LIFE11 ENV GR 975

FLIRE: Floods and fIre Risk assessment and management

Model validation report

31/3/2015

Project location	Greece – Attiki region
Project starting date:	01/10/2012
Project ending date:	30/09/2015
Coordinating Beneficiary	National Technical University of Athens

Associated Beneficiary responsible for the Action	National Observatory of Athens
Contact person	Mrs. Vassiliki Kotroni
Postal address	V. Pavlou & I. Metaxa, P. Penteli, 15127, Greece
Telephone	00-30-210-8109126
Fax :	00-30-210-8103236
E-mail	kotroni@noa.gr
Other Associated Beneficiaries involved in the Action	ICL
Contact persons	ICL: Cedo Maksimovic

Table of Contents

Extended Summary

Action B5 focuses on the provision of short-term weather forecasting data as well as of dense meteorological observations for the study area. Both weather forecasts and observations serve as an input in the Weather Information Management Tool (WIMT) of the Decision Support System.

The present report is devoted to the detailed description of the evaluation of MM5 forecast skill in predicting precipitation over the Greater Athens area. For that purpose, the verification covered the period from March 2013 up to December 2014, a period that comprises 37 rain episodes, with at least one station recording more than 20 mm of rain within 24 hours. For the verification period, 44 rain gauges were selected, operated by the National Observatory of Athens and the National Technical University of Athens, both partners of the project. Following the methodology widely accepted for evaluation of precipitation, a contingency table (yes/no for observed/modelled rain) was constructed for the totality of the 37 episodes and several statistical scores were calculated. Various 24-h rain thresholds were set, in order to evaluate the model performance for light, moderate and high precipitation amounts.

Calculation of the statistical scores revealed a decreasing trend of the Probability of Detection (POD) with increasing rain threshold, with a POD of 0.42 for the highest precipitation amounts, a score that is close to that referenced in the literature for similar activities of high-resolution rain forecasts in the Mediterranean area, taking also into account that in our report the second day of simulation (DAY2) is evaluated. On the other hand, the calculated False Alarm Ratio (FAR) was very low for all rain thresholds, indication thus that the model has no tendency to provide false alarms. Verification of the quantity of forecasted rain against observations (calculation of mean error and mean absolute errors) showed scores that are close or even better than scores reported in the literature for the Mediterranean region.

Overall, the verification procedure showed that MM5 rain forecasts showed a good skill in predicting correctly rain events for the next day, although it was found that for the high precipitation amounts, models still suffer from underestimation of the rain amounts. Therefore it is proposed that model forecasts have to be jointly used with observations (surface stations, lightning measurements, radar data, if available) in order to achieve a high skill in forecasting high precipitation events.

1. Introduction

The aim of Action B5 is the provision of short-term weather forecasting data as well as of observed data that will serve as an input in the Weather Information Management Tool (WIMT) of the Decision Support System. This report describes in detail the strategy adopted in order to set-up an operational weather forecasting chain that will provide the necessary input to other actions of FLIRE project.

For that purpose, NOA implemented the MM5 modelling system that offers a great flexibility of choice of physical parameterization schemes and it is used worldwide for a great variety of applications. Indeed, MM5 model is run operationally, once per day, following a three- nest strategy with 24-km, 8-km and 2-km horizontal grid increment (details on the model setup are given in the following section). This modelling effort follows the worldwide trend to use increasingly higher resolutions with NWP models at operational basis, following the significant improvement of computing capabilities at prices that are continuously decreasing.

Many recent studies, found in literature, deal with the use of very-high resolution modelling. Doyle (1997) and Colle and Mass (2000), among others, have shown that increasing horizontal resolution could be very advantageous, especially in cases of circulations forced by topography and surface contrasts. Bernadet et al. (2000) argued that a grid spacing of 2 km should be used in order to explicitly reproduce convective activity. Benoit et al. (2000) also verified the results of simulations with a grid spacing varying from 35 down to 3 km in order to reproduce the flow produced by a midlatitude cyclone and found a considerable improvement in the results by increasing the grid spacing. Most of the studies, presented in literature, were devoted to the analysis of selected case studies and only a few publications were based on an objective verification of very high-resolution forecasts over longer periods. Longer period objective verification studies have been performed mainly over the United States by Colle et al., (2000), (2003a,b), and Mass et al. (2002), among others. Mass et al. (2002) based on the verification results of a 2-years period over the Pacific Northwest in the United States, found that although there is a noticeable improvement of forecast skill when increasing the grid spacing from 36-km to 12-km, there is a minimal forecast improvement when decreasing the grid spacing from 12 to 4 km; Colle et al. (2003a, b) also supported these findings based on the verification of MM5 model forecasts for almost 2 years in the northeastern United States. On the other hand these authors suggested that the increase in horizontal grid spacing to less than 10-15 km improves the realism of the results, especially as it concerns the precipitation fields, without necessarily improving the traditional objective verification scores.

Quantitative precipitation forecasting (QPF) is recognized as one of the most challenging tasks in numerical weather prediction (NWP). Despite substantial improvements in the forecast skill of sea level pressure, wind and temperature over the years, QPF skill has not improved accordingly, remaining thus a challenge for forecasters (e.g. Fritsch et al., 1998; Mahoney and Lackman, 2006), especially during the warm period of the year (Gallus and Segal, 2001; Fritsch and Carbone, 2004). According to many authors (e.g. Fowle and Roebber, 2003; Fritsch and Carbone, 2004) the uncertainties in the initial conditions, the limited knowledge about the precipitation processes in general, and cloud microphysical issues in particular and the lack of capability of the operational NWP models to explicitly resolve convection due to their coarse resolution are connected with serious errors in QPF.

The aim of this work is to evaluate MM5 forecast skill over the expanded urban area of Athens which is characterized by complex terrain for the important rainfall events that occurred since the operational setup of the model for the needs of FILRE (March 2013) up to the end of 2014. The assessment of the precipitation forecast skill is a key issue in an operational environment. The MM5 precipitation forecasts skill at 2 km horizontal resolution is evaluated over the complex terrain of the Athens Basin.

2. The study area

The Athens urban basin includes the cities of Athens and Piraeus and is characterized by a relatively complex terrain (Fig. 1b). It is surrounded by mountains on the three sides, while on the fourth side there is a major opening to the sea in the southwest (the Saronic Gulf). The three main mountains are Hymettus (1050 m) to the East, Penteli (1100 m) to the North and Parnitha (1400 m) to the Northwest. These mountains with only small gaps between them along with numerous hills located downtown Athens play an important role on the modification of the flow in the area. The Korinthian Gulf (to the west of Athens basin) is surrounded by high mountains both to the north and to the south, which under specific conditions can result in a channelling of the flow towards the Saronic Gulf.

3. The model setup

Three one-way nested grids are defined and used at an operational basis for the needs of FLIRE project (Fig. 1a). Grid 1 has 24-km horizontal grid increment, covering the major part of Europe, the Mediterranean and the northern African coast. Grid 2 has 8-km horizontal grid increment, covering the Greek territory and all the Greek islands. Finally, Grid 3 has a 2-km horizontal grid increment, covering the entire Athens area and the adjacent water bodies (Fig. 1b), including the study area of FLIRE. The horizontal extension of the defined operational grids is shown in Fig.1. In the vertical twenty-three unevenly spaced full sigma levels are selected (σ = 1., 0.99, 0.98, 0.96, 0.93, 0.89, 0.85, 0.80, 0.75, 0.70, 0.65, 0.60, 0.55, 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10, 0.05, 0.00).

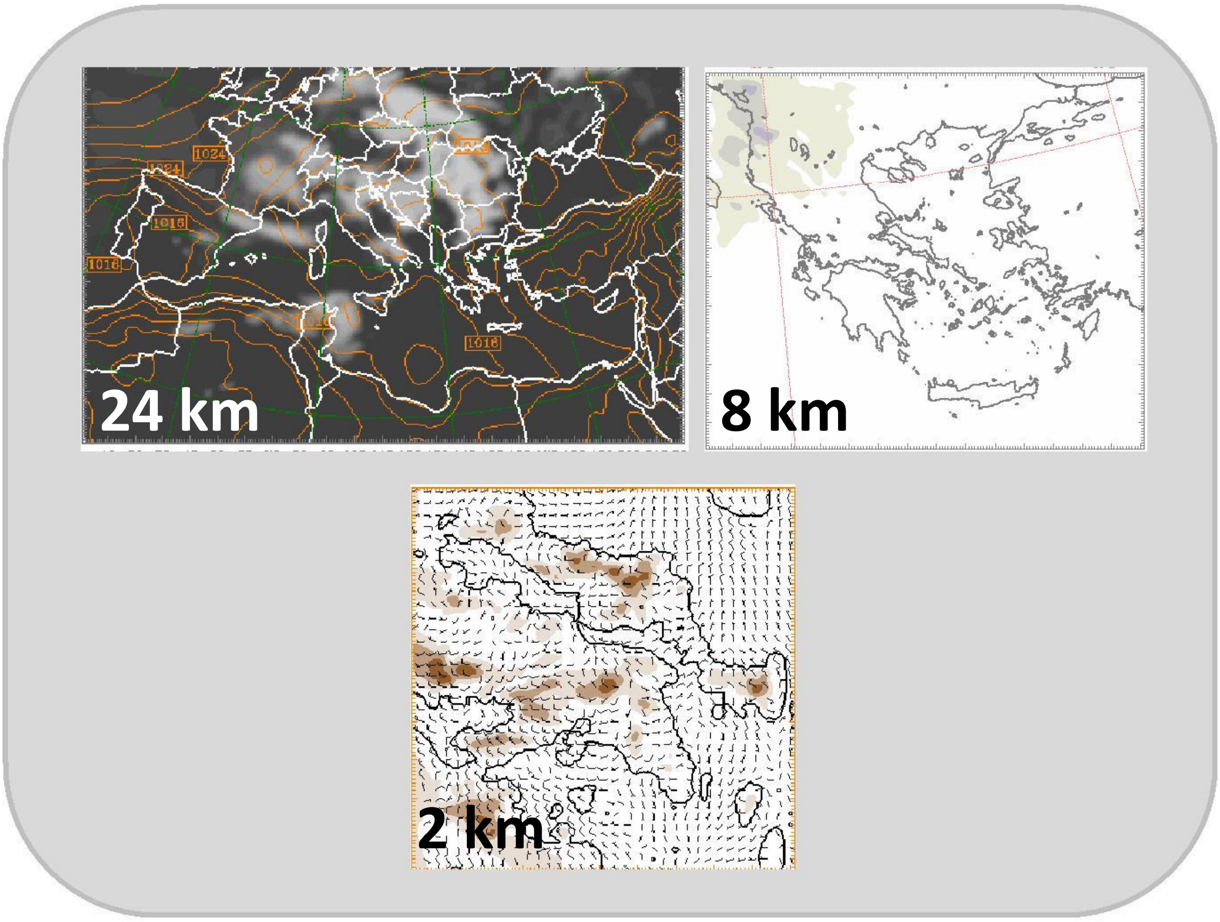

Figure 1: Horizontal extension of MM5 grids.

MM5 model is run once daily, initialized at 0000 UTC. Grid 1 simulation lasts 72 hours, Grid 2 starts at t+6 with a total simulation time of 66 h and finally Grid 3 starts at t+6, with a total simulation time of 42 h. Therefore Grid 3 provides every day detailed weather forecasts for the same day and the following day, at 1-h interval.

The 0000 UTC Global Forecast System (GFS, provided by the National Centres for Environmental Predictions-NCEP, USA) gridded analysis fields and 6-hour interval forecasts, at 0.5-degree lat/lon horizontal grid increment, are used to initialize the model and to nudge the boundaries of Grid 1 during the simulation period. No pre-forecast spin up period or assimilation of additional observations is used in the operational MM5 model chain. The sea surface temperature is initialized from the same GFS dataset. For land use and topography the 30arcsec resolution files provided by USGS are used.

4. Data Sets

The forecast verification is conducted for the rainfall events that have occurred from March 2013 up to December 2014 in the Greater Athens Area. For the selection of the events the following criteria have been applied: (a) at least one station to have recorded more than 20 mm of rain in 24 hours, (b) more than 3 surface meteorological stations have recorded rainfall (in order to avoid very local events). The selected events and the maximum recorded 24-h rainfall for each event are listed in Table 1.

Table 1: Selected events and max 24-h rainfall for each event

	date	max 24hrain	#stations with rain
1	31 December 2014	38	6
2	30 December 2014	24.6	13
3	17 December 2014	23.4	30
4	12 December 2014	60.4	37
5	11 December 2014	60.6	38
6	08 December 2014	68.6	36
7	14 November 2014	21	8
8	08 November 2014	61.8	36
9	24 October 2014	101.8	30
10	23 October 2014	40.4	8
11	16 September 2014	58.6	11
12	09 September 2014	34.4	7

13	07 September 2014	30.6	6
14	19 July 2014	55.6	12
15	15 July 2014	62	8
16	27 April 2014	55.8	15
17	07 April 2014	24.4	6
18	06 April 2014	46	9
19	03 March 2014	75.2	33
20	01 March 2014	22.8	29
21	28 February 2014	20.2	4
22	28 January 2014	62.4	35
23	26 January 2014	28	30
24	25 January 2014	28.8	16
25	24 January 2014	37.6	20
26	15 January 2014	53	14
27	27 December 2013	56.2	35
28	02 December 2013	28.2	27

29	24 November 2013	55.8	32
30	21 November 2013	36.2	17
31	14 November 2013	40	35
32	11 November 2013	92	24
33	06 November 2013	45.8	34
34	01 October 2013	26.2	17
35	12 June 2013	50.4	14
36	08 June 2013	45.4	4
37	12 May 2013	20	4

For the statistical evaluation, the precipitation observations from 32 rain gauges operated by the National Observatory of Athens (red bullets), and 12 rain gauges operated by the National Technical University of Athens (green bullets) are used for the verification (Figure 2).

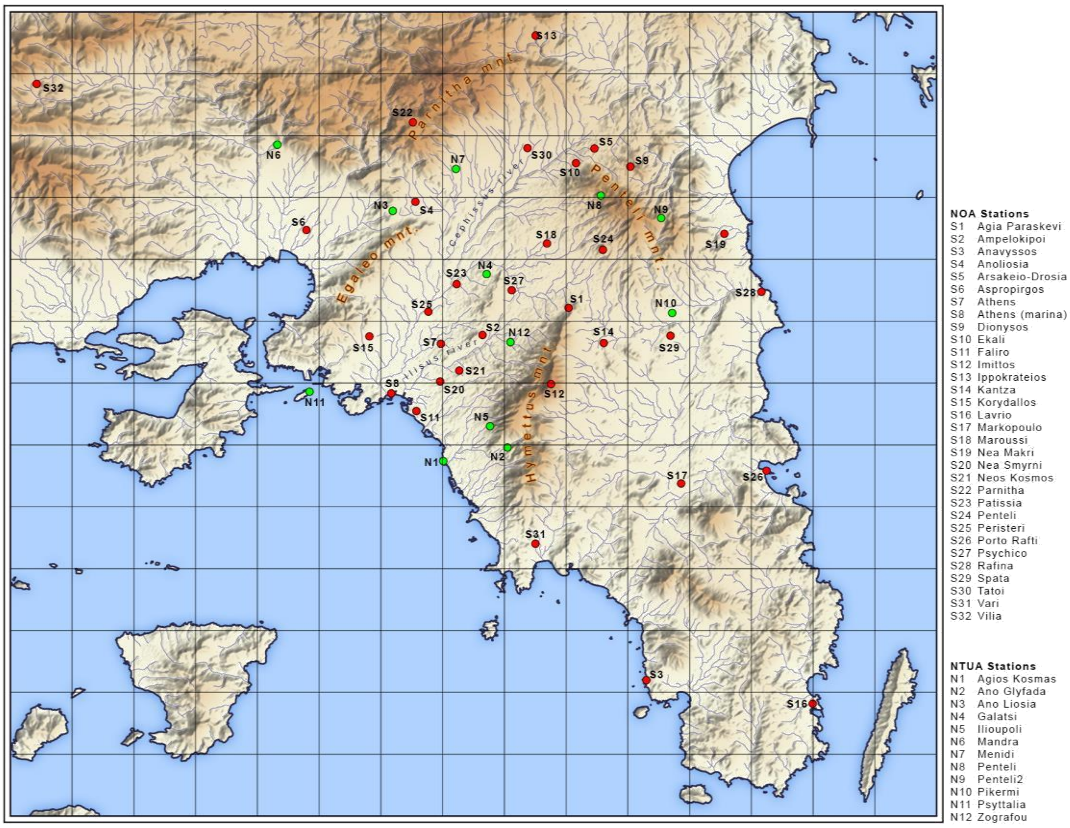

Figure 2: Location and names of the automated meteorological stations used for the model validation; *Red dot*: stations operated by the National Observatory of Athens; *Green dot*: stations operated by the National Technical University of Athens.

