



Three-model ensemble wind prediction in southern Italy

Rosa Claudia Torcasio¹, Stefano Federico², Claudia Roberta Calidonna¹, Elenio Avolio¹, Oxana Drofa³, Tony Christian Landi³, Piero Malguzzi³, Andrea Buzzi³, and Paolo Bonasoni³

¹CNR-ISAC, Zona Industriale Comparto 15, 88046 Lamezia Terme, Italy

²CNR-ISAC, Via del Fosso del Cavaliere 100, 00133 Rome, Italy

³CNR-ISAC, Via Gobetti 101, Bologna, Italy

Correspondence to: Stefano Federico (s.federico@isac.cnr.it)

Received: 5 December 2015 – Revised: 9 March 2016 – Accepted: 10 March 2016 – Published: 21 March 2016

Abstract. Quality of wind prediction is of great importance since a good wind forecast allows the prediction of available wind power, improving the penetration of renewable energies into the energy market. Here, a 1-year (1 December 2012 to 30 November 2013) three-model ensemble (TME) experiment for wind prediction is considered. The models employed, run operationally at National Research Council – Institute of Atmospheric Sciences and Climate (CNR-ISAC), are RAMS (Regional Atmospheric Modelling System), BOLAM (BOlogna Limited Area Model), and MOLOCH (MOdello LOcale in H coordinates). The area considered for the study is southern Italy and the measurements used for the forecast verification are those of the GTS (Global Telecommunication System). Comparison with observations is made every 3 h up to 48 h of forecast lead time.

Results show that the three-model ensemble outperforms the forecast of each individual model. The RMSE improvement compared to the best model is between 22 and 30 %, depending on the season. It is also shown that the three-model ensemble outperforms the IFS (Integrated Forecasting System) of the ECMWF (European Centre for Medium-Range Weather Forecast) for the surface wind forecasts. Notably, the three-model ensemble forecast performs better than each unbiased model, showing the added value of the ensemble technique. Finally, the sensitivity of the three-model ensemble RMSE to the length of the training period is analysed.

Keywords. Meteorology and atmospheric dynamics (mesoscale meteorology)

1 Introduction

Wind farm power prediction is of great importance for renewable energy applications (Giebel et al., 2011; Monteiro et al., 2009; Pinson et al., 2009). Typically, for the prediction of wind energy on timescales longer than 3–6 h, the models for power prediction include the output of a numerical weather prediction model, which is eventually scaled at the turbine hub height and/or interpolated or averaged over an area representative of the wind farm location (Monteiro et al., 2009; Frías et al., 2009). Hence, the quality of the power forecast at different forecasting ranges depends on the quality of the wind prediction over the area of the wind farm (Alessandrini et al., 2013; Pinson et al., 2007; Von Bremen, 2007). The capability of the wind forecast to enhance the penetration of wind energy into the energy market is particularly important in areas where the grid, i.e. the interconnected system for the distribution of the electricity especially at high and medium tension, is poorly developed (as in southern Italy, where this study is focused) and originally not built to support the distributed input of wind power (Alessandrini et al., 2011).

This paper reports the performance of a three-model ensemble (hereafter referred to as TME) experiment (Krishnamurti et al., 1999, 2000) in southern Italy for 10 m wind prediction. The TME is calculated starting with the forecasts of three different models: RAMS (Regional Atmospheric Modelling System; Cotton et al., 2003; Tiriolo et al., 2015), BOLAM (BOlogna Limited Area Model; Malguzzi et al., 2006), and MOLOCH (MOdello LOcale in H coordinates; Malguzzi et al., 2006). All models are operational at National Research Council – Institute of Atmospheric Sciences and Climate (CNR-ISAC). In their real-time implementation, MOLOCH is nested into the BOLAM

model, which uses initial and boundary conditions taken from the 00:00 UT run of the National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS) global model, while the RAMS model is initialized and driven by initial and boundary conditions obtained from the 12:00 UT run of the ECMWF (European Centre for Medium-Range Weather Forecasts) IFS (Integrated Forecasting System) model. The RAMS forecast lasts 60 h and starts at 12:00 UT: the first 12 h are spin-up time and are discarded, so the RAMS initial validity time is 00:00 UT of the day following the initial day of run. The initial time is 00:00 UT for BOLAM and 03:00 UT for MOLOCH, i.e. 3 h later than BOLAM. This choice is made to avoid the initialization of the high-resolution model directly over global analyses.

Results are shown for southern Italy because the RAMS model in the configuration employed here attains its maximum resolution over that area. The forecasts' final validity is 48 h, and the evaluation comprises a whole year (1 December 2012–30 November 2013). Model outputs, as well as TME evaluations, are available every 3 h. To consider the performance of the TME regarding the wind direction forecast, the comparison between the TME and individual model forecasts is made separately for the horizontal wind speed and the zonal (u) and meridional (v) wind components.

Because of the annual variability of the Mediterranean climate, results are stratified seasonally. Southern Italy is characterized by a typical Mediterranean climate with a high seasonal dependence.

In the summertime anticyclonic conditions prevail over the area and large-scale winds are generally weak. In such conditions breeze circulations develop along the coasts and inland, representing the dominant circulation. Winter cyclones and synoptic-scale circulations dominate in the cold season (from November to March; Bolle, 2012; Federico et al., 2008; Hurrell, 1995).

Another point considered in this paper is the evaluation of the added value of the TME technique compared to the forecast of each unbiased model (Carter et al., 1989; Glahn and Lowry, 1972; Wilks, 2006). For this study, the TME forecast is compared with the forecasts obtained from each individual model after the systematic error (bias) has been removed.

The sensitivity of the TME performance to the choice of the training period is also studied. Finally, the performance of the TME is compared with that of the global IFS model (<http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/cycle-41r1>) of the ECMWF, which is taken here as benchmark, over the same target area.

It is important to note that there are many differences in the physics, dynamics and numerical schemes among the models involved in this study that have an important impact on the models' performance. Among these differences two are of particular importance: (a) BOLAM and IFS are hydrostatic models, while RAMS and MOLOCH are non-hydrostatic, and (b) the horizontal grid spacings of the models involved

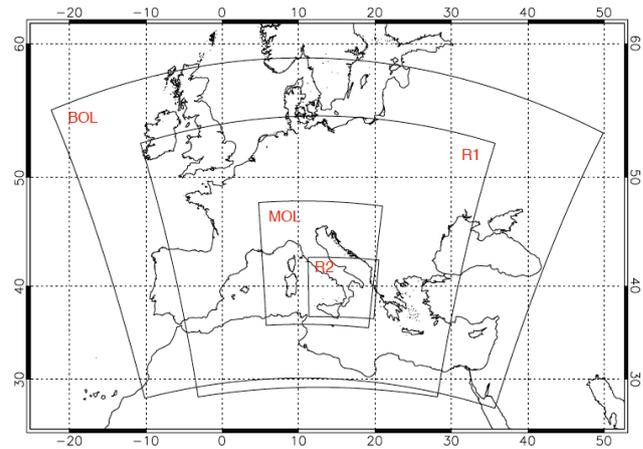


Figure 1. Model domains: BOL shows the BOLAM domain, MOL shows the MOLOCH domain, R1 shows the RAMS first domain, and R2 shows the RAMS second domain. The R2 grid is nested into the R1 grid with a two-way nesting.

are quite different (IFS, ~ 25 km; BOLAM, 10 km; RAMS, 3 km; MOLOCH, 2 km).

For motions in the meso- β and meso- γ (i.e. between 10 and 100 km as breeze circulations, thunderstorms, and mountain gravity waves, among others), the hydrostatic approximation weakens and has an impact on the IFS and BOLAM performance.

For the point (b) above it is important to note that the orographic complexity of southern Italy, involving both the sea-land contrast and mountainous areas, is better resolved as the horizontal grid spacing becomes higher, penalizing the performance of BOLAM and IFS compared to RAMS and MOLOCH.

The paper is organized as follows: in Sect. 2, the model configurations, the data set used for the verification, and the TME procedure are described; results are presented in Sect. 3, while conclusions are given in Sect. 4.

2 Data and methodology

2.1 Models configuration

This study, in addition to the IFS output, uses three meteorological models operational at CNR-ISAC: RAMS, BOLAM, and MOLOCH.

A detailed description of the RAMS model, which is operational in southern Italy (Tiriolo et al., 2015) with a configuration similar to that used in this paper, is given in Cotton et al. (2003). Two two-way nested domains are used (Fig. 1) at 12 and 3 km grid spacing, respectively. The first grid covers the central Mediterranean Basin, while the second extends over the whole of southern Italy. Thirty-five vertical levels, up to about 22 000 m in a terrain-following coordinate system, are used for both domains. Levels are not equally

spaced: layers below 1500 m a.g.l. are between 50 and 200 m thick, whereas layers in the middle and upper troposphere (> 7000 m a.g.l.) are 1000 m thick. The first level is 20 m above the surface and the wind speed is scaled to a height of 10 m with a logarithmic profile. For the experiment in this paper, initial and (6 hourly¹) boundary conditions are given by the ECMWF operational analysis and forecast cycle at 12:00 UT.

The Land Ecosystem–Atmosphere Feedback model (LEAF) is used to calculate the exchange between soil, vegetation, and atmosphere (Walko et al., 2000). LEAF is a representation of surface features, including vegetation, soil (eight levels), lakes and oceans, and snow cover, and their influence on each other and on the atmosphere. Sea surface temperature is kept fixed for each RAMS simulation and is interpolated from the IFS analysis at 12:00 UT of the starting day of the simulation.

Explicitly resolved precipitation is computed by a bulk microphysical scheme, which considers the mixing ratios of seven water categories: cloud water, rain, pristine ice, snow, ice aggregates, graupel, and hail (Walko et al., 1995). The scheme uses a generalized gamma size spectrum and uses a stochastic collection rather than a continuous accretion. It includes a heat budget equation for each hydrometeor class, allowing heat storage and the existence of mixed phase hydrometeors. The sub-grid-scale effect of convective and non-convective clouds is parameterized following Molinari and Corsetti (1985) who modified the Kuo scheme (Kuo, 1974) to account for updrafts and downdrafts. Unresolved vertical transport is parameterized by the *K* theory, in which the covariance is evaluated as the product of an eddy mixing coefficient and the gradient of the transported quantity. The turbulent mixing in the horizontal directions is parameterized following Smagorinsky (1963); it relates the mixing coefficients to the fluid strain rate and includes corrections for the influence of the Brunt–Väisälä frequency and the Richardson number (Pielke, 2002). The radiation scheme detailed in Chen and Cotton (1983) is used for short- and long-wave radiation. The scheme accounts for the total condensate present in the atmosphere.

BOLAM is a hydrostatic, primitive-equation, limited-area model that has been developed at CNR-ISAC since the early 1990s. It has been compared with other mesoscale models in the course of the Comparison of Mesoscale Prediction and Research Experiments (COMPARE), a multi-annual project organized by the World Meteorological Organization (Nagata et al., 2001). MOLOCH is a non-hydrostatic, convection-permitting model developed by CNR-ISAC since early 2000s (Malguzzi et al., 2006) and designed for high horizontal resolution with hybrid terrain-following coordinates. MOLOCH is nested into the BOLAM model with a

one-way procedure. Both models employ rotated latitude–longitude coordinates to optimize computational efficiency.

BOLAM and MOLOCH have in common the physical parameterization schemes, with some adjustments to account for their different resolution. Radiation is based on the ECMWF radiation scheme (Morcrette, 1991). Large-scale precipitation is computed according to single moment bulk microphysics of the Kessler type, with five water species and pristine ice parameterization. Deep convection is parameterized in the hydrostatic BOLAM model only, using a modified Kain–Fritsch scheme (Kain, 2004). The sub-grid turbulence is taken into account with a 1.5 closure that contains a prognostic equation for the turbulent kinetic energy (*E-l* scheme; Zampieri et al., 2005). Surface layer is parameterized according to the Monin–Obukov similarity theory. Finally, a parameterization of soil water and temperature exchange, very similar to the Hydrology-Tiled ECMWF Scheme for Surface Exchange over Land (H-TESSEL), is implemented with seven soil layers.

BOLAM and MOLOCH are used for experimental, real-time forecasting activity under a cooperative agreement with the Italian Civil Protection Agency. Short-term forecasts are made every day with initial and 3-hourly boundary conditions obtained from the GFS model of NCEP, accessible at http://www.isac.cnr.it/_dinamica/projects/forecasts/. Model output of selected surface fields is available every hour and archived in General Regularly-distributed Information in Binary form (GRIB2) format. Current horizontal resolution is 8.5 km for BOLAM (in the European area; 10 km for the experiment considered in this paper) and 1.5 km (2.2 km for the experiment considered in this paper) for MOLOCH over the entire Italian territory (Fig. 1). Both models are implemented with 50 vertical levels. The lowermost BOLAM (MOLOCH) level is located at about 30 m (70 m) above orography, while the model top is about 36 000 m high.

Starting from RAMS, BOLAM, and MOLOCH 10 m wind forecasts, a TME forecast was computed for the period ranging from 1 December 2012 to 30 November 2013. For each day and model, 2-day forecasts are considered, with model outputs and TME forecasts available every 3 h. As stated in the Introduction, RAMS and the BOLAM–MOLOCH chain use different initial and boundary condition data sets.

2.2 Observational data set

The SYNOP reports distributed through the GTS (Global Telecommunication System) are used to evaluate the performance of the models. Data for wind speed and direction are considered, which are converted into the zonal (i.e. west–east direction, *u*) and meridional (i.e. south–north direction, *v*) wind components to assess the ability of the three-model ensemble to improve the forecast of the wind direction.

Verification is made over southern Italy; the stations distributed through the GTS are shown in Fig. 2. The SYNOP data were downloaded from the MARS (Meteorological

¹In this paper the model performance is evaluated every 3 h. However, when simulations were performed with the RAMS model, the boundary conditions were downloaded every 6 h.

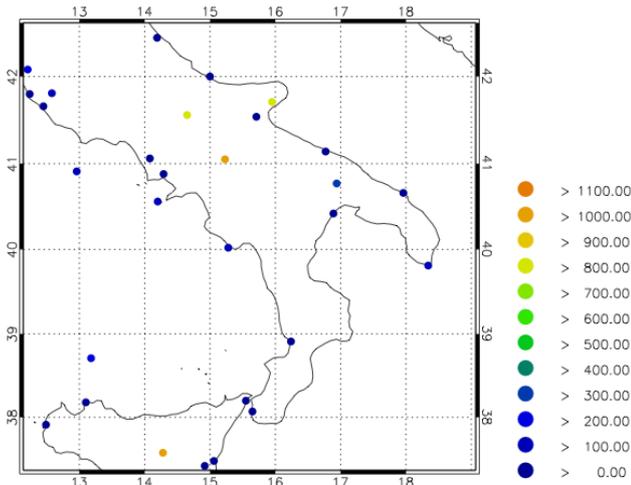


Figure 2. The SYNOP stations distributed through the GTS over southern Italy. The colours represent the station elevation.

Archive and Retrieval System) of the ECMWF. Available reports for each station vary depending on the season and forecasting time because either they are not transmitted through the GTS or they are flagged as missing data. A detailed description of the data available (all stations) for each of the four seasons considered in this work can be found in Table 1. Only the first day is shown because the number of data has (nearly) a 24 h periodicity.

The RAMS, BOLAM, and MOLOCH (hereafter referred to as M1, M2, and M3, respectively) forecasts are interpolated bilinearly to the positions of the SYNOP observations, for all stations available at a given forecast time. We consider the forecast–observed wind speed and the forecast–observed horizontal wind components. Then, bias and RMSE statistics are calculated for each model, as well as for the TME, considering all the forecast–observed pairs available over southern Italy (Fig. 2) at each forecasting time.

2.3 The multi-model approach

In order to improve the performance of each model, the TME post-processing technique is adopted. In this technique (see Krishnamurti et al. (1999, 2000) for a detailed description) several models are weighted with an adequate set of weights computed during a training period. More specifically, the TME forecast at each station location is given by

$$S = \bar{O} + \sum_{i=1}^N a_i (F_i - \bar{F}_i), \tag{1}$$

where N is the number of the models forming the TME, a_i is the weight of the i th model, F_i is the forecast of the i th model, \bar{F}_i is its mean value over the training period, and \bar{O} is the mean observation over the training period.

Table 1. Number of data available at different forecasting times over southern Italy from the SYNOP reports of wind speed and direction for each season. The table shows the first day. For the second day the available data show similar values (compare 0 and 24 h).

Hour	Winter	Spring	Summer	Fall
0	2133	2167	1947	1887
3	2196	2223	2010	1924
6	2802	2855	2585	2569
9	2815	2988	2889	2563
12	2882	2991	2966	2738
15	2927	3050	3000	2684
18	2866	2945	2889	2515
21	2167	2194	2074	1934
24	2133	2166	1947	1890

The calculation of the weight a_i is given by the minimization of the mean square distance D^2 :

$$D^2 = \sum_{k=1}^L (S_k - O_k)^2, \tag{2}$$

where L is the training period length. In this paper, $N = 3$ and the TME is calculated for each station of the GTS of Fig. 2. In principle, the weights could be disaggregated according to the forecast validity or at least the hour of the day. However, lumping together all forecast validities gave comparatively better results and is considered in the following. This issue is likely caused by the need for a longer training data set to give a more reliable estimate of the TME weights, disaggregated according to the forecast validity, compared to that possible in this study.

In the following sections, the TME performance is quantified considering 20 forecast attempts. The performance of the TME computed with weights disaggregated according to the forecast validity gave similar, yet worse, results compared to those shown in the following sections.

The training data set is 80 % of the available data and, to consider the natural variability of the Mediterranean climate, statistics are stratified according to seasons. Therefore, for each season and station, the available data set is divided into two parts: a training period containing 80 % of the data and a forecasting period with the remaining 20 %. The weights a_i are evaluated over the training period and are then used to compute the forecast (“cross evaluation”). Moreover, to get a more robust assessment of the impact of the TME on the wind forecast, Eqs. (1) and (2) are applied 20 times, randomly selecting the 80 % data set of the training period and the 20 % of the forecast period.

In addition to the wind speed, in order to extend the TME forecast to the wind direction as well, we applied the TME to the zonal and meridional wind components separately. For this analysis, wind speed and direction observed by the GTS were converted into horizontal components.

Table 2. RMSE and bias for the TME and for each single model for the four seasons. Scores are relative to the wind speed (WSP) and to the zonal and meridional wind components. The scores of the IFS model are reported as reference. Values in parentheses show the error in percentage of the seasonally averaged wind speed.

