The Meteorological Model BOLAM at the National Observatory of Athens: Assessment of Two-Year Operational Use

K. LAGOUVARDOS, V. KOTRONI, A. KOUSSIS, AND H. FEIDAS

Institute of Environmental Research and Sustainable Development, National Observatory of Athens, Athens, Greece

A. BUZZI AND P. MALGUZZI

Institute of Atmospheric Sciences and Climate, CNR, Bologna, Italy

(Manuscript received 29 March 2002, in final form 27 May 2003)

ABSTRACT

Since November 1999, the hydrostatic meteorological Bologna Limited-Area Model (BOLAM) has been running operationally at the National Observatory of Athens. The assessment of the model forecast skill during the 2-yr period included (a) calculation of the root-mean-square errors (model vs gridded analyses) of geopotential height and temperature at 850 and 500 hPa, (b) evaluation of the model's quantitative precipitation forecast skill for the most important events, and (c) evaluation of the model skill in the prediction of surface winds in comparison with buoy observations. Comparison of the verification results with those provided in the literature showed that BOLAM has a high forecast skill, even for precipitation, which is the most difficult parameter to forecast. Especially for precipitation, the comparison between coarse (~21 km) and fine (~6.5 km) grid spacing forecasts showed that for the low and medium precipitation amounts, the finer-grid forecasts are not as good as the coarse-grid forecasts. For the large precipitation amounts, the calculated statistical scores provide only little support of the idea that the fine-grid forecasts are better than those of the coarse grid because the fine-grid forecasts give better scores only for the quantity bias and the mean absolute error.

1. Introduction

Regional real-time numerical weather prediction is currently performed at a large number of sites worldwide using mesoscale models. Following the rapid advances of computer capabilities, limited-area models with improved spatial and temporal resolution are used. Moreover, the range and quality of weather services provided by limited-area models is being maximized, including their capability to provide reliable precipitation, temperature, and wind forecasts at resolutions that tend to meet the needs of specific applications (e.g., agriculture, navigation, hydrology, flood forecasting).

In the summer of 1999, the Institute of Environmental Research and Sustainable Development of the National Observatory of Athens (NOA hereinafter) initiated a project for the implementation and operational use of a limited-area model over the area of the eastern Mediterranean. This initiative was part of the European Union (EU) funded demonstration program Telematics Assisted Handling of Flood Emergencies in Urban Areas (TELEFLEUR) that was aimed at the development of a comprehensive operational system for handling urban flood emergencies. This system synthesizes cutting-edge telematics technology with advanced forecasting of meteorology and hydrology encapsulated in a decision support system. The utility of the validated, integrated system has been demonstrated in two selected urban areas in the Mediterranean: Athens, Greece, and Genoa, Italy. Results from the application of the integrated system for the handling of flood emergencies at these sites are presented in Koussis et al. (2003).

One of the main components of the integrated system is the meteorological model. For that reason NOA selected the Bologna Limited-Area Model (BOLAM), which is a hydrostatic limited-area model with the ability to perform one-way nested simulations. After a 3-month period for installation and testing, BOLAM became operational at NOA in November 1999, providing daily 48-h forecasts (72 h as for the summer of 2001) for the eastern Mediterranean region.

The verification of weather forecasts is considered an essential part of any forecasting system. Thus, the present paper is devoted to the verification of model forecasts after its 2-yr (November 1999–March 2001) operational use. The verification procedure entails the comparison of BOLAM forecasts with verifying anal-

Corresponding author address: Dr. K. Lagouvardos, Institute of Environmental Research and Sustainable Development, National Observatory of Athens, Lofos Koufou, P. Pendeli, 15236 Athens, Greece.
E-mail: lagouvar@meteo.noa.gr

yses and available observations. Special attention is paid to the verification of precipitation. The precipitation forecast is considered to be one of the most important model output fields provided by numerical weather prediction models because precipitation has direct (and often disastrous) impacts on human activities. Further, a comparison of wind forecasts with observations provided by a number of buoys deployed in the maritime areas of Greece is also performed. Accuracy of wind forecasts over the maritime areas of Greece is also important for a large number of activities, such as maritime transport and fishery. The verification procedure is of paramount importance because BOLAM is being further tested because it will be used operationally by NOA during the Athens 2004 summer Olympics.

The rest of the paper is organized as follows. Section 2 provides a brief description of BOLAM in which additional information is provided concerning the special settings for the operational use at NOA. Section 3 presents results from the verification procedure performed for the 850- and 500-hPa geopotential height and temperature, the precipitation forecasts of autumn and winter cases of the period 1999–2001, and wind forecasts. The concluding remarks and the prospects for future work are presented in section 4.

2. Description of BOLAM and model settings

The operational weather forecasts at NOA are performed with BOLAM. The most recent version of BOLAM is based on previous versions of the model described by Buzzi et al. (1994, 1997, 1998) and Buzzi and Foschini (2000). The main features of the model are as follows:

- hydrostatic primitive equations;
- dependent variables: surface pressure, wind components u and v, potential temperature, specific humidity, and five microphysical variables;
- an Arakawa-C grid (rotated lat–lon coordinates) and σ vertical coordinate;
- a forward–backward (FB) 3D Eulerian advection scheme (Malguzzi and Tartaglione 1999) and semi-Lagrangian advection of hydrometeors;
- a split-explicit time scheme (FB for gravity modes);
- fourth-order horizontal diffusion and second-order divergence diffusion; and
- a Davies–Källberg–Lehmann (Lehmann 1992) relaxation scheme for lateral boundary conditions.

Physical parameterizations include

- dry adiabatic adjustment;
- radiation: infrared and solar, interacting with clouds (Ritter and Geleyn 1992);
- vertical diffusion (surface layer and planetary boundary layer parameterization), depending on Richardson number (Louis 1979; Louis et al. 1982);
- surface thermal and water balance (three soil layers);
- an explicit microphysical scheme with two water and three ice species; and
- a convective parameterization scheme proposed by Kain and Fritsch (1990, 1993), with implementation (see Buzzi and Foschini 2000) of the modifications suggested by Spencer and Stensrud (1998), which concern the delay of downdrafts in newly developed convection.

BOLAM has been validated with very encouraging results in different intercomparison exercises among many limited-area models in the context of the Comparison of Mesoscale Prediction and Research Experiment (COMPARE) project (Gyakum et al. 1996; Georgelin et al. 2000; Nagata et al. 2001). In particular, BOLAM has given comparatively very good results in simulating episodes of heavy orographic precipitation in the Alps (Bacchi and Ranzi 2000).

It is instructive to give here some more information on the schemes used in BOLAM to simulate the resolved- and subgrid-scale precipitation. The microphysical scheme was coded mainly on the basis of the transformation process models described in Schultz (1995). This is a simple and computationally efficient approach to represent cloud processes in an operational model. The scheme includes five hydrometeor categories: cloud ice, cloud water, rain snow, and graupel. In his paper, Schultz compares the results of his scheme against both the results of a well-documented *research* microphysics algorithm and observations. He found that both schemes provided similar and generally skillful precipitation forecasts, with the advantage that his scheme is about 7–10 times as fast as the research-oriented algorithm. Recently, Schultz's scheme has been implemented into other numerical models with worldwide use [e.g., the fifth-generation Pennsylvania State University (PSU)–National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5)].

