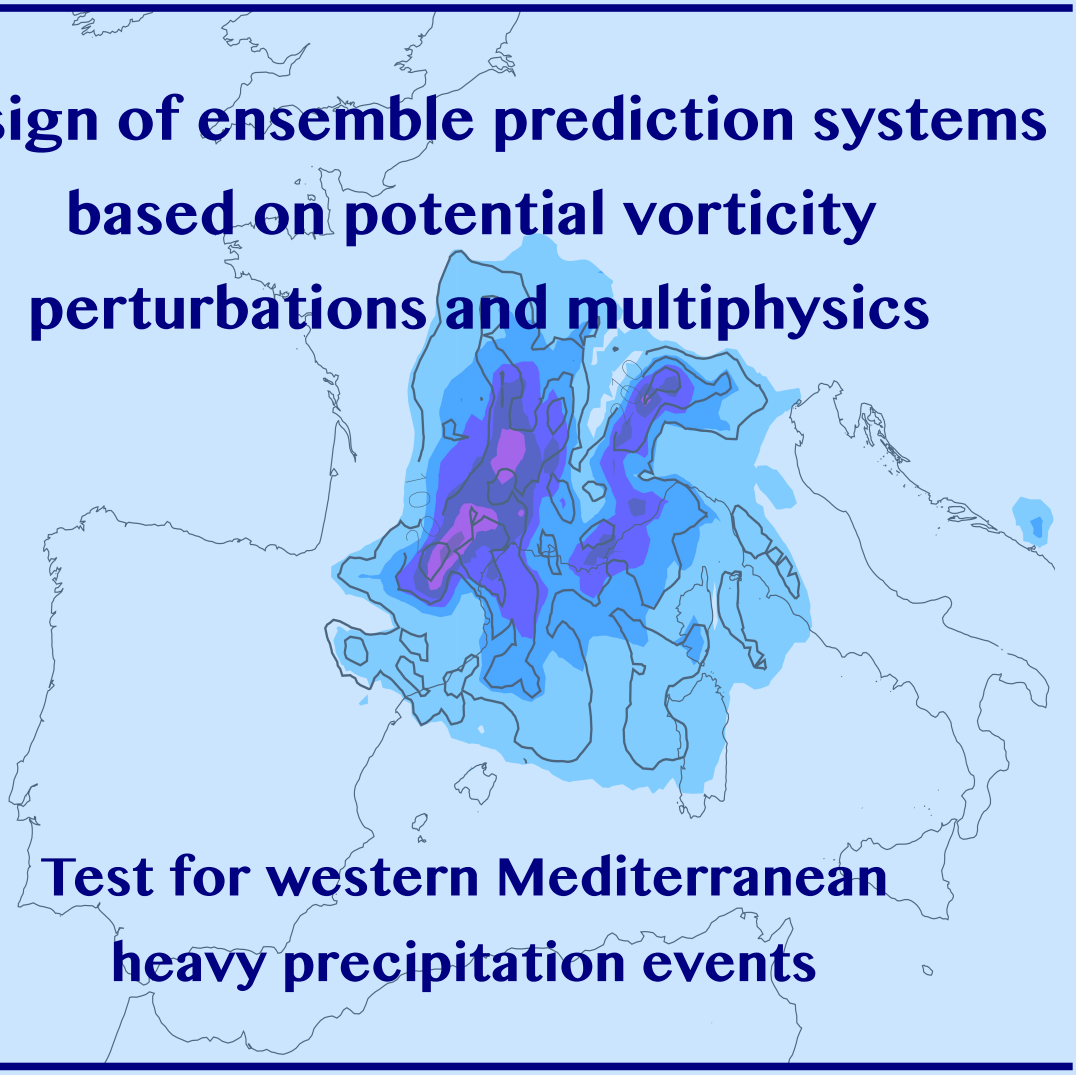
UNIVERSITAT DE LES ILLES BALEARS
DEPARTAMENT DE FÍSICA

Design of ensemble prediction systems based on potential vorticity perturbations and multiphysics



Test for western Mediterranean heavy precipitation events

Maria del Mar Vich Ramis

Grau de Doctora en Física
Programa Oficial de Postgrau en Física

Universitat de les Illes Balears
Departament de Física

Design of ensemble prediction systems based on potential vorticity perturbations and multiphysics

Test for western Mediterranean heavy precipitation events

Tesi presentada per **Maria del Mar Vich Ramis** *per optar al Grau de Doctora en Física per la Universitat de les Illes Balears*

Director de la tesi, Dr. Romualdo Romero March, *Catedràtic de la Universitat de les Illes Balears*

Programa Oficial de Postgrau en Física
Universitat de les Illes Balears

A en Roberto,

Als meus pares,

Agraïments - Acknowledgements

Primer de tot, vull agrair al meu director, prof. Romualdo Romero, l'oportunitat de poder dedicar-me a la recerca dins un camp tan apassionant com és la meteorologia. Des del primer dia, la seva gran capacitat investigadora i comunicadora ha estat un gran exemple a seguir. Sense el seu suport i guiatge, aquesta tesi no hauria estat possible. Moltes gràcies, Romu, per tot el que m'has ensenyat durant aquests anys.

Voldria destacar el gran ambient de feina que caracteritza el grup de Meteorologia de la UIB. Al Dr. Víctor Homar, li vull agrair la seva inestimable ajuda durant aquests anys, com quan em va ensenyar a usar el model adjunt o quan ens vàrem trobar a Norman; al Dr. Joan Cuxart i a la Dra. Maria Antònia Jiménez, l'experiència del primer article publicat; als professors Sergio Alonso i Climent Ramis, el suport i la proximitat. En aquests agraïments, no poden faltar-hi la resta de companys que completen el grup: els estudiants i els tècnics informàtics. Sempre hem estat una gran família. Especialment la ja Dra. Lorena Garcies: moltes gràcies, Lorena, per totes les petites converses de despatx, tant les professionals com les personals, sempre enormement enriquidores.

I am very grateful to Dr. Harold Brooks for the unique opportunity of visiting the National Severe Storm Laboratory in Norman, Oklahoma. During my three-month stay, he taught me how to verify a forecast and also how to take shelter if there was a tornado warning... all fascinating! Even now, each time we meet in a conference, you teach me something new. So, thank you Harold.

I am also very glad to thank Dr. Evelyne Richard for inviting me to visit the Laboratoire d'Aérologie. Some of the results included in this thesis come from this fruitful stay in Toulouse, which was very positive from both a personal and a scientific perspective. *Merci beaucoup*, Evelyne, for your help and advice during my stay and for the nice working environment.

També vull agrair a tota la gent de l'AEMET, amb els quals he col·laborat en diversos projectes comuns, especialment al Dr. Agustí Jansà i a Joan Campins, el suport i el tracte excel·lent que m'han dispensat sempre que he necessitat la seva ajuda, com ara el subministrament de les observacions utilitzades al llarg de tota aquesta tesi. També vull agrair al Ministerio de Ciencia e Innovación la concessió d'una beca de formació de personal investigador (BES-2006-14044), la

qual m'ha permès desenvolupar part d'aquesta tasca, així com el suport econòmic que m'ha ofert a través dels projectes PRECIOSO (CGL2005-03918/CLI) i MEDICANES (CGL2008-01271/CLI). Tampoc no puc oblidar el suport del Departament de Física de la UIB durant aquests anys, en especial pel que fa a la plaça de professora ajudant que m'ha permès acabar aquesta tesi i m'ha donat l'oportunitat d'adquirir una valuosa experiència docent.

El meu darrer agraïment, però segurament el més important, és per a la meva família. A ma mare i mon pare, pel suport incondicional al llarg dels anys i les paraules d'encoratjament quan m'han fet falta. Als meus padrins, que sempre han estat disposats a donar una mà quan ho he necessitat. A en Roberto perquè sempre hi és, tant per celebrar el èxits com per superar els entrebancs, i pel suport i la motivació constants. Moltíssimes gràcies a tots, sense vosaltres aquesta tesi no seria el que és.

Maria del Mar Vich Ramis
Palma, 12 de març de 2012

Resum

La regió de la Mediterrània occidental es veu sovint afectada per esdeveniments de pluja intensa. Aquests esdeveniments acostumen a tenir un gran impacte en la societat a causa de les pèrdues econòmiques, les lesions personals i les morts que causen. S'ha establert una relació entre alguns d'aquests esdeveniments de pluja intensa i la presència d'un cicló sobre la regió. La Mediterrània occidental és una àrea molt ciclogènica a causa, principalment, de la topografia complexa que envolta un mar quasi tancat. Per tant, qualsevol millora de la nostra capacitat de predir aquests esdeveniments potencialment perillosos suposaria importants beneficis per a les societats afectades. Els serveis de protecció civil podrien aprofitar aquestes previsions millorades per implementar mesures de prevenció més eficients i advertir la població amb més antelació. Una resposta més ràpida podria marcar la diferència entre danys major i menors, o fins i tot entre la vida i la mort.

L'objectiu principal d'aquesta tesi és millorar l'actual capacitat de predicció d'aquests fenòmens meteorològics potencialment perillosos. Desenvolupam aquesta tasca mitjançant un sistema de predicció per conjunts (SPC) -*ensemble*, en anglès. Més concretament, desenvolupam i verificam tres sistemes de predicció per conjunts, tots centrats en els fenòmens meteorològics de gran impacte en la Mediterrània occidental. Entre els tres conjunts es tenen en compte tant les incerteses presents en els models numèrics de predicció meteorològica com les presents en les condicions inicials. Consideram les deficiències del model en el conjunt de multifísica. Els membres del conjunt de multifísica es generen variant els esquemes de parametrització física. A més del conjunt de multifísica, utilitzam dos SPC que generam pertorbant les condicions inicials i les condicions de contorn del model, que tenen en compte l'impacte de les incerteses presents en aquestes condicions. Aquestes pertorbacions s'apliquen sobre el camp de vorticitat potencial (VP; *potential vorticity*, PV, en anglès) mitjançant una tècnica d'inversió, aprofitant la connexió entre les estructures de VP i els ciclons. Per això anomenam *PV-perturbed* aquests dos conjunts. Treballar amb el camp de VP té l'avantatge de poder definir les pertorbacions sobre una sola variable, VP, i, juntament amb la tècnica d'inversió de VP, ens asseguram que la resta de camps meteorològics siguin pertorbats sense comprometre el balanç atmosfèric que hi ha entre els camps de vent i de temperatura. Per evitar pertorbacions no realistes, desenvolupam una climatologia d'error de VP. Aquesta climatologia proporciona un marge d'error de VP consis-

tent en el rang d'incertesa propi del camp de VP. Malgrat que els dos conjunts de *PV-perturbed* es basen en pertorbar el camp de VP, el criteri que segueixen per decidir on aplicar les pertorbacions és diferent. Breument, en un conjunt les pertorbacions s'apliquen sobre les zones del camp de VP amb valors més intensos i major gradient (*PV-gradient*), mentre que en l'altre les pertorbacions s'apliquen a les zones de sensibilitat calculades pel model adjunt MM5 (*PV-adjoint*).

Avaluam l'actuació dels tres conjunts mitjançant una col·lecció de denou ciclons d'alt impacte associats a precipitació i a vents forts (tots el ciclons provenen de la base de dades del projecte MEDEX). La verificació se centra en la precipitació a causa de l'impacte social, tot i que som conscients que les seves característiques, és a dir, forts gradients en l'espai i variacions ràpides en el temps, dificulten el procés de verificació. En altres paraules, és difícil obtenir bons resultats a la verificació quan avaluam el camp de precipitació. Per altra banda, la verificació també es veu afectada per la naturalesa extrema i poc freqüent dels episodis d'interès. Els valors extrems de precipitació no sempre són ben capturats pel model, i la base de dades d'observacions de fenòmens extrems i rars no sempre és prou àmplia (mida de la mostra insuficient).

Els nostres resultats, tal com esperàvem, mostren que el sistema de predicció per conjunts proporciona una predicció més hàbil que una predicció determinista. A més, quan els comparam, el conjunt *PV-gradient* es mostra més hàbil que el conjunt de multifísica. Entre els conjunts *PV-perturbed*, l'SPC *PV-gradient* també supera en habilitat el conjunt *PV-adjoint*. Això implica que el cost computacional addicional d'utilitzar el model adjunt MM5 en el conjunt *PV-adjoint* no és compensat en habilitat de predicció. Per tant, el conjunt *PV-gradient* ha resultat ser una estratègia rendible, tant per la simplicitat com per l'habilitat que demostra.

Posteriorment, els tres conjunts s'apliquen a tres situacions diferents per tornar a posar a prova el potencial i la utilitat que tenen. L'èxit d'aquestes aplicacions ha estat divers. La primera aplicació consisteix a construir un *superensemble* basat en el nostre conjunt de multifísica. Els resultats mostren que el cost addicional de construir un *superensemble* en comptes d'una mitjana amb el biaix corregit no es veu compensat en habilitat, ja que les dues prediccions obtenen puntuacions similars en la verificació. La segona aplicació utilitza el conjunt *PV-gradient* per calcular el camp de precipitació necessari per a les simulacions hidrometeorològiques d'escolament. En aquest cas, els resultats mostren que l'ús d'una estratègia basada en la predicció per conjunts dins un sistema hidrometeorològic de

predicció d'escolament millora la dita predicció. Per tant, en aquesta aplicació l'ús de tècniques basades en predicció per conjunts es mostra com un bon mètode per millorar el resultat de les simulacions d'escolament. La tercera aplicació mostra el potencial d'altres mètodes basats en la pertorbació del camp de VP. En aquesta aplicació modificam el camp de VP en base a la informació que ens proporciona el canal de vapor d'aigua del satèl·lit METEOSAT-7. Les modificacions es fan aprofitant la relació teòrica entre el camp de VP a nivells alts i el camp de temperatura de brillantor de vapor d'aigua. Els resultats mostren que quan aplicam una tècnica de pertorbació obtenim una millor predicció que quan no l'aplicam. A més, quan comparam el resultat d'aquesta tècnica de pertorbació amb els resultats dels membres de la predicció per conjunts *PV-gradient* i *PV-adjoint*, veim que es manté dins els rangs dels dos conjunts i és estatísticament indistingible.

Els resultats d'aquesta tesi mostren la utilitat i la idoneïtat dels mètodes de predicció basats en la pertorbació dels tàlvegs de nivells alts, precursors de les situacions ciclòniques. Els resultats i les estratègies presentades pretenen ser un punt de partida per a futurs estudis que facin ús d'aquests mètodes.

Contents

Summary

The western Mediterranean region is often affected by heavy rain events. These events generally have quite an impact on society due to the economic losses, personal injuries and fatalities they cause. The connection between some of these heavy rain events and the presence of a cyclone over the region has been established. The western Mediterranean is recognized as a very cyclogenetic area as a consequence of the complex topography surrounding a quasi-closed sea. As one would expect, any improvement of our capability to predict these potentially dangerous events would represent an important gain for the affected societies. Civil protection services could take advantages of these improved forecasts, so that issuing better preventive measures and earlier warnings would be possible. These quicker reactions could make the difference between major and minor damages or even between life and death.

The main goal of this Thesis is to improve the current prediction skill of these potentially hazardous weather events. We tackle this task by using an ensemble forecasting system approach. We develop and test three different ensemble prediction systems (EPSs) that target western Mediterranean high-impact weather events. By means of these three ensembles we account for uncertainties present in the numerical weather models and in the initial conditions. We deal with the model deficiencies by building a multiphysics ensemble. The members of the multiphysics ensemble are generated by varying the physical parameterization schemes. In addition to the multiphysics ensemble, we use two EPSs built by perturbing the initial conditions and the boundary conditions of the model, so that the impact of the uncertainties of these conditions are accounted for. These perturbations are applied over the potential vorticity (PV) field by means of an inversion technique, exploiting the connection between PV structures and cyclones. So, we call these two ensembles the PV-perturbed ensembles. Working with the PV field has the advantage of defining the perturbations on a single variable, PV, while the PV inversion technique ensures that the rest of meteorological fields are consistently perturbed without compromising the mass-wind balance of the atmosphere. To avoid unrealistic perturbations, we develop a PV error climatology. This climatology provides a PV error range consistent with the PV field uncertainty range. Although the two PV-perturbed ensembles are based on perturbing the PV field, the decision on where to apply the perturbations is different. Briefly, in one en-

semble the perturbations are applied over the most intense values and gradients PV zones (PV-gradient) while in the other the perturbations are applied along the MM5 adjoint model calculated sensitivity zones (PV-adjoint).

We evaluate the performance of the three ensembles on a collection of 19 high-impact cyclones associated with heavy precipitation and strong winds (all cyclones come from the MEDEX database). The verification procedure focuses on the precipitation field due to its direct impact on society, even though we are aware that its features, i.e., strong gradients in space and rapid variations in time, make the verification process difficult. In other words, it is hard to obtain good verification scores when dealing with precipitation. The verification procedure is also affected by the extreme and rare nature of the targeted events. Extreme precipitation values are not always well captured by the model, and the observational database of extreme and rare events is not always as extensive as desired, leading to the well known sampling problem.

Our results show the expected improvement in forecast skill when using an ensemble prediction system instead of a determinist forecast. Moreover, when compared, the PV-gradient ensemble outperforms the multiphysics ensemble. Between the PV-perturbed ensembles, the PV-gradient EPS also outperforms the PV-adjoint ensemble. This means that the extra computational cost derived from running the MM5 adjoint model in the PV-adjoint ensemble is not compensated later in forecast skill. Therefore the PV-gradient ensemble reveals itself as a profitable strategy, owing to both its low computational cost and its competitive skill.

Later, the three ensembles are applied on three different situations to further test their potential and usefulness. These applications result in varying degrees of success. In the first application, we build a superensemble fed by our multiphysics ensemble. The results show that the extra cost of building a superensemble instead of a cheaper bias corrected ensemble mean is not returned in skill since both forecasts obtain similar scores in the verification results. In the second application, we use the PV-gradient ensemble to produce the precipitation fields needed to drive hydrometeorological runoff simulations. Here, the results show that applying an ensemble strategy into a hydrometeorlogical forecasting chain leads to more skillful runoff simulations. Therefore, in this particular application the use of ensemble-based techniques is a successful method that improves the performance of the runoff simulations. The third application exposes the potentiality of further extending the PV perturbing approach. In this application, we modify the PV

field using the information provided by the METEOSAT-7 satellite water vapor (WV) channel for guidance. This guide comes from the existing relationship between upper-level PV features and WV brightness temperature. The results show that the perturbed run performs better than a non-perturbed run. Besides, when confronted with PV-gradient and PV-adjoint ensembles members, the tested perturbing technique remains within the range of both ensembles members scores and is statistically indistinguishable.

The results of this Thesis show the utility and suitability of forecasting methods based on perturbing the upper-level precursor trough present in cyclonic situations. The results and strategies here discussed aim to be a basis for future studies making use of these methods.

Preface

Storm Ending

by Jean Toomer, 1992

Thunder blossoms gorgeously above our heads,
Great, hollow, bell-like flowers,
Rumbling in the wind,
Stretching clappers to strike our ears...
Full-lipped flowers
Bitten by the sun
Bleeding rain
Dripping rain like golden honey-
And the sweet earth flying from the thunder.

Nature marvels us with these breathtaking phenomenons and we cannot help but stare in wonder. The human race has been fascinated by meteorological events since time in memorial. We have written poems, composed paintings, sung songs,... while wondering about their whys and hows.

Seeking answers for these questions may perfectly represent the main pillar of meteorology, the science devoted to the study of the atmospheric processes and dynamics that govern the weather. In our quest for answers, new questions arise: when and where can we witness the next major weather event? Will it be dangerous?... So the desire to predict the event is then added to the longing to diagnose it.

This PhD Thesis is a result of succumbing to these desires and longings and at the same time, an attempt to provide a meaningful contribution to the field of meteorology.

Objectives and Outline of this Thesis

The main goal of this Thesis is to improve the current prediction capability of high impact weather events that often strike the western Mediterranean. We undertake this task by developing and testing several ensemble forecasting systems. It is appropriate to use this kind of strategy since extreme and rare events, like the ones we are targeting, are often not well resolved by a deterministic prediction given their inherently unpredictable nature and their high sensitivity to model deficiencies and initial conditions errors.

This Thesis is organized as follows. Part I provides the necessary background for this work. In Chapter 1 we deepen on the reasons behind focusing on the western Mediterranean region. Chapter 2 describes general issues on numerical weather models and provides details on the models used here. The ensemble forecasting strategy as well as some examples and related issues are discussed in Chapter 3, while the necessary tools to verify the quality of the forecasts are revised in Chapter 4. An overview of the Potential Vorticity Thinking approach is provided in Chapter 5.

Throughout Part II we develop a multiphisics ensemble that combines different model physical parameterization schemes and two EPSs based on perturbing the initial and boundary conditions (Chapter 6, 7 and 8). These perturbed initial and boundary conditions ensembles are accomplished by perturbing the PV field, thus exploiting the strong connection between PV streamers and cyclogenesis observed in the troposphere. These two PV-perturbed ensembles differ in the criteria used to locate the perturbing zones.

The performance of the three ensembles is evaluated with a thorough verification process. Different quality attributes of the forecast will be explored by a wide range of verification scores and indices. This evaluation is done for a trial set consisting of 19 MEDEX cyclones collection. These events are associated with floods and strong winds over the western Mediterranean and represent the kind of phenomena we are targeting.

In Part III we apply these techniques in three different contexts to test further their properties and take advantage of their skill. The multiphysics ensemble is used to build a superensemble forecasting system in Chapter 9. The performance of the superensemble is compared with that of a simpler bias corrected ensemble

mean. Chapter 10 describes the utility of using the PV-gradient ensemble in a hydrometeorological context with the aim of accounting for the uncertainty in the initial and boundary conditions. In Chapter 11, a different method for modifying the PV field in the model initial state is developed and tested for two case studies. The results are examined and compared with both PV-gradient and PV-adjoint ensembles modification techniques.

Finally, general conclusions and several potential future studies are presented and discussed in Chapter 12.

Part I

Introduction

Chapter 1

Why focus on the western Mediterranean region?

The Mediterranean region is the region in between, in between cultures, lands, environments, etc. In fact, it literally means "sea in the middle of the earth"[1]. This middleman position makes its a place of confluence: atmosphere, sea, orography and topography together in close quarters, leading to very interesting weather phenomenons. Because, while it is true that Mediterranean climate is mainly characterized by hot dry summers and wet cool winters, it can also be notoriously capricious with sudden torrential downpours or bouts of high wind occurring at various times of the year.

The Mediterranean basin covers portions of three continents: Europe, Asia, and Africa, and is surrounded by prominent mountain ranges, like the Pyrenees (dividing Spain from France), the Alps (dividing Italy from Central Europe) and the Balkan mountains. A shallow submarine ridge between the island of Sicily, Italy, and the coast of Tunisia divides the sea in two main subregions, the western Mediterranean and the eastern Mediterranean. The western Mediterranean, our target region, covers an area of about 0.85 million km^2 (see Fig. 1.1).

The western Mediterranean is often hit by heavy rain events that have a high socio-economical impact (Llasat and Sempere-Torres, 2001; Llasat et al., 2010). For example, on 10 June 2000 a storm flooding caused five victims, more than 500 evacuated and material losses, including the destruction of a highway bridge,

[1]See `http://www.etymonline.com/index.php?term=Mediterranean`

Figure 1.1. *The western Mediterranean region and its orography. (Image source: Google Maps.)*

were estimated to exceed 65 million euros by the local media on Catalonia, Spain (Llasat et al., 2003; Martín et al., 2007) and on 10 November 2001 a storm flooding caused 886 victims on Algiers (Tripoli et al., 2005) and 4 casualties in the island of Mallorca, Spain, accompanied by more than 200.000 uprooted trees and records of wind gusts exceeding 140 km/h and heavy precipitation up to 400 mm in two days (Romero et al., 2002). In fact, several studies have established a connection between some of these heavy rain events and the presence of a cyclone over the region (e.g. Jansà et al. 2001). This area, the Western Mediterranean, is a very cyclogenetic area as a consequence of the complex topography surrounding a quasi-closed sea (Reiter, 1975; Meteorologial Office, 1962).

Some examples of these cyclones associated to heavy rain events are:

- shallow weak disturbances with a warm core over land masses of thermal origin (Romero et al., 2001),
- shallow weak lows with a warm core over the sea at the lee of important mountain ranges, linked to the orographic effect on the atmospheric flow (Romero et al., 2000), and
- baroclinic systems with great vertical amplitude (Homar et al., 2002), developed along frontal zones under the intrusion in the Mediterranean region of an upper-level through.

The capability to predict such hazardous events is still limited. However many efforts are being devoted to successfully increase it in hopes of preventing and reducing the damages they cause. Some of these efforts are endorsed on international programs like MEDEX[2] and HyMeX[3] and are in complete agreement with the final objective of this Thesis: to improve the capability of predicting accurately and timely the western Mediterranean high-impact weather events by means of several ensemble prediction systems developed for this purpose.

[2]MEDEX is the Mediterranean Experiment on cyclones that produce high impact weather in the Mediterranean, a project endorsed by the WMO under the THORPEX WWRP (`http://medex.aemet.uib.es`).

[3]HyMeX is the Hydrological cycle in the Mediterranean Experiment, a project also endorsed by the WMO under the GEWEX WCRP and THORPEX WWRP (`http://www.hymex.org/`).

Chapter 2

Numerical weather prediction models

A numerical weather prediction (NWP) model is a mathematical model of the atmosphere that uses the fluid dynamics and thermodynamics equations to estimate the state of the atmosphere at a future time once the state of the fluid has been sampled. The first attempt to use a numerical weather prediction model was done in 1922 by Lewis Fry Richardson that worked by hand on a local example, after Vilhelm Bjerknes had laid out the fundamental principles of the computation of atmospheric evolution in 1904 (Shuman, 1989). However the task was too big to solve by hand, so it was thirty years later with the invention of computers when numerical weather prediction was able to take off. The first successful NWP was done in 1950 by a working group leaded by Jule Charney and assisted by the mathematician John von Neuman using a barotropic model. Since then, numerical weather prediction models have greatly improved hand in hand with computer capacity. Nowadays, several NWP models are run operationally worldwide to forecast the weather and are also used in many research centers as a tool to study the atmosphere and improve the prediction, among other tasks. A thorough and recent review of the origins of computer weather prediction can be found in Lynch (2008).

The NWP models predict the future state of the atmosphere solving the *basic equations* on a three-dimensional grid representing the atmosphere after gathering the current state of the atmosphere into the defined grid. The basic equations, used to approximate global atmospheric flow and also used in most atmospheric

models, consist of the following set of nonlinear differential equations:

- Conservation of momentum (the Navier-Stokes equations)
- Thermodynamics laws
- Mass continuity equation
- Equation of state (Ideal gas law)
- Water species predictive equations

These nonlinear partial differential equations cannot be solved analytically, instead they must be solved using numerical techniques (an approximation) requiring a huge computational power. The most common used *numerical methods* are the finite difference methods for all three spatial dimensions, spectral methods for the horizontal dimensions and finite difference methods in the vertical. It is worth to note that the physical process that can not be explicitly solved owning to their small scale, their complexity, the lack of understanding or their high computationally cost (e.g. turbulence, convection, diffusion, etc) have to be parameterized.

Certainly the finer the *model grid resolution* the more accurately the prediction will be. However, finer resolution requires more computational power and better *physical parameterizations* schemes that can not always be achieved, either due to excessive computational cost and/or inability to correctly parameterize the complex sub-grid physical processes. To overcome the computational limitation, two kind of models coexist: *global models* that cover the whole Earth with a coarser resolution and regional models that cover local areas with a finer resolution. If *regional models* do not get information from outside their domain, they can only predict for short times since as time passes they are not able to account for weather phenomena occurring outside their domain but affecting the weather inside. Therefore, limited-area models need to be nested inside a larger coarser domain run by a global model, in other words, needs to be provided with boundary conditions or lateral forcing. The effects of this *nesting* is widely discussed in Section 3.2 on the EPS framework.

As already stated, a NWP model needs to know the current state of the atmosphere to initialize the predictive equations. The *initial conditions* needed to run the models are usually the weather analysis: the best available representation of the current

atmospheric state. The analysis are obtained processing irregularly spaced observations by data assimilation and objective analysis methods that perform quality control into a evenly spaced grid usable by the model (See Section 3.1.3 for an overview on data assimilation techniques). The analysis are worldwide generated by operational centers like the National Centers for Environmental Prediction, the European Centre for Medium-Range Weather Forecasts, Japan Meteorological Agency, Météo-France or UK Met Office. Several operational centers have done a re-analysis to include all the now available observations, measures and advanced numerical techniques, for example the ECMWF has available the ERA-40 reanalysis that cover the period between 1957 and 2002 and has just produced the ERA-Interim, a more refined reanalysis covering from 1989 to the present time (ERA-40: Uppala et al. 2005 and ERA-Interim: Berrisford et al. 2011; Dee et al. 2011).

The *observations* used in atmospheric models come mainly from state weather services all over the globe. The World Meteorological Organization[1] has issued many Manuals and Guides providing a standardization for observational measurements (e.g. WMO 2008) and a suitable framework worldwide for meteorological observations, especially the Global Observing System (WMO, 1989, 2003), aeronautical meteorology (WMO, 1990), hydrology (WMO, 1994), agricultural meteorology (WMO, 1981) and climatology (WMO, 1983). The main observational inputs are surface observations from automated weather stations at ground level over land and from weather buoys at sea. Other observations come from radiosondes launches that measure through the troposphere until the stratosphere, aircraft and ships reports and weather satellites. Additionally, research projects launch special campaigns to increase the number of observations over a region of interest like the forthcoming HyMeX 2012 Special Observation Period (SOP) campaign that plans to improve the observational network over the northwestern part of the Mediterranean to study key processes of the water cycle, especially heavy precipitation systems, intense air-sea fluxes and dense water formation (see `http://www.hymex.org` for details).

Another key component in a NWP model is the *terrain* representation. In order to produce a plausible forecast the model needs information about the orography, topography, land uses, etc to better describe phenomena like downslope winds, mountain waves, thermally-driven circulations and so on. The land use data allows

[1]WMO: `http://www.wmo.int/`

to describe the interaction between weather and ecosystems by proving information such as the kind of vegetation or the presence of desert. This information is crucial to determine the land properties like the albedo or the moisture required in the model radiation scheme, for example.

Several efforts have also been directed to couple atmospheric models with oceanographic ones to remove the need of artificially adjust fluxes across the ocean surface interface. These *coupled atmosphere-ocean models* are basically used on future climate predictions, like the HadCM3[2] coupled atmosphere-ocean general circulation model, used in the IPCC[3] Third Assessment Report in 2001.

2.1 MM5 non-hydrostatic mesoscale model

The non-hydrostatic mesoscale MM5 model is a high resolution short-range weather forecast model developed by the National Center for Atmospheric Research (NCAR) and the Pennsylvania State University (PSU; Dudhia 1993; Grell et al. 1995). The main characteristics of the MM5 are described below with the aid of the MM5 Users' Guide (Dudhia et al., 2005).

2.1.1 The MM5 model horizontal and vertical grid

The MM5 acquires and analyzes the data on pressure surfaces. This information is previously interpolated to the model's vertical coordinate. The model vertical coordinate, σ, follows the terrain at the lower grid levels and flattens at the upper surface (see Fig. 2.1). This σ coordinate is defined by

$$\sigma = \frac{p - p_t}{p_s - p_t} \ , \tag{2.1}$$

where p_t is upper boundary of pressure and p_s is the surface pressure. The values of σ go from 0 at the atmosphere's top to 1 at the earth's surface.

The horizontal grid has an Arakawa-Lamb B-staggering of the velocity variables with respect to the scalars. This is shown in Fig. 2.2 where it can be seen that

[2]Hadley Centre Coupled Model, version 3, developed at the Hadley Centre in the United Kingdom (e.g. Collins et al. 2001).

[3]Intergovernmental Panel on Climate Change; `http://www.ipcc.ch/`

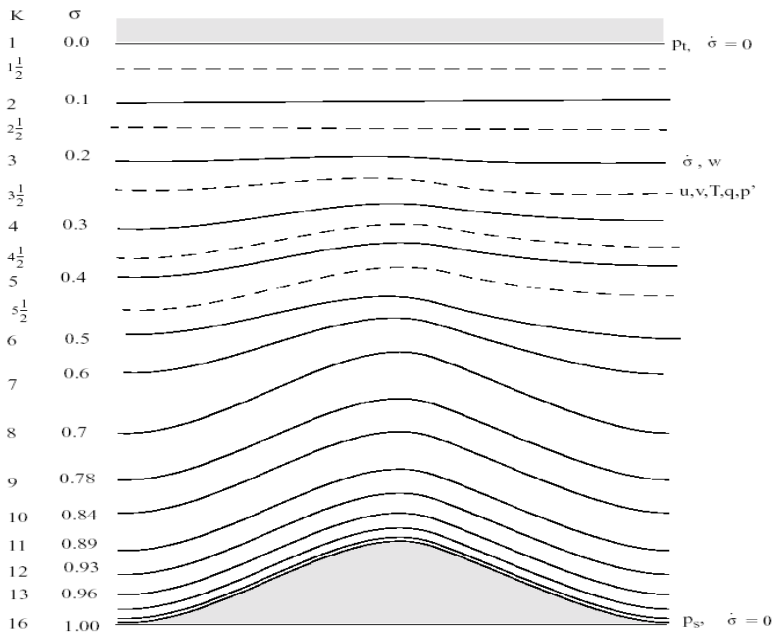


Figure 2.1. *Schematic representation of the vertical structure of the model. This example is for 15 vertical layers. Dashed lines denote half-sigmalevels, solid lines denote full-sigma levels.*

the scalars, like temperature, are defined at the center of the grid square, while the eastward and northward velocity components are collocated at the corners. All these variables are defined in the middle of each model vertical layer, referred to as half-levels, and vertical velocity is carried at the full levels (see Fig. 2.1).

2.1.2 The nesting capability

The MM5 model contains a capability of multiple nesting with up to nine domains running at the same time and completely interacting. A possible configuration is shown in Fig. 2.3. The nesting ratio is always 3:1 for two-way interaction, and is not restricted for one-way nesting. The one-way nesting differs from two-way nesting in having no feedback and coarser temporal resolution at the boundaries. Each sub-domain has a mother domain in which it is completely embedded. Moving a domain and turn on and off a nest at any time in the simulation is also possible. There are three ways of doing the two-way nesting. First, nest interpolation, where the nest is initialized by interpolating coarse-mesh fields and requires no additional input files. Second, nest analyses input, which requires a model input file to be prepared for the nest in addition to the coarse mesh. Third, nest terrain input

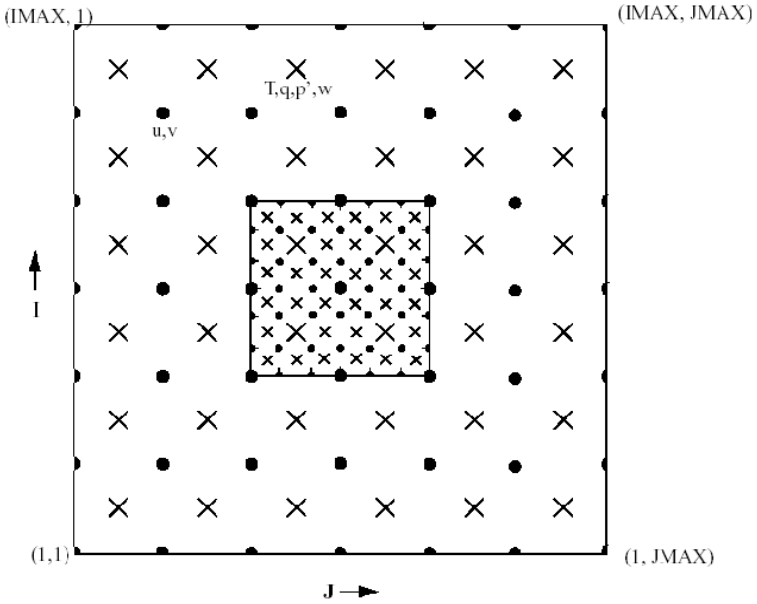


Figure 2.2. *Schematic representation of the horizontal Arakawa-Lamb B-staggering staggering of the dot and cross grid points. The smaller inner box is representative of a mesh staggering for a 3:1 coarse-grid distance fine-grid distance ratio.*

that requires just land use input file, so the meteorological fields are interpolated from the coarse mesh and vertically adjusted to the new terrain.

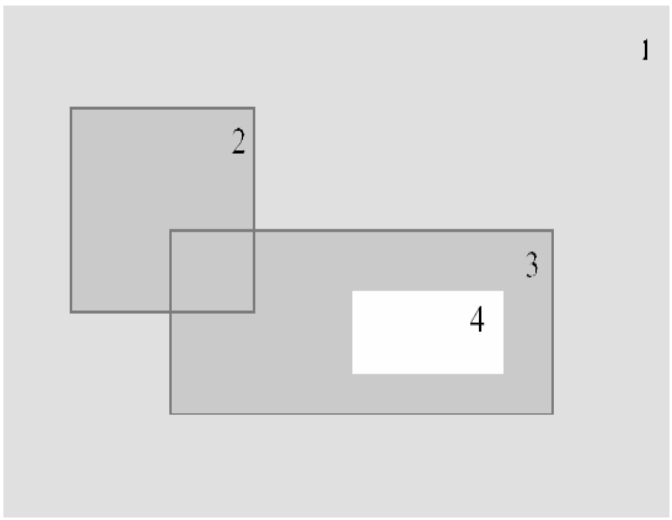


Figure 2.3. *Example of nesting configuration. The shading shows three different levels of nesting.*

2.1.3 The lateral boundary conditions

A regional numerical weather prediction model requires lateral boundary conditions. In MM5 all four boundaries have specified horizontal winds, temperature, pressure and moisture fields, and can have specified microphysical fields, such as cloud and precipitation species, if these are available. These boundary values have to be set in the simulations in addition to initial values for these fields.

The boundary values can come from analyses at the future times, from a previous coarser-mesh simulation, and from another model's forecast. For real-time forecasts the lateral boundaries will ultimately depend on a global-model forecast. In studies of past cases the analyses providing the boundary conditions may be enhanced by observation analysis in the same way as initial conditions are. The MM5 uses these discrete-time analyses by linearly interpolating them in time to the model time. In two-way nest, the boundaries are updated every coarse-mesh timestep.

2.1.4 The nonhydrostatic dynamics

In the mesoscale models, the hydrostatic approximation can be applied when the typical horizontal grid sizes are comparable with or greater than the vertical depth of features of interest. Then, the pressure is completely determined by the overlying air's mass, like the hydrostatic relation shows,

$$dP = -\rho_0 g \; dz \; . \tag{2.2}$$

However, when the scale of resolved features in the model have aspect ratios nearer unity, or when the horizontal scale becomes shorter than the vertical scale, nonhydrostatic dynamics can not be neglected. The nonhydrostatic dynamic introduces an additional term, the vertical acceleration that contributes to the vertical pressure gradient, so the hydrostatic balance is no longer exact.

The reference state in the nonhydrostatic model

The reference state is an idealized temperature profile in hydrostatic equilibrium, given by

$$T_0 = T_{s0} + A \; ln \left(\frac{p_0}{p_{00}} \right) \; , \tag{2.3}$$

where T_0 is specified by: sea-level pressure, p_{00}, the reference temperature at p_{00}, T_{s0}, and a measure of lapse rate, A, representing the temperature difference between p_{00} and p_{00}/e. Usually, just T_{s0} needs to be selected based on a typical sounding in the domain.

The surface reference pressure, therefore, depends entirely upon the terrain height, and can be derived from Eq. (2.3), using the hydrostatic relation (2.2),

$$Z = -\frac{RA}{2g}\left(ln\frac{p_0}{p_{00}}\right)^2 - \frac{RT_{s0}}{g}\left(ln\frac{p_0}{p_{00}}\right) \tag{2.4}$$

and this quadratic can be solved for p_0(surface) given the terrain elevation, Z. Once this is done, the heights of the model σ levels are found from

$$p_0 = p_{s0}\sigma + p_{\text{top}} \ , \tag{2.5}$$

where

$$p_{s0} = p_0(\text{surface}) - p_{\text{top}} \ , \tag{2.6}$$

and then Eq. (2.4) is used to find Z from p_0. It can be seen that since the reference state is independent of time, the height of a given grid point is constant.

2.1.5 The land use categories

The MM5 provides three sets of land use categorizations that are assigned along with elevation. These have various categories like: type of vegetation, desert, urban, water, ice, and others. Each grid cell of the model is assigned one of the categories, and this determines surface properties such as albedo, roughness length, longwave emissivity, heat capacity and moisture availability. Additionally, if a snow cover dataset is available, the surface properties may be modified accordingly. These values are also variable according to summer or winter season, for the northern hemisphere. It is important to note that the values are climatological and may not be optimal for a particular case, especially moisture availability.

2.1.6 The map projections and map-scale factors

The modeling system has a choice of several map projections. Lambert Conformal is suitable for mid-latitudes, Polar Stereographic for high latitudes and Mercator for low latitudes. These transformations are accounted for in the model pre-processors that provide data on the model grid, and post-processors.

The map scale factor, m, is defined by

$$m = \frac{\text{distance on grid}}{\text{actual distance on earth}} \ , \tag{2.7}$$

and its value is usually close to one varying with latitude. The projections in the model preserve the shape of small areas, but the grid length varies across the domain to allow a representation of a spherical surface on a plane surface. Map scale factors need to be accounted for in the model equations wherever horizontal gradients are used.

2.1.7 The basic equations of the MM5 model

As expected, the MM5 basic equations are nonhydrostatic and are given in terms of terrain following coordinates (x,y,σ). These equations without moisture terms are given by

- the equation of state:

$$p = \rho R T \ , \tag{2.8}$$

  where p is the pressure, ρ is the air density, R is the gas constant (which depends on the composition of the air), and T is absolute temperature.

- the pressure equation:

$$\frac{\partial p'}{\partial t} - \rho_0 g w + \gamma p \nabla \cdot \overrightarrow{V} = -\overrightarrow{V} \cdot \nabla p' + \frac{\gamma p}{T}\left(\frac{\dot{Q}}{c_p} + \frac{T_0}{\theta_0} D_\theta\right) \ , \tag{2.9}$$

  where p' is the nonhydrostatic perturbed pressure, p is the hydrostatic pressure p_0 plus p', ρ_0 is the air density, g is the gravity constant, w is the vertical velocity, $\gamma = \frac{C_p}{C_v}$, where C_p is the air calorific heat at constant pressure and C_v is at constant volume, $\overrightarrow{V}$ is the velocity vector, $\dot{Q}$ is the heat exchange

with the environment, T_0 is the temperature of the buoyancy term, θ_0 is the reference potential temperature and D_θ is the heat loss owning to friction and turbulence. This equations shows that pressure temporal variations are due to the rising and subsidence fluid motions, variations produced by converge and divergence, pressure advection and variations provided by heat exchanges.

– the momentum equations:

a) component x:

$$\frac{\partial u}{\partial t} + \frac{m}{\rho}\left(\frac{\partial p'}{\partial x} - \frac{\sigma}{p^*}\frac{\partial p^*}{\partial x}\frac{\partial p'}{\partial \sigma}\right) = -\overrightarrow{V}\cdot\nabla u + v\left(f + u\frac{\partial m}{\partial y} - v\frac{\partial m}{\partial x}\right) - ew\cos\alpha - \frac{uw}{r_{earth}} + D_u \ , \quad (2.10)$$

where m is the map scale factor, $p^* = p_{surface} - p_{top}$ is the difference between pressures, f and $e = 2\Omega\cos\lambda$ are the Coriolis terms where λ is the latitude, $\alpha = \phi - \phi_c$, ϕ is the longitude and ϕ_c is the central longitude, $u\frac{\partial m}{\partial y}$, $v\frac{\partial m}{\partial x}$ and r_{earth} are the curvature effect terms, and D_u is the heat loss term due to friction and turbulence in the component x direction. The component x momentum temporal variations are due to spacial variations in the pressure field, u velocity advection and curvature and Coriolis effects.

b) component y:

$$\frac{\partial v}{\partial t} + \frac{m}{\rho}\left(\frac{\partial p'}{\partial y} - \frac{\sigma}{p^*}\frac{\partial p^*}{\partial y}\frac{\partial p'}{\partial \sigma}\right) = -\overrightarrow{V}\cdot\nabla v + u\left(f + u\frac{\partial m}{\partial y} - v\frac{\partial m}{\partial x}\right) + ew\sin\alpha - \frac{uw}{r_{earth}} + D_v \ . \quad (2.11)$$

c) component z:

$$\frac{\partial w}{\partial t} - \frac{\rho_0}{\rho}\frac{g}{p^*}\frac{\partial p'}{\partial \sigma} + \frac{g}{\gamma}\frac{p'}{p} = -\overrightarrow{V}\cdot\nabla w + g\frac{p_0}{p}\frac{T'}{T_0} - \frac{gR_d}{c_p}\frac{p'}{p} + e(u\cos\alpha - v\sin\alpha) + \frac{u^2+v^2}{r_{earth}} + D_w \ , \quad (2.12)$$

where R_d is the dry air universal constant.

– the thermodynamic equation:

$$\frac{\partial T}{\partial t} = -\overrightarrow{V} \cdot \nabla T + \frac{1}{\rho C_p}\left(\frac{\partial p'}{\partial t} + \overrightarrow{V} \cdot \nabla p' - \rho_0 g w\right) + \frac{\dot{Q}}{C_p} + \frac{T_0}{\theta_0} D_\theta \ . \quad (2.13)$$

The temperature temporal variations are due to thermal advection, density variations, heat exchanges and heat loss owning to friction and turbulence.

2.2 Adjoint model

An adjoint model is a powerful tool to estimate the sensitivity of a given forecast feature, limited to the linear framework. In other words, the adjoint models allow to evaluate the effect of any perturbation to one particular response function[4]. Formally, an adjoint model is defined as the transposition of a linear operator that is constructed tangent to the phase space trajectory that is followed by the forward nonlinear deterministic forecast. This tangent linear approximation can be affected by the timespan of the adjoint run. The longer the evolution analyzed, the farther away from a linear evolution the perturbations evolve in the nonlinear model. The linear assumption is valid for smooth integrated response functions defined at lead times up to 48 h (e.g. Rabier et al. 1992; Vukićević 1991), while decreasing to 24 h when diabatic processes significantly affect the response function (Homar and Stensrud, 2004). Also, the response function definition is critical since a response function highly influenced by nonlinear forecast features, e.g. rain, may severely constrain the tangent linear approximation.

2.2.1 Applications

Sensitivity analysis estimates the impact of the uncertainty contribution to different initial conditions aspect on the uncertainty of a given forecast aspect. Therefore a sensitivity analysis is very suitable to evaluate and investigate the existent relationship between different atmospherics features and their predictability given

[4]The term response function is commonly used to denote a forecast feature of interest.

certain conditions. The simplest form of sensitivity analysis is to create a given set of scenarios where the sensitivity analysis can determine how changes in one forecast feature will impact a response function (e.g. Romero 2001; Martín et al. 2007). The problem with this kind of sensitivity analysis is that it is possible that another perturbation with similar characteristics and properties of the first one may produce a very different impact, leading to a different interpretation of the sensitivity. The adjoint has the advantage of being able to evaluate the effect of any perturbation to one particular response function instead of being limited to evaluate the effect of one perturbation to any number of response functions (Homar and Stensrud, 2004). Therefore, the use of an adjoint model is a more efficient and direct way to tackle sensitivity analysis (see, e.g., Homar and Stensrud 2004; Errico 1997; Errico and Vukićević 1992; Rabier et al. 1992 and Hall et al. 1982).

Although sensitivity analysis is an adjoint model main application, there are other applications that can be derived from applying sensitivity analysis (see Errico (1997) for a more extensive review of adjoint models and their applications). One example are optimal analyses like the one used by data assimilation to adjust the model initial fields to incorporate observations. The optimal solution can be efficiently determined using the sensitivity analysis provided by the adjoint model (e.g. Lewis and Derber 1985). The adjoint approach has the advantage of solving the optimization analysis in a reasonable time, enough to be applied in real-time forecasting. An analogous approach can be used to efficiently determine model parameters, like a model physic parameterization scheme, where instead of optimizing the initial conditions to obtain an accurate forecast, the model parameters are the ones being optimized.

Another useful application of adjoints is *targeting* (e.g. Langland et al. 1999). This technique aims to determine where it is most important to produce an accurate initial condition for a forecast using the sensitivity fields derived from a sensitivity analysis. The method takes advantage of the fact that any significant error in the initial conditions where the sensitivity is large, will have a significant impact on the forecast response function and therefore on the accuracy of its prediction and viceversa; similar errors where the sensitivity is low will have little effect on the forecast response function. These locations will depend on the accuracy of the nonlinear model used to generate the forecast and on the accuracy of its adjoint, used to determine the sensitivity (see, e.g., Palmer et al. 1998). Maintaining initial errors at minimum in the sensitive regions is therefore crucial. This can be

achieved by, for example, increasing the number of observations in those regions, or improving the initial conditions on specific synoptic situations by modifying them to obtain another plausible initial condition better representing the true state of the atmosphere. This later application is exploited on Part II by means of an ensemble prediction system that uses a sensitive analysis as guide to build sets of initial conditions perturbations. This approach has already been successfully applied on several studies, for example, Molteni et al. (1996); Xu et al. (2001) and Vich et al. (2011b).