5. Statistics

In order to evaluate the model skill in providing accurate precipitation forecasts during the studied period, a verification procedure has been undertaken for the studied area. The 24-h accumulated precipitation values (from t+24 up to t+48), so the model forecast for DAY2, are verified against the rain gauges available. For the verification, the nine model grid points surrounding each rain gauge are considered and as forecast value the closest to the observed one is selected. A contingency table is built from the observed and forecast values as shown in the following:

2x2 Contingency Table		**Event Observed**	
		Yes	No
Event Forecast	Yes	*A*	*B*
	No	*C*	*D*

where A is the number of stations for which the model-forecast precipitation and the observed precipitation equalled or exceed a threshold (hits), B is the number of stations for which only the model-forecast precipitation equalled or exceed a threshold (false alarm), C is the number of the stations for which only the observed precipitation equalled or exceed a threshold (misses) and D is the number of the stations for which neither the model-forecast precipitation nor the observed precipitation equalled or exceed a threshold (correct negatives). The following measures are calculated:

- Areal Bias, $AB = \frac{A+B}{A+C}$
- Probability Of Detection, $POD = \frac{A}{A+C}$
- False Alarm Ratio, $FAR = \frac{B}{A+B}$
- Critical Success Index, $CSI = \frac{A}{A+B+C}$, and

In the framework of this study, the aforementioned statistical scores are calculated for five distinct thresholds: 1, 2.5, 5, 10, 20 mm. In addition, the following quantitative measures have been calculated:

- quantity bias $QB = \overline{P_f} - \overline{P_o}$, where P_f is a single precipitation forecast,

P_o is the corresponding precipitation observation, and the overbar represents a mean over the sample (number of stations that observed precipitation amounts within a range); and

- mean absolute error $MAE = \frac{\sum |P_f - P_o|}{n}$, where n is the number of observing stations.

QB and MAE are calculated for five ranges: 0.1 – 2.5, 2.5 – 5, 5 – 10, 10 – 20, and >20 mm. All of the aforementioned skill scores were averaged for all cases and the results are discussed in the following section.

6. Discussion of the validation results

The Areal Bias measures the ratio of the frequency of forecast events to the frequency of observed events and indicates whether the model has a tendency to underestimate (*AB*<1) or overestimate (*AB*>1) events. Table 2 shows the Areal Bias scores for the five selected precipitation thresholds. For all thresholds the areal bias is lower that one, indicating an underestimation of the areal coverage of rain, especially for the medium (>10 mm) and high (>20 mm) precipitation thresholds. These values are however better than the values reported by Mazarakis et al (2009) who analysed MM5 forecasts of precipitation for selected cases of summer convection over Greece.

As far as the ability of the model to correctly forecast the observed precipitation events is concerned, probability of detection (POD) has been calculated and it is shown in Table 2. POD shows a decreasing trend for increasing rain thresholds, reaching the value of 0.51 for 24-h accumulated rain exceeding 20 mm. Concerning the false alarm ratio (FAR) shown in the same table, it is worth to notice the low values (lower than 0.04 for all threshold except the highest for which the value is 0.17), especially for the light rain thresholds, indicating that the model has no tendency to provide false alarms, which is also important for the prompt operation of a warning system for the occurrence of flood events. It should be noted that the values of POD and FAR shown in Table 2 are in good agreement with recent published work on the evaluation of high resolution precipitation forecasts in other areas of Europe (e.g. Oberto et al., 2012 for Italy, using two high-resolution weather prediction models: WRF and COSMO).

Finally, the critical success index (CSI) is examined, a score that measures the fraction of observed and/or forecast events that were correctly predicted. CSI can be thought of as the accuracy when correct negatives have been removed from consideration and is sensitive to hits, penalizes both misses and false alarms and the unity is the perfect score, while 0 is the lowest possible value. The CSI scores shown in Table 2 decay with increasing threshold values and range from 0.88 for the lowest threshold to ~0.4 for the >20 mm threshold. Again these results are similar or even better than equivalent scores found in the literature (Mazarakis et al, 2009).

The aforementioned scores only give a measure of the model accuracy based on the frequency of precipitation occurrence at or above a threshold and do not account for the magnitude of precipitation errors. Investigation of the quantitative bias of forecast precipitation is performed through inspection of the Mean Absolute Error (MAE) and QB shown in Table 3. MM5 forecasts underestimate the amounts of rain for all ranges (negative QB) with the exception of the first range of 0.2 – 2.5 mm. The greatest errors are obtained for the large amounts of rainfall (for >20 mm the mean value of QB is -14.36 mm). The same comments apply for the MAE values that range from 0.66 mm for the lowest range, 3 mm for the medium precipitation amounts (5-10 mm range) and an error value of ~15 mm for the high precipitation amounts. The numbers reported in Table 3 are better that the values reported in previous studies over Greece (Kotroni and Lagouvardos, 2004, Mazarakis et al, 2009). It should be noted however that these previous studies report only to summer period cases (when rain forecast is in general a more demanding task) while the analysis performed in the frame of this project spans on cases throughout the year.

Table 2: Results for the calculated statistical scores for various thresholds of 24-h accumulated rainfall.

	Rain Thresholds in mm				
	1	***2.5***	***5***	***10***	***20***
Areal Bias	0.90	0.83	0.73	0.65	0.51
POD	0.89	0.82	0.72	0.63	0.42
FAR	0.01	0.01	0.01	0.04	0.17
CSI	0.88	0.82	0.72	0.61	0.39

Table 3: Results for the calculated statistical scores for various ranges of 24-h accumulated rainfall.

	Rain Ranges in mm				
	0.2-2.5	***2.5-5***	***5.-10.***	***10. - 20.***	***>20.***
QB	0.28	-0.52	-1.65	-3.59	-14.36
MAE	0.66	1.02	3.05	5.97	15.95

REFERENCES

Benoit, R., P. Pellerin, N. Kouwen, H. Ritchie, N. Donaldson, P. Joe, and E.D. Soulis, 2000: Toward the use of coupled atmospheric and hydrologic models at regional scale. *Mon. Wea. Rev.,* 128, 1681-1706.

Bernadet, L.R., L. D. Grasso, J.E. Nachamkin, C.A. Finley, and W.R. Cotton, 2000: Simulating convective events using a high resolution mesoscale model. *J. Geophys. Res.,* 105, 14963-14982.

Colle, B.A. and C.F. Mass, 2000: The 5-9 February 1996 flooding event over the Pacific Northwest: Sensitivity studies and evaluation of the MM5 precipitation forecasts. *Mon. Wea. Rev*., 128, 593-617.

Colle, B.A., C.F. Mass, and K.J. Westrick, 2000: MM5 precipitation verification over the Pacific Northwest during the 1997-99 cool seasons. Wea. Forecasting, 15, 730-744.

Colle, B.A., J.B. Olson and J.S. Tongue, 2003a: Multiseason verification of the MM5, Part I: Comparison with the Eta Model over the Central and Eastern United States and Impact of MM5 Resolution. *Wea. Forecasting*, 18, 431-457.

Colle, B.A., J.B. Olson and J.S. Tongue, 2003b: Multiseason verification of the MM5, Part II: Evaluation of high resolution precipitation forecasts over the northeastern United States. *Wea. Forecasting*, 18, 458-480.

Doyle, J.D., 1997: The influence of mesoscale orography on a coastal jet and rainband prediction. *Mon.Wea. Rev.,* **125**, 1465-1488.

Fowle, M. A. and P. J. Roebber, 2003: Short-range (0–48 h) numerical prediction of convective occurrence, mode, and location. *Wea. Forecasting,* **18,** 782–794.

Fritsch, J. M., and Co-authors, 1998: Quantitative precipitation forecasting: Report of the Eighth Prospectus Development Team, U.S. Weather Research Program. *Bull. Amer. Meteor. Soc.,* **79,** 285-299.

Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts n the warm season: A USWRP research and development strategy. *Bull. Amer. Meteor. Soc.,* **85,** 955-965.

Gallus, W. A., and M. Segal, 2001: Impact of improved initialization of mesoscale features on convective system rainfall in 10-km Eta simulations. *Wea. Forecasting,* **16,** 680–696.

Kotroni, V., and K. Lagouvardos, 2004: Evaluation of MM5 high-resolution real-time forecasts over the urban area of Athens, Greece. *J. Appl. Meteor.,* **43,** 1666–1678.

Mahoney, K. M. and G. M. Lackman, 2006: The sensitivity of Numerical Forecasts to Convective Parameterization: A Case study of the 17 February 2004 East Coast Cyclone. *Wea. Forecasting,* **21,** 465–488.

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

Mazarakis, N., V. Kotroni, K. Lagouvardos and A.A. Argiriou, 2009: The sensitivity of numerical forecasts to convective parameterization during the warm period and the use of lighting data as an indicator for convective occurrence. *Atmospheric Research* **94**, 704-714.

Oberto E., M. Milelli, F. Pasi and B. Gozzini, 2012: Intercomparison of two meteorological limited area models for quantitative precipitation forecast verification. *Nat. Hazards Earth Syst. Sci.*, **12**, 591–606.