The subgrid-scale precipitation is treated in BOLAM, following the Kain–Fritsch convective parameterization scheme (Kain and Fritsch 1993). The Kain–Fritsch scheme is based on the Fritsch–Chappel scheme, with improvements on the detrainment effect and the cloud model. It has been developed for mesoscale models with a grid size of a few tens of kilometers. In this scheme, convection is triggered by lifting a lower-level slab layer with an impetus heating as a function of grid-scale vertical motion at the lifting condensation level. The convective adjustment is based on convective available potential energy (CAPE) and, thus, once convection is triggered, CAPE is assumed to be removed in a grid column within a convective timescale. This timescale is in the range of 30–60 min, depending on the averaged wind speed between the lifting condensation level and 500 hPa. The triggering vertical velocity is automatically adjusted to the grid spacing. In the version of the Kain–Fritsch scheme implemented in BOLAM, an additional modification, regarding the delaying of down-

draft occurrence (Spencer and Stensrud 1998) has been introduced. Namely, the first downdraft is started not before 30 min of initiation of new convection. The Kain–Fritsch scheme has shown considerable success in simulating the development and evolution of convection under a variety of convective and synoptic environments (Kuo et al. 1996; Wang and Seaman 1997; Ferretti et al. 2000). Wang and Seaman (1997) have compared the performances of several cumulus convective parameterization schemes, for six different cases over the United States for both the warm and cold seasons, and have established the skill of the Kain–Fritsch scheme, especially during the cold season and for heavy precipitation events. Most recently, Ferretti et al. (2000) tested the skill of different convective parameterization schemes in predicting the precipitation patterns on the Alpine region and concluded that the Kain–Fritsch scheme is better in reproducing the precipitation over the area, especially for the cases of strong local convection. Additionally, the Kain–Fritsch scheme has shown the most consistent behavior among a number of convective parameterization schemes implemented in MM5 for the simulation of cold-period precipitation cases over Greece (Kotroni and Lagouvardos 2001).

BOLAM has the capability to perform one-way-nested simulations. A first simulation is performed with a coarse-grid interval, and then the outputs of this coarse simulation are used as initial and boundary conditions on a subsequent run with finer grid spacing. For the operational use at NOA, 2 one-way nested grids have been used:

- The coarse grid consists of 135 × 110 points with a 0.21° horizontal grid interval (~23 km) centered at 41°N latitude and 15°E longitude, covering the area of the eastern Mediterranean (Fig. 1a). This configuration was valid for the second year of operational use, while during the first year the coarse grid covered about one-half of this domain (90 × 84 grid points).
- The fine grid consists of 140 × 128 points with a 0.06° horizontal grid interval (~6.5 km), centered at 38°N latitude and 24°E longitude (approximately the position of Athens). The fine grid covers the Greek peninsula with its maritime areas expanding from the Ionian Sea in the west up to the Turkish coasts in the east (Fig. 1b).

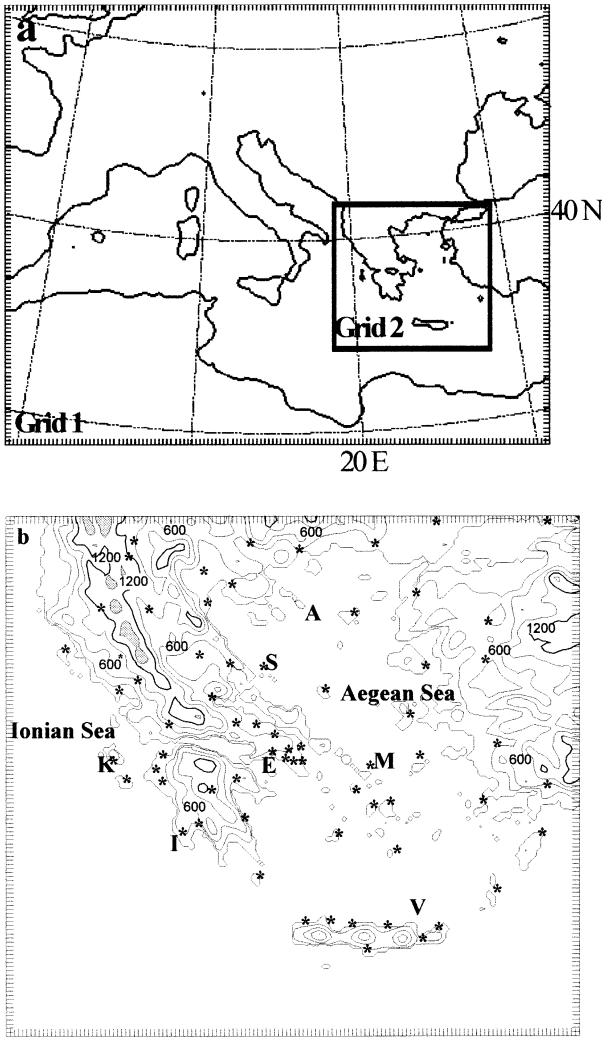

FIG. 1. (a) Horizontal extension of BOLAM grids. The rectangle denotes the position of the fine grid. (b) Orography of the fine grid. Elevation contours are given at 300-m interval, boldface line denotes the 1200-m elevation contour, and gray shading is the heights exceeding 1500 m. Stars denote the position of the rain gauges. Some locations are also reported (the buoy location is denoted with A for Athos, M for Mykonos, E for Egina, and V for Avgo, while the rain gauge at Kefalonia Island is denoted with K, the gauge at Methoni with I, and the gauge at Skiathos island with S).

In the vertical 30 levels are used in the coarse grid and 40 levels in the fine grid, while the model top has been set at about 10 hPa on both nests. The vertical resolution is higher in the boundary layer and, to a lesser extent, at the average tropopause level.

The Medium-Range Forecast (MRF) model [aviation run (AVN), provided by the National Centers for Environmental Predictions (NCEP)] gridded analysis fields and 6-h interval forecasts, at 1.25° lat × lon horizontal grid increment, are used to initialize the model and to nudge the boundaries of the coarse grid during the simulation period. These fields are objectively interpolated on sigma levels from which they are then interpolated onto the model grid points. The orography fields are derived from a terrain data file with 30 arc s resolution provided by the U.S. Geological Survey (USGS).

The operational run is initialized every day with the 0000 UTC AVN analysis. The duration of the simulation is 48 h (72 h as of the summer of 2001) for the coarse grid, and 42 h (66 h as of the summer of 2001) for the inner grid, starting at 0600 UTC of the same day. The main meteorological products are displayed online at http://www.noa.gr/forecasts, together with maps of available observations (surface reports, lightning im-

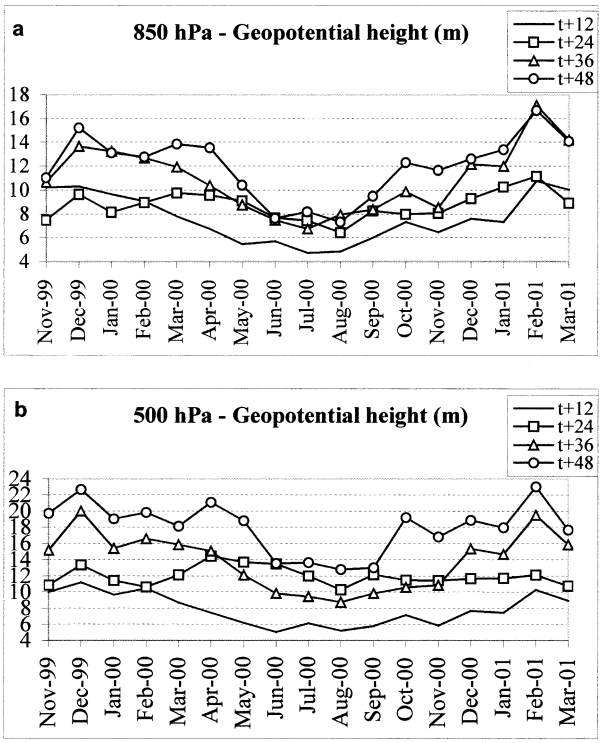

FIG. 2. Rmse of BOLAM coarse-grid geopotential height at (a) 850- and (b) 500-hPa pressure levels. calculated against AVN verifying analyses.

