

AROME-WMED, a real-time mesoscale model designed for the HyMeX Special Observation Periods

N. Fourrié¹, É. Bresson¹, M. Nuret¹, C. Jany¹, P. Brousseau¹, A. Doerenbecher¹, M. Kreitz², O. Nuissier¹, E. Sevault¹, H. Bénichou³, M. Amodei³, and F. Poupponeau³

¹CNRM-GAME, Météo-France and CNRS, Toulouse, France

²École Nationale de la Météorologie, Météo-France, Toulouse, France

³Météo-France, Toulouse, France

Correspondence to: N. Fourrié (nadia.fourrie@meteo.fr)

Abstract

During autumn 2012 and winter 2013, two Special Observation Periods (SOPs) of the Hydrological cycle in the Mediterranean Experiment (HyMeX) took place. For the preparatory studies and to support the instrument deployment during the field campaign, a dedicated version of the operational convective-scale AROME-France model was developed: the AROME-WMED model. It covers the western Mediterranean basin with a 48 h forecast range. It provided real time analyses and forecasts which were sent daily to the HyMeX operational centre to forecast high precipitation events and to help decision makers on the deployment of **meteorological** instruments. This paper presents the main features of this numerical weather prediction system in terms of data assimilation and forecast. Some specific data of the HyMeX SOP were assimilated in real time.

The forecast skill of the AROME-WMED is then assessed with objective scores and compared to the operational AROME-France model, for both autumn 2012 (5 September to 6 November 2012) and winter 2013 (1 February to 15 March 2013) SOPs. **The overall performance of AROME-WMED is good for SOP1 (i.e. mean 2m temperature root mean square error (RMSE) of 1.7 °C and mean 2 m relative humidity RMSE of 10% for the 0-30-h forecast ranges) and similar to those of AROME-France for the 0 to 30 h common forecast range (maximal absolute difference of 2 m temperature RMSE of 0.2 °C and 0.21% for the 2m relative humidity). For the 24 to 48 h forecast range is of course less accurate (relative loss between 10% and 12% in 2 m temperature and relative humidity RMSE, and ETS for 24-h accumulated rainfall) but it remains useful for scheduling observation deployment.** The characteristics of parameters such as precipitation, temperature or humidity, are illustrated by one heavy precipitation case study that occurred over the south of Spain.

1 Introduction

HyMeX (HYdrological cycle in the Mediterranean EXperiment, Drobinski et al., 2014) is a 10 year international programme. It aims at a better understanding and quantification of

the hydrological cycle and related processes in the Mediterranean Sea, with an emphasis on high-impact weather events, inter-annual to decennial variability of the Mediterranean coupled system, and associated trends in the context of global climate change.

The first HyMeX Special Observation Period (SOP1) took place from 5 September to 6 November 2012 over the north-western Mediterranean Sea including the French, Italian and Spanish coastal regions. SOP1 was dedicated to heavy precipitation and flash-flooding (Ducrocq et al., 2014). The second field campaign (SOP2) was set for February–mid-March 2013 in the Gulf of Lion and aimed at studying intense air-sea exchanges and dense water formation (Drobinski et al., 2014). In this region, dense water is formed at the surface by a drop in temperature and an increase in salinity. This is due to the action of strong heat and water losses (strong evaporation), generally associated with strong northerly/north-westerly wind events (Mistral/Tramontane) during the autumn and winter (Marshall and Schott, 1999). If the surface water is dense enough, a violent mixing occurs (i.e. deep ocean convection), sometimes reaching as far as the sea-floor (2000 m-depth).

AROME-WMED (Application of Research to Operations at MESoscale, West MEDiteranean Sea) is the HyMeX-dedicated version of the French operational convective-scale AROME-France system (Seity et al., 2011). It was developed in 2009 and constantly improved since then by CNRM-GAME (Centre National de Recherches Météorologiques–Groupe d'études de l'Atmosphère Météorologique) to study heavy precipitation and to provide the HyMeX operational centre (HOC) during the several SOP campaigns with real time analyses and forecasts. AROME-France, from which AROME-WMED is derived, leaves a large part of Mediterranean Sea out of its domain. As several studies have shown the importance of an accurate description of the moist low-level flow (that feeds meso-scale convective systems which can result in heavy precipitating events over the Mediterranean Sea, (Duffourg and Ducrocq, 2011; Bresson et al., 2012; Ricard et al., 2012)), we choose to extend further south the AROME-WMED domain (Fig. 1). In addition, AROME-WMED ran with its own assimilation cycle in quasi real time from the beginning of September 2012 to mid-March 2013.

In the past, the AROME model had been already used during several field campaigns, such as the Mesoscale Alpine Programme-Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events (MAP-D-PHASE, Rotach et al., 2009), the Convective and Orographically induced Precipitation Study (COPS, Wulfmeyer et al., 2008) or the Boundary-Layer Late Afternoon and Sunset Turbulence (BLLAST, Lothon et al., 2014). In the latter experiment, only specific AROME forecast output was made available every hour over the small BLLAST domain. To be able to make forecast during MAP-D PHASE and COPS experiments, an AROME domain was created over the Alps. This model was initialised using ALADIN-France, which was at the time the operational regional Météo-France model, taking its lateral boundary conditions from ARPEGE and its initial state from a three-dimensional variational data assimilation (3D-Var) scheme (Fischer et al. (2006)). This AROME model was run during 6-months (June-November 2007). Only during the last month of the COPS experiment (August 2007), this specific AROME was run in real-time with an assimilation cycle.

During the SOPs, some AROME-WMED fields (surface fields and precipitation) were used for driving hydrological and ocean models (Lebeauupin-Brossier et al., 2014). AROME-WMED simulations were also used for decision making on observation deployments and safety management of the instruments.

During the campaign, innovative observations came from Boundary Layer Pressurized Balloons (BLPBs) developed by CNES (Centre National d'Etudes Spatiales, Doerenbecher et al., 2013) and from flights of the German Dornier and the French SAFIRE Falcon 20 and ATR-42 aircraft. Flight plans and dropsonde launches were planned on the eve following AROME-WMED forecasts. Radiosondes were also launched from mobile platforms along the French and Italian Mediterranean coasts and in Corsica according to meteorological situations. Moreover additional operational radiosondes were activated on request at 06:00 and 18:00 UTC through the Data Targeting System (DTS) implemented by ECMWF (European Centre for Medium-range Weather Forecasts, Prates et al., 2009) in the EUCOS (EUMETNET Composite Observing System) framework (former EUMETNET EIG Observation Programme).

The aim of this paper is to review the main characteristics of the AROME-WMED real-time version in terms of data assimilation and forecast. The AROME-WMED model configuration is described in Sect. 2. In Sect. 3, the performances of AROME-WMED and AROME-France models are evaluated during the SOP1 over a common area. The comparison is based on Météo-France operational scores and on scores computed with additional surface observations from the HyMeX database. Section 4 is dedicated to the model evaluation during the SOP2. Section 5 illustrates the AROME-WMED forecast for the Intensive Observation Period 8 (IOP8) case that affected southern Spain with heavy precipitation and flash flooding. The last section concludes this paper.

2 Description of the AROME-WMED model

2.1 Model configuration

AROME-WMED is based on AROME-France, which is a limited area model that rests upon non-hydrostatic equations (Bénard et al., 2010). Both models have a 2.5 km grid and 60 vertical levels ranging from 10 m above ground to 1 hPa. Both models use a 1-moment microphysical parameterization (Pinty and Jabouille, 1998; Caniaux et al., 1994), which takes into account five classes of hydrometeors (cloud liquid water, cloud ice, rain, snow and graupel). The vertical turbulent transport in the boundary layer is represented by two schemes: an eddy diffusivity part based on a prognostic turbulent kinetic energy parametrization following Cuxart et al. (2000) and a dry thermal and shallow convection mass flux scheme following Pergaud et al. (2009). The deep convection is explicitly resolved in the AROME numerical weather prediction system. The surface scheme, called Externalized Surface (SURFEX, Masson et al., 2013), diagnoses the 2 m temperature, 2 m humidity, and 10 m wind at every time step due to a specific algorithm called Surface Boundary Layer (SBL) (Masson and Seity, 2009). Please refer to Seity et al. (2011) for further details.