		TME		M1		M2		M3		IFS	
		RMSE (ms ⁻¹)	Bias (ms ⁻¹)								
Winter	WSP	2.0 (44)	-0.0 (0)	2.7 (59)	-0.4 (-8)	3.1 (67)	-0.2 (-5)	3.0 (66)	0.3 (6)	3.0 (64)	-0.3 (-6)
	<i>u</i>	2.6	0.0	3.0	0.2	3.3	0.2	3.5	0.6	3.2	0.4
	<i>v</i>	2.6	0.0	3.0	0.2	3.3	0.3	3.4	0.2	3.1	0.0
Spring	WSP	1.9 (43)	0.0 (0)	2.5 (56)	-0.4 (-8)	2.7 (62)	-0.4 (-9)	2.7 (61)	0.1 (2)	2.6 (60)	-0.5 (-10)
	<i>u</i>	2.3	-0.0	2.8	-0.1	3.0	-0.3	3.1	0.1	2.9	-0.0
	<i>v</i>	2.4	0.0	2.8	0.2	2.8	0.1	3.0	0.2	2.7	0.1
Summer	WSP	1.4 (40)	-0.0 (0)	2.0 (56)	-0.5 (-13)	2.0 (58)	-0.5 (-15)	2.0 (57)	-0.2 (-6)	2.0 (56)	-0.7 (-19)
	<i>u</i>	1.9	-0.0	2.1	-0.0	2.2	0.0	2.3	0.0	2.1	-0.0
	<i>v</i>	1.9	0.0	2.2	0.0	2.2	0.0	2.3	-0.1	2.1	-0.0
Fall	WSP	1.8 (46)	0.0 (0)	2.3 (58)	-0.6 (-14)	2.6 (66)	-0.6 (-16)	2.5 (64)	-0.2 (-5)	2.5 (64)	-0.7 (-18)
	<i>u</i>	2.2	0.0	2.5	-0.0	2.8	-0.2	2.9	-0.0	2.7	-0.0
	<i>v</i>	2.1	-0.0	2.5	0.3	2.6	0.2	2.7	0.3	2.5	0.2

3 Results

3.1 Multi-model performance

In this section, the performance of the TME, as well as that of each individual model, is presented and discussed.

Figure 3 shows the RMSE and bias of the TME and of each model of the wind speed for the winter season. The same quantities computed for the IFS are also reported for comparison. The RMSE (Fig. 3a) of the models M1–M3 varies from 2.6 to 3.3 ms⁻¹ depending on the forecast time, with M1 showing the best performance on average, while the RMSE of the TME varies between 1.9 and 2.2 ms⁻¹. From Fig. 3a it is apparent that the TME reduces the RMSE for all forecasting times. This is confirmed by the difference between the RMSE of the best model (of M1–M3) and that of the TME, which is larger than 0.6 ms⁻¹ for all times (lower curve in Fig. 3a). It is worth noting that this difference is, for several forecast times, larger than 20 % of the RMSE of M1–M3 (≈ 3.0 ms⁻¹), showing the important improvement of the TME ensemble on the wind speed forecast.

Regarding the variability of the RMSE for the 20 attempts, a rather stable result is noted, with the interval between the 25th and 75th percentile less than 0.5 ms⁻¹ for all forecasting times.

The increase in the RMSE (and in the RMSE variability over the 20 attempts) at specific times is caused by the distribution of the wind values used for the verification. When the distribution of the wind speed spans a comparatively

larger data interval, as for $t + 18$ h compared, for example, to $t + 15$ h, the RMSE spread increases. Moreover, because of the difficulty to forecast high wind speeds, the RMSE averaged over the 20 attempts increases for wind distributions involving larger winds, as for $t + 18$ h compared to $t + 15$ h.

The IFS performance for surface wind forecast is similar to that of M1–M3, showing the goodness of the IFS forecast despite its coarser horizontal resolution and the fact that IFS is a hydrostatic model. While the use of the GFS initial and dynamic boundary conditions could determine, at specific forecast times, a worse performance of M2 and M3 compared to IFS, we know from previous experience that M2 and M3 underperform compared to IFS also when IFS analyses and boundary conditions (BC) are used. This aspect has to do with the parameterization of surface layer or planetary boundary layer (PBL) and in particular with the fact that the M3 model is poorly resolved above the surface (lowest level at 70–80 m). Moreover, the RMSE favours smoother large-scale fields compared to higher-resolution smaller-scale fields.

However, the TME outperforms M1–M3 and IFS. The RMSE improvement of TME is further quantified in Table 2, which shows the averaged value of the RMSE over all forecast times for each model and season. Focusing on winter, we note that the RMSE of M1–M3 is between 59 and 67 % of the averaged wind speed in winter, while the RMSE of TME is reduced to 44 % of the averaged wind speed in winter.

Another important aspect of the wind speed forecast is its bias. Figure 3b shows the winter bias for M1–M3, TME, and

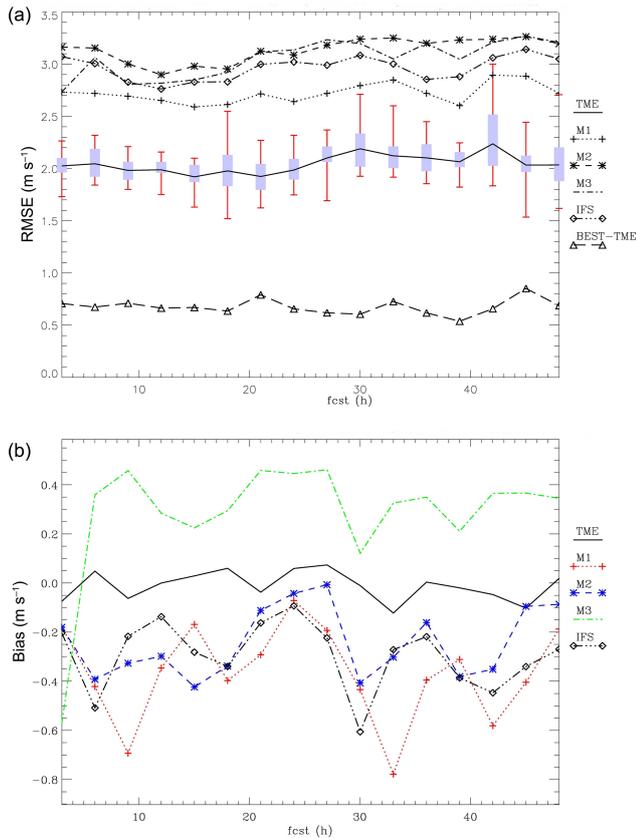


Figure 3. RMSE (a) and bias (b) for the winter wind speed versus forecast time. M1, M2, and M3 are the three models forming the TME, while the IFS model is shown for comparison. The difference between the RMSE of the best model (of M1–M3) and that of the TME is also shown in (a). The boxes on the TME RMSE curve show the 25th and 75th percentile of the RMSE distribution for the 20 attempts, while the error bars extend between the maximum and minimum value of the RMSE for the 20 attempts.

IFS. The TME bias is closer to 0 compared to other models for each forecast time. This is also shown by the values of Table 2, which shows a bias of about 0.3 ms^{-1} for all models (varying between -8 and 6% of the averaged wind speed value in winter). The bias is negative for all models but M3, with an absolute value less than 0.05 ms^{-1} for the TME case. The values of Table 2 show that the TME wind speed forecast is approximately unbiased.