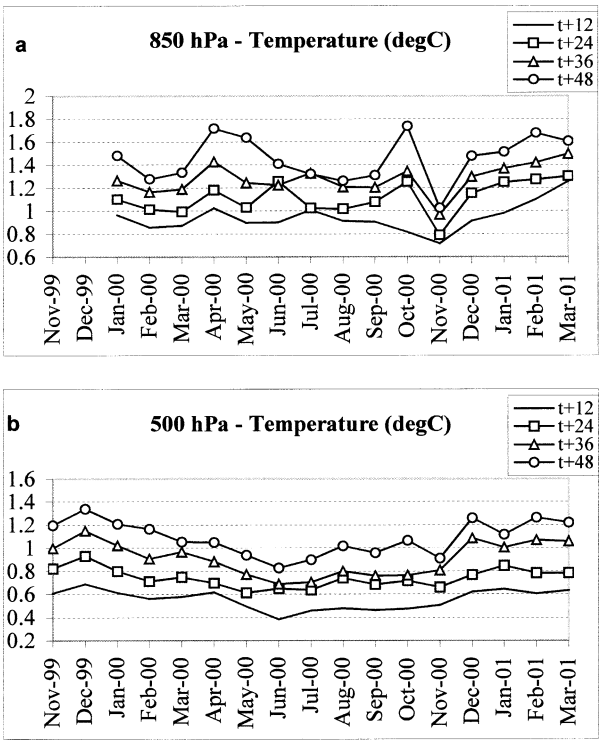

FIG. 3. Rmse of BOLAM coarse-grid temperature at (a) 850- and (b) 500-hPa pressure levels. calculated against AVN verifying analyses.

pacts, satellite images) in order to have a first subjective verification of model results.

3. Statistical verification of BOLAM forecasts

The statistical evaluation of BOLAM forecasts was considered a first priority task—one to be performed with the start of the operational use of the model. As Brooks and Doswell (1996) pointed out, ''*producing forecasts without verifying them systematically is an implicit admission that the quality of the forecasts is a low priority.*'' A verification procedure has been developed in order to assess the ability of the model to provide accurate forecasts. This procedure comprises classical calculations of the root-mean-square error (rmse) of geopotential height and temperature at different pressure levels and is performed automatically every day. For all cases with important amounts of accumulated precipitation over the Greek peninsula that occurred during autumn and winter 1999–2001, model precipitation forecasts have been statistically verified against the available rain gauge data of the network of Greece and of neighboring countries. Moreover, wind forecasts over selected maritime locations have been verified against buoy observations (see Fig. 1b for the location of the buoys).

a. Geopotential height and temperature

BOLAM forecast fields (geopotential height and temperature at the 850- and 500-hPa levels from the coarse grid of simulation) are verified against upper-air gridded analysis fields provided by the AVN model. Each day, the automatic verification procedure calculates the rmse of geopotential height and temperature, given by the formula

$$\text{rmse} = \left[\frac{1}{N}\sum_{i=1}^{N}(F_i^{\text{fc}} - F_i^{\text{obs}})^2\right]^{1/2}, \qquad (1)$$

where F^{obs} is the value of the parameter as it is provided by AVN analysis, F^{fc} is the corresponding forecast value provided by BOLAM, and N is the number of grid points. Because AVN analysis fields are available at 12-h intervals, the verification procedure is applied at t + 12, t + 24, t + 36, and t + 48 BOLAM forecast hours.

Figures 2a and 2b show the rmse of geopotential height at the 850- and 500-hPa levels for 12, 24, 36, and 48 forecast hours, for the time period 1 November 1999–31 March 2001, while Figs. 3a and 3b show the rmse of temperature at the same levels and for the same period. The rmse of geopotential height increases with forecast time and has a considerable seasonal variation. Namely, from June through September, the rmse of geopotential height at 850 hPa is less than 10 m at all forecast times (with a minimum of 4.7 m in July 2000 at t + 12 and of 7.3 m in August 2000 at t + 48). At the 500-hPa level and from June through September, the rmse of geopotential height is less than 16 m for all

forecast times (with a minimum of 4.7 m in August 2000 and 2001 at $t + 12$, and 13 m in September 2000 at $t + 48$). During the cold period of the year, the errors are more important, ranging from 6.5 to 13 m at $t + 12$ and from 8.5 to 17.7 m at $t + 48$ at 850 hPa and ranging from 6 to 11.6 m at $t + 12$ and from 14.4 to 23 m at $t + 48$ at 500 hPa. The fact that during the warm period of the year the rmse are better than during the cold period of the year is reasonable, because during summer the geopotential heights present in general a smaller day-to-day variability.

The rmse of temperature at the same levels also evolves almost linearly with forecast time (Figs. 3a,b). A seasonal variation is evident at the 500-hPa level (Fig. 3b) with higher rmse during the cold than during the warm period of the year. The rmse variation at the 850-hPa level (Fig. 3a) shows no systematic seasonal dependence. The 850-hPa level is influenced by the orographic features (part of the model domain orography intersects the 850-hPa isosurface) and this is reflected to daily forecasts, partially masking the seasonal variability evident at the 500-hPa level. The rmse of the 850-hPa temperature range from 0.7° (in November 2000) to 1.2°C (in March 2001) at $t + 12$ and from 1° (in November 2000) to 1.7°C (in November 2001) at $t + 48$. At the 500-hPa level, the rmse ranges from 0.4° (in June 2000) to 0.7°C (in December 1999) at $t + 12$ and from 0.8° (in August 2001) to 1.3°C (in December 1999) at $t + 48$.

It would be interesting to compare the rmse values calculated for BOLAM forecasts with similar verification scores available from other sources. The Canadian Meteorological Service Web site (available online at http://www.msc-smc.ec.gc.ca) presents rmse calculations of geopotential height for various global-scale models over the Northern Hemisphere. Inspection of the rmse of 500-hPa geopotential height reveals that for the 24-h forecast, the best rmse (provided by the ECMWF global spectral model) is ~13 m in winter and ~12 m in summer, while for the 48-h forecast, the rmse varies between 22 (winter) and 18 (summer) m. Similar results, with the ECMWF forecasts outperforming other models, over the northern European area (an area comparable in size with the area covered by BOLAM coarse grid) are also provided by the Swedish Meteorological Service in the verification report of ECMWF products (ECMWF 2000). Saulo et al. (2001) report rmse calculations from a limited-area model and MRF global model over the southern part of South America for a 2-month period. The authors report an rmse of 500-hPa geopotential height of 15 (MRF model) and 22 (limited-area model) m at $t + 24$, increasing to 23 and 35 m, respectively, at $t + 48$ (see Saulo et al. 2001, their Fig. 7). For the 850-hPa temperature, rmse of 1° (MRF model) and 1.5°C (limited-area model) at $t + 24$, and 1.5° and 2.2°C, respectively, at $t + 48$ are reported. The rmse of 500-hPa geopotential height from BOLAM, shown in Fig. 3b, average at 12 m for both winter and summer for the $t + 24$ h forecast and at 15 m during summer and 19.5 m during winter for the $t + 48$ h forecast. BOLAM rmse of 850-hPa temperature average at 1.1°C for both winter and summer for the $t + 24$ h forecast, and at 1.5° during winter and 1.3°C during summer at $t + 48$ h. At this point it should be noted that limited conclusions could be derived by this comparison because the results of such type of verification highly depend on the geographical location and also the time period used.

b. Accumulated precipitation

The statistical verification of precipitation fields has been performed by applying the statistical measures described by Schultz (1995), part of which are also used by the NCEP for verification of their operational products (e.g., Mesinger et al. 1990). Four statistical measures are calculated as follows.

- The areal bias B is defined as

$$B = \frac{F}{O},$$

  where F is the number of stations for which the model-predicted precipitation amount exceeded a certain threshold and O is the number of stations that observed at least the threshold amount.
- The threat T is defined as

$$T = \frac{\mathrm{CF}}{F + O - \mathrm{CF}},$$

  where CF is the number of stations with correct forecasts from the model. A threat equal to 1 is a perfect score, while 0 is the lowest possible value. These statistical measures are used extensively for evaluation of model forecasts of precipitation (e.g., Mesinger et al. 1990; Mesinger 1996; Belair et al. 2000). In the framework of this study, bias and threat are calculated for six distinct threshold values of precipitation (0.1, 1, 2.5, 5, 10, and 20 mm).
- The quantity bias (QB) is defined as

$$\mathrm{QB} = \overline{P_f} - \overline{P_o},$$

  where P_f is a single precipitation forecast, P_0 is the corresponding precipitation observation, and the overbar represents a mean over the sample.
- The mean absolute error (MAE) is defined as

$$\mathrm{MAE} = \frac{\sum |P_f - P_o|}{n},$$

  where n is the number of observing stations.

The quantity bias and mean absolute error are calculated for five ranges of the observed precipitation amounts (01–2.5, 2.5–5, 5–10, 10–20, and >20 mm).

Last, Schultz (1995) presented calculations of the rank correlation, which can assess the model ability to

TABLE 1. Summary of the synoptic-scale characteristics of the 11 precipitation cases.

Case	Main characteristics
7–8 Nov 1999	Low pressure system of ~998 hPa north of Sicily at 1200 UTC 7 Nov, moving slowly toward southeast and filling (1001 hPa over east Sicily at 0600 UTC 8 Nov, associated with a cold front along the western Greek coasts). The 24-h accumulated precipitation exceeded 100 mm at two stations.
5–6 Dec 1999	Low pressure system of ~1006 hPa over the central Adriatic at 1200 UTC 5 Dec 1999, moving eastward over the Ionian Sea (~1009 hPa at 0600 UTC 6 Dec 1999). The 24-h accumulated precipitation exceeded 60 mm at one station.
10–11 Dec 1999	Low pressure system of ~1002 hPa over Corsica at 1200 UTC 10 Dec moving eastward (~1005 hPa over the central Adriatic at 1200 UTC 11 Dec, associated with a cold front over the Ionian Sea). The 24-h accumulated precipitation exceeded 70 mm at one station and 50 mm a five stations.
21–22 Dec 1999	Low pressure system of ~1003 hPa over Sicily at 0600 UTC 21 Dec, moving eastward over Ionian Sea (1010 hPa at 1200 UTC 22 Dec). The 24-h accumulated precipitation exceeded 75 mm at one station and 40 mm at three stations.
30–31 Dec 1999	Low pressure system of ~1016 hPa east of Sicily at 1200 UTC 30 Dec, moving northeastward and deepening (1004 hPa over the north Aegean at 1800 UTC 31 Dec). The 24-h accumulated precipitation exceeded 40 min at two stations.
9–10 Feb 2000	Low pressure system of ~1013 hPa over central Italy at 1200 UTC 9 Feb movnig southward and deepening (1009 hPa over Sicily at 1200 UTC 10 Feb, associated with a cold front along the western Greek coasts). The 24-h accumulated precipitation exceeded 40 mm at two stations.
19–20 Feb 2000	Low pressure system of ~1006 hPa over north Adriatic Sea (1005 hPa at 1200 UTC 19 Feb) moving southward (~1005 hPa over south Adriatic at 1200 UTC 20 Feb associated with a cold front along the western Greek coasts). Small-scale low of ~1007 hPa formed over north Aegean Sea. The 24-h accumulated precipitation exceeded 40 mm at two stations.
2–3 Mar 2000	Low pressure system over southern Italy (1006 hPa at 1200 UTC 2 Mar) moving rapidly toward the east (1008 hPa over north Aegean at 0600 UTC 3 Mar, associated with a cold front oriented north–south over the Aegean). The 24-h accumulated precipitation exceeded 25 mm at two stations.
24–25 Nov 2000	Low pressure system in the Gulf of Lion (1005 hPa at 0000 UTC 24 Nov) moving eastward (1007 hPa over north Adriatic at 0000 UTC 25 Nov). The system and the associated cold front crossed Greece later during the day. The 24-h accumulated precipitation exceeded 60 mm at one station and 30 mm at three stations.
26–27 Nov 2000	Low pressure system over central Adriatic (1003 hPa at 0000 UTC 27 Nov) moving rapidly eastward. Twelve hours later the low pressure center was located over the northern Aegean. The 24-h accumulated precipitation exceeded 100 mm at one station and 40 mm at two stations.
31 Dec 2000–1 Jan 2001	Low pressure system over Sicily (992 hPa at 1200 UTC 31 Dec) moving rapidly eastward while filling (1001 hPa west of Crete at 1200 UTC 1 Jan). The 24-h accumulated precipitation exceeded 50 mm at one station and 40 mm at two stations.

place the largest forecast amounts where the largest amounts occurred. This statistical measure is considered of great importance because it can help to identify if the model places the bulk of precipitation correctly and, therefore, can assist the forecaster to issue a corresponding warning. The rank correlation ρ is given by the formula

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)},$$

where d is the difference in rank between the observations and forecasts and n is the number of observing stations. The perfect rank correlation equals 1; when the largest forecast amounts are found in places where the smallest amounts were observed, and vice versa, rank correlation equals -1.

Although many verification studies restrict their analyses to the calculation of bias and threat scores, there is a strong need to calculate parameters that evaluate the accuracy of the model quantitative precipitation forecasts and the accurate placement of precipitation maxima and minima. For that purpose, all of the aforementioned statistical measures are applied in all cases of significant precipitation that occurred over the southern Balkans and, especially, over Greece during the cold periods of 1999 and 2000. In general, the cold seasons of 1999 and 2000 were relatively dry over Greece, with no significant problems related to floods. The only case with reported flooding was on 7–8 November 1999 with moderate damage in areas over eastern continental Greece. This case is verified in the following together with 10 other cases with important precipitation amounts over Greece, namely, 5–6 December, 10–11 December, 21–22 December, and 30–31 December in 1999, and 9–10 February, 19–20 February, 2–3 March, 24–25 November, 26–27 November, and 31 December in 2000. These cases correspond to the passage of low pressure systems and associated frontal activity over Greece. Table 1 summarizes the synoptic-scale characteristics and gives the reported maxima of accumu-

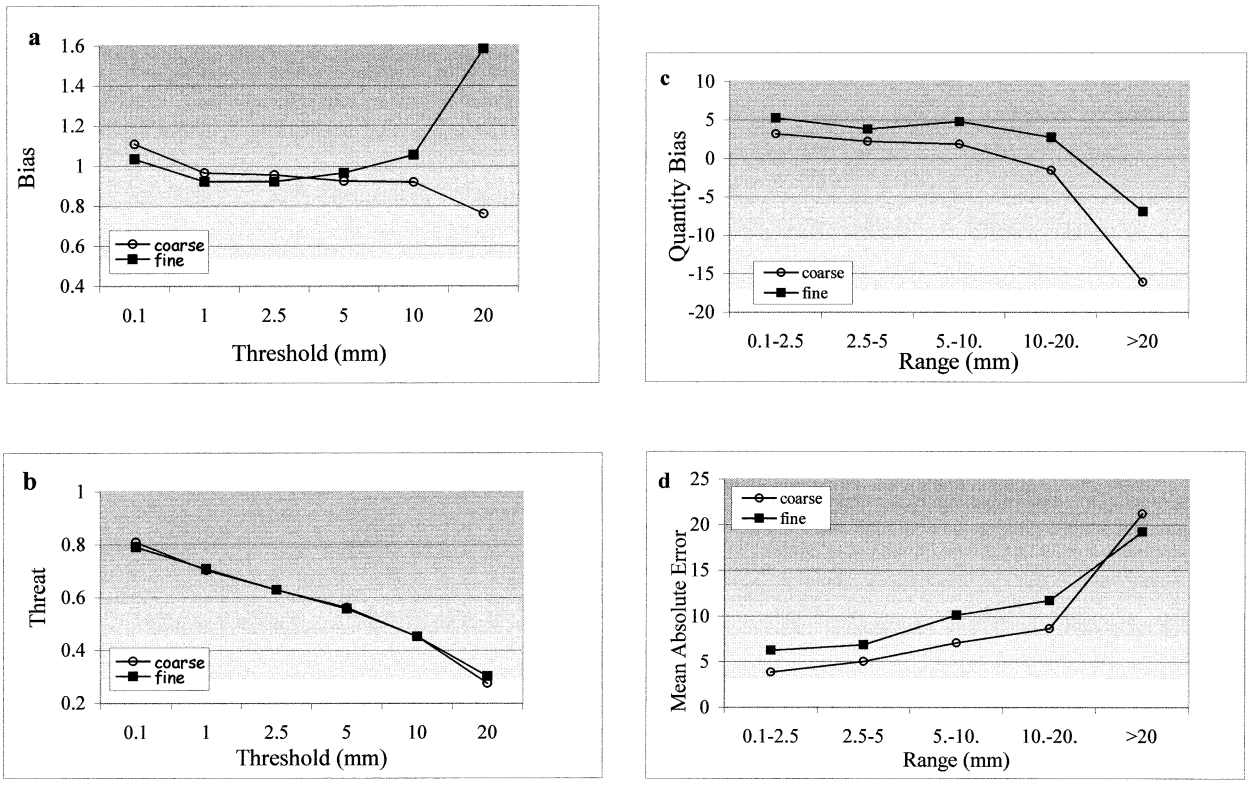

FIG. 4. (a) Bias and (b) threat scores for various precipitation thresholds (mm) and (c) QB and (d) MAE for various precipitation ranges (mm), for 24-h accumulated precipitation of BOLAM, averaged over the 11 analyzed cases.

lated precipitation of each event. The verification of accumulated precipitation has been performed for a 24-h period. In all cases, the observed precipitation (from the network of rain gauges of the Hellenic Meteorological Service) is verified against both the coarse- and the fine-grid forecast precipitation accumulated between $t + 18$ and $t + 42$ forecast hours. For the verification procedure the model forecasts at the four closest points to each rain gauge site are averaged, weighted by the inverse of their squared distance from the rain gauge. On average, 65 rain gauge stations are used for each rain event (see Fig. 1b for the location of rain gauges). The average number of observations per event is 11 for the 01–2.5-mm precipitation range, 7 for the 2.5–5-mm range, 10 for the 5–10-mm range, 10 for the 10–20-mm range, and 12 for the >20-mm range.

Figure 4 presents the results of the verification statistics for the 11 cases analyzed in this study for both the outer- and the fine-grid model forecasts. Inspection of Fig. 4 leads to the following remarks:

- Bias score is very close to 1 (perfect score) in the thresholds from >0.1 to >5 mm and shows a similar behavior for both the coarse- and the fine-grid forecasts (Fig. 4a). The bias ranges from 0.92 to 1.11 for the coarse grid and from 0.92 to 1.06 for the fine grid for these ranges. For higher amounts of rain the bias curves of the coarse and fine grids diverge, indicating an underprediction of the areal extent of precipitation by the coarse grid (a bias of 0.76 for >20 mm) and an overprediction by the fine grid (a bias of 1.58 for >20 mm).
- The threat score shows an overall good forecast skill for both the coarse- and the fine-grid forecasts, which have an almost identical behavior. The threat scores are greater than 0.55 for all thresholds up to 5 mm (Fig. 4b). Even for large precipitation amounts the model still has skill, with threat scores ~0.3.
- The calculation of the quantity bias (Fig. 4c) and of the mean absolute error (Fig. 4d) reveals that the forecasts of both grids overpredict the rain amounts (QB > 0) for the low and medium precipitation ranges. This overprediction is more important for the fine-grid forecasts. At the high precipitation categories, the coarse-grid forecasts substantially underpredict the rain amounts (QB = −16 at the >20 mm range). This can be attributed to the coarse-grid model resolution (0.21°), which does not permit the correct representation of fine-scale convective motions that usually give the highest precipitation amounts. This is a common problem with numerical models (Schultz 1995; Wang and Seaman 1997; Belair et al. 2000). The fine-grid forecasts still underpredict the high precipitation amounts (QB = −7 at the >20 mm range) but to a lower degree, indicating that both the grid-scale re-

solved precipitation from the explicit scheme and the subgrid-scale parameterized precipitation from the convective scheme are better reproduced at that resolution (~0.06°).

Further, one of the most encouraging results is reflected in the calculation of rank correlation that is very high, suggesting that the model is able to place correctly the maxima and minima of rain. Indeed the rank correlation has an average value (over the 11 analyzed cases) of 0.57 for both the coarse- and the fine-grid forecasts. This result is of great importance for an accurate prediction of areas affected by rain and for the early warning by relevant authorities.

From the comparison of the results on the forecast skill of both grids, it is obvious that, for the low and medium precipitation amounts, the finer-grid forecasts are not as good as the coarse-grid forecasts. For precipitation amounts exceeding 20 mm, the forecast skill of the finer grid is poorer than the coarse grid because it concerns the bias score, while it is better than the coarse grid because it concerns the quantity bias and the mean absolute error. Namely, inspection of the case-by-case results showed that, at the precipitation range >20 mm, for 9 out of the 11 analyzed cases the fine-grid quantity bias is significantly smaller than the coarse grid, while for 7 out of the 11 cases the mean absolute error of the fine-grid forecast is better than that of the coarse grid. Based on the statistical scores used in this study, it is not easy to reach a conclusion on the necessity of the use of a fine grid for the more accurate prediction of the high precipitation amounts because the fine-grid skill is better than the coarse-grid skill in some, but not all, of the statistical scores. It is, however, encouraging that for the high precipitation amounts the fine-grid forecasts present better results for a quantitative measure of precipitation such as the quantity bias and the mean absolute error. At this point, one should also consider the adequacy of these commonly used statistical scores for the verification of fine-resolution precipitation forecasts. Ducrocq et al. (2002) discussed this issue and pointed out that spatial errors of the fine-grid precipitation forecasts produce poorer standard objective scores than coarser-grid forecasts, even though the subjective impression of the forecaster is that the fine-grid precipitation fields are much closer to the real precipitation fields. Mass et al. (2002) also argued that high-resolution forecasts will result in lower objective skill scores, even if structures are more realistic in the fine-grid precipitation forecasts, because of small timing and placement errors.

Comparison of the aforementioned results with those provided in the literature implies that BOLAM precipitation forecasts are promising. Kotroni and Lagouvardos (2001) performed an analysis on MM5 forecast skill for the cold period of 1999–2000 over Greece for various combinations of convective and microphysical schemes and found threat scores of 0.66 for the 2.5-mm threshold dropping to 0.26 for the >20-mm threshold, while their calculated MAE was ~7 mm for the precipitation ranges <10 mm, and it increased to 21 mm for the range >20 mm, with the combination of the Kain–Fritsch and Schultz schemes. Wang and Seaman (1997) used MM5 for three cold-season cases and provided threat scores of 0.39 for the 2.5-mm threshold and 0.14–0.18 for the 12.7-mm threshold, while their calculated MAE ranged from 16 to 18 mm, depending on the applied convective parameterization scheme. Belair et al. (2000) presented bias and threat scores of precipitation forecasts from the Canadian operational regional model, with bias values fluctuating around 1.4–1.5, while threat was about 0.35 for the 5-mm threshold and 0.2 for the 20-mm threshold. Mesinger (1996) verified the 48-km Eta Model, and presented bias scores between 1–1.1, for all thresholds up to 25 mm. These results showed an overall lower (or similar) forecast skill than the evaluated precipitation forecast skill of BOLAM. With regard to the model skill to place the maxima and minima of rain, the calculated rank correlation is in most of the cases higher (better) than the rank correlation calculated by Schultz (1995) for a late February storm over the United States, which was ~0.37.

Analysis of the fine-grid precipitation forecasts also revealed a tendency to overestimate the precipitation in the windward slopes of mountains. This characteristic has been identified in some coastal stations in western Greece where the orography is placed upstream from the low-level flow of the precipitating systems usually affecting the area during winter (south-southwestern flow over the Ionian Sea). Such a flow-dependent structure on errors has been also identified by Ghelli and Lalaurette (2000). Colle et al. (2000) and Saulo et al. (2001) also found that an increase in horizontal resolution results in heavier precipitation over the windward slopes. Figure 5 presents the 24-h accumulated precipitation observed and forecasted by both grids for all analyzed events at two coastal stations in western Greece placed upstream from the low-level flow of the analyzed precipitating systems (Kefalonia Island and Methóni, see Fig. 1b for their location). In the same figure, the 24-h-accumulated precipitation observed at a station placed near eastern continental Greece (Skiathos island, see Fig. 1b for location) is given. The fine grid enables the prediction of large precipitation events (e.g., 10 December 1999 at Kefalonia), but occasionally it may largely overestimate local precipitation. Locally, such overestimation can represent a more systematic error due to specific orographic features. For example, the total precipitation over the analyzed cases at Kefalonia from the coarse grid is 98% of the observed, while from the fine grid is 302% of the observed. At the Methóni station (Fig. 5b) the total amounts are 148% and 348% for the coarse and the fine grids, respectively. Both are coastal stations, where upslope motions are enhanced by frictional convergence and terrain features, in cases of wind blowing at low levels from the sea. A

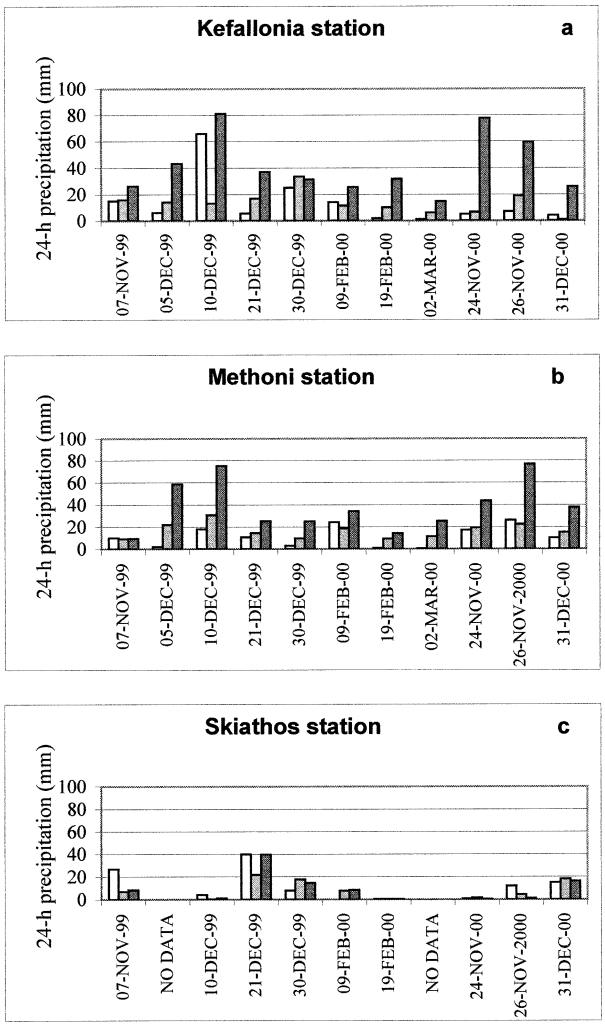

FIG. 5. The 24-h accumulated precipitation observed (white) and forecast by the coarse (light gray) and the fine (dark gray) grids for all analyzed events at two coastal stations in western Greece: (a) Kefalonia island and (b) Methoni and one station in eastern Greece: (c) Skiathos.

closer inspection has revealed that the parameterized convective precipitation, rather than stratiform precipitation, mostly contributes to such a coastal rainfall anomaly. At the coastal station of Skíathos near eastern continental Greece (Fig. 5c), this behavior is not evident. Indeed, the total precipitation over the analyzed cases at Skíathos from the coarse grid is 74% of the observed, while that from the fine grid is 84% of the observed. This behavior is indicative that further investigation and development should be performed for the model treatment of orography and boundary layer physics, as well as for the convective parameterization schemes.

c. Near-surface wind

The near-surface forecasts of wind speed provided by the BOLAM fine grid are compared with the wind speed measured by four buoys deployed by the National Center of Marine Research (NCMR). The buoys are located (Fig. 1b) in the northern Aegean Sea (Athos), the central Aegean Sea (Mýkonos), the Saronic Gulf (Egina), and the southern Aegean Sea (Avgó). The verification of the model performance concerning the wind forecasts over the maritime areas is necessary because of the importance and the implication of such forecasts in two major activities over Greece: maritime transportation (of people and goods) and fishery. The verification procedure includes computation of mean error (ME), MAE, rmse, and correlation coefficient between model and observed winds. The daily observations at 0000 and 1200 UTC for the period September 2000–March 2001 are used, and the model output for the same period is verified for 12-, 24-, 36-, and 48-h forecast lead time. It should be noted that the buoy's sensors measure the winds at ~2 m while the model wind speed is reported at 10 m from the surface. For that reason, the model wind speed has been reduced to 2-m height using the logarithmic wind profile.