The AROME-France domain is comprised between 38° N– 11° W and 55° N– 15° E (Fig. 1). AROME-WMED covers an extended domain (+11 % more grid points than the operational

one) from Portugal to Italy and from North-Africa to France (34° N– 11° W/ 48° N– 20° E, Fig. 1). It is particularly well suited for the study of high precipitation events which occur in eastern Spain, in the Cévennes area (France) and in Liguria-Tuscany (Italy), all of which are located in the middle of the domain, hence, the upstream atmospheric flow and air-sea exchanges leading to heavy precipitation are also included in the area. For AROME-France, the grid is defined by 750×720 points from the centre at 46.4° N 2.2° E, whereas the AROME-WMED one is defined with 960×640 points from 41.5° N 4.1° E. 47 % of grid points are covered by sea surface versus 41 % for the AROME-France model.

As both domain and grid points of AROME-WMED and the AROME-France differ, it was felt necessary to use new climatologies for the SURFEX scheme (Masson et al., 2013). The same method as the one used in AROME-France was used to set up these climatologies (Seity et al., 2011). Physiographic data are initialized over the AROME-WMED domain using the so-called ECOCLIMAP database (Masson et al., 2003) at 1 km resolution. As in AROME-France, the SURFEX orography comes from the one of the atmospheric model to avoid issues with altitude inconsistencies. It is extracted from the Global 30 Arc-Second Elevation Data Set (GTOPO30) database.

For forecast purposes, lateral boundary conditions are hourly provided by the French operational global model ARPEGE (Action de Recherche Petite Echelle Grande Echelle, Courtier et al., 1991) forecasts. Once per day, a 48 h forecast was started at 00:00 UTC with AROME-WMED. Only one 30 h forecast was made by the operational AROME-France model at 00:00 UTC but also at 06:00, 12:00 and 18:00 UTC. This choice for AROME-WMED was motivated by computational costs and by the forecasters' requirements during HyMeX SOPs. Indeed, from this 48 h forecast, it was possible to study the weather conditions for the next day during the 07:30 UTC briefing while, at Météo-France, it is made routinely from the 12:00 UTC forecast.

2.2 Data assimilation

The initial atmospheric state of the AROME-WMED is an analysis performed every 3 h with a three dimensional variational data assimilation system (3D-Var) at the horizontal resolu-

tion of 2.5 km (Seity et al., 2011). Each 3 h step computes an analysis using observations within a ± 1 h 30m assimilation window and a 3 h forecast to serve as first guess for the next step. The wind horizontal components together with the temperature, the specific humidity and the surface pressure are analyzed, other model fields such as cloud parameters are cycled from the AROME guess. The surface analysis (soil temperature, soil humidity over land and sea surface temperature) is produced with an optimal interpolation scheme every 3 h, taking into account the data measured by surface stations and buoys.

2.2.1 Background error statistics

The background error covariance matrix is an essential component of any variational data assimilation system. It strongly controls the spread of the observation impact within a variational mesoscale analysis. The background error covariances of AROME-France cannot be used, as both domains are different. The error covariances of the AROME-WMED data assimilation system, specially developed for the HyMeX campaign, has been computed using the ensemble approach proposed by Brousseau et al. (2011). Several studies have indeed demonstrated the strong dependence of the background error statistics on weather regimes. For instance, using the Weather Research and Forecasting (WRF) model at 3 km resolution, Michel et al. (2011) showed that the specific humidity standard deviation of the model background error is larger in rain affected areas than in non rainy conditions.

With AROME-France (Brousseau et al., 2012), daily and spatially averaged background errors show some important variations: the specific humidity (resp. temperature) values of the background error variance in the low troposphere are higher in warm (resp. anticyclonic) conditions than in cold (resp. low pressure) ones. These authors also demonstrated an improvement on the analyses and the forecasts of AROME-France when background error covariances are consistent to the meteorological situations.

However, accuracy and computational costs of representing background-error statistics are major issues in data assimilation, therefore, compromises are a necessity when it comes to operational applications. To introduce the flow dependency into the background error covariances proves to be too costly, hence, a climatological background error repre-

sentation (i.e. spatially and temporally averaged statistics over 2 week period) is used in AROME-France and in AROME-WMED instead. The period chosen for the computation of the operational AROME-France model background was characterized by various meteorological conditions (warm and anticyclonic conditions during the first week (25 April to 2 May 2010), cold and perturbed conditions during the second one (3 to 10 May 2010), giving a good meteorological sample). As one of the goals of the HyMeX campaign was to observe and forecast heavy precipitation, and bearing in mind these variations, an autumnal period in October 2010, characterized by convective systems over Catalonia and the south of France, was selected to compute the AROME-WMED background error covariances.

The comparison of the statistics, standard deviation and correlation length, of AROME-WMED with the AROME-France ones, shown in Fig. 2, is quite consistent with findings by Michel et al. (2011) and Brousseau et al. (2012). This temperature background error shape in AROME-France for the lowest levels is due to the estimation of the error covariances, which occurred during a period with more anticyclonic and stable meteorological conditions than for the AROME-WMED one. The warmer period, which had been chosen to compute the background error covariances of AROME-WMED leads to larger standard deviations of the specific humidity background error (above 1 g kg^{-1} in the lowest levels) as compared to AROME-France (below 1 g kg^{-1} in the lowest levels). This increase of the standard deviation holds true for the whole troposphere (Fig. 2a).

The background error standard deviation for the vorticity and divergence fields (not shown) presents similar patterns. No systematic difference was found for the temperature background error (Fig. 2b), except in the lowest levels, below 850 hPa. This AROME-France temperature background error shape in the lowest levels is due to the estimation of the error covariances, which occurred during a warmer and more stable atmospheric period than for AROME-WMED one. The vertical profiles of correlation length scales for humidity and temperature (Fig. 2c and d) exhibit shorter horizontal correlations for AROME-WMED above 850 hPa, in accordance with the higher values in variance spectra for horizontal length-scales below 300 km (not shown).

The vertical correlation reflects the relationship between the various layers of the atmosphere. AROME-WMED forecast vertical error correlations are larger in the troposphere than the AROME-France ones (Fig. 3), especially in the lower troposphere, below 850 hPa, due to a more important vertical mixing present during the period of the AROME-WMED background error covariance computation.

All these results are consistent with the findings of Brousseau et al. (2012) in convective situations and should help the AROME-WMED data assimilation system to extract more information on convective phenomena from observations.

2.2.2 Observations

The observations assimilated are of the same type as the ones used in the operational AROME-France system: radiosondes, automated land surface stations, wind profilers, ships, buoys and aircraft reports. Satellite data are also used: infra-red radiances from SEVIRI (Spinning Enhanced Visible and InfraRed Imager) on board the geostationary satellite Meteosat Second Generation, infra-red and microwave radiances from polar orbiting satellites, satellite winds from Atmospheric Motion Vectors (AMVs) and scatterometer surface winds. The GPS Zenith Total Delay (GPS-ZTD) observations from the EUMETNET EIG GNSS water VApour Programme (E-GVAP) network are assimilated as well. Doppler winds (Montmerle and Faccani, 2009) and reflectivities which provide information on relative humidity profiles (Caumont et al., 2010; Wattrelot et al., 2014) from the French Radar network are also used.

Some observations from the HyMeX campaign, when distributed through the Global Telecommunication System (GTS), were assimilated in AROME-France, AROME-WMED and in the global ARPEGE model. The radiosondes deployed for the DTS and the additional SOP1 ones were assimilated by AROME-WMED and AROME-France in real-time as they were transmitted through the GTS. Thus, 373 out of the 403 additional EUCOS radiosondes launched during the SOP1 were assimilated in AROME-WMED.

Data from 18 BLBPs were also assimilated during SOP1 and 16 during SOP2 (Fig. 4). These BLBPs were launched on request from the HyMeX Operational Centre. These data

were available in real-time only in the Météo-France database, temperature and wind data were thus assimilated in AROME-France and AROME-WMED. This represents a total of nearly 850 reports (351 in SOP1 and 497 in SOP2). From these BLPBs, it was possible to collect data close to the sea surface, which is generally a data void area. The high rate of observations (two per minute) was filtered and averaged ($10'$) to generate representative observational reports, from a meteorological point of view, which were then coded in the form of commercial aircraft data with distinctive identifiers. These reports included humidity (mixing ratio) in addition to temperature, pressure, wind and geographical elevation. The winds were derived from the drift of balloons, thanks to a 3-D GPS positioning. The average range of the altitude was 400 to 900 m. **Only data from stable parts of the flights were used to generate this special dataset, which has been derived from raw BLPB data by time filtering. Some balloons ascended rapidly when encountering strong up-drafts which were generated by deep convection. The data corresponding to these rapid vertical variations were discarded. The sampling ratio of the BLPBs is not meant to capture very rapid changes in the measured parameters.** In AROME-WMED, only temperature and wind data (at pressure levels) were assimilated in real-time, whereas humidity was not, because the format used for the observation dissemination did not allow it.

Ancillary surface data from AEMET (Agencia Estatal de METeorología), also have been assimilated in real-time only in AROME-WMED because these data were available on a ftp site.

Figure 5 shows the amount of assimilated observations by both models over their specific domain and over their common area, ranging from 39 to 47.5° N and from 8° W to 11° E for SOP1 and SOP2. Aircraft data are fewer in the AROME-WMED model as most major European airports lie outside the domain. The same holds true for GPS data which are denser over northern Europe. Despite real time constraints, many more satellite microwave observations were assimilated through an increase of the horizontal **sampling** (one observation per 80 km² square box, instead of one per 125 km² box in AROME-France) which gave a better observation coverage over the sea similar to the operational 80 km **sampling** of IASI (**Infra-red Atmospheric Sounding Interferometer**) (Guidard et al., 2011) in AROME-France.

Additional Spanish surface data were used in the AROME-WMED analyses to improve coverage over Spain as shown in Fig. 5.

In total, over the common area, 3.3 % of additional observations are taken into account in AROME-WMED for SOP1 and 5.2 % for SOP2. The discrepancy is due to the increased horizontal resolution for the microwave radiance assimilation (+35.7 % for SOP1 and +42.4 % for SOP2) and to a later AROME-WMED analysis (half an hour later than the AROME-France). It follows that more observations were available in the meteorological database for the assimilation. Buoys provided around 5 % additional data for both SOPs as the assimilation was made after that of AROME-France. However fewer data from Atmospheric Motion Vectors (AMVs), scatterometers, profilers and ground GPS stations were used during SOP1. The lack of assimilated ground GPS data in AROME-WMED (−22.4 and −21.4 % for SOP1 and SOP2 respectively) was caused by an inadequate bias correction in AROME-WMED, which discarded a lot of GPS observations from assimilation.

The main difference between the results of SOP2 and those of SOP1 lies in the activation of 4 additional wind profilers (+27 %) in southern France. A quality control was made against the AROME-WMED model during SOP1. Once assessed, these data were assimilated in AROME-WMED in real-time, as from December 2012.

2.3 Runs in real-time

AROME-WMED was run in real-time on the Météo-France NEC-SX8 computer dedicated to research numerical activities (i.e. not on the computer dedicated to operations). Though the model configuration was developed in a research environment, it was necessary to modify it in order to run it more quickly and without service break. Hence, it was monitored 24 h a day during the 9 month period (September 2012–March 2013) by the desk operator team to rerun it in case of model crash. The AROME-WMED data assimilation task was launched about 2 h after the analysis time and it took approximately 1 h and 30 min on 8 dedicated nodes to perform a 48 h forecast (assimilation took 30 min). Finally, for 00:00 UTC slot, model outputs were available at around 04:30 UTC on Météo-France forecaster display workstations. This schedule was reached in 95 % of cases for SOP1 and 93 % for SOP2.

Worthy to be mentioned, the cycle 36 code version of AROME-WMED was upgraded to cycle 37 on 25 September 2012 at 06:00 UTC in a way similar to the operational model upgrade (AROME-France and ARPEGE). This means, during SOP1, the meteorological fields were produced by two different versions of the code whereas the outputs of SOP2 come from only one AROME-WMED version. For technical reasons, it was not possible to have a single version of the AROME-WMED during SOP1. The main changes in the AROME model reside in the revision of the cloud scheme with a realistic increase of clouds with partial cloud cover, in addition to changes concerning observation use. As it was not possible to run simultaneously both cycles of the AROME-WMED, we could not quantify the impact of the code change on the AROME-WMED forecast during SOP1. However, the evaluation of the code change in the operational AROME-France suite has shown that low-level cloud fields were altered and the precipitation were slightly improved.

AROME-WMED outputs were extensively used for the deployment of the BLPBs (Ducrocq et al., 2014). For instance, atmospheric profiles from AROME-WMED were used to determine the boundary layer structure and to build atmograms over the launch site in Menorca. Strong surface winds and heavy rain are problematic when launching the balloons. Moreover, a systematic forecast of the balloon trajectories starting every 3 h and at 3 different levels was made every day. This includes the prediction of location of some simulated balloons but also their environmental parameters such as pressure, temperature, humidity and occurrence of rain. When flying, all balloon trajectories were extrapolated from the latest observed position using the latest AROME-WMED forecast in order to predict the next steps of the balloon trajectories. This update procedure was done every 20 min.

For safety reason, the dropsondes were not allowed to reach inhabited land. The drift of the dropsondes launched by the research Falcon 20 aircraft was thus extrapolated from AROME-WMED forecasts.

A data subset composed of precipitation, temperature, wind, wet bulb temperature, relative humidity, brightness temperature, convective available potential energy, integrated water vapour, mean sea level pressure, sea surface temperature, radar reflectivity was also made available in real-time to HyMeX Operational Centre website for secondary opera-

tional centres (L'Aquila, Italy; Palma de Mallorca; Mahon, Menorca; San Giuliano, Corsica; Toulouse, France). The real-time fields of AROME-WMED were also stored in the HyMeX database (doi:10.6096/HYMEX.AROME_WMED.2012.02.20) for future process studies. They are available in NetCDF and Grib formats at the following address: <http://mistrals.sedoo.fr/HyMeX/>.

3 Forecast evaluation during the first Special Observation Period (SOP1)

Forecasters and scientists daily monitored AROME-WMED and AROME-France forecasts. As domains and forecast ranges of both models are different, the comparison focussed on a common area defined by 38°N – 47.5°N / 8°W – 12°E . Forecasts were compared up to the 30 h range, being the last available one for AROME-France at the time of the HyMeX campaign. Models were evaluated against observation data, which are subject to errors and biases but in this study they are used as a reference and assumed to represent the truth.

3.1 Temperature, relative humidity at 2 m and wind at 10 m

The observations used for the evaluation came from the HyMeX database, where the SYNOP (surface synoptic observations) network over the HyMeX area was stored. To complete this network, more 2 m temperature and 2 m relative humidity were also provided each hour by France, AEMET and Catalonia. The wind at 10 m was also provided from some surface stations. These data are spatially unevenly distributed (Fig. 6). The network is denser over Spain, France and Switzerland than elsewhere. Over the ocean, it was supplemented by data from ships, buoys and marisondes. The forecast error bias and the forecast root mean square errors (RMSE) were computed for these surface parameters for each forecast lead time with respect to the surface stations (Fig. 7). The values of temperature were corrected to account for the difference in altitude between observations and models which assume that variations in temperature are linear with respect to the altitude (-0.65°C per dam). The number of observations used for the comparison is also plotted and varies as a function of time, with fewer data during night-time whereas during daytime this number

is quite constant at around 120 000 data for temperature. Temperature biases and RMSE are affected by the diurnal cycle. The bias in both models is positive during night-time with a maximum at 06:00UTC and negative during day-time with a minimum at 15:00UTC (Fig. 7a). The absolute bias values are slightly larger for AROME-WMED than for AROME-France, between 0.02 and 0.03 °C on average for AROME-WMED. The positive bias in 2 m temperature during night-time is the result of an excessive coupling of the Masson and Seity, 2009 scheme between the surface and the lowest level of the model. The negative bias during day-time corresponding to the delay in the temperature increase is also well known but remains yet to be explained. The 2 m temperature RMSE are similar for AROME-WMED and AROME-France. The biases for the 24–48 h ranges follow the same pattern as those of the first 24 h, but RMSE increases for the 24–48 h range (about 0.2 °C). A minimum in relative humidity bias is found at 6 h and maximum at 15 h (Fig. 7b). In that case, the error difference between both models nearly reaches 1 %. As for 2 m temperature, the RMSE are similar between both models and increase for the 24–48 h range. Concerning the wind at 10 m (Fig. 7c), the AROME-WMED mean bias is lower at 12:00UTC and larger during night-time between 18:00 and 06:00UTC with an 0.2 m s⁻¹ overestimation. The RMSE, varying between 1.5 and 1.9 m s⁻¹, is also a little larger than the one in AROME-FRANCE.

Rainaud et al. (2015) evaluated AROME-WMED over the sea against measurements of the freighter Marfret-Niolon, moored buoys (“Côte d’Azur” buoy and “Gulf of Lion” buoy) and Marisonde drifting buoys during SOP1. A good agreement with these observations was found for the mean sea level pressure and the 2 m air temperature. The scores for 10 m wind and 2 m relative humidity are slightly worse. The most significant biases for the 2 m air temperature were found locally in the Gulf of Lion during a severe Mistral/Tramontane wind event (28 October 2012). The values obtained by Rainaud et al. agree with those we have obtained with land surface stations.

After each surface station had provided data for two whole months, statistics were computed for each of them, thus enabling the study of the spatial differences between models. Figures 8 and 9 exhibit maps of biases in temperature at 2 m for AROME-WMED and AROME-France for the 6, 12, 18 and 24 h forecast ranges computed over the period from 7

September to 5 November 2012. As many Spanish stations were absent during the first two days of SOP1 (Fig. 6), these days were discarded from the computation period. The observation coverage varies with the analysis time, because the dataset is made from SYNOP stations and climatological stations which are not automated. For example, there are fewer data over Spain at 06:00 UTC (Fig. 8a and b) than at 00:00 (Fig. 9c and d), 12:00 and 18:00 UTC. Only SYNOP station data are available for each analysis time. Whereas the bias is slightly negative over the Iberian Peninsula and France at 12:00 UTC, it becomes positive at 18:00 UTC over south-western France and central Spain, excepting over the mountains in south-eastern France and in Corsica. This pattern is also found for the 42 h forecast range (not shown). This positive bias in 2 m temperature of AROME-France during night-time is well known and it is due to the excessive coupling of the scheme between the surface and the lowest level of the model. The Masson and Seity (2009) surface scheme tends, in fact, to overestimate the surface temperature during night-time and summer. Figure 9 at 18:00 and 00:00 illustrates this behaviour for night-time. The delay in the increase in temperature at 2 m during summer during day-time (i.e. at 12:00 UTC in Fig. 8c and d) is also well known but not yet explained.

3.2 24 h accumulated precipitation

The evaluation was carried out using the 24 h accumulated precipitation (5 September to 5 November 2012), from the HyMeX database available on 1 January 2014. These data had been subject to a quality control prior to computing scores. Duplicates and outliers were removed from the validation dataset. For France and Spain, the 24 h accumulated precipitation data were directly stored in the HyMeX database, whereas for Italy, they were computed by adding each 1 h precipitation data for 24 h periods. Hence, if one 1 h interval was missing, the total 24 h precipitation could not be calculated. A good coverage is obtained over France, Spain and the north-west of Italy, where stations had provided data from 06:00 to 06:00 UTC next day, for the whole 62 days of the SOP1 period (Fig. 10). The coverage is less extensive for the accumulated precipitation between 00:00 and 00:00 UTC the following day because neither SYNOP nor climatological data over France are available.

Half of the stations had at least one missing data and 7% of the whole dataset only had one missing data.

The performance of quantitative precipitation forecast was evaluated with frequency bias (ratio between the frequency of forecast to the observed frequency) and Equitable Threat Scores (ETS) for the 24 h accumulated precipitation (Fig. 11), computed with all the available surface data plotted in Fig. 10. The mathematical formulae used for the computation come from the appendix of Ebert (2008) and are given in Appendix A. The closer to 1 the ETS is, the better the prediction. The quantitative scores are similar for AROME-WMED and AROME-France up to the 5 mm threshold. The ETS is slightly larger for AROME-WMED model for thresholds over 3 mm but is not statistically significant since these cases are fewer than for lower thresholds. Indeed, for small thresholds ($0.1\text{--}3\text{ mm }24\text{ h}^{-1}$), the number of available observations varies between 30 000 and 65 000. For a 60 mm threshold, there are only 683 observations and only 108 observations for the 100 mm threshold. For the first 24 h period, AROME-WMED and AROME-France have a frequency bias close to 1.15 (ideal value is 1.0), it is even 5% higher in AROME-WMED than in AROME-France (Fig. 11). This leads to a higher false alarm rate in the prediction of the 24 h precipitation in AROME-WMED, but also to a better probability of detection as compared to AROME-France. For day +1 (precipitation predicted between 24 and 48 h), the ETS is smaller and follows the same pattern as the 24 h precipitation ETS which decreases with an increasing threshold. For day +1 forecast, the AROME-WMED frequency bias decreases to around 1.1 for thresholds lower than 10 mm per 24 h.

The Brier skill score (BSS) is a probabilistic score relative to persistence for different thresholds of accumulated precipitation. It has been computed operationally with French surface data by converting the deterministic rain forecasts into frequencies beyond discrete thresholds in the neighbourhood of a 54 km radius in order to apply a fuzzy verification method (Amodei and Stein, 2009). Its formula is given in Appendix A. Observations are processed in the same way. This method introduces a tolerance for faulty rainfall location and alleviates the problem of the double penalty, whereby an event that is observed and

predicted with the correct intensity but at the wrong place may be penalized twice in a conventional point-by-point score. The closer to 1 the BSS is, the better is the prediction.

The BSS for the total 24 h accumulated rain between 6 and 30 h ranges decreases with the threshold for both models but it is above 0.8 for AROME-WMED (Fig. 12). As for standard scores, results are similar for both models with a slight improvement by AROME-WMED for thresholds exceeding 5 mm per 24 h.

Figure 13 shows the accumulation of the 24 h precipitation between 06:00 and 06:00 UTC the next day, for the whole SOP1 measured by surface stations and predicted by AROME-WMED and AROME-France models over the period 5 September 2012 to 5 November 2012. Only surface stations with daily precipitation for all the 62 days of the SOP1 are plotted. France is well covered by surface observations followed by Spain thanks to AEMET, which provided not only data for the assimilation but also additional climatological surface data (rain rate, temperature, humidity and wind) over the Iberian peninsula from the beginning of September 2012 to Spring 2013. This dataset is more comprehensive than the one used for the real-time assimilation. **Some raingauge data were missing sometimes during the SOP1, the date depending of the station.** A thorough comparison cannot be undertaken by simply removing a single date. To be able to compare observations and model we choose to suppress all stations with missing data from the map comparison. Noteworthy, the observational coverage of the evaluation area is uneven and the scores are mainly driven by the comparison over France. At first glance, despite a slight over-estimation by models, precipitation is well predicted, especially over the mountains. The precipitation maxima are observed over the relief of the Cévennes, the Pyrénées, the Alps and Corsica. AROME-WMED better fitted the amount of rainfall over the Cévennes and the Côte d'Azur. **Another maximum is present over the north-east of Catalonia in AROME-WMED, while the maximum over the Valencia area is larger in AROME-France than in AROME-WMED.**

3.3 Other parameters

Simulated brightness temperatures lower than 240 K in the SEVIRI 10.8 μm channel (related to high level cloud discrimination) and wind gusts, with speed in excess of 40 km h^{-1} ,

were chosen to highlight the capability of both models to simulate convective systems. Figure 14 shows the BSS for the SEVERI 10.8 μm channel brightness temperature simulated from both models and compared with the satellite data at various forecast ranges and within a 54 km neighbouring distance. The slight improvement brought by AROME-WMED is statistically significant at a 90% level for ranges above 12 h, as shown by the stars on the graph. This score has also been applied to the 1 h wind gust exceeding 40 km h^{-1} (Fig. 15) for the 5–6 h, 11–12 h, 17–18 h and 23–24 h ranges in a 54 km radius. The main outcome is that AROME-WMED 12 h forecast provides better wind gust estimation which is statistically significant at a 90% level. However the evaluation network used for wind gust is smaller (about 700 stations over France) than the one used for the 24 h accumulated precipitations (about 4000 stations), that is why it is more difficult to obtain robust results for this parameter.

4 Forecast evaluation during the second Special Observation Period (SOP2)

Contrary to the first SOP which was dedicated to the heavy precipitation study, the second SOP focussed on the study of strong winds and dense water formation on the northwestern Mediterranean basin (Drobinski et al., 2014).

4.1 Temperature, relative humidity at 2 m and wind at 10 m

Statistics of temperature and relative humidity at 2 and 10 m wind speed during SOP2, with the number of observations used for the validation at each forecast range, are shown in Fig. 16. The minimum in the number of observations is found at the 00:00 UTC slot. Though the 2 m temperature RMSE are similar (around 2C) between AROME-WMED and AROME-France, they are slightly larger (5.6% on average) for AROME-WMED beyond the 24 h forecast range. The negative 2 m temperature error bias value becomes larger for AROME-WMED beyond the 12 h range. The difference in temperature biases of both models is around 0.1 $^{\circ}\text{C}$ for forecast ranges over 12 h (Fig. 7a). The pattern of the negative bias follows a diurnal cycle, which is less pronounced than during SOP1 (Fig. 7a). Its values are

however identical between the 0–24 h and 24–48 h forecast ranges. Concerning the relative humidity, the bias cycle with respect to time is stronger than the temperature bias cycle (Fig. 7b). The minimum is obtained at 6 h and the maximum at 15 h. Moreover the RMSE in relative humidity is larger for ranges from 24 to 48 h than for the day-1 range. The RMSE maximum is reached at 15:00UTC (15 and 39 h ranges). AROME-France and AROME-WMED have a quite similar behaviour, with a better fit for AROME-France, as shown by a smaller RMSE. As for other parameters, the wind error RMSE in AROME-WMED is larger, ranging from 1.8 to 2.2 m s^{-1} during SOP2 (Fig. 7c). The differences in error bias are more pronounced.

4.2 24 h accumulated precipitation

The frequency bias and the ETS for the 24 h precipitation, computed with surface data over the 0–24 h range for both models and the 24–48 h range for AROME-WMED only, are presented in Fig. 17. The AROME-WMED frequency bias shows larger positive values for SOP2 than for SOP1 (1.3 vs. 1.2). For the lowest x-axis thresholds (i. e. <20mm/24hr) corresponding to comparison samples exceeding 1000, the frequency bias is fairly similar for both SOP1 and SOP2, that is around at 1.2 - 1.3. For higher thresholds, AROME-WMED always exhibits a higher frequency bias reaching 1.8 for the 60 mm/24h. However, for threshold above 20 mm/24h, the smaller the sample (between 100 and 50 verification data), the larger the error margin, hence the larger observed differences are not significant. On the contrary, during SOP1, there were 700 verification data for the 60 mm/24h threshold, making the comparison more robust. The difference between both model biases is similar to the one found during SOP1 (Figs. 11 and 17). In contrast with results obtained during SOP1, ETS is also less good for AROME-WMED than for AROME-France. These results may be due to disparate meteorological events, to domains being different, and to a special computation of background error covariances during convective events. These covariances favour the performance of AROME-WMED during SOP1 with respect to the AROME-France ones but it is less well suited to the forecast of rainy events driven by the large scales during SOP2 as compared to the one of AROME-France.

4.3 Other parameters

The Brier skill score was also computed for the SOP2 period for both models. A few examples are shown in Figs. 18 and 19, respectively for simulated brightness temperatures and 1 h wind gust in excess of 40 km h^{-1} . As for the SOP1 period, BSSs are very similar for both models, with a significant improvement of brightness temperature simulation by AROME-WMED for the threshold of 240 K and for the 6–18 h forecast ranges. As more radiances are assimilated in AROME-WMED than in AROME-France, this leads to a better agreement between AROME-WMED and SEVIRI observations during the first forecast ranges.

In summary, both AROME-WMED and AROME-France forecast performances are equally good. AROME-France proves to be better for 2 m humidity, 10 m wind and small rain rates when AROME-WMED is better in predicting heavy precipitation, strong convection and wind gusts.

5 The case of the **Intensive Observation Period 8 (IOP 8): 28–29 September 2012**

An example of strong precipitation simulation by AROME-WMED is given with IOP8. IOP8 is a case of deep convection associated with a mesoscale convergence line (Ducrocq et al., 2014) which occurred on 28–29 September 2012 in Southern Spain. Heavy rainfall during IOP8 caused severe damages which resulted in 13 casualties in Andalusia and Murcia (southern Spain, neither of these areas is a specific HyMeX target but they are included in the AROME-WMED domain). The synoptic situation was characterized by an upper level cut-off low over southern Portugal at 00:00 UTC on 28 September (Fig. 20a), which first affected Andalusia, then progressed eastward to finally reach eastern Spain at 12:00 UTC on 29 September (Fig. 20b). **In the north easterly flank of the cut-off low, where there is upwards forcing, favouring the triggering of the convection, low level depression and convergence were created, reinforcing the convection and heavy precipitation. The low-level convergence zone shifted from inner Andalusia (at 06:00UTC on 28 September) to Catalonia at 12:00UTC on 29 September.** Most of the heavy precipitation that fell on the Murcia

region was caused by a mesoscale convective system between 10:00 and 13:00 UTC on 28 September. It was generated along a convergence line between the warm and moist easterly low-level flow in the Balearic basin and the rapid westerly low-level flow between southern Spain and North Africa, in the Aliborean basin (Fig. 21a and b). This convergence line was located ahead of the deep upper level trough (Fig. 20b). Both progressed north eastwards reaching the north of the Balearic Islands at 00:00 UTC on 29 September (Figs. 20c and 21c). Both progressed north eastwards passing over Valencia and reaching the north of the Balearic Islands at 00:00 UTC on 29 September (Figs. 19c and 20c). They finally reached Catalonia on 29 September 12:00 UTC.

The measured amount of daily precipitation exceeded 200 mm in Andalusia and in Murcia (28 September, Fig. 22a). Another precipitation maximum was observed in the Valencia area. Worthy to be mentioned, AROME-WMED was able to forecast accurately the 24 h accumulated precipitation amount (Fig. 22b) at 24 h and even at the 48 h forecast range (Fig. 22c). At the 24 h range, the Andalusia precipitation maximum is indeed underestimated, it has a very small horizontal extension, located at around 37° N and 4° W (Fig. 22b). Also to be noted that the 48 h precipitation forecast from 27 September 00:00 UTC seems to be better than the 24 h one as it isolates three precipitation maxima over Andalusia, the Murcia area and in Valencia, even though the first two are not located precisely at the observation locations. The analysis of the 1-h precipitation accumulation in the observations showed that heavy precipitation over Andalusia at 5 W 36.5 N occurred as soon as 27 September 21:00 UTC. In AROME-WMED forecast from 28 September 00:00 UTC the corresponding system is located more westerly with lower values of rainfall (up to 15 mm/1h) instead of more than 50 mm/1h in the observations. In that case, AROME-WMED model had difficulty in reproducing heavy precipitation in the early 4h forecast ranges. Noteworthy to mention that this area is close to the boundary of the model domain and contains few assimilated data. This wrong location is illustrated when comparing the SEVIRI brightness temperature observations against the simulations from the 9 h forecast starting on 28 September 2012 00:00 UTC and from the 33 h forecast starting on 27 September 2012 00:00 UTC (Fig. 23). At first sight, the low brightness temperature values over Spain, es-

pecially in the Andalusian area and the Murcia region are higher in the simulation than in the observations (Fig. 23a). They are associated with two convective systems present over these areas. The system over Murcia is associated with low values around -60°C . The brightness temperature simulation from the 9 h forecast indicates that the system is less developed and extended (Fig. 23b) over the Andalusia region. In the 33 h forecast simulation, the system over Andalusia is less intense, its spatial coverage is smaller than in the observations (Fig. 23c). The Murcia system is also shifted to the north-east and is located over the coast instead of in-land.

For 29 September 2012, the maximum values of rain accumulated between 24 and 48 h are overestimated (Fig. 24a and c) and located over the Cévennes Vivarais area and the south-west of Catalonia. In the 24 h forecast range (Fig. 24c), the maximum is lower than at day-2 forecast and more centred over Catalonia as displayed by the observations (Fig. 24a). The brightness temperature simulation for 29 September 2012 at 06:00 UTC shows that the system over Catalonia is more developed and extended over land in the 30 h simulation than in the 6 h one and in the observations (Fig. 25). This poor location of prediction at 28 September 2012 affects only 00:00 UTC forecast and concerns both models, the AROME-WMED model itself and the coupling ARPEGE model during the 24–36 h range. **The persistence and the stationary position of the convective system in the 24-30-h forecasts lead to an overestimation of precipitation over Catalonia in the first 6-h on 29 September.** For 29 September 12:00 UTC, the global ARPEGE model (not shown) and AROME-WMED (Fig. 26) forecast too much warm air over the Balearic Islands and Catalonia. This feature is associated to a low with two minima off-shore the Var region (south-east of France) and over the Balearic Islands in the 36 h forecast (not shown), which leads to overestimate precipitation on 29 September over Catalonia. The system over south-east of France is already established in 6 h forecast simulation and is ahead compared to the observations (Fig. 25a). Finally, the convective system affecting Liguria is present in the 30 and 6 h simulations, even though the associated brightness temperatures are lower than in the observations.

To summarize, AROME-WMED is able on this case to simulate more than 150 mm/24h even if the location and the temporal evolution are not perfect.

6 Concluding remarks

The AROME-WMED model was initially developed to forecast heavy precipitation events around the north-western Mediterranean basin during the HyMeX Special Observation Periods (SOPs). For the first time, a research version of AROME with its own data assimilation system and with a wider domain than the operational AROME-France one was run in real time over a long period (9 months). Moreover, some experimental data from the HyMeX campaign were assimilated in real-time. This project reached its main goal with an availability in real-time of 95 % during SOP1 and 93 % during SOP2, which helped the decision process on the observation deployment, such as dropsondes launched from aircraft or Boundary Layer Pressurized Balloons.

Once the field campaign **was** over, the quality of forecasts was compared with those of AROME-France, thanks to specific scores that used additional surface observations made available from the HyMeX database and to standard probabilistic scores used operationally at Météo-France. The quality of both models was similar over the common area, in terms of routine scores. During SOP1, AROME-WMED is slightly better than AROME-France for parameters associated to convective system prediction (i. e. surface parameters, large rain rate thresholds and brightness temperature simulation). For SOP2, AROME-France gives better results in terms of temperature, humidity, wind and precipitation. AROME-WMED is found to be better only for brightness temperature simulation and 1 h wind gust. **During winter period, AROME-France model benefited from a B-matrix, computed over different meteorological situations (including anticyclonic and stable situations), more representative of the meteorological conditions encountered during SOP2 (see paragraph 2.2.1).** One of the advantages of AROME-WMED was the availability of its 24 to 48 h forecasts. In fact, the 2 day period of simulation enabled forecasters to build a scenario for high precipitation events and to follow their development in the simulation. As a consequence, it facilitated the scheduling of the observation deployment. In point of fact, the model was able to simulate heavy precipitation above 150 mm per 24 h. An illustration was given by the IOP8 case, which was quite well predicted by AROME-WMED at the 48 h range with three isolated max-

ima of precipitation. The differences observed in the simulation between both models are due to (i) their own specific domain, AROME-WMED being well suited for the simulation of intense precipitation, (ii) to their background error covariance matrix, which was computed specifically for convective events and (iii) to a different dataset of assimilated observations.

5 The large maritime domain of AROME-WMED is able to well describe the atmospheric circulation over the Mediterranean Sea and the neighbouring mountains that constrain this circulation.

As mentioned in Sect. 2.3, the operational version of ARPEGE and AROME-France was upgraded during SOP1 on 25 September 2012 at 06:00 UTC. In order to provide a homogeneous field dataset in the HyMeX database, a first reanalysis with a unique version of the forecasting system, taking into account additional data, is already available to the scientific community. The main changes are the following: more observations have been assimilated; the cut-off time for the available observations has been removed for the data selection; the assimilation of some mobile soundings which could not be used in real-time (dissemination issue) has been performed; the final forecast range has been extended up to 54 h in order to enable comparison of the 2 day precipitation forecast to the 06:00–06:00 UTC 24 h observed precipitation accumulation. However a second reanalysis is also scheduled in order to improve the previous AROME-WMED versions. It will be produced from the current operational (2014) version of the AROME software and with additional observations from the campaign. These data will comprise several Spanish radars data provided in real-time by AEMET, wind profilers, dropsonde data, reprocessed GPS ground stations, high-resolution radiosoundings, research aircraft in-situ measurements and lidar observations. The orography of the model will also be revised with a more recent physiographic database. The definition of background error covariance matrix for the SOP1 period will be also addressed. This new AROME-WMED dataset should be available in 2015. The availability of AROME-WMED analyses and forecasts to the HyMeX scientific community will allow dedicated studies towards an improved understanding of high precipitation events over the north-western Mediterranean Sea.

10
15
20
25

Code availability

The source code of AROME-WMED, being derived from the AROME operational one, can not be obtained but the model fields are available in the HyMeX database (<http://mistrals.sedoo.fr/HyMeX/>, doi:10.6096/HYMEX.AROME_WMED.2012.02.20).

5 Appendix A: Skill scores

The scores used in this study are presented below.

A1 Frequency bias (FBIAS) and Equitable Threat Score (ETS)

Frequency bias (FBIAS) and Equitable Threat Score (ETS) formulae come from appendix of Ebert (2008) and use the results of a contingency table (Table A1). This latter is made
 10 from forecast interpolated on the observation grid, for each threshold chosen, we count hits, false alarms, misses and correct rejections (Table A1). This table allows to define several scores such as frequency bias and equitable threat score.

The frequency bias is then defined as the ratio of the forecast frequency to the observed frequency.

$$15 \text{ FBIAS} = \frac{\text{hits} + \text{false alarms}}{\text{hits} + \text{misses}} \quad (\text{A1})$$

The ETS gives the fraction of all events forecast and/or observed that were correctly diagnosed but it accounts for the hits that would occur purely due to a random chance.

$$\text{ETS} = \frac{\text{hits} - \text{hits}_{\text{random}}}{\text{hits} + \text{misses} + \text{false alarms} - \text{hits}_{\text{random}}} \quad (\text{A2})$$

where

$$20 \text{ hits}_{\text{random}} = \frac{1}{N} (\text{observed yes} \times \text{forecast yes}) \quad (\text{A3})$$

For further details, please refer to Ebert (2008).

A2 Brier Skill Score

The Brier Skill Score with neighbourhood observations is used in this study. It has been defined in Amodei and Stein (2009). It is deduced from a Brier score with neighbourhood observations BS_NO:

$$5 \quad BS_NO = \frac{1}{dd} \sum_{d=1}^{dd} \frac{1}{N(dd)} \sum_{i=1}^{N(dd)} (P(rr > t) - v(rr > t))^2 \quad (A4)$$

where $P(rr > t)$ is the probability for the parameter rr to exceed a given threshold t , computed at every verification point as the ratio of the number of points where this event occurs in the neighborhood to its total number points, $v(rr > t)$ is the observed frequency of parameter higher than the threshold t in the neighbourhood of the verification points.

10 The Brier Skill Score BSS_NO is defined by :

$$BSS_NO = 1 - \frac{BS_NO}{BS_NO_p} \quad (A5)$$

where BS_NO_p is the Brier score for a persistence forecast. In depth documentation can be found in Amodei and Stein (2009).

Appendix B: List of Acronyms

15 **AEMET:** Agencia Estatal de METeorología

ALADIN-France: Aire Limitée Adaption Dynamique et développement InterNational

AMV: Atmospheric Motion Vector

AROME-France: Application of Research to Operations at MEsocale, France

AROME-WMED: Application of Research to Operations at MEsocale, WestMEDiterranean
Sea

20

ARPEGE: Action de Recherche Petite Echelle Grande Echelle

BLPB: Boundary Layer Pressurized Balloon

BLLAST: Boundary-Layer Late Afternoon and Sunset Turbulence

BSS: Brier skill score

5 **CNES:** Centre National d'Etudes Spatiales

CNRM-GAME: Centre National de Recherches Météorologiques-Groupe d'études de l'Atmosphère
Météorologique

COPS: Convective and Orographically-induced Precipitation Study

DTS: Data Targeting System

10 **ECMWF:** European Centre for Medium-range Weather Forecasts

ECOCLIMAP:

E-GVAP: EUMETNET EIG Global navigation Satellite System water VApour Programme

EUMETNET EIG: EUMETNET Economic Interest Grouping

ETS Equitable Threat Score

15 **EUCOS:** EUMETNET Composite Observing System, former EUMETNET EIG Observation
Programme

GPS-ZTD: GPS Zenith Total Delay

GTOPO30 : Global 30 Arc-Second Elevation Data Set

GTS: Global Telecommunication System

20 **HyMeX :** HYdrological cycle in the Mediterranean EXperiment

IASI: Infrared Atmospheric Sounding Interferometer

IOP: Intensive Observation Period

MAP-D-PHASE: Mesoscale Alpine Programme-Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Event

5 **SBL:** Surface Boundary Layer

SEVIRI: Spinning Enhanced Visible and InfraRed Imager

SOP(1/2): Special Observation Period (1: Autumn 2012 / 2: Winter 2013)

SURFEX: Externalized Surface (surface scheme)

SYNOP: surface synoptic observations

10 **WRF:** Weather Research and Forecasting

Acknowledgements. The authors would like to acknowledged the MISTRALS/HyMeX programme and the funding by ANR under contract IODA-MED ANR-11-BS56-0005. The development of the AROME-WMED model was initiated by Véronique Ducrocq and Geneviève Jaubert. The setup and the validation of AROME-WMED configuration was due to the support and implication of many
15 Météo-France services. The CNRM/GMAP group is firstly acknowledged for its support in many aspects of the AROME-WMED run. J. Stein and GCO team are also thanked for helping to the setup of the AROME-WMED model. The success of the AROME-WMED run in real-time was enabled by H. Lebras, N. Merlet, J.-L Robert, M. Pottier and S. Stamatiou with the desk operator team who supervised 24 h a day the AROME-WMED model as well as their routine work on operational
20 models. José A. Guijarro and Agustí Jansà from AEMET provided us Spanish data in real-time for assimilation and observations maps, Toni Barrera for Catalonian data and Maurizio Maffei for Italian ones. The authors also thank Dr Klaus Gierens and the anonymous reviewers for their fruitful comments. Véronique Ducrocq, Jean-Francois Mahfouf and Jean Maziejewski are warmly thanked for helping to improve the manuscript.

References

- Amodei, M. and Stein, J.: Deterministic and fuzzy verification methods for a hierarchy of numerical models, *Met. Appl.*, 16, 191–203, 2009.
- 5 Bénard, P., Vivoda, J., Mašek, J., Smolliková, P., Yessad, K., Smith, C., Brožková, R., and Geleyn, J.-F.: Dynamical kernel of the Aladin-NH spectral limited area model: revised formulation and sensitivity experiments, *Q. J. Roy. Meteor. Soc.*, 136, 155–169, 2010.
- Bresson, E., Ducrocq, V., Nuissier, O., Ricard, D., and de Saint-Aubin, C.: Idealized numerical simulations of quasi-stationary convective systems over the Northwestern Mediterranean complex terrain, *Q. J. Roy. Meteor. Soc.*, 138, 1751–1763, 2012.
- 10 Brousseau, P., Berre, L., Bouttier, F., and Desroziers, G.: Background error covariances for a convective-scale data-assimilation system : AROME-France 3D-Var, *Q. J. Roy. Meteor. Soc.*, 137, 409–422, 2011.
- Brousseau, P., Berre, L., Bouttier, F., and Desroziers, G.: Flow-dependent background-error covariances for a convective-scale data assimilation system, *Q. J. Roy. Meteor. Soc.*, 138, 320–322, doi:10.1002/qj.920, 2012.
- 15 Caniaux, G., Redelsperger, J.-L., and Lafore, J.-P.: A numerical study of the stratiform region of a fast-moving squall line, Part I: General description and water and heat budgets, *J. Atmos. Sci.*, 51, 2046–2074, 1994.
- Caumont, O., Ducrocq, V., Wattrelot, E., Jaubert, G., and Pradier-Vabre, S.: 1D+3D-Var assimilation of radar reflectivity data: a proof of concept, *Tellus*, 62, 173–187, 2010.
- 20 Courtier, P., Freydl, C., Geleyn, J.-F., Rabier, F., and Rochas, M.: The Arpege project at Meteo-France. Proceedings of the ECMWF workshop on numerical methods in atmospheric models, Reading, UK, 9–13 September, pp. 193–231, 1991.
- Cuxart, J., Bougeault, P., and Redelsperger, J.-L.: A turbulence scheme allowing for mesoscale and large eddy simulations, *Q. J. Roy. Meteor. Soc.*, 126, 1–30, 2000.
- 25 Doerenbecher, A., Basdevant, C., Bernard, F., Cocquerez, P., Verdier, N., and Durand, P.: Boundary Layer Pressurized Balloons Deployed in the Mediterranean, in: Proceedings of the 21st ESA Symposium European Rocket & Balloon Programmes and Related Research, Thun Switzerland, 9–13 June 2013 (ESA SP-721, October 2013), 99–106, 2013.
- 30 Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C., Boubrahmi, N. F., Font, J., Grubisic, V., Gualdi, S., Homar, V., Ivancan-Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello, P., Llasat, M.,

- Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero, R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I., Tintore, J., Uijlenhoet, R., and Wernli, H.: HyMeX, a 10 year multidisciplinary program on the Mediterranean water cycle, *BAMS*, 95, 1063–1082, doi:10.1175/BAMS-D-12-00242.1, 2014.
- 5 Ducrocq, V., Braud, I., Davolio, S., Ferretti, R., Flamant, C., Jansa, A., Kalthoff, N., Richard, E., Taupier-Letage, I., Ayrat, P.-A., Belamari, S., Berne, A., Borga, M., Boudevillain, B., Bock, O., Boichard, J.-L., Bouin, M.-N., Bousquet, O., Bouvier, C., Chiggiato, J., Cimini, D., Corsmeier, U., Coppola, L., Cocquerez, P., Defer, E., Delanoë, J., Girolamo, P. D., Doerenbecher, A., Drobinski, P., Dufournet, Y., Fourrié, N., Gourley, J. J., Labatut, L., Lambert, D., Coz, J. L., Marzano, F. S., Molinié, G., Montani, A., Nord, G., Nuret, M., Ramage, K., Rison, B., Roussot, O., Said, F., Schwarzenboeck, A., Testor, P., Van-Baelen, J., Vincendon, B., Aran, M., and Tamayo, J.: HyMeX-SOP1, the field campaign dedicated to heavy precipitation and flash-flooding in the northwestern Mediterranean, *B. Am. Meteorol. Soc.*, 95, 1083–1100, doi:10.1175/BAMS-D-12-00075.1, 2014.
- 10 Duffourg, F. and Ducrocq, V.: Origin of the moisture feeding the Heavy Precipitating Systems over Southeastern France, *Nat. Hazards Earth Syst. Sci.*, 11, 1163–1178, doi:10.5194/nhess-11-1163-2011, 2011.
- Ebert, E.E.: Fuzzy verification of high-resolution gridded forecast: a review and proposed framework, *Meteorol. Appl.*, 15, 51–64, 2008.
- Fischer C., T. Montmerle, L. Berre, L. Auger and S.E. Stefanescu: An overview of the variational assimilation in the ALADIN/France NWP system. *Q. J. Roy. Meteor. Soc.*, 613, 3477–3492, 2006.
- 20 Guidard, V., Fourrié, N., Brousseau, P., and Rabier, F.: Impact of IASI assimilation at global and convective scales and challenges for the assimilation of cloudy scenes, *Q. J. Roy. Meteor. Soc.*, 137, 1975–1987, doi:10.1002/qj.928, 2011.
- 25 Lebeaupin-Brossier, C., Arsouze, T., Béranger, K., Bouin, M.-N., Bresson, E., Ducrocq, V., Giordani, H., Nuret, M., Rainaud, M., and Taupier-Letage, I.: Ocean mixed layer responses to intense meteorological events during HyMeX-SOP1 from a high resolution ocean simulation, *Ocean Model.*, 84, 84–103, 2014.
- Lothon, M., Lohou, F., Pino, D., Couvreux, F., Pardyjak, E. R., Reuder, J., Vilà-Guerau de Arellano, J., Durand, P., Hartogensis, O., Legain, D., Augustin, P., Gioli, B., Lenschow, D. H., Faloona, I., Yagüe, C., Alexander, D. C., Angevine, W. M., Bargain, E., Barrié, J., Bazile, E., Bezombes, Y., Blay-Carreras, E., van de Boer, A., Boichard, J. L., Bourdon, A., Butet, A., Campistron, B., de Coster, O., Cuxart, J., Dabas, A., Darbieu, C., Deboudt, K., Delbarre, H., Derrien, S., Flament, P., Four-

- mentin, M., Garai, A., Gibert, F., Graf, A., Groebner, J., Guichard, F., Jiménez, M. A., Jonassen, M., van den Kroonenberg, A., Magliulo, V., Martin, S., Martinez, D., Mastrorillo, L., Moene, A. F., Molinos, F., Moulin, E., Pietersen, H. P., Pigué, B., Pique, E., Román-Cascón, C., Rufin-Soler, C., Saïd, F., Sastre-Marugán, M., Seity, Y., Steeneveld, G. J., Toscano, P., Traullé, O., Tzanos, D., Wacker, S., Wildmann, N., and Zaldei, A.: The BLLAST field experiment: Boundary-Layer Late Afternoon and Sunset Turbulence, *Atmos. Chem. Phys.*, 14, 10931–10960, doi:10.5194/acp-14-10931-2014, 2014.
- Marshall, J. and Schott, F.: Open-ocean convection: observations, theory, and models, *Rev. Geophys.*, 37, 1–64, 1999.
- Masson, V. and Seity, Y.: Including atmospheric layers in vegetation and urban offline surface schemes, *J. Appl. Meteorol. Clim.*, 48, 1377–1397, 2009.
- Masson, V., Champeaux, J.-L., Chauvin, F., Meriguet, C., and Lacaze, R.: A global database of land surface parameters at 1 km resolution in meteorological and climate models, *J. Climate*, 16, 1261–1282, 2003.
- Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouysse, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdran, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V., and Voldoire, A.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, *Geosci. Model Dev.*, 6, 929–960, doi:10.5194/gmd-6-929-2013, 2013.
- Michel, Y., Auligné, T., and Montmerle, T.: Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables, *Mon. Weather Rev.*, 139, 2994–3015, 2011.
- Montmerle, T. and Faccani, C.: Mesoscale assimilation of radial velocities from Doppler radars in a preoperational framework, *Mon. Weather Rev.*, 137, 1939–1953, 2009.
- Pergaud, J., Masson, V., Malardel, S., and Couvreur, F.: A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction, *Bound.-Lay. Meteorol.*, 132, 83–106, 2009.
- Pinty, J.-P. and Jabouille, P.: A mixed-phased cloud parameterization for use in a mesoscale non-hydrostatic model: simulations of a squall line and of orographic precipitation, in: *Proceeding, Conference on Cloud Physics*, Everett, WA, Amer. Meteor. Soc., pp. 217–220, 1998.

- Prates, C., Richardson, D., and Sahin, C.: Final report of the PREVIEW observation Data Targeting System (DTS), ECMWF Tech Memo, p. 581, 2009.
- Rainaud, R., Lebeaupin-Brossier, C., Ducrocq, V., Giordani, H., Nuret, M., Fourrié, N., Bouin, M.-N., Taupier-Letage, I., and Legain, D.: Characterization of air-sea exchanges over the Western Mediterranean Sea during the HyMeX SOP1 using the AROME-WMED model, *Q. J. Roy. Meteor. Soc.*, online first, doi:10.1002/qj.2480, 2015.
- Ricard, D., Ducrocq, V., and Auger, L.: A Climatology of the Mesoscale Environment Associated with Heavily Precipitating Events over a Northwestern Mediterranean Area, *J. Appl. Meteor. Clim.*, 51, 468–488, 2012.
- Rotach, M., Ambrosetti, P., Ament, F., Appenzeller, C., Arpagaus, M., Bauer, H. S., Bouttier, F., Buzzi, A., Corazza, M., Davolio, S., Denhard, M., Dorninger, M., Fontannaz, L., Frick, J., Fundel, F., Germann, U., Gorgas, T., Hegg, C., Hering, A., Keil, C., Liniger, M., Marsigli, C., McTaggart-Cowan, R., Montani, A., Mylne, K., Ranzi, R., Richard, E., Rossa, A., Santos-Muoz, D., Schar, C., Seity, Y., Staudinger, M., Stoll, M., Volkert, H., Walser, A., Wang, Y., Werhahn, J., Wulfmeyer, V., and Zappa, M.: MAP-D-PHASE: real-time demonstration of weather forecast quality in the Alpine region, *B. Am. Meteorol. Soc.* 90, 1321–1336, 2009.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective scale operational model, *Mon. Weather Rev.*, 139, 976–991, 2011.
- Wattrelot, