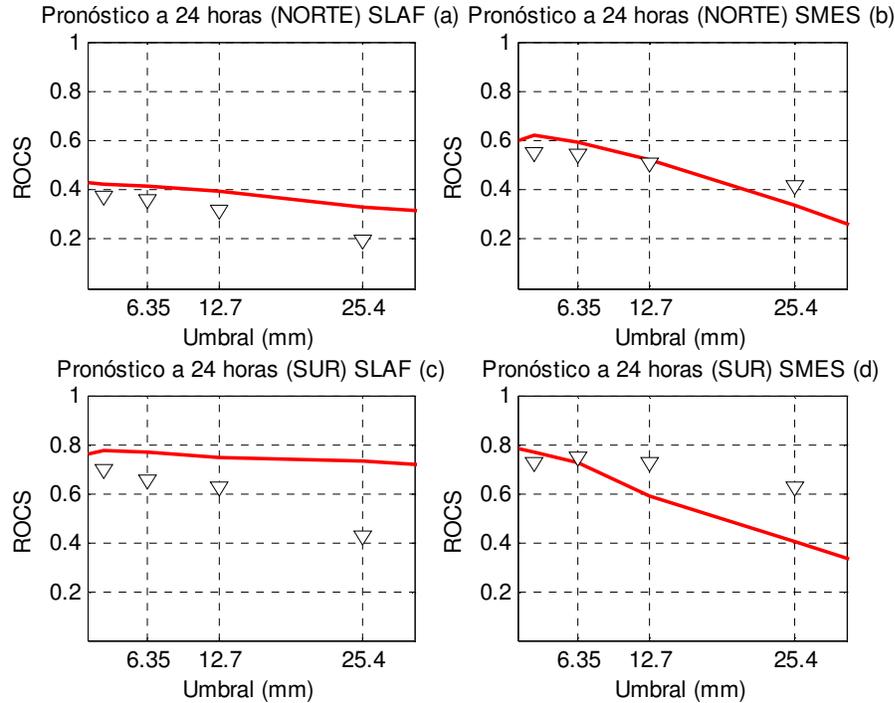
**Figura 3.12: Como en la Figura 3.11 pero para la componente de confiabilidad del Índice de acierto de Brier (BSS).**



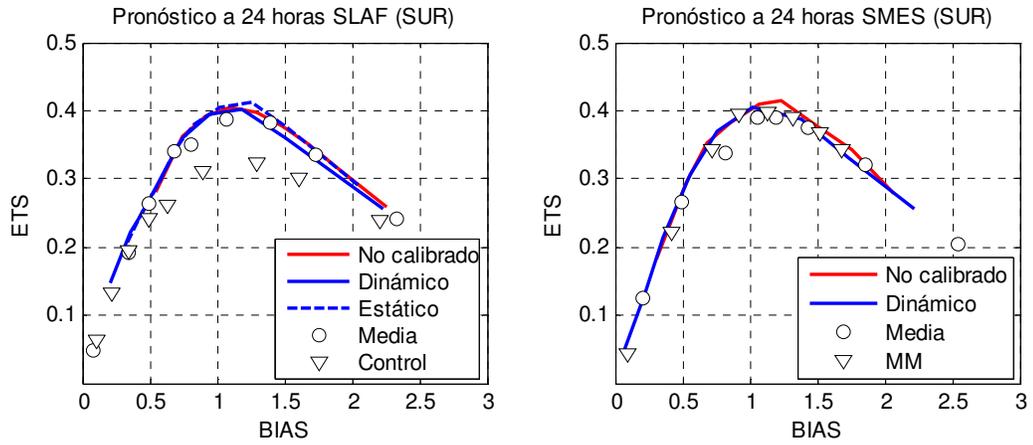
**Figura 3.13: Como en la Figura 3.11 pero para la componente de resolución del Índice de acierto de Brier (BSS).**



**Figura 3.14: Diagramas R.O.C para el umbral de 2.5 mm. Los colores de las curvas son como en la Figura 3.11.**

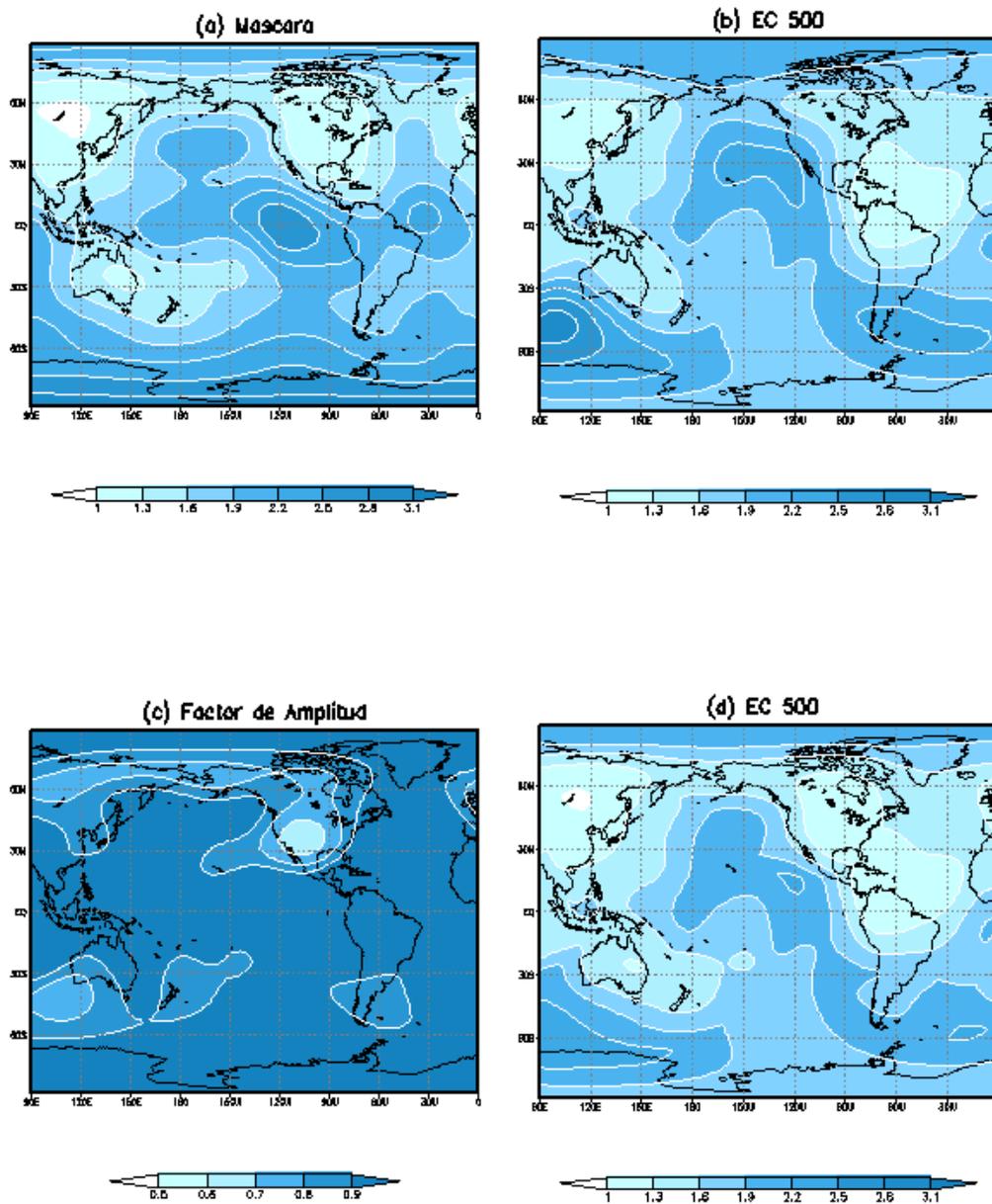


**Figura 3.15: Índice ROCS para el pronóstico no calibrado (curva roja), el pronóstico control SLAF (triángulos) (a) y (c) y el pronóstico MM (triángulos) (b) y (d). (a) Ensamble SLAF región norte, (b) ensamble SMES región norte, (c) ensamble SLAF región sur y (d) ensamble SMES región sur.**

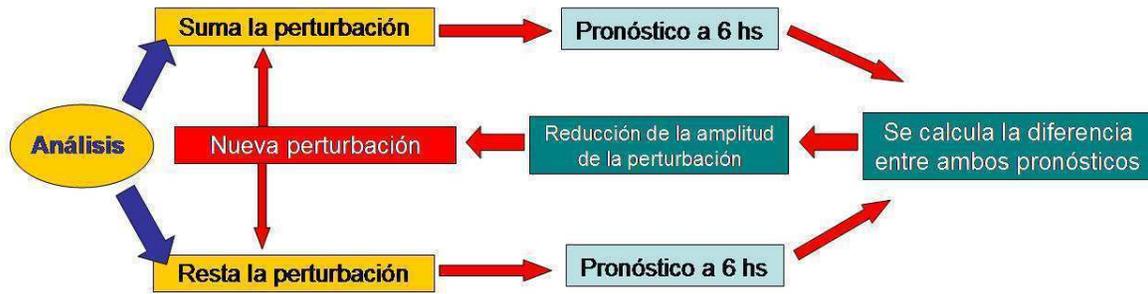


**Figura 3.16: Diagrama E.T.S-bias para (a) el ensamble SLAF y (b) el SMES, para el pronóstico a 24 horas sobre la región sur. Los colores son como en la Figura 3.14.**

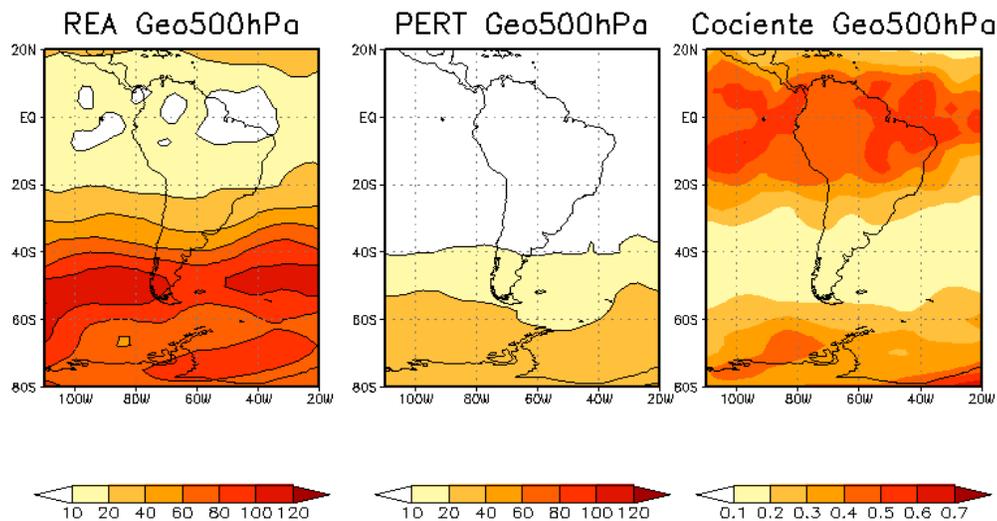
# Figuras Capítulo 4



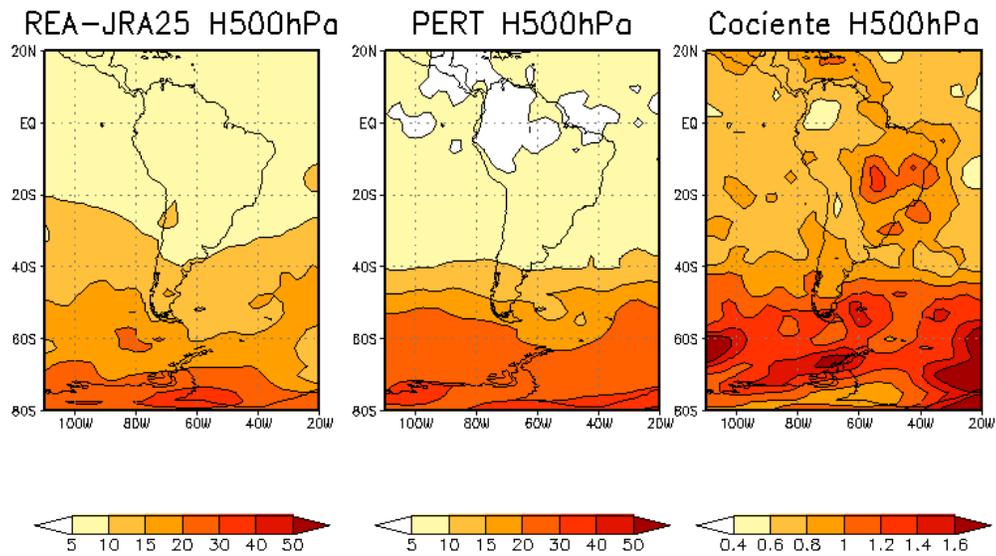
**Figura 4.1:** (a) Máscara global utilizada en el proceso de reducción de amplitud de las perturbaciones. (b) Energía cinética rotacional de la perturbación en 500 hPa suavizada. (c) Factor de reducción de amplitud de la perturbación y (d) Energía cinética rotacional de la perturbación en 500 hPa suavizada, luego de la reducción de la amplitud.



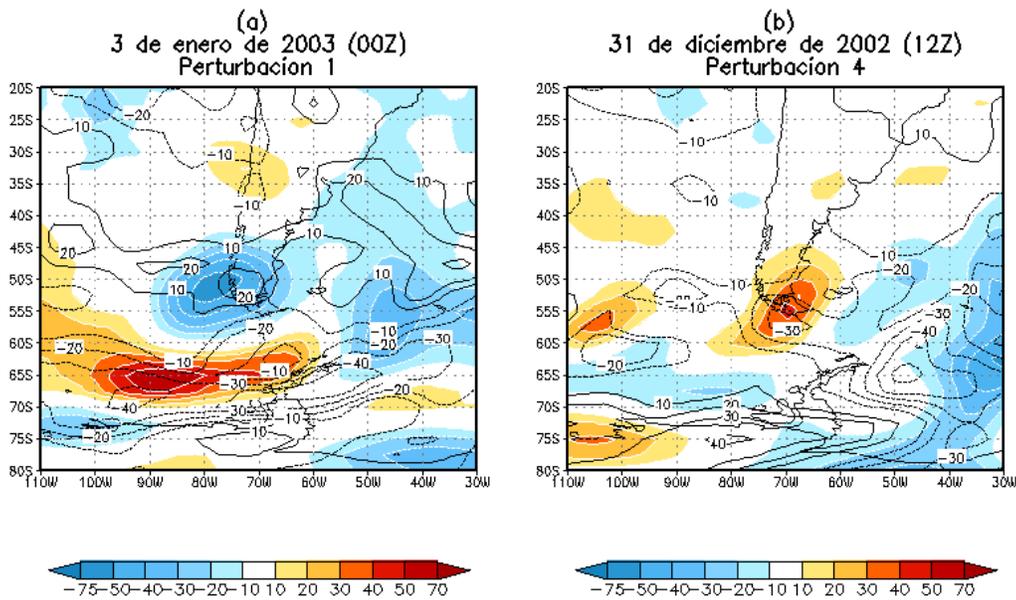
**Figura 4.2: Esquema del ciclo de generación de perturbaciones utilizando el método de los vectores criados.**



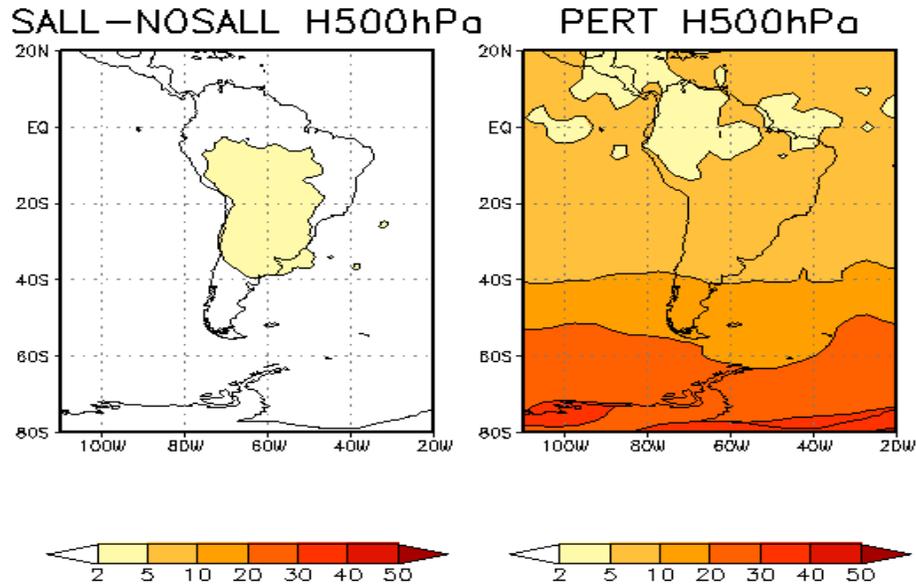
**Figura 4.3: Desviación estándar de la altura geopotencial de 500 hPa. para los reanálisis del NCEP (izquierda), diferencia cuadrática media para la perturbación de dicha variable generada con el método de los vectores criados (centro) y el cociente entre ambas variables (sigma perturbación / sigma reanálisis), para el período comprendido entre el 15 de diciembre de 2002 y el 15 de febrero de 2003.**



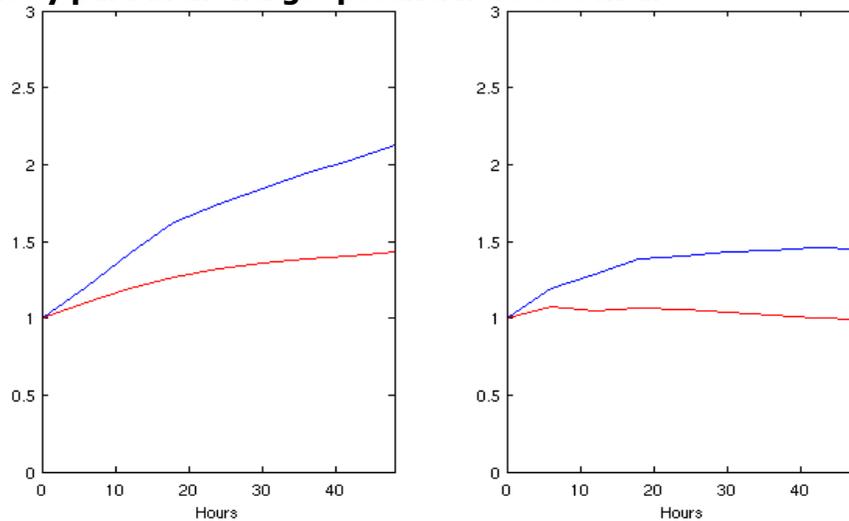
**Figura 4.4:** (a) RMSE de las diferencias entre los reanálisis del NCEP y los reanálisis JRA-25 para la altura geopotencial de 500 hPa. (b) RMSE para las perturbaciones en esa misma variable. (c) Cociente entre ambas. (RMSE perturbaciones / RMSE análisis.)



**Figura 4.5:** Perturbación en el campo de geopotencial en 500 hPa (m<sub>gp</sub>) (sombreado) e incerteza presente en el análisis (contornos) para (a) la perturbación 1 a las 00 UTC del día 3 de enero de 2003 y (b) para la perturbación 4 a las 12 UTC del día 4 de diciembre de 2002.

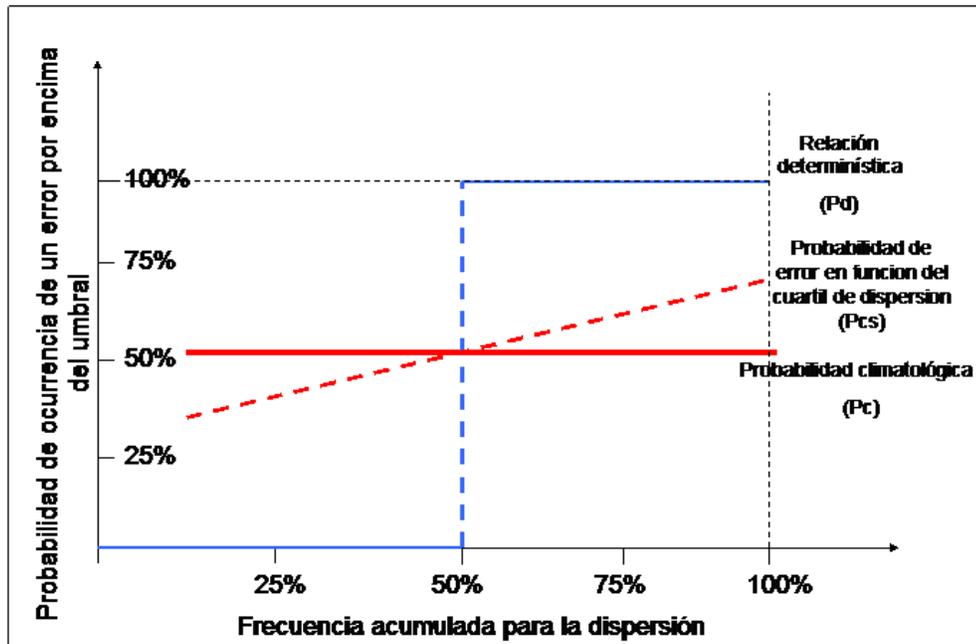


**Figura 4.6:** (a) RMSE de las diferencias entre los análisis con los datos SALLJEX asimilados y los análisis operativos, (b) RMSE del vector criado correspondiente a la perturbación 1. En ambos casos para todo el período del experimento y para la altura geopotencial de 500 hPa.

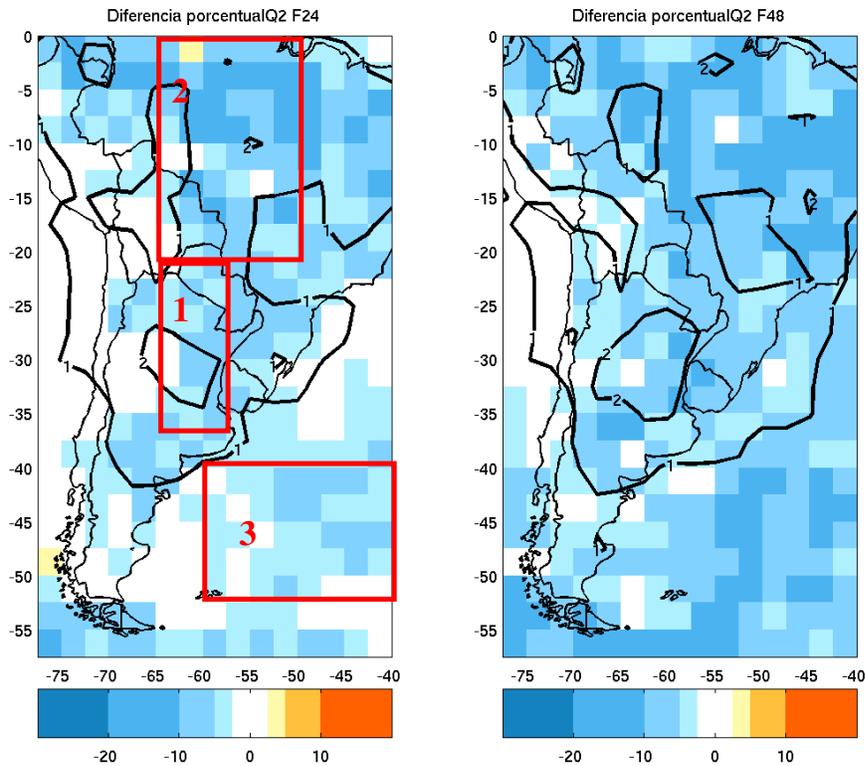


**Figura 4.7:** Cociente entre la amplitud de las diferencias entre un par de pronósticos para el viento en 500 hPa. (a) un par de pronósticos en donde la diferencia inicial entre ambos es un vector criado y (b) en donde la diferencia es el análisis GDAS con datos SALLJEX asimilados o sin datos SALLJEX asimilados. En ambos casos se muestran los resultados para latitudes tropicales (al norte de 20°S, línea roja) y latitudes medias y altas (entre 20° y 50°S de latitud línea azul).

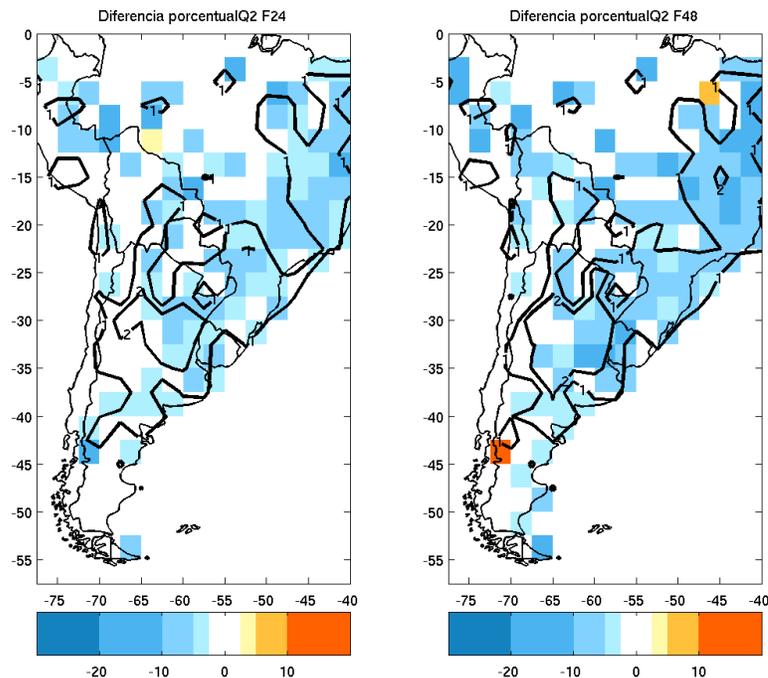
## Figuras Capítulo 5



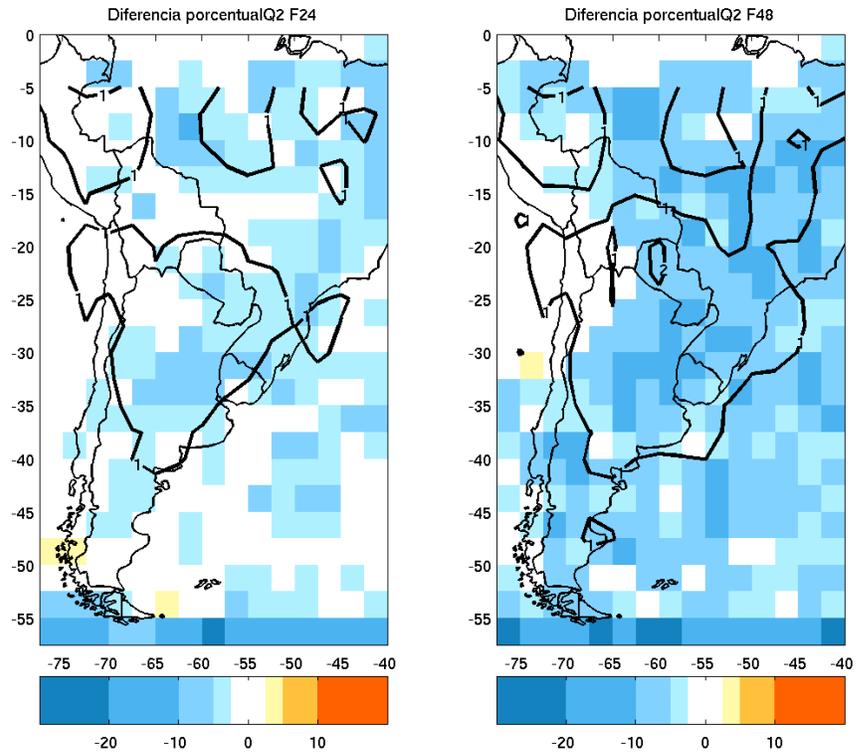
**Figura 5.1:** Esquema de relación entre la dispersión del ensamble y el error. La línea roja continua representa la probabilidad climatológica de ocurrencia de un determinado umbral de error. La línea azul representa la probabilidad condicional de ocurrencia de errores como función de la dispersión si la relación entre ambas variables fuera determinística y perfecta. La línea roja a trazos representa esta misma probabilidad condicional para un pronóstico real.



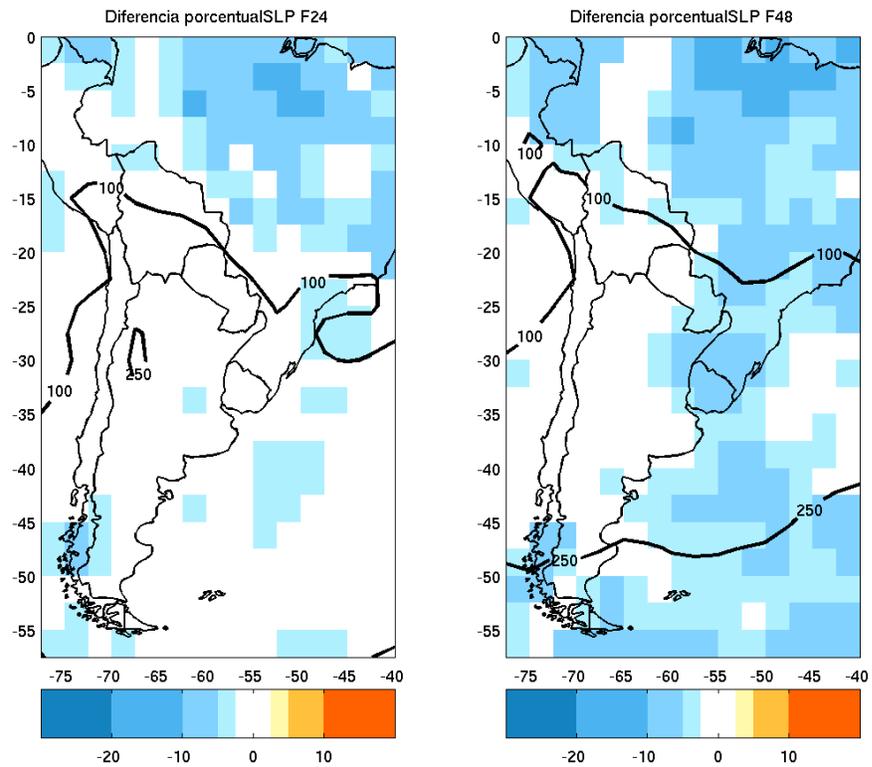
**Figura 5.2:** Diferencia entre el error de la media del ensamble y el error del pronóstico control, dividido por el error del pronóstico control (sombreado) (%). Los contornos muestran el error del pronóstico control ( $\text{g Kg}^{-1}$ ) para Q2M para el pronóstico a 24 horas (izquierda) y el pronóstico a 48 horas (derecha). Las cajas numeradas indican la posición y número de las regiones que serán discutidas más adelante en el texto. En ambos casos para el ensamble global.



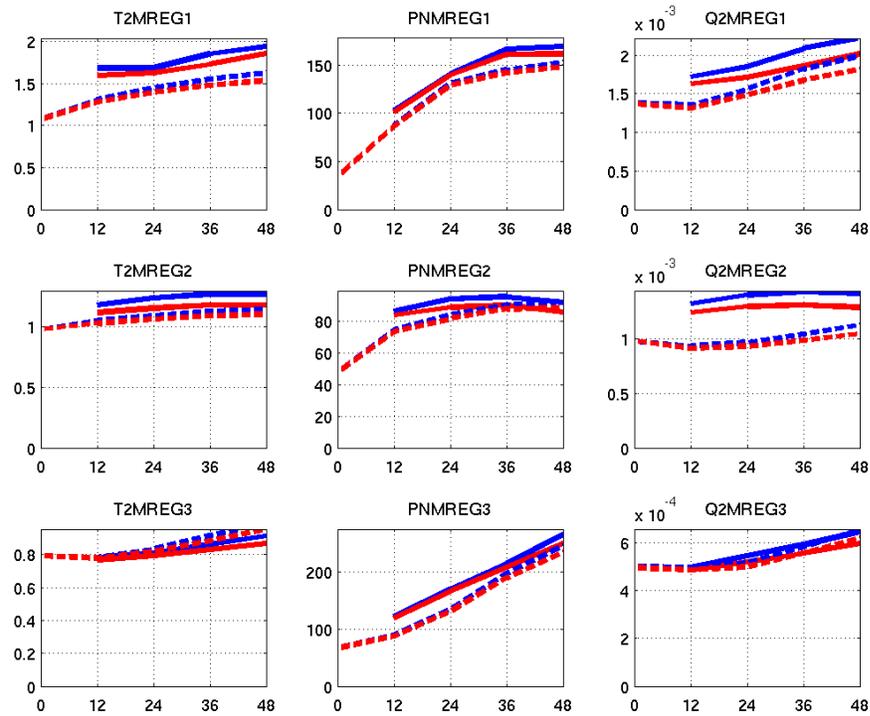
**Figura 5.3:** Como en la Figura 5.2, pero utilizando los datos del GTS como verificación.



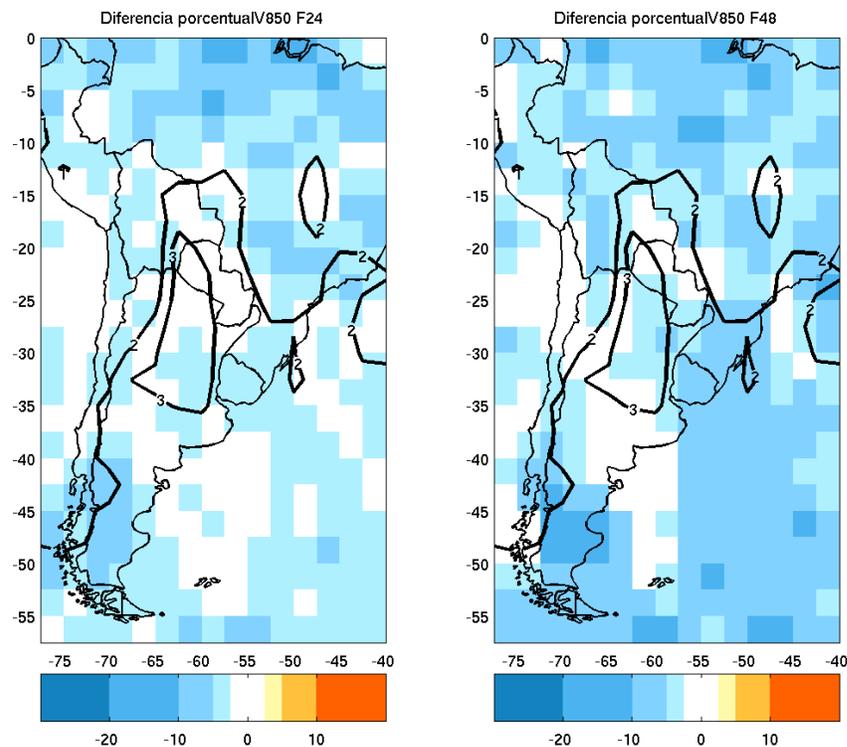
**Figura 5.4:** Como en la Figura 5.2, pero para el ensamble regional.



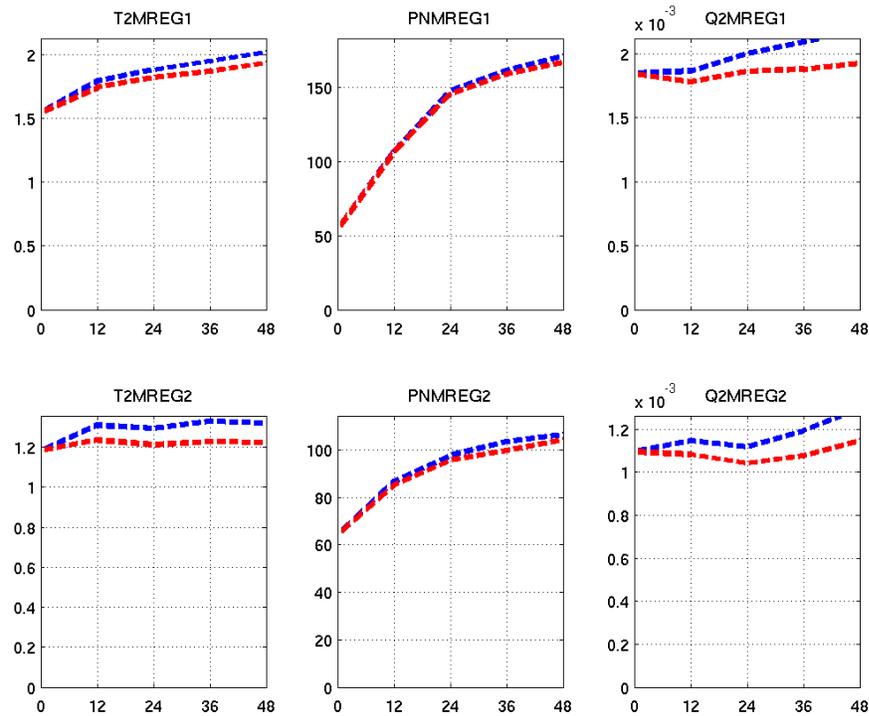
**Figura 5.5:** Como en la Figura 5.2, pero para la presión a nivel del mar (Pa) en el ensamble global.



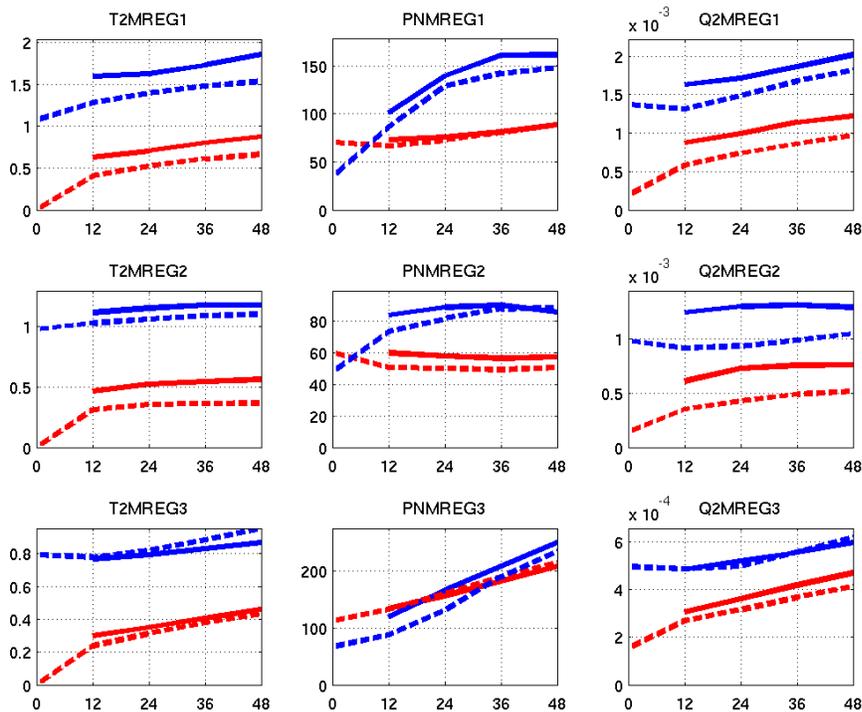
**Figura 5.6: Evolución de los errores en el ensamble regional y global para la T2M (K), PNM (Pa) y Q2M ( $\text{kg kg}^{-1}$ ) para las diferentes regiones analizadas. La línea azul representa el error en el pronóstico control y la línea roja representa el error en la media del ensamble. Las líneas continuas corresponden al ensamble global y las discontinuas al ensamble regional.**



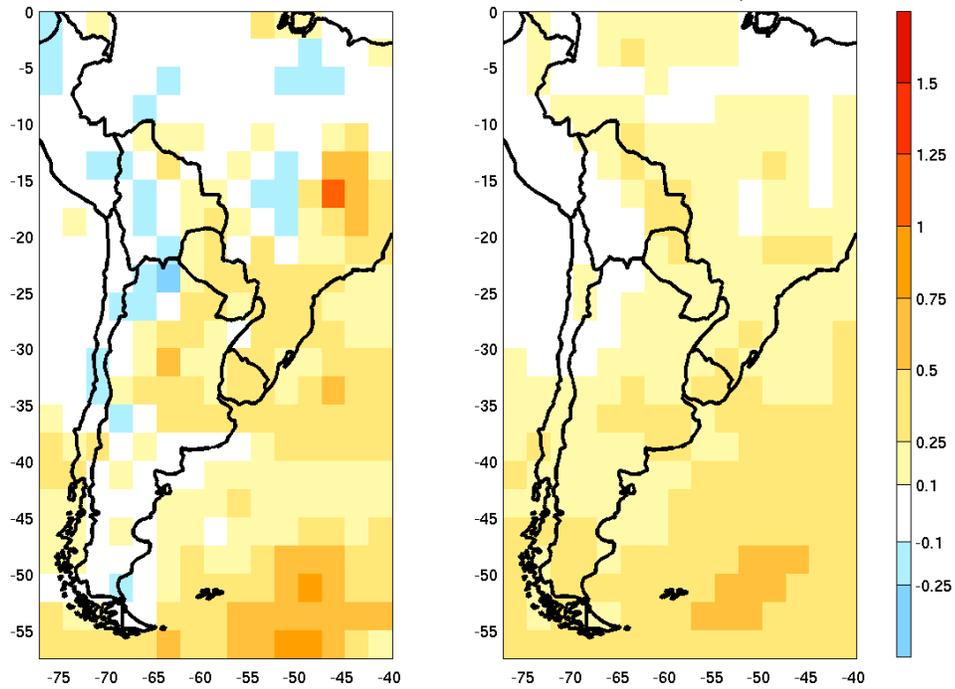
**Figura 5.7: Como en la Figura 5.2, pero para la componente V del viento en 850 hPa.**



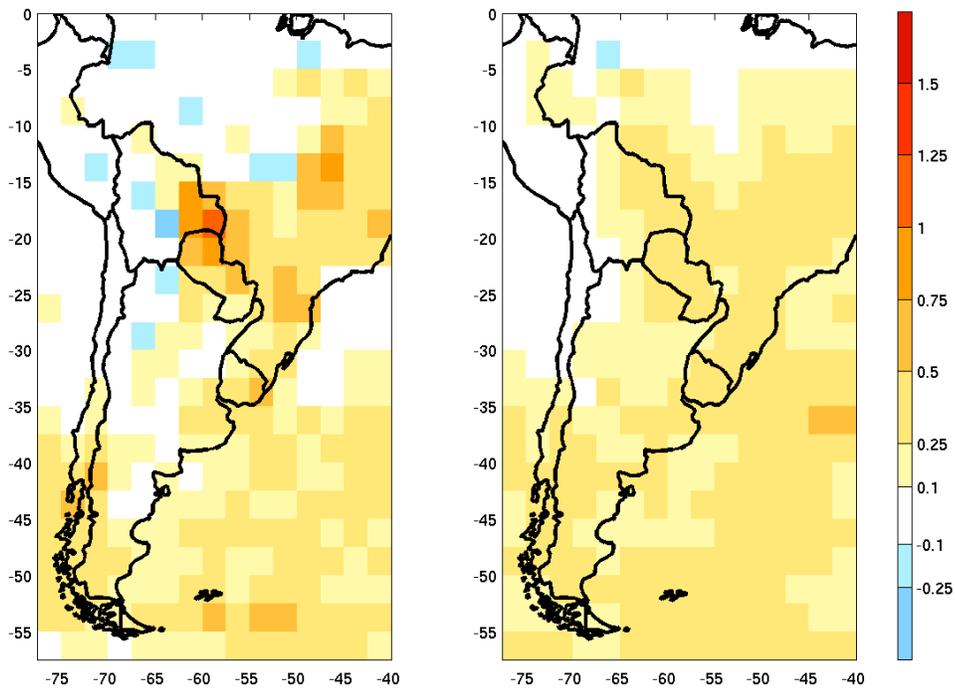
**Figura 5.8:** Como en la Figura 5.6 pero para el ensamble regional en alta resolución comparado contra los datos de superficie del GTS.



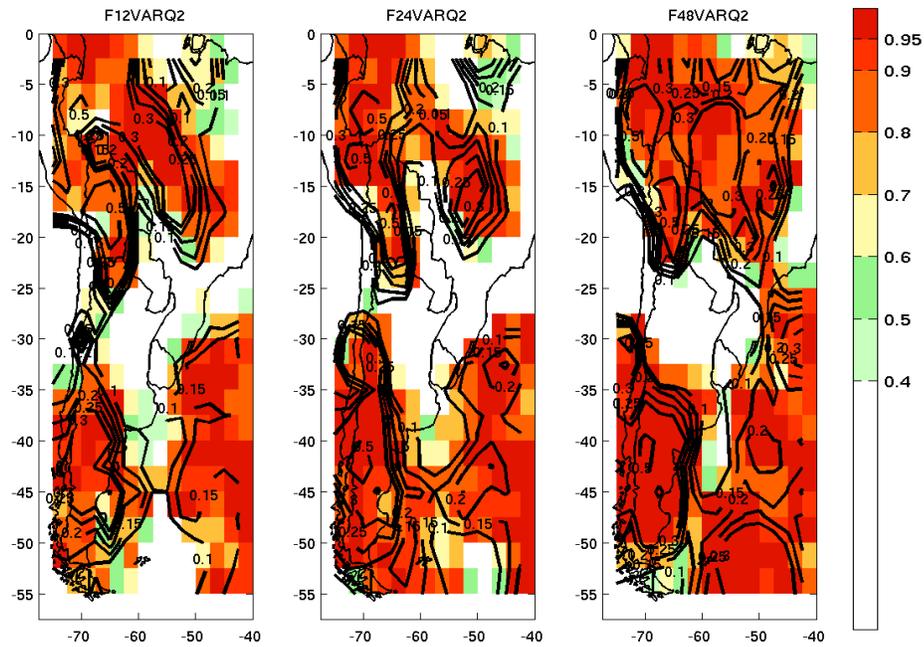
**Figura 5.9:** Error (azul) y dispersión (rojo) medios sobre diferentes regiones como función del plazo del pronóstico (horas) para el ensamble global (línea continua) y el ensamble regional (línea discontinua). La primera columna corresponde a la temperatura a T2M (K), la segunda a PNM (Pa) y la tercera a Q2M ( $\text{Kg Kg}^{-1}$ ). La primera fila corresponde a la región 1, la segunda a la región 2 y la tercera a la región 3.



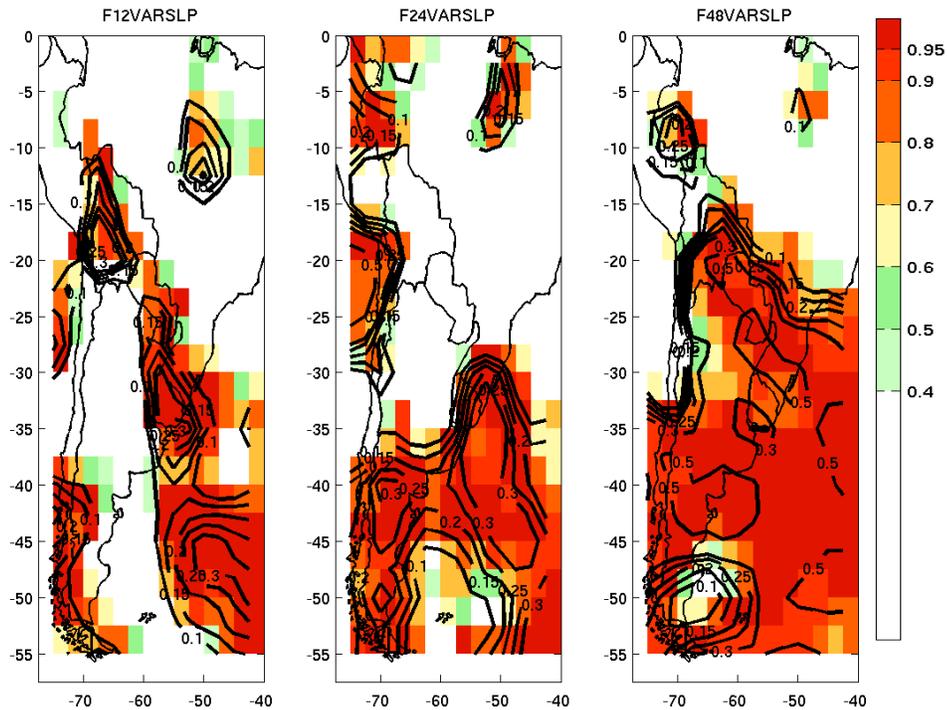
**Figura 5.10: Incremento relativo del error entre 12 y 48 horas de pronóstico para la Q2M (izquierda) e incremento relativo de la dispersión entre 12 y 48 horas de pronóstico para la misma variable para el ensamble global (derecha).**



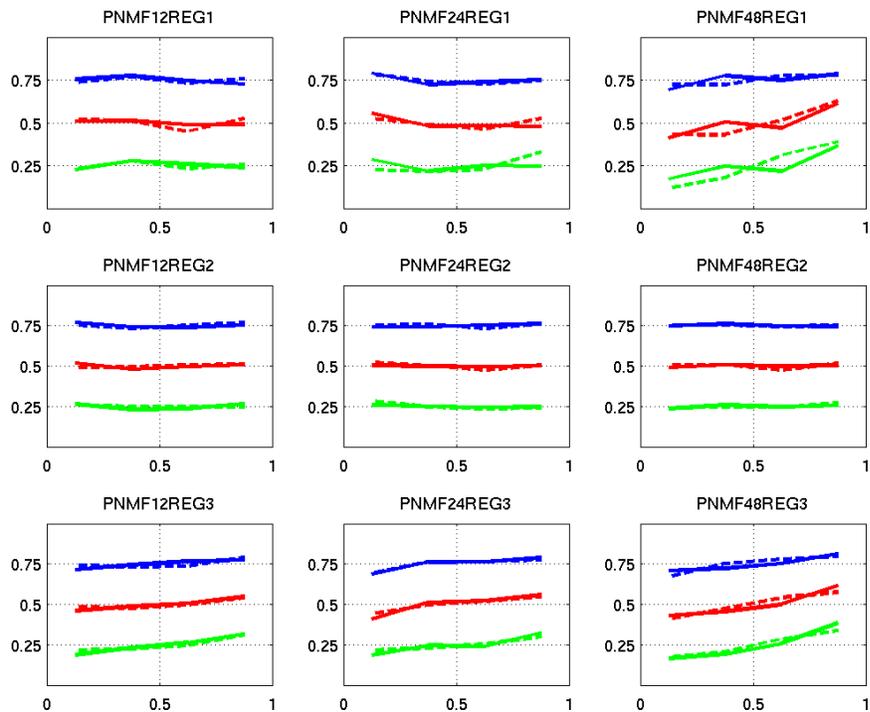
**Figura 5.11: Como en la Figura 5.10 pero para el ensamble regional.**



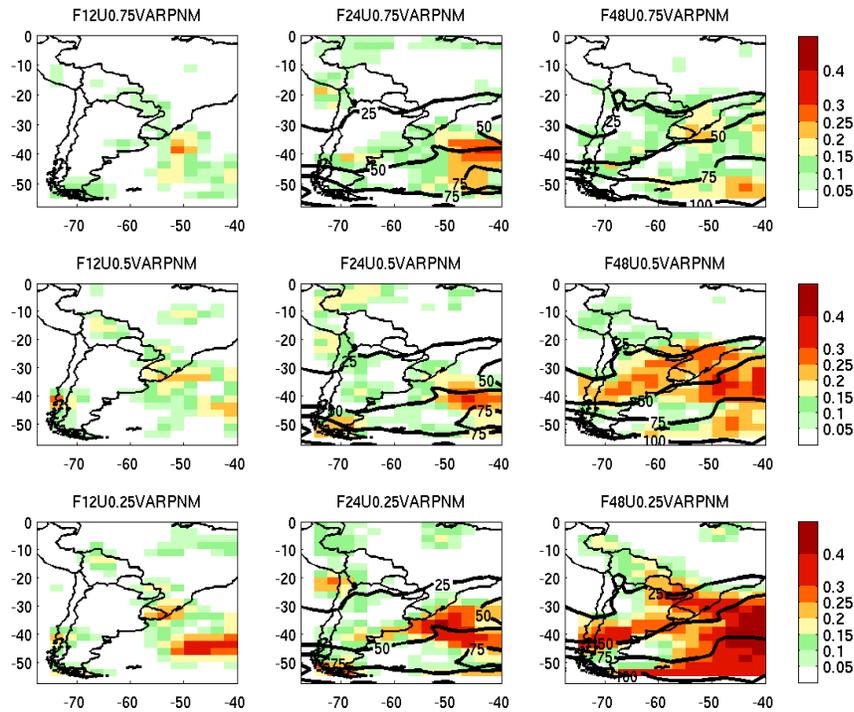
**Figura 5.12: Correlación lineal entre la dispersión del ensamble y la dispersión de los errores para el pronóstico de Q2M (sombreado) a 12 horas (izquierda), 24 horas (centro) y 48 horas (derecha) para el ensamble global. Los contornos negros indican la pendiente de la recta de regresión entre ambas variables.**



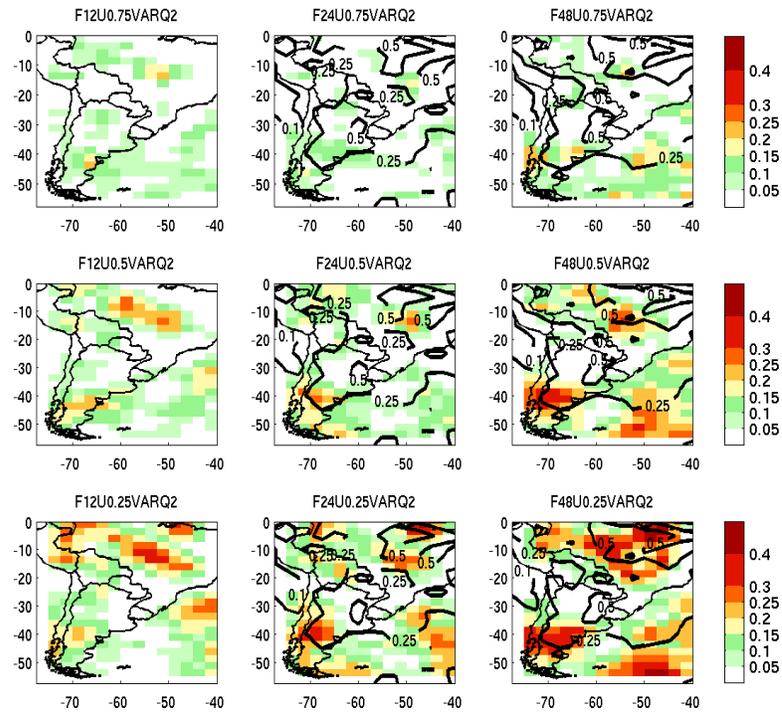
**Figura 5.13: Como la Figura 5.12, pero para la presión a nivel del mar.**



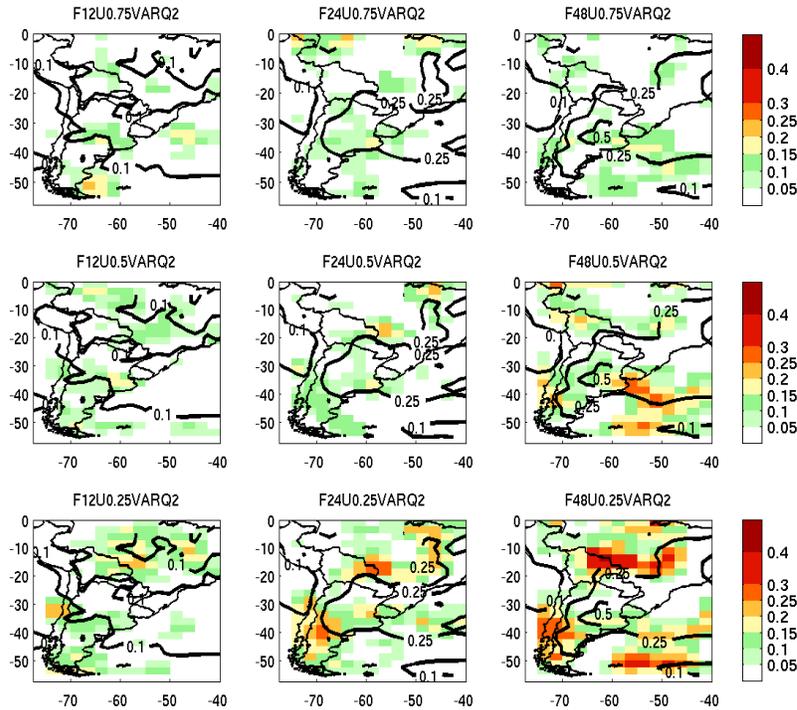
**Figura 5.14: Probabilidad de ocurrencia del error por encima de un determinado umbral como función de la categoría de dispersión (Pcs) para errores cuya probabilidad de ocurrencia climatológica es 25% (línea verde), 50% (línea roja) y 75% (línea azul). La línea continua representa al modelo global, mientras que la línea discontinua al modelo regional. Para la presión a nivel del mar en las 3 regiones seleccionadas, región 1 (primera fila), región 2 (segunda fila) y región 3 (tercera fila) y a distintos plazos de pronóstico 12 (primera columna), 24 (segunda columna) y 48 horas (tercera columna). El eje x designa la frecuencia acumulada correspondiente a los valores de dispersión y el eje y, la probabilidad de ocurrencia de errores por encima del umbral asociado a cada una de las curvas.**



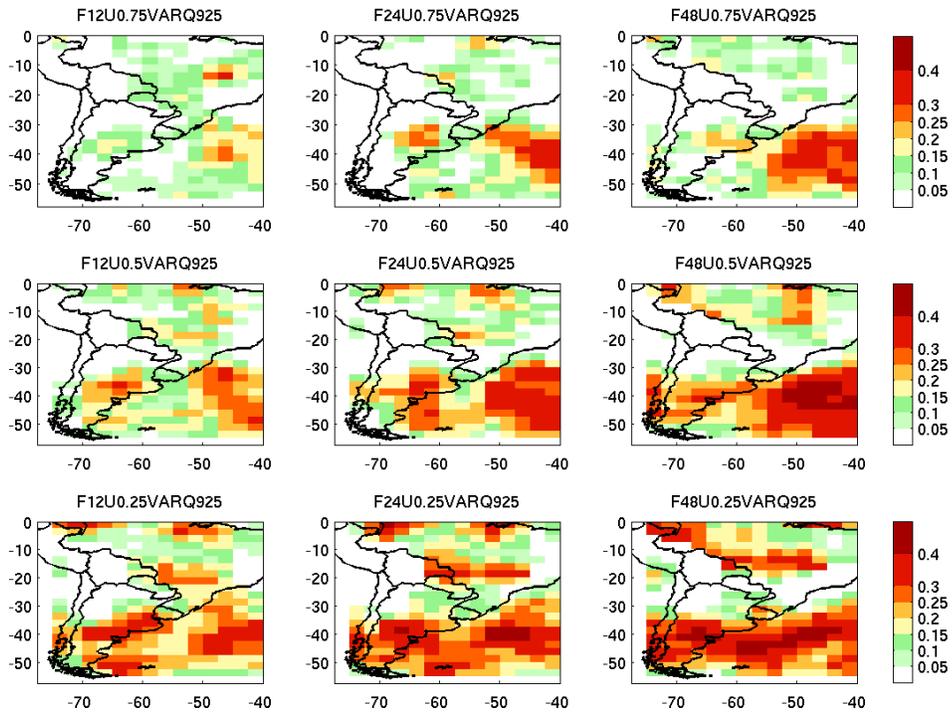
**Figura 5.15: IRSS en sombreado para el ensamble global y para diferentes umbrales de error en la PNM, Pc de 75% (primera fila), Pc de 50% (segunda fila) y Pc de 25 % (tercera fila). Las diferentes columnas muestran los resultados para diferentes tiempos: 12 (primera columna), 24 (segunda columna) y 48 horas (tercera columna). En contornos se muestra la desviación estándar temporal de la dispersión.**



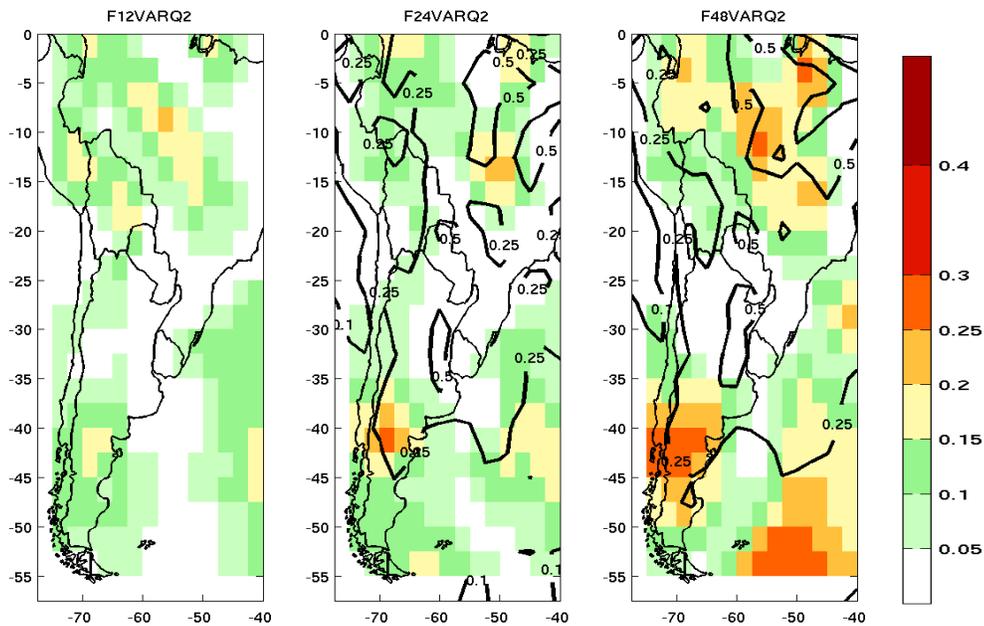
**Figura 5.16: Como en la Figura 5.15, pero para Q2M.**



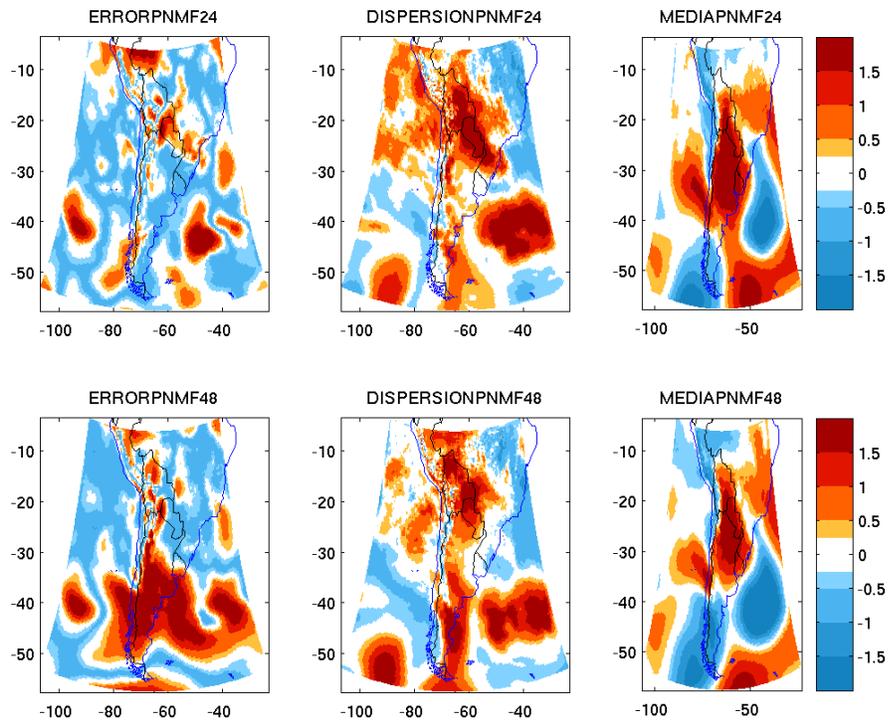
**Figura 5.17: Como en la Figura 5.16 pero para el ensamble regional.**



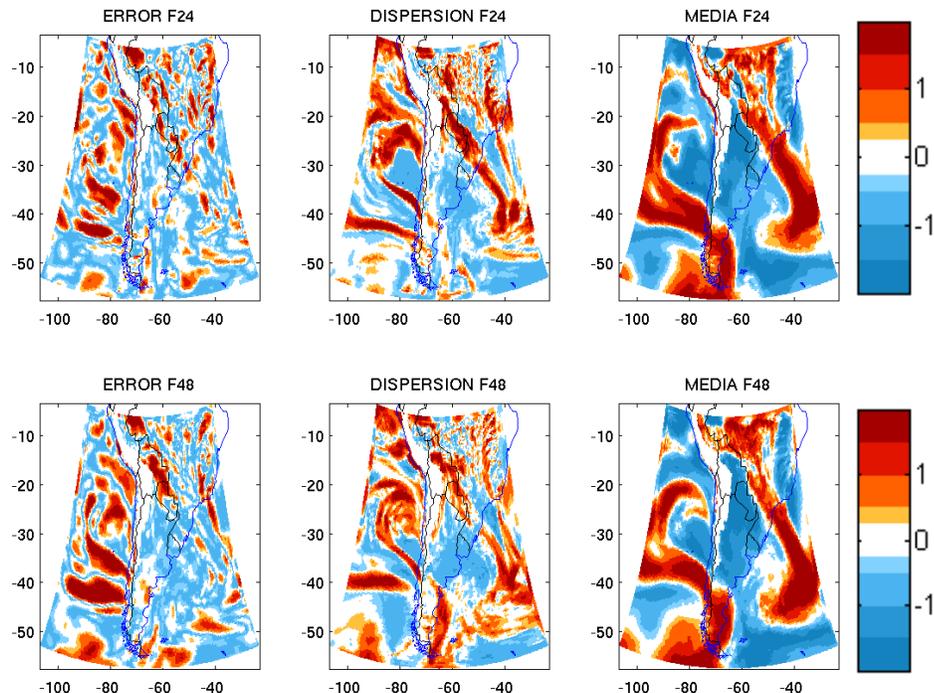
**Figura 5.18:** Como en la Figura 5.17 pero para la humedad específica en el nivel de 925 hPa.



**Figura 5.19:** Correlación lineal (sombreado) entre la dispersión y el error de la media. Para Q2M y para los pronósticos a 12 (izquierda), 24 (centro) y 48 horas de pronóstico.

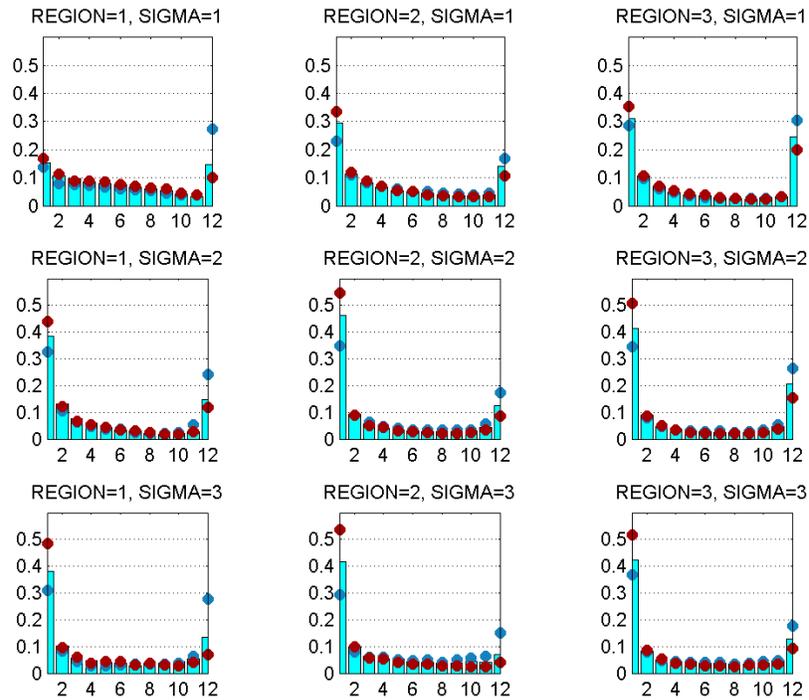


**Figura 5.20: Ejemplo de la relación entre la dispersión y el error estandarizados para un caso particular en el que existió una buena relación entre ambas variables para la PNM (sin dimensiones). La primera columna corresponde al error estandarizado, la segunda corresponde a la dispersión estandarizada del ensamble y la tercera a la media del ensamble estandarizada. Todos los valores corresponden al día 24 de enero a las 00 UTC.**

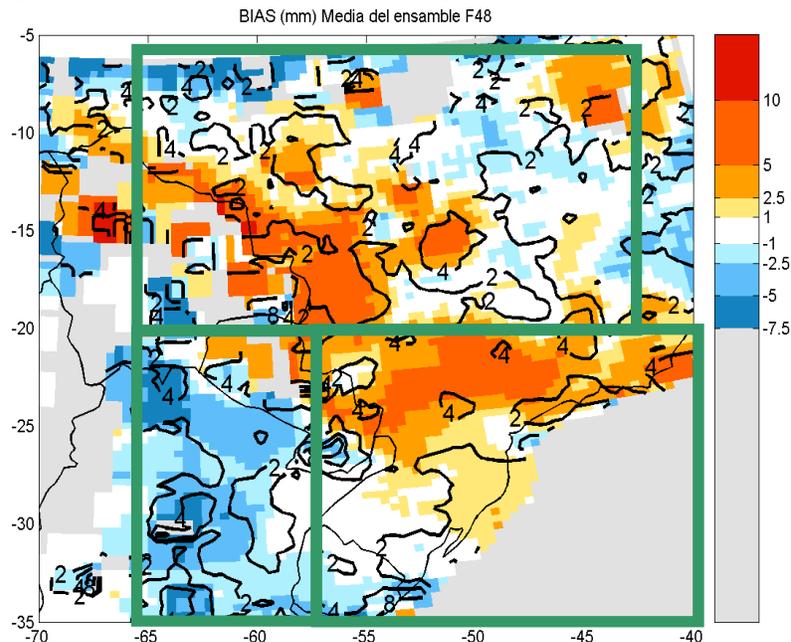


**Figura 5.21: Como en la Figura 5.20, pero para la humedad específica estandarizada (sin dimensiones) en 850 hPa.**

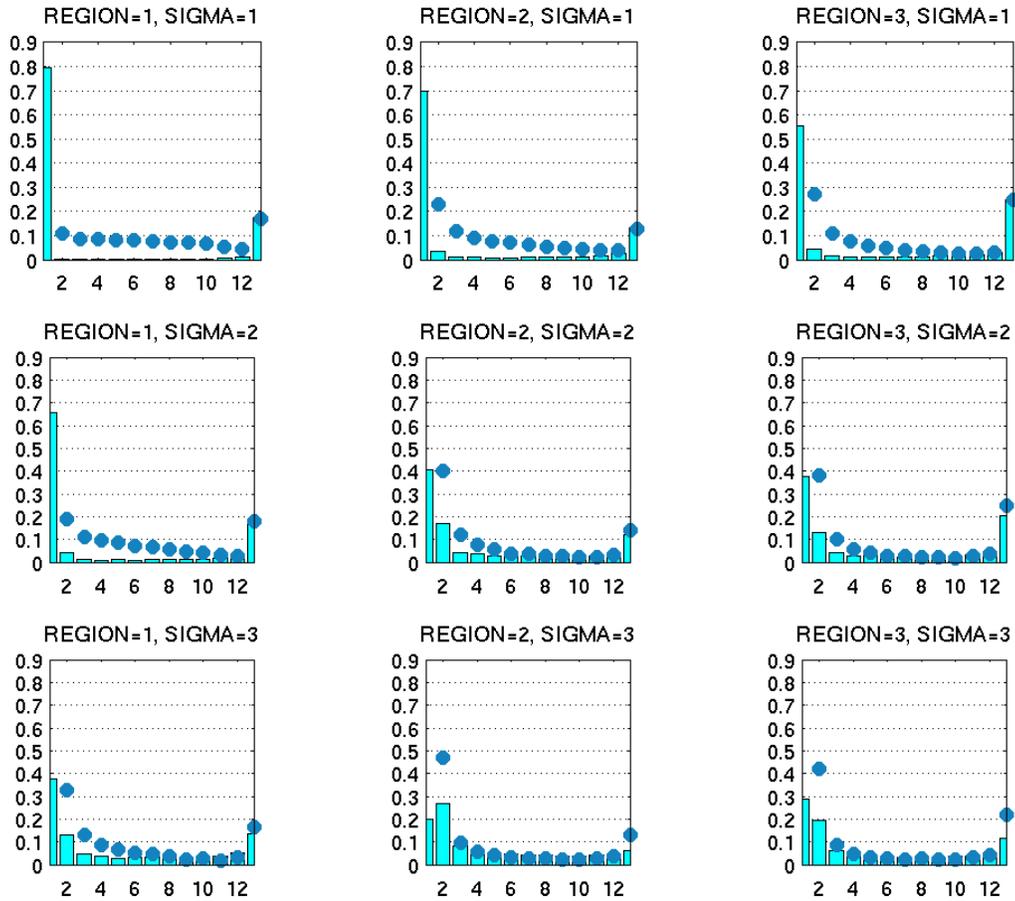
## Figuras Capítulo 6



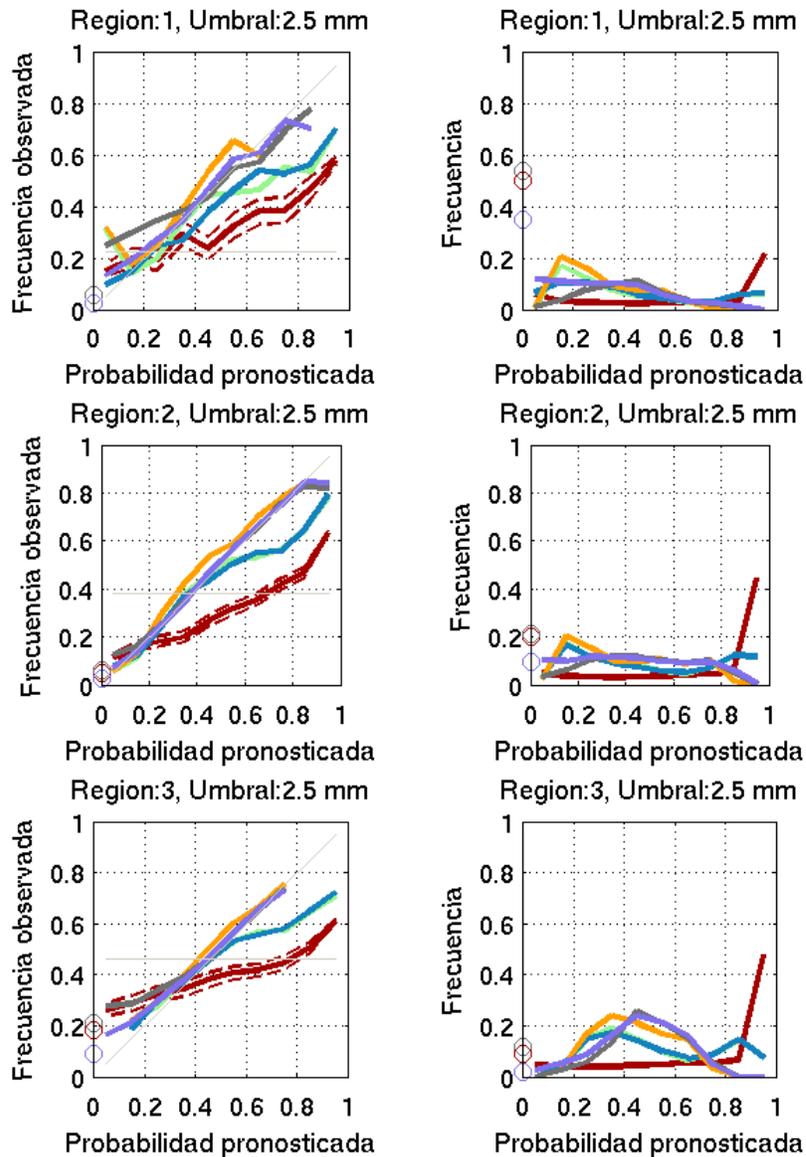
**Figura 6.1: Histograma de rangos obtenido utilizando 3 categorías de dispersión (filas) y 3 categorías de errores sistemáticos del modelo: bajo (puntos azules), medio (barras) y alto (puntos rojos) para el pronóstico a 48 horas en las 3 regiones: norte de Argentina (columna 1), sudeste de Brasil (columna 2) y norte de Brasil (columna 3).**



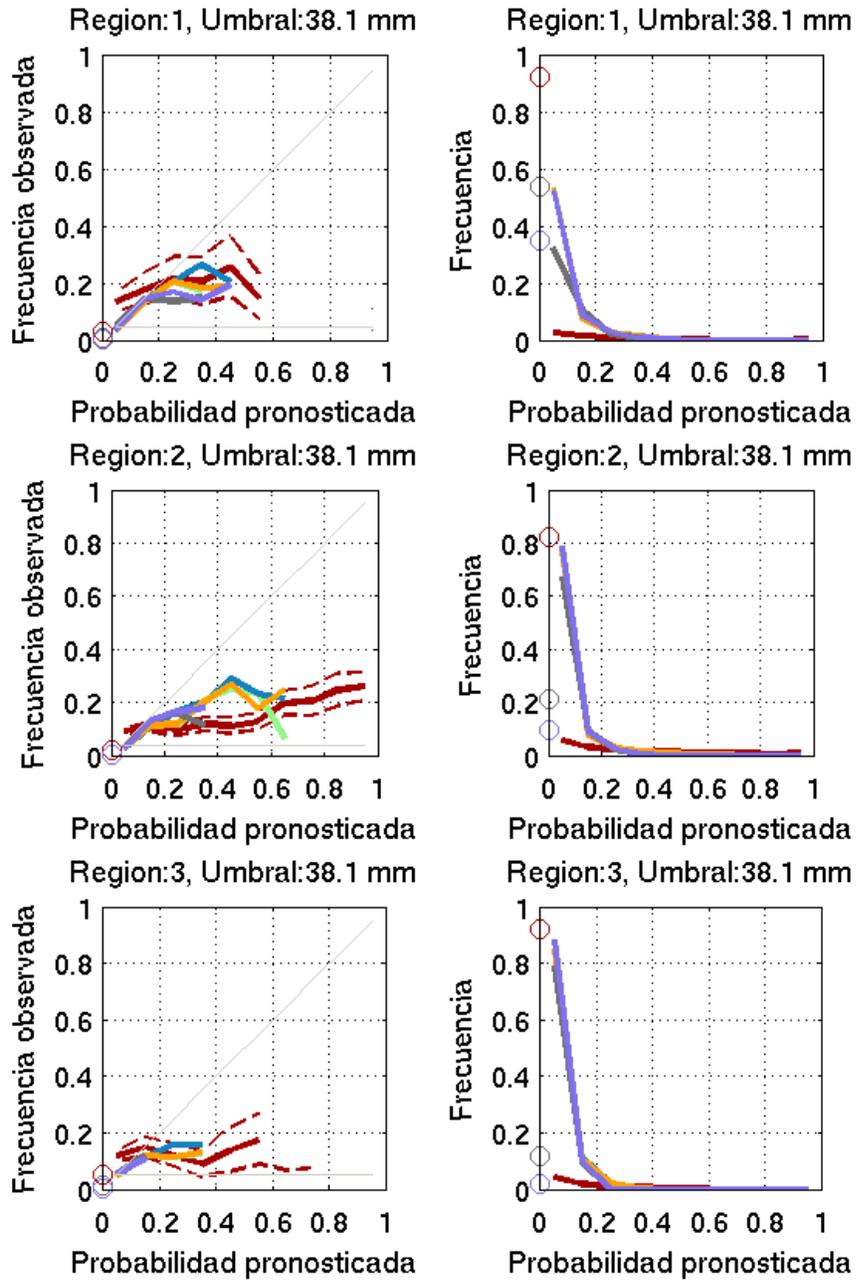
**Figura 6.2: Error sistemático de precipitación para todo el período SALLJEX para los pronósticos a 48 horas ( $\text{mm dia}^{-1}$ ) (sombreado), los contornos indican la desviación estándar de los errores sistemáticos calculados para los períodos de entrenamiento (mm) (contornos).**



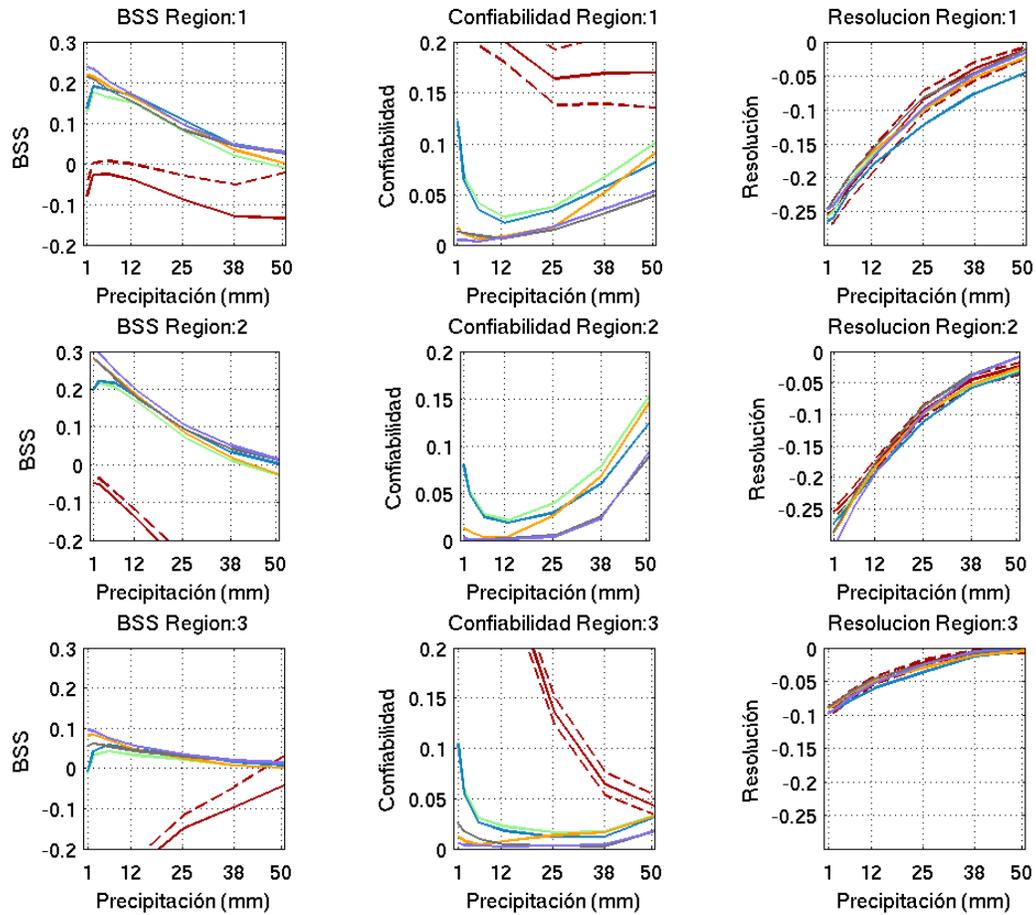
**Figura 6.3:** Histograma de rangos construido como en HC98 (puntos celestes) e histograma de rangos considerando por separado la frecuencia de ocurrencia de ceros en la verificación (barras celestes). La primera del histograma es la frecuencia de ocurrencia de ceros. Las filas representan las diferentes categorías de dispersión y las columnas las diferentes regiones.



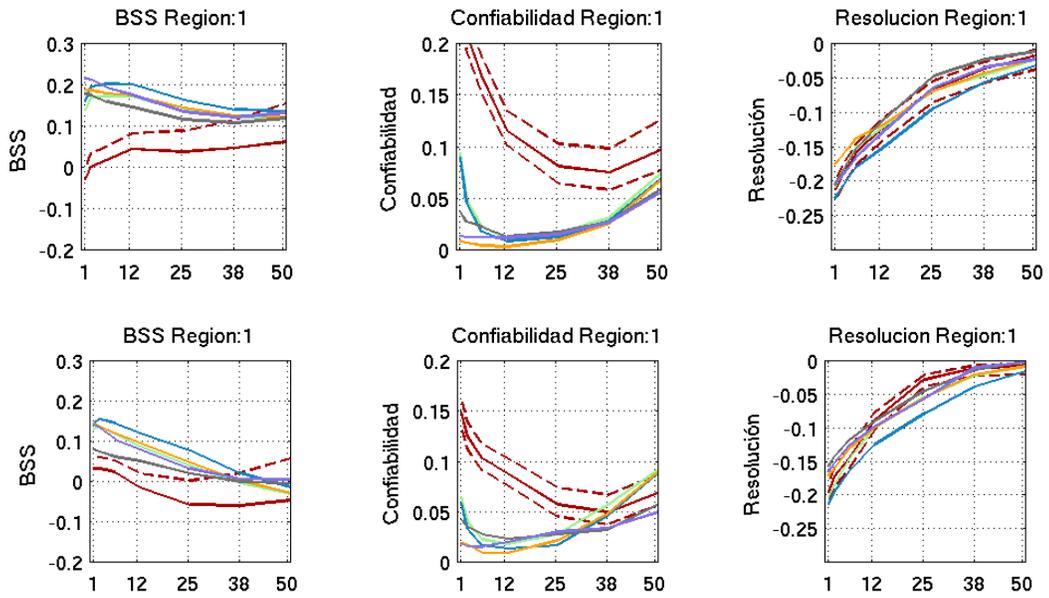
**Figura 6.4: Diagrama de confiabilidad para los pronósticos probabilísticos a 24 horas y el umbral de 2.5 mm. La columna 1 muestra la confiabilidad y la segunda la frecuencia con la que se pronostican los diferentes valores de probabilidad. La fila 1 representa la región norte de Argentina, la fila 2 la región sudeste de Brasil y la fila 3 la región norte de Brasil. Pronóstico no calibrado (línea bordó continua) y sus intervalos de confianza al 5% (línea bordó discontinua), calibración 1 (línea verde), 2 (línea naranja), 3 (línea turquesa), 4 (línea violeta) y 5 (línea gris). Las líneas grises discontinuas marcan las líneas de no resolución y la línea correspondiente al pronóstico perfecto. Los círculos en el lado izquierdo del gráfico representan lo que ocurre cuando la probabilidad de precipitación pronosticada es 0.**



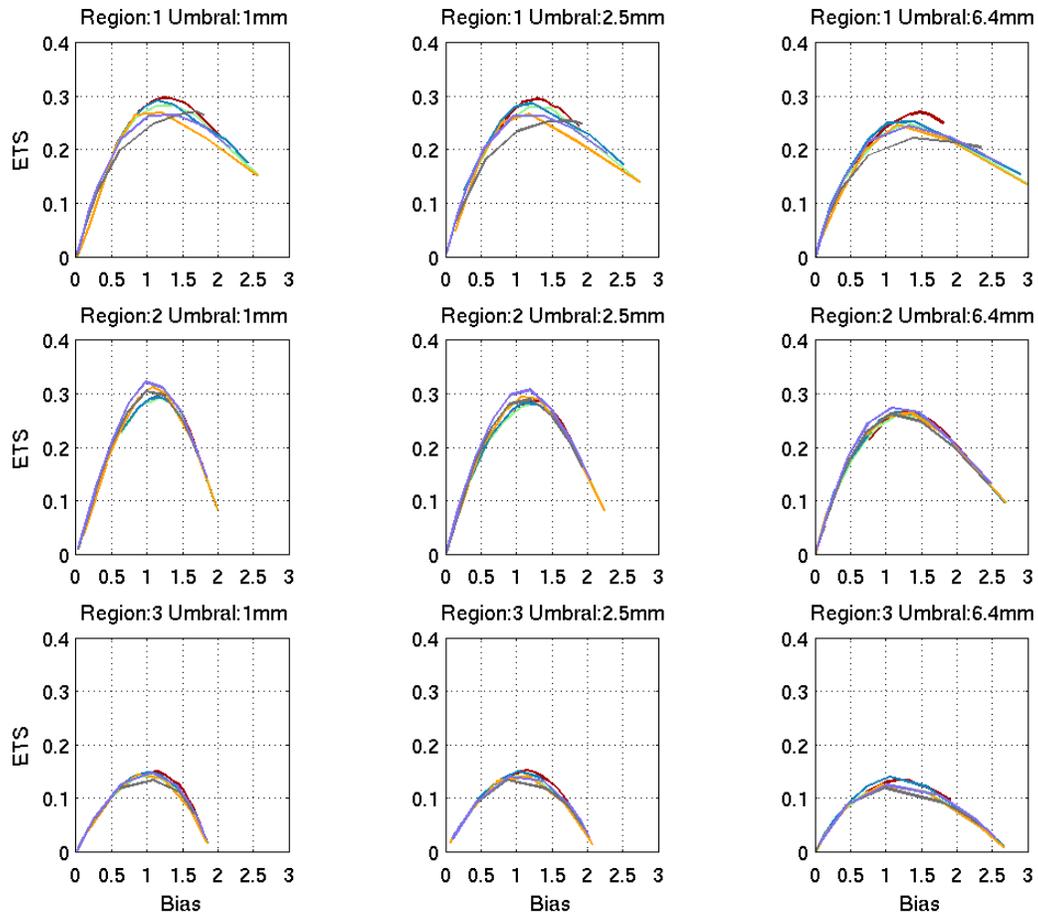
**Figura 6.5: Como en la Figura 6.4 pero para el umbral de 38.1 mm.**



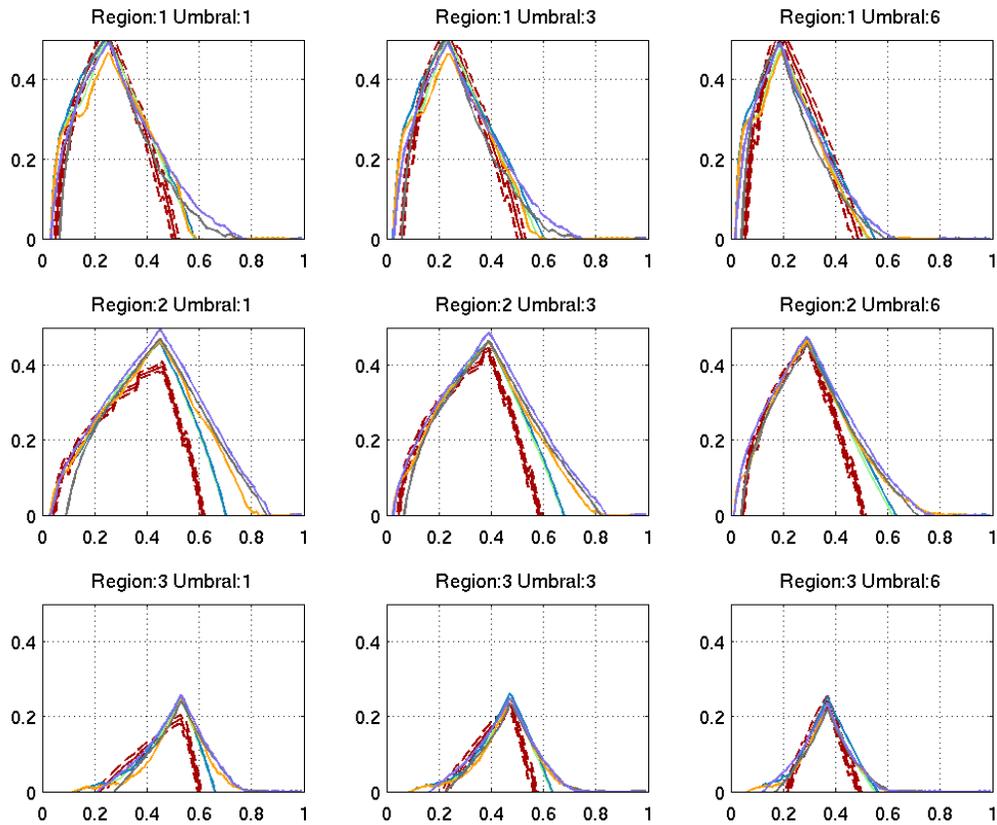
**Figura 6.6: Brier Skill Score como función del umbral (columna izquierda), componente de confiabilidad (columna del centro) y componente de resolución (columna de la derecha) para la región 1 (primera fila), región 2 (segunda fila) y región 3 (tercera fila). Para el pronóstico calibrado y las diferentes calibraciones a 24 horas de plazo. Los colores de las líneas son como en la Figura 6.4.**



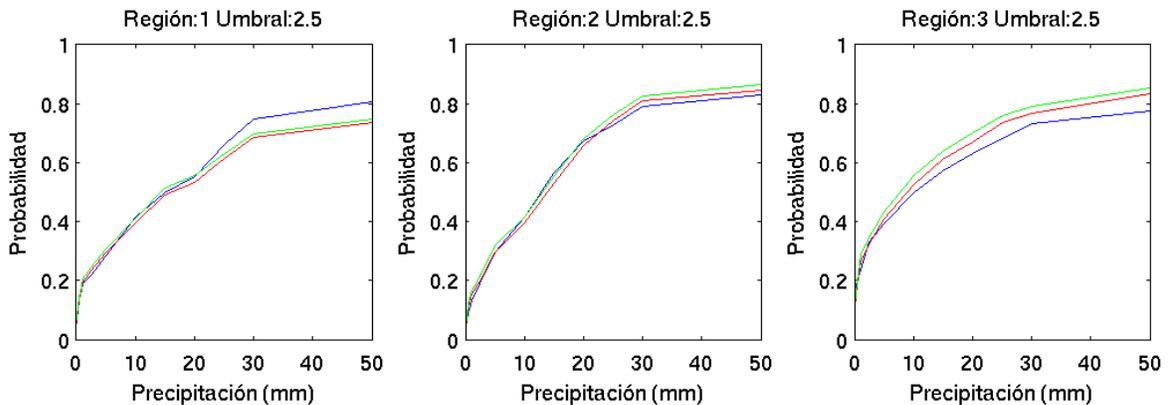
**Figura 6.7:** Como en la Figura 6.6, pero para la región 1. La fila superior muestra los resultados a 36 horas de pronóstico y la fila inferior los resultados a 48 horas de pronóstico.



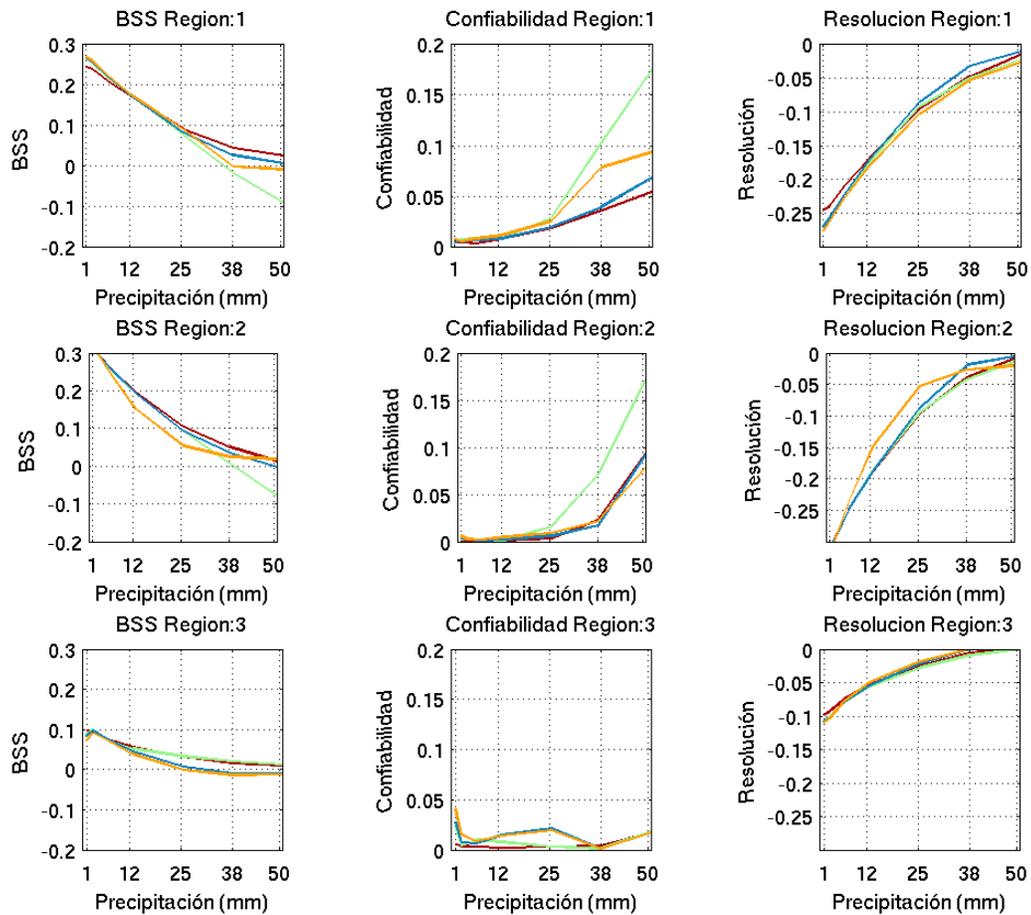
**Figura 6.8: Diagrama ETS-bias para las calibraciones 1 a 5. Umbral de 1 mm (primera columna), 2.5 mm (segunda columna) y 6.35 mm (tercera columna). La primera fila corresponde a la región 1, la segunda a la región 2 y la tercera a la región 3. En todos los casos los resultados son para el pronóstico a 24 horas de plazo.**



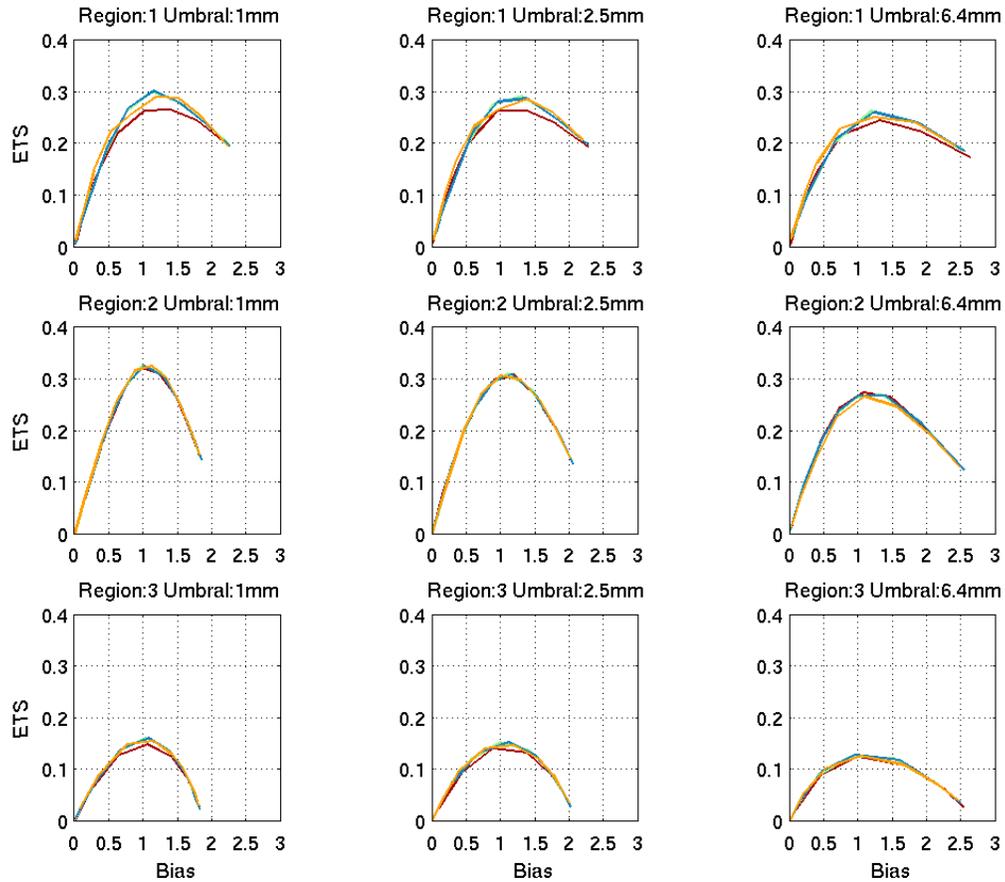
**Figura 6.9:** Valor relativo de los diferentes pronósticos probabilísticos considerados como función de la relación C/L, para el umbral de 1 mm (primera columna) 2.5 mm (segunda columna) y 6.35 mm (tercera columna) y para las 24 horas de pronóstico. Región 1 (panel superior), región 2 (panel del medio) y región 3 (panel inferior). Los colores de las líneas son como en la Figura 6.4.



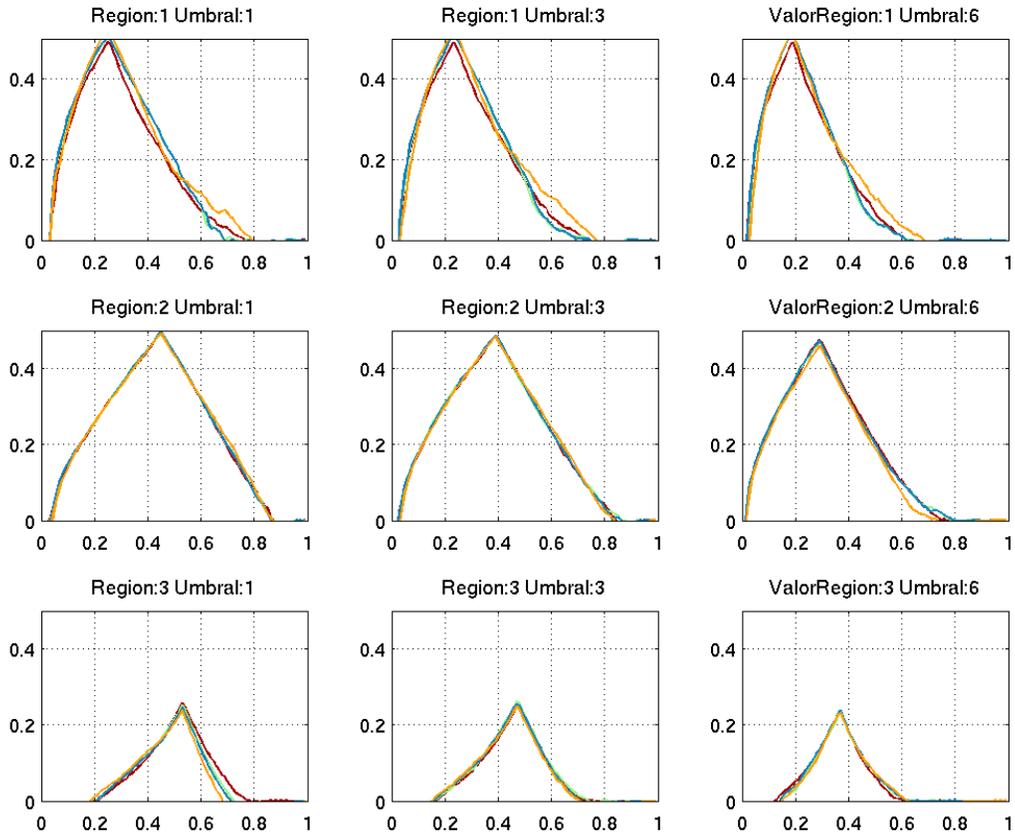
**Figura 6.10:** Curvas de calibración para la media del ensemble calculadas a partir de los datos pluviométricos (curva azul), las estimaciones CMORPH (curva roja) y las estimaciones CMORPH calibradas (curva verde). Para la región 1 (izquierda), 2 (centro) y 3 (derecha).



**Figura 6.11: Como en la Figura 6.5, pero para la calibración en base a pluviómetros (rojo), CMORPH sin calibrar (verde), CMORPH calibrados (azul) y CMORPH local (amarillo). En todos los casos para el pronóstico a 24 horas y utilizando la calibración 4.**

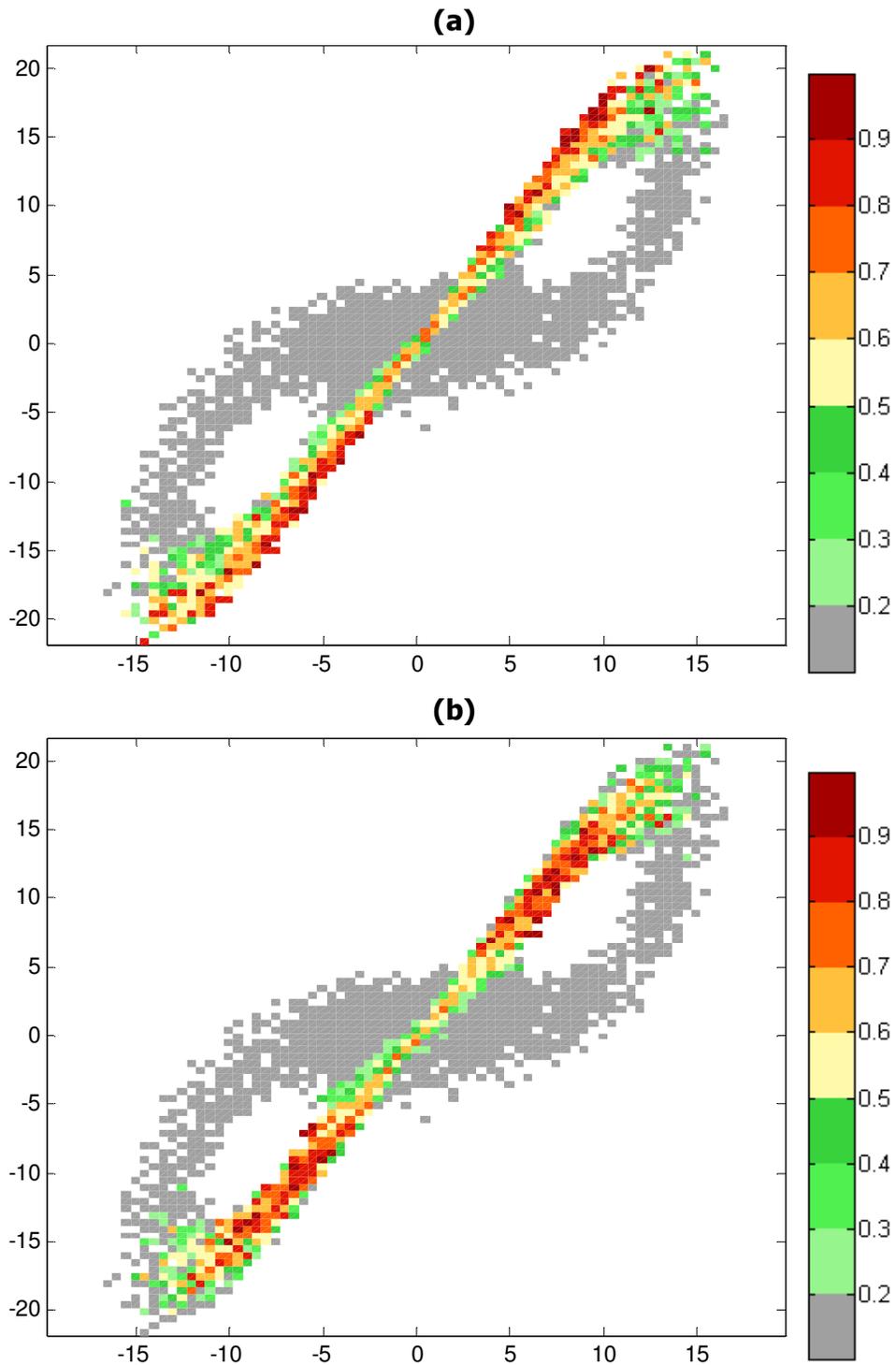


**Figura 6.12: Diagrama ETS-bias para el pronóstico calibrado con datos pluviométricos (rojo), CMORPH (verde), CMORPH calibrado (azul) y calibración local (amarillo). Umbral de 1 mm (primera columna), 2.5 mm (segunda columna) y 6.35 mm (tercera columna). La primera fila corresponde a la región 1, la segunda a la región 2 y la tercera a la región 3. En todos los casos los resultados son para el pronóstico a 24 horas de plazo.**

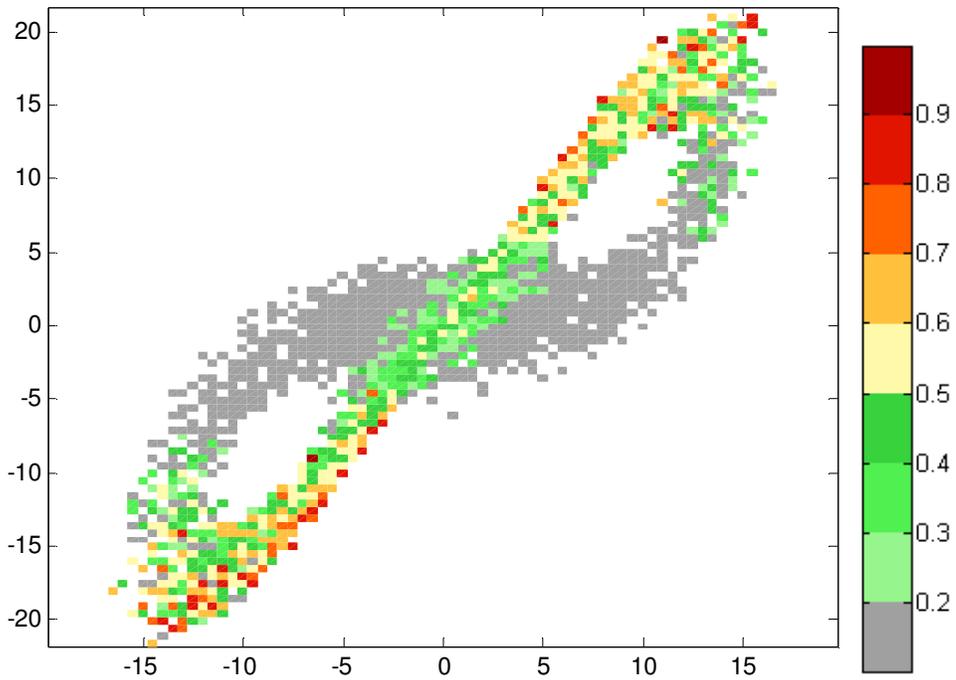


**Figura 6.13: Valor relativo como función del cociente C/L. Umbral de 1 mm (primera columna), 2.5 mm (segunda columna) y 6.35 mm (tercera columna). La primera fila corresponde a la región 1, la segunda a la región 2 y la tercera a la región 3. En todos los casos los resultados son para el pronóstico a 24 horas de plazo. Los colores de las curvas son como en la Figura 6.12.**

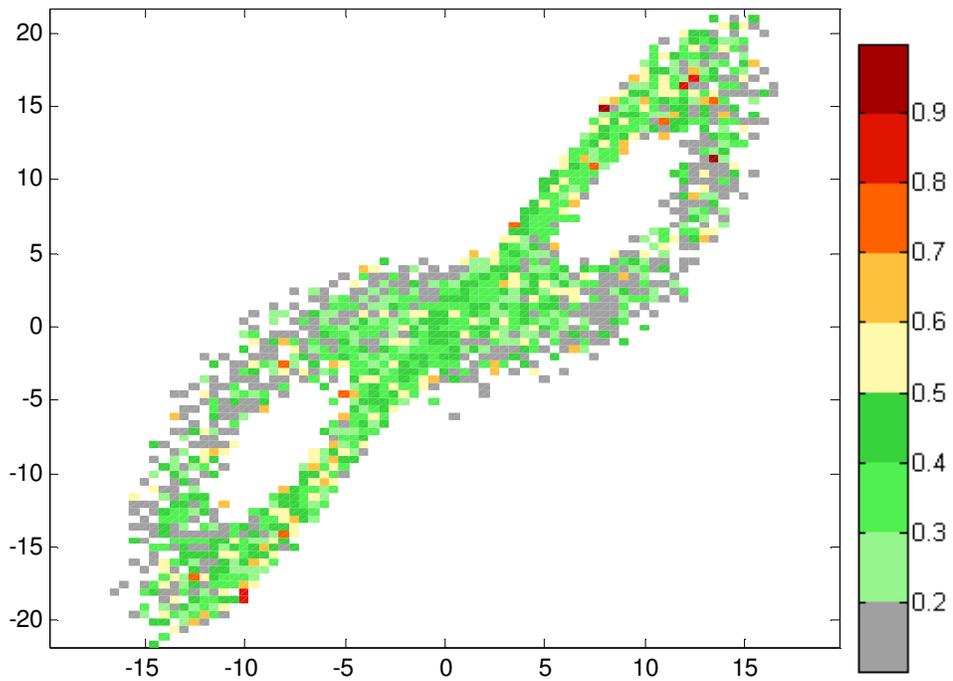
## Figuras Apéndice B



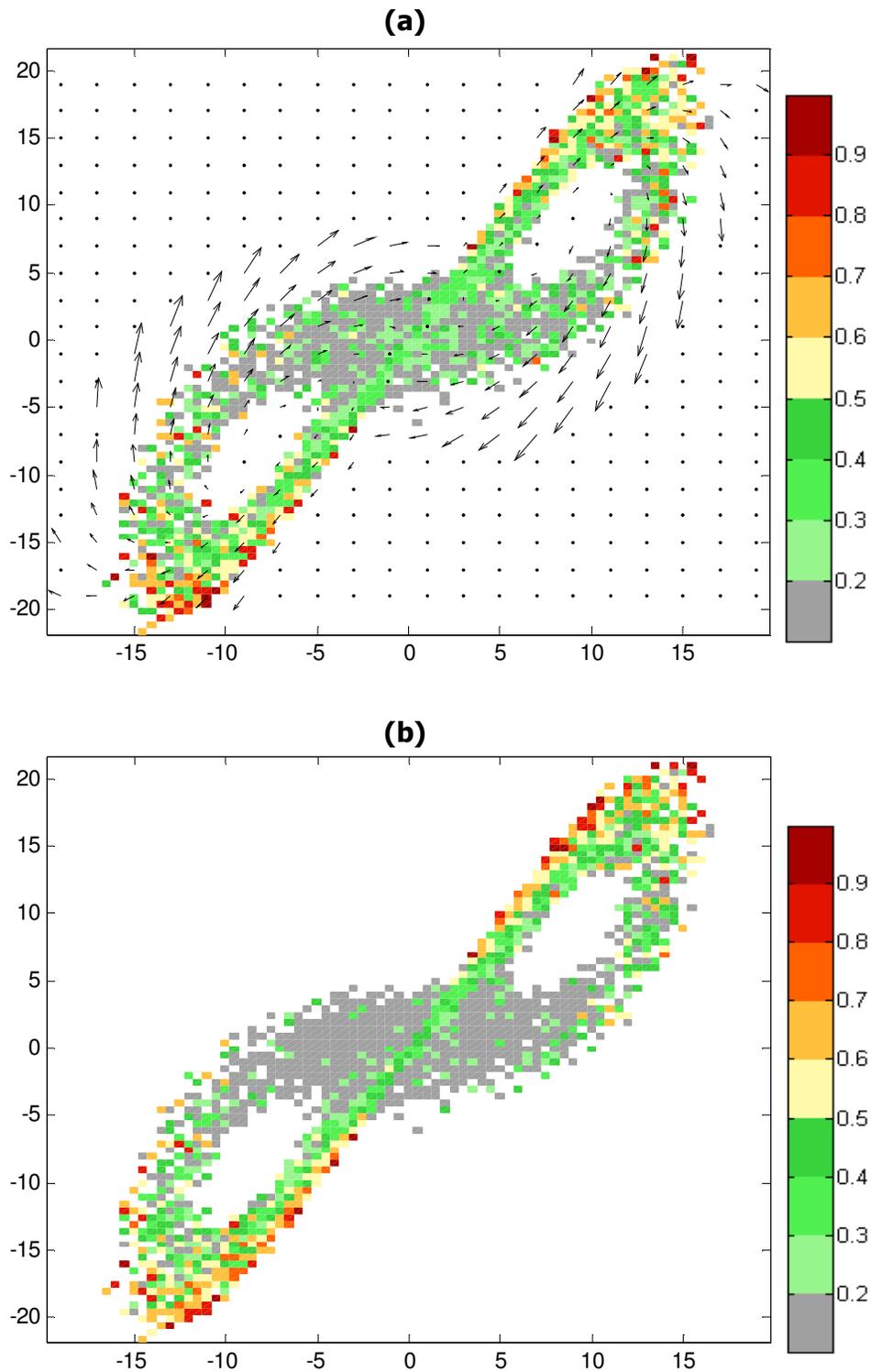
**Figura B.1:** (a) probabilidad de que el crecimiento del error en  $X_b$  sea alto y (b) probabilidad de que el crecimiento en el VC sea alto. En ambos casos como función de la posición en el plano  $X_1$ - $X_2$ .



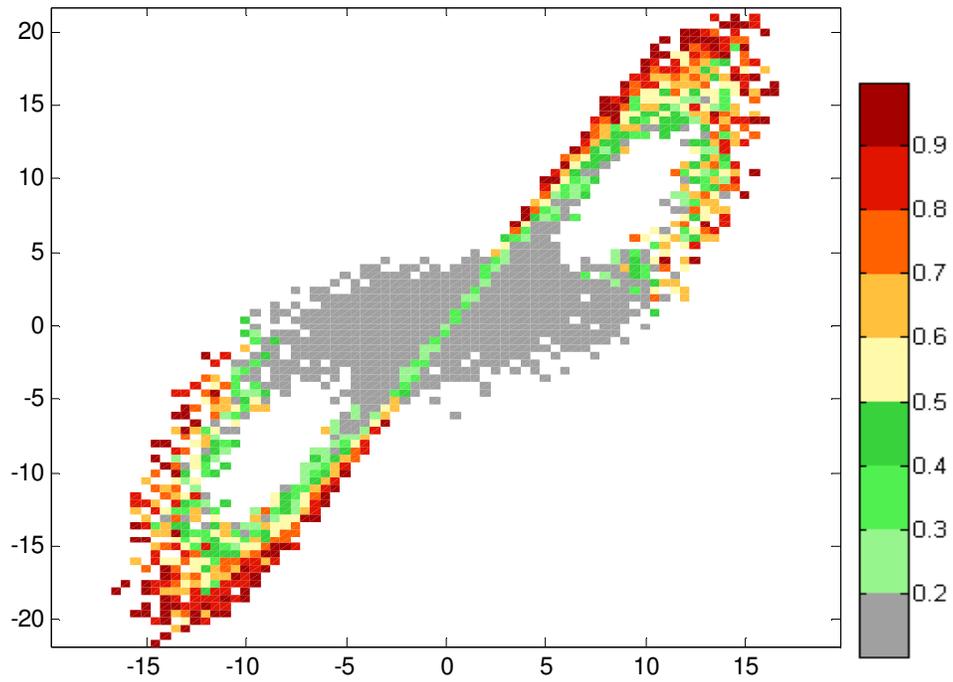
**Figura B.2: Como en la Figura B.1 pero para el crecimiento de la perturbación al azar.**



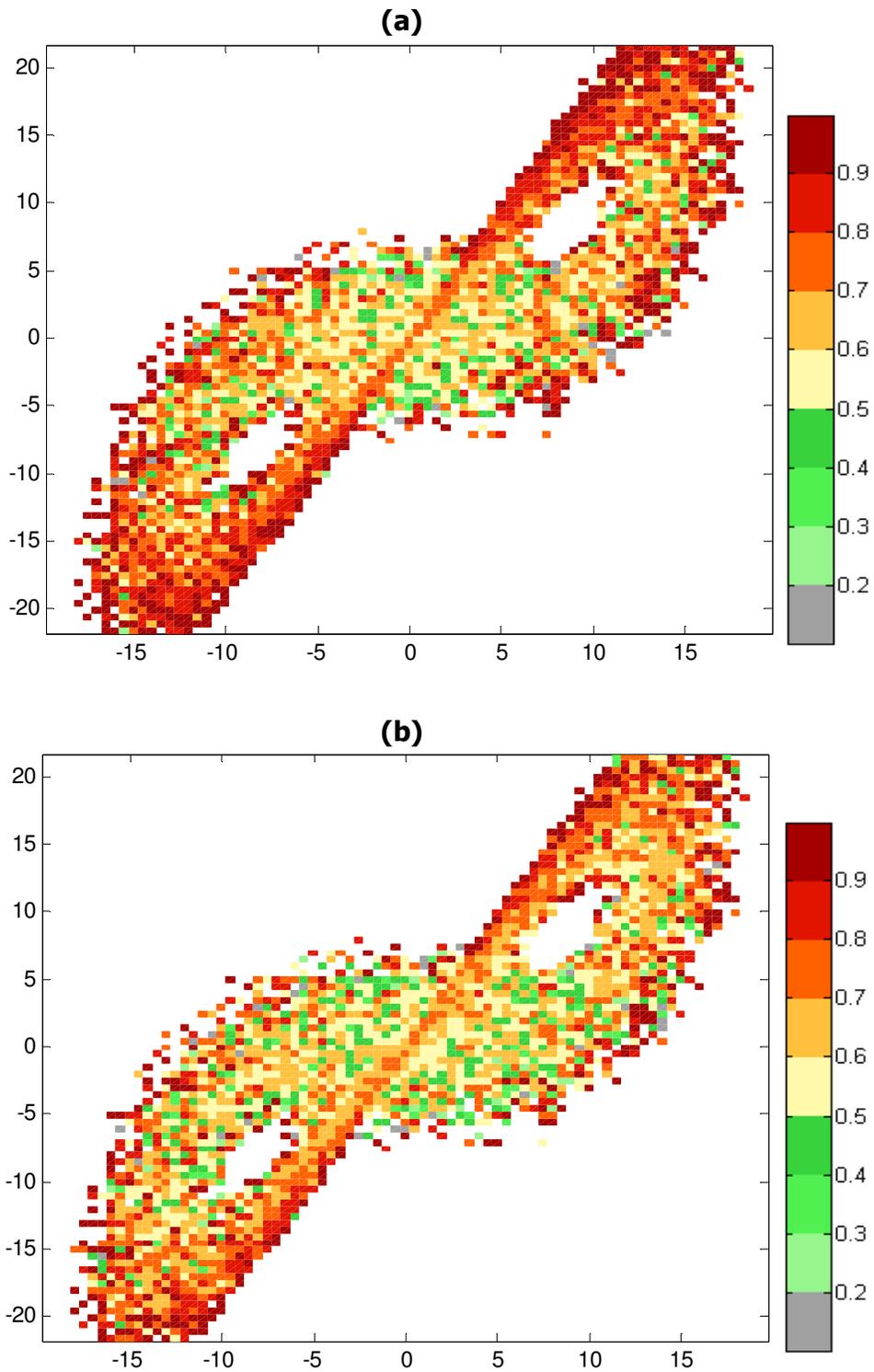
**Figura B.3: Ídem Figura B.1 pero para el crecimiento en el error del análisis.**



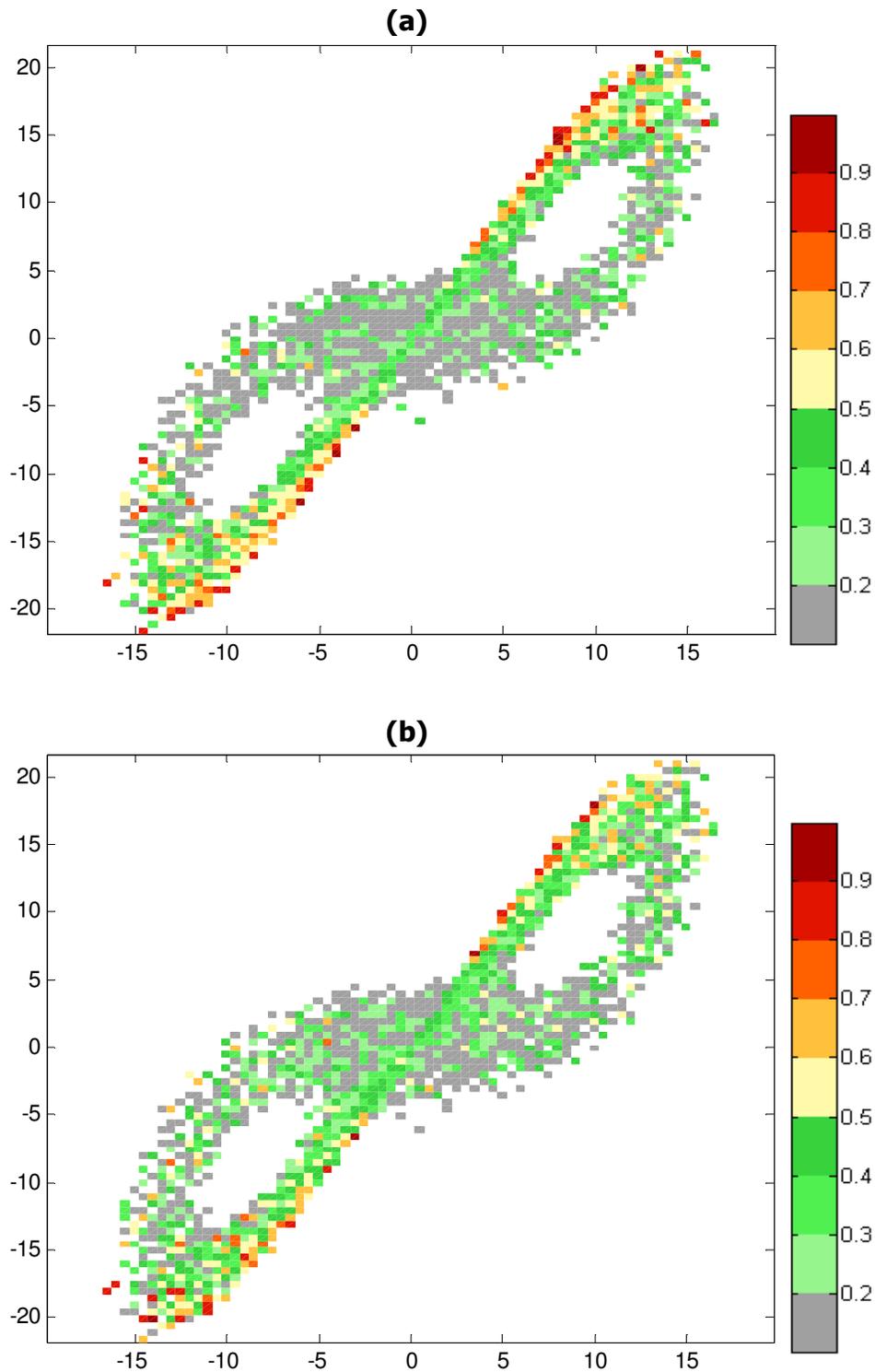
**Figura B.4: (a) Probabilidad de tener un error elevado en el análisis como función de la posición en el plano  $X_1$ - $X_2$  (sombreado) y velocidad de desplazamiento media de las trayectorias del sistema en el plano  $X_1$ - $X_2$  (vectores), (b) ídem (a) pero para la probabilidad de tener un error elevado en el campo preliminar.**



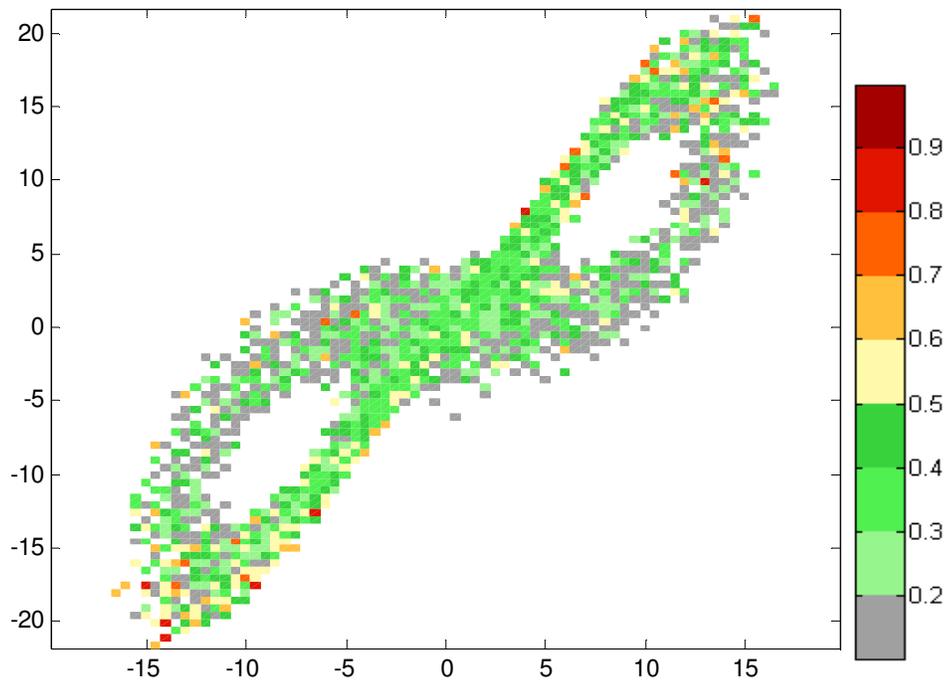
**Figura B.5: Como en la Figura B.4, pero para la probabilidad de tener un vector criado de módulo alto.**



**Figura B.6: Valor medio de la proyección entre (a) el VC y los errores del campo preliminar y (b) el VC y los errores del análisis.**



**Figura B.7: Probabilidad de tener un valor de proyección alto entre (a) el VC y el error en  $X_b$  y (b) el VC y el error en  $X_a$ .**



**Figura B.8: Probabilidad de tener un valor de proyección alto entre la perturbación al azar y el error en  $X_a$ .**