To consider the variability of the results with the season, Fig. 4a and b show the RMSE and the bias, respectively, of the summer wind speed. Figure 4a shows that the best model changes with forecasting time among M1, M2, and M3. The RMSE values for M1–M3 and IFS are between 1.8 and 2.3 ms^{-1} , depending on forecast time.

From Fig. 4a it is apparent that the TME outperforms the models M1–M3. In particular, the difference between the RMSE of the best model and that of the TME is above

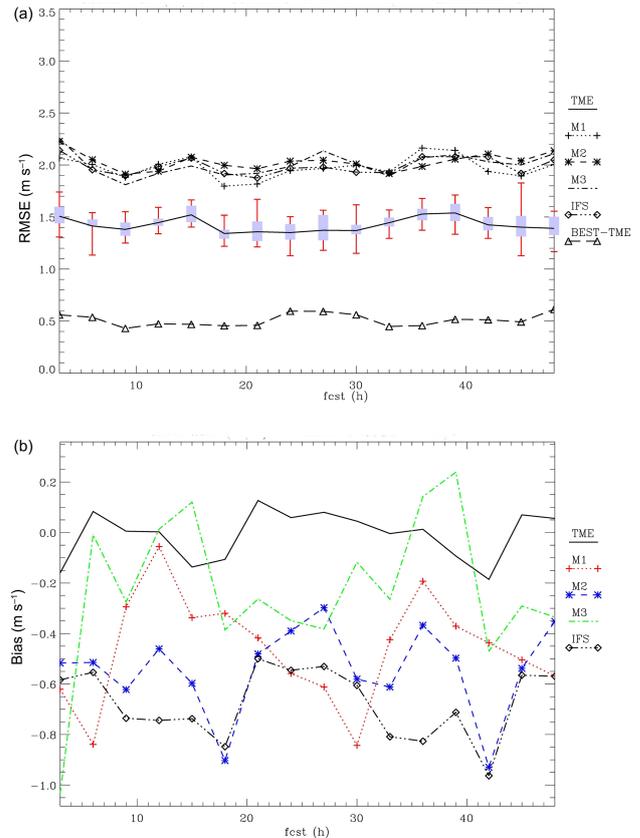


Figure 4. Same as Fig. 3 but for the summer season.

0.4 ms^{-1} for all forecasting times, showing that the TME RMSE is about 20 % less than those of M1–M3.

Considering the RMSE of the TME for the 20 attempts, we note that the interval between the 25th and 75th percentile is lower than 0.3 ms^{-1} for all forecast times, showing the statistical robustness of the multi-model. A diurnal cycle of the RMSE with larger values in the afternoon is also noted. This is caused by the development of sea breeze circulations, which increases both the surface wind speed and its variability among the stations involved.

The results of Fig. 4a are further quantified in Table 2. In summer, considering the average over all forecasting times, the RMSE for M1–M3 and IFS is about 2.0 ms^{-1} (55–60 % of the summer averaged wind speed). The RMSE reduces to 1.4 ms^{-1} (40 % of the summer averaged wind speed) for TME.

The bias in summer shows again that the TME is approximately unbiased. More specifically the averaged value of the bias over the forecasting times varies between -0.2 and -0.5 ms^{-1} for M1–M3, while its absolute value is less than 0.05 ms^{-1} for the TME.

The results for spring and fall are considered in Table 2. We note that the RMSEs for M1–M3 and IFS are between the values of summer and winter. In all cases, the TME has a relevant impact on the wind speed forecast because the RMSE

is reduced from 56 % (best model) to 43 % of the averaged wind speed in spring and from 58 % (best model) to 46 % of the averaged wind speed in fall.

Regarding the bias, the results of Table 2 show that the TME bias is, in absolute value, less than 0.05 ms^{-1} , while it varies between 0.1 and 0.6 for M1–M3. It is also noted that the TME outperforms the IFS model both for the bias and RMSE in all seasons and that the IFS errors are similar to those of M1–M3 showing the skill of the IFS forecast.

Table 2 shows that the RMSE of the wind speed has its maximum in winter and its minimum in summer, while spring and fall RMSEs are between those of summer and winter. This seasonal behaviour is determined by the different seasonal wind regimes in the central Mediterranean (Bolle, 2012). More specifically, winds over southern Italy are dominated by the breeze circulation in summer. This is especially valid for the SYNOP stations of Fig. 2, which are mainly located by the sea. On the other hand, the incidence of the cyclones over southern Italy is at a maximum in winter. As a consequence, the winds are less intense and follow a more regular pattern in summer compared to winter, and the wind forecast shows a comparatively better performance. Spring and fall share characteristics with both summer and winter and their wind forecast performance is between that of summer and fall.

This characteristic of the Mediterranean climate is well represented by the year considered in this paper. Figure 5 shows the wind speed averaged for all SYNOP stations of Fig. 2 for the different seasons of the year. It is apparent that the largest wind speeds occur in winter (with the exception of 09:00 UT when the maximum is attained in spring), when the incidence of the cyclones is at a maximum (Bolle, 2012), while the lowest wind speeds occur in summer, when the wind regime is dominated by the breeze (Tiriolo et al., 2015). Spring and fall share characteristics with both summer and winter and wind speeds have values between those of summer and winter.

Figure 5 also shows a diurnal cycle in all seasons due to the thermally forced local circulations (Federico et al., 2010; Mangia et al., 2004). The amplitude of the cycle is largest in summer and smallest in winter because of both the greater insolation in summer and the larger number of cyclones that cross southern Italy in winter. Again the fall and spring behaviour is between that of summer and winter.

To show the impact of the TME on the forecast of the wind direction, we consider the results for the ensemble applied to each horizontal wind component (Table 2). It can be seen that the improvement of the TME is found for all seasons and for both RMSE and bias. For the RMSE the improvement is larger than 10 % of the M1–M3 RMSE, showing a sizable impact of the TME on the zonal wind components forecast, while the bias evaluation indicates that the TME is approximately unbiased. The comparison between the TME RMSE and that of the best model for each forecasting time

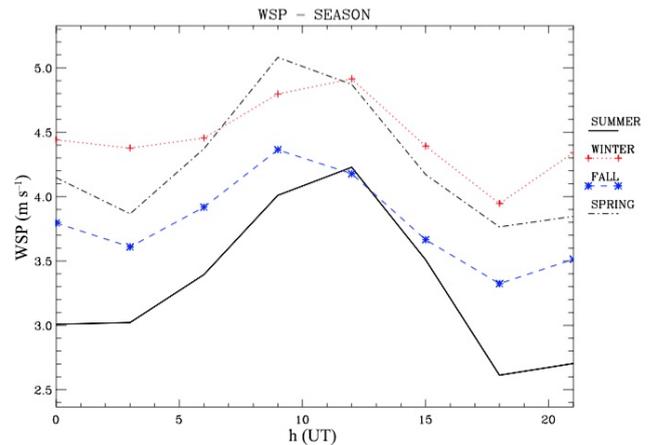


Figure 5. Averaged wind speed for the different seasons as a function of the hour of the day (every 3 h). The average has been computed grouping all the available data of the SYNOP stations for the period 1 December 2012–30 November 2013.

shows a reduction of 0.4 ms^{-1} for both components and for all forecasting times (figures not shown).

A common feature evidenced in Figs. 3 and 4 is the small error growth with increasing forecast lead time. In fact, almost all the RMSE is already present at the first instant (+3 h) and at the time of initial condition (00:00 UT, not shown). This is true for all the models considered here, regardless of the originating centre (GFS or IFS) from which the initial conditions are taken, and for the 10 m wind downloaded from the ECMWF MARS archive. The mismatch between analyses and SYNOP observations may be due to many reasons, such as approximations in the model orography, surface layer parameterization, and overweighting of background with respect to observations in the assimilation procedures.

Overall, this section shows that the TME technique represents an important improvement in the forecasting of the wind speed and horizontal wind components for the model configurations in the area considered in this paper.

3.2 Comparison between multi-model and unbiased models

In this section we investigate the added value of the TME compared to the unbiased forecast (Carter et al., 1989; Glahn and Lowry, 1972; Wilks, 2006). More specifically, the TME forecast (the TME is computed from biased models) is compared with the forecast of M1–M3 (and IFS for reference) after applying bias removal from all model outputs.

Figure 6 shows the RMSE versus forecast time of the wind speed for the winter season. We note that the RMSE of M1–M3 is between 2.2 and 2.7 ms^{-1} , significantly lower than that reported in Fig. 3a (about 0.5 ms^{-1} less, i.e. 17 % of the M1–M3 and IFS RMSE). This improvement, which is the consequence of the bias removal from single-model forecasts, has

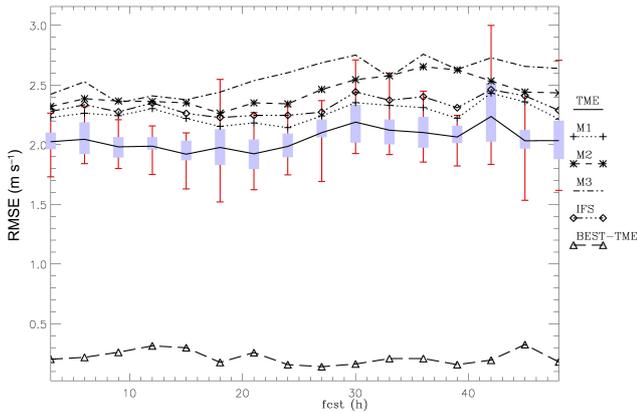


Figure 6. Winter RMSE for the wind speed after bias elimination. The difference between the RMSE of the best model (depending on the forecast time) and that of the TME is also shown (BEST-TME).

an important impact on the forecast of wind speed. In particular, the relative error computed in respect of the seasonally averaged wind speed decreases by more than 10 % for M1–M3 (compare Tables 2 and 3). Nonetheless, the TME forecast for wind speed is still better than that of all unbiased forecasts, the TME RMSE being between 2.0 and 2.2 ms⁻¹ for all forecast times. This is confirmed, in Fig. 6, by the difference between the best model and the TME forecast, which is between 0.2 and 0.3 ms⁻¹.

While the error of the unbiased forecast of M1–M3 ranges between 47 and 57 % of the seasonal average wind speed value, the TME relative error is between 43 and 46 %, showing the added value of the TME approach. Table 3 shows that the TME outperforms the unbiased IFS forecast, which, however, shows performance comparable to the best model of M1–M3.

Similar results are found for the zonal and the meridional wind components (not shown), whose difference between the RMSE of the best model and that of the TME is between 0.3 and 0.5 ms⁻¹, depending on the forecast time.

Considering the variability of the TME RMSE for the 20 attempts, a variability less than 1 ms⁻¹ is noted for most forecasting times, while the interval between the 25th and 75th percentile is in most cases less than 0.5 ms⁻¹.

Table 3 shows the average RMSE over all forecasting times of the unbiased models. Focusing on winter wind speed, we note that the RMSE of the best unbiased model is 2.3 ms⁻¹, while the RMSE of the TME is 2.0 ms⁻¹. Table 3 further shows that the improvement is resilient because it occurs in all seasons. The TME RMSE for winter is 2.6 ms⁻¹ for both wind components, while that of the best M1–M3 model is 2.9 ms⁻¹.

Comparing the results of Tables 2 and 3, it is apparent that the unbiased forecast for the zonal and meridional wind components has a lower improvement with respect to that of the wind speed. This occurs in all seasons. This smaller improve-

Table 3. Same as Table 2 but for the unbiased models (RMSE only). Values in parentheses show the error in percentage of the seasonally averaged wind speed.

		TME RMSE (ms ⁻¹)	M1 RMSE (ms ⁻¹)	M2 RMSE (ms ⁻¹)	M3 RMSE (ms ⁻¹)	IFS RMSE (ms ⁻¹)
Winter	WSP	2.0 (44)	2.3 (47)	2.4 (52)	2.6 (55)	2.3 (50)
	<i>u</i>	2.6	2.9	3.2	3.4	3.1
	<i>v</i>	2.6	2.9	3.2	3.3	3.0
Spring	WSP	2.0 (45)	2.2 (50)	2.3 (52)	2.4 (55)	2.2 (50)
	<i>u</i>	2.3	2.8	2.9	3.1	2.9
	<i>v</i>	2.4	2.7	2.7	2.9	2.7
Summer	WSP	1.4 (43)	1.7 (49)	1.7 (49)	1.8 (51)	1.6 (47)
	<i>u</i>	1.9	2.1	2.1	2.2	2.0
	<i>v</i>	1.9	2.1	2.1	2.2	2.1
Fall	WSP	1.8 (46)	2.0 (51)	2.1 (54)	2.2 (57)	2.0 (51)
	<i>u</i>	2.2	2.5	2.8	2.8	2.7
	<i>v</i>	2.1	2.3	2.4	2.6	2.4

ment is caused by the smaller bias of the zonal and meridional wind forecast, so that bias removal is less effective for these quantities. Table 3 shows that, in all cases, the TME performs better than the unbiased models, and that the TME outperforms the unbiased IFS forecast as well.

Overall, the results of this section show that the TME improves the forecast compared to the unbiased models, providing an additional value.

3.3 Variability with the training period

In this section, the sensitivity of the TME forecast to the length of the training period is analysed. Figure 7 shows the RMSE difference between the best model and the TME, for the winter wind speed, for a different training data set, namely 90, 80, 60, and 40 % of the available data. It is important to point out that the RMSE differences shown in Fig. 7 are computed over the same data set, so that the curves of Fig. 7 are directly comparable. This data set is that remaining when the TME is trained with 90 % of the available data, which is the largest data set not used for training any TME (90, 80, 60, and 40 %).

Figure 7 shows that the TME forecast improves the wind speed forecast, also using 40 % of the available data set for training. These results indicate the potential of the TME for wind speed forecast even when using a short training period (36 days for each season in this case).

As expected, the TME performance worsens for shorter training periods. The difference among the different training periods may become substantial for specific forecast times. For example, at 12 h of forecast time, the TME RMSE in-

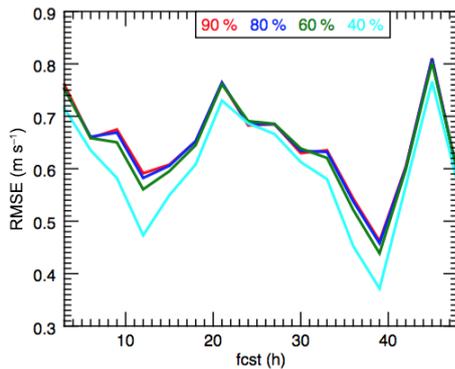


Figure 7. Difference between the RMSE (best model minus TME) of the winter wind speed forecast for different percentages of training data.

creases by more than 0.1 ms^{-1} when 40 % of the training data set is used instead of 80 %.

4 Conclusions

This paper shows the performance of a TME forecasting system for the near-surface wind prediction over southern Italy. The study was motivated by the need to improve the quality of near-surface wind forecast, in particular to enhance the penetration of renewable energies into the energy market.

The TME is formed by the models RAMS, BOLAM, and MOLOCH, which are all currently operational at CNR-ISAC. The study extends over a 1-year span (1 December 2012 to 30 November 2013), and the results are disaggregated on a seasonal basis to account for the natural variability of the Mediterranean climate. In addition to the wind speed, the impact of the TME technique on the zonal and meridional wind components is considered in order to show the potential of the TME prediction of the wind direction.

In general, it is found for all models considered that the error increase with forecast lead time is small relative to the average error present in the initial conditions. Results show that the TME forecast improvement for the wind speed is larger in winter (0.7 ms^{-1} of RMSE reduction with respect to that of the best M1–M3 model) and smaller in fall (0.5 ms^{-1}). In all seasons, the RMSE of the TME is reduced by at least 10 % with respect to the best model, this improvement often being around 20 %. The TME improvement is resilient because it is found for all seasons and for all forecasting times. These results show the important impact of the TME on the wind speed forecast.

Another important result is that the TME bias for the wind speed is lower than those of the single models. Averaged over all forecasting times, the TME bias is, in absolute value, less than 0.05 ms^{-1} . Similar results are found for the zonal and meridional wind components, so that we can confirm that the TME is also effective for wind direction prediction.

The TME forecast is compared with the IFS forecast, which is taken here as a benchmark. Despite the lower spatial horizontal resolution, the IFS performance for surface wind forecast is similar to that of RAMS, BOLAM, and MOLOCH, showing the skill of the IFS forecast. It is found that the TME outperforms the IFS forecast for all seasons and forecast times.

The added value of the TME is further studied by comparing the wind forecast made by each unbiased model and the TME forecast. It is shown that the TME forecast outperforms each single unbiased model for all seasons and forecast times and that the TME RMSE is at least 10 % less of the RMSE of the best unbiased model. Again, the TME outperforms the IFS unbiased forecast for all seasons and forecast times.

The final point considered in this paper is the stability of the TME evaluation with the length of the training period. It is found that a training period of 36 days, i.e. 40 % of the available data for each season, already gives a sizable improvement to the TME RMSE. Nevertheless, the performance difference between training performed with 40 and 80 % of the available data is significant for specific forecasting times, showing the importance of using the largest possible data set for training.

The TME forecast quality may be further increased, at least in principle, by making the TME weights a function of the forecast time. Such a disaggregation would, however, require a longer training data set in order to compute reliable model–model and model–observation covariances. This generalization will hopefully be accomplished in the near future when more years of forecast will be available.

Acknowledgements. The authors acknowledge the CNMCA (Centro Nazionale di Meteorologia e Climatologia Aeronautica) and the ECMWF for access to the MARS (Meteorological Archive and Retrieval System) database. This work is partially supported by projects PON04a2_E Sinergreen-ResNovae-“Smart Energy Master for the energetic government of the territory” and PONa3_00363 “High Technology Infrastructure for Climate and Environment Monitoring” funded by the Italian Ministry of University and Research (MIUR) PON 2007–2013. The second author thanks Claudio Tranterici for the maintenance of the computing resources. Two anonymous reviewers and the editor are acknowledged for their useful comments and support.

The topical editor, Vassiliki Kotroni, thanks two anonymous referees for help in evaluating this paper.

References

- Alessandrini, S., Pinson, P., Hagedorn, R., Decimi, G., and Sperati, S.: An application of ensemble/multi model approach for wind power production forecasting, *Adv. Sci. Res.*, 6, 35–37, doi:10.5194/asr-6-35-2011, 2011.
- Alessandrini, A., Sperati, S., and Pinson, P.: A comparison between the ECMWF and COSMO Ensemble Prediction Systems applied

- to short-term wind power forecasting on real data, *Appl. Energy*, 107, 271–280, 2013.
- Bolle, H. J.: *Mediterranean Climate, trends and variability*, Springer, Berlin, 384 pp., 2012.
- Carter, G. M., Dallavalle, J. P., and Glahn, H. R.: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system, *Weather Forecast.*, 4, 401–412, 1989.
- Chen, C. and Cotton, W. R.: A One-Dimensional Simulation of the Stratocumulus-Capped Mixed Layer, *Bound.-Lay. Meteorol.*, 25, 289–321, 1983.
- Cotton, W. R., Pielke Sr., R. A., Walko, R. L., Lista, D. E., Tremback, C. J., Jiang, H., McAnelly, R. L., Harrington, J. Y., Nicholls, M. E., Carrio, G. G., and McFadden, J. P.: RAMS 2001: Current status and future directions, *Meteorol. Atmos. Phys.*, 82, 5–29, 2003.
- Federico, S., Avolio, E., Pasqualoni, L., and Bellecci, C.: Atmospheric patterns for heavy rain events in Calabria, *Nat. Hazards Earth Syst. Sci.*, 8, 1173–1186, doi:10.5194/nhess-8-1173-2008, 2008.
- Federico, S., Pasqualoni, L., Sempreviva, A. M., De Leo, L., Avolio, E., Calidonna, C. R., and Bellecci, C.: The seasonal characteristics of the breeze circulation at a coastal Mediterranean site in South Italy, *Adv. Sci. Res.*, 4, 47–56, doi:10.5194/asr-4-47-2010, 2010.
- Frías, L., Pascal, E., Irigoyen, U., Cantero, E., Loureiro, Y., Lozano, S., Fernandes, P. M., and Martí, I.: Support Vector Machines in the wind energy framework. A new model for wind energy forecasting, *Proc. of the 2009 European Wind Energy Conference EWEC'09*, Marseille, France, 9 pp., 2009.
- Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., and Draxl, C.: The State of the Art in Short-Term Prediction of Wind Power, *Anemos.plus deliverable report*, 110 pp., [http://orbit.dtu.dk/files/5277161/GiebelEtAl-StateOfTheArtInShortTermPrediction_ANEMOSplus_2011_\(2\).pdf](http://orbit.dtu.dk/files/5277161/GiebelEtAl-StateOfTheArtInShortTermPrediction_ANEMOSplus_2011_(2).pdf) (last access: 17 March 2016), 2011.
- Glahn, H. R. and Lowry, D. A.: The use of model output statistics in objective weather forecasting, *J. Appl. Meteorol.*, 11, 1203–1211, 1972.
- Hurrell, J. W.: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation, *Science*, 269, 676–679, 1995.
- Kain, J.S.: The Kain–Fritsch convective parameterization: an update., *J. App. Meteorol.*, 43, 170–181, 2004.
- Krishnamurti, T. N., Kishtawal, C. M., LaRow, T., Bachiochi, D., Zhang, Z., Williford, C. E., Gadgil, S., and Surendran, S.: Improved weather and seasonal climate forecasts from multi-model superensemble, *Science*, 285, 1548–1550, 1999.
- Krishnamurti, T. N., Kishtawal, C. M., Zhang, Z., LaRow, T., Bachiochi, D., Williford, E., Gadgil, S., and Surendran, S.: Multi-model Ensemble Forecasts for Weather and Seasonal Climate, *J. Climate*, 13, 4196–4216, 2000.
- Kuo, H. L.: Further Studies of the Parameterization of the Influence of Cumulus Convection on Large-Scale Flow, *J. Atmos. Sci.*, 31, 1232–1240, 1974.
- Malguzzi, P., Grossi, G., Buzzi, A., Ranzi, R., and Buizza, R.: The 1966 “century” flood in Italy: A meteorological and hydrological revisitation, *J. Geophys. Res.*, 111, D24106, doi:10.1029/2006JD007111, 2006.
- Mangia, C., Martano, P., Miglietta, M. M., Morabito, A., and Tanzarella, A.: Modelling local winds over the Salento Peninsula, *Meteorol. Appl.*, 11, 231–244, 2004.
- Molinari, J. and Corsetti, T.: Incorporation of cloud-scale and mesoscale down-drafts into a cumulus parametrization: results of one and three-dimensional integrations, *Mon. Weather Rev.*, 113, 485–501, 1985.
- Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., and Conzelmann, G.: Wind Power Forecasting: State-of-the-Art 2009, Argonne National Laboratory ANL/DIS-10-1, November 2009, Argonne National Laboratory, 198 pp., <http://ceeesa.es.anl.gov/pubs/65613.pdf> (last access: 17 March 2016), 2009.
- Morcrette, J. J.: Radiation and cloud radiative properties in the ECMWF operational weather forecast model, *J. Geophys. Res.*, 96D, 9121–9132, 1991.
- Nagata, M., Leslie, L., Kamahori, H., Nomura, R., Mino, H., Kurihara, Y., Rogers, E., Elsberry, R. L., Basu, B. K., Buzzi, A., Calvo, J., Desgagne, M., D'Isidoro, M., Hong, S. Y., Katzfey, J., Majewski, D., Malguzzi, P., McGregor, J., Murata, A., Nachamkin, J., Roch, M., and Wilson, C.: A mesoscale model intercomparison: A case of explosive development of a tropical cyclone (COMPARE III), *J. Meteorol. Soc. Jpn.*, 79, 999–1033, 2001.
- Pielke, R. A.: *Mesoscale Meteorological Modeling*, Academic Press, San Diego, 676 pp., 2002.
- Pinson, P., Nielsen, H. A., Madsen, H., and Kariniotakis, G.: Skill forecasting from different wind power ensemble prediction methods, *J. Phys. Conf. Ser.*, 75, 012046, doi:10.1088/1742-6596/75/1/012046, 2007.
- Pinson, P., Nielsen, H. A., Madsen, H., and Kariniotakis, G.: Skill forecasting from ensemble predictions of wind power, *Appl. Energy*, 86, 1326–1334, 2009.
- Smagorinsky, J.: General circulation experiments with the primitive equations. Part I, The basic experiment, *Mon. Weather Rev.*, 91, 99–164, 1963.
- Tiriolo, L., Torcasio, R. C., Montesanti, S., and Federico, S.: Verification of a Real Time Weather Forecasting System in Southern Italy, *Adv. Meteorol.*, 2015, 758250, doi:10.1155/2015/758250, 2015.
- Von Bremen, L.: Combination of deterministic and probabilistic meteorological models to enhance wind farm forecast, *J. Phys. Conf. Ser.*, 75, 012050, doi:10.1088/1742-6596/75/1/012050, 2007.
- Walko, R. L., Cotton, W. R., Meyers, M. P., and Harrington, J. Y.: New RAMS cloud microphysics parameterization part I: the single-moment scheme, *Atmos. Res.*, 38, 29–62, 1995.
- Walko, R. L., Band, L. E., Baron, J., Kittel, T. G., Lammers, R., Lee, T. J., Ojima, D., Pielke Sr., R. A., Taylor, C., Tague, C., Tremback, C. J., and Vidale, P. L.: Coupled Atmosphere–Biosphere–Hydrology Models for environmental prediction, *J. Appl. Meteorol.*, 39, 931–944, 2000.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, 2nd Edn., Academic Press, Amsterdam, 627 pp., 2006.
- Zampieri, M., Malguzzi, P., and Buzzi, A.: Sensitivity of quantitative precipitation forecasts to boundary layer parameterization: a flash flood case study in the Western Mediterranean, *Nat. Hazards Earth Syst. Sci.*, 5, 603–612, doi:10.5194/nhess-5-603-2005, 2005.