2.2.2 MM5 adjoint model

The MM5 adjoint mesoscale model was developed at the National Center for Atmospheric Research and is based on the MM5 model, described in Section 2.1. The current version of the MM5 adjoint modeling system includes model dynamics, diffusion, bulk planetary-boundary-layer processes, surface friction, a semi-implicit time-split integration scheme, dry convective adjustment, a cumulus parameterization scheme and a resolvable-scale precipitation process. Moreover, since the MM5 is a regional model, the adiabatic MM5 adjoint has been developed with flexibility of controlling initial conditions and/or lateral boundary conditions. An extensive description of the mathematical and numerical formulation of the MM5 adjoint model can be found in Zou et al. (1997) and a guide of how to run the programs that made up the model is gathered on its User's manual (Zou et al., 1998).

Chapter 3

Ensemble prediction systems

Numerical weather prediction models have greatly improved since their origins (Lynch, 2008). Likewise the quality, quantity and availability of the observations required to estimate the model initial conditions have also been improved (WMO, 2008). However, neither numerical weather prediction models nor observations are perfect. Numerical weather predictions need to deal with the uncertainties present on the observations (Lorenz, 1963) and the inaccuracies of the models (Frank, 1983). As a response to the constraints imposed by these uncertainties the concept of probabilistic forecast arises, since the uncertainties associated with a forecast can be better conveyed through the use of probability levels.

An ensemble prediction systems (EPS) produces probabilistic forecast based on a set of deterministic forecasts valid at the same time. The underlying uncertainties present in both initial conditions and model formulations allow to build a set of plausible states that captures the true state. In fact, the ensemble average or mean is generally more accurate than a single deterministic forecast (Leith, 1974; Zhang and Krishnamurti, 1997; Du et al., 1997; Buizza and Palmer, 1998). Fig. 3.1 illustrates a conceptual ensemble consisting of a control forecast initialized with an analysis and two additional forecasts initialized with two perturbations of the analysis. The first ensemble (Fig. 3.1.a) is able to track the true state while the second one (Fig. 3.1.b) does not capture the true state inside the range predicted by the ensemble members, most probably due to model defects and/or to implausible initial perturbations. However, capturing the true state is not enough, an ensemble forecasting system also needs to be reliable and to have resolution (Murphy, 1973).

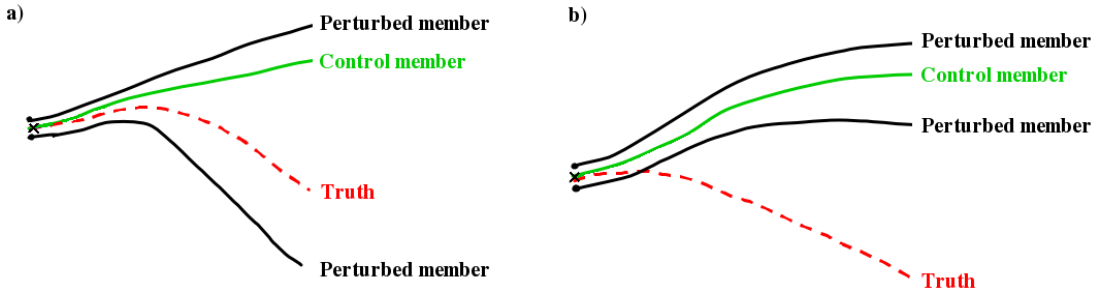


Figure 3.1. *Illustration of a conceptual EPS consisting of three members, a control and two perturbed states. The control forecast (green line) is initialized with the analysis (cross symbol), and two initial perturbations of the analysis (dot symbol) are used to initialize the two perturbed forecasts (black line). The truth or verifying analysis is represented by the red dashed line. a) Represents an ensemble in which the truth has been captured by the ensemble members. In contrast, b) represents an ensemble in which the ensemble members do not manage to capture the true state. (Adapted from Palmer and Hagedorn 2006.)*

The forecast probability of an event is estimated by the fraction of the forecasts predicting the event among all forecasts considered (number of ensemble members), assuming that each forecast is an independent realization. An estimate of the EPS uncertainty is given by the spread of the forecasts, usually defined as the standard deviation about its mean. Moreover, the ensemble can be interpreted as an estimate of the evolution of the probability density function (pdf) of the selected states over time. It is worth to note that ensemble forecasting only estimates the forecast probabilities through sampling the pdf initial state and evolving it over time. The analytical evolution of the pdf over time involves, in addition to knowing the initial pdf, computing the Liouville equations (e.g. Ehrendorfer 1994). Currently, while theoretically it should be doable, in a realistic context it becomes unattainable due to the extremely high dimensionality of phase space that must be considered and its consequent computational cost.

In terms of probability density function, the initial pdf is sampled by the ensemble members initial conditions reflecting the likely uncertainty in the analysis. This narrow pdf is widened over time as the forecast lead-time increases and the growth of initially small perturbations and model deficiencies increments the uncertainty. An skillful EPS evolves the pdf capturing the true state through time and provides a narrower pdf than the climatological distribution, in other words, the ensemble pdf resolves better whether or not an event is likely to occur than climatology (see Fig. 3.2). When evolving over time, one needs to be careful to not exceed the predictability threshold, the forecast uncertainties/errors will then grow non-linearly until all predictability is lost and the forecast pdf evolves into the pdf

invariant distribution of the attractor. This point is illustrated in Fig. 3.3 by means of the Lorenz attractor schematics (Lorenz, 1963) presented on Palmer (1999).

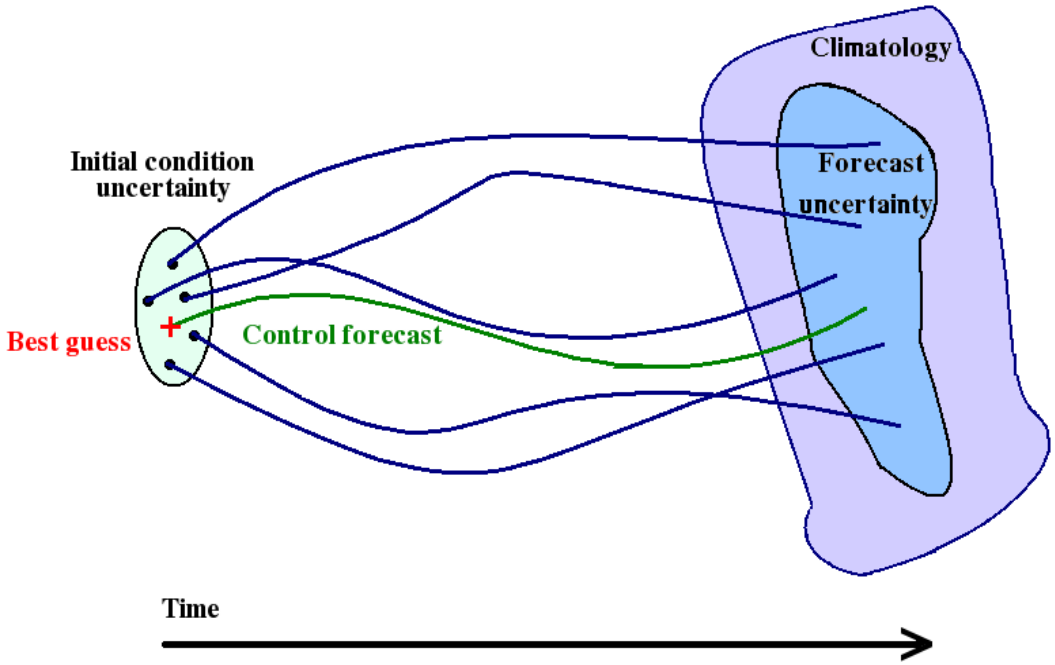


Figure 3.2. *Schematic of how the forecast uncertainty is sampled by an EPS, assuming that the model is perfect. Even with a perfect model, the forecast can not be perfect since the true initial state remains unknown due the initial conditions uncertainties. If we use our best guess (red cross) as substitute for the true state we generate a forecast that can be inaccurate (commonly known as the* control *member of the EPS; green line). If we sample the uncertainty in the initial conditions and run several ensemble members with the model (blue lines), we obtain an estimation of the forecast uncertainty. If the EPS is well designed, the forecast pdf will be contained within the climatological pdf but will have more resolution.*

3.1 Ensemble forecasting building techniques

The generation of an ensemble forecasting system is achieved by exploiting the uncertainties present on the initial state, the model deficiencies or both uncertainty sources simultaneous. In fact, several weather services have implemented operational ensemble prediction systems by adding perturbations to the analyses to define the initial conditions, thus assuming that forecast errors arise from analysis errors while the model is considered perfect. The National Centers for Environmental Prediction (NCEP) define these perturbations through bred modes (Toth and Kalnay, 1993, 1997) which represent the fastest growing perturbations and

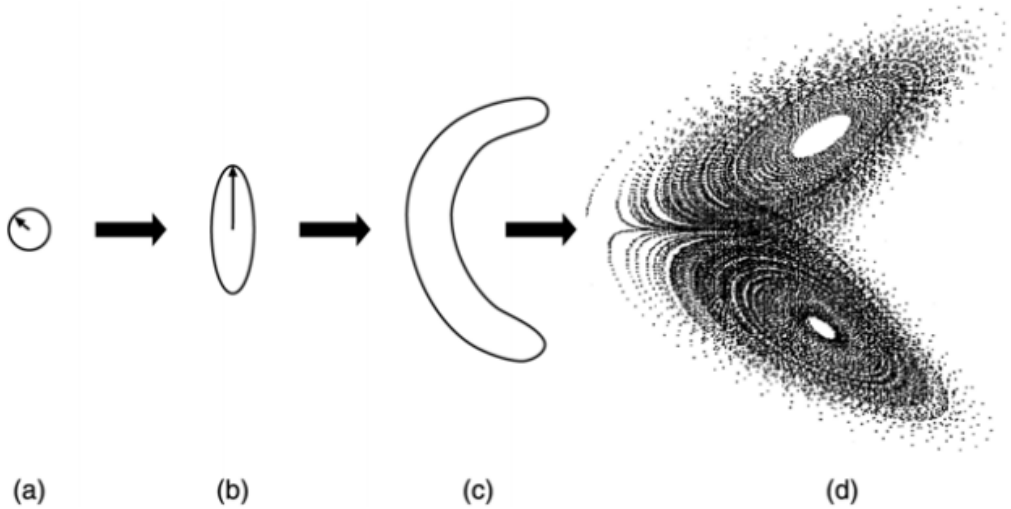


Figure 3.3. *Schematic evolution of an isopleth of the probability density function (pdf) of initial and forecast error in N-dimensional phase space. (a) At initial time, (b) during the linearised stage of evolution. A (singular) vector pointing along the major axis of the pdf ellipsoid is shown in (b), and its pre-image at initial time is shown in (a). (c) The evolution of the isopleth during the nonlinear phase is shown in (c); there is still predictability, though the pdf is no longer Gaussian. (d) Total loss of predictability, occurring when the forecast pdf is indistinguishable from the attractor's invariant pdf. (From Palmer 1999.)*

simulate the development of growing errors in the analysis cycle. The European Center for Medium-Range Weather Forecasts (ECMWF) used singular vectors (Buizza and Palmer, 1995; Molteni et al., 1996) which maximize the linear growth of energy over a specified domain until June 2010, when it was replaced by an Ensemble Data Assimilation (EDA) system that perturbs observations, sea-surface temperature fields and model physics (Fisher, 2003; Tan et al., 2007). The Canadian Meteorological Center (CMC) applies the ensemble Kalman filter (EnKF; Evensen 2003; Houtekamer et al. 2005) which provides an ensemble of initial conditions thanks to the assimilation of an ensemble of perturbed observations. Other EPSs building techniques, that do include model errors, involve the use of different forecast models (e.g. Evans et al. 2000; Krishnamurti et al. 2000a; Stensrud 2001), different physical parameterization schemes (e.g. Houtekamer et al. 1996; Stensrud et al. 1999) or stochastic physical parameterization (e.g. Palmer 2001; Grell and Devenyi 2002).

An overview of the cited generating techniques is provided below.

3.1.1 Singular vector

The singular vector EPS generation techniques is just one of many useful applications in signal processing and statistics of the singular value decomposition[1]. In meteorology, this technique is used to estimate the most linearly quickly growing perturbations over a given initial forward time period. Since these perturbations are consistent with uncertainties in the observed properties of the atmosphere, they are suitable to build an ensemble forecasting system. Despite this, it is worth to remember that no-method is perfect, Isaksen et al. (2005) showed that perturbations constructed by the singular vector method could give rise to structures rarely observed in the atmosphere.

The ECMWF EPS singular vectors (SV) generation method assumes that the perturbations grow linearly in time, a generally valid assumption in the atmosphere for small perturbations for one or two days lead-time. In the ECMWF EPS the singular vectors are calculated over a 48-hour period (Buizza and Palmer, 1995), a compromise between the constraints of the linearity approximation and the requirement to provide an operational EPS for medium-range forecasting, using a tangent linear forward and adjoint model during this 48-hour period. It is worth mentioning, that the ECMWF computes the singular vector decomposition solving the eigenvalue problem with a Lanczos code[2] (Golub and Van Loan, 1983) due to the very large dimension of the system (Buizza, 2001).

On the other hand, the size of the perturbation at both initial and final time of the forwarded period needs to be measured. Ideally, the norm should be related to the spatial distribution of expected errors in the analysis (at the initial time) and to the forecast errors of interest (at the final time). In practice, this measure is done by a norm provided by the use of total energy for both initial and final time since it fits the general requirements of the ECMWF EPS (Buizza and Palmer, 1995).

Moreover, a targeting region can be implement in the singular vector computations. If there is a particular region of interest the singular vector can be calculated to have the greatest impact in the target region even if they originated outside this

[1] a factorization of a real or complex matrix in linear algebra

[2] The Lanczos algorithm is specifically suitable for decomposing very large matrices. This iterative algorithm finds the singular value decomposition of a rectangular matrix with an adaptation power method.

region. The ECMWF EPS singular vector system aims to produce perturbations localized over the extratropics, in fact, a separate calculation is made for each hemisphere. See Buizza et al. (2000) for a comprehensive review of ECMW EPS singular vector.

3.1.2 Bred vector

The Bred vectors (BVs) are designed to build a set of initial states in agreement with atmospheric realistic structures. While the singular vector method optimizes the growth over time of the perturbations, the Breeding technique generates the perturbations integrating a dynamical model forward in time. Magnusson et al. (2008) states that BVs present a larger perturbation growth than SVs for medium and large time-scales and viceversa for short time-scales, when an orthogonal set of initial perturbations for the mode are used.

The BV are calculated on a breeding cycle (Toth and Kalnay, 1993, 1997) that consists on:

1. Adding a random perturbation to the analysis.
2. Integrating the initial conditions from the breeding run (perturbed) and the analysis (not perturbed) for 12 hours.
3. Normalizing the differences between these two nonlinear runs.
4. Adding the difference to the new analysis, valid at 12 h later than the previously used analysis.
5. Repeating steps 2 to 4.

The Bred vectors are defined as the normalized differences between breeding and the 12 hour forecast runs. After a transitory period of the order of the dominant instability time scale (3-5 days for baroclinc instabilities, Toth and Kalnay 1993), the breeding cycle converges in a statistical sense[3]. Additional BV can be generated using different initial perturbations to start the breeding cycle. Therefore, the

[3]On some strongly nonlinear models the bred vectors remain distinct instead of converging to a single leading bred vector, presumably due to the nonlinear terms and physical parameterizations introducing sufficient stochastic forcing to avoid such convergence (Kalnay et al., 2002).

Lyapunov vectors (LVs) of a dynamical system are closely related to Bred vectors (Kalnay et al., 2002). In fact, the bred vectors would be identical to the leading LVs if an infinite breeding time and infinitesimal amplitudes were used, even though in a practical context, bred vectors are local in space and time and not globally orthogonalized, and they are not infinitesimal vectors.

An alternative method is *self-breeding* that uses pairs of ensemble forecasts to generate the perturbation at the next time (Toth and Kalnay, 1997). In this procedure, the difference is scaled down as before and then added and subtracted to the valid analysis. This technique, illustrated in Fig. 3.4, maintains the linearity of the perturbation to second order instead of the first order of linearity obtained with the one-side generation. It is also worth to note that the sef-breeding technique is cost-free in an ensemble forecasting context since the pair of ensemble forecasts is already built.

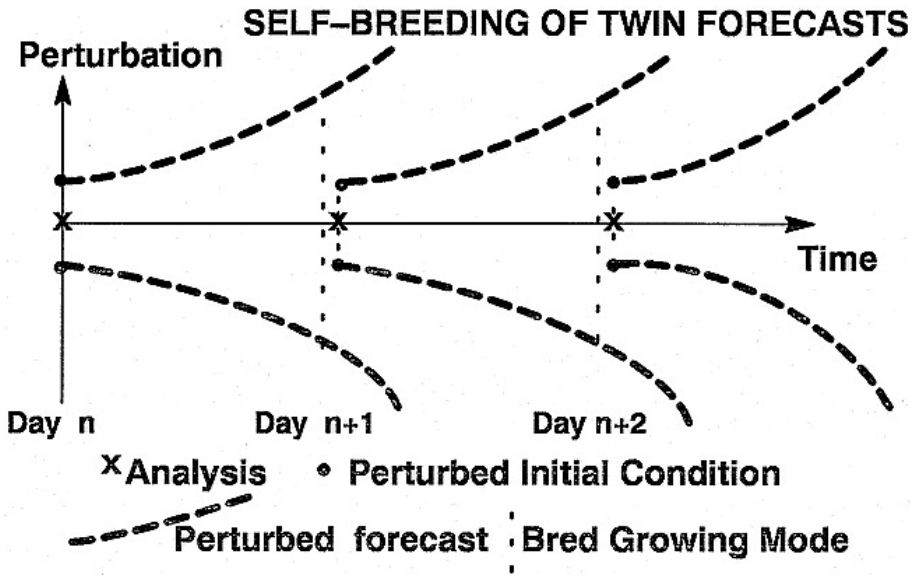


Figure 3.4. *Schematic of a self-contained breeding pair of ensemble forecasts. (From Toth and Kalnay 1997.)*

3.1.3 Ensemble Data Assimilation

Data assimilation (DA) is a common analysis practice in weather forecasting since Gandin (1963) introduced the concept of optimal interpolation[4]. The DA approach combines the model state and the observational data at each time step in order to provide a detailed analysis of the current state of the system. These combinations help to overcome limitations in both the model and the data, and to better understand the physical processes underlying the observational data leading to a more accurate initial state of a forecast run.

The three-dimensional variational data assimilation (3D-Var, i.e., Parrish and Derber 1992; Courtier et al. 1998; Rabier et al. 1998; Andersson et al. 1998) was the logical next step after Lorenc (1981) generalized the optimal interpolation (OI) to three spatial dimensions. The 3D-Var technique is an improved version of the OI, for example: the analysis variables are spectral coefficients instead of gridpoint values and all observations are used at once instead to solve a single global problem. The objective is to minimize the cost function build as a combination of forecast and observation deviations from the desired analysis (weighted by means of the corresponding forecast- and observation-error covariance matrices). A graphical schematic of the technique is shown in the dashed square of Fig. 3.5.

The four-dimensional variational data assimilation (4D-Var, i.e., Courtier et al. 1994; Rabier et al. 2000; Mahfouf and Rabier 2000) is a generalization of the 3D-Var technique. The added dimension is time meaning that, in the 4D-Var framework, the observations are distributed in time. The equations are the same, provided the observation operators are generalized to include a forecast model that will allow a comparison between the model state and the observations at the appropriate time. The 4D-Var DA technique is illustrate in Fig. 3.5 with an schematic representation, as well as, the transition from 3D-Var to 4D-Var by adding the time dimension.

In an ensemble prediction system framework, the DA methodology can be used to provide an ensemble of initial conditions thanks to the assimilation of an ensemble of perturbed observations. This approach has been widely tested by many operational centers. In fact, the 3D-Var technique was implemented in ECMWF

[4]Optimal interpolation interpolates arbitrarily located observations to a regular grid using a background field as a first guess. The merged field is optimal in the sense that it has the lowest error variance. Both fields, observations and background, may contain errors.

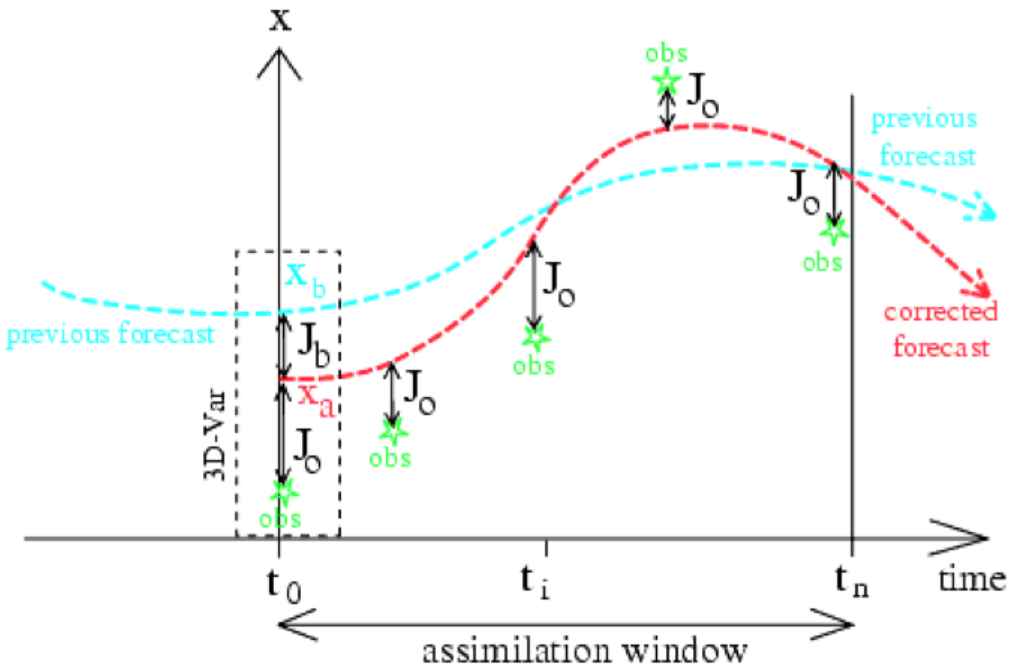


Figure 3.5. *Example of 4D-Var intermittent assimilation in a numerical forecasting system. The technique is applied every 6 hours to assimilate the most recent observations, using a segment of the previous forecast as background. This updates the initial model trajectory for the subsequent forecast. Schematics notation: X_a analysis, X_b background, J_b cost function first term (as in the 3D-Var), J_o cost function second term and obs observations. (From Bouttier and Courtier 1999.)*

operations on January 1996, and was replaced on November 1997 by the 4D-Var technique (Rabier et al., 2000). The same occurred at the Meteorological Service of Canada (MSC) that implemented the 3D-Var system in 1997 (Gauthier et al., 1999) to replace it by the 4D-Var on 2007 (Gauthier et al., 2007). Other centers have also implemented a 4D-Var global assimilation system like, the MetOffice (Rawlins et al., 2007), the Japan Meteorological Agency (Kadowaki, 2007) and Météo-France (Gauthier and Thépaut, 2001).

3.1.4 Ensemble Kalman filter

The Ensemble Kalman filter (EnKF) is a sequential data assimilation method that has gained a lot of popularity in the weather forecasting community. It was proposed by Evensen (1994) as an alternative to the computationally extremely demanding approximate error covariance equation used in the extended Kalman

filter[5]. The EnKF technique address the errors and non-linearities of the initial conditions at a valid time through explicitly describing the flow-dependent forecast error structures, in contrast to the 4D-Var approach where their description is implicit. Many papers have further developed and examined this technique like Evensen and van Leeuwen (1996), Houtekamer and Mitchell (1998), Hamill and Snyder (2000), Bishop et al. (2001), Houtekamer and Mitchell (2001), Whitaker and Hamill (2002), Keppenne and Rienecker (2002), Tippett et al. (2003) and Evensen (2003) which provides a review and overview of the technique and has become a reference document for the basic methodology.

The EnKF can be tackled by two different approaches (Palmer and Hagedorn, 2006):

1. *Perturbed observations* (Evensen, 1994; Houtekamer and Mitchell, 1998; Hamill and Snyder, 2000): An ensemble of data assimilations is built using the same observations to which random noise has been added. The ensemble is then used to estimate the forecast error covariance. This approach has proven a worthy competitor to the operational 3D-Var technique, a well supported technique by the community (Houtekamer et al., 2005).

2. *Square Root Filters* (Tippett et al., 2003; Bishop et al., 2001; Whitaker and Hamill, 2002): This approach does not require to perturb the observations, instead ensemble forecasts are used to obtain a background error covariance at the time of the analysis. The analysis increment and the analysis error covariance are obtained thanks to introducing the new observations using the full Kalman Filter. Then, the new initial analysis perturbations are obtained by solving a square root filter that relates the background and analysis error covariance. The EnKF analysis cycle is made up by the steps just described, Fig. 3.6 schematically represents this sequence.

A key point in the popularity gained by the EnKF is its simple conceptual formulation and relative ease of implementation, e.g. it requires no derivation of a tangent linear operator or adjoint equations and no integrations backward in time (Evensen, 2003).

[5]The extended Kalman filter is the nonlinear version of the Kalman filter which linearizes an estimate of the current mean and covariance.

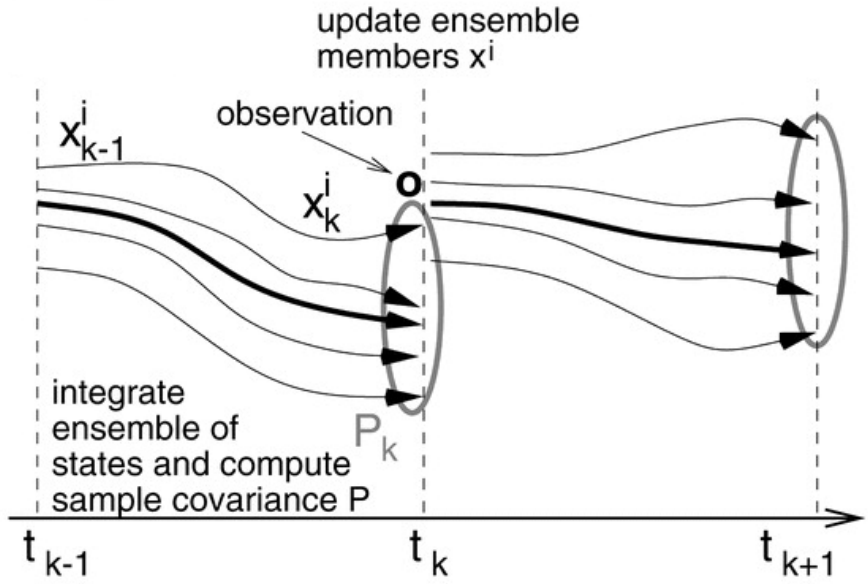


Figure 3.6. *Schematic of the ensemble Kalman filter (Square Root Filters approach). (Adapted from Reichle and Koster 2003.)*

3.1.5 Multimodel, multiphysics and stochastic physical parameterizations

The previous techniques accounted for uncertainties in the initial conditions but not for model deficiencies. In order to include model uncertainties when building ensembles several strategies have been developed:

1. *Multimodel*
   A different forecast model is used to generate each ensemble member in order to account for structural modelling uncertainty inherent in each model construction (e.g. Evans et al. 2000; Krishnamurti et al. 2000a; Stensrud 2001).

2. *Multiphysics*
   Each ensemble member uses a different set of physical parameterization schemes to take into account the present uncertainty in model parameters controlling the best-estimate outputs of parameterizations of sub-grid-scale processes, for example cloud physics (e.g. Houtekamer et al. 1996; Stensrud et al. 1999).

3. *Stochastic physical parameterizations*
   Each ensemble member is perturbed by a stochastic forcing term that represents the statistical fluctuations in the subgrid-scale fluxes (stochastic diabatic tendencies) as well as altogether unrepresented interactions between the resolved an unresolved scales (stochastic kinetic energy backscatter). This approach has shown its potential in numerous studies since Buizza et al. (1999) introduced it, some examples are Palmer (2001); Grell and Devenyi (2002) and Shutts (2005).

3.1.6 Hybrid approach: Accounting for model and initial conditions uncertainties

Another approach to build ensembles is to mix the previous exposed techniques in order to simultaneously account for model and initial conditions uncertainties. Meng and Zhang (2007) is an example where different physical parameterizations are introduced after perturbing the initial state. Another example, that has already proven its value (Callado et al., 2011), is the Multimodel Short Range Ensemble Prediction System (SREPS) developed by the Spanish Meteorological Service (AEMET) generated by five different limited area models and initialized by initial and boundary conditions provided by five different global deterministic model and multiple analysis sources from weather services worldwide (García-Moya et al., 2011).

3.2 From global to regional area modeling

The use of limited-area numerical weather prediction model (LAM) has been greatly expanded during the last decades due to the increase in computational power of computers and workstations at an affordable price. A LAM covers only a part of the Earth in contrast to a global model, that cover the entire Earth, allowing to use a finer or smaller grid spacing than global models. Thanks to this finer grid space, the regional modes are able to resolve explicitly smaller-scale meteorological phenomena that cannot be represented on the coarser grid of a global model. On the downside, LAMs need to be initialized and forced at lateral boundaries by a global model in order to allow systems from outside the regional

model domain to move into their area. Meaning that, when dealing with regional models, apart from the uncertainty and errors attributable to the regional model itself, one also needs to account for the uncertainty and errors attributable to the global model used for the boundary conditions (Warner et al., 1997).

Several studies have already addressed this issue: for example Torn and Hakim (2008) and Stensrud et al. (2009) apply one of the various methods presented in Torn et al. (2006) for creating an ensemble of lateral boundary conditions (with analysis errors in the interior domain comparable to a global ensemble) to examine a pseudo-operational EnKF; other studies use various analyses from different operational forecast centers as initial and boundary conditions, like Eckel and Mass (2005) for their study of a multimodel and multiphysics mesoscale EPS; and others use a multimodel formed by a combination of several operational forecast centers control runs plus perturbations of these runs, like Fujita et al. (2007) in their exploration of various ensembles dealing with the influence of initial conditions and multiphysics, or Meng and Zhang (2007) in their test of an EnKF for data assimilation.

Hence, even though the boundary condition error has not yet been defined unequivocally, studies using LAM EPSs take this uncertainty as an additional part of the total uncertainty of the problem. On the other hand, a recent study (Vié et al., 2011) stated that a better solution would be to use a higher-resolution global EPS to circumvent the need for an intermediate downscaling in the ensemble configuration or an intermediate downscaling ensemble taking into account all the uncertainty sources. A disadvantage of these alternatives is that both require much longer computing time.

3.3 Superensemble

The ensemble prediction systems outputs often undergo statistical postprocessing in an attempt to improve the prediction. As previously stated at the beginning of this chapter, the ensemble mean of an EPS, also known as poor-man ensemble mean, is as good or even better than a single forecast. Fritsch et al. (2000) stated that this improvement comes from overlapping differences in the sign of the errors associated with individual traveling disturbances. Moreover, if each ensem-

ble member is bias[6] corrected, the resulting ensemble is generally more accurate. For example, Stensrud and Yussouf (2003) showed that a simple bias corrected ensemble mean performance is comparable to the performance of a nested grid model output statistics based upon a multiple linear regression approach (Glahn and Lowry, 1972; Jacks et al., 1990) , with the advantage that it does not require a long data archive to produce good results.

The superensemble (Krishnamurti et al., 1999, 2000a,b, 2001) can be seen as the next generation of bias corrected ensemble mean forecasting, where not only the biases of individual members are recognized and corrected, but past performances of members are also recognized. Whereas the poor-man ensemble mean and bias corrected mean involves a straight average of all the members involved, the superensemble assigned weights to each member. The actual weights are determined through a multiple linear regression technique, where members forecasts are regressed against the observed state in order to reward better-performing members with higher weights and poorer-performing members with lower weights.

The superensemble generation comprises a training phase and a forecast phase. The training phase is made up of previous forecasts from the ensemble members and the corresponding observed states. As previous outlined, the ensemble members variables of latitude, longitude, and intensity are regressed against the observed states through a linear regression technique. The linear regression technique involves a minimization function that acts to limit the spread between the variables of the members and the observed state. This minimization function is described by

$$G = \sum_{t=1}^{T} \left(S_t - O_t\right)^2 \ , \tag{3.1}$$

where G is the minimization function, T is the length of the training period, S_t is the superensemble prediction, and O_t is the observed state. The length of a particular training set is crucial in achieving high-skill forecasts, in fact, between 50 and 75 forecast cases are vital for attaining good forecasts.

For the forecast phase of the superensemble, S_t is derived by using the data gathered through the training phase and current ensemble members forecasts. The forecast phase of the superensemble constructed with bias corrected data is given

[6]The bias is a term which refers to how far the average statistic lies from the parameter it is estimating, that is, the error which arises when estimating a quantity.

by

$$S_t = \bar{O} + \sum_{i=1}^{N} a_i \left(F_{t_i} - \bar{F}_i \right) \ , \tag{3.2}$$

where $\bar{O}$ is the observed mean value in the training phase, N is the number of ensemble members, a_i represents the regression coefficient, weight, for member i. F_{t_i} is the variable forecast made by member i, and $\bar{F}_i$ is the mean of a particular variable over all the forecasts in the entire training period. The susperensemble, S_t, is computed for each variable at each forecast hour. A schematic view of operational application of the superensemble is shown in Fig 3.7.

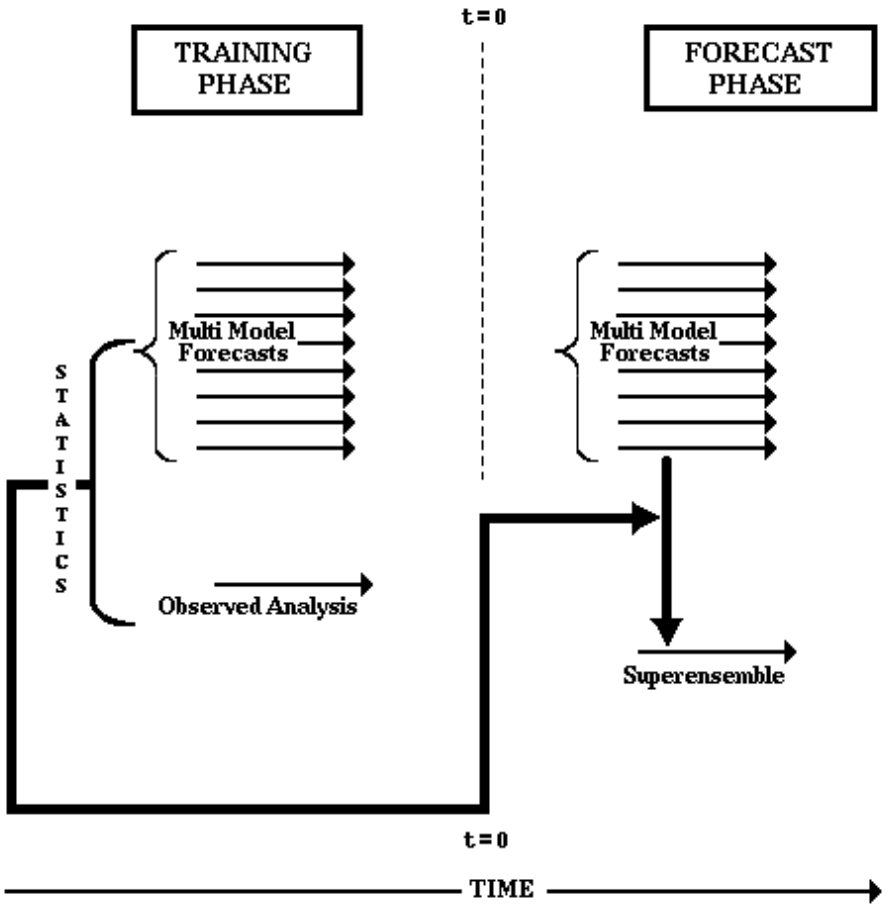


Figure 3.7. *The vertical line in the center denotes time $t = 0$, and the area to the left denotes the training area where a large number of forecasts experiments are carried out by multianalysis-multimodel system. During the training period, the observed fields provide statistics that are then passed on to the area on the right, where $t > 0$. Here the multianalysis-multimodel forecasts along with the aforementioned statistics provide the superensemble forecasts. (From Krishnamurti et al. 2001.)*

More recent applications have been carried out by several authors testing the methodology in different contexts. As examples, Cane and Milelli (2005, 2006) apply the superensemble approach on variables like temperature and wind, while Yun et al. (2005) focuses on seasonal precipitation, and Cane and Milelli (2010) on average precipitation over warning areas and Krishnamurti et al. (2008) on the diurnal cycle of the precipitation. Furthermore, a recent application of the superensemble applied on a muliphysics ensemble will be thoroughly explored on Chapter 9.

Chapter 4

Forecast Verification

Quoting Murphy and Winkler (1987): "*Forecast verification is the process and practice of determining the quality of forecast, and it represents an essential component of any scientific forecasting system.*" It is an essential component because it allows to monitor and improve the quality of a forecast, as well as compare its quality with a different forecast system, all in an objective way.

But *What is a good forecast?*, Dr. Allan Hunt Murphy (1931-1997) established the guidelines to answer this question in his essay: "What is a good forecast? An essay on the nature of goodness in weather forecasting" (Murphy, 1993). In this paper, Dr. Murphy points out three kinds of forecast goodness:

1. Consistency: correspondence between forecasts and judgments. A forecast is consistent when the uncertainty inherent in forecasters' judgments is properly reflected in their forecasts.

2. Quality: correspondence between forecasts and observations. The traditional approach to forecast verification.

3. Value: incremental benefits of forecasts to users. A forecast has no value on its own, it only acquires value when is being used.

If we focus on the quality of the forecast, Murphy (1993) describes nine different quality attributes of the forecast that need to be evaluated on a verification procedure. These nine attributes are[1]:

[1]These list can be found on the publicly available verification web-page created and maintained by the Joint Working Group on Verification (JWGW) under the WMO/World Weather

1. Bias: the correspondence between the mean forecast and mean observation.

2. Association: the strength of the linear relationship between the forecasts and observations (for example, the correlation coefficient measures this linear relationship)

3. Accuracy: the level of agreement between the forecast and the truth (as represented by observations). The difference between the forecast and the observation is the error. The lower the errors, the greater the accuracy.

4. Skill: the relative accuracy of the forecast over some reference forecast. The reference forecast is generally an unskilled forecast such as random chance, persistence (defined as the most recent set of observations, *persistence* implies no change in condition), or climatology. Skill refers to the increase in accuracy due purely to the *smarts* of the forecast system. Weather forecasts may be more accurate simply because the weather is easier to forecast - skill takes this into account.

5. Reliability: the average agreement between the forecast values and the observed values. If all forecasts are considered together, then the overall reliability is the same as the bias. If the forecasts are stratified into different ranges or categories, then the reliability is the same as the conditional bias, i.e., it has a different value for each category.

6. Resolution: the ability of the forecast to sort or resolve the set of events into subsets with different frequency distributions. This means that the distribution of outcomes when A was forecast is different from the distribution of outcomes when B is forecast. Even if the forecasts are wrong, the forecast system has resolution if it can successfully separate one type of outcome from another.

7. Sharpness: the tendency of the forecast to predict extreme values. To use a counter-example, a forecast of *climatology* has no sharpness. Sharpness is a property of the forecast only, and like resolution, a forecast can have this attribute even if it's wrong (in this case it would have poor reliability).

Research Program and the WMO Working Group on Numerical Experimentation (`http://www.cawcr.gov.au/projects/verification/`).

8. Discrimination: ability of the forecast to discriminate among observations, that is, to have a higher prediction frequency for an outcome whenever that outcome occurs.

9. Uncertainty: the variability of the observations. The greater the uncertainty, the more difficult the forecast will tend to be.

There is a wide range of verification methods that can be used to evaluate the quality of a forecast system, deterministic or probabilistic, measuring the relationship between a forecast, or set of forecasts, and the corresponding observations. Both Jolliffe and Stephenson (2003) and Wilks (1995) provide extensive details on many verification scores, as well as the already mentioned verification web-page[2]. A brief introduction and description of the verification measures that are used in this Thesis is done bellow.

4.1 Point-to-point verification scores

Any point-to-point verification procedure involves a comparison between matched pairs of forecasts and the observations to which they pertain. In order to analyze this relationship, for dichotomous forecasts, contingency tables are used. A dichotomous forecast refers to the yes/no nature of the forecast at each point, for example a rainfall threshold defines the transition between a rain event against a nonrain. The contingency table, shown in Table 4.1, counts for each of the four possible outcomes for the observed and forecast event: Hit (a), False alarm (b), Miss (c) and Correct rejection (d).

		Observed	
		Yes	No
Forecast	Yes	Hit (a)	False alarm (b)
	No	Miss (c)	Correct rejection (d)

Table 4.1. *Contingency table for observed event and forecast event in a 2x2 problem.*

[2] `http://www.cawcr.gov.au/projects/verification/`

4.1.1 Relative Operating Curve (ROC)

The ROC was introduced in meteorology by Mason (1982) from the signal detection theory and measures the ability of the forecast to discriminate between two alternative outcomes, thus measuring resolution. The ROC is obtained by plotting, as illustrated in Fig. 4.1, probability of detection [POD=a/(a+c)] against the probability of false detection [POFD=b/(b+d)] using a set of increasing probability thresholds (for example, 5%, 15%, 25% and so on) to make the yes/no decision (see Schwartz et al. (2010) for a more complete description of how to calculate a ROC for a probabilistic forecast). The area under the ROC curve (ROC area) is frequently used as score, in fact an area of 0.5 indicates no skill and of 1 a perfect skill. Moreover, according to Buizza et al. (1999) and Stensrud and Yussouf (2007) forecasting systems with ROC areas greater than 0.7 not only have better skill than a random forecast but are useful.

It is worth to note that the ROC curve can also be computed for a deterministic forecast using a different criteria, such as the median, certain percentiles, or chosen thresholds of interest instead of the probability thresholds used in the probabilistic forecast.

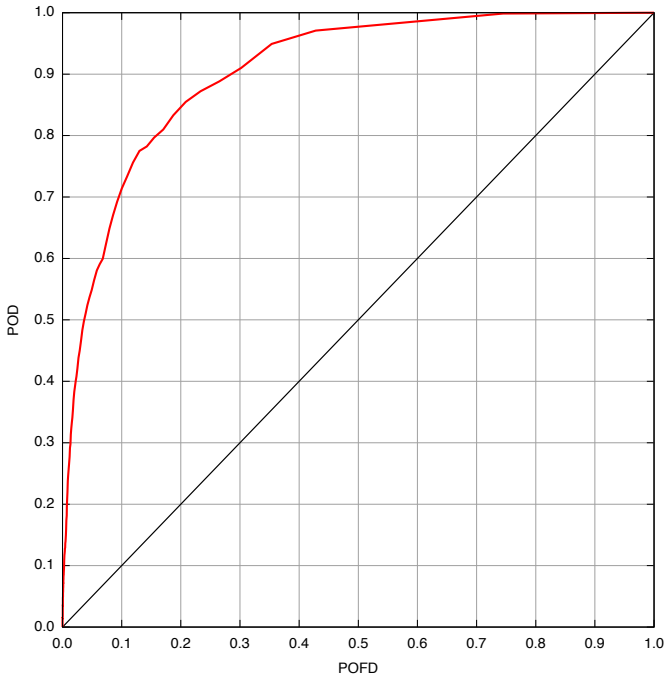


Figure 4.1. *ROC curve, an example.*

4.1.2 Frequency Bias score

The Frequency Bias, referred to as Bias from now on, indicates how the forecast event frequency, F, compares to the observed event frequency, O.

$$\text{BIAS} = \frac{F}{O}$$

An unbiased forecast has a value of 1, while Bias > 1 indicates an overpredicting forecast and Bias < 1 underpredicting.

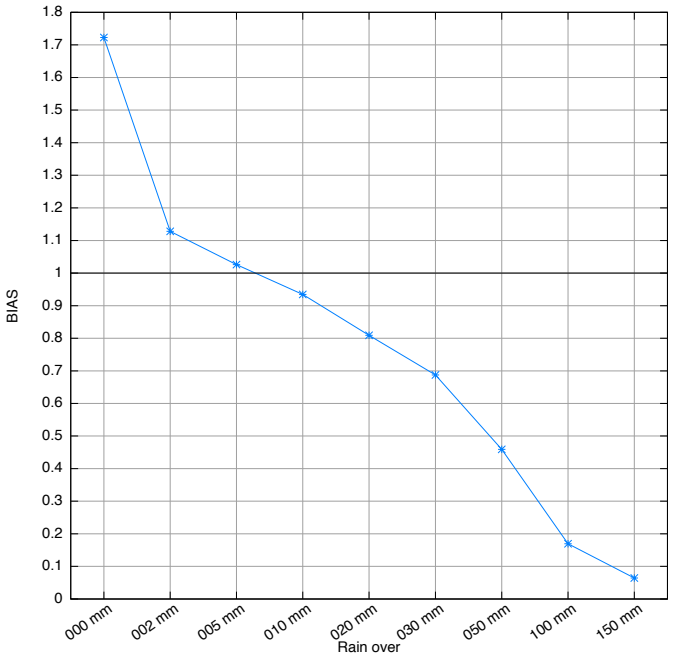


Figure 4.2. *Bias for a test forecast, an example.*

4.1.3 Taylor diagram

Taylor diagrams (Taylor, 2001) provide a way of graphically summarizing how closely a pattern (or a set of patterns) matches observations. The similarity between two patterns is quantified in terms of their correlation, their centered root-mean-square difference and the amplitude of their variations (represented by their standard deviations). The perfect score is obtained when the data point representing the forecast field matches up with the observed one. The radial distance from the origin is proportional to the standard deviation of a pattern. The centered

RMS difference between the observed and forecast field is proportional to their distance apart. The correlation between the two fields is given by the azimuthal position of the forecast field. The standard deviation and centered RMS difference units are rainfall millimeters. It is worth to note that the means of the fields are subtracted, so the diagram does not provide information about overall biases, but solely characterizes the centered pattern error. An example of a Taylor diagram is displayed on Fig. 4.3. .