The statistics of the comparison are given in Table 2. Model wind speed is in good agreement with observations. The magnitude of the wind speed mean error

TABLE 2. Statistics of the comparison of BOLAM fine-grid wind speed with the wind speed provided by the four buoys, including ME, MAE, rmse, and correlation coefficient (CORR). Boldface is used for the best performance.

	Valid at 0000 UTC—t + 24				Valid at 0000 UTC—t + 48			
Buoy	ME	MAE	Rmse	CORR	ME	MAE	Rmse	CORR
Athos	0.5	1.9	2.6	**0.78**	0.9	2.2	2.9	0.71
Mykonos	−1	2.1	**2.6**	**0.78**	−0.5	2.1	2.6	**0.79**
Egina	0.9	1.9	2.4	0.61	1	2.1	2.7	0.56
Avgo	**−0.2**	**1.7**	**2.2**	0.77	**0.3**	**1.9**	**2.5**	0.73
	Valid at 1200 UTC—t + 12				Valid at 1200 UTC—t + 36			
Athos	1	**1.6**	**2.3**	**0.85**	1.1	2.1	2.9	0.66
Mykonos	−1.1	2	2.6	0.82	−0.7	1.9	2.4	**0.81**
Egina	1.4	2.1	2.7	0.64	1.2	2.1	2.7	0.61
Avgo	**−0.6**	1.8	**2.3**	0.69	**−0.1**	**1.7**	**2.2**	0.74

ranges from −1.1 to 1.4 m s^{-1} with magnitudes of the ME better than 1 m s^{-1} in general. Positive bias denotes overestimation of wind speed by the model. The rmse varies from 2.2 to 2.9 m s^{-1}, and the correlation coefficient varies from 0.61 to 0.85. The forecast skill decreases slightly with increasing forecast lead time. Comparison of the forecast skill among stations shows that the best skill is provided at the station in the maritime area north of the island of Crete (referred to as Avgó), with the second best skill at the station north of the island of Mýkonos in the central Aegean (referred to as Mýkonos). As can be seen in Fig. 1, where the station locations are reported, these buoys suffer the least from eventual modification of the flow from the interaction with local circulations. The worst skill is calculated for the bouy at Egina, located within the Saronic Gulf. The wind flow at this station, especially under weak synoptic forcing, is highly modulated by the local circulations developed because of the proximity of land surfaces characterized by highly complex terrain. The model fine-grid horizontal grid increment of ~6.5 km is then too coarse to accurately resolve such type of circulations. Indeed, if the data at Egina station for measured winds less than 3 m s^{-1} are excluded from the comparison the statistics improve considerably. At 1200 UTC validation time the ME reduces from 1.4 to 0.3 m s^{-1} and the rmse reduces from 2.7 to 2.4 m s^{-1} at t + 12 h forecast lead time, while for the t + 36 forecast lead time, the ME reduces from 1.2 to 0.9 m s^{-1} and the rmse reduces from 2.7 to 2.3 m s^{-1}.

The statistics from the comparison of model-predicted with observed winds provided promising scores in comparison with the scores given in the literature. Bromwich et al. (2001) compared near-surface winds predicted by MM5 with observations over Greenland for a 2-month period. The reported ME ranged from −0.03 to +1.59 m s^{-1}, with magnitudes of the ME generally better than −1 m s^{-1}. The rmse in their study varied from 1.40 to 3.08, while the calculated correlation coefficient varied from 0.43 to 0.84. Hanna and Yang (2001) evaluated simulations of near-surface winds over land by four mesoscale models. They reported rmse consistently about 2 m s^{-1} for the wind speed when four-dimensional data assimilation (FDDA) is used. On the other hand, over complex terrain (Iraq area) and without FDDA, the authors reported rmse of about 5–6 m s^{-1}.

4. Concluding remarks—Prospects

This paper was devoted to the presentation of the results from the operational use of BOLAM at the National Observatory of Athens, the aim of which was to produce detailed weather forecasts over the eastern Mediterranean, and especially over the Greek peninsula. Verification of BOLAM weather forecasts after 2 yr of operational use has shown the model's ability to provide skillful forecasts. Indeed, rmse calculations of geopotential height and temperature at 850 and 500 hPa give scores that are among the best reported in the literature. The rmse of 500-hPa geopotential height average 12 m for both winter and summer at the t + 24 h forecast, and 15 m during summer and 19.5 m during winter at the t + 48 h forecast. For the 850-hPa temperature forecasts, the rmse averages at 1.1°C for both winter and summer at t + 24 h, and at 1.5°C during winter and 1.3°C during summer at t + 48 h. These results further support the notion that regional-scale models present good skill in forecasting upper-air meteorological parameters.

There is still a challenge in the forecast of surface parameters, such as precipitation and winds, because of their important socioeconomic impact. Recent advances in computer power availability give the possibility to test the forecast skill at fine grid spacing in an operational setting. The precipitation forecast skill of BOLAM has been evaluated using two grids: a coarse (~23 km) and a fine (~6.5 km) one, for 11 cases with heavy precipitation over Greece. The verification statistics showed that the threat and bias precipitation scores for the forecasts on both grids are satisfactory for a regional forecast. The calculated rank correlation was also good, suggesting that BOLAM forecasts the placement of maxima and minima of precipitation correctly. Comparison between the coarse- and the fine-grid model forecasts showed that for the low and medium precipitation amounts, the finer-grid forecasts are not as good as the coarse-grid forecasts. For the large precipitation amounts the calculated statistical scores provide only a little support of the idea that the fine-grid forecasts are better than those of the coarse grid because the fine-grid forecasts give better scores only for the quantity bias and the mean absolute error. As was also pointed out by Ducrocq et al. (2002) and Mass et al. (2002), small time and spatial errors in the fine-grid forecasts can have a significant degrading impact in the statistical score results. Thus, in order to verify the fine-grid forecasts, new verification procedures (such as pattern/structure comparison, hydrological response comparison, etc.) may be needed in order to support the necessity of the operational use of finer model grids, which nowadays is more feasible because of the increasing availability of powerful computer infrastructure.

From the analysis it was also evident that the fine-grid forecasts may produce spuriously heavy precipitation over the windward slopes. This latter feature is a challenge in regional-scale modeling for further improvement of moist physics (e.g., use of a shallow convection scheme in order to consume some of the convective available potential energy, convective schemes suitable for fine grid spacings), but also for the treatment of orography and the boundary layer physics included in the model.

The evaluation of the near-surface wind forecasts by the fine-grid simulations with BOLAM over maritime areas of Greece also showed a consistent behavior. The calculated rmse range from 2.2 to 2.9 m s^{-1}, and the

correlation coefficient from 0.61 to 0.85. These results are promising, especially for Greece where maritime transport and fishery are important socioeconomic domains. It should be noted that in the maritime areas affected by proximity to complex terrain (e.g., Saronic Gulf) the forecast errors were larger, and this can be improved by the use of finer horizontal resolutions and also through assimilation of local-scale observations.

The overall behavior of BOLAM, as has been evaluated after 2 yr of operational use, qualifies it as a valuable tool for both forecasters and end users. Some critical features that would improve the model forecast skill include the improvement of model physics and the quality and resolution of the initial and boundary conditions. Indeed, AVN analysis and forecast fields that are currently used have a coarse horizontal resolution of 1.25° and a time interval availability of 6 h. The access to ECMWF analysis and forecast fields, which today offer a better spatial (~40 km) and temporal (3 h) resolution, would considerably benefit the operational forecasts.

Acknowledgments. This work was supported by European Commission/Information Science and Technology (Contract EN-4003) and by the General Secretariat for Research and Technology (Excellence in Research Centers—MIS64563). The Hellenic National Meteorological Service and the Hellenic National Center of Marine Research are kindly acknowledged for providing the observed precipitation and wind data used in this study. The authors are also grateful to NCEP for providing AVN initial and forecast field data, which allowed the operational use of the meteorological model at NOA. Special thanks are given to the three anonymous reviewers of this paper. Prof. D. Lalas (National Observatory of Athens) is also acknowledged for fruitful discussions.

REFERENCES

Bacchi, B., and R. Ranzi, Eds., 2000: RAPHAEL—Runoff and atmospheric processes for flood hazard forecasting and control. Final Rep. to EC, Directorate General XII, Programme Environment and Climate 1994–1998, Contract ENV4-CT97-0552.

Belair, S., A. Methot, J. Mailhot, B. Bilodeau, A. Patoine, G. Pellerin, and J. Cote, 2000: Operational implementation of the Fritsch–Chappell convective scheme in the 24-km Canadian regional model. *Wea. Forecasting,* **15,** 257–274.

Bromwich, D., J. Cassano, T. Klein, G. Heinemann, K. Hines, K. Steffen, and J. Box, 2001: Mesoscale modeling of katabatic winds over Greenland with the Polar MM5. *Mon. Wea. Rev.,* **129,** 2290–2309.

Brooks, H. E., and C. A. Doswell, 1996: A comparison of measures-oriented and distributions-oriented approaches to forecast verification. *Wea. Forecasting,* **11,** 288–303.

Buzzi, A., and L. Foschini, 2000: Mesoscale meteorological features associated with heavy precipitation in the southern Alpine region. *Meteor. Atmos. Phys.,* **72,** 131–146.

——, M. Fantini, P. Malguzzi, and F. Nerozzi, 1994: Validation of a limited area model in cases of Mediterranean cyclogenesis: Surface fields and precipitation scores. *Meteor. Atmos. Phys.,* **53,** 137–153.

——, R. Cadelli, and P. Malguzzi, 1997: Low level jet simulation over the Antarctic Ocean. *Tellus,* **49A,** 263–276.

——, N. Tartaglione, and P. Malguzzi, 1998: Numerical simulations of the 1994 Piedmont flood: Role of orography and moist processes. *Mon. Wea. Rev.,* **126,** 2369–2383.

Colle, B., 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.

Ducrocq, V., D. Ricard, J.-P. Lafore, and F. Orain, 2002: Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Wea. Forecasting,* **17,** 1236–1256.

ECMWF, 2000: Application and verification of ECMWF products in member states and co-operating states. ECMWF Rep., 172 pp.

Ferretti, R., T. Paolucci, W. Zheng, G. Visconti, and P. Bonelli, 2000: Analyses of the precipitation pattern on the Alpine region using different cumulus convection parameterizations. *J. Appl. Meteor.,* **39,** 182–200.

Georgelin, M., and Coauthors, 2000: The second COMPARE exercise: A model intercomparison using a case of a typical mesoscale orographic flow, the PYREX IOP3. *Quart. J. Roy. Meteor. Soc.,* **126,** 991–1030.

Ghelli, A., and F. Lalaurette, 2000: Verifying precipitation forecasts using upscaled observations. *ECMWF Newsl.,* **87,** 9–17.

Gyakum, J. R., and Coauthors, 1996: A regional model intercomparison using a case of explosive oceanic cyclogenesis. *Wea. Forecasting,* **11,** 521–543.

Hanna, S. R., and R. Yang, 2001: Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths. *J. Appl. Meteor.,* **40,** 1095–1104.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.,* **47,** 2784–2802.

——, and ——, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus in Numerical Models, Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 165–177.

Kotroni, V., and K. Lagouvardos, 2001: Precipitation forecast skill of different convective parameterization and microphysical schemes: Application for the cold season over Greece. *Geophys. Res. Lett.,* **28,** 1977–1980.

Koussis, A. D., and Coauthors, 2003: Flood forecasts for urban basin with integrated hydrometeorological model. *J. Hydrol. Eng.,* **8,** 1–11.

Kuo, Y. H., R. J. Reed, and Y. Liu, 1996: The ERICA IOP5 storm. Part III: Mesoscale cyclogenesis and precipitation parameterization. *Mon. Wea. Rev.,* **124,** 1409–1434.

Lehmann, R., 1992: On the choice of relaxation coefficients for Davies lateral boundary scheme for regional weather prediction models. *Meteor. Atmos. Phys.,* **52,** 1–14.

Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.,* **17,** 187–202.

——, M. Tiedtke, and J. F. Geleyn, 1982: A short history of the PBL parameterization at ECMWF. *Proc. ECMWF Workshop on PBL Parameterization,* Reading, United Kingdom, ECMWF, 59–80.

Malguzzi, P., and N. Tartaglione, 1999: An economical second order advection scheme for numerical weather prediction. *Quart. J. Roy. Meteor. Soc,* **125,** 2291–2304.

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

Mesinger, F., 1996: Improvements in quantitative precipitation forecasts with the Eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Meteor. Soc.,* **77,** 2637–2649.

——, T. L. Black, D. W. Plummer, and J. H. Ward, 1990: Eta Model precipitation forecasts for a period including Tropical Storm Allison. *Wea. Forecasting,* **5,** 483–493.

Nagata, M., and Coauthors, 2001: A mesoscale model intercomparison: A case of explosive development of a tropical cyclone (COMPARE III). *J. Meteor. Soc. Japan,* **79,** 999–1033.

Ritter, B., and J. F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.,* **120,** 303–325.

Saulo, A. C., M. Seluchi, C. Campetella, and L. Ferreira, 2001: Error evaluation of NCEP and LAHM regional model daily forecasts over southern South America. *Wea. Forecasting,* **16,** 697–712.

Schultz, P., 1995: An explicit cloud physics parameterization for operational numerical weather prediction. *Mon. Wea. Rev.,* **123,** 3331–3343.

Spencer, P. L., and D. J. Stensrud, 1998: Flash flood events: Importance of the subgrid representation of convection. *Mon. Wea. Rev.,* **126,** 2884–2912.

Wang, W., and N. L. Seaman, 1997: A comparison study of convective parameterization schemes in a mesoscale model. *Mon. Wea. Rev.,* **125,** 252–278.