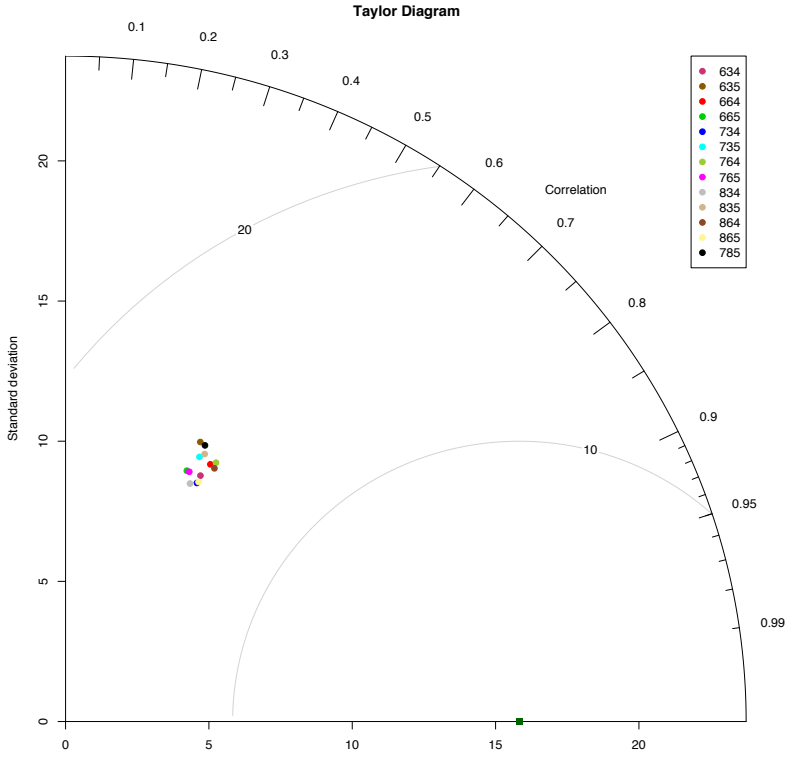


Figure 4.3. *An example of a Taylor diagram for 13 different forecasts valid at the same time over the same region. The observed field is represented by the green solid square.*

4.1.4 Brier score

The Brier score (Brier, 1950) assesses the accuracy of a probability forecast, in terms of predicting whether or not an event occurred. Thus defined by the mean squared error of probabilistic forecasts with events assigned a value of 1 and non-events zero, can be decomposed into the sum of three individual parts related to

reliability, resolution and the underlying uncertainty of the observations (Murphy, 1973). Therefore, if the verification sample is partitioned into categories according to the forecast probabilities, the Brier score can be defined as follows:

$$\mathrm{BS} = \frac{1}{n}\sum_{i=1}^{k}\left(f_i - o_i\right)^2 =$$
$$\frac{1}{n}\sum_{i=1}^{L} n_i\left(f_i - \overline{o}_i\right)^2 - \frac{1}{n}\sum_{i=1}^{L} n_i\left(\overline{o}_i - \overline{o}\right)^2 + \overline{o}\left(1 - \overline{o}\right)$$

where k is the number of forecast/event pairs, L is a discrete number of forecast values or categories (often 11, corresponding to 0,10,..,100%, deciles of forecast probability), n the number of realizations of the forecast, f_i the forecast probability corresponding to the i category, n_i the number of times f_i is forecast, o_i the observation probability corresponding to the i category that is 1 or 0 depending on whether the event occurred or not, $\overline{o}_i$ the mean occurrence of event for forecast category i and $\overline{o}$ the base rate or climatological mean. The Brier score ranges from 0 to 1 with a perfect score of 0. The first term on the right-hand side is a measure of reliability, the second term refers to resolution, and the third term represents the uncertainty of the observations, so it is independent of the forecast system. A perfect reliability means a perfect agreement between the forecast values and the observed values, while a perfect resolution means that the forecast system can successfully separate one type of outcome from another. This score can be used to define a positively oriented index, the Brier skill score, that measures the difference between the Brier score for the forecast and the Brier score for the unskilled standard forecast normalized by the total possible improvement that can be achieved. The unskilled standard forecast is often the climatology, then the BSS can be written as

$$\mathrm{BSS} = 1 - \frac{BS}{BS_{climatology}}$$

which ranges from $-\infty$ to 1; 0 indicates no skill compared to the climatology and the perfect score is 1. Fig. 4.4 displays an example of the Brier Score and its components obtained by a forecast.

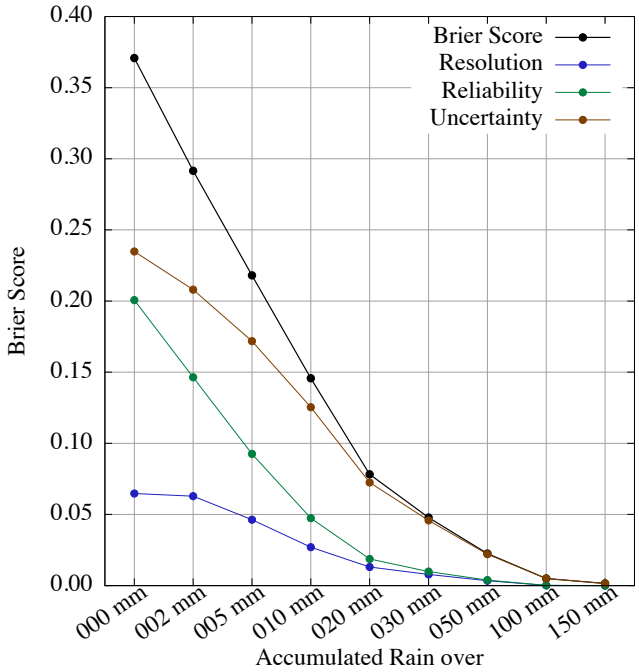


Figure 4.4. *Brier Score and its resolution, reliability and uncertainty terms for an illustrative forecast.*

4.1.5 Attribute diagram

The attribute diagram plots the observed frequency against the forecast probability including the no-resolution and no-skill line and is a measure of how well the predicted probabilities of an event correspond to their observed frequencies (Fig. 4.5). Reliability is indicated by the proximity of the plotted curve to the diagonal. The deviation from the diagonal gives the conditional bias. If the curve lies below the line, this indicates overforecasting (probabilities too high) while points above the line indicate underforecasting (probabilities too low). Then the perfect score is represented by a curve that matches the diagonal and the climatological frequency of the event is represented by both horizontal and vertical lines (no-resolution line), while the 0.5 slope that crosses the no-resolution and perfect score lines represents the no-skill line. Points between the no-skill line and the diagonal contribute positively to the Brier skill score.

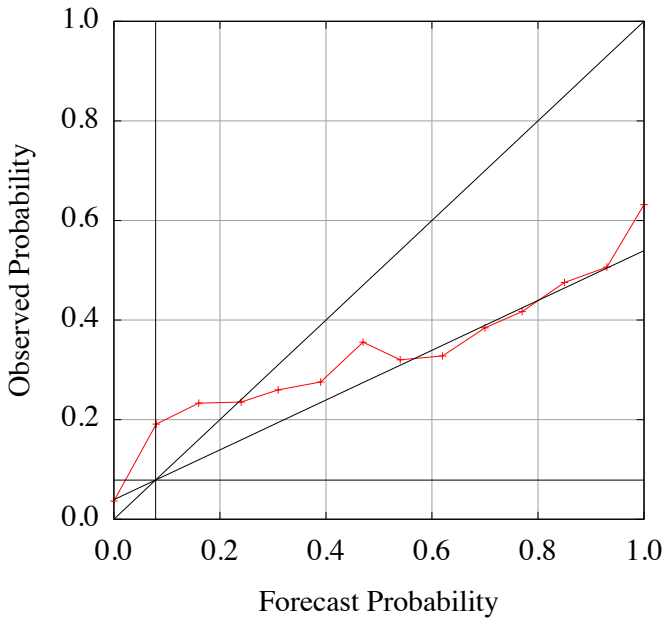


Figure 4.5. *An example of a forecast attribute diagram.*

4.1.6 Rank histogram

The rank histograms (Talagrand et al., 1997) indicate how well the ensemble forecast spread represents the true variability or uncertainty of the observations and are constructed by combining the verification data with the ensemble members and determining what rank represents the verification data in the combined data. The rankings from the forecast samples are then represented in a histogram just like in Fig. 4.6. A flat histogram means that verification is equally likely to be found anywhere within the ensemble, implying that the ensemble has the right spread. An U-shape means that the verification is more likely to be found outside of the ensemble range, so the ensemble spread is too small. On the other hand a dome-shape means that the ensemble spread is too large and most observations are falling near the center of the ensemble. An asymmetric curve indicates that the ensemble contains bias. More detailed information on rank histogram interpretation can be found in Hamill (2001).

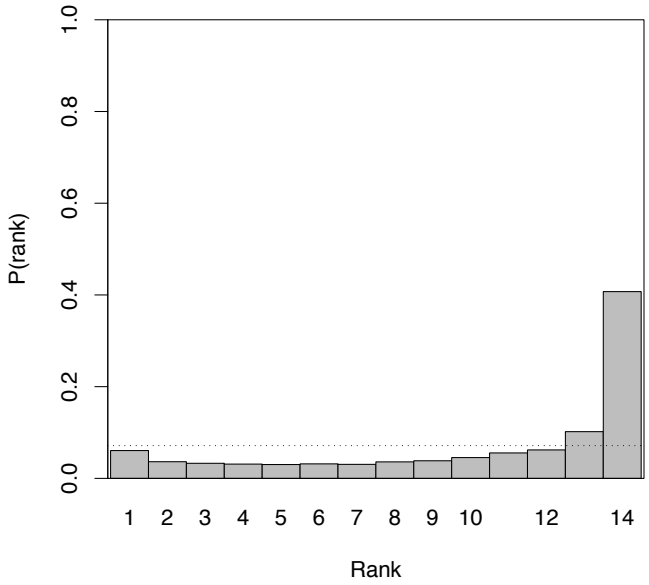


Figure 4.6. *Rank histogram, an illustration.*

4.2 Spatial verification techniques

The use of traditional point-to-point verification scores to evaluate a field, like the precipitation field, characterized by highly localized gradients (with large fraction of zones and times where the precipitation is zero) presents several problems (Mass et al., 2002). The double penalty is one of them, for example when the shape and pattern of the precipitation field is accurate but is displaced in space (Ebert and McBride, 2000; Baldwin and Kain, 2006).

Several spatial verification techniques have been developed to help deal with these issues, for example scale decomposition, object-based verification and fuzzy verification (see Casati et al. (2008) and Gilleland et al. (2009) for reviews). The main drawback of these methods is that even though they are more advanced and precipitation oriented, at their current development stage, they rely on observations defined continuously over a spatial domain. Therefore if the observations are on a sparse network, instead of on a dense or radar/satellite-based network, the verification procedure has to lean on point-to-point scores. These point-to-point verification methods can operate with an observational network not defined continuously over a spatial domain, the kind of observational raingauge network available for this Thesis.

Chapter 5

Potential Vorticity thinking

The *Potential Vorticity thinking* provides a conceptual framework in which to interpret numerical analyses and prognoses. Even though the concept of Potential Vorticity (PV) was introduced by Prof. Carl-Gustaf Arvid Rossby and Prof. Hans Ertel in the early 40's, it was not until Prof. Brian John Hoskins and his colleges wrote the paper called "On the use and significance of isentropic potential vorticity maps" (Hoskins et al., 1985) that the understating and use of the PV thinking really took off.

For a hydrostatic atmosphere, with potential temperature as a the vertical coordinate the Ertel's potential vorticity is given by

$$q = \frac{1}{\rho} \vec{\eta} \cdot \vec{\nabla} \theta \ , \tag{5.1}$$

which is conserved following three-dimensional, adiabatic, frictionless motion (Rossby, 1940; Ertel, 1942). Here $\vec{\eta}$ is the absolute vorticity vector, θ the potential temperature, and ρ the density. In essence, the potential vorticity is proportional to the product of vorticity and stratification that, following a parcel of air, can only be changed by diabatic or frictional processes.

The properties that make the potential vorticity an important and very used variable for representing dynamical processes in the atmosphere are:

1. *Conservation Principle*: If contributions from diabatic and turbulent mixing an neglected (i.e., adiabatic and frictionless motion) the potential vorticity is conserved along its trajectory. The potential vorticity represents the vorticity

that would potentially manifest an air parcel if it were brought adiabatically and without friction to a standard latitude and static stability.

2. *Invertibility Principle*: given the proper boundary conditions and the balance condition[1] imposed on the wind, it is possible to determine, uniquely, the distribution of both vorticity and static stability associated with a PV field. In other words, if the distribution of PV is known, then the wind and temperature fields are also known (Bluestein, 1993).

3. *Climatological distribution of PV*: a useful property when describing and understanding atmospheric dynamics (e.g. dynamical tropopause and PV anomaly).

The climatological PV distribution, on Fig. 5.1, shows that on average the PV ranges from 0.3 to 0.5 PVU[2] in the low and middle troposphere and reaches 1 PVU in the upper troposphere and 3 PVU in the stratosphere after a rapid increase. This rapidly increase to 3 PVU is due to the strong increase of static stability characteristic of the stratosphere. The discontinuity provoked by this rapid increase together with the PV conservation property lead to define the 1.5-PVU surface (sometimes the 2-PVU) as the *dynamical tropopause* surface separating the troposphere, characterized by weak and quasi-uniform PV, from the stratosphere, with strong PV.

An abrupt folding or lowering of the dynamical tropopause can also be called an upper PV anomaly. When this occurs, stratospheric air penetrates into the troposphere resulting in high values of PV with respect to the surroundings, creating a positive PV anomaly. In the lower levels of the troposphere, strong baroclinic[3] zones often occur which can be regarded as low level PV anomalies. Under the assumption of a three dimensional balance between the fields of mass (potential temperature), pressure and wind, positive PV anomalies are connected with cyclonic vorticity and negative PV anomalies with anticyclonic vorticity (Hoskins et al., 1985).

- In the case of a positive PV anomaly, the isentropes are characterized by higher values than in the surrounding areas above the anomaly (indicating

[1] The balance condition links the wind field to the temperature field (a stability measure).

[2] 1 PVU = $10^{-6}\mathrm{m}^2\mathrm{s}^{-1}\mathrm{Kkg}^{-1}$

[3] A baroclinic atmosphere is one for which the density depends on both the temperature and the pressure.

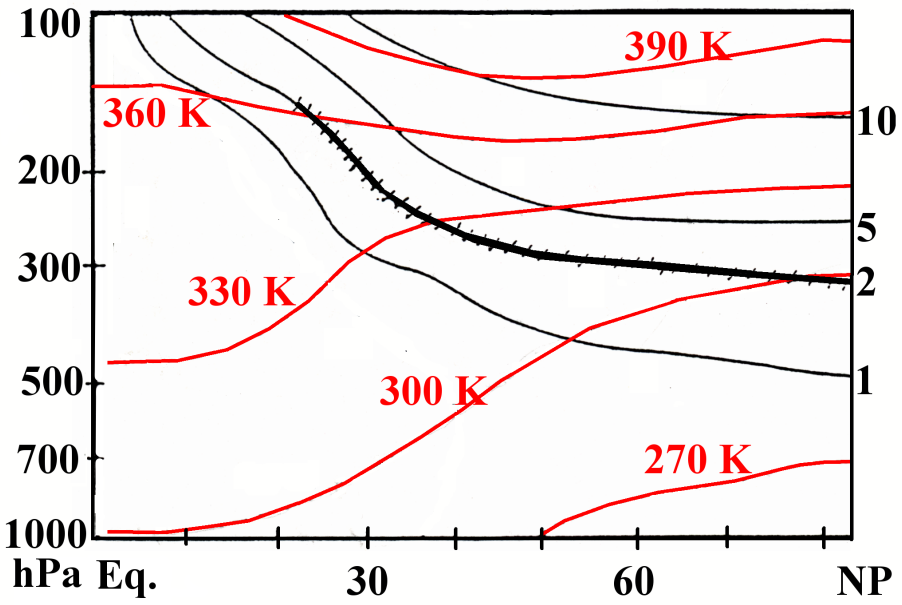
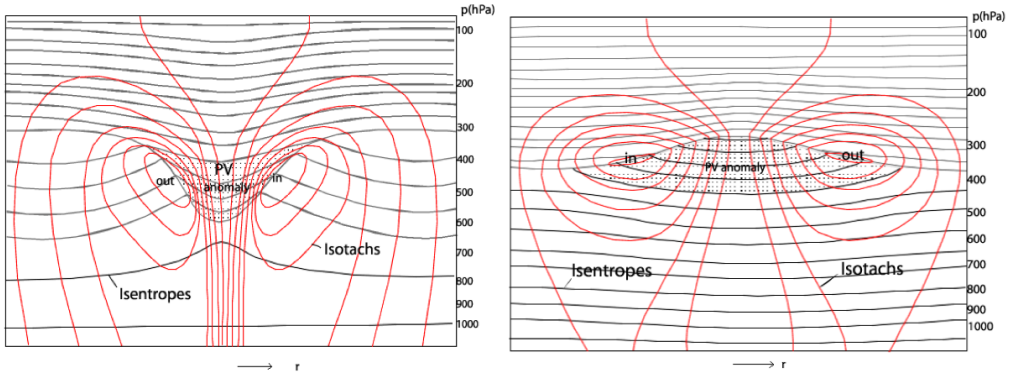


Figure 5.1. *Climatological distribution of potential vorticity. The black lines represent lines of constant Ertel potential vorticity while the red lines represent lines of constant potential temperature. The dynamic tropopause is represented by the bold black line (2-PVU surface). (From Hoskins 1990.)*

warmer air) and lower values below (indicating colder air). The height of the tropopause has a local minimum. The corresponding pressure and wind field shows an area with low pressure and a cyclonic circulation (Fig. 5.2.a).

- In the case of a negative PV anomaly, the isentropes are characterized by lower values than in the surrounding areas above the anomaly (indicating colder air) and lower values below (indicating warmer air). The height of the tropopause has a local maximum. The corresponding pressure and wind field shows an area with high pressure and an anticyclonic circulation (Fig. 5.2.b).

The upper PV anomaly perturbs the wind and pressure fields throughout the whole depth of the troposphere (Fig. 5.2). The effect of this perturbation is proportional to the horizontal scale of the anomaly and inversely proportional to static stability. Observations and model fields show the existence of low level PV anomalies. These low levels PV anomalies do not have stratospheric origin but are being formed in strong baroclinic zones where much release of latent heat takes place. A more thorough description of PV thinking and an extensive review of its applications can be found in Hoskins et al. (1985) and Bluestein (1993), among others.



(a) *Positive PV anomaly* (b) *Negative PV anomaly*

Figure 5.2. *Circularly symmetric flows induced by simple, isolated, PV anomalies, whose locations are shown dotted. In (a) the sense of azimuthal wind is cyclonic and in (b) it is anticyclonic. (From Hoskins et al. 1985.)*

5.1 Applications: understanding cyclogenesis

The material conservation, combined with the advantage of invertibility is very useful to describe complex dynamic processes. For example, cyclogenesis can be explained by an interaction between low and upper-level PV anomalies. When the phase difference between two PV anomalies has an optimum value, interaction followed by mutual amplification takes place. Of course the opposite can also occur, in the case of decaying cyclones.

If an upper-level PV anomaly develops, then as a consequence of the PV conservation, positive vorticity is released and high PV values from the stratosphere influence the less stable environment of the troposphere. Fig. 5.3 illustrates cyclogenesis by baroclinic interaction by means of the PV thinking: an upper-level PV anomaly caused by advecting downstream an area of PV maximun to the mid-troposphere induces vorticity through the depth of the troposphere over a baroclinic zone. Fig. 5.3.a shows such a situation with the solid plus sign indicating the upper-level PV anomaly. The thick solid arrow around the PV maximum indicates the cyclonic rotation. This rotation is induced at lower levels of the baroclinic zone as shown by the thin solid circulation arrow. This low level circulation causes warm advection ahead leading to a low level positive temperature anomaly indicated by the open plus sign in Fig. 5.3.b. This temperature anomaly is associated with a cyclonic vortex which is marked by the open arrow at low levels. In turn, this circulation has a positive feedback to the upper troposphere,

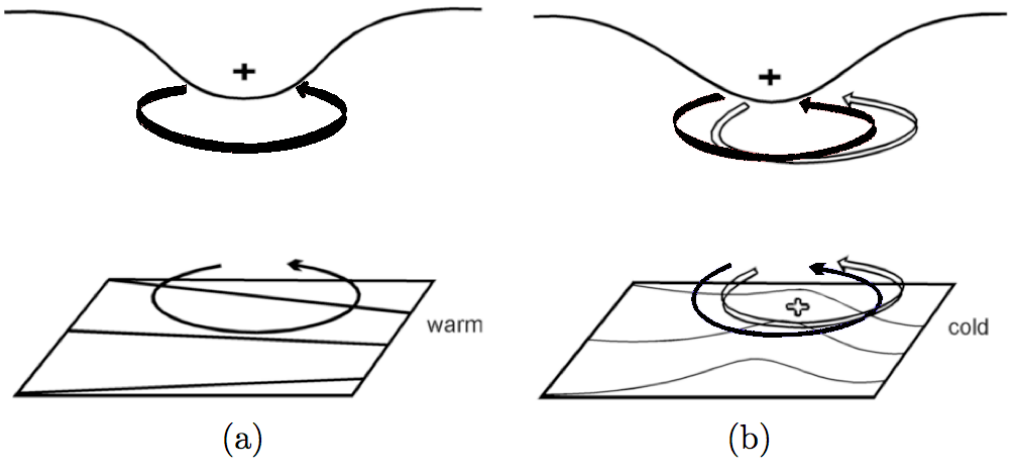


Figure 5.3. *A schematic picture of cyclogenesis associated with the arrival of an upper air PV-anomaly (solid '+' sign) over a low-level baroclinic region. The circulation induced by the anomaly is indicated by a solid arrow, and the potential temperature contours are shown at the lower boundary by thin lines. A low-level PV anomaly (open '+' sign in (b)), can also induce a cyclonic circulation (open arrow), that reinforces the circulation pattern induced by the upper-level PV anomaly. (From Hoskins et al. 1985.)*

shown by an open circulation arrow at higher levels. In parallel a second process is taking place, the induced low level vortex results in a strong equatorward wind component under the upper-level PV anomaly. This southward component also influences the higher levels and leads to an equatorward advection of the upper-level PV anomaly which in turn intensifies the upper-level wave. Within this increased flow, higher PV values to the west of the PV anomaly are advected southward and lower PV values to the east of the PV anomaly are advected northward. As a consequence of the latter process, the eastward movement of the PV anomaly is decreased. Hence, the interaction between low and upper-level circulations and the already ongoing cyclogenesis process will strengthen.

5.2 Potential Vorticity inversion technique

The invertibility principle can be applied to develop a diagnostic system that obtains the wind and temperature perturbations associated with a given PV perturbation distribution. This application is called potential vorticity inversion technique and can be very useful in order to discuss the relevance of the PV-anomalies in the atmospheric behavior (Romero, 2001, 2008; Argence et al., 2008). For exam-

ple, the PV inversion allows to diagnose developments in which the PV-anomalies are not growing, but are merely changing their relative positions or their shapes. Furthermore, the conservation law for PV will help isolate those disturbances growing chiefly by non-conservative processes while the invertibility principle allows direct calculation of their associated circulation. Several PV inversion techniques have been developed, some examples are the one proposed by Davis and Emanuel (1991), that is also applied in this Thesis and it is outlined bellow, the one by Viúdez and Dritschel (2004), that inverts without any balance equation but based on backward and forward model integrations to recover the balanced solution associated with a given PV distribution and that by Arbogast et al. (2008), that inverts applying a digital-filtering method to a stratified primitive-equation model.

The potential vorticity inversion technique, developed by Davis and Emanuel (1991), uses the balance condition derived by Charney (1955) to link the wind field to the temperature field. This balance condition is very accurate in flows with large curvature because is quite similar to gradient wind balance. The Charney balance equation is obtained taking the horizontal divergence of the horizontal momentum equations and decomposing the wind field into a nondivergent and an irrotational part. After a scaling, the resulting equation may be written in spherical coordinates (λ, ϕ, a), following Davis and Emanuel (1991) formulation (see Romero (2008) for a version in (x,y) map coordinates):

$$\nabla^2\Phi = \nabla\cdot(f\nabla\Psi) + \frac{2}{a^4\cos^2\phi}\frac{\partial\left(\frac{\partial\Psi}{\partial\lambda},\frac{\partial\Psi}{\partial\phi}\right)}{\partial(\lambda,\phi)}\ , \tag{5.2}$$

with Φ the geopotential, Ψ the nondivergent stream function, λ the longitude, ϕ the latitude, and a the earth's radius. Relation (5.2) reduces to geostrophic balance if f is constant and the Jacobian term is neglected.

One more diagnostic equation relating Ψ to Φ is needed in order to close the system. This is obtained from an approximate definition of Ertel's Potential Vorticity (ErPV, Eq. 5.1):

$$q = -\frac{g\kappa\pi}{p}\left(\eta\frac{\partial\theta}{\partial\pi} - \frac{1}{a\ \cos\phi}\frac{\partial v}{\partial\pi}\frac{\partial\theta}{\partial\lambda} + \frac{1}{a}\frac{\partial u}{\partial\pi}\frac{\partial\theta}{\partial\phi}\right)\ , \tag{5.3}$$

where $\kappa = R_d/C_p$, p is the pressure, π is the Exner function $[C_p(p/p_0)^\kappa]$ and serves as vertical coordinate, η is the vertical component of absolute vorticity and the

hydrostatic approximation has been made. Performing the same kind of scaling done to obtain Eq. 5.2 on to Eq. 5.3, results in a relation between the potential vorticity, Φ and Ψ,

$$q = -\frac{g\kappa\pi}{p}\left[\left(f + \nabla^2\Psi\right)\frac{\partial^2\Phi}{\partial\pi^2} - \frac{1}{a^2\cos^2\phi}\frac{\partial^2\Psi}{\partial\lambda\partial\pi}\frac{\partial^2\Phi}{\partial\lambda\partial\pi} - \frac{1}{a^2}\frac{\partial^2\Psi}{\partial\phi\partial\pi}\frac{\partial^2\Phi}{\partial\phi\partial\pi}\right] . \quad (5.4)$$

The approximation used here replaces the vertical derivative of the total wind by the vertical derivative of the nondivergent wind (i.e ., the irrotational component of the wind is very small relative to the non-divergent wind) and uses the Hypsometric equation, that relates the thickness between two isobaric surfaces to the mean temperature of the layer, to replace the θ. Equations (5.2) and (5.4) form a complete system for the unknowns Φ and Ψ, given q.

For boundary conditions, Φ and Ψ are defined in the lateral boundaries (Dirichlet conditions) and their vertical derivatives on the horizontal boundaries are specified (Neumann-type conditions). The observed geopotential serves as Φ on the lateral edges and $\partial\Phi/\partial\pi = -\theta$ is applied at the top and bottom. The gradient of Ψ along the edge is forced to match the normal wind component, and the horizontal boundaries are given by $\partial\Psi/\partial\pi = -\theta$, applied at both the top and the bottom. The obtained solutions are fairly insensitive to this choice of boundary conditions.

The technique used to solve the system is an iterative numerical method of successive overrelaxation (SOR) applied to each vertical level. As long as the ErPV is everywhere positive the method consistently converges. Tests with different initial guesses, done by Davis and Emanuel (1991), showed that the solution obtained was apparently unique, or at least other existing solutions were unreachable.

Part II

Ensemble prediction of Mediterranean high-impact events using PV perturbations

Chapter 6

Stage 1: Building the perturbed initial and boundary conditions ensemble*

As established in the Introduction, the ensemble forecasting technique stands out as a suitable approach to improve the current prediction of high impact weather events associated with Mediterranean cyclones, our target. The proposed ensemble introduces the perturbations on the PV field instead of using a more standard technique (e.g. breeding, EnKF). A method focused on the PV field, which encompasses information regarding the structure and location of the associated surface cyclone, is more appropriate for our target. Indeed, previous studies have already highlighted the sensitivity of cyclones to PV perturbations (e.g. Huo et al. 1999; Romero 2008; Argence et al. 2008).

6.1 Datasets

Ideally, an EPS needs to be a real time fully operational system. The UIB Meteorology Group has been running the MM5 model on a daily basis for some years (see `http://mm5forecasts.uib.es`). Owing to computational limitations, these

*This Chapter is based on the published paper Vich, M., R. Romero, and H. E. Brooks, 2011: *Ensemble prediction of Mediterranean high-impact events using potential vorticity perturbations. Part I: Comparison against the multiphysics approach.* Atmos. Res., 102, 227-241. doi:10.1016/j.atmosres. 2011.07.017.

48 h numerical forecasts, centered over the western Mediterranean region, are initialized with global coarse resolution 24 h forecast fields valid at 00 UTC and forced at the lateral boundaries with the subsequent data (i.e. with 30, 36, ..., 72 h global forecasts). Thus, the mesoscale prediction system can not be considered as real-time in the sense that forecasts, rather than analyses, are used to nest the MM5 model. With such a testbed in mind, the same type of data is used to initialize and force the numerical weather model in the present Chapter. These meteorological data are provided by the European Center for Medium-Range Weather Forecasts.

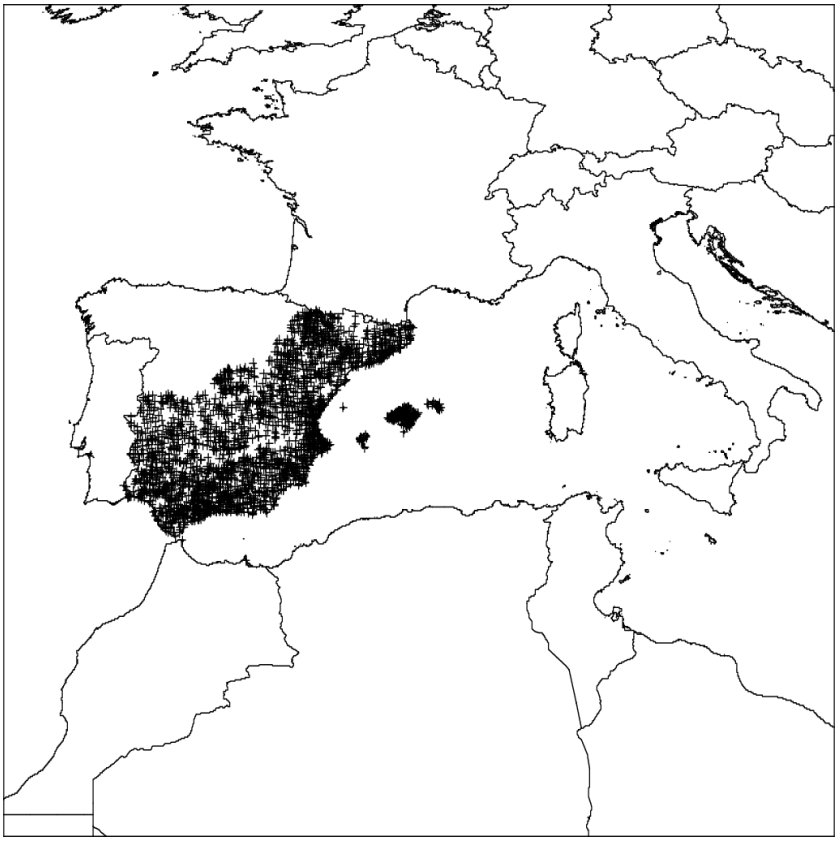


Figure 6.1. *Geographical domain used for the MM5 numerical simulations. The spatial distribution of the AEMET raingauge network used in the verification procedure is plotted using crosses.*

The Mediterranean Experiment on Cyclones that produce High Impact Weather in the Mediterranean, MEDEX, is a project designed to contribute to a better understanding and short-range forecasting of high impact weather events in the Mediterranean, mainly heavy rain and strong winds. Therefore MEDEX provides a suitable database for our study, so the ensembles trial set consists of a collection of 19 MEDEX cyclonic episodes comprising 56 different days between September

1996 and October 2002 (Table 6.1). The observations come from the AEMET (Agencia Estatal de Meteorología - Spanish Weather Service) climatological rain-gauge network which provides 24 h accumulated precipitation from 06 UTC to 06 UTC the next day. Fig. 6.1 shows the spatial distribution of this network over the Mediterranean influenced regions of Spain, with more than 2300 stations depending on the event.

	Date	**Country mainly affected**
01	11-12 Sep. 1996	Spain
02	06-09 Oct. 1996	Italy, Spain
03	14 Oct. 1996	Spain, Italy
04	04-06 Nov. 1997	Portugal, Spain, France
05	11-14 Nov. 1999	Italy, Spain, France
06	10 Jun. 2000	Spain
07	21-26 Oct. 2000	Spain
08	02-05 Nov. 2001	Spain
09	09-13 Nov. 2001	Algeria, Spain, Croatia, Morocco
10	14-16 Nov. 2001	Spain
11	14-15 Dec. 2001	Spain
12	11 Apr. 2002	Spain
13	06-08 May. 2002	Spain
14	12-15 Jul. 2002	Spain, Croatia
15	31 Jul. - 01 Aug. 2002	Spain
16	08-10 Sep. 2002	France
17	12-13 Sep. 2002	Spain
18	23-24 Sep. 2002	Spain
19	08-10 Oct. 2002	Spain

Table 6.1. *List of MEDEX episodes used for this study. A more detailed description can be found at* `http://medex.aemet.uib.es`.

6.2 Potential vorticity error climatology

To assure that the perturbations, introduced into the initial and boundary conditions of the EPS that is being generated, are consistent with the PV field uncertainty range, a quantitative assessment of this uncertainty is required. This is estimated by means of a Potential Vorticity error climatology (PVEC) that compares the ECMWF 24 h forecast and ECMWF analysis PV fields at all available times: 00, 06, 12 and 18 UTC (the alternative approach of comparing two differ-

ent analysis sources, like NCEP and ECMWF, seems to offer essentially the same results; Horvath 2008). For this comparison it is assumed that the main error sources are the displacement between PV structures and differences in intensity following the concept of Hoffman et al. (1995) who decomposed forecast error into displacement, amplitude and residual errors. The PVEC is derived over the same larger domain used in the following Section 6.3 (200x200 nodes with an horizontal resolution of 22.5 km).

In order to obtain statistically significant results, 19 MEDEX cyclonic episodes (described in Table 6.1) containing 56 days are used for the calculations. The PV fields of each 56 days are sampled over the 200x200 mesh at the following pressure levels: 100, 200, 300, 400, 500, 700, 850, 925 and 1000 hPa. The goal is to obtain the climatological (for the 19 MEDEX cyclones sample) displacement and intensity differences between ECMWF 24 h forecast and ECMWF analysis PV structures, fit a function on the displacement and intensity error percentiles found for each pressure level and finally tabulate the coefficients of these functions.

6.2.1 Displacement error

The displacement error (DE) is calculated for each grid point of the 200x200 nodes domain, providing then a statistical significant number of occurrences. At a certain pressure level, for each grid point an area of 450x450 km centered on the grid point is used to compare the ECMWF 24 h forecast with the ECMWF analysis PV field. This area is displaced up to 225 km (10 nodes) grid point by grid point (i.e. in 22.5 km intervals), in the four cardinal directions (441 displacements in total). The correlation between both fields over the 450x450 km area is obtained in each displacement. Therefore, for each grid point we have 441 correlation values. The minimum displacement of the forecast field providing a local (or relative) maximum of correlation is defined as the DE for that grid point. Fig. 6.2 illustrates this process schematically. This process is repeated for each grid point of the domain and for each pressure level. The choice of the area dimension and the maximum displacement tested are not entirely arbitrary, actually the 450x450 km matrix size ensures that, according to the PV field typical scales (Bluestein, 1993), subsynoptic structures will be effectively sampled and the possible displacements up to 225 km ensure that plausible lags between forecast and analysis PV fields can be identified.

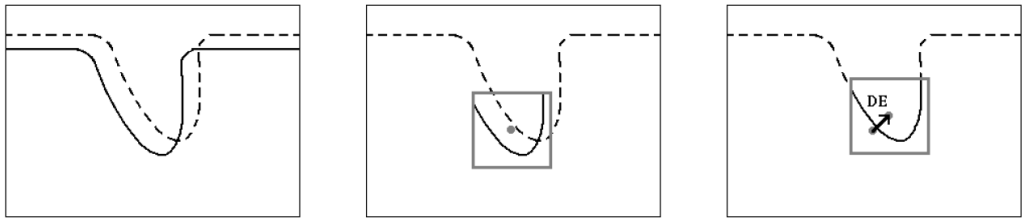


Figure 6.2. *A schematic on how the displacement error is obtained for an arbitrary grid point. a) Initial PV field spatial pattern of the ECMWF 24 h forecast (solid line) and analysis (dashed line), b) The 450x450 km area (grey square) associated to the grid point (grey dot), c) Minimum displacement of the ECMWF 24 h forecast PV field providing local maximum correlation with the analysis PV field.*

6.2.2 Intensity error

As in the DE, the intensity error (IE) is also calculated for each grid point at each pressure level with the aid of the 450x450 km subdomains explained above. The IE is defined at each grid point as the difference in PV field value between the ECMWF 24 h forecast (once displaced the corresponding DE of the grid point) and the ECMWF analysis. This difference is averaged over the 450x450 km area associated to the grid point in question and then divided by the same average of the analysis PV field, the result is the IE value. Again this process is repeated for each grid point of the domain and for each pressure level. The IE represents the difference between the displaced forecast and analysis PV field as an error percentage.

6.2.3 DE and IE percentile levels

The representation of the displacement error density function, expressed as function of number of grid lengths, shows a very clear symmetry along South-North and West-East directions (see Fig. 6.3 for the 300 hPa pressure level). This fact allows to express the error as an absolute value in both directions. Likewise the intensity error presents a very high symmetry between positive and negative values (not shown), so the absolute value is used too. Regarding the dependence on the local PV value, DE appears to be insensitive to this value, while IE exhibits an appreciable dependence. It could be argued that this later case owes to the fact of expressing IE in relative terms, but a similar behavior is found when expressed as

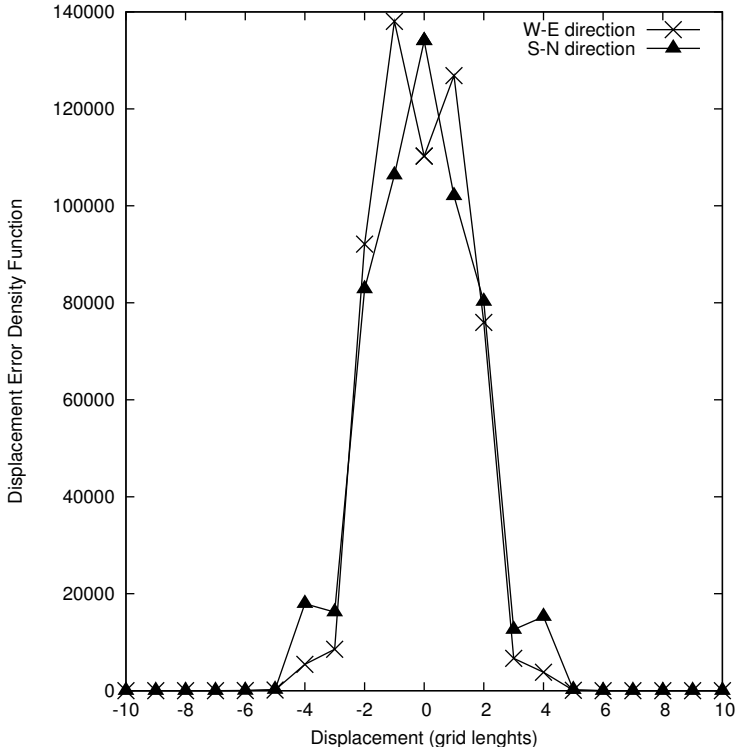


Figure 6.3. *Displacement error density function at 300 hPa. It represents the number of occurrences (grid points) as function of the displacement error in grid lengths. Results for the South-North (solid line) and West-East (dotted line) directions are represented separately.*

an absolute error. Analytical functions have been fitted to the percentile levels of the displacement error percentiles (Fig. 6.4) and the intensity error (Fig. 6.5) using linear-like and power-law function, respectively. These DE and IE percentiles shape the PV error climatology that is used in the ensemble building generation to introduce the perturbation into the initial and boundary states. It should be noted that in the ensemble building generation a superior limit is imposed to the IE ($IE_{max} = 200\%$) to avoid inconsistencies because this relative error tends to infinity when PV tends to zero due to the power-law function properties, as Fig. 6.5 shows.

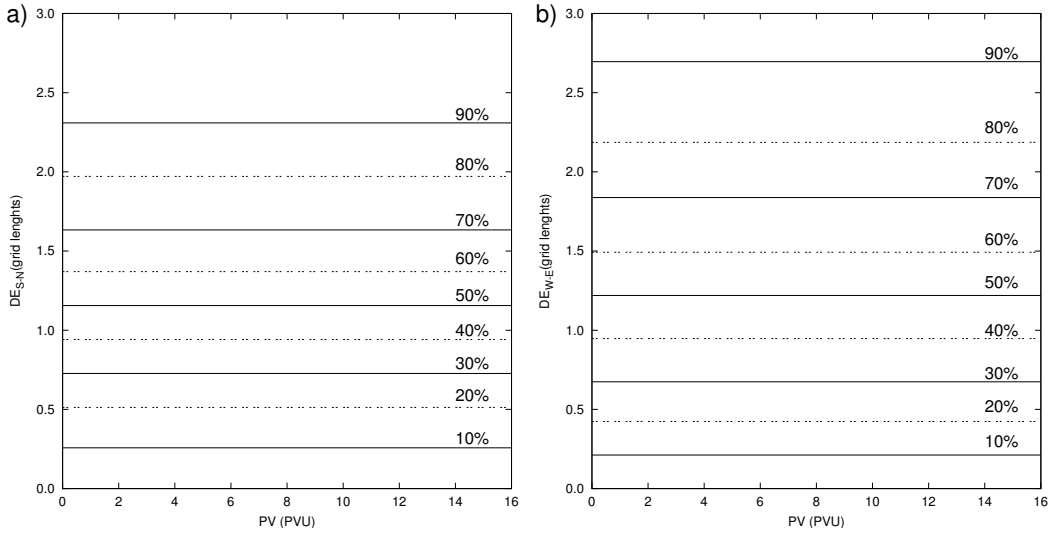


Figure 6.4. *Displacement Error percentile levels along a) South-North and b) West-East directions at 300 hPa.*

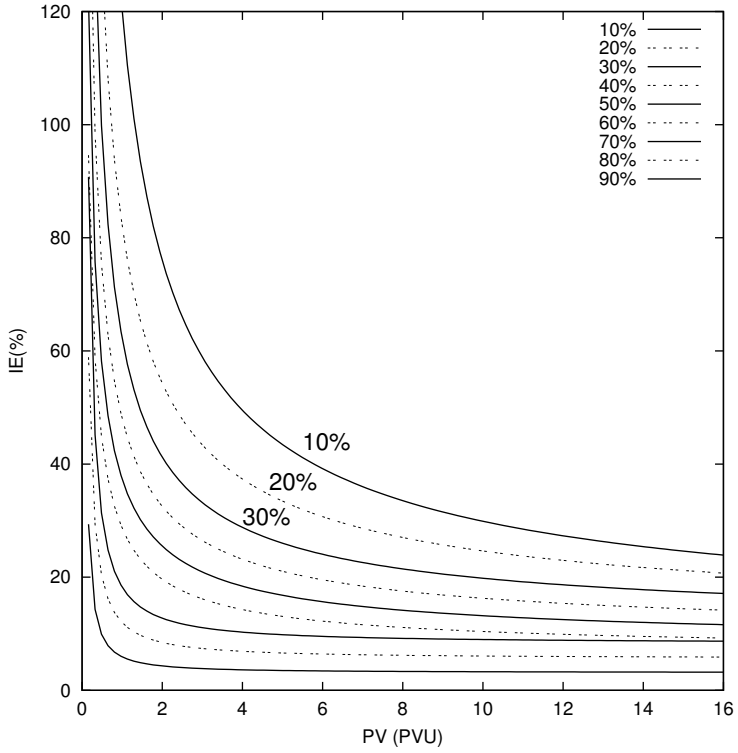


Figure 6.5. *Intensity Error percentile levels at 300 hPa.*

6.3 Perturbed initial and boundary conditions ensemble

The ensemble members are run using the non-hydrostatic mesoscale model MM5, presented in Section 2.1. The simulation domain, Fig. 6.1, covers the whole area affected by the selected MEDEX cyclones, and is defined by a 22.5 km resolution horizontal grid mesh with 120x120 nodes, centered at 39.8° latitude and 2.4° longitude, with 30 sigma levels in the vertical grid mesh. Provided that the observed daily rainfall accumulations span over 06-06 UTC, the forecasting period starts at 00 UTC and extends for 54 h. To be able to compare the forecasts with the observations in the verification procedure, the forecast gridded fields are interpolated over the raingauge locations. Also, since the ensemble only varies the initial and boundary conditions the same physical parameterization set is used on each member, the UIB operational model set made up by the explicit moisture scheme of Reisner graupel (Reisner et al., 1998), the cumulus parameterization scheme of Kain-Fritsch 2 (Kain, 2004), the PBL scheme of MRF (Troen and Mahrt, 1986; Hong and Pan, 1996), the cloud-radiation scheme of Dudhia (1989) and the five-layer soil model described in Dudhia (1996).

The varying initial and boundary conditions are accomplished by perturbing the PV field, thus exploiting the strong connection between PV streamers and cyclogenesis observed in the atmosphere. Moreover, the potential vorticity inversion technique developed by Davis and Emanuel (1991) allows to convert the perturbations introduced into the PV field to the wind and temperature fields in a dynamically consistent manner. This PV inversion technique has already been used successfully by different authors (e.g. Romero 2001, Swarbrick 2001, Rosting et al. 2006, Romero 2008, Amengual et al. 2009 and Baxter et al. 2011) to understand quantitatively the relevance of the PV-anomalies on the atmospheric behavior. This is a neat approach of building physically-consistent perturbations of the primitive mass and wind fields.

As detailed in the previous Section, the introduced perturbations are consistent with the PV field uncertainty range thanks to the developed PV error climatology. Summarizing its results, the displacement errors are found to be symmetric along South-North and West-East directions and show no dependence on the magnitude of PV (Fig. 6.4). Conversely, a strong dependence on the PV value is found for

the intensity error (Fig. 6.5), while both PV errors depend on pressure levels.

These ensemble perturbations are introduced into the PV field using the zones of most intense values and gradients of this field as guidance. This choice assumes that these are the most sensitive zones of the subsequent atmospheric evolution (Garcies and Homar, 2009; Romero et al., 2006). This idea is reinforced by Plu and Arbogast (2005) who suggest that an ensemble based on extracting the PV coherent structures from a given analysis and modifying their features (shape, amplitude and position) would account for part of the analyses errors, and also by Snyder et al. (2003) who stated that the behavior of PV perturbations and their statistics appears to be fundamentally controlled by the reference-state PV gradients. From now on this ensemble will be referred to as the PV-gradient ensemble.

In practical terms, the field given by the difference of the three-dimensional PV field and a highly smoothed version of it is defined as the guidance field. Then, a threshold is defined as the average over the whole domain of the absolute value of the guidance field. Finally, the 3-dimensional regions where the guidance field exceeds, in absolute value, the threshold is where the perturbations are introduced. On each of these 3-dimensional regions or volumes, the intensity -both magnitude and sign- and displacement -magnitude in zonal and meridional axes- perturbations are assigned randomly accordingly to the PV error climatology statistics. However, to avoid discontinuities in the perturbed PV field, the displacement directions of the volumes lie all in the same randomly chosen quadrant. It is worth to note that this randomly chosen quadrant is different from one ensemble member to another. After perturbing, difference between the original and perturbed balanced fields (obtained applying the PV inversion technique to both the original and perturbed PV fields separately) is added to the non-perturbed, mass and wind fields, to produce the initial and boundary conditions of the corresponding ensemble member. A similar strategy has been used to analyze the effects of upper-level PV anomalies on the numerical simulation of Mediterranean severe weather events (Homar et al., 2002, 2003).

It is worth to note that the ensemble building procedure is realized over a domain larger than the simulation domain to avoid inconsistencies in the boundary conditions provided to the ensemble, since the method keeps the boundary values of the inverted fields fixed. While the simulation horizontal domain is made up of 120x120 nodes, the domain used to perturb the initial and boundary states is made up of 200x200 nodes. Since the horizontal resolution is 22.5 km and both domains

are centered at the same geographical point, a symmetric frame of 900 km (40 nodes) exists between these domains for the perturbed fields to effectively reach the simulation domain boundaries. To illustrate the perturbing procedure a case study is presented in the following section.

6.4 Illustration of the methodology: 9-10 June 2000 case study

This Section illustrates the PV-gradient ensemble design step by step through the case study of 9-10 June 2000. The event was mainly characterized by the entrance of an Atlantic low-level cold front and an upper-level trough over the Iberian peninsula. A mesoscale cyclone was generated in the Mediterranean Sea east of mainland Spain leading to a mesoscale convective system that remained quasi-stationary. The most affected region was the northeastern part of the Iberian Peninsula. The heavy rains produced severe floods over densely populated areas and very substantial material losses over Catalonia. The synoptic situation on 10th June 2000 at 00 UTC is shown in Fig. 6.6.a, a more detailed dynamical description of the event can be found in Chapter 11.

Fig. 6.7 shows a horizontal section at 300 hPa of the perturbation volumes, which follow closely the structure of the upper-level trough as one would expect given its definition. Now the locations of the perturbations are known but the characteristics of each perturbation still need to be determined. All perturbation volumes are displaced at once in the same quadrant (chosen randomly), in this case, the southeast sector. Regarding the intensity perturbation (magnitude and sign) and displacement magnitude (zonal and meridional) of each perturbation volume, a PV error climatology percentile level is chosen randomly. As an example, on one volume its PV structure is displaced 1.9 grid lengths in the N-S direction (percentile 80 % from Fig. 6.4.a) and 0.5 grid lengths in the W-E direction (percentile 20 % from Fig. 6.4.b) while its PV value is increased 80 % for PV values of 2 PV units and 60 % for PV values of 3 PVU (percentile 10 % from Fig. 6.5). Note that since the correspondence between grid lengths/PV values and percentile levels depends on the pressure level, the grid lengths and PV values vary with pressure even though the percentile levels are the same for the whole volume.

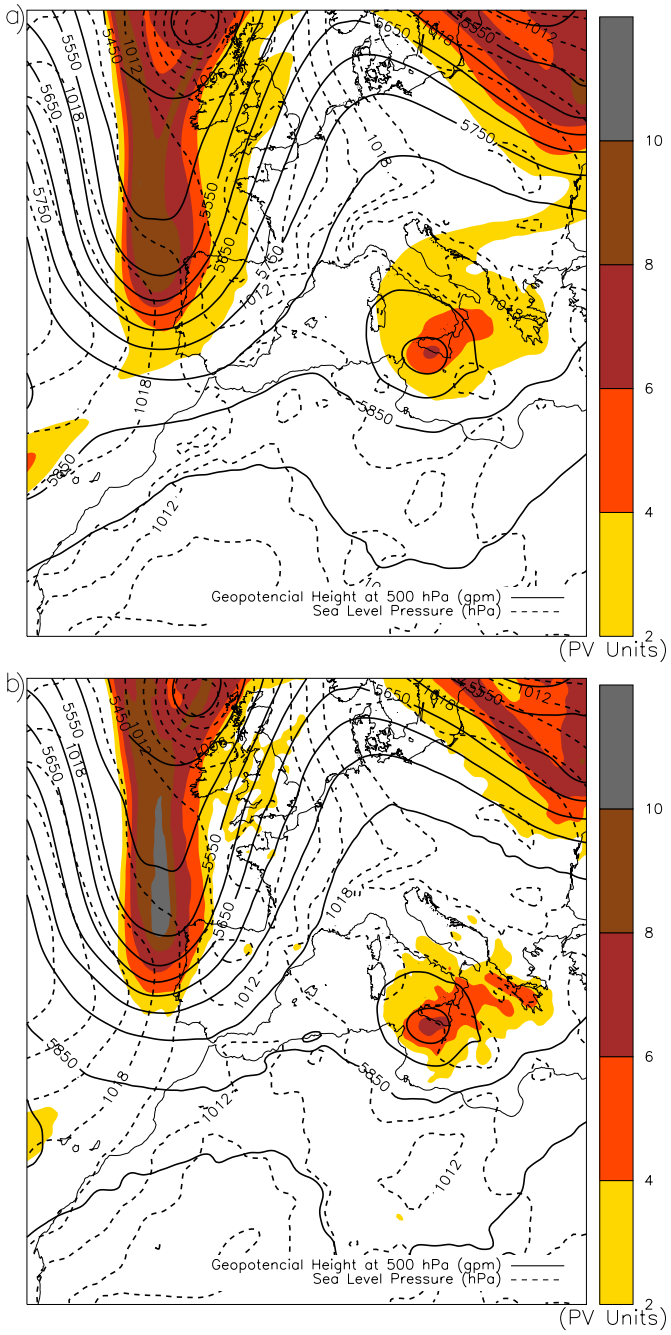


Figure 6.6. *Synoptic situation for 9 June 2000 at 00 UTC. a) Non-perturbed initial conditions (ECMWF 24 h forecast) and b) a PV-gradient ensemble member perturbed initial conditions. Geopotential height (continuous line, in gpm) at 500 hPa, sea level pressure (dashed line, in hPa), and Potential Vorticity on the 330 K surface (color contours, in PV units).*

The PV inversion technique is applied to the PV perturbed field obtained in the previous step and to the non-perturbed PV field. The difference between these two balanced fields is added to the original non-perturbed mass and wind fields

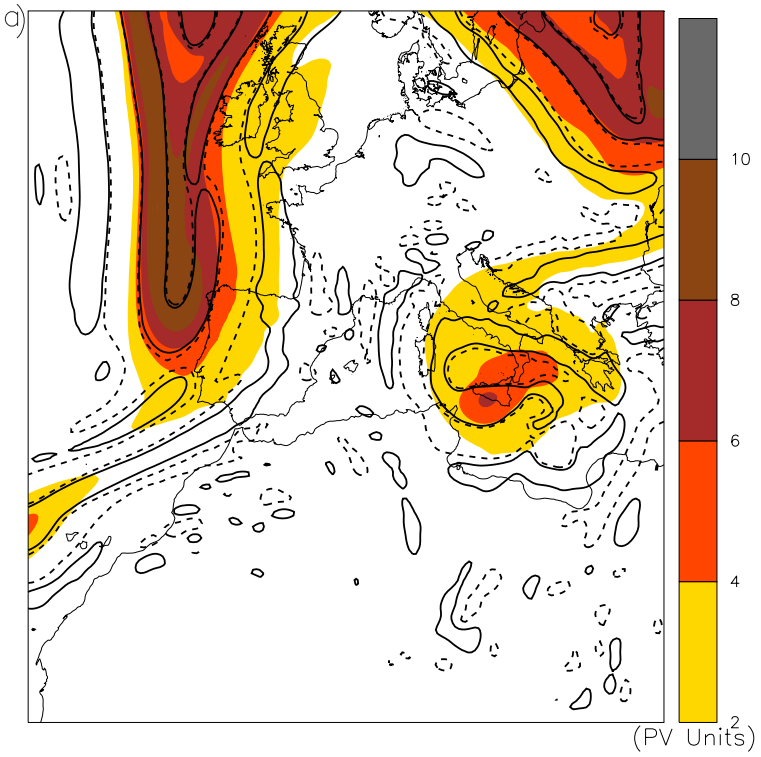


Figure 6.7. *PV-gradient EPS horizontal section of perturbation volumes at 300 hPa, solid line positive value (guidance field values over the threshold) and dashed line negative value (guidance field values below the negative of the threshold). As color contours in PV units Potential Vorticity field on 330 K surface.*

defining then the initial and boundary conditions of the corresponding ensemble member. Fig. 6.6 shows the visible differences on the PV, geopotential and sea level pressure fields between the non-perturbed and the PV-gradient perturbed initial conditions. The MM5 model is initialized and forced with each ensemble member fields and used to run the 54 h forecasts.

Chapter 7

Stage 2: Comparison against the multiphysics approach*

In this Chapter, the building procedure of our multiphysics ensemble will be detailed as it was done in the previous chapter with the PV-gradient ensemble. Afterwards, we will proceed to compare them. It is worth to keep in mind that each ensemble accounts for a different source of forecast uncertainty: model deficiencies on the multiphysics EPS and initial conditions uncertainties on the PV-gradient ensemble.

7.1 Multiphysics ensemble

The multiphysics ensemble is generated using a variety of physical parameterization schemes available in MM5 which are adequate for mesoscale simulations of precipitation systems. Specifically, the multiphysics sets are the result of combining three explicit moisture schemes (Goddard microphysics, Reisner graupel and Schultz microphysics), two cumulus parameterizations (Grell and Kain-Fritsch), two PBL schemes (Eta and MRF) and the cloud-radiation scheme and the five-layer soil model, plus the set used in the operational model run by our group, hereafter the control run (the explicit moisture scheme Reisner graupel, the cumulus

*This Chapter is also based on the published paper Vich, M., R. Romero, and H. E. Brooks, 2011: *Ensemble prediction of Mediterranean high-impact events using potential vorticity perturbations. Part I: Comparison against the multiphysics approach.* Atmos. Res., 102, 227-241. doi:10.1016/j.atmosres. 2011.07.017.

parameterization Kain-Fritsch 2, the PBL scheme MRF and the cloud-radiation scheme and the five-layer soil model). Therefore the ensemble consists of 13 members, exceeding the threshold given by Du et al. (1997) which indicates that 90% of the improvement due to ensemble averaging is obtainable with ensembles sizes as small as 8 to 10 members.

In order to be consistent when comparing both EPSs, both the PV-gradient and the multiphysics ensemble use the MM5 to run the forecast, have the same simulation domain, verification dataset and are both made up of 13 members.

7.2 Comparison: Verification results

Before starting the evaluation of both EPSs through a verification procedure, we return to the 9-10 June 2000 event used to illustrate the methodology in Section 6.4. Fig. 7.1 shows the 30-54 h forecast accumulated rainfall mean and standard deviation (STD) of each 13-member ensemble, i.e., multiphysics and PV-gradient. The multiphysics ensemble produces the highest rainfall values more northeastwards than the PV-gradient EPS and a more meridionally elongated pattern across France, while both STD fields present relative maxima over the highest rainfall values. A more extensive examination of the ensembles performance is done in the following.

The evaluation of the ensembles is done over the collection of 19 MEDEX presented in Section 6.1. This trial set is representative of the kind of events that this study aims: cyclones producing floods and strong winds over the western Mediterranean. The heavy rain associated with this kind of events awakes the interest of the public so, eventhough the precipitation field is hard to predict and verify due to its complex nature, it is the subject of our verification.

The precipitation, is characterized by highly localized gradients, meaning that there is a large fraction of zones and times where the precipitation is zero. As mentioned in Section 4.2 of the Introduction, this fact increases the difficulty of matching the forecast and the verifying data as the forecasts provide regularly spaced grid fields while the verifying observations usually are in irregular spaced networks of stations. Even though there are several spatial verification methods, like object-based verification, devoted to surpass these issues they unfortunately lean on a continuous gridded observation network that is not always available.

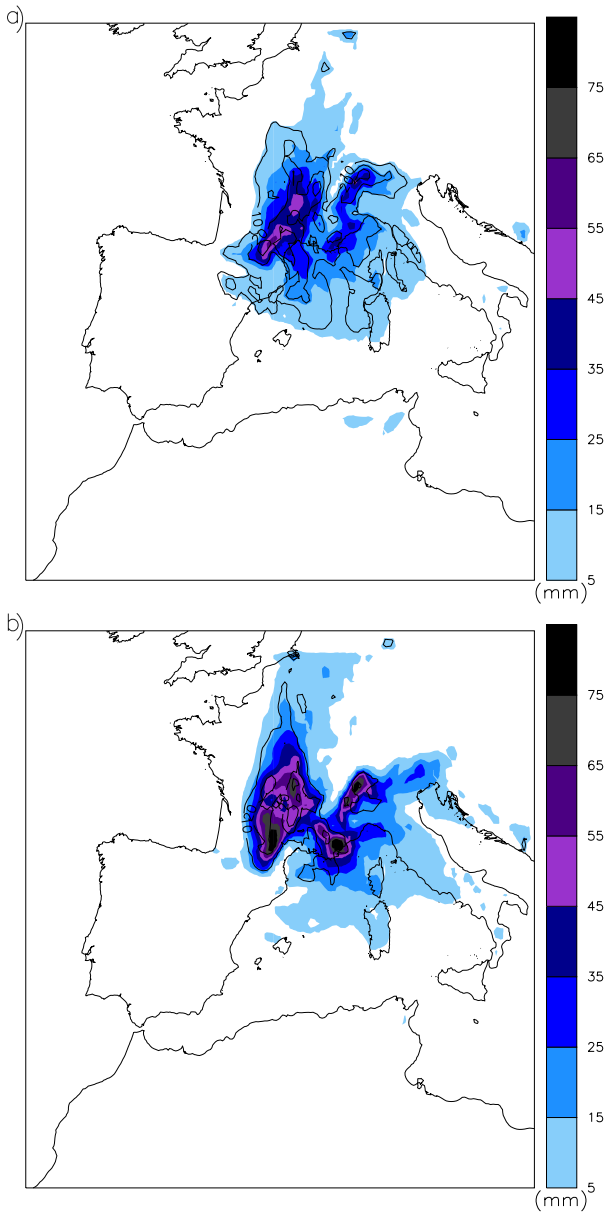


Figure 7.1. *Ensemble mean (color contours, in mm) and ensemble standard deviation (solid line, in mm at 10 mm intervals) for the 24 h accumulated precipitation over the forecast region (from 10 to 11 June 2000 at 06 UTC). a) PV-gradient ensemble and b) multiphysics EPS.*

So, since point-to-point verification methods can operate with an irregular spaced observational network, like the one available for this study (an irregular raingauge network) and the spatial verification cannot, we will base our verification process on them. Despite of leaning on traditional point-to-point verification scores we are very aware of their limitations, like for example the double penalty that penalizes the forecast even when the shape and pattern of the precipitation field is accurate

but is displaced in space.

Before evaluating each EPS as a whole using probabilistic scores, the performance of the ensemble members is assessed in order to assure that all of them are good enough to be included within the ensemble. Since this study is not focused on verifying a single observation threshold but on evaluating the general performance of the ensembles, the definition of the observed event is not fixed. Therefore several rainfall amount thresholds have been defined as observed events. All verification results showed here correspond to the second day of simulation (30-54h simulation time).

A brief description of each verification measure used next can be found in Chapter 4. The ROC area results, Fig. 7.2, show that the multiphysics EPS members are better at discriminating between events and non-events than the PV-gradient ones for all thresholds, nevertheless both EPSs present ROC areas above 0.76, a very satisfying result. The results for the Frequency Bias (Fig. 7.3) show that all members of both ensembles overpredict (Bias > 1) rainfall amounts less than 5 mm while underpredict (Bias < 1) the larger rainfalls amounts. At approximately 5 mm threshold, the transitional region between over and underprediction, both EPS are almost at the perfect score (Bias $= 1$). The Bias fast decay towards zero can be due to a sample problem, since the base rate shows a rapid decrease of the number of event samples for extreme precipitation values (Fig. 7.3). The Taylor diagrams (Fig. 7.4) show similar results for both EPSs, approximately all ensemble members present a root-mean-square (RMS) difference of 15 mm and a correlation coefficient of 0.5, while the forecast and observation standard deviations are approximately of 10 and 16 mm, respectively. Taking into account that the RMS errors, correlation coefficients and standard deviations are sensitive to steep gradients, noise and outliers, the obtained results, for the rainfall field, show a reasonable skill.

The fact that all members of the EPSs have a reasonable skill is a positive feature since they make a good basis to build a valuable ensemble system. To verify the EPSs as a whole, the forecast probability of an event is calculated as the fraction of the EPS members predicting the event among all members. With the purpose of establishing a baseline and corroborate the advantages of using a probabilistic forecast against a deterministic forecast, the performance of both types of forecasts is compared, as well as each EPS against its control run. The chosen approach to define the deterministic forecast is to consider a one-member ensemble made of

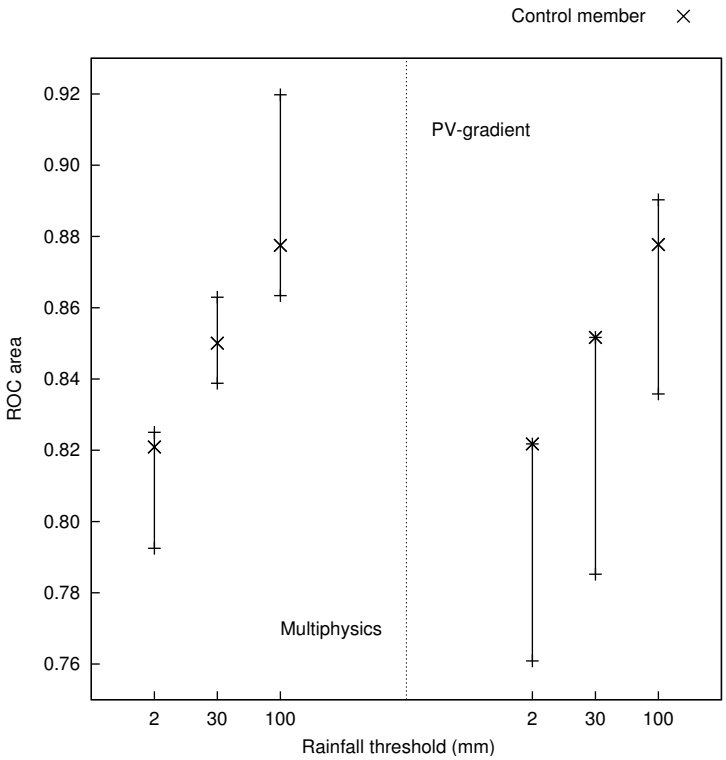


Figure 7.2. *ROC area range for the multiphysics and PV-gradient ensemble members, as functions of different rainfall event thresholds. The vertical lines represent these ROC area ranges where the extremes, maximum and minimum, values of the range correspond to the ensemble member associated with the highest/lowest value of ROC area.*

the control member. The control member corresponds to the non-perturbed run using the physical parameterizations of our operational model, a member present on both ensembles by construction.

The ensembles ROC areas results (Fig. 7.5) are very encouraging according to Buizza et al. (1999) and Stensrud and Yussouf (2007) who stated that forecasting systems with ROC areas greater than 0.7 are useful in discriminating between events and non-events. In fact, our area values lie over 0.8 for rainfall thresholds below 10 mm and over 0.7 below 50 mm. The low ROC area values above 50 mm are largely insignificant due to the following issues: first, a sample problem associated with the small number of observed events, already noted for the individual ensemble members; and second, the difficulty of the EPS in predicting extreme rainfall values due to the 22.5 km grid length. The PV-gradient ensemble shows better results than the multiphysics in almost all thresholds while both EPSs exhibit significantly better results than the control member, as expected.

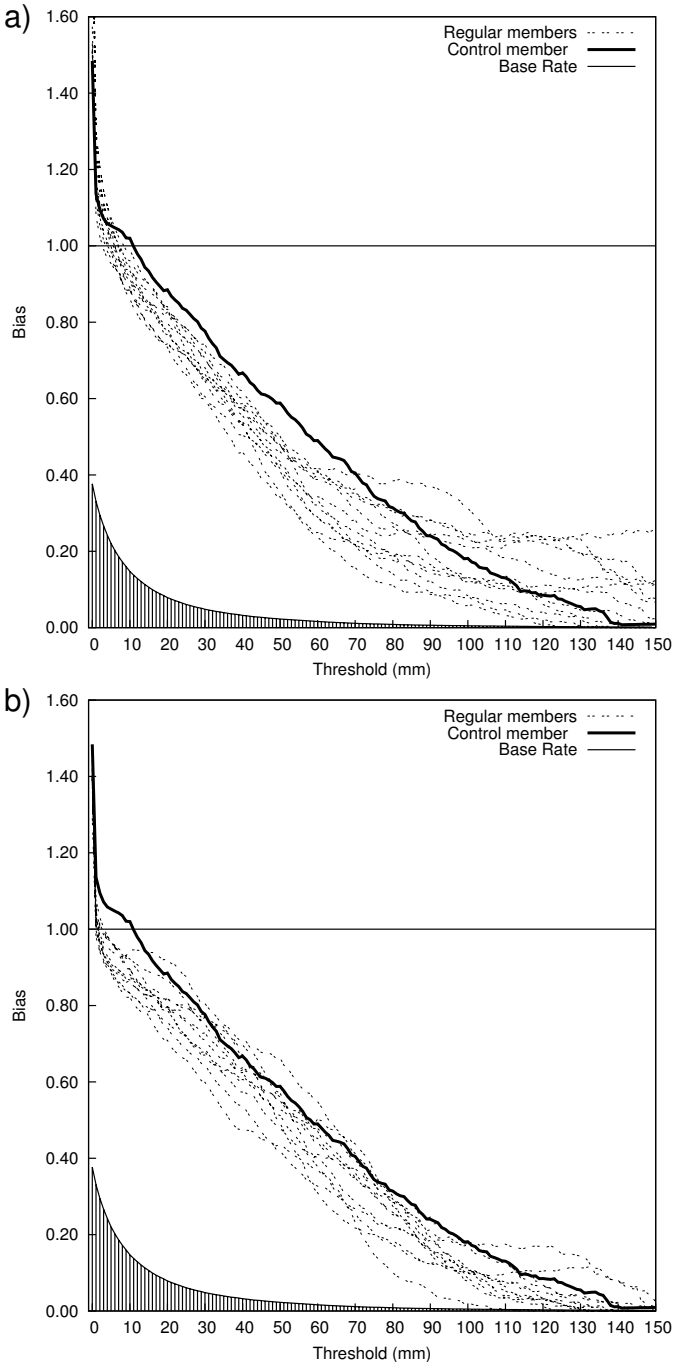


Figure 7.3. *Bias and base rate of the events, for a) multiphysics and b) PV-gradient EPS.*

The Bias results (Fig. 7.6) reveal a similar behavior to the EPS individual members, in the sense that both EPSs overpredict small rainfall amounts while underpredict large amounts. The highest score for both EPSs lies between 2 and 10 mm rainfall values. Like the ROC area, the Bias shows a slightly better performance

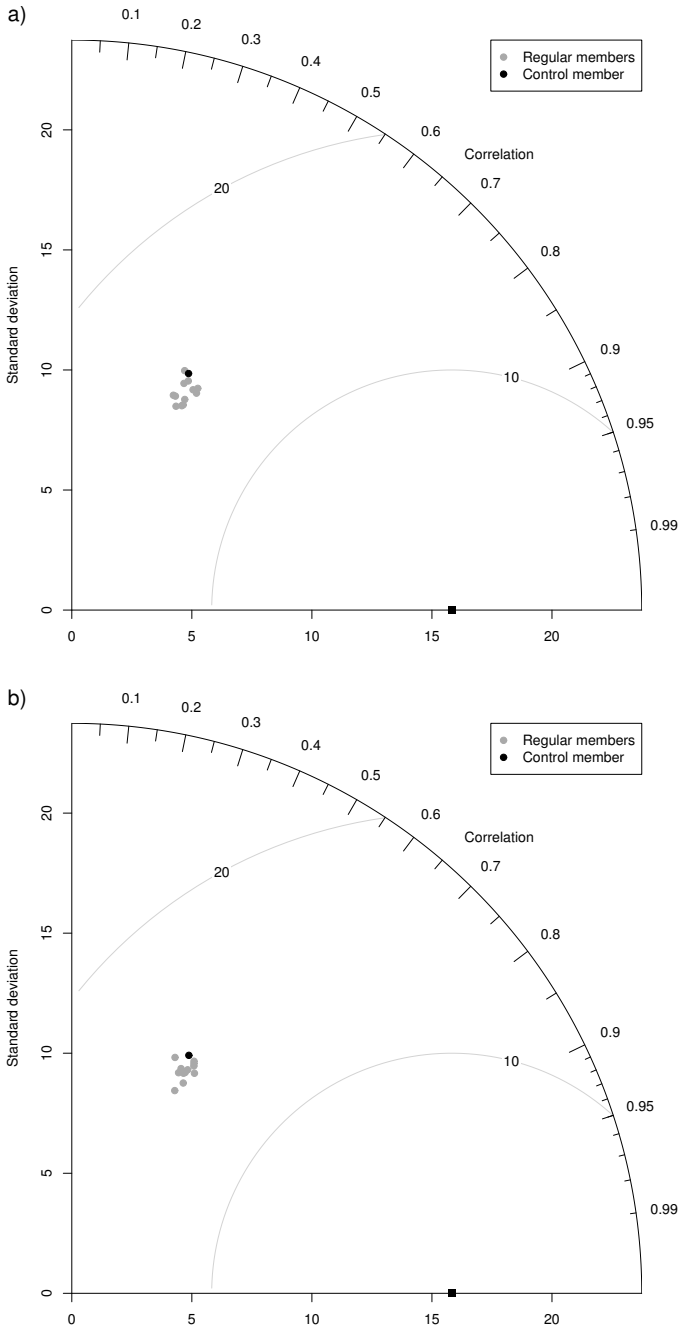


Figure 7.4. *Taylor diagrams for a) multiphysics and b) PV-gradient EPS. The black square represents the observed field and the dots the forecast fields. The perfect score is obtained when the data point representing the forecast field matches up with the observed one.*

of the PV-gradient ensemble than the multiphysics system and a fast decay in skill for rainfall values above 50 mm. The verification procedure, from now on, is focused on the thresholds ranging from 0 to 50 mm since out of this range statis-

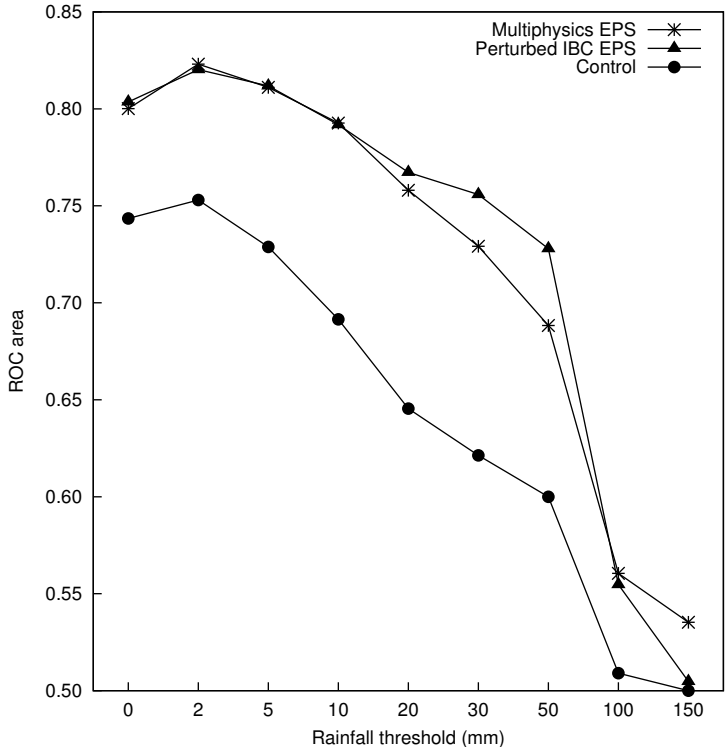


Figure 7.5. *ROC area for the multiphysics and PV-gradient ensembles and the control one-member ensemble, as functions of different rainfall event thresholds.*

tical significance is low and the difficulties for quantitative precipitation forecasts increase.

The Brier Skill Score (BSS) indicates the relative skill of the probability forecast over that of climatology, in terms of predicting whether or not an event occurred. Fig. 7.7 shows that the BSS is indistinguishable for both EPSs and the control run, while the BS terms show different behaviors depending on the EPS and the control. Both EPSs and control run present good skill for small rainfall thresholds that decreases as the rainfall threshold increases. The BS uncertainty term does not depend on the forecast but on the observations uncertainties so it is the same for both EPSs and the control run. The results obtained on the BSS are explained through the BS reliability and BS resolution terms. These two terms are oppositely oriented and they almost compensate each other in the total score. Taking this into account, the multiphysics ensemble shows better performance than the PV-gradient for the resolution term, while both ensembles show better performance over the control run for the BS resolution term, and viceversa for the BS reliability term.

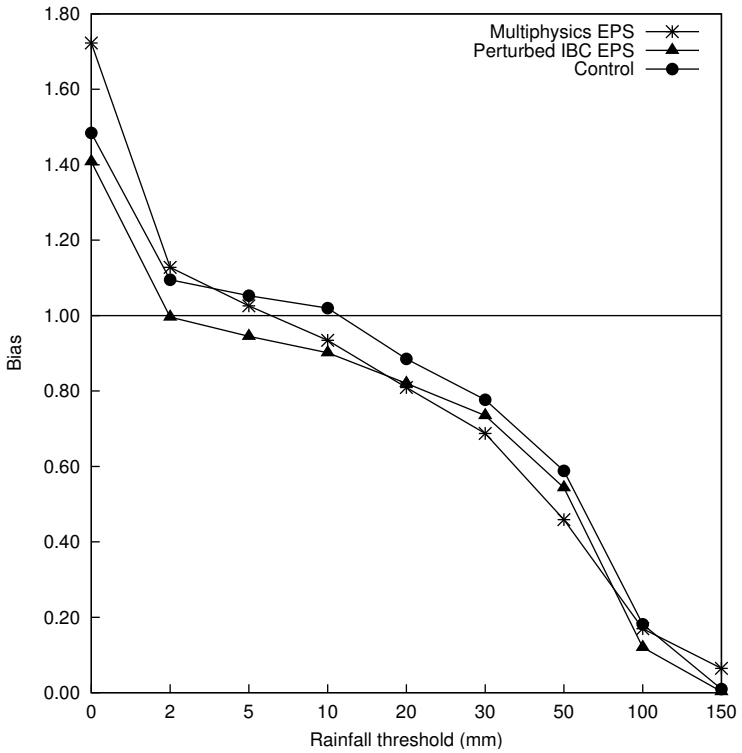


Figure 7.6. *Bias for the multiphysics and PV-gradient ensemble and their respective control one-member ensemble, as functions of different rainfall event thresholds.*

The attribute diagram plots the observed frequency against the forecast probability including the no-resolution and no-skill line. Fig. 7.8 shows the results for rainfall thresholds of 2 and 30 mm which states how the skill decreases as the threshold increases for both EPSs and the control run. Nevertheless the curves do not migrate out of the regions of skill. Although both EPSs lie within the skill zone, the PV-gradient ensemble presents better skill than the multiphysics EPS in almost all regions. The control run is clearly not skillful at producing probabilistic forecasts. Examining now the information of the BS decomposition terms given the attribute diagram, Fig. 7.8.a shows a better reliability and resolution for high forecast probability (over 0.4) than for low forecast probability (below 0.4) for both EPSs, meaning that they underforecast lower probability thresholds and overforecast higher probability thresholds (conditional bias). A conditional bias was already revealed by the Bias results (Fig. 7.6). From moderate rainfall thresholds (Fig. 7.8.b) the interpretation is more complex and highly dependent on the probability region.

The rank histograms indicate how well the ensemble spread represents the true

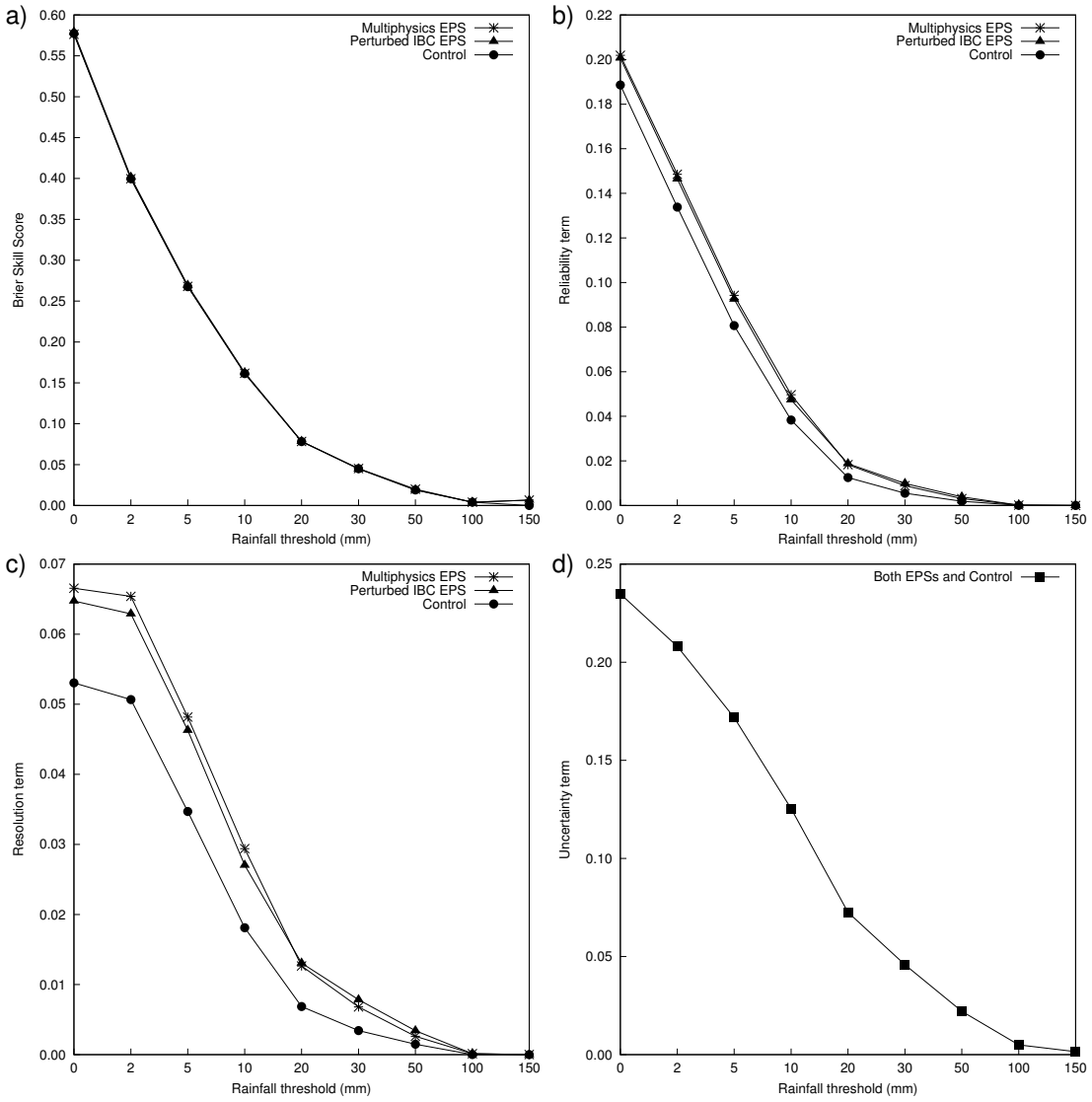


Figure 7.7. *Brier Skill Score and the three terms of Brier Score for the multiphysics and PV-gradient ensembles and the control one-member ensemble, as function of different rainfall event thresholds. a) Brier Skill Score, b) Brier Score Reliability term, c) Brier Score Resolution, and d) Brier Score Uncertainty.*

variability or uncertainty of the observations. The rank histograms for rainfall thresholds of 2 and 30 mm (Fig. 7.9) present an U-shaped profile which indicates that the spread in both EPSs is too small, as most of the observations fall outside the extremes of the ensemble. The 30 mm threshold presents a form between an U-shape combined with a strong right-asymmetry meaning that the EPSs results fall outside the ensemble extreme with a negative bias. The PV-gradient ensemble presents slightly better rank histograms than the multiphysics EPS.

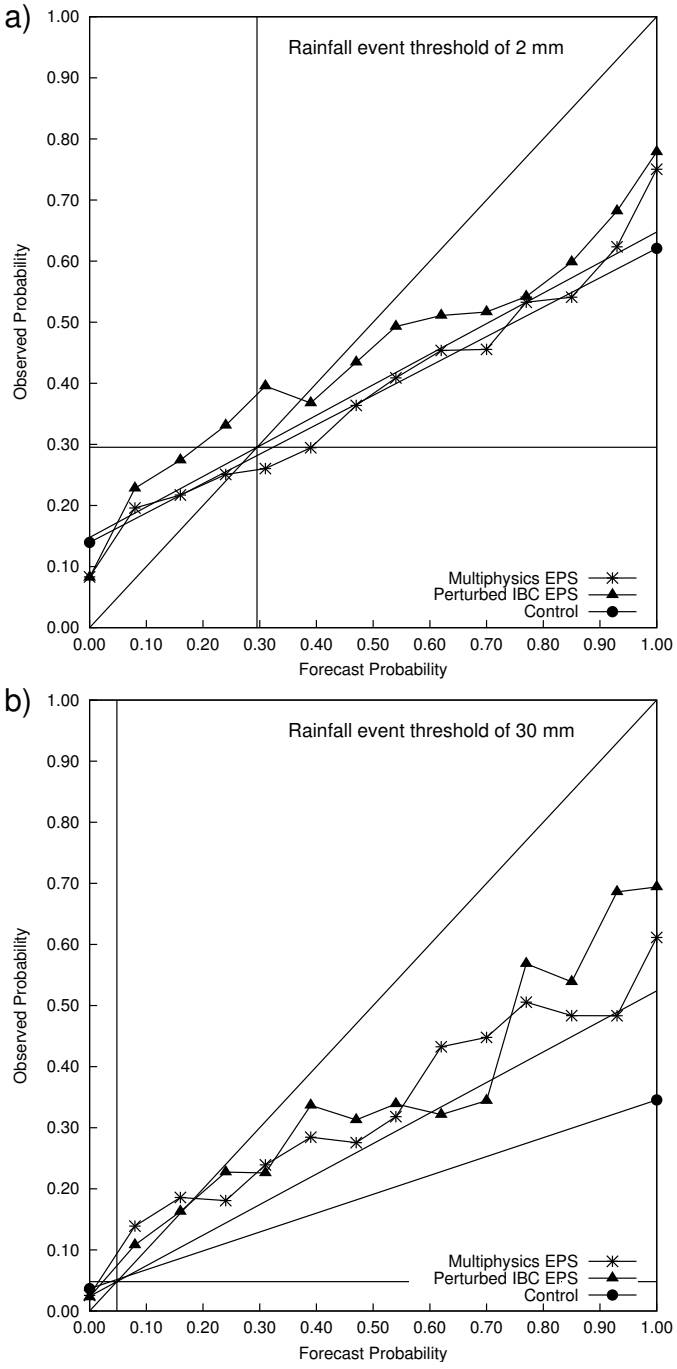


Figure 7.8. *Attribute diagrams for the multiphysics and PV-gradient ensemble and the control one-member ensemble, for a) 2 mm and b) 30 mm rainfall event thresholds.*

As a complement to the rank histograms, the ensemble mean error and spread of both ensembles were calculated (Jolliffe and Stephenson, 2003). These values are, respectively, 5.21 and 2.44 mm for the multiphysics ensemble and 5.62 and 3.46

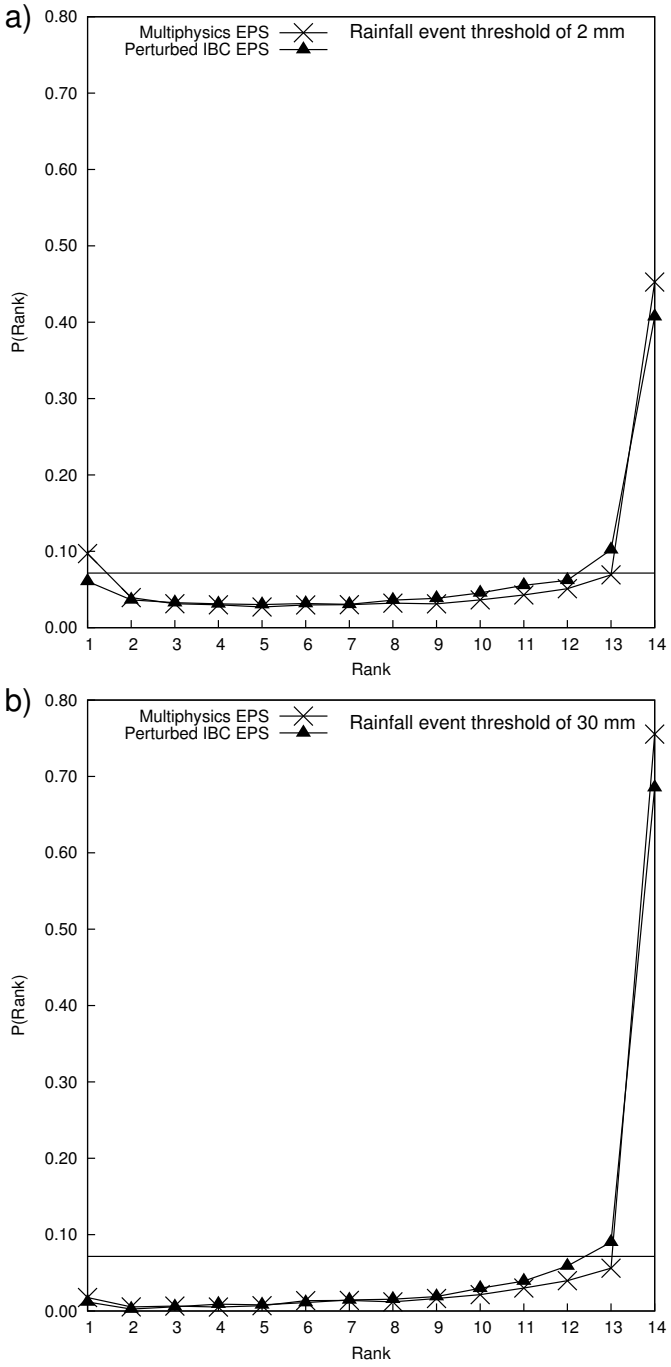


Figure 7.9. *Rank Histograms for the multiphysics and PV-gradient ensemble, for a) 2 mm and b) 30 mm rainfall event thresholds.*

mm for the PV-gradient, confirming that the lack of spread is more accentuated in the multiphysics EPS.

7.3 Conclusions

The designed multiphysics and perturbed initial and boundary conditions ensemble prediction systems have proved to be a good strategy to improve the short-range numerical forecasts of heavy rain events associated with western Mediterranean cyclones. The methodology developed for building the PV-gradient ensemble appears to be a promising tool. On the one hand, the use of a single variable (potential vorticity) on which to define perturbations, combined with the PV inversion technique, keeps the method simple while ensures modifications of all the meteorological fields without compromising the mass-wind balance. On the other hand, the verification results prove the usefulness of the method to yield skillful probabilistic prediction of Mediterranean heavy precipitation events.

The verification procedure highlights the difficulties of evaluating the rainfall field because it has strong gradients in space, rapid variations in time, and is observed over irregularly spaced networks, as well as the sample problems associated with extreme rainfall values and difficulties of the model to forecast extreme precipitations. In spite of these difficulties the verification ratifies the advantages of an EPS over a deterministic forecast. Moreover, the verification results also point out that the PV-gradient EPS performs slightly better than the multiphysics ensemble.

Chapter 8

Stage 3: Adjoint-derived sensitivity zones*

Encouraged by the better performance of the PV-gradient EPS over the more traditional multiphysics approach, a new ensemble is developed and tested in this Chapter. This new ensemble introduces the perturbations over the MM5 adjoint model calculated sensitivity zones (hereafter PV-adjoint), while the PV-gradient ensemble perturbs the PV field along the zones of the three-dimensional PV structure presenting the local most intense values and gradients of the field. An overview on adjoint models can be found in Section 2.2 of the Introduction.

Both PV-perturbed ensembles (PV-gradient and PV-adjoint) introduce perturbations into the initial and boundary potential vorticity field and propagate them to the temperature and wind fields using the PV inversion algorithm. Perturbing the boundary conditions prevents the ensemble from losing variance as lead time increases (Nutter et al., 2004a,b). The sensitivity of cyclones and associated high-impact weather to PV perturbations has been established in several studies (e.g. Huo et al. 1999; Romero 2008; Argence et al. 2008 and Vich et al. 2011a). Perturbations in the initial conditions are shown to be crucial for the accurate simulation of severe convective events over the western Mediterranean (Cohuet et al., 2011).

*This Chapter is based on the published paper Vich, M., R. Romero, and V. Homar, 2011: *Ensemble prediction of Mediterranean high-impact events using potential vorticity perturbations. Part II: Adjoint- derived sensitivity zones.* Atmos. Res., 102, 311–319. doi:10.1016/j.atmosres.2011.07.016.

The PV-adjoint EPS takes advantage of the main application of an adjoint model: sensitivity analysis, that determines the sensitivity of a particular forecast feature of interest to the initial condition. The tangent linear approximation made on the adjoint formulation can be affected by the timespan of the adjoint run. The longer the evolution analyzed, the farther away from a linear evolution the perturbations evolve in the nonlinear model. As established in Section 2.2, this linear assumption is valid for smooth integrated response functions defined at lead times up to 48 h, while decreasing to 24 h when diabatic processes significantly affect the response function. Moreover, the response function definition is also critical since forecast features like rain (our final target) are nonlinear, i.e. they may invalidate the tangent linear approximation. In a study like ours, focused on the rainfall associated with intense cyclones, a response function involving a precursor larger-scale dynamical feature like the intensity of the cyclone (the vertical component of the relative vorticity near the surface) allows to circumvent the mentioned limitation in the response function definition without compromising the objective of the adjoint calculations.

Following Homar and Stensrud (2008) that classifies a sensitivity estimation as *objective* if it is based on the tangent linear and adjoint models, and *subjective* if it is based on human interpretation of the atmospheric fields and the links between the chosen forecast aspect and the initial structures as derived from the conceptual models, each ensemble can be tagged accordingly, the PV-gradient as subjective and the PV-adjoint as objective. Therefore, in other words, this study aims to explore the advantages or disadvantages of using an objective method like MM5 adjoint model instead of the conceptual model that links the Mediterranean cyclogenesis to an upper-level precursor PV anomaly, the basis of the PV-gradient ensemble.

8.1 Ensemble prediction system design

The building of the PV-adjoint ensemble is analogous to the construction of the PV-gradient EPS, except for the criterion used to create the PV perturbations. Both ensembles are made up of 13 members (12 perturbed members plus a non-perturbed) using the same MM5 configuration applied on the previous PV-gradient ensemble.

The approach proposed on both PV-perturbed ensembles exploit the well-known connection between PV structures and our target, high impact weather in a limited area associated with cyclones, instead of using a more standard technique. For example, Beare et al. (2003) stated that cyclogenesis is sensitive to PV structures at different spatial scales, Snyder et al. (2003) found a strong relation between the reference state of the PV field and the evolution of PV perturbations, while Plu and Arbogast (2005) pointed the likelihood of replicating the system variability modifying the PV coherent structures present in the analyses. But, while both PV-gradient ensembles are based on the idea of perturbing the initial and boundary conditions through the PV field, they differ in the criteria that locate the perturbed zones. The PV-gradient adopts the zones of most intense values and gradients of the PV field as guidance, the most sensitive zones (Garcies and Homar, 2009; Romero et al., 2006) of the subsequent atmospheric evolution, like the cyclogenesis process that occurs over the western Mediterranean. On the other hand, the PV-adjoint uses the PV sensitivity field calculated with the MM5 adjoint model, assuming these objectively-obtained zones are the most sensitive areas of the later cyclonic evolution.

The MM5 adjoint sensitivity field has already proved its value in computing sensitivity areas of intense Mediterranean cyclones, as several previous studies sustain (e.g. Homar and Stensrud 2004; Homar et al. 2006; Homar and Stensrud 2008, among others). In this study, the MM5 adjoint simulation timespan is 24 h, which assures that the tangent linear assumption made by the adjoint run is valid. The linear assumption can also be hampered if nonlinear processes are dominant, so the explicit nonlinear processes present in a high resolution adjoint run may lead to unreliable sensitivity results (Homar and Stensrud, 2004). A 90 km horizontal resolution for the adjoint run resolves well the dynamical features of interest (MEDEX cyclones) and does not diminish the physical capabilities of the adjoint while also keeping the computational cost low. Therefore, the simulation domain is defined as a 90 km resolution horizontal grid mesh with 86x120 nodes, centered at 41° latitude and 3° longitude. The vertical grid mesh is defined by 30 sigma levels. As previously explained, our response function is the vertical component of the relative vorticity near the surface, strongly related to the associated mid-to large-scale low-pressure system, our target. Specifically, the response function is defined as the surface vorticity averaged over the 1530x1530 km^2 square centered on the minimum sea level pressure grid point of the corresponding ECMWF 24 h forecast

cyclone. The size of the response function is set as a balance between the precise target of the cyclone center (aiming at capturing the factors influencing its intense deepening) and the stabilizing effect on the sensitivity results of widening the area to a larger extent based on our experience on both Mediterranean intense cyclones and the definition of response functions in sensitivity experiments. Since the perturbations are introduced on the PV field, the sensitivity to the quasi-geostrophic PV is used as a proxy for the sensitivity to the full PV field. This sensitivity can be written as function of the sensitivity to the stream function which in turn can be formulated in terms of the sensitivity to the wind, both expressions obtained by means of a simple hand transposition strategy (Errico and Vukićević, 1992). The physical parameterizations of adjoint models are not as well resolved as in the standard model version since they are based on a linear version of the model equations, so in this study the adjoint configuration considers a dry atmosphere to assure the Courant-Friedrichs-Lewy condition[1] (CFL condition) with our available computational resources for all the needed MM5 adjoint runs, one run per day of study.

8.1.1 Practical implementation

As detailed in the PV-gradient building procedure, the guidance field is defined by the difference between the three-dimensional PV field and a highly smoothed version of itself, thus highlighting the most intense values and gradients. On the contrary, in the PV-adjoint procedure the guidance field corresponds to the PV sensitivity field obtained with the MM5 adjoint model. Once the guidance field is known, the following generation stages for both ensembles are the same. A threshold is then defined as the average of the guidance field over the whole domain, in absolute value. The three-dimensional regions where the perturbations are introduced are then defined by the 3D regions where the guidance field exceeds, in absolute value, the threshold. On each of these volumes the intensity -both magnitude and sign- and displacement perturbations are assigned randomly accordingly

[1] A condition in numerical equation solving which states that, given a space discretization, a time step bigger than some computable quantity should not be taken. The condition can be viewed as a sort of discrete "light cone" condition, namely that the time step must be kept small enough so that information has enough time to propagate through the space discretization. (Weisstein, Eric W. "Courant-Friedrichs-Lewy Condition." From MathWorld-A Wolfram Web Resource. `http://mathworld.wolfram.com/Courant-Friedrichs-LewyCondition.html`)

to the PV error climatology[2]. On the other hand, the displacement direction perturbation is assigned using the same randomly chosen value to all volumes to avoid discontinuities in the perturbed PV field. After perturbing, the difference between the original and perturbed balanced fields (obtained applying the PV inversion technique) is added to the non-perturbed, mass and wind fields, to produce the initial and boundary conditions of the corresponding ensemble member.

The different various steps of the perturbing procedure are illustrated again by means of the 9-10 June 2000 case study, like in the PV-gradient ensemble generation. Its synoptic situation on 9th June 2000 at 00 UTC is shown in Fig. 8.1.a.

Fig. 8.2 shows a horizontal section at 300 hPa of the perturbation volumes. The PV-gradient zones follow closely the structure of the upper-level trough as expected given its definition (Fig. 8.2.a), while the PV-adjoint zones highlight different regions including the upper-level trough region (Fig. 8.2.b).

Fig. 8.1 gathers the results of the perturbing technique showing the visible differences on the PV, geopotential and sea level pressure fields between the non-perturbed, the PV-gradient and PV-adjoint perturbed initial conditions. The MM5 model is initialized and forced with each perturbed initial condition and used to run the 54 h forecasts. Fig. 8.3 shows the 30-54 h forecast accumulated rainfall mean and standard deviation (STD) of each 13-member ensemble. The PV-gradient ensemble presents the highest rainfall values located farther south than in the PV-adjoint results and with a more meridionally elongated pattern across France, while both STD fields present relative maxima over the highest rainfall values. An extensive examination of the performance of each ensemble is done in the following section.

[2]PV error climatology derived on Section 6.2 assures that the perturbations are consistent with the PV field uncertainty range.

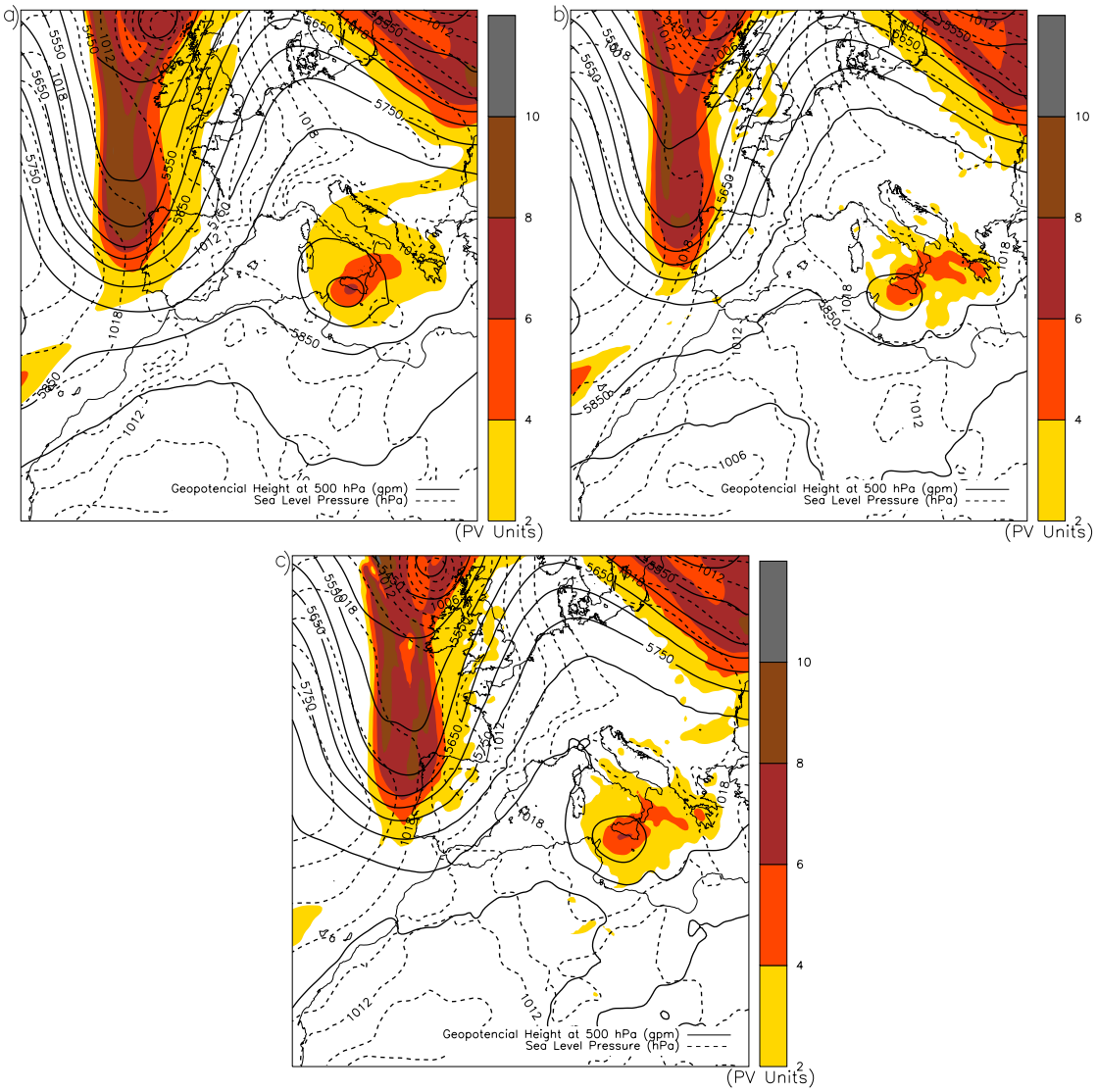


Figure 8.1. *Synoptic situation for 9 June 2000 at 00 UTC. a) Non-perturbed initial conditions (ECMWF 24 h forecast), b) PV-gradient and c) PV-adjoint ensemble member perturbed initial conditions for a randomly-chosen ensemble member. Note that the PV-gradient ensemble member is different than the one displayed in Fig.6.6, page 75. Geopotential height (continuous line, in gpm) at 500 hPa, sea level pressure (dashed line, in hPa), and Potential Vorticity on the 330 K surface (color contours, in PV units).*

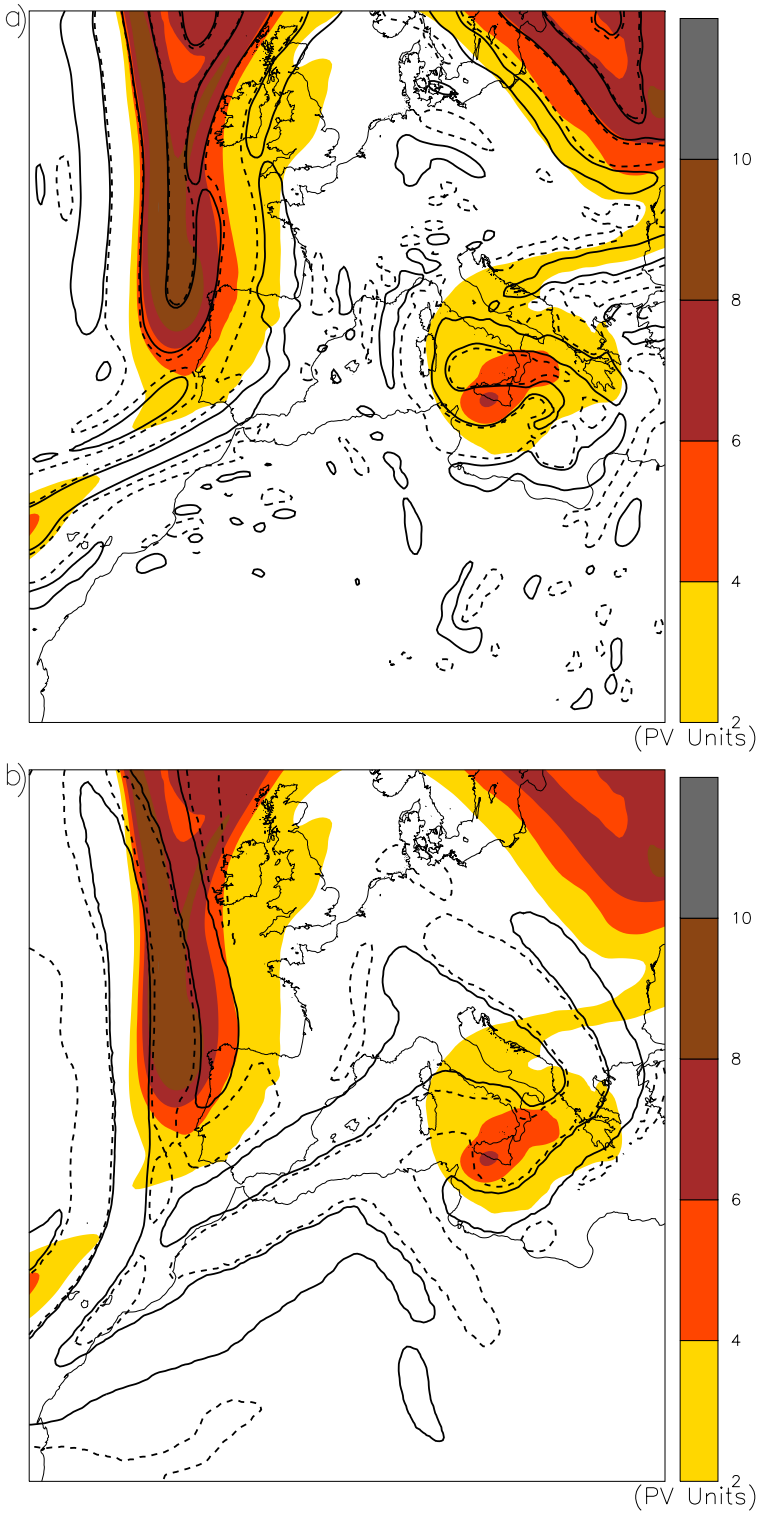


Figure 8.2. *a) PV-gradient EPS and b) PV-adjoint ensemble horizontal section of perturbation volumes at 300 hPa, solid line positive value (guidance field values over the threshold) and dashed line negative value (guidance field values below the negative of the threshold). Unperturbed PV field on 330 K surface, color contours in PV units.*

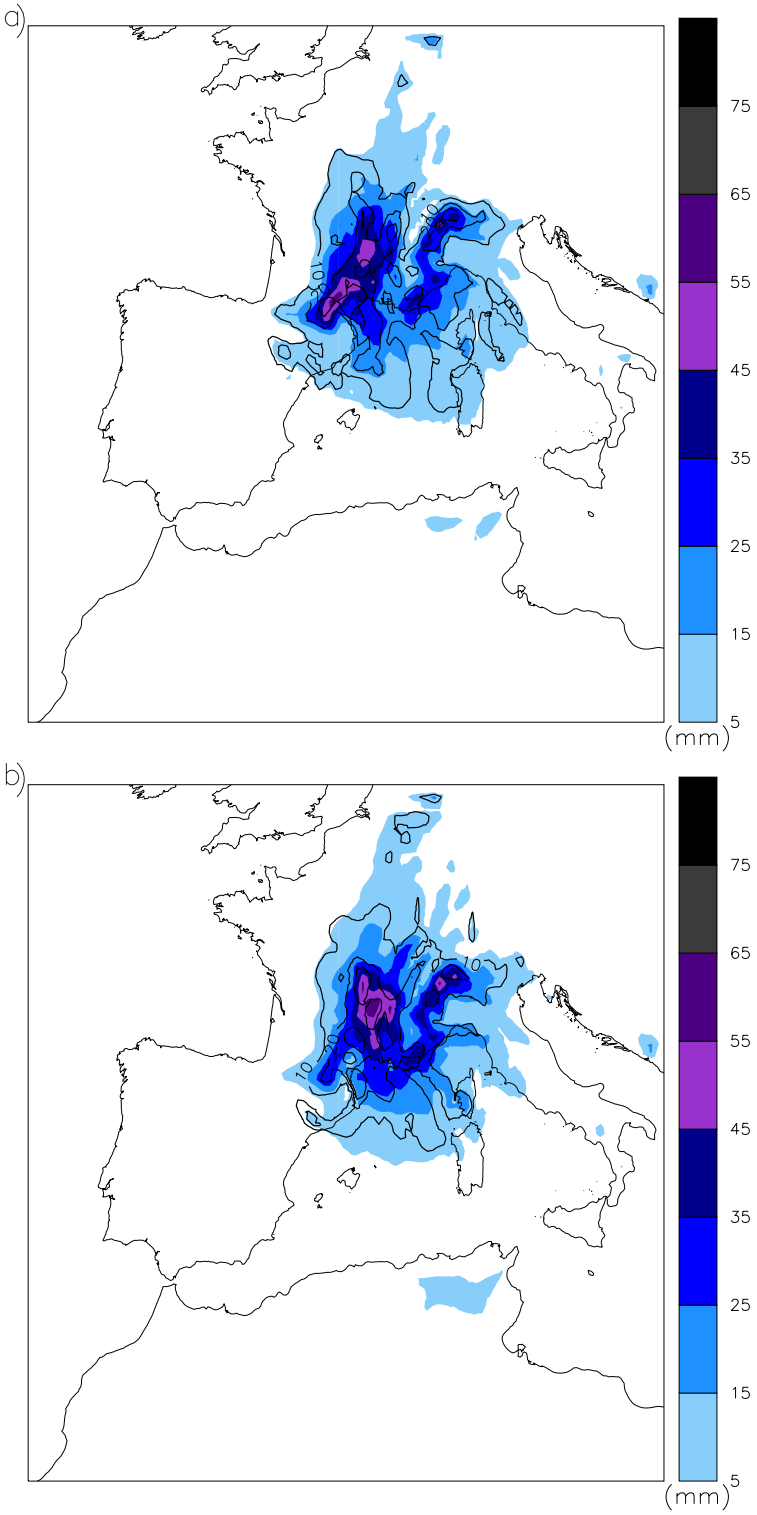


Figure 8.3. *Ensemble mean (color contours, in mm) and ensemble standard deviation (solid line, in mm at 10 mm intervals) for the 24 h accumulated precipitation over the forecast region (from 10 to 11 June 2000 at 06 UTC). a) PV-gradient ensemble and b) PV-adjoint EPS.*

8.2 Comparison: Verification results

The verification framework is like the one used in the multiphysics and PV-gradient ensemble comparison. As detailed in Chapter 6, the verification testbed consists of a 56-day period from the 19 MEDEX episodes collection between September 1996 and October 2002 and the 24 h accumulated precipitation (from 06 UTC to 06 UTC daily) provided by the AEMET climatological raingauge network.

An evaluation of the predictive skill of each EPS is done using probabilistic scores and indices, assuming each deterministic forecast (ensemble member) as an independent realization of the same underlying process. The verification is done for the 24 h accumulated precipitation period corresponding to the second day of simulation and addresses the general performance of the ensembles, not a unique observation threshold. Therefore, nine rainfall amount thresholds (0, 2, 5, 10, 20, 30, 50, 100 and 150 mm) have been defined as observed events.

The Bias score plotted on Fig. 8.4 shows a slightly better performance of the PV-gradient ensemble over the PV-adjoint system considering that both EPSs overpredict (Bias > 1) rainfall amounts less than 5 mm while underpredict (Bias < 1) the larger rainfall amounts. Between 2 and 10 mm thresholds both EPSs are almost unbiased. Like in the previous chapter results, the fast Bias decay towards zero for greater thresholds, above 50 mm, is most probably due to a sample problem as indicated by the rapid decrease in the number of events of extreme precipitation values (Base Rate in Fig. 8.4). It may also be due to the difficulties of the EPSs to forecast extreme precipitation values while running over a 22.5 km horizontal resolution domain. From now on our verification procedure focuses on the thresholds ranging from 0 to 50 mm due to the lack of statistical significance outside this range and the low capability of mesoscale models to handle extreme precipitation forecasts.

The ROC areas obtained for both ensembles (Fig. 8.5) are very skillful since all forecasts lie well above 0.7, the threshold established by Stensrud and Yussouf (2007) which indicates usefulness of a forecasting system on discriminating between events and non-events. Both EPSs exhibit good results for all rainfall thresholds presenting their maximum value at 2 mm. The PV-gradient ensemble shows better skill than the PV-adjoint for all rainfall thresholds, except for the 10 mm where their skill is similar.

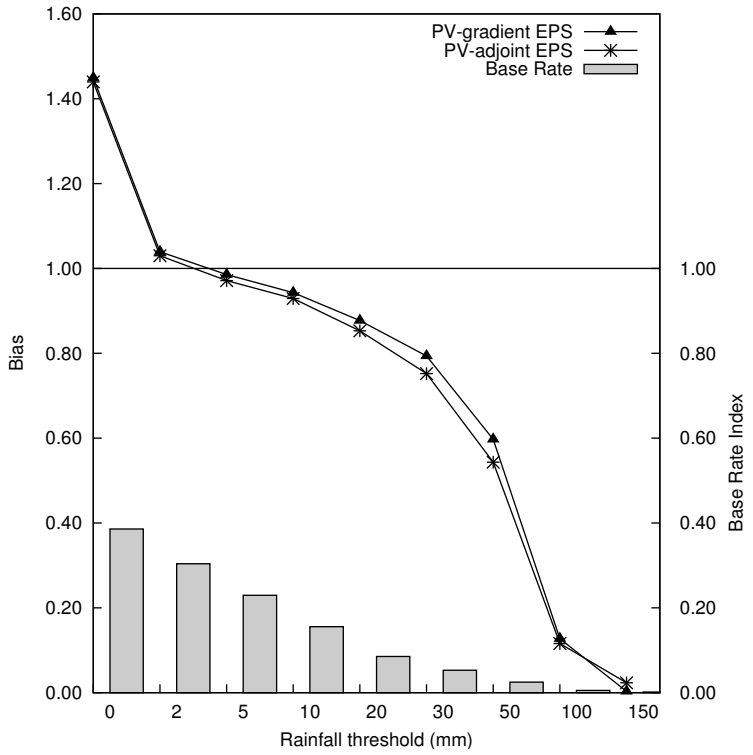


Figure 8.4. *Frequency Bias for the PV-gradient and PV-adjoint ensembles, as function of different rainfall thresholds. The Base Rate index shown by the histogram represents the observed event probability (the sample size is 109276).*

Fig. 8.6 shows that the Brier Skill Score is almost the same for both EPSs, while the BS terms show different behaviors depending on the EPS. Both EPSs present good skill for low rainfall thresholds that decreases as the rainfall threshold increases. As the BS uncertainty term exclusively depends on the observations uncertainties and not on the forecast, the results for all the EPSs are identical as they share the same observational database. The almost coincidence of both EPSs on the BSS are due to the BS reliability and BS resolution terms, two terms of opposite sign, that almost compensate each other in the total score. The BS reliability is almost the same for both ensembles (indistinguishable in the graph) indicating that both ensembles present very similar reliability skill, while the BS resolution term is slightly different between ensembles but not enough to show a meaningful difference in how the different forecast events are classified by the forecast system. Despite this, the PV-gradient ensemble slightly outperforms the PV-adjoint.

Fig. 8.7 gathers the attribute diagram results for the 2 and 30 mm rainfall thresholds and states how the skill of both ensembles decreases as the threshold increases.

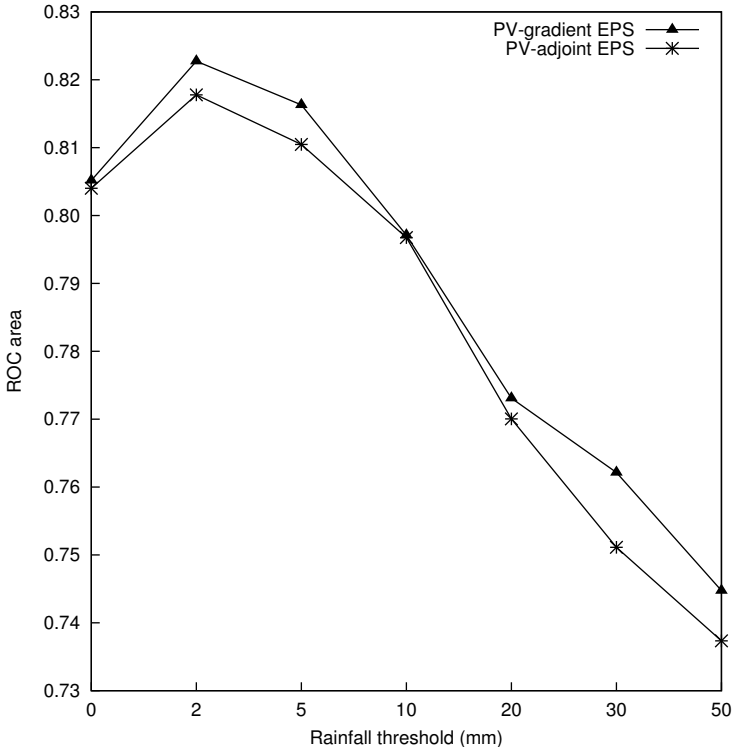


Figure 8.5. *ROC area for the PV-gradient and PV-adjoint ensembles, as function of different rainfall event thresholds.*

In spite of this skill decrease, the curves remain inside the skill region. Even though both ensembles lay inside the skill region, both ensembles exhibit a tendency to underforecast the Observed Probability for low values of Forecast Probability and overforecast it for high values, indicating a conditional bias. Nevertheless, the PV-gradient is closer globally to the perfect score than the PV-adjoint.

All rank histograms (Fig. 8.8) present a U-shaped form combined with a pronounced right-asymmetric profile due to an excessive population within the extremes rank, revealing that the ensembles clearly underestimate the higher precipitation values and slightly overestimate the lower precipitation values (an extended discussion on rank histogram interpretation can be found in Hamill (2001)). This behavior agrees with the Bias results (Fig. 8.4) which also show overprediction for lower thresholds and underprediction for higher. Once again, the PV-gradient ensemble shows slightly better skill than the PV-adjoint.

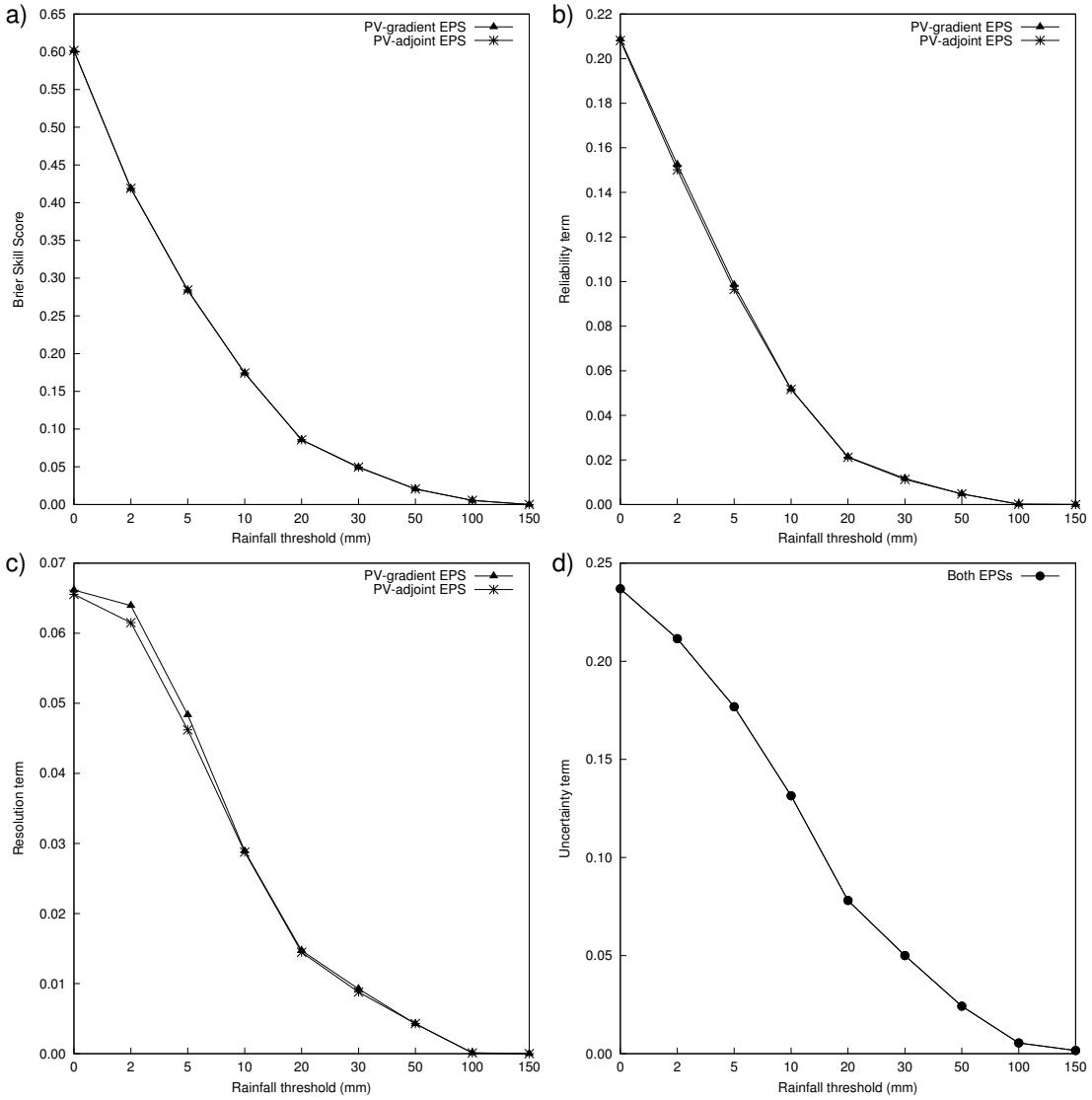


Figure 8.6. *Brier Skill Score and the Brier Score three components for the PV-gradient and PV-adjoint ensembles, as function of different rainfall event thresholds. a) Brier Skill Score, b) Brier Score Reliability term, c) Brier Score Resolution term, and d) Brier Score Uncertainty term (only depends on the observations).*

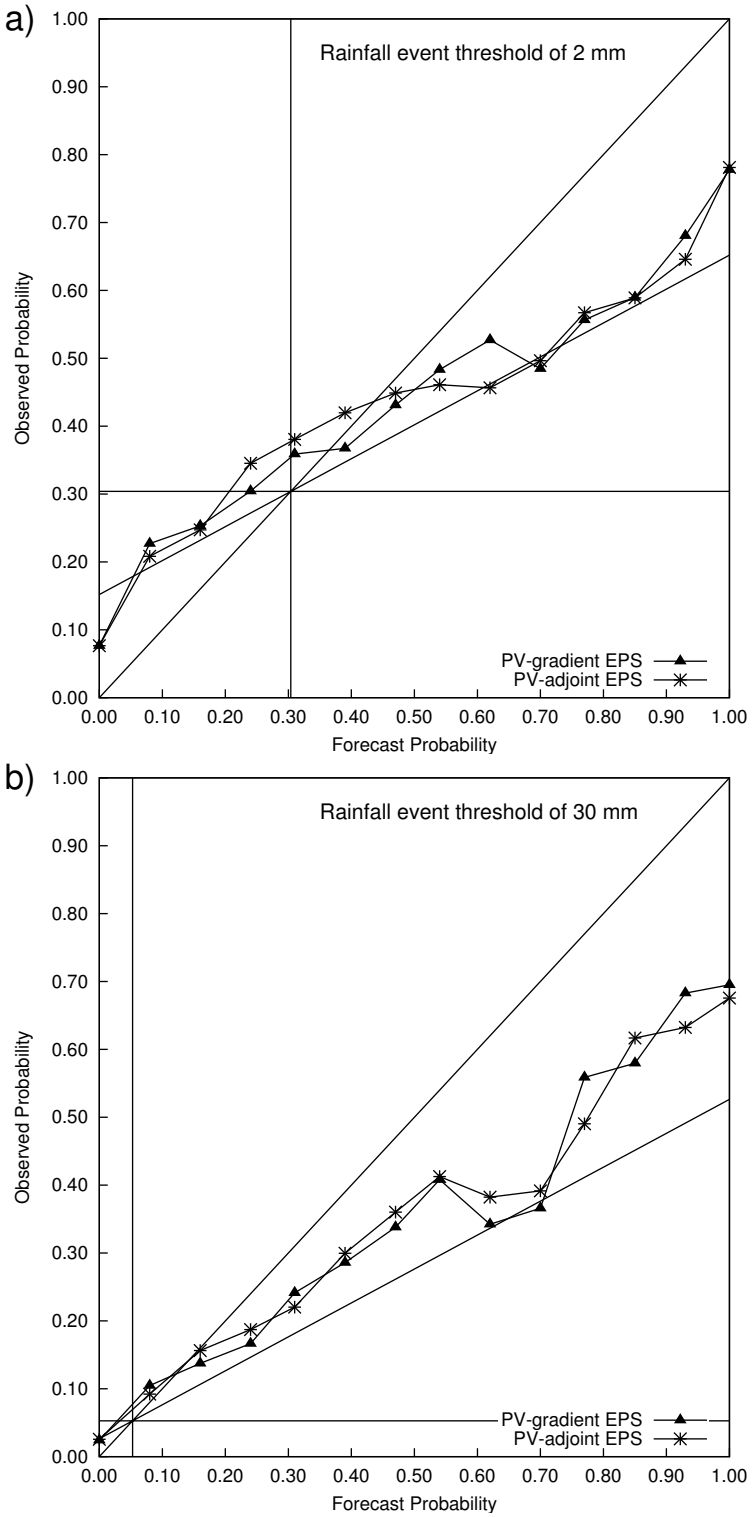


Figure 8.7. *Attribute diagrams for the PV-gradient and PV-adjoint ensembles, for a) 2 mm and b) 30 mm rainfall event thresholds. The perfect score is represented by a curve that matches the diagonal. The vertical and horizontal lines represent the no-resolution, also known as the climatological frequency of the event, and the 0.5 slope that crosses the no-resolution and perfect score lines represents the no-skill line.*

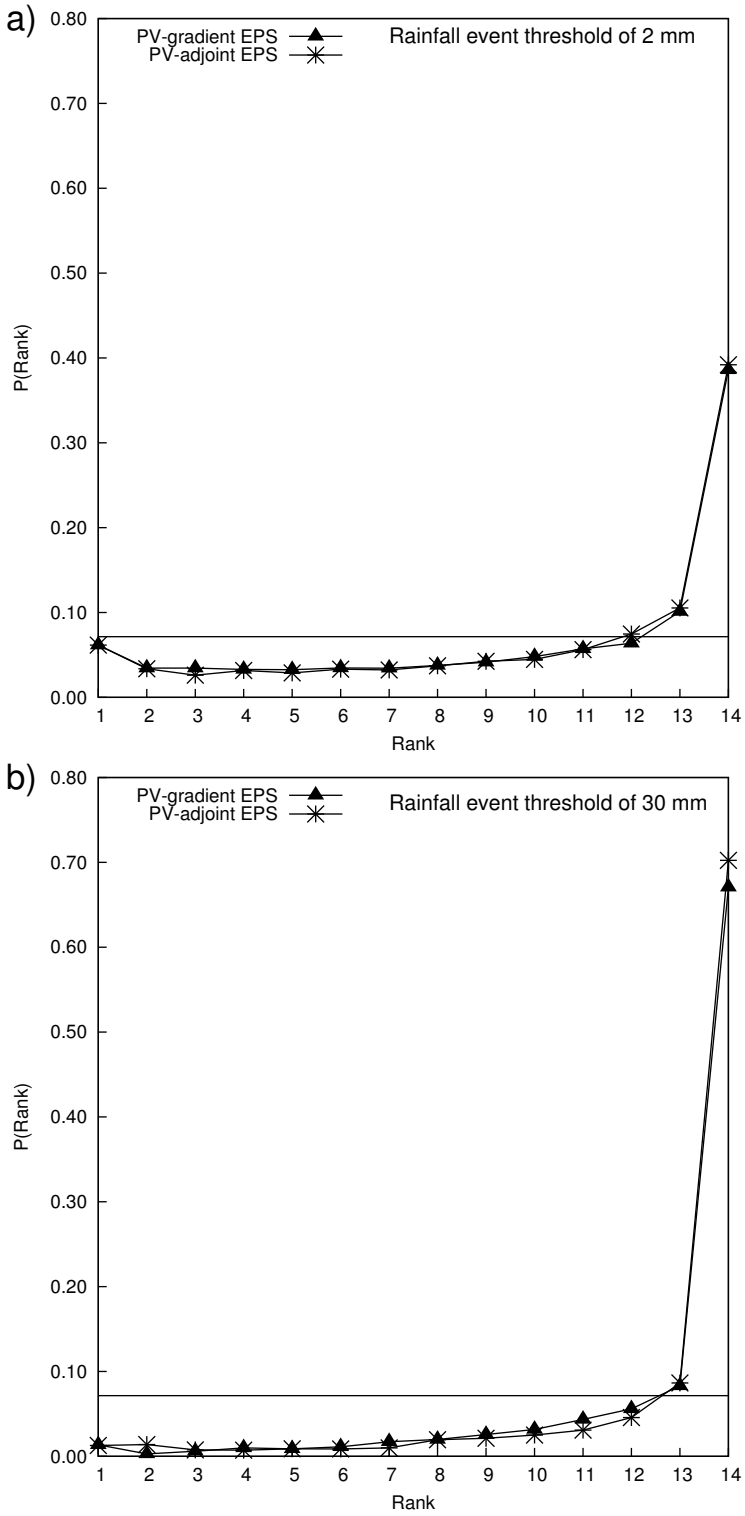


Figure 8.8. *Rank histograms for the PV-gradient and PV-adjoint ensembles, for a) 2 mm and b) 30 mm rainfall event thresholds. The horizontal line or flat histogram represents the perfect score.*

8.3 Conclusions

The performance of two different EPSs is evaluated using a thorough verification procedure. Considering that the aim of this study is to improve the short-range numerical forecasts of cyclones associated with heavy rain events in the western Mediterranean, the verification setup is focused on the 24 h accumulated precipitation field; a field observed over non gridded networks and highly discontinuous in both space and time that makes the evaluation harder and more demanding. In addition, the study deals with extreme events which are difficult to predict in nature and rare by definition.

Both ensembles are built using a single variable (PV) to define perturbations combined with the PV inversion technique, keeping the method simple while ensuring modifications of all the meteorological fields without compromising the mass-wind balance. The only difference between both EPSs lies in the location where the perturbations are introduced. The PV-gradient ensemble introduces the perturbations in the areas corresponding to the PV zones of most intense values and gradients (in essence a subjective choice based on our experience) while the PV-adjoint does it in the MM5 adjoint model calculated sensitivity zones (an objective method).

The high computational cost of the PV-adjoint ensemble (which implies running the MM5 adjoint model for each simulation day) versus the low cost of the PV-gradient is not compensated later in ensemble skill. Even though both EPSs are skillful and present a more than adequate performance, the results obtained by the PV-gradient ensemble are generally better than those obtained by the PV-adjoint EPS. Thus, the PV-gradient ensemble is proved to be a more profitable strategy than the PV-adjoint EPS, which is also useful but computationally more expensive for our testbed. This result is in agreement with Homar and Stensrud (2008) results, which stated that for intense cyclogenesis over the western Mediterranean, adjoint-estimated sensitivity is comparable or slightly inferior to subjective (gradient and human) sensitivity estimates.

Part III

Applications

Chapter 9

A multiphysics superensemble forecast*

In pursue of our final goal, to improve the prediction skill of the heavy precipitation events that characterize the western Mediterranean coastal countries, the superensemble technique is tested in the region. As exposed in Section 3.3, the superensemble forecast technique is a powerful post-processing method for the estimation of weather forecast parameters, like precipitation. Throughout Part II the multiphysics ensemble, as well as both PV-perturbed EPSs, were tested in the same fashion with good results. Therefore encouraged by the skill shown by the multiphysics ensemble prediction system, the superensemble is fed with it instead of the common approach of using a multimodel ensemble. Note that the superensemble technique requires an ensemble made up by distinguishable members, so from our three ensembles only the multiphysics qualifies.

9.1 Superensemble Construction

As part of the superensemble building procedure, a different databases for each phase is required. In this study, the forecast phase consists of the same collection of 19 MEDEX cyclones between September 1996 and October 2002 used in Part II,

*This Chapter is based on the published paper Vich, M. and R. Romero, 2010: *Multiphysics superensemble forecast applied to Mediterranean heavy precipitation situations*. Nat. Haz. and Earth. Syst. Sci., 10, 2371–2377. doi:10.5194/nhess-10-2371-2010.

while the training phase consists of a very wet four-month period, September - December 2001, characteristic of the precipitation climatology of the region during Autumn. Notice that both phases focus on the precipitation field, as in previous chapters.

The group of forecast members also needed to build the superensemble is provided by applying the multiphysics EPS generation techinque to the 56- and 120-day period comprising these 19 MEDEX cyclones between September 1996 and October 2002 and the four moths between September and December 2001, respectively. Fig. 9.1 shows the area coverage of the MM5 domain and the available observational data, the same we used in Part II to build and test our three ensembles.

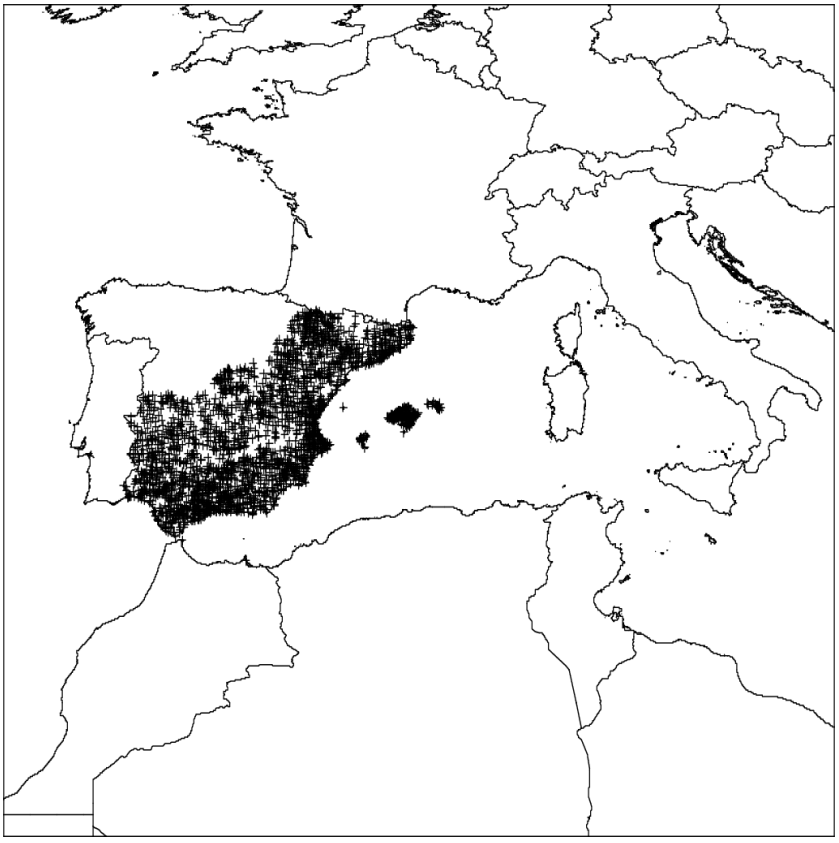


Figure 9.1. *Geographical domain used for the MM5 numerical simulations. The spatial distribution of the AEMET raingauge network used in the verification procedure is plotted using crosses.*

9.2 Experiments and results

The evaluation of the superensemble performance for the rainfall field is done thanks to the wide range of verification indices presented in Chapter 4 and also to comparing the superensemble results with that of the ensemble mean (a simple average of all the members), the bias corrected ensemble mean and the multiphysics EPS control member. It is worth noticing that for both phases, training and forecast, all computations have been done for the 24 h accumulated precipitation field considering both the 06-30 h and 30-54 h accumulated periods indistinctly, this implies that the number of days for both phases are doubled. Even though Krishnamurti et al. (2000b) separates both time windows we have checked that for this study merging them does not affect negatively the superensemble performance and we gain statistical significance. Since we are not focused on verifying a single observation threshold but on evaluating the general performance of the superensemble, the definition of the observed event is not fixed. For example, if a catchment gets flooded when it rains more than 50 mm/day this would be the observed event, since such threshold separates safety from disaster. Here nine rainfall amount thresholds (0, 2, 5, 10, 20, 30, 50, 100 and 150 mm) are defined as observed events.

The ROC area results plotted on Fig. 9.2 show that the bias corrected ensemble mean performs better than the other forecasts followed by the ensemble mean, the control member and finally by the superensemble, nevertheless all forecasts present ROC areas above 0.7, a very satisfying result according to Stensrud and Yussouf (2007) who establish that forecasting systems with ROC area greater than the mentioned threshold are useful to discriminate between events and non-events.

The Bias results (Fig. 9.3) show that both ensemble means, poor man and bias corrected, overpredict (Bias > 1) rainfall amounts less than 40 mm and underpredict (Bias < 1) the larger rainfalls amounts, while the control member presents the same behavior at a transitional threshold of 70 mm. On the other hand the superensemble overpredicts rainfall amounts less than 80 mm and keeps steady around the perfect score (Bias $= 1$) for larger rainfalls amounts.

The Taylor diagram (Fig. 9.4) shows similar results for both ensemble means, poor man and bias corrected, approximately both present a RMS difference of 12 mm and a correlation coefficient of 0.5, while the standard deviations of the forecast

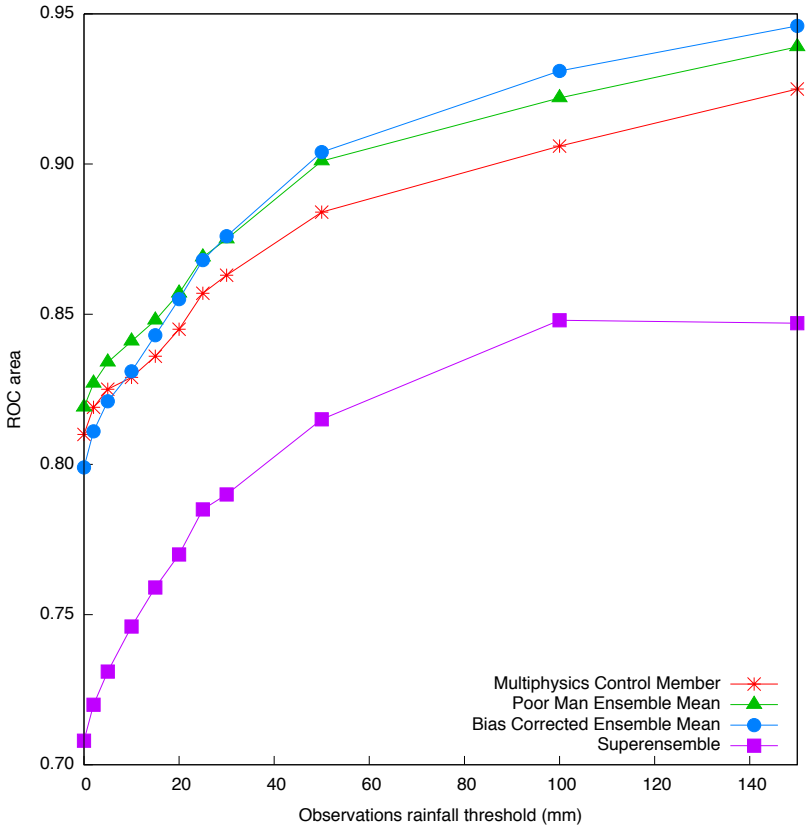


Figure 9.2. *ROC area for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds.*

are between 11 and 13 mm and the observed standard deviation is approximately 13 mm. The control member and the superensemble show a higher RMS difference, lower correlation and higher forecast and observation standard deviation. Note that the statistics used on the Taylor diagram are negatively affected owing to the discontinuities, noise and outliers characteristic of the rainfall field.

The Q-Q plots compare the observed and forecast distributions in terms of quantiles. A diagonal line indicates a perfect skill, while below the diagonal the forecast underestimates the observation and overestimates it over the diagonal (a more detailed description can be found in Wilks (1995)). Fig. 9.5 shows that over the 100 mm rainfall threshold all forecasts except the superensemble underpredict the observed precipitation distribution, while the superensemble captures the observed precipitation distribution (perfect score).

The obtained superensemble scores at the ROC area and Taylor diagrams are lower

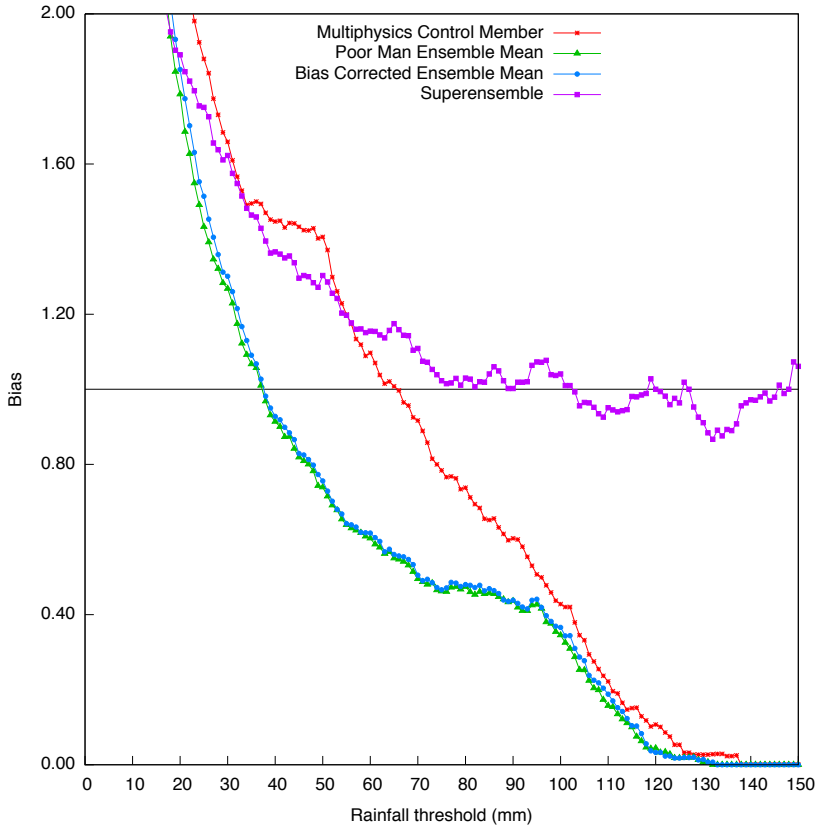


Figure 9.3. *Bias for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds.*

than expected. The cause of these low scores could be related to the superensemble dependency on the assumption that the performance of the members past forecasts accurately represent the performance of those members in the forecast period. Since we are dealing with extreme and rare events this assumption might not be achieved. Also it is worth to notice that ECMWF forecasts, our initial and boundary conditions, have undergone severals updates during the period this study is focused on (from 1996 to 2002). Although these changes could also affect the superensemble skill, the possible effects will be neglected at this stage under the assumption that the model physical parameterizations are the dominant source of variability in heavy precipitation simulations. Bearing this in mind a new test in done exchanging the training and forecast datasets in order to examine the stability of the results. In this new experiment the superensemble is trained for the MEDEX cyclones collection and tested for the 4-month period in the forecast phase.

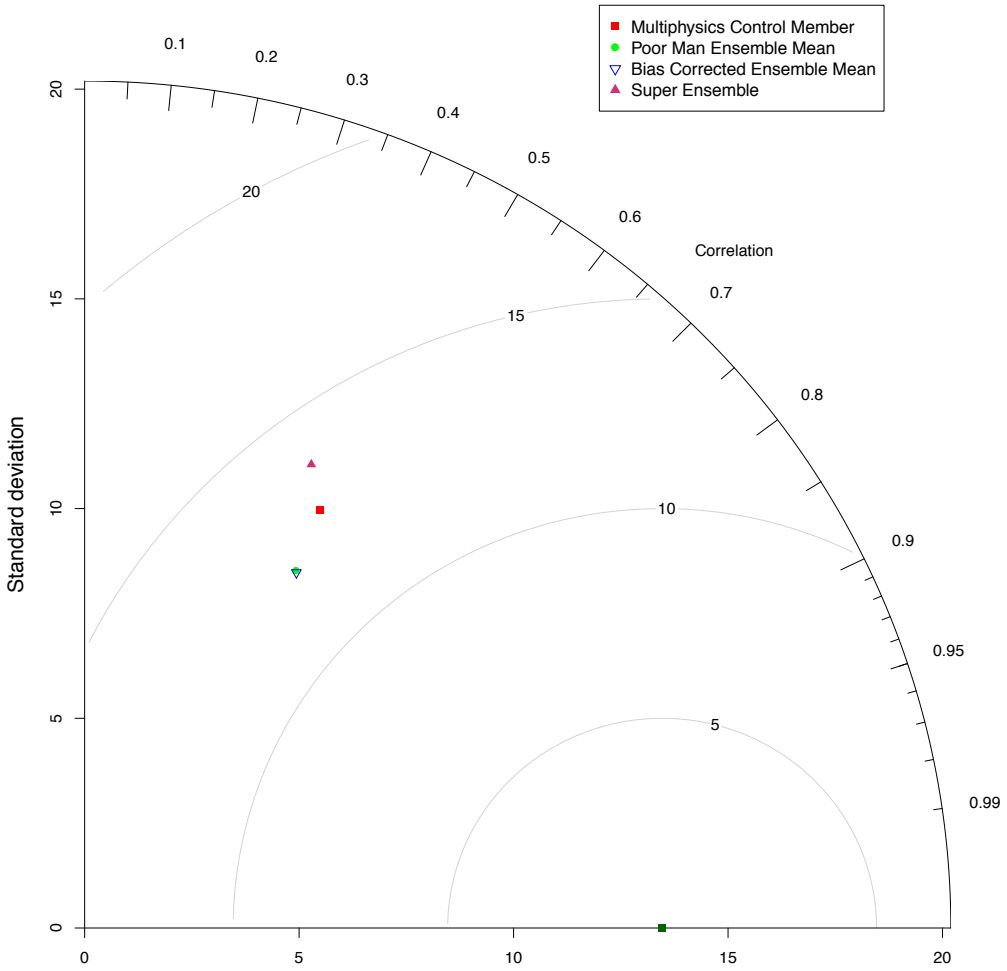


Figure 9.4. *Taylor diagram for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds. The perfect score is obtained when the data point representing the forecast field matches up with the observed one (green solid square).*

The ROC area (Fig. 9.6) shows that the superensemble is tuned for the 100 m rainfall threshold being the forecast with the highest score. The Bias (Fig. 9.7) also shows that the superensemble is the nearest to the perfect score for a wider range of rainfall thresholds than the other forecasts, and while the others underpredict for higher thresholds the superensemble slightly overpredicts them. The Taylor diagram (Fig. 9.8) behaves as in the previous experiment: both ensemble means (poor man and bias corrected) are the nearest to the perfect score, followed by the control forecast and the superensemble. The Q-Q plot (Fig. 9.9) also shows

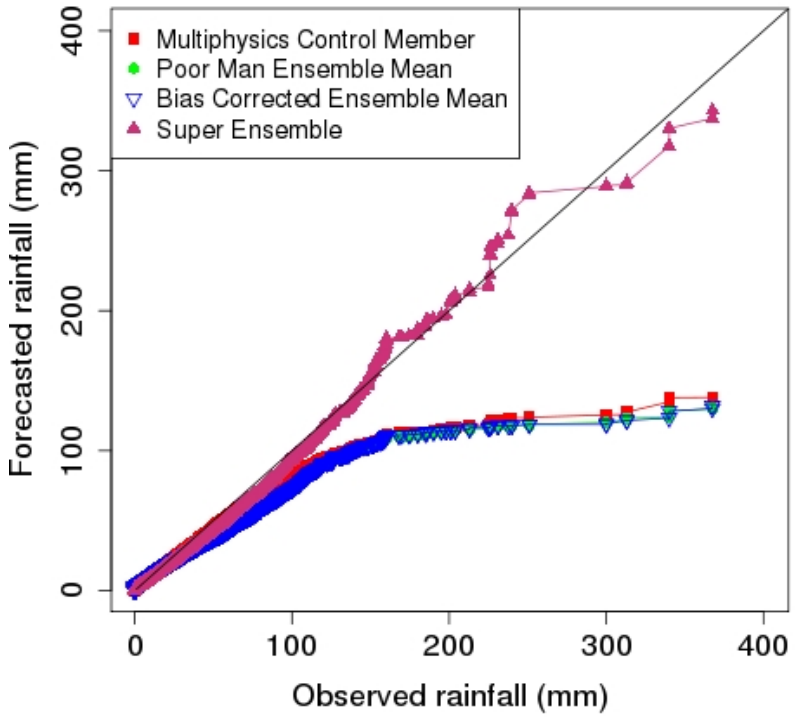


Figure 9.5. *Q-Q plot for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds.*

that the superensemble is the nearest to the perfect score as in the previous test. These results seem to indicate that exchanging the superensemble datasets make the superensemble more attuned to the higher precipitation thresholds.

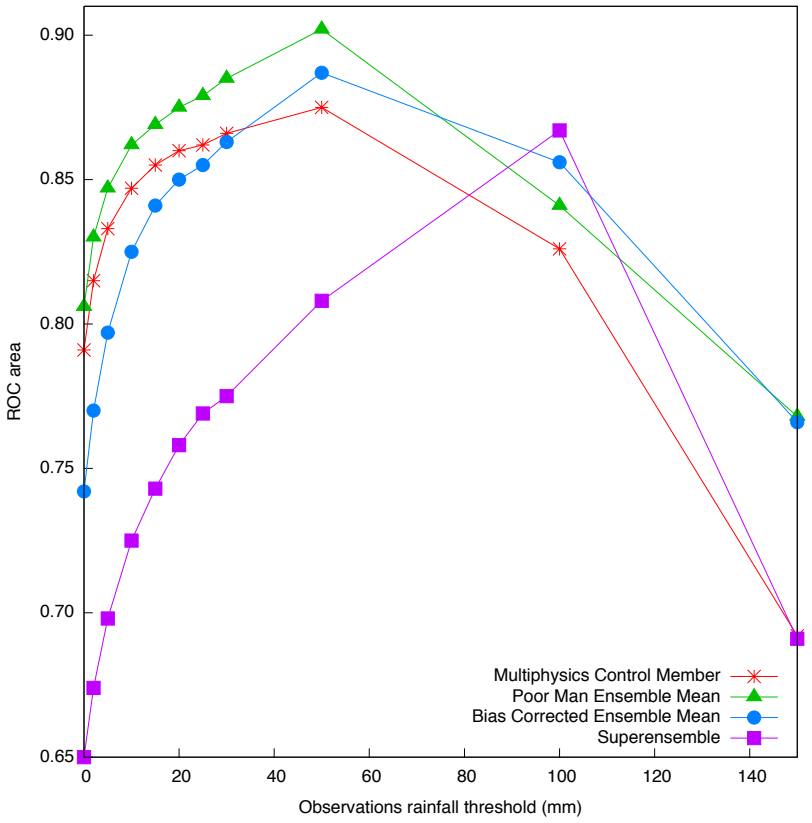


Figure 9.6. *ROC area for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds. Note: Training period 19 MEDEX cases and forecast period Sep-Dec 2001.*

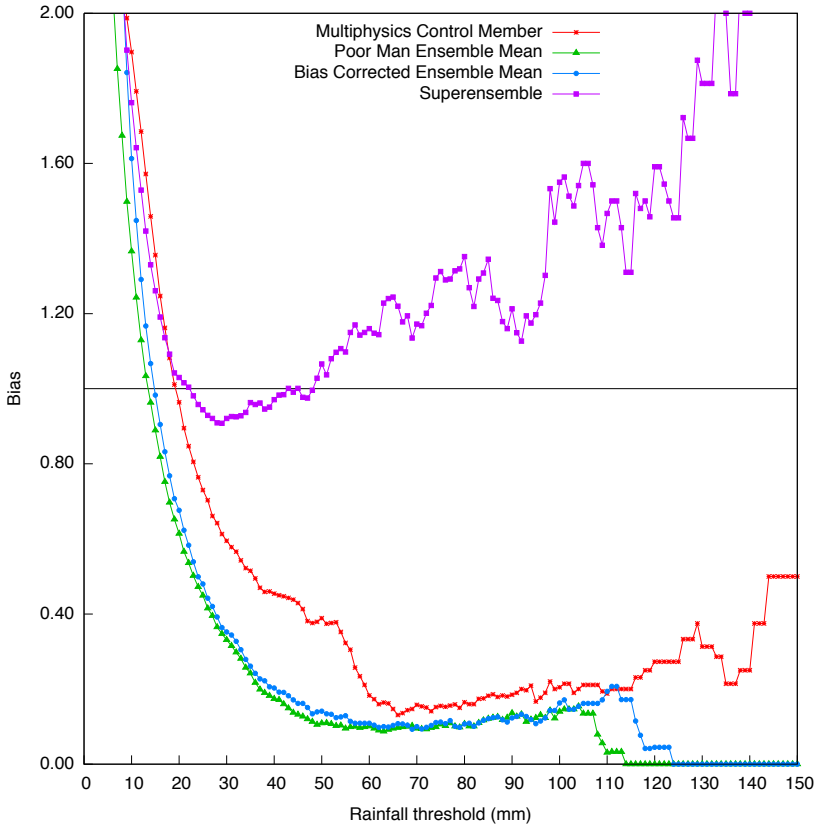


Figure 9.7. *Bias for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds. Note: Training period 19 MEDEX cases and forecast period Sep-Dec 2001.*

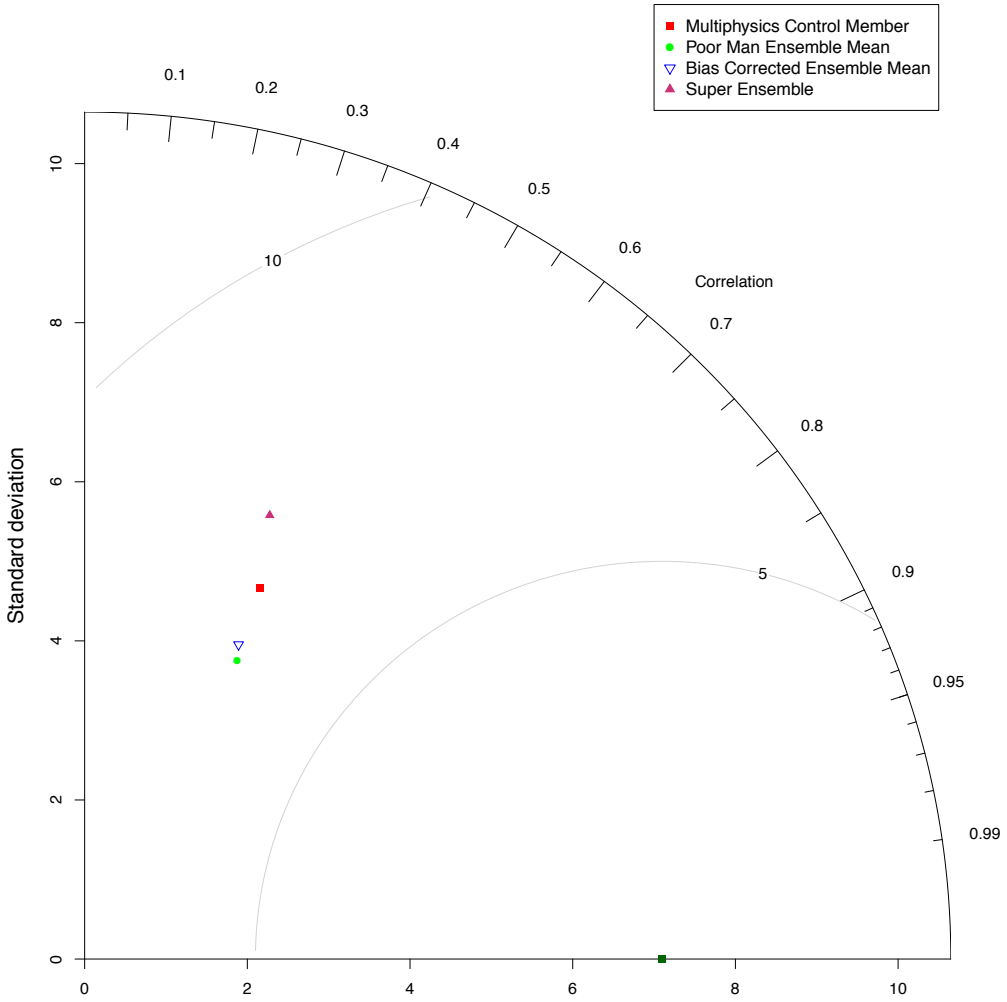


Figure 9.8. *Taylor diagram for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds. The perfect score is obtained when the data point representing the forecast field matches up with the observed one (green solid square). Note: Training period 19 MEDEX cases and forecast period Sep-Dec 2001.*

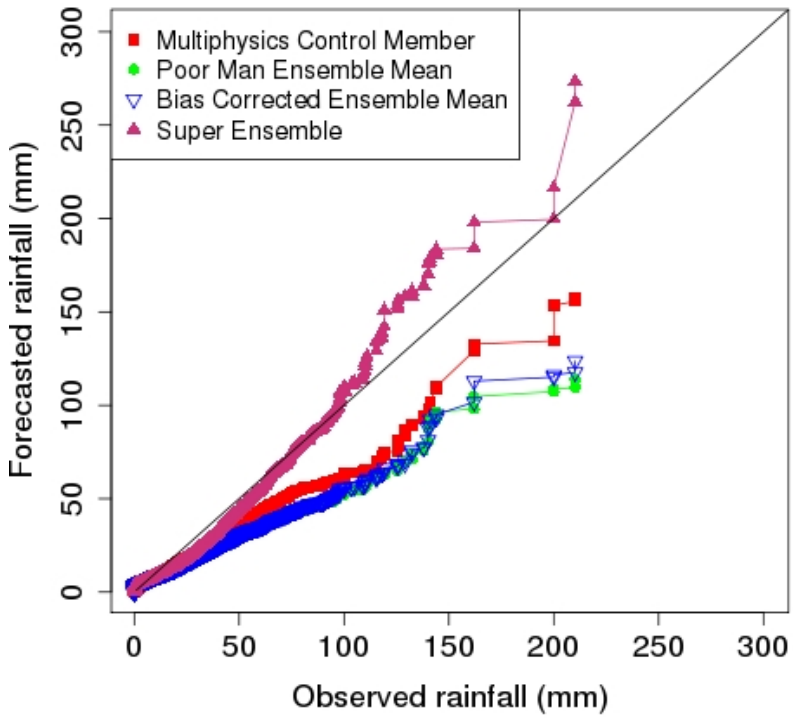


Figure 9.9. *Q-Q plot for the multiphysics ensemble control member, the multiphysics ensemble mean, the bias corrected ensemble mean and the superensemble, as functions of different rainfall event thresholds. Note: Training period 19 MEDEX cases and forecast period Sep-Dec 2001.*

9.3 Conclusions

The superensemble based on a multiphysics EPS instead of a multimodel ensemble and applied to rare and extreme events has not performed as expected, although the superensemble has proved its value in previous studies dealing with ordinary situations. The fact that we are dealing with a multiphysics ensemble may lead to more correlation between ensemble members and therefore affect the multi-linear regression technique used to calculated the superensemble weights. Another fact that is worth to note is that the superensemble technique assumes that the past behavior of each ensemble member is representative of the present behavior, and this assumption may not be accurate for the kind of events tested in this study, cyclone-induced heavy rain events, rare and extreme by definition.

The bias corrected ensemble mean and the poor-man ensemble mean show a clear improvement over the control member, as expected. Furthermore, the superensemble is the best in the bias and Q-Q plots scores but is not good enough in the ROC area and Taylor diagrams scores. The second experiment points out that inverting the superensemble datasets attunes the superensemble better for higher rainfall thresholds. In spite of the difficulties derived from focusing the verification on the rainfall field, its results stress the good performance of the forecasts, specially how the superensemble captures the quantile distribution of the precipitation.

Additional experiments can be envisaged to help to improve the performance of the superensemble approach regarding extreme events. This line of research will be revisited in the last chapter of this Thesis.

Chapter 10

Inclusion of PV uncertainties into a hydrometeorological forecasting chain*

An ensemble prediction system can be regarded as an appropriate approach to introduce the uncertainties found in the initial and boundary conditions of mesoscale model forecasts into the hydrometeorological modeling chain. Hydrological models driven by the ensemble-derived quantitative precipitation forecasts (QPFs) are better fitted to deal with the difficulties in providing precipitation over the correct locations and times, owing to both the small scale nature of the responsible atmospheric features and model physical parameterization deficiencies.

It has been established that traditional warning systems based on hydrological models driven by rainfall observations do not provide the timely predictions required to implement the precautionary civil protection measures on short-time scales of flash floods events (Siccardi, 1996) and that the use of high-resolution numerical weather prediction models further extend the lead times associated to discharge predictions (with lead times up to 12-48 h; e.g. Benoit et al. 2003). So the use of QPFs derived by an ensemble prediction system could further enhance this improvement in discharge predictions lead times and have a positive impact

*This Chapter is based on the published paper Amengual, A., R. Romero, M. Vich, and S. Alonso, 2009: *Inclusion of potential vorticity uncertainties into a hydrometeorological forecasting chain: application to a medium size basin of Mediterranean Spain.* Hydrol. Earth Syst. Sci., 13, 793–811. doi: 10.5194/hess-13-793-2009.

on the flash flood warning systems.

10.1 The targeted region

As emphasized in Chapter 1, the Spanish Mediterranean area (Fig.10.1), as part of the western Mediterranean, is often hit by high impact weather phenomenon like flash flood, especially in autumn. The coast steep streams and high urbanization rates typical of the region lead to short hydrological response times and potentially high personal and material damages. Moreover, the characteristic semiarid environment is associated with many ephemeral small and medium steep streams prone to high flow velocities and large volumes of sediments causing unexpected and extensive flood damage. Not surprisingly, many efforts, like ours, are being devoted to flood forecasting tools, a key point in implementing reliable warning systems before these kinds of events.

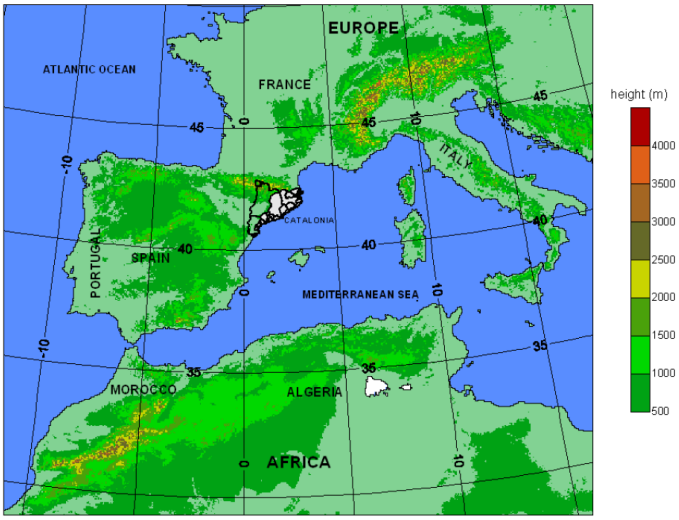


Figure 10.1. *Geographical locations and main mountain systems of the Western Mediterranean region. Major topographic features are shown according to the colour scale at 500 m intervals starting at 500 m. The thick continuous line shows Catalonia and its internal basins (highlighted with shaded grey).*

The Llobregat river basin is a medium size catchment located in Catalonia, northeastern Spain (Figs.10.1 and 10.2), and it is periodically affected by intense precipitation events resulting in floods which sometimes produce hazardous effects. In fact, four notable hydrometeorological episodes occurred over the 1996-2004 period affecting the Llobregat river basin. These four events are the perfect testbed

to test the viability of the Hydrologic Engineering Center's Hydrological Modeling System (HEC-HMS) runoff model discharge predictions driven by the QPFs provided by:

– the kriging interpolation of the rain-gauge data available,

– the MM5 mesoscale model forecast nested within the ECMWF large-scale forecast fields, and

– the PV-gradient ensemble prediction system developed in Chapter 6.

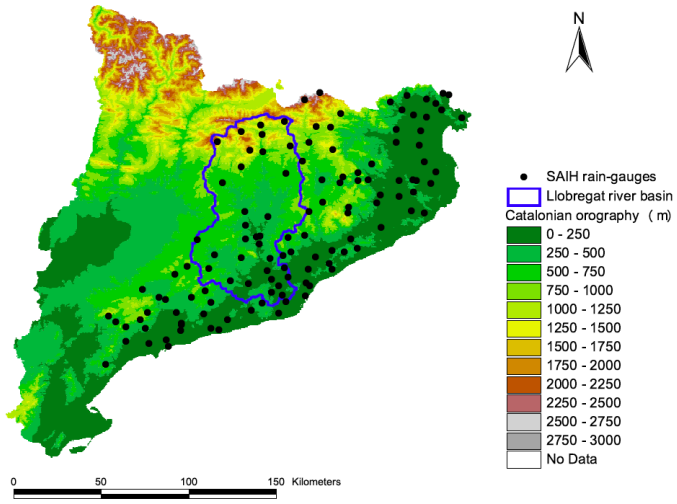


Figure 10.2. *Distribution of the rain-gauges from the Automatic Hydrological Information System in the internal basins of Catalonia (IBC; shown in Fig. 10.1). It includes a total of 126 automatic rainfall stations distributed over an area of 16000 km^2. Llobregat basin is highlighted.*

The selected events produced heavy precipitation which resulted in severe floods on the period covering 1996 until 2004. These case studies are a sample of different intense rainfall episodes produced by prototype atmospheric dynamical situations. The first two cases (16-17 November 1996 and 17-18 December 1997) were linked with sustained and stratiform-like rainfalls over long periods of time, which resulted in remarkable discharges at the Llobregat basin outlet. The last two cases (9-10 June 2000 and 29-30 August 2004) produced important and sudden rising flows owing to their convective nature, bearing heavy precipitations in short temporal scales.

Note that from now on, we will focus our attention on the MM5 deterministic simulation and on the PV-gradient EPSs approach, presenting their results for

the 9-10 June 2000 event. Additionally, the results from the PV-gradient EPSs approach for all four episodes are gathered on several verification scores, like for example, the Bias or the probability of detection. Readers interested in the whole study can find it in Amengual et al. (2009).

The 9-10 June 2000 event, also known in the literature as the *Montserrat* flash flood event, is part of the 19 MEDEX event collection used to verify our ensembles and it was used to illustrate the methodology of both the PV-gradient and PV-adjoint ensembles prediction systems. On this event, many internal basins of Catalonia (IBC, Fig. 10.1) were overwhelmed as a result of the heavy precipitation produced by a quasi-stationary mesoscale convective system. The following flood produced five fatalities and material losses estimated at about 65 million euros. A more detailed description of the event can be found in Section 11.1.

10.2 Experiment setup

The Llobregat river basin is the most important internal hydrographic catchments of Catalonia regarding to its size, river length, mean flow and population living inside (Figs. 10.2 and 10.3). This catchment is formed by the Llobregat river and its main tributaries, the Anoia and the Cardener. Llobregat basin extends from the Pyrenees - with heights over 3000 m, through the pre-Pyrenees – with a height transition from 750 m to 200 m in the pre-coastal range. The last section of the river flows through the Mediterranean orographic systems which consist of two mountainous aligments almost parallel to the coast line: the pre-coastal and coastal ranges. The basin has a total drainage area of 5040 km^2 and a maximum length close to 170 km (for a more detailed description of the watershed see Amengual et al. 2007).

The available raw precipitation is made up by 5-min rainfall data recorded at 126 stations inside the Internal Basins of Catalonia distributed over an area of 16000 km^2 (Fig. 10.2). This network of rain-gauge stations belongs to the Automatic Hydrological Information System (SAIH) of the Catalan Agency of Water (ACA). Approximately 45 of this 126 stations lie inside the Llobregat basin or near its boundaries.

This study is carried out using the physically-based HEC-HMS rainfall-runoff model (USACE-HEC, 1998). The model has been implemented in a semi-distributed

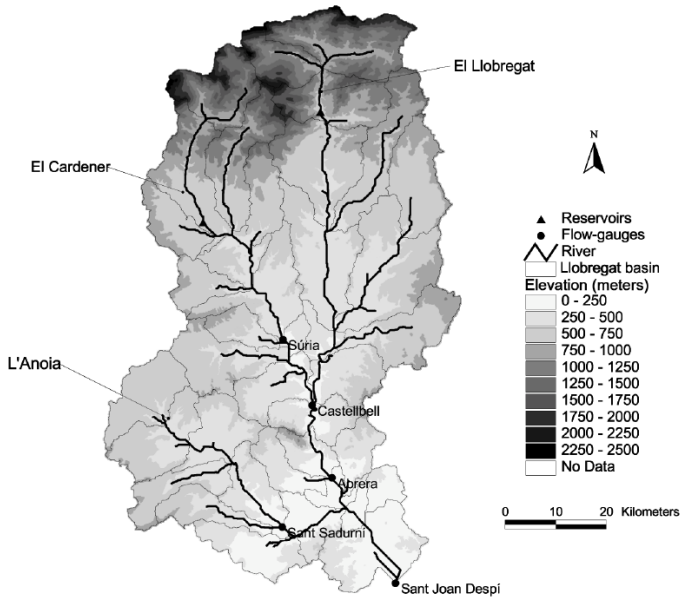


Figure 10.3. *Digital terrain model of the Llobregat river basin with a cell size of 50m. It displays the basin segmentation, main tributaries, stream-gauges (circles) and reservoirs (triangles).*

and eventbased configuration. HEC-HMS utilizes a graphical interface to build the semi-distributed watershed model and to set up rainfall and control variables for the simulations. Fig. 10.3 shows the digital terrain model for the Llobregat river basin together with the main watercourses and its tributaries, the considered division in sub-basins and the location of the available river gauges.

The hydrological model is forced by using a single hyetograph[1] for each sub-basin. Rainfall spatial distributions are generated applying the kriging interpolation method with a horizontal grid resolution of 1000 m on the hourly accumulated values from: 1) the SAIH rain-gauges records (irregular grid) and 2) from the model regular gridded QPFs. Then, the hourly rainfall series are calculated for each sub-basin as the area-averaged of the gridded rainfall within each subcatchment.

The MM5 model set up used to perform the meteorological simulations is the one used by both PV-perturbed ensembles. A complete description can be found in Section 6.3, page 72. The only addition is a nested second domain, with two way interaction and a horizontal resolution of 7.5 km. As Fig. 10.4 shows, the second domain spans the entire Catalonian territory and contiguous land and oceanic areas, and is used to supply the high-resolution rainfall fields to drive the

[1] A graphical representation of the distribution of rainfall over time.

hydrologic simulations.

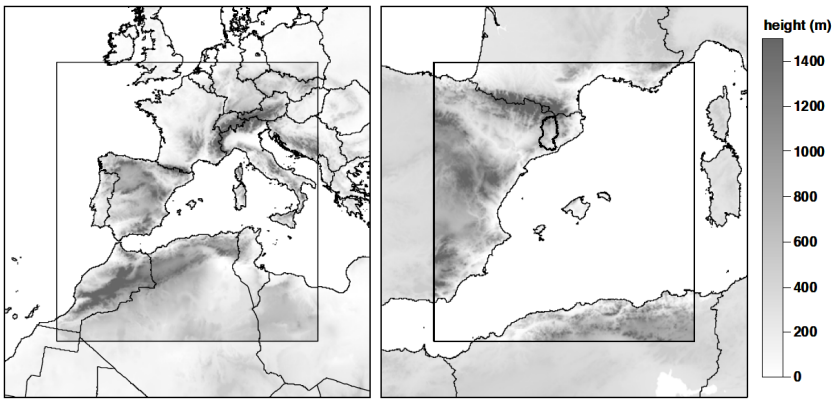


Figure 10.4. *Configuration of the two computational domains used for the MM5 numerical simulations (inner squares, with horizontal resolutions of 22.5 and 7.5 km for left and right, respectively). The Llobregat river basin is highlighted as thick line on the smaller domain.*

The ensemble simulations for each hydrometeorological episode are also run with the same MM5 model configuration. Since the inner domain is two way nested into the outer domain, the PV-gradient initial and boundary perturbations are only applied on the outer domain. All MM5 simulations cover a a 54 h period for each of the flood episodes under study, starting at 00:00 UTC 16 November 1996 for Case 1; 00:00 UTC 17 December 1997 for Case 2; 00:00 UTC 9 June 2000 for Case 3; and 00:00 UTC 29 August 2004 for Case 4.

10.3 Results

10.3.1 Runoff simulations driven by MM5 deterministic QPFs

The deterministic forecasts for the 9-10 June 2000 event displays poor skill in reproducing the observed accumulated precipitations and its spatial and temporal distributions over the catchment (see Fig. 10.5). The MM5 control simulation is clearly deficient as the maximum quantity of precipitation is located quite far away from the Llobregat basin. The set of verifications indices used to evaluate the skill of the MM5, correlation coefficient (r), root-mean-square error (RMSE) and mean absolute error (MAE), is also in agreement as Table 10.1 clearly shows. In fact, the actual maximum quantity is located in the north-western part of the Pyrenees range (not shown here).

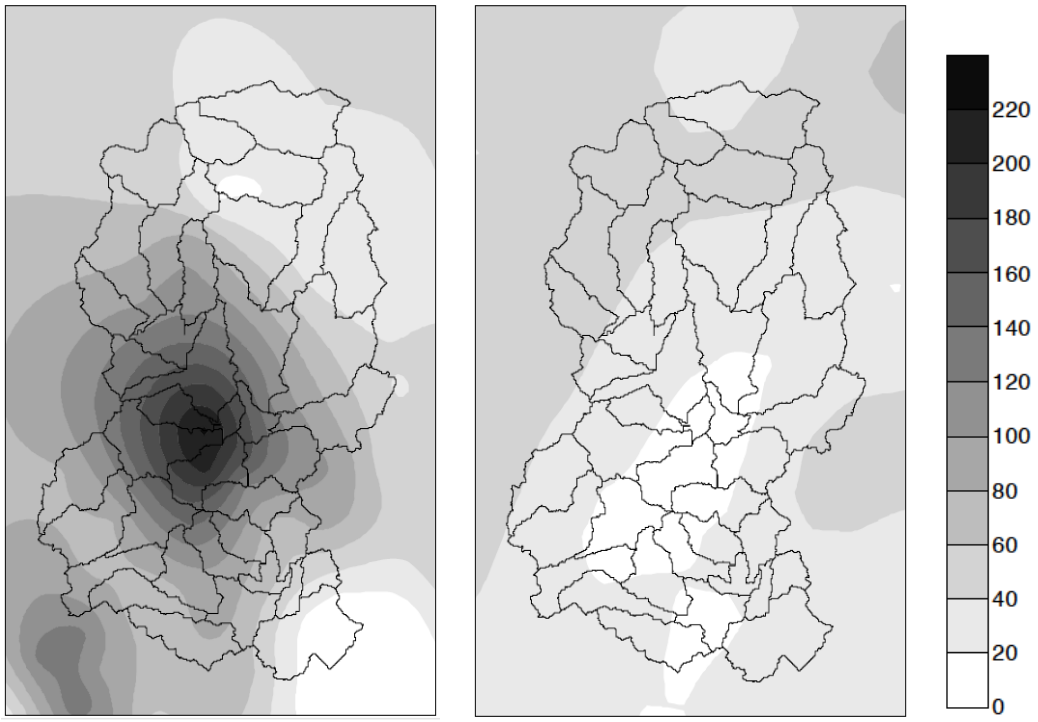


Figure 10.5. *Spatial distribution of the accumulated precipitation (in mm according to the scale) in the Llobregat basin for 9-10 June 2000 episodes: a) Observed and b) MM5 control forecast simulation.*

	spatial	temporal
r	0.28	0.11
RMSE	8.7	24.2
MAE	6.7	15.8

Table 10.1. *Correlation coefficient, RMSE (in Hm^3) and MAE (in Hm^3) of the spatial and temporal rainfall volume distributions yielded by the 9-10 June event MM5 control simulation.*

These deficiencies affect the runoff simulation driven by the MM5 model results. In fact, the severe underestimation in the QPFs means that the runoff simulation misses the observed peak flow as Fig. 10.6 clearly exhibits.

Amengual et al. (2007) studied the same flash-flood event but initializating the MM5 model with the NCEP analyses. It was found that the control simulation reasonably reproduced the spatial and temporal distributions of the observed rainfall and the total precipitated volume. So it would appear that that there is a high dependence of the hydrometeorological prediction on the initial and boundary conditions of the databases, aside from the dependence on the flood event itself.

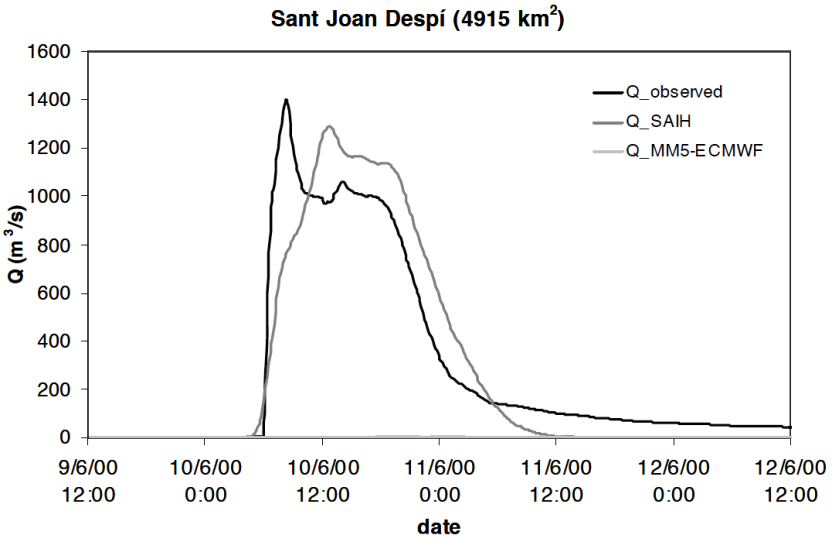


Figure 10.6. *Observed SAIH rain-gauge driven and MM5 control simulation driven runoff discharge for 9-10 June 2000 episode. This hydrograph corresponds to the stream-gauges of Sant Joan Despí where the maximums of rainfall (224 mm) and flow (1400 $m^3 s^{-1}$) were observed.*

Its is worth mentioning that here we have only shown one of the four events explored in the complete study, in fact one out of the two events that were not captured by the MM5 deterministic simulation. The runoff prediction driven by the well simulated events (November 1996 and December 1997) could be considered suitable to use in an emergency management directive (e.g. Anderson et al. 2002).

10.3.2 Runoff simulations driven by MM5 probabilistic QPFs ensemble

A PV-gradient ensemble of 20 members is built following the generation procedure detailed in Section 6.3, plus the control member (the previous non-perturbed MM5 deterministic run). This 21-member ensemble is used to drive a new set of runoff simulations as means of introducing the atmospheric initial and boundary conditions uncertainty into the hydrometeorological chain.

The discussion of the results is done for the runoff peak flow by means of cumulative distribution functions (CDFs) plotted in a Gumbel chart (e.g Ferraris et al. (2002); Fig. 10.8). Although no hydrometeorological forecasting chain is currently implemented for civil-protection purposes in the Llobregat river basin, we have considered suitable for the present study the introduction of two hypothetical

warning discharge thresholds (the maximum peak discharges for return periods of 5 and 10 years). These peak flows have been extracted from the technical report issued by ACA (ACA, 2001). The first and second warning discharge thresholds are 266.3 m^3s^{-1} and 555.4 m^3s^{-1} at Sant Joan Despí hydrometric section. On the other hand, a the quality of the PQPFs at sub-basin scale has been tested by a set of categorical verification indices: the probability of detection (POD), the false alarm rate (F), the frequency-bias score (Bias) and the relative-operating-characteristics (ROC) score[2], computed for diverse precipitation and runoff volume thresholds. In order to increase the statistical significance of the results, these skill scores are calculated using all the hydrometeorological episodes as follows:

1. the rainfall and driven runoff volumes of the MM5 control runs are compared against the observed rainfall and SAIH rain-gauge driven runoff volumes, respectively;
2. the rainfall and driven runoff volumes of the ensemble means are employed;
3. the rainfall and driven runoff volumes by the individual members of the ensembles are used for the comparison with the observed rainfall and rain-gauge driven runoff volumes.

Note that all these rainfall and runoff volumes are accumulated at hourly time-steps and using the sub-basins as accumulation units. For the driven discharge volumes, the scores are calculated using only the hourly data produced at each sub-basin that is non-zero in at least one of the two compared series in order to prevent an artificial improvement of the ROC score values. As commented in Chapter 3, statistically speaking the ensemble mean should provide a better forecast that any individual ensemble member, as errors in the individual forecasts tend to cancel when averaged, so some benefits from a simple ensemble average over the control experiment is expected.

Fig. 10.7 displays the ensemble mean and standard deviation of the accumulated forecast rainfall over the MM5 inner domain. The rainfall pattern clearly shows an important spread of the rainfall cumulative values over north-eastern Spain within the mesoscale model ensemble.

It seems that the *Montserrat* hydrometeorological event is highly sensitive to the PV perturbations, recall the control run missed it completely. Although this

[2]ROC score, also known as ROC area. See more details in Chapter 4.

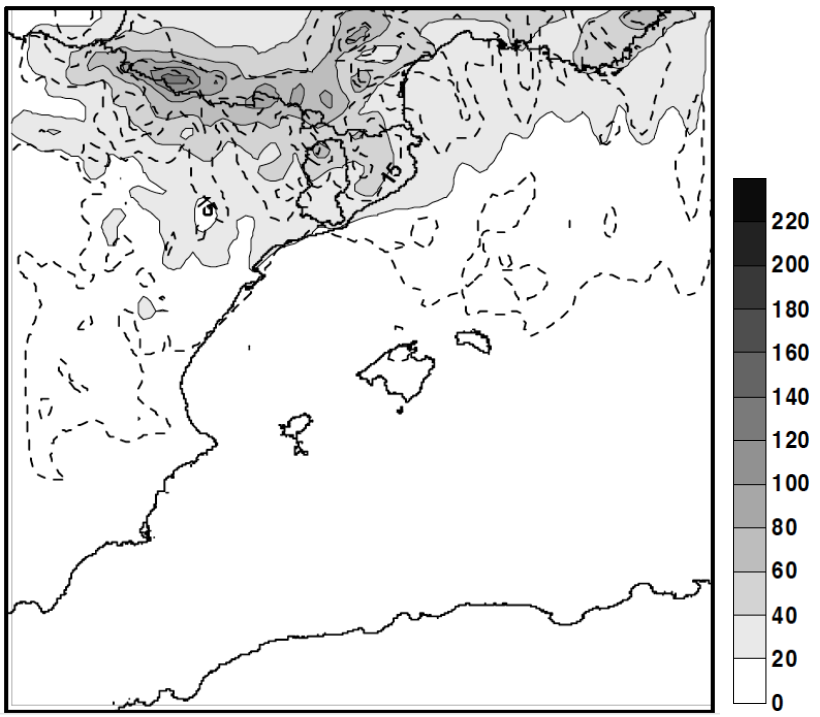


Figure 10.7. *Ensemble mean (shaded contours, in mm) and ensemble standard deviation (dashed line, in mm at 7.5mm intervals) for the accumulated precipitation over the second MM5 computational domain for the 9-10 June 2000 event. The Llobregat river basin is highlighted as thick line.*

episode was characterized by an important orographic enhancement mechanism in the organization, development and duration of the quasi-stationary mesoscale convective system, the atmospheric dynamical forcing played an essential role as well. These dynamical factors are linked to the upper-level synoptic trough and mid-to-upper-tropospheric flow, which are emphasized by the PV perturbation approach. As consequence, the biases found in the control run are partially corrected, and in spite that none of the members of the ensemble is able to forecast the observed peak flow for this exceptional event, some of them yield important peak discharges (Fig. 10.8). The probabilities of exceeding the first and second warning discharge thresholds are close to 0.3 and 0.15, respectively. So at least, a first theoretical warning could have been issued, even though the magnitude of the *Montserrat* episode would have remained undetected.

The aforementioned verification indices results for the rainfall and runoff volumes and their 95% confidence intervals are present on Table 10.2 and Figs. 10.9 and 10.10. This confidence intervals have been computed for each experiment by a bootstrap test (Diaconis and Efron, 1983) with 1000 repetitions in order to better

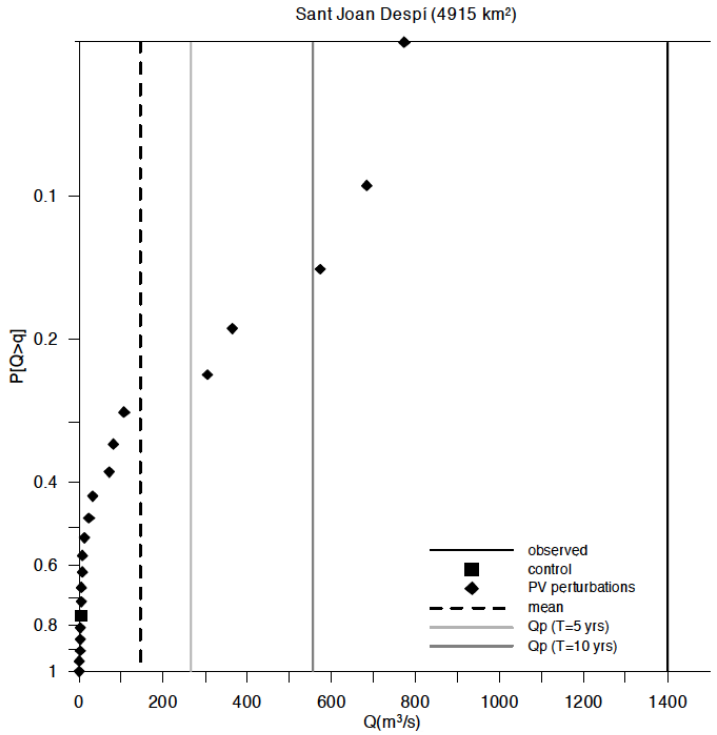


Figure 10.8. *Peak discharge exceedance probability plotted on a Gumbel chart for 9–10 June 2000 episode at the stream-gauge of Sant Joan Despí. The vertical black line indicates the observed maximum peak flow, the vertical dashed black line indicates the maximum peak discharge from the ensemble mean. Light and dark grey vertical lines show the maximum peak flows of 5 and 10 years return periods, respectively.*

establish the uncertainties around the statistical indices.

- *Rainfall volumes* results: The ensemble means exhibit the best skill for the POD at low and medium thresholds at expenses of a decrease in skill for the F score. In other words, as the capacity of detecting the event increases the false alarm rate also increases. However, the ensemble means ROC score, that relates POD and F, is also the best of the three, indicating that the ensemble means have smoothed the individual forecasts errors when averaging the rainfall fields. Also, the ROC score for the control runs exceeds the ROC score obtained from the ensembles, as the deterministic runs present a higher probability of detection for all the volume thresholds than the full ensembles while both sets of experiments present a similar false alarm rates for all volume thresholds. The Bias score results indicate that a systematic underforecasting of the precipitation amounts at all thresholds is occurring on the deterministic and the probabilistic experiments.

- *Runoff volumes* results: In this case, the best ROC score is obtained by the full ensembles as they also obtain the smallest false alarm rate scores while their POD skill is comparable to the one obtained by the ensemble mean. However, the control simulations are the ones showing the highest POD score. Notice the overall underforecasting of all simulations, and that the less biased are the control runs. This general underprediction of the runoff volumes clearly benefits the false alarm rates which become smaller, and therefore, lead to higher ROC scores. In addition, no improvement of the ROC scores is found when comparing the ensemble means against the deterministic simulations.

	control	mean	ensemble
rainfall volumes	0.62 (0.56-0.68)	0.65 (0.61-0.69)	0.58 (0.54-0.62)
runoff volumes	0.79 (0.73-0.84)	0.79 (0.73-0.84)	0.87 (0.83-0.89)

Table 10.2. *ROC scores and the 95% percentile confidence intervals for the control simulation, ensemble mean and ensemble members of all the hydrometeorological experiments, for hourly rainfall and runoff volumes.*

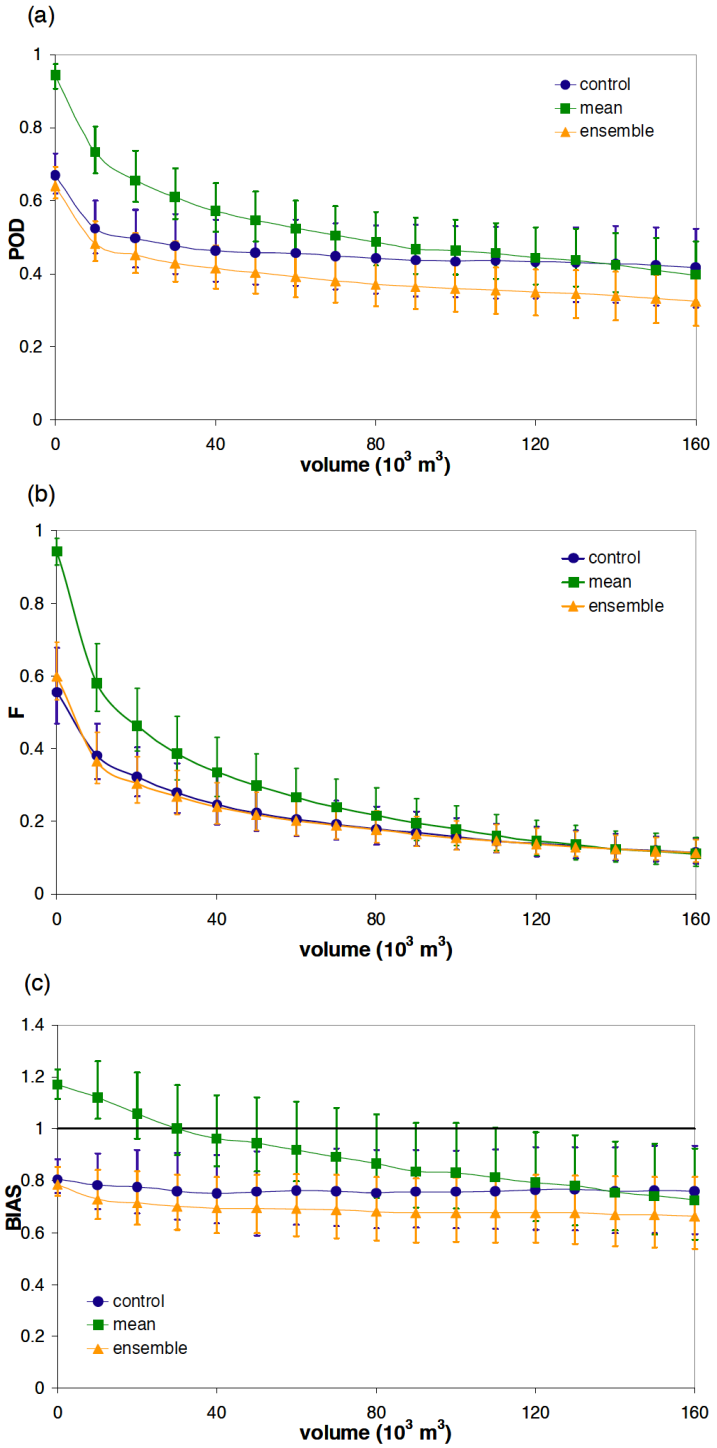


Figure 10.9. *POD, F and BIAS skill scores and the 95% percentile confidence intervals for different* ***precipitation*** *volume thresholds obtained by the ensemble of MM5 simulations for all the hydrometeorological experiments.*

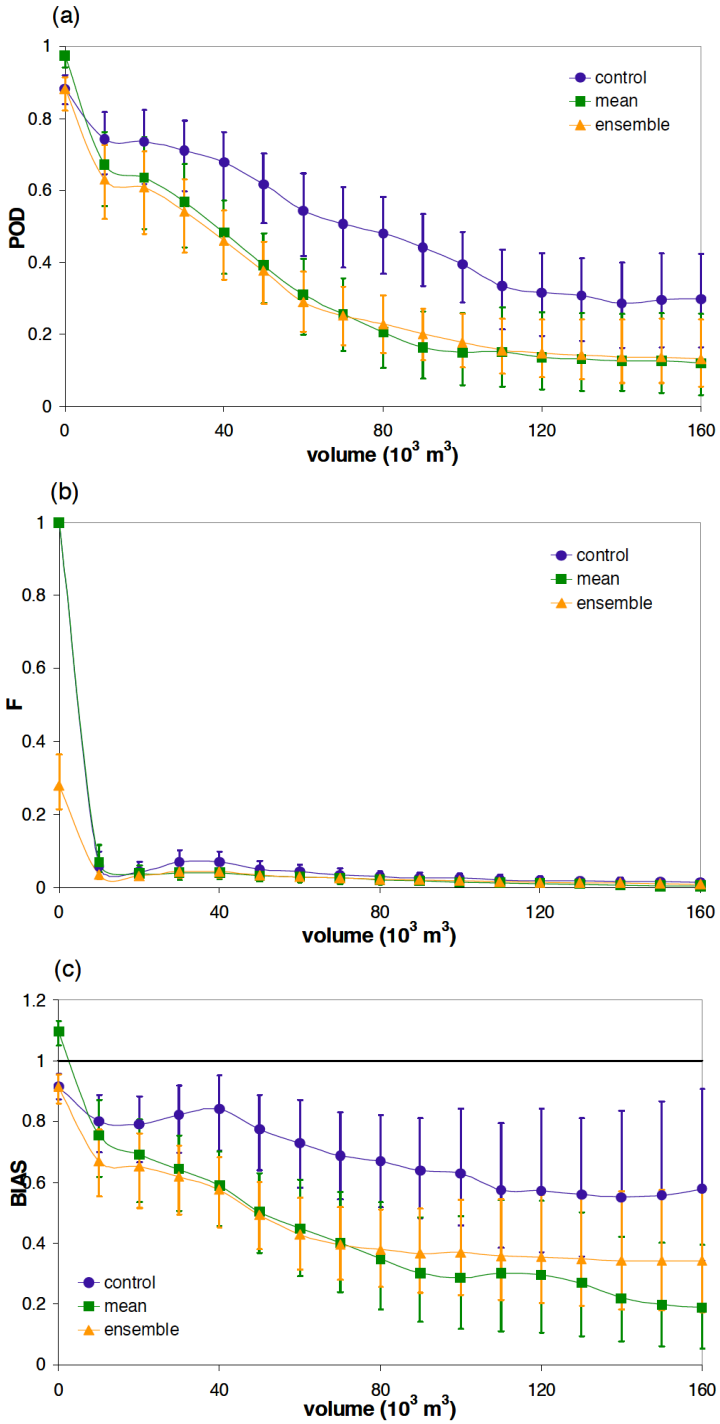


Figure 10.10. *POD, F and BIAS skill scores and the 95% percentile confidence intervals for different* ***runoff*** *volume thresholds obtained by the ensemble of MM5 driven runoff discharge simulations for all the hydrometeorological experiments.*

10.4 Summary and conclusions

Triggering an early warning to implement the right emergency procedures before extreme floods is critical to reduce their possible hazardous consequences. This process can benefit from a runoff forecast model driven by QPFs provided by a mesoscale model and an ensemble prediction system. That is why this study assesses the use of a MM5 model and the PV-gradient ensemble forecasts to force the HEC-HMS runoff model, focusing on four intense rainfall episodes which resulted in flood events of different spatial and temporal scales over the Llobregat medium size river basin. These kinds of intense precipitation events, often highly localized and convectively driven, present short recurrence periods in the Spanish Mediterranean area and deserve particular attention.

The MM5 model is nested within the ECMWF large-scale forecasts in an attempt to reproduce an operational hydrometeorological context. The PV-gradient ensemble forecasting system is also run for each case study. The performance of the control/non-perturbed QPFs has been evaluated from a comparison of the forecast and observed rainfall volumes in space and time over the sub-basins by means of a set of continuous verification scores. On the other hand, the probabilistic QPFs have been evaluated following the same procedure, but using a set of categorical verification indices. This methodology allows to assess the performance of the forecast rainfall patterns at the catchment scale. The one-way coupling between the meteorological and hydrological models has been regarded as a validation tool as well.

Most of the deterministic quantitative precipitation forecasts show significant deficiencies over the Llobregat river basin in terms of the rainfall amounts, their timing and localization. Obviously, these deficiencies have a major impact on the one-way coupling between the meteorological and hydrological models. Our ensemble strategy succeeds in reducing these biases at the analyzed hydrometric sections (Castellbell, not shown here, and Sant Joan Despí, partially shown), for most of the hydrometeorological episodes. However, the probabilistic approach indicates a better achivement of pre-defined theoretical warning runoff thresholds for only the 17-18 December 1997 case (not shown). It is worth to note that due to the restricted number of important floods available in the databases, the procedure presented here has been applied to a limited number of cases but representative of different dynamical atmospheric situations bearing important precipitation

amounts.

The value of using an EPS that accounts for the intrinsic uncertainties in initial and boundary conditions has been well proven. It has also pointed out the high/low sensitivity of different dynamical and orographic factors to the PV perturbations. The external-scale meteorological uncertainties have been reflected through moderate spatial and temporal variations of the forecast rainfall patterns as well as in significant changes of the precipitation amounts.

Chapter 11

Perturbing the PV field using the Water Vapor brightness temperature field as guidance*

In agreement with our previous line of work, a different approach to perturbing the PV field is implemented in this Chapter. Our goal remains unchanged: to minimize or even prevent the dramatic consequences on society that the western Mediterranean high impact weather events cause by improving their mesoscale numerical forecasts.

This study proposes to modify the potential vorticity (PV) field of the model initial state taking advantage of the information provided by the water vapor (WV) channel of METEOSAT-7 satellite. This technique (hereafter PV-WV technique) uses the known relation between the upper-level PV features and the WV brightness temperature field, under some assumptions, as a guidance to modify the PV field in order to reduce the mismatch between these two fields.

We have established throughout this Thesis that the uncertainty in numerical model initial conditions plays a major role in forecasting any event and that the forecast performance benefits from improvements in the quality of the initial conditions. Extreme weather events can be especially sensitive to the uncertainties of atmospheric state tampering with their predictability. Many studies have shown

*This Chapter is based on a submitted paper Vich, M., R. Romero, E. Richard, P. Arbogast and K. Maynard, 2012: *Perturbing the potential vorticity field in mesoscale forecasts of two Mediterranean heavy precipitation events.* Tellus A, Conditionally accepted.

that enhancing initial conditions accuracy results in a better forecast (e.g. Rabier et al. 1996; Nutter et al. 1998; Nuissier et al. 2007). As highlighted in previous chapters and other studies like Huo et al. (1999); Homar et al. (2003) to cite a few, forecasts are very sensitive to PV modifications. More precisely, several studies show that modifying the PV field according to WV imagery has a positive effect on reducing the forecast error (e.g. Dermitas and Thorpe 1999; Hello and Arbogast 2004; Guérin et al. 2006; Røsting and Kristjánson 2006, 2008; Manders et al. 2007; Argence et al. 2009).

11.1 Meteorological description of the events

This study targets the 9-10 June 2000 and 9-10 October 2002 high impact weather events gathered in our 19-MEDEX events database as a testbed for evaluating the effectiveness of the PV-WV technique. Both events produced large amounts of precipitation over Spain and France, up to 223 and 200 mm 24-h accumulated rainfall, respectively, and presented a similar synoptic situation.

The event of 9-10 June 2000[1] was characterized by the entrance of an Atlantic low-level cold front and an upper-level trough that contributed to the generation of a mesoscale cyclone in the Mediterranean Sea in front of the Catalonia Coast (northeast Spain; Fig. 11.1). The circulation associated with this mesoscale cyclone advected warm and moist air towards Catalonia from the Mediterranean Sea. The combination of the mediterranean air mass at low levels with the cold air aloft is reflected in Figs. 11.1.b and d as moderate values of convective instability along the Spanish mediterranean coast. The cyclone induced maritime flow together with the prominent orography of northeast Spain produced strong convergence of the water vapor flux in the lower troposphere during the whole episode (same figures). As a result of this favorable synoptic-mesoscale environment, two long-lived mesoscale convective systems developed over Catalonia and later merged and remained quasistationary nearby Barcelona city for nearly 2 hours. A more detailed diagnostic description of the event, including a sequence of radar images, can be found in Martín et al. (2007).

[1]This chapter uses the 9-10 June 2000 event as a case study in combination with another event and not only as an illustrative example (recall Chapters 6, 8 and 10). Therefore, we need to describe it more deeply than in previous chapters.

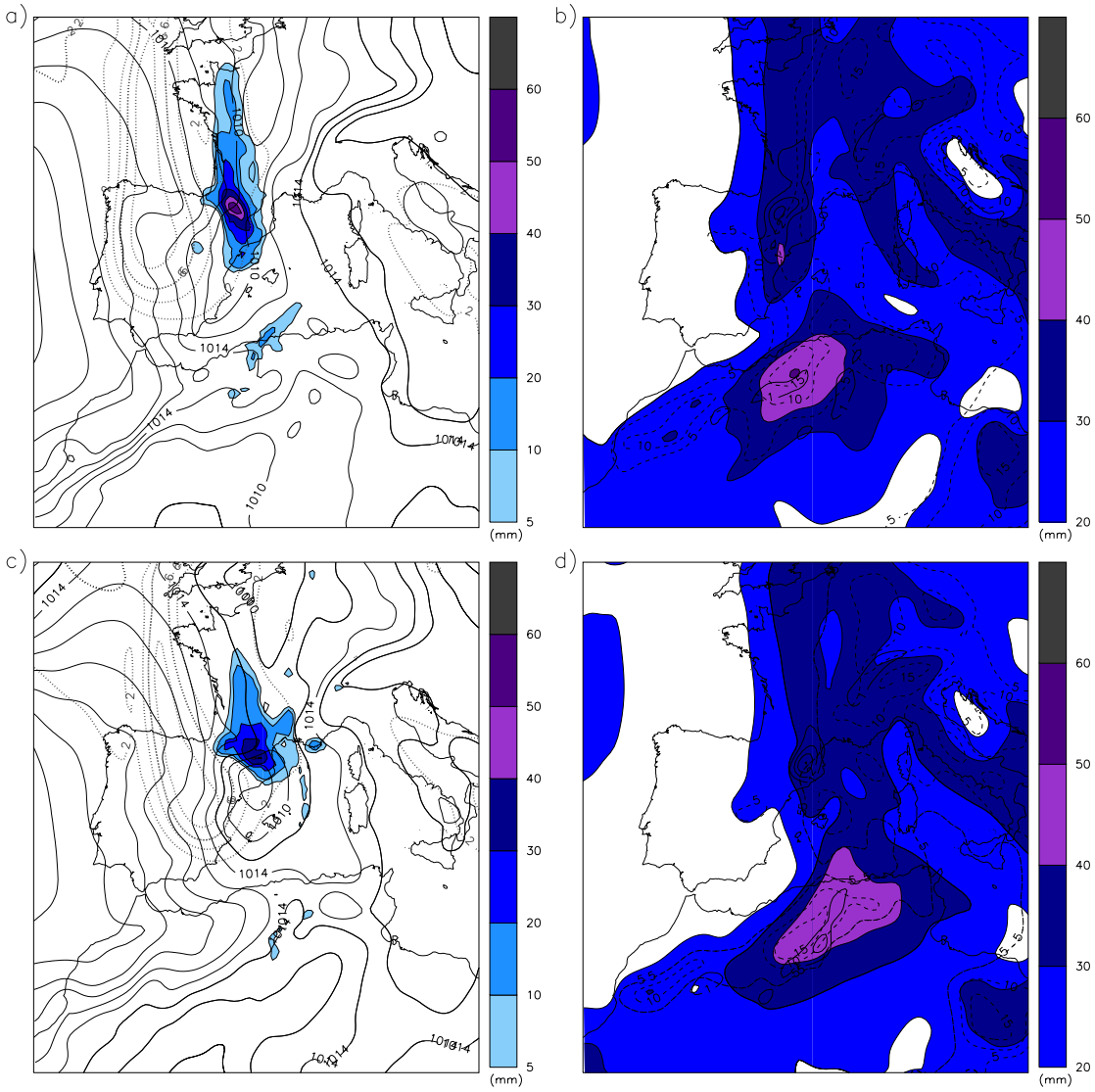


Figure 11.1. *MM5 control forecast from 00 UTC 9 June to 06 UTC 11 June 2000. (left) Potential Vorticity on the 330 K isentropic surface (dashed line, in PV units), sea level pressure (continuous line, in hPa), and 6 h accumulated rainfall (color contours, in mm according to scale) at (a) 9 June at 18 UTC and (c) 10 June at 00 UTC. (right) Water vapor flux convergence in the 1000-700 hPa (continuous line, contour interval is 1 g m^{-2} s^{-1}, starting at 1 g m^{-2} s^{-1}), convective instability (as measured by the equivalent potential temperature difference between 1000 and 500 hPa, at intervals of 5°C starting at 5°C; dashed line) and precipitable water (color contours, in mm according to scale) at (b) 9 June at 18 UTC and (d) 10 June at 00 UTC.*

Although the event of 9-10 October 2002 (Fig. 11.2) shares many large-scale characteristics with the June 2000 event, some remarkable differences can be highlighted. For example, the upper-level though has a longer wave length and remains negatively tilted with respect to the horizontal wind shear during the episode, the surface low-pressure area is much larger, spreading over the whole western Mediterranean, and its minimum is located further north (Fig. 11.2.a and c). The resulting maritime flow and thermal advection over the Catalan coast are weaker, but the impinging mediterranean winds over the south coast of France, in the form of a well-defined low-level jet (LLJ), are more notable. The diagnostic indicators summarized in Figs. 11.2.b and d evidence that, with respect to the previous event, basic ingredients for deep moist convection (large precipitable water, convective instability and low-tropospheric water vapor flux convergence) are shifted towards the northeast in this case, affecting mainly the south coast of France.

11.2 PV modifications

Briefly summarizing, the PV field is a useful meteorological parameter to study the dynamical structures at the synoptic scale thanks to the conservative and invertibility principles (Hoskins et al., 1985). In fact, both principles make the PV field a suitable tracer of upper-level dynamics, which plays an important role in mid-latitude synoptic developments. Upper-level positive PV anomalies can be interpreted as upper-level disturbances penetrating into the upper troposphere linked to the undulation of the tropopause. As Santurette and Joly (2002) showed, only one level of the dynamical tropopause is needed to diagnose upper-level dynamics if balance in the atmosphere and a monotonic vertical PV gradient are assumed, generally the 1.5 PVU surface or the PV field on a isobaric surface between 200 and 400 hPa.

The water vapor imagery provided by the 6.3 μm METEOSAT-7 channel is an important tool for synoptic-scale analyses since it mostly reveals the WV content in the mid- and upper troposphere. There is a close relationship between PV distribution and satellite WV images in dynamically active regions, so their joint examination provides an excellent framework to assess the numerical model behavior and/or analyses quality. The key point to our PV-WV technique is that the dark (light) features in the WV images can be associated to positive (negative) PV

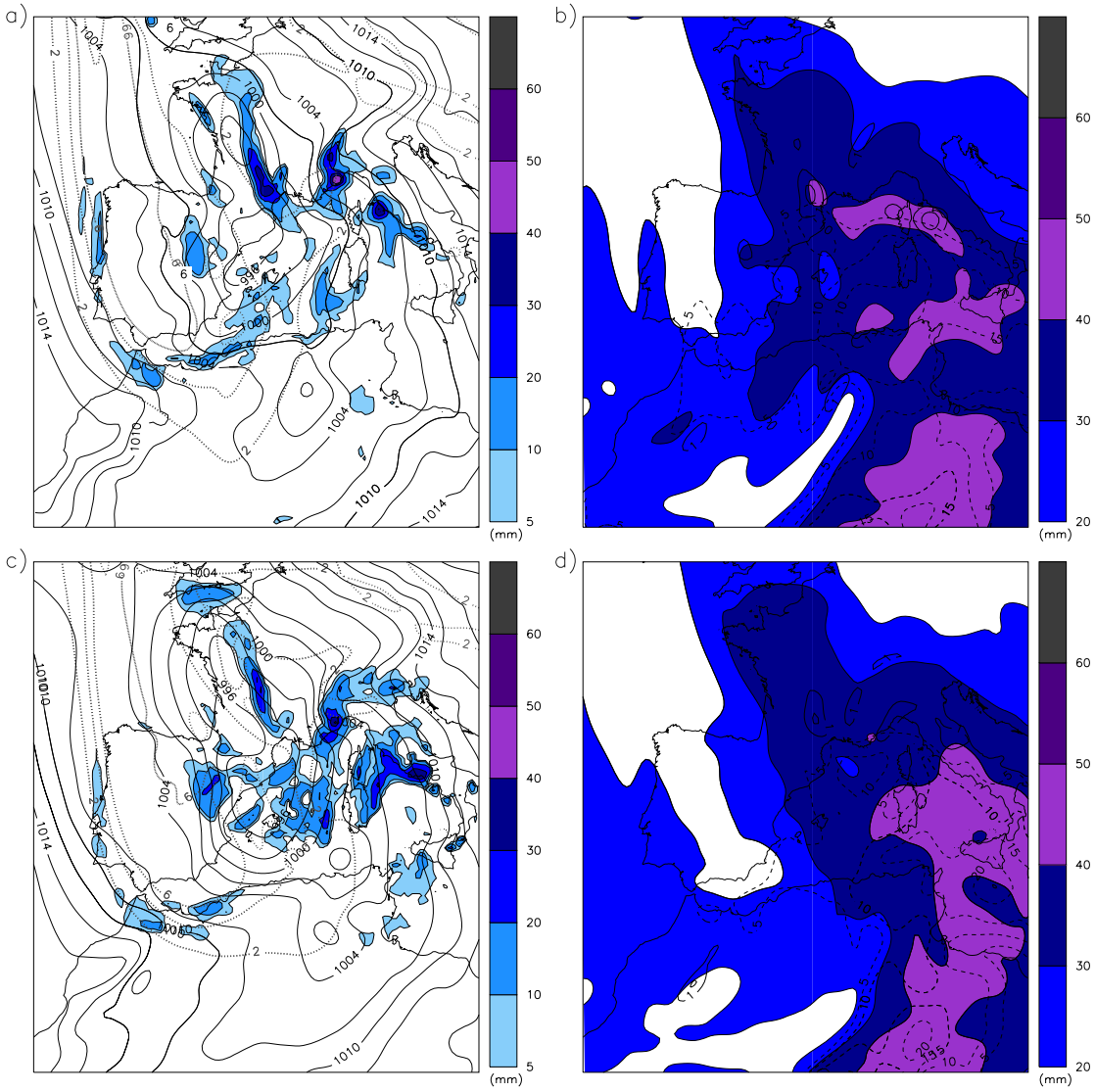


Figure 11.2. *MM5 control forecast from 00 UTC 9 October to 06 UTC 11 October 2002. (left) Potential Vorticity on the 330 K isentropic surface (dashed line, in PV units), sea level pressure (continuous line, in hPa), and 6 h accumulated rainfall (color contours, in mm according to scale) at (a) 9 October at 18 UTC and (c) 10 October at 00 UTC. (right) Water vapor flux convergence in the 1000-700 hPa (continuous line, contour interval is 1 g m^{-2} s^{-1}, starting at 1 g m^{-2} s^{-1}), convective instability (as measured by the equivalent potential temperature difference between 1000 and 500 hPa, at intervals of 5°C) starting at 5°C; dashed line) and precipitable water (color contours, in mm according to scale) at (b) 9 October at 18 UTC and (d) 10 October at 00 UTC.*

anomalies, descending (ascending) motions and low (high) geopotential heights of the dynamical tropopause. Santurette and Georgiev (2005) details the use of WV imagery side by side with the PV field, as well as an individual analysis of each of them.

Taking advantage of the known relationship between the METEOSAT-7 WV images and the PV field close to the dynamical tropopause level, some modifications can be applied to the PV field in the analyses to improve their matching. These modifications are applied to the dynamical tropopause surface defined by the height of the 1.5 PVU surface. Following the procedure used in Argence et al. (2009) and Arbogast et al. (2012) these next steps are followed:

1. Compare the dynamical tropopause (1.5 PVU surface height) with the METEOSAT-7 WV brightness temperature distribution at the simulation initial time.
2. Reduce the mismatch by modifying the topography of the dynamical tropopause accordingly.
3. For each horizontal grid point, the new vertical height of the dynamical tropopause defines a PV correction which is actually the difference between the original PV value at that height and 1.5 PVU. A 1DVAR method based on known forecast error statistics is then applied to build a vertical profile of PV modification over the corresponding grid point.
4. Invert the modified and control (unmodified) PV fields given some mass-wind balance conditions.
5. Identify the initial perturbation as the difference between the PV-control and PV-modified inverted fields.
6. Calculate the perturbed atmospheric state by adding this perturbation to the original model initial conditions.
7. Perform the corresponding numerical run using this *improved* initial state.

The PV field modifications consist of adding/subtracting PV structures as well as shifting them to minimize the mismatch between the PV field and the WV brightness temperature. It is also worth to note that these modifications are confined to the layer between 150 and 500 hPa to target the corrections on the tropopause topography. The PV inversion technique used in this Chapter is the one presented in Section 5.2 and used in Chapters 6 and 8 as part of the PV-perturbed ensembles construction.

A practical implementation of the PV-WV technique is illustrated in Fig. 11.3. This figure shows the comparison of the METEOSAT-7 WV brightness temperature and the PV field at 300 hPa for the initial time of the June 2000 case, calculated using the ECMWF grid analyses. Using the known relationship between these two fields one can decide how to modify the PV field to improve their mismatch. The locations of these modifications are indicated in Fig. 11.3 by capital letters, where G means that the initial PV gradient has been increased and C that the curvature has been increased to match better the curvature of the PV feature present on the WV image. An equivalent procedure has been done on the initial state of the October 2002 case (not shown).

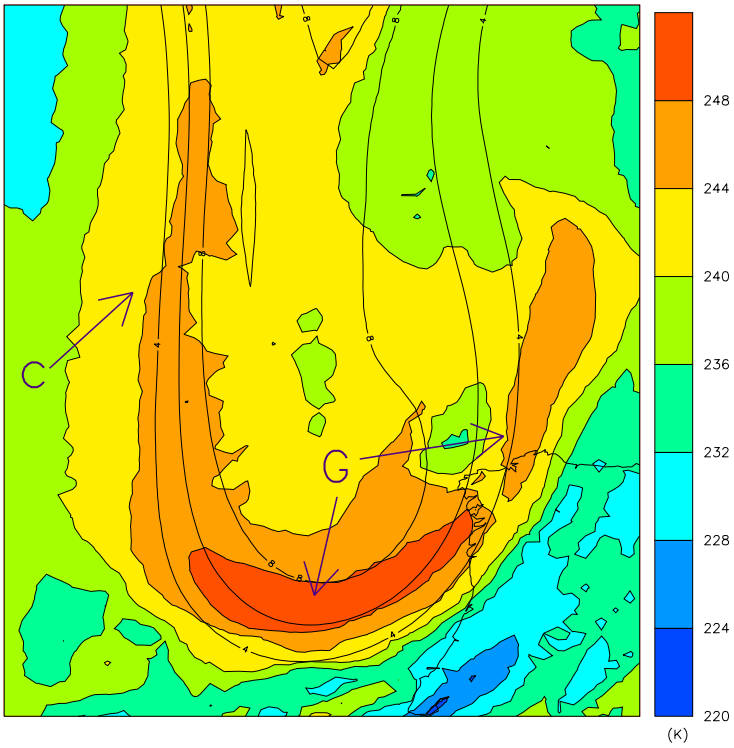


Figure 11.3. *Overlap for 9 June 2000 at 00 UTC of the potential vorticity field at 300 hPa (solid lines, contour interval 2 PVU, above 4 PVU) and METEOSAT-7 water vapor brightness temperatures (shading, in K according to scale). The letters G and C indicate the locations where the PV modifications are applied (see the details on the text).*

11.3 Results examination

Once the agreement between PV field and WV brightness temperatures is strengthened by modifying the PV field accordingly, this modified field is incorporated into the model initial state using the PV inversion technique cited before. Then two forecasts per event are run, the non-perturbed and perturbed, using the same MM5 configuration used on all PV-perturbed ensembles runs. In this study, the observations come from the climatological raingauge networks maintained by Météo-France and AEMET.

The results obtained after implementing the PV modifications are shown in Fig. 11.4 for the June 2000 case and in Fig. 11.5 for the October 2002 case. For example, in comparison with Fig. 11.1, Fig. 11.4 shows that the PV gradient around the trough at the designed points has been increased, as well as a more pronounced curvature on the western flank of the trough. Although these changes could be considered irrelevant and, in fact, the meteorological scenario has been essentially unaltered at synoptic scale, the mesoscale details do have a tangible impact on the forecast precipitation (see also Fig. 11.6). Comparatively, the PV modifications dictated by the PV-WV visual mismatch in the October 2002 case are larger (compare the Fig. 11.6.a-c evolution with the Fig. 11.2.a-c sequence) but the resulting precipitation forecast turns to be less sensitive to the initial state perturbation (Fig. 11.7). These results are consistent with the findings of Romero et al. (2005) who, based on perturbed numerical simulation of mediterranean heavy precipitation events, concluded that cases driven by sub-synoptic cyclones (like the June 2000 event) are more sensitive to the uncertainties of the precursor upper level disturbance than those cases governed by a mediterranean LLJ embedded in a large-scale low (e.g. October 2002).

Figs. 11.6 and 11.7 show the 30-54 h forecast accumulated rainfall for the non-perturbed and perturbed runs and the corresponding observation, for each case. The perturbed simulation of both cases present a forecast rainfall field closer to the observed field than the non-perturbed one, even though it does not completely match the observed pattern. A closer examination shows that the highest rainfall value for the June 2000 case predicted by the perturbed run is located further southeastwards than in the non-perturbed forecast and nearer to the observed maximum (Fig. 11.6). A similar affirmation is true for the October 2002 case (Fig. 11.7), where the maximum rainfall center (southern tip of the French

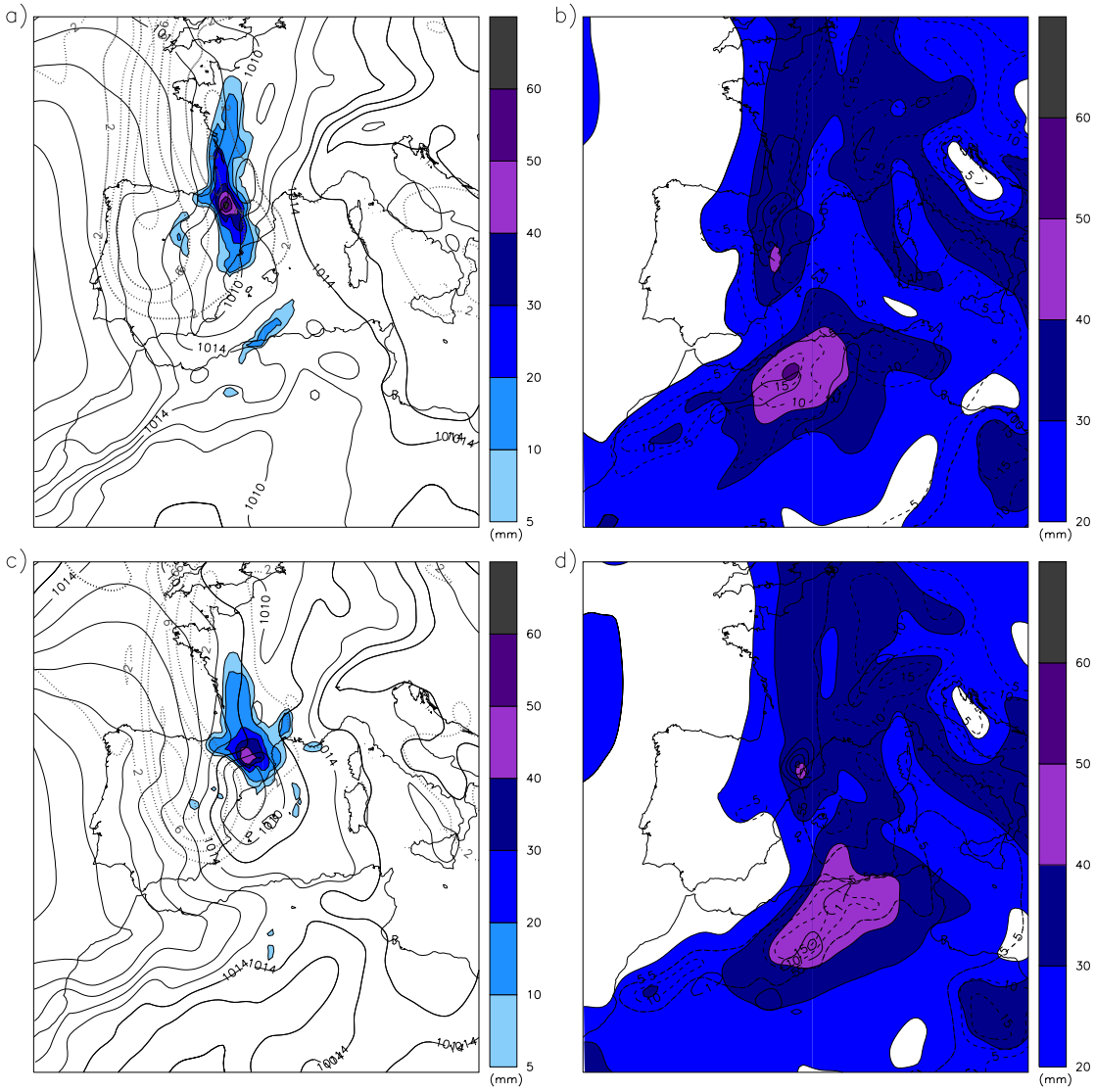


Figure 11.4. *As in Fig. 11.1 once the potential vorticity modifications are applied (a, b) 9 June at 18 UTC and (c, d) 10 June at 00 UTC.*

Riviera) from the perturbed run is shifted northeastwards with respect to the non-perturbed and nearer to the observed maximum.

A different approach for evaluating the improvement accomplished by the PV-WV technique is to use objective verification scores and compare their values with the results obtained for the same case studies by the PV modifications used on the PV-perturbed EPSs presented in Part II. Briefly, the PV-gradient ensemble applies perturbations in the areas corresponding to the PV zones of most intense values and gradients while the PV-adjoint perturbs in the MM5 adjoint model calculated

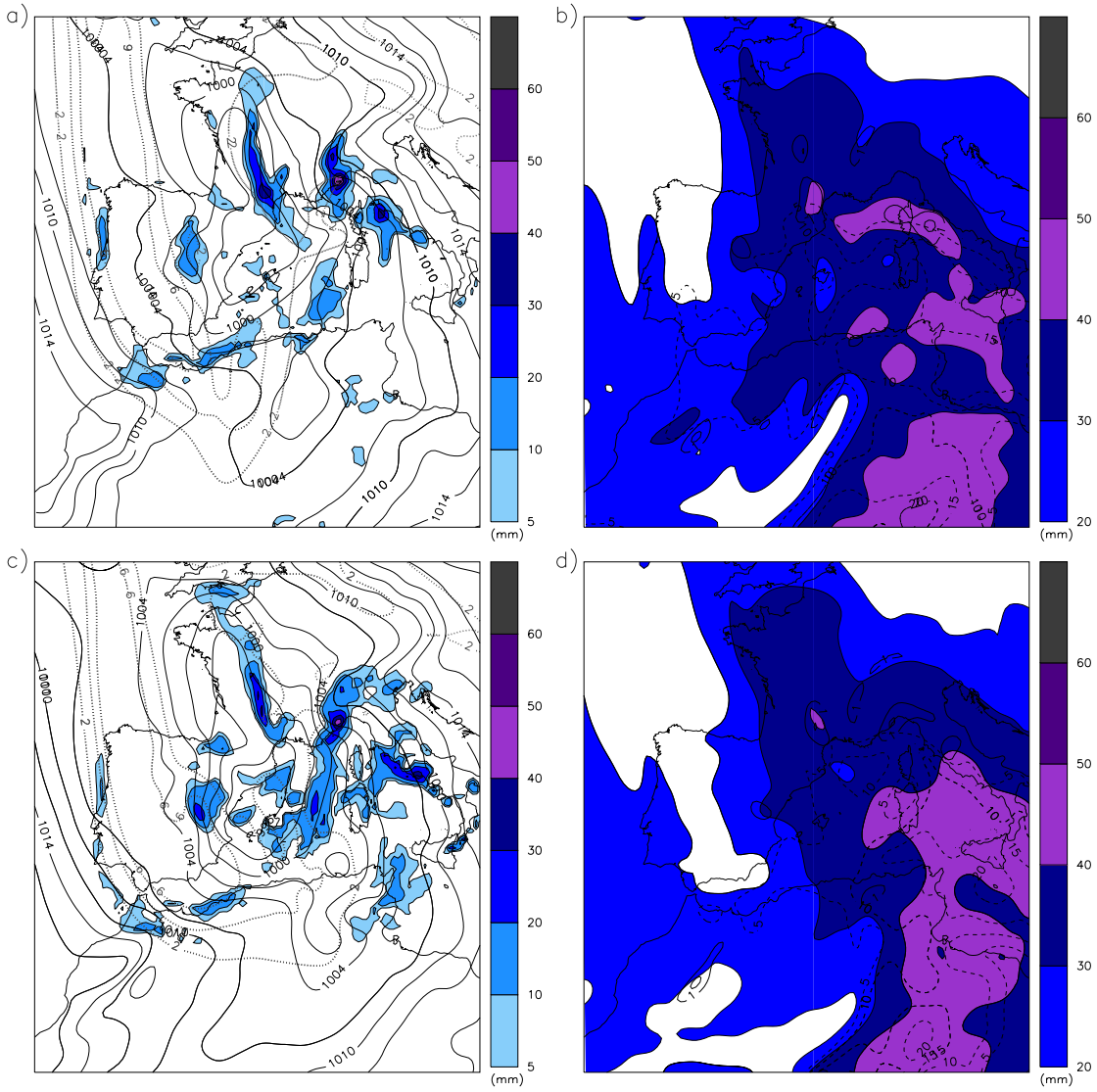


Figure 11.5. *As in Fig. 11.2 once the potential vorticity modifications are applied(a, b) 9 October at 18 UTC and (c, d) 10 October at 00 UTC.*

sensitivity zones. Both use a PV error climatology to randomly increase/decrease the intensity and displace the PV features on the regions highlighted by each method.

The ROC areas obtained by the PV-WV technique and the two ensembles for both case studies (Fig. 11.8) are very skillful since all forecasts lie well above 0.7. Moreover, on both events, the non-perturbed and the PV-WV perturbed runs lie within the range of both ensembles, the perturbed one being more skilled than the non-perturbed run. It is worth mentioning that these results correspond to

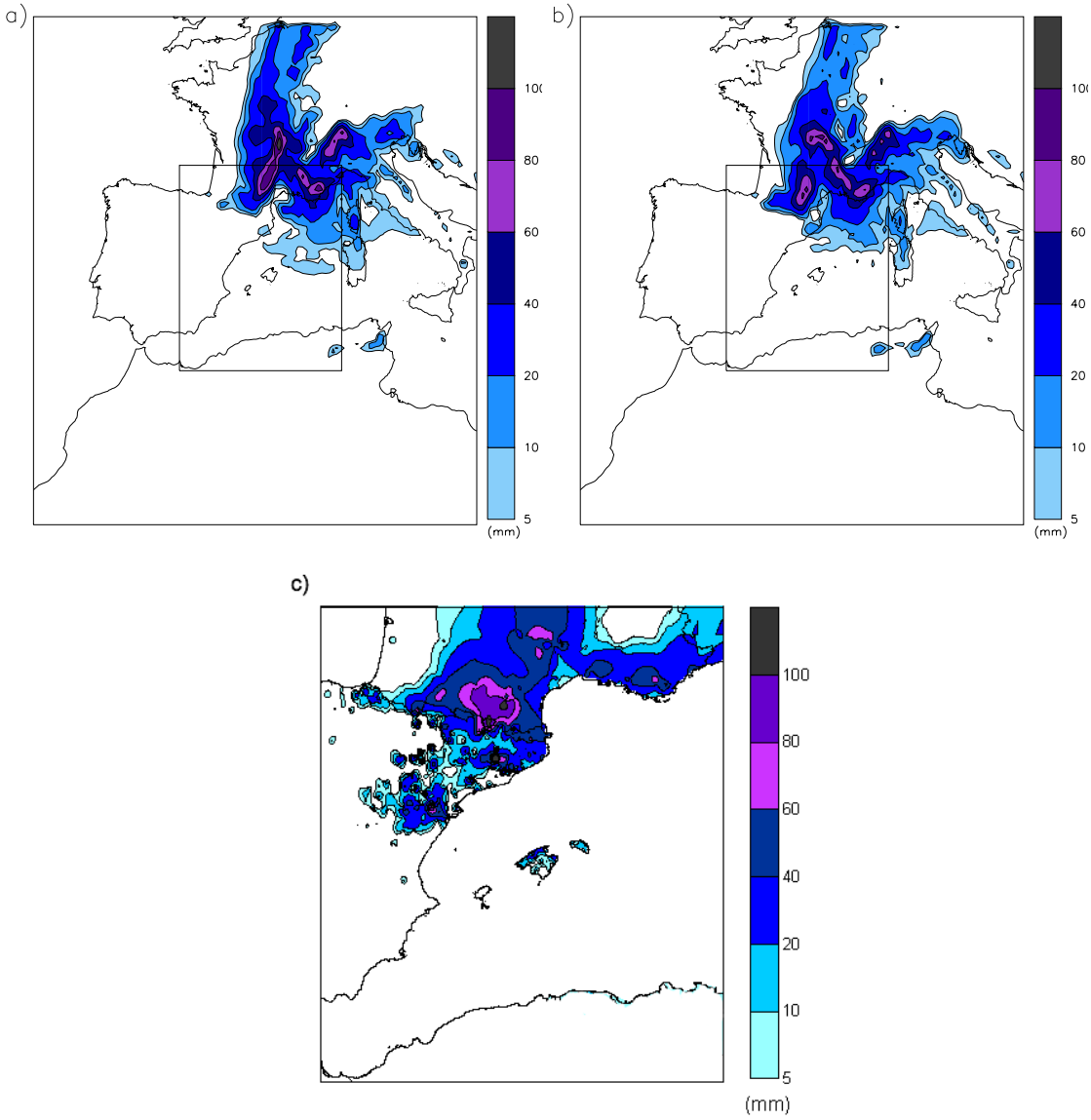


Figure 11.6. *Accumulated rainfall (color, in mm according to scale) between 10 June 2000 at 06 UTC and 11 June 2000 at 06 UTC. a) Non-perturbed run, b) PV-WV run and c) Observed.*

the 30-54 h forecast accumulated rainfall, but analogous results were found for the 06-30 h forecast period (not shown). The same behavior was observed on the rest of verification scores explored, therefore only the 30-54 h rainfall verification results are shown.

The results of comparing the observed and forecast distributions in terms of quantiles by means of a Q-Q plot (Fig. 11.9) show that, as in the ROC area results, the non-perturbed and the PV-WV perturbed runs lie within the range of both ensem-

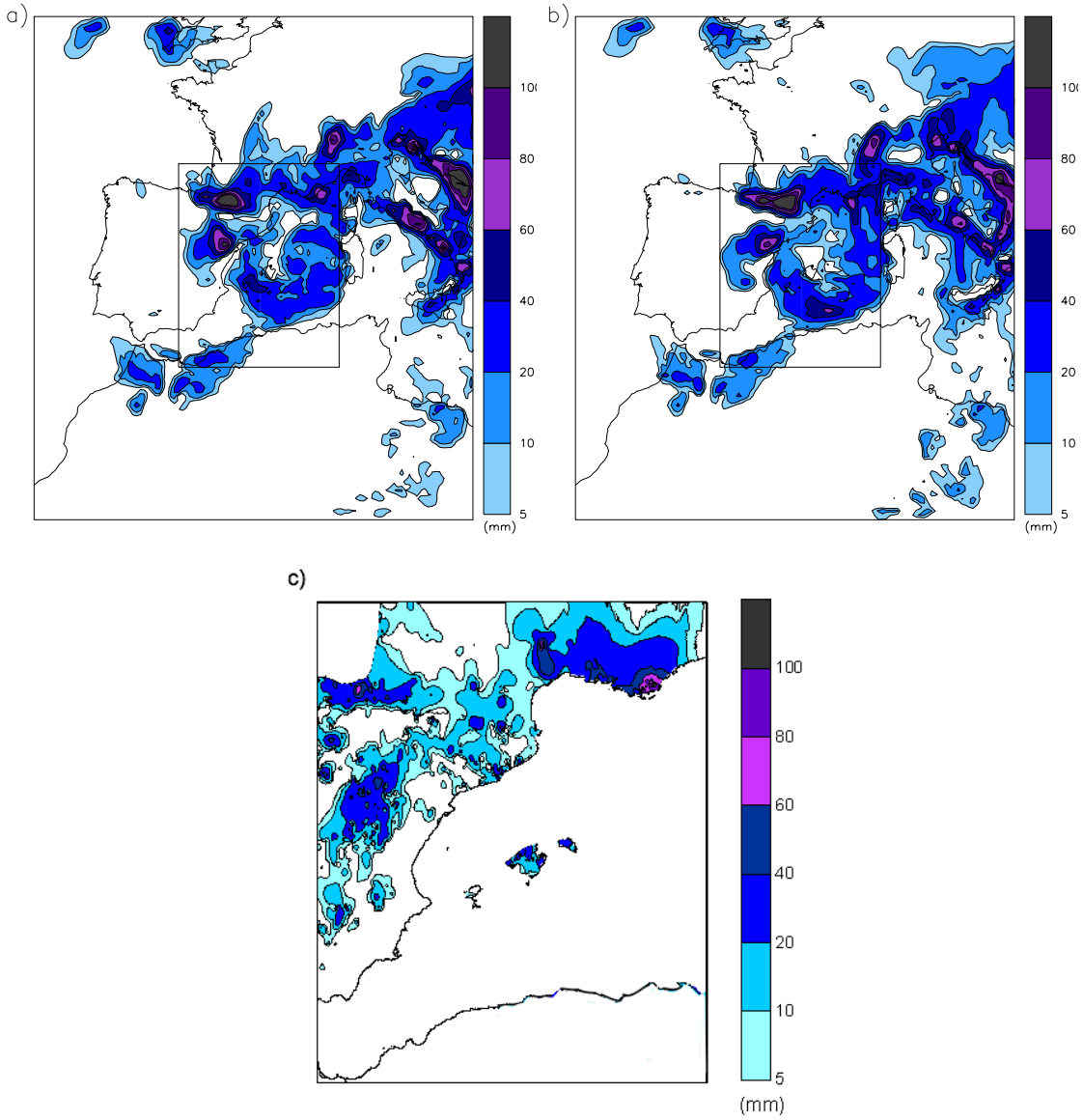


Figure 11.7. *Accumulated rainfall (color, in mm according to scale) between 10 October 2002 at 06 UTC and 11 October 2002 at 06 UTC. a) Non-perturbed, b) PV-WV run and c) Observed.*

bles, being the perturbed forecast more skilled than the non-perturbed run. Each event falls into a different region of the diagram, the June 2000 case overpredicts the observed precipitation while the October 2002 case tends to underpredicts it.

The Taylor diagram diagrams of both events (Fig. 11.10) show the same tendency seen in the other verification scores, that is the non-perturbed and the PV-WV perturbed results are contained within both ensembles results. One of the differ-

ences between the two events is the pattern displayed by the whole collection of runs in the Taylor diagram, the June 2000 pattern is more elongated than the October 2002 cluster of positions meaning that the almost all June 2000 runs exhibit a RMS difference around 9 mm, while the RMS difference shown by the October 2002 runs range from 8 to 18 mm. Another difference between the cases is in the correlation coefficient, on the June 2000 events it ranges from 0.4 to 0.7, higher values that those obtained for the October 2002 event (from 0.15 to 0.55). Even though the observed precipitation field is different in each event, both exhibit a similar standard deviation almost 10 mm on the June 2000 event and around 9 mm on the October 2002 case. The forecasts standard deviations range from 0 to 15 mm for the June 2000 case and from 8 to 18 mm for the October 2002 event. It is also worth mentioning that the statistics displayed on the Taylor diagram are negatively affected by the discontinuities, noise and outliers typically present on the rainfall field.

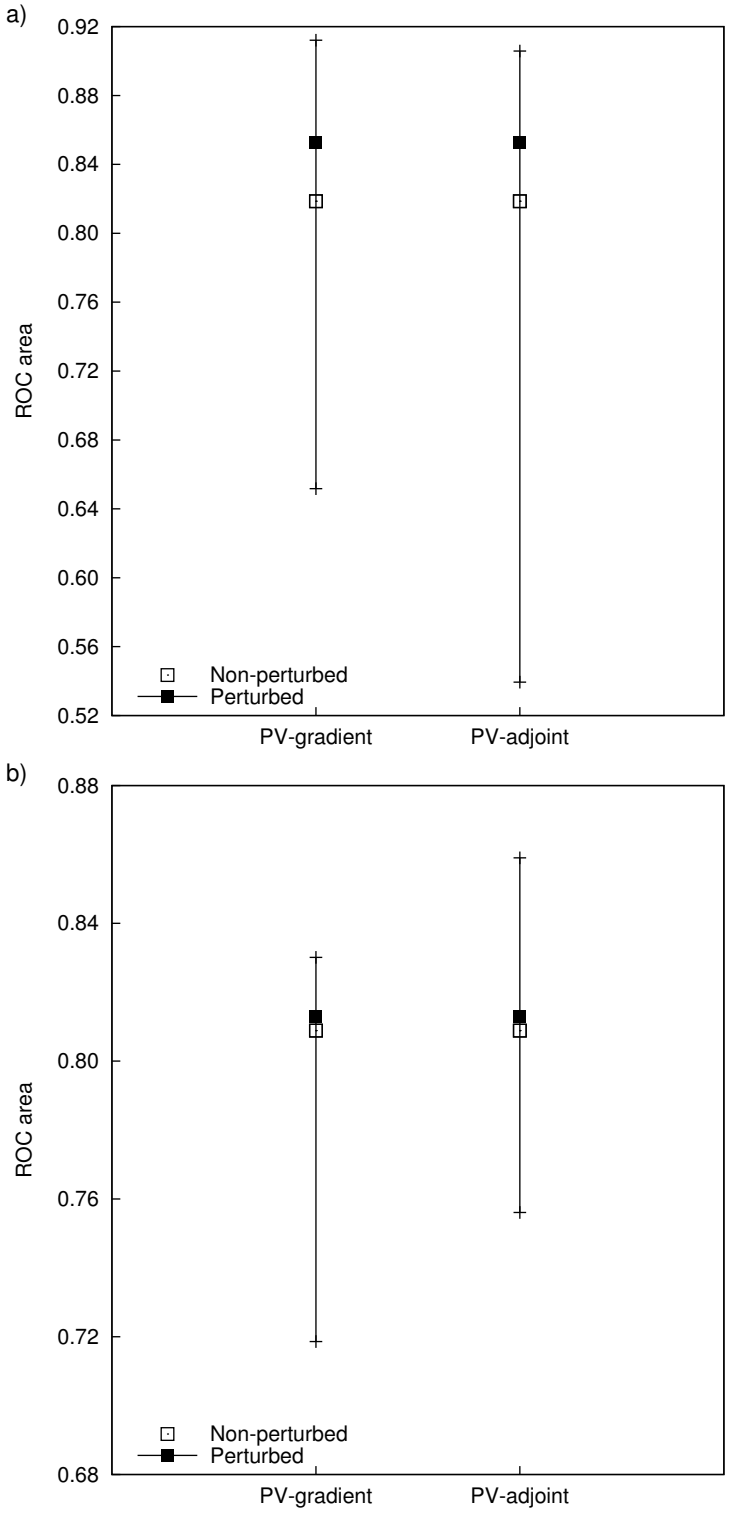


Figure 11.8. *ROC area range for the PV-gradient and PV-adjoint ensemble members, where the empty square corresponds to the non-perturbed run and the filled square to the PV-WV perturbed run. a) June 2000 event and b) October 2002 case. The vertical lines represent these ROC area ranges where the extremes, maximum and minimum, of the range correspond to the ensemble member associated with the highest/lowest value of ROC area.*

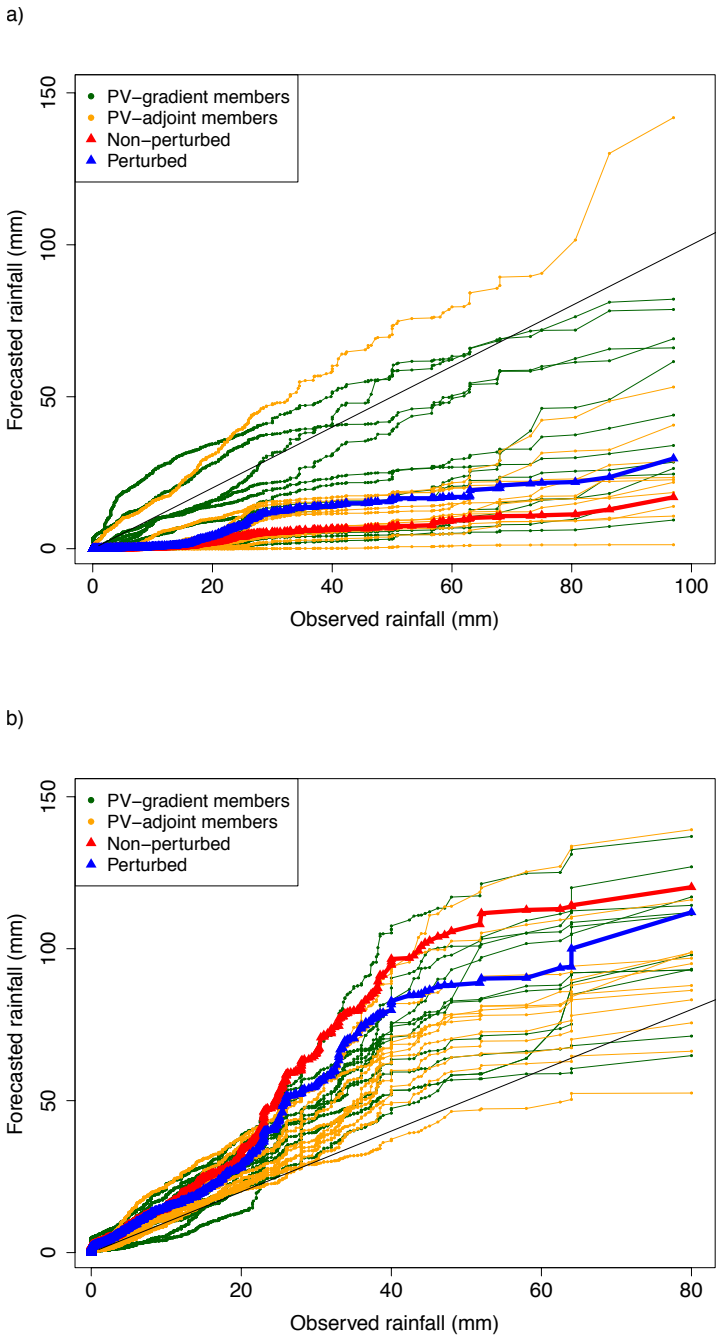


Figure 11.9. *Q-Q plot for the PV-gradient and PV-adjoint ensemble members, the non-perturbed and the PV-WV perturbed run. a) June 2000 event and b) October 2002 case. The diagonal line represents the perfect score and the region below (above) the diagonal indicates that the forecast underestimates (overestimates) the observation.*

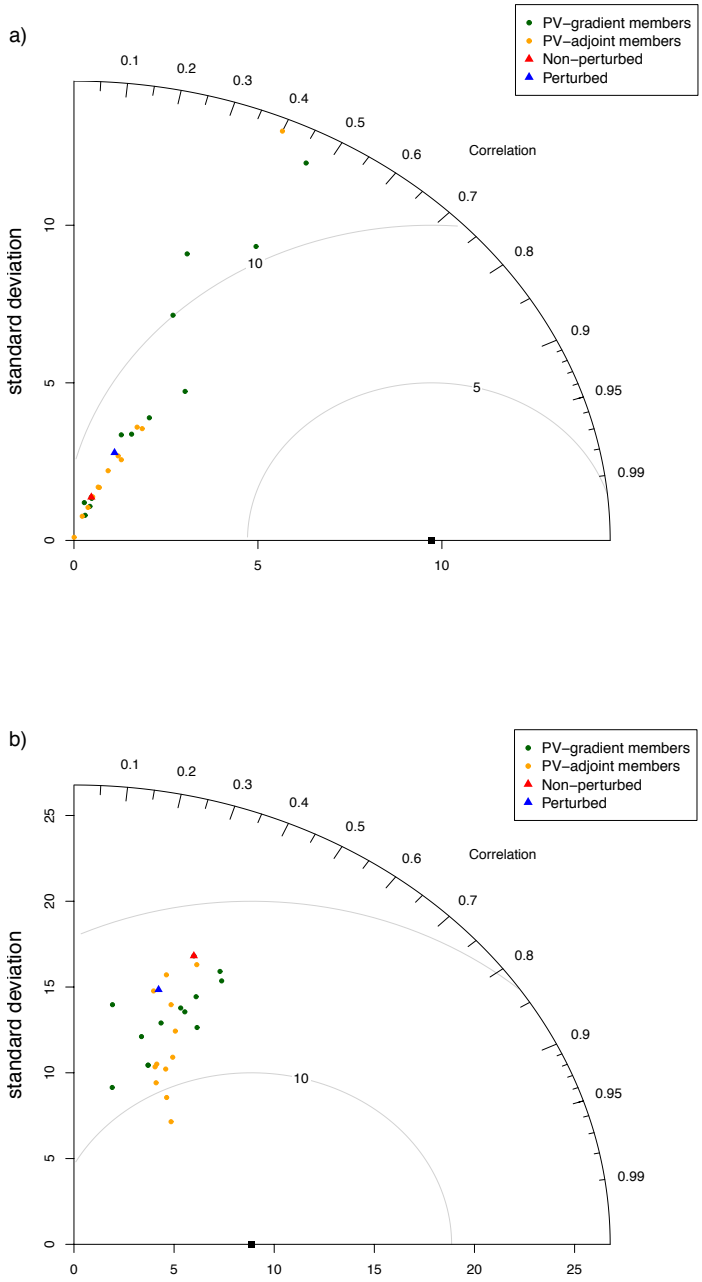


Figure 11.10. *Taylor diagram for the PV-gradient and PV-adjoint ensemble members, the non-perturbed and the PV-WV perturbed run. a) June 2000 event and b) October 2002 case. The filled square corresponds to the observations. The perfect score is obtained when the data point representing the forecast field matches up with the observed one (solid square).*

11.4 Conclusions

This Chapter implements a PV modifying technique applied on the initial state of two Mediterranean high-impact weather events and tests its impact on the corresponding mesoscale forecasts. The applied technique is based on correcting the mismatch between the upper-level PV field and the observed WV brightness temperature given by the METEOSAT-7 satellite. The forecast precipitation fields, our feature of interest due to its regional socio-economical impact, obtained after perturbing the initial state show a clear improvement compared with the non-perturbed forecast, even though there seems to be still room for improvement on the two selected events.

The examined verification scores also reflect the improvement in skill when the initial state of the event is perturbed using the PV-WV technique described. In fact this improvement is observed on both events and on all the verification scores presented, the perturbed run always achieves a higher score than the non-perturbed run. When the perturbed and non-perturbed runs are compared with the PV-adjoint and PV-gradient ensemble members, they remain within the range of both EPS members scores and are statistically indistinguishable.

Conclusions and further work

In this Thesis, we have focused on the characteristic high impact weather events that often affect the western Mediterranean region. To improve the capability in forecasting these phenomena is of paramount importance, provided the high impact these events have on our society.

We have developed this task by designing three different ensemble prediction systems and by testing them for this type of events. The multiphysics ensemble focuses on the uncertainties derived from model deficiencies while the PV-gradient and PV-adjoint ensembles target the uncertainties of the initial and boundary conditions. These PV-perturbed ensembles exploit the connection between PV structures and high impact weather events associated with cyclones. This fact permits to implement a simple perturbation method based on perturbing a single variable, i.e., the PV field, that ensures consistent perturbations of all the meteorological fields without compromising the mass-wind balance. In addition, to ensure consistency between the PV perturbations and the PV field uncertainty range we have developed a PV error climatology that assesses these PV errors present in the initial and boundary conditions. The two PV-perturbed ensembles differ by the fact that, the PV-gradient ensemble basically perturbs the PV zones of most intense values and gradients (in essence a subjective choice based on our experience) while the PV-adjoint ensemble perturbs along the sensitivity zones calculated by the MM5 adjoint model (an objective method).

During the ensembles verification, the precipitation field picked our interest because of its direct impact on society. The downside of this choice is that the precipitation field has both strong gradients in space and rapid variations in time, and is observed over irregularly spaced networks. These facts cause the verification procedure to be difficult, i.e., it is hard to obtain good verification scores when dealing with precipitation. In addition, we cannot forget that we are dealing with

extreme events often linked with observational sample problems and with the difficulties of the model to forecast extreme precipitations. The verification testbed is basically fed by the MEDEX cyclones database. All the used cyclones caused heavy precipitation over the western Mediterranean and were often accompanied by strong winds.

In spite of the aforementioned difficulties, the verification procedure ratified the advantages of an EPS over a deterministic forecast and highlighted the skill of all three ensembles to provide a useful probabilistic prediction for our testbed. It has also pointed out that the PV-gradient EPS performs slightly better than the multiphysics ensemble. Regarding the two PV-perturbed ensembles, we find that the high computational cost of the PV-adjoint ensemble (which implies running the MM5 adjoint model for each simulation day) versus the low cost of the PV-gradient is not compensated later in ensemble skill. Even though both PV-perturbed EPS perform more than adequate, the PV-gradient EPS is generally more skillful than PV-adjoint ensemble. This fact turns the PV-gradient ensemble into a more favorable strategy since it provides a more useful forecast than the PV-adjoint at a better cost, at least when our testbed is concerned. This conclusion is not surprising if we take into account that, using a similar testbed, Homar and Stensrud (2008) concluded that the adjoint-estimated sensitivity is comparable or even slightly inferior to subjective (gradient and human) sensitivity estimates.

To further explore the potential of the three developed ensembles, we have tested their capabilities on three different applications. Next, we recapitulate our findings regarding these applications:

1. Using the multiphysics ensemble to feed a superensemble calibration did not result as profitable as expected. For our testbed, the results show that a cheaper approach like a bias corrected ensemble mean is as beneficial as the superensemble one. We propose several further studies that may help to track down the reasons behind these results. For example, to repeat the experiments feeding the superensemble with a multimodel ensemble in which its members are less correlated than on multphysics ensemble by construction. Another possibility is to redo the experiment using the ERA-INTERIM fields as initial and boundary conditions in order to avoid any possible influence of the updates undergone by the ECMWF forecasts between the training period and earlier MEDEX cases.

2. In the second application, we used the PV-gradient ensemble to introduce the initial and boundary conditions uncertainties into a hydrometeorological forecasting chain. This procedure gave very successful results. The hydrometeorological runoff simulations driven by the quantitative precipitation forecasts provided by the PV-gradient ensemble generally performed better than deterministic simulations. This is so because an ensemble, by definition, is better equipped to deal with the intrinsic difficulties in determining accurately the intensity, timing and location of the precipitation.

3. The third application was designed to further explore the potential exhibited by the PV perturbing approach. To do so, we chose a method that modifies the PV field using the METEOSAT-7 satellite water vapor (WV) channel information to decide where and in which amount the PV field has to be perturbed. This technique proved to be worthy on both examined events. The PV-WV perturbed run performed better than a non-perturbed run. Moreover, when compared with the PV-gradient and PV-adjoint ensembles members, the PV-WV technique run remains within the range of both ensembles members scores and it is statistically indistinguishable. In the future it seems appropriate to explore whether the PV-WV technique outputs lie systematically within the range of best ensembles members and whether this technique can be used to generate additional ensemble members that would increase the ensemble spread. These future checks should be done statistically over a large number of events. In this sense, we should at least repeat the experiment for the whole set of MEDEX cases used in the PV-gradient and PV-adjoint ensemble generation.

In view of all the results summarized above, a plausible future step would consist in designing a hybrid approach as, e.g., that in Meng and Zhang (2007) where different physical parameterizations are introduced after perturbing the initial state. An ensemble made up of the multiphysics, PV-gradient and PV-adjoint ensemble members should increase the accuracy and usefulness of the ensemble by creating greater divergence in the ensemble trajectories (see, e.g., Stensrud et al. (2000), among others). On the downside, this mixing of uncertainty sources makes it very difficult to investigate each error type separately, as done here. Another feasible future step is to expand the current testbench; statistically speaking, an increase in case studies leads to more robust and general conclusions.

Regarding the PV perturbation line of work, we should continue exploring additional ensemble generation methods based upon perturbing the initial and boundary conditions through the PV field. After this Thesis experience, we suggest a PV perturbation method that targets the precursor upper-level trough characteristic of the mid-latitude cyclonic situations as in our PV-gradient ensemble, since it was shown that these are the most fruitful dynamical structures. Another strategy may be based on *flow dependent* PV perturbation amplitudes. Our current set up can be categorized as partially flow dependent, in the sense that while the perturbed zones depend on the situation/event, the PV perturbation amplitudes do not. Instead, they rely on a fixed climatology, i.e., the PV error climatology. Flow dependent PV perturbations (both in amplitude and location) will require an appropriate estimate of the PV error for each case study. Some ideas on how to obtain this PV analysis error are: 1) to use a proxy like the standard deviation of an operational ensemble, for example the one run by the ECMWF, and 2) to use a data assimilation system that by definition will provide a flow dependent initial time PV analysis error.

Finally, bearing in mind a more operational framework, campaigns like the forthcoming HyMeX 2012 Special Observation Period (SOP; see `http://www.hymex.org` for details) provide an excellent real-time framework to further explore the capabilities and application of perturbation techniques like the ones studied here. Moreover, at the time this Thesis was finished, we have already begun to implement our three ensembles operationally. Once they are ready, the results will be added to the UIB Meteorology Group-generated daily forecast products, and publicly available on a web-page similar to `http://mm5forecasts.uib.es/`.

Bibliography

ACA, Agència Catalana de l'Aigua., 2001: *Delimiting of flood plain for the draft of the INUNCAT. Internal Basins of Catalonia. Hydrological calculations and determination of the flood wave discharge, Vol. 2.* Departament de Medi Ambient, Generalitat de Catalunya, 88 pp.

Amengual, A., R. Romero, M. Gómez, A. Martín, , and S. Alonso, 2007: A hydrometeorological modeling study of a flash flood event over Catalonia, Spain. *J. Hydrometeor.*, **8**, 282–303.

Amengual, A., R. Romero, M. Vich, and S. Alonso, 2009: Inclusion of potential vorticity uncertainties into a hydrometeorological forecasting chain: application to a medium size basin of Mediterranean Spain. *Hydrol. Earth Syst. Sci.*, **13**, 793–811. doi: 10.5194/hess-13-793-2009.

Anderson, M. L., Z. Q. Chen, M. L. Kavvas, and A. Feldman, 2002: Coupling HEC-HMS with atmospheric models for prediction of watershed runoff. *J. Hydrol. Eng.*, **7**, 312–318.

Andersson, E., J. Haseler, P. Undén, P. Courtier, G. Kelly, D. Vasiljevic, C. Brankovic, C. Gaffard, A. Hollingsworth, C. Jakob, P. Janssen, E. Klinker, A. Lanzinger, M. Miller, F. Rabier, A. Simmons, B. Strauss, P. Viterbo, C. Cardinali, and J.-N. Thépaut, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III: Experimental results. *Quart. J. Roy. Meteor. Soc.*, **124**, 1831–1860. doi:10.1002/qj.49712455004.

Arbogast, P., K. Maynard, and F. Crepin, 2008: Ertel potential vorticity inversion using a digital filter initialization method. *Quart. J. Roy. Meteor. Soc.*, **134**, 1287–1296. doi:10.1002/qj.283.

Arbogast, P., K. Maynard, and C. Fougère-Piriou, 2012: Manual PV modifications; a measure of forecaster's expertise. *Wea. Forecasting.* Under revision.

Argence, S., D. Lambert, E. Richard, J.-P. Chaboureau, and N. Söhne, 2008: Impact of initial condition uncertainties on the predictability of heavy rainfall in the Mediterranean: a case study. *Quart. J. Roy. Meteor. Soc.*, **134**, 1775–1788. doi:10.1002/qj.314.

Argence, S., D. Lambert, E. Richard, J. Pierre Chaboureau, J. Philippe Arbogast, and K. Maynard, 2009: Improving the numerical prediction of a cyclone in the Mediterranean by local potential vorticity modifications. *Quart. J. Roy. Meteor. Soc.*, **135**, 865–879. doi:10.1002/qj.422.

Baldwin, M. and J. Kain, 2006: Sensitivity of several performance measures to displacement error, bias, and event frequency. *Wea. Forecasting.*, **21**, 636–648.

Baxter, M. A., P. N. Schumacher, and J. M. Boustead, 2011: The use of potential vorticity inversion to evaluate the effect of precipitation on downstream mesoscale processes. *Quart. J. Roy. Meteor. Soc.*, **137**, 179–198. doi: 10.1002/qj.730.

Beare, R., A. Thorpe, and A. White, 2003: The predictability of extratropical cyclones: Nonlinear sensitivity to localized potential vorticity perturbations. *Quart. J. Roy. Meteor. Soc.*, **129**, 219–237.

Benoit, R., N. Kouwen, W. Yu, S. Chamberland, and P. Pellerin, 2003: Hydrometeorological aspects of the Real-Time Ultrafinescale Forecast Support during the Special Observing Period of the MAP. *Hydrol. Earth Syst. Sci.*, **7**, 877–889. URL `http://www.hydrol-earth-syst-sci.net/7/877/2003/`.

Berrisford, P., P. Kållberg, S. Kobayashi, D. Dee, S. Uppala, A. J. Simmons, P. Poli, and H. Sato, 2011: Atmospheric conservation properties in ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **137**, 1381–1399. doi:10.1002/qj.864.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420. doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Bluestein, H. B., 1993: *Synoptic-dynamic meteorology in midlatitudes*, volume I and II. Oxford University Press.

Bouttier, F. and P. Courtier, 1999: *Data assimilation concepts and methods.* ECMWF Training Courses Notes.

Brier, G. W., 1950: Verification of forecats expressed in terms of probabilities. *Mon. Wea. Rev.*, **78**, 1–3.

Buizza, R., 2001: Chaos and weather prediction - A review of recent advances in Numerical Weather Prediction: Ensemble forecasting and adaptive observation targeting. *Nuovo Cimento C Geophysics Space Physics C*, **24**, 273–302.

Buizza, R., J. Barkmeijer, T. N. Palmer, and D. S. Richardson, 2000: Current status and future developments of the ECMWF Ensemble Prediction System. *Meteor. Appl.*, **7**, 163–175.

Buizza, R., A. Hollingsworth, F. Lalaurette, and A. Ghelli, 1999: Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System. *Weather and Forecasting*, **14** (2), 168–189. doi: 10.1175/1520-0434(1999)014<0168:PPOPUT>2.0.CO;2.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908. doi:10.1002/qj.49712556006.

Buizza, R. and T. N. Palmer, 1995: The singular-vector structure of the atmospheric general circulation. *J. Atmos. Sci.*, **52**, 1434–1456.

Buizza, R. and T. N. Palmer, 1998: Impact of Ensemble Size on Ensemble Prediction. *Mon. Wea. Rev.*, **126**, 2503. doi:10.1175/1520-0493(1998)126<2503:IOESOE>2.0.CO;2.

Callado, A., C. Santos, P. Escribà, D. Santos-Muñoz, J. Simarro, and J. A. García-Moya, 2011: Performance of multi-model AEMET-SREPS precipitation probabilistic forecasts over Mediterranean area. *Adv. Geosci.*, **26**, 133–138. doi:10.5194/adgeo-26-133-2011.

Cane, D. and M. Milelli, 2005: Use of Multimodel SuperEnsemble Technique for Mountain-area weather forecast in the Olympic Area of Torino. *Croat. Met. J.*, **40**, 215–219.

Cane, D. and M. Milelli, 2006: Weather forecast with Multimodel SuperEnsemble Techinique in a complex orography region. *Meteorol. Z.*, **15(2)**, 1–8.

Cane, D. and M. Milelli, 2010: Multimodel SuperEnsemble techinique for quantitative precipitation forecasts in Piemonte region. *Nat. Haz. and Earth. Syst. Sci.*, **10**, 265–273.

Casati, B., L. Wilson, D. Stephenson, P. Nurmi, A. Ghelli, M. Pocernich, U. Damrath, E. Ebert, B. Brown, and S. Mason, 2008: Forecast verification: Current status and future directions. *Meteor. Appl.*, **15**, 3–8.

Charney, J. G., 1955: The use of primitive equation of motion in numerical prediction. *Tellus*, **7**, 22–26.

Cohuet, J. B., R. Romero, V. Homar, V. Ducrocq, and C. Ramis, 2011: Initiation of a severe thunderstorm over the Mediterranean Sea. *Atmos. Res.*, **100**, 603–620. doi:10.1016/j.atmosres.2010.11.002.

Collins, M., S. F. B. Tett, and C. Cooper, 2001: The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dynam.*, **17**, 61–81. doi:10.1007/s003820000094.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1807. doi:10.1002/qj.49712455002.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387. doi:10.1002/qj.49712051912.

Davis, C. A. and K. A. Emanuel, 1991: Potential vorticity diagnostics of cyclogenesis. *Mon. Wea. Rev.*, **119**, 1928–1953.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597. doi:10.1002/qj.828.

Dermitas, M. and A. Thorpe, 1999: Sensitivity of Short-Range Weather Forecasts to Local Potential Vorticity Modifications. *Mon. Wea. Rev.*, **127**, 922–939.

Diaconis, P. and B. Efron, 1983: Computer-intensive methods in statistics. *Sci. Am.*, **248**, 116–130.

Du, J., S. 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 Expriment Using a Mesoscale Two-Dimensional Model. *J. Atmos. Sci.*, **46**, 3077–3107. doi:10.1175/1520-0469(1989)046(3077:NSOCOD)2.0.CO;2.

Dudhia, J., 1993: A non-hydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**, 1493–1513.

Dudhia, J., 1996: A multi-layer soil temperature model for MM5. *Preprints from the Sixth PSU/NCAR Mesoscale Model Users' Workshop 22-24*. `http://www.mmm.ucar.edu/mm5/lsm/soil.pdf`.

Dudhia, J., D. Gill, K. Manning, W. Wang, and C. Bruyere, 2005: *PSU/ NCAR mesoscale modeling system tutorial class notes and user's guide: MM5 modeling system version 3.7*. NCAR Tech. `http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html`.

Ebert, E. and J. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, **239**, 179–202.

Eckel, F. and C. Mass, 2005: Aspects of effective mesoscale, short-range ensemble forecasting. *Wea. Forecasting.*, **20**, 328–350.

Ehrendorfer, M., 1994: The Liouville equation and its potential use fulness for the prediction of forecast skill. Part I: Theory. *Mon. Wea. Rev.*, **122**, 703–713.

Errico, R. M., 1997: What is an adjoint model? *Bull. Amer. Meteor. Soc.*, **78**, 2577–2591.

Errico, R. M. and T. Vukićević, 1992: Sensitivity analysis using an adjoint of the PSU-NCAR mesoscale model. *Mon. Wea. Rev.*, **120**, 16644–1660.

Ertel, H., 1942: Ein neuer hydrodynamischer wirbelsatz. *Meteor. Z.*, **59**, 271–281.

Evans, R., S. Harrison, R. Graham, and K. Mylne, 2000: Joint medium-range ensembles from the Met Office and ECMWF systems. *Mon. Wea. Rev.*, **128**, 3104–3127.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162. doi:10.1029/94JC00572.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean. Dyn.*, **53**, 343–367.

Evensen, G. and P. J. van Leeuwen, 1996: Assimilation of Geosat Altimeter Data for the Agulhas Current Using the Ensemble Kalman Filter with a Quasigeostrophic Model. *Mon. Wea. Rev.*, **124**, 85. doi:10.1175/1520-0493(1996)124<0085:AOGADF>2.0.CO;2.

Ferraris, L., R. R., and F. Siccardi, 2002: The uncertainty in the prediction of flash floods in the Northern Mediterranean environment. *J. Hydrometeor.*, **3**, 714–727.

Fisher, M., 2003: Background error covariance modelling. *In Proceedings of the ECMWF Seminar on Recent developments in Data Assimilation for Atmosphere and Ocean*, pages 8–12 September 2003, Reading, UK, 45–63.

Frank, W. M., 1983: The cumulus parameteritzation problem. *Mon. Wea. Rev.*, **111**, 1859–1871.

Fritsch, J. M., J. Hilliker, J. Ross, and R. L. Vislocky, 2000: Model Consensus. *Wea. Forecasting.*, **15**, 571–582. doi:10.1175/1520-0434(2000)015<0571:MC>2.0.CO;2.

Fujita, T., D. Stensrud, and D. Dowell, 2007: Surface data assimilation using ensemble Kalman filter approach with initial condition and model physics uncertainties. *Mon. Wea. Rev.*, **135**, 1846–1868.

Gandin, L. S., 1963: Objective Analysis of Meteorological Fields . *Hydromet Press, Leningrad.*

García-Moya, J.-A., A. Callado, P. Escribà, C. Santos, D. Santos-Muñoz, and J. Simarro, 2011: Predictability of short-range forecasting: a multimodel approach. *Tellus A*, **63**, 550–563. doi:10.1111/j.1600-0870.2010.00506.x.

Garcies, L. and V. Homar, 2009: Ensemble sensitivities of the real atmosphere: application to Mediterranean intense cyclones. *Tellus A*, **61**, 394–406. doi:10.1111/j.1600-0870.2009.00392.x.

Gauthier, P., C. Charette, L. Fillion, P. Koclas, and S. Laroche, 1999: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmos.–Ocean*, **37**, 103–156.

Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada. *Mon. Wea. Rev.*, **135**, 2339. doi:10.1175/MWR3394.1.

Gauthier, P. and J.-N. Thépaut, 2001: Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.*, **129**, 2089–2102.

Gilleland, E., D. Ahijevych, B. Brown, B. Casati, and E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting.*, **24**, 1416–1430.

Glahn, H. R. and D. A. Lowry, 1972: The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *J. Appl. Meteor.*, **11**, 1203–1211. doi:10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Golub, G. H. and C. F. Van Loan, 1983: *Matrix computation.* North Oxford Academic Publ. Co. Ltd. pp. 476.

Grell, G. and D. Devenyi, 2002: A generalized approach to parameterizating convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**, 1693, doi:10.1029/2002GL015311. doi:10.1029/2002GL015311.

Grell, G., J. Dudhia, and D. R. Stauffer, 1995: *A description of the fifth-generation of the Penn State/NCAR mesoscale model (MM5).* NCAR Tech. NCAR/TN-398+STR.

Guérin, R., G. Desroziers, and P. Arbogast, 2006: 4D-Var analysis of potential vorticity pseudo-observations. *Quart. J. Roy. Meteor. Soc.*, **132**, 1283–1298.

Hamill, T. M. and C. Snyder, 2000: A Hybrid Ensemble Kalman Filter 3D Variational Analysis Scheme. *Mon. Wea. Rev.*, **128**, 2905. doi:10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Hall, M. C. G., D. G. Cacuci, and M. E. Schlesinger, 1982: Sensitivity Analysis of a Radiative-Convective Model by the Adjoint Method. *J. Atmos. Sci.*, **39**, 2038–2050. doi:10.1175/1520-0469(1982)039<2038:SAOARC>2.0.CO;2.

Hamill, H. C., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550–560.

Hello, G. and P. Arbogast, 2004: Two different methods to correct the initial conditions applied to the storm of 27 December 1999 over southern France. *Meteor. Appl.*, **11**, 51–57.

Hoffman, R., Z. Liu, J.-F. Louis, and C. Grasotti, 1995: Distorsion representation of forecast errors. *Mon. Wea. Rev.*, **123**, 2758–2770.

Homar, V., C. Ramis, and S. Alonso, 2002: Numerical study of the October 2000 torrential precipitation event over eastern Spain: Analysis of the synoptic-scale stationarity. *Ann. Geophys.*, **20**, 2047–2066.

Homar, V., R. Romero, D. J. Stensrud, C. Ramis, and S. Alonso, 2003: Numerical diagnosis of a small, quasi-tropical cyclone over the Western Mediterranean: Dynamical vs. boundary factors. *Quart. J. Roy. Meteor. Soc.*, **129**, 1469–1490.

Homar, V. and D. J. Stensrud, 2004: Sensitivities of an intense Mediterranean cyclone: Analysis and validation. *Quart. J. Roy. Meteor. Soc.*, **130**, 2519–2540.

Homar, V. and D. J. Stensrud, 2008: Subjective versus objective sensitivity estimates: application to a North African cyclogenesis. *Tellus A*, **60**, 1064–1078. doi: 10.1111/j.1600-0870.2008.00353.x.

Homar, V., D. J. Stensrud, J. J. Levit, and D. R. Bright, 2006: Value of Human-Generated Perturbations in Short-Range Ensemble Forecasts of Severe Weather. *Wea. Forecasting.*, **21**, 347–363.

Hong, S.-Y. and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339.

Horvath, K., 2008: *Dynamical processes in the upper-troposphere and lee cyclogenesis in the western Mediterranean.* Ph.D. thesis, Dep. Geophysics, Faculty of Science, University of Zagreb.

Hoskins, B. J., 1990: *Theory of extratropical cyclones.* Extratropical Cyclones: The Erik Palmen Memorial Volume. American Meteorological Society.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111** (470), 877–946.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritcie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.

Houtekamer, P. L. and H. L. Mitchell, 1998: Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon. Wea. Rev.*, **126**, 796. doi:10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houtekamer, P. L. and H. L. Mitchell, 2001: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Wea. Rev.*, **129**, 123. doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Mon. Wea. Rev.*, **133**, 604. doi:10.1175/MWR-2864.1.

Huo, Z., D.-L. Zhang, and J. Gyakum, 1999: Interaction of potential vorticity anomalies in extratropical cyclogenesis. Part I: Static piecewise inversion. *Mon. Wea. Rev.*, **127**, 2546–2560.

Isaksen, L., M. Fisher, E. Andersson, and J. Barkmeijer, 2005: The structure and realism of sensitivity perturbations and their interpretation as 'Key Analysis Errors'. *Quart. J. Roy. Meteor. Soc.*, **131**, 3053–3078. doi:10.1256/qj.04.99.

Jacks, E., J. Brent Bower, V. J. Dagostaro, J. Paul Dallavalle, M. C. Erickson, and J. C. Su, 1990: New NGM-Based MOS Guidance for Maximum/Minimum Temperature, Probability of Precipitation, Cloud Amount, and Surface Wind. *Wea. Forecasting.*, **5**, 128–138. doi:10.1175/1520-0434(1990)005<0128:NNBMGF>2.0.CO;2.

Jansà, A., A. Genovés, M. A. Picornell, J. Campins, R. Riosalido, and O. Carretero, 2001: Western Mediterranean cyclones and heavy rain. Part 2: Statistical approach. *Meteor. Appl.*, **8**, 43–56.

Jolliffe, I. T. and D. B. Stephenson, 2003: *Forecast verification: A practioner's guide in atmospheric science.* John Wiley and Sons.

Kadowaki, T., 2007: *A 4-dimensional variational assimilation system for the JMA Global Spectrum Model.* Research Activities in Atmospheric and Oceanic Modelling. Working Group for Numerical Experimentation (WGNE), Blue Book Series (J.Cote, Ed).

Kain, J. S., 2004: The Kain-Fritsch Convective Parametrization: An Update. *J. Appl. Meteor.*, **43**, 170–181.

Kalnay, E., M. Corazza, and M. Cai, 2002: Are bred vectors the same as Lyapunov vectors? . *Preprints, Symp. on Observations, Data Assimilation and Probabilistic Prediction, Orlando, FL, Amer. Meteor. Soc.*, pages 173–177.

Keppenne, C. L. and M. M. Rienecker, 2002: Initial Testing of a Massively Parallel Ensemble Kalman Filter with the Poseidon Isopycnal Ocean General Circulation Model. *Mon. Wea. Rev.*, **130**, 2951. doi:10.1175/1520-0493(2002)130<2951:ITOAMP>2.0.CO;2.

Krishnamurti, T. N., C. Gnanaseelan, A. K. Mishra, and A. Chakraborty, 2008: Improved Forecasts of the Diurnal Cycle in the Tropics Using Multiple Global Models. Part I: Precipitation. *J. Climate*, **21(16)**, 4029–4043.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548–1550.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 2000a: Multi-model superensemble forecasts for weather seasonal climate. *J. Climate*, **13**, 4196–4216.

Krishnamurti, T. N., C. M. Kishtawal, D. W. Shin, and C. E. Williford, 2000b: Improving tropical precipitation forecasts from multianalysis superensemble. *J. Climate*, **13**, 4217–4227.

Krishnamurti, T. N., S. Surendran, D. W. Shin, R. Correa-Torres, T. S. V. Vijaya Kumar, C. E. Williford, C. Kummerow, R. F. Adler, J. Simpson, R. Kakar, W. Olson, and F. J. Turk, 2001: Real time Multianalysis/Multimodel Superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.*, **129**, 2861–2883.

Langland, R. H., R. Gelaro, G. D. Rohaly, and M. A. Shapiro, 1999: Targeted observations in FASTEX: Adjoint-based targeting procedures and data impact experiments in IOP17 and IOP18. *Quart. J. Roy. Meteor. Soc.*, **125**, 3241–3270. doi:10.1002/qj.49712556107.

Leith, C. E., 1974: Theoretical Skill of Monte Carlo Forecasts. *Mon. Wea. Rev.*, **102**, 409. doi:10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2.

Lewis, J. M. and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A*, **37**, 309. doi:10.1111/j.1600-0870.1985.tb00430.x.

Llasat, M., M. Llasat-Botija, M. Prat, F. Porcú, C. Price, A. Mugnai, K. Lagouvardos, V. Kotroni, D. Katsanos, S. Michaelides, Y. Yair, K. Savvidou, and K. Nicolaides, 2010: High-impact floods and flash floods in Mediterranean countries: the FLASH preliminary database. *Adv. Geosci.*, **23**, 1–9.

Llasat, M., T. Rigo, and M. Barriendos, 2003: The 'Montserrat-2000' flash flood event: a comparison with the floods that have ocurred in the northeast Iberian peninsula since the 14th century. *Int. J. Climatol.*, **23**, 453–469.

Llasat, M. and D. Sempere-Torres, 2001: Heavy rains and floods in west mediterranean areas: a climatic feature. *Geophysical Research Abstr acts. European Geophysical Society*, **3**.

Lorenc, A. C., 1981: A Global Three-Dimensional Multivariate Statistical Interpolation Scheme. *Mon. Wea. Rev.*, **109**, 701. doi:10.1175/1520-0493(1981)109<0701:AGTDMS>2.0.CO;2.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Atmos. Sci.*, **20**, 130–141.

Lynch, P., 2008: The origins of computer weather prediction and climate modeling. *J. Comp. Phys.*, **227**, 3431–3444. doi:10.1016/j.jcp.2007.02.034.

Magnusson, L., M. Leutbecher, and E. Källén, 2008: Comparison between Singular Vectors and Breeding Vectors as Initial Perturbations for the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **136**, 4092. doi:10.1175/2008MWR2498.1.

Mahfouf, J. F. and F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1171–1190. doi:10.1002/qj.49712656416.

Manders, A., W. Verkley, J. Diepeveen, and A. Moene, 2007: Application of a potential vorticity inversion method to a case of rapid cyclogenesis over the Atlantic Ocean. *Quart. J. Roy. Meteor. Soc.*, **133**, 1755–1770.

Martín, A., R. Romero, V. Homar, A. De Luque, S. Alonso, T. Rigo, and M. C. Llasat, 2007: Sensitivities of a flash flood event over Catalonia: a numerical analysis. *Mon. Wea. Rev.*, **135**, 651–669.

Mason, I., 1982: A model for assessment of weather forecasts. *Austr. Meteor. Mag.*, **30**, 291–203.

Mass, C., D. Ovens, K. Westrick, and B. Colle, 2002: Does increasing horitzontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430.

Meng, Z. and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135**, 1403–1423.

Meteorologial Office, 1962: *Weather in the Mediterranean.* Air Ministry. Vol. 1, 362 pp.

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.

Murphy, A., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 534–537.

Murphy, A., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting.*, **8**, 281–293.

Murphy, A. H. and R. L. Winkler, 1987: A General Framework for Forecast Verification. *Mon. Wea. Rev.*, **115**, 1330. doi:10.1175/1520-0493(1987)115<1330:AGFFFV>2.0.CO;2.

Nuissier, O., V. Ducroq, D. Ricard, C. Lebeaupin, and S. Anquetin, 2007: A numerical study of three catastrophic events over Souther France. Part I: Numerical framework and synoptic ingredients. *Quart. J. Roy. Meteor. Soc.*, **134**, 11–130.

Nutter, P., D. Stensrud, and M. Xue, 2004a: Effects of Coarsely Resolved and Temporally Interpolated Lateral Boundary Conditions on the Dispersion of Limited-Area Ensemble Forecasts. *Mon. Wea. Rev.*, **132**, 2358–2377.

Nutter, P., M. Xue, and D. Stensrud, 2004b: Application of Lateral Boundary Condition Perturbations to Help Restore Dispersion in Limited-Area Ensemble Forecasts. *Mon. Wea. Rev.*, **132**, 2378–2390.

Nutter, P. A., S. L. Mullen, and D. P. Baumhefner, 1998: The Impact of Initial Condition Uncertainty on Numerical Simulations of Blocking. *Mon. Wea. Rev.*, **126**, 2482–2502. doi:10.1175/1520-0493(1998)126<2482:TIOICU>2.0.CO;2.

Palmer, T., 1999: *Predicting uncertainty in forecasts of weather and climate. Technical Memorandum 294.* ECMWF, Shinfield Park, Reading.

Palmer, T., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate predicion models. *Quart. J. Roy. Meteor. Soc.*, **127**, 279–304.

Palmer, T. and R. Hagedorn, 2006: *Predictability of Weather and Climate.* Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511617652.

Palmer, T. and R. Hagedorn, 2006: *Predictability of Weather and Climate.* Cambridge University Press.

Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular Vectors, Metrics, and Adaptive Observations. *J. Atmos. Sci.*, **55**, 633–653. doi:10.1175/1520-0469(1998)055<0633:SVMAAO>2.0.CO;2.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.*, **120**, 1747. doi:10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Plu, M. and P. Arbogast, 2005: A cyclogenesis evolving into two distinct scenarios and its implications for short-term ensemble forecasting. *Mon. Wea. Rev.*, **133**, 2016–2029.

Rabier, F., P. Courtier, and O. Talagrand, 1992: An application of adjoint models to sensitivity analysis. *Beitr. Phys. Atmos.*, **65**, 177–192.

Rabier, F., H. Jarvinen, E. Klinker, J. F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1148–1170.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.*, **122**, 121–150. doi:10.1002/qj.49712252906.

Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Quart. J. Roy. Meteor. Soc.*, **124**, 1809–1829. doi:10.1002/qj.49712455003.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362. doi:10.1002/qj.32.

Reichle, R. H. and R. D. Koster, 2003: Assessing the Impact of Horizontal Error Correlations in Background Fields on Soil Moisture Estimation. *J. Hydrol.*, **4**, 1229. doi:10.1175/1525-7541(2003)004<1229:ATIOHE>2.0.CO;2.

Reisner, J., R. Rasmussen, and R. Bruintjes, 1998: Explicit Forecasting of Supercooled Liquid Water in Winter Storms Using the MM5 Mesoscale Model. *Quart. J. Roy. Meteor. Soc.*, **124**, 1071–1107. doi:10.1002/qj.49712454804.

Reiter, E., 1975: *Handbook for forecasters in the Mediterranean. Part 1: General description of the meteorological processes.* Naval Environmental Research Facility: Monterey, California.

Romero, R., 2001: Sensitivity of a heavy-rain-producing western Mediterranean cyclone to embedded potential-vorticity anomalies. *Quart. J. Roy. Meteor. Soc.*, **127**, 2559–2597. doi:10.1002/qj.49712757805.

Romero, R., 2008: A method for quantifying the impacts and interactions of potential-vorticity anomalies in extratropical cyclones. *Quart. J. Roy. Meteor. Soc.*, **134**, 385–402. doi:10.1002/qj.219.

Romero, R., C. A. Doswel III, and R. Riosalido, 2001: Observations and fine-grid simulations of a convective outbreak in northeastern Spain: Importance of diurnal forcing and convective cool pools. *Mon. Wea. Rev.*, **129**, 2157–2182.

Romero, R., C. A. Doswell III, and C. Ramis, 2000: Mesoscale numerical study of two cases of long-lived quasistationary convective systems over eastern Spain. *Mon. Wea. Rev.*, **128**, 3731–3751.

Romero, R., V. Homar, C. Ramis, and S. Alonso, 2002: The 10-12 November 2001 Balearic Superstorm: An Assessment of Baroclinic and Diabatic Contributions Through Pv Diagnosis. In *EGS General Assembly Conference Abstracts*, volume 27 of *EGS General Assembly Conference Abstracts*, page 1436.

Romero, R., A. Martín, V. Homar, S. Alonso, and C. Ramis, 2005: Predictability of prototype flash flood events in the western Mediterranean under uncertainties of the precursor upper-level disturbance: The HYDROPTIMET case studies. *Nat. Haz. and Earth. Syst. Sci.*, **5**, 505–525.

Romero, R., A. Martín, V. Homar, S. Alonso, and C. Ramis, 2006: Predictability of prototype flash flood events in the western Mediterranean under uncertainties of the precursor upper-level disturbance: The HYDROPTIMET case studies. *Adv. Geosci.*, **7**, 55–63.

Rossby, C., 1940: Planetary flow patterns in the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **66**, 68–87.

Røsting, B. and J. Kristjánson, 2006: Improving simulations of severe winter storms by initial modification of potential vorticity in sensitive regions. *Quart. J. Roy. Meteor. Soc.*, **132**, 2625–2652.

Røsting, B. and J. Kristjánson, 2008: A successful resimulation of the 7-8 January 2005 winter storm through initial potential vorticity modification in sensitive regions. *Tellus A*, **60**, 604–619.

Rosting, B., J. E. Kristjansson, and J. Sunde, 2006: The sensitivity of numerical simulations to initial modifications of potential vorticity - a case-study. *Quart. J. Roy. Meteor. Soc.*, **129**, 2697–2718. doi: 10.1256/qj.01.112.

Santurette, P. and C. G. Georgiev, 2005: *Weather Analysis and Forecasting: Applying satellite water vapor imagery and potential vorticity analysis*. Elsevier Academic Press.

Santurette, P. and A. Joly, 2002: ANASYG/PRESYG, Météo-France's new graphical summary of the synoptic situation. *Meteor. Appl.*, **9**, 129–154.

Schwartz, C., J. Kain, S. Weiss, M. Xue, D. Bright, F. Kong, K. Thomas, J. Levit, M. Coniglio, and M. Wandishin, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting.*, **25**, 263–280.

Shuman, F. G., 1989: History of Numerical Weather Prediction at the National Meteorological Center. *Wea. Forecasting.*, **4**, 286–296. doi:10.1175/1520-0434(1989)004<0286:HONWPA>2.0.CO;2.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102. doi:10.1256/qj.04.106.

Siccardi, F., 1996: Rainstorm hazards and related disasters in the western Mediterranean region. *Rem. Sens. Rev.*, **14**, 5–21.

Snyder, C., T. Hamill, and S. Trier, 2003: Linear evolution of error covariances in quasi-geostrophic model. *Mon. Wea. Rev.*, **131**, 189–205.

Stensrud, D., H. Brooks, J. Du, M. Tracton, and E. Rogers, 1999: Using ensembles for short-range forecasting. *Mon. Wea. Rev.*, **127**, 433–446.

Stensrud, D. and N. Yussouf, 2007: Reliable probabilistic quantitative precipitation forecasts from a short-range ensemble forecasting system. *Wea. Forecasting.*, **22**, 3–17.

Stensrud, D. J., 2001: Using short-range ensemble forecasts for predicting severe weather events. *Atmos. Res.*, **53**, 3–17. doi: 10.1016/S0169-8095(00)00079-X.

Stensrud, D. J., J. Bao, and T. T. Warner, 2000: Using initial conditions and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107.

Stensrud, D. J. and N. Yussouf, 2003: Short-Range Ensemble Predictions of 2-m Temperature and Dewpoint Temperature over New England. *Mon. Wea. Rev.*, **131**, 2510. doi:10.1175/1520-0493(2003)131<2510:SEPOMT>2.0.CO;2.

Stensrud, D. J., N. Yussouf, D. C. Dowell, and M. C. Coniglio, 2009: Assimilating surface data into a mesoscale model ensemble: Cold pool analyses from spring 2007. *Atmos. Res.*, **93**, 207–220. doi: 10.1016/j.atmosres.2008.10.009.

Swarbrick, S. J., 2001: Applying the relationship between potential vorticity fields and water vapour imagery to adjust initial conditions in NWP. *Meteor. Appl.*, **8**, 221–228. doi: 10.1017/S1350482701002109.

Talagrand, O., R. Vautard, and B. Strauss, 1997: Evaluation of probabilistic prediction systems. *Proceedings, ECMWF Workshop on Predictability,* ECMWF, pages 1–25.

Tan, D. G. H., E. Andersson, M. Fisher, and L. Isaksen, 2007: Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM–Aeolus wind profiling mission. *Quart. J. Roy. Meteor. Soc.*, **133**, 381–390. doi: 10.1002/qj.43.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble Square Root Filters. *Mon. Wea. Rev.*, **131**, 1485. doi:10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.

Torn, R. and G. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.*, **136**, 3947–3963.

Torn, R., G. Hakim, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490–2502.

Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.

Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319.

Tripoli, G. J., C. M. Medaglia, S. Dietrich, A. Mugnai, G. Panegrossi, S. Pinori, and E. A. Smith, 2005: The 9 10 November 2001 Algerian Flood: A Numerical Study. *Bull. Amer. Meteor. Soc.*, **86**, 1229–1235. doi:10.1175/BAMS-86-9-1229.

Troen, I. and L. Mahrt, 1986: A simple model of the Atmospheric Boundary Layer: Sensitivity to Surface Evaporation. *Bound-Lay Meteorol.*, **37**, 129–148.

Uppala, S. M., P. W. KÅllberg, A. J. Simmons, U. Andrae, V. D. C. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. V. D. Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, R. Jenne, A. P. McNally, J.-F. Mahfouf, J.-J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012. doi:10.1256/qj.04.176.

USACE-HEC, US Army Corps of Engineers Hydrologic Engineering Center., 1998: *HEC-HMS Hydrologic Modeling system user's manual.* Davis, California.

Vich, M. and R. Romero, 2010: Multiphysics superensemble forecast applied to Mediterranean heavy precipitation situations. *Nat. Haz. and Earth. Syst. Sci.*, **10**, 2371–2377. doi:10.5194/nhess-10-2371-2010.

Vich, M., R. Romero, and H. E. Brooks, 2011a: Ensemble prediction of Mediterranean high-impact events using potential vorticity perturbations. Part I: Comparison against the multiphysics approach. *Atmos. Res.*, **102**, 227–241. doi:10.1016/j.atmosres.2011.07.017.

Vich, M., R. Romero, and V. Homar, 2011b: Ensemble prediction of Mediterranean high-impact events using potential vorticity perturbations. Part II: Adjoint-derived sensitivity zones. *Atmos. Res.*, **102**, 311–319. doi:10.1016/j.atmosres.2011.07.016.

Vich, M., R. Romero, E. Richard, P. Arbogast, and K. Maynard, 2012: Perturbing the potential vorticity field in mesoscale forecasts of two Mediterranean heavy precipitation events. *Tellus A*. Conditionally accepted.

Vié, B., O. Nuissier, and V. Ducrocq, 2011: Cloud-resolving ensemble simulations of mediterranean heavy precipitating events: uncertainty on initial conditions and lateral boundary conditions. *Mon. Wea. Rev.*, **139**, 403–423. doi: 10.1175/2010MWR3487.1.

Viúdez, Á. and D. G. Dritschel, 2004: Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.*, **521**, 343–352. doi:10.1017/S0022112004002058.

Vukićević, T., 1991: Nonlinear and Linear Evolution of Initial Forecast Errors. *Mon. Wea. Rev.*, **119**, 1602. doi:10.1175/1520-0493(1991)119<1602:NALEOI>2.0.CO;2.

Warner, T., R. Peterson, and R. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **78**, 2599–2617.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Mon. Wea. Rev.*, **130**, 1913. doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Wilks, D., 1995: *Statistical methods in the atmospheric sciences: An introduction*. Academic. Press.

WMO, World Meteorological Organization., 1981: *Guide to Agricultural Meteorological Practices. Second edition, WMO-No. 134*. Secretariat WMO, Geneva, Switzerland.

WMO, World Meteorological Organization., 1983: *Guide to Climatological Practices. Second edition, WMO-No. 100*. Secretariat WMO, Geneva, Switzerland.

WMO, World Meteorological Organization., 1989: *Guide on the Global Observing System. WMO-No. 488*. Secretariat WMO, Geneva, Switzerland.

WMO, World Meteorological Organization., 1990: *Guide on Meteorological Observation and Information Distribution Systems at Aerodromes. WMO-No. 731*. Secretariat WMO, Geneva, Switzerland.

WMO, World Meteorological Organization., 1994: *Guide to Hydrological Practices. Fifth edition, WMO-No. 168*. Secretariat WMO, Geneva, Switzerland.

WMO, World Meteorological Organization., 2003: *Manual on the Global Observing System. Volume I – Global Aspects, WMO-No. 544*. Secretariat WMO, Geneva, Switzerland.

WMO, World Meteorological Organization., 2008: *Guide to Meteorological Instruments and Methods of Observation, No 8., 7th edition*. Secretariat WMO, Geneva, Switzerland, 1th edition on 1996.

Xu, M., D. J. Stensrud, J.-W. Bao, and T. T. Warner, 2001: Applications of the adjoint technique to short-range ensemble forecasting of mesoscale convective systems. *Mon. Wea. Rev.*, **129**, 1395–1418.

Yun, W. T., L. Stefanova, A. K. Mitra, T. S. V. V. Kumar, W. Dewar, and T. N. Krishnamurti, 2005: A multi-model superensemble algorithm for seasonal climate prediction using DEMETER forecasts. *Tellus A*, **57**, 280–289. doi: 10.1111/j.1600–0870.2005.00131.x.

Zhang, Z. and T. N. Krishnamurti, 1997: Ensemble Forecasting of Hurricane Tracks. *Bull. Amer. Meteor. Soc.*, **78**, 2785–2795. doi:10.1175/1520-0477(1997)078<2785:EFOHT>2.0.CO;2.

Zou, X., W. Huang, and Q. Xiao, 1998: *A User's Guide to the MM5 Adjoint Modeling System*. NCAR Tech., Note NCAR/TN-437+IA, Boulder, USA.

Zou, X., F. Vandenberghe, M. Pondeca, and Y.-H. Kuo, 1997: *Introduction to adjoint techniques and the MM5 adjoint modeling system*. NCAR Tech., Note NCAR/TN-435-SRT, Boulder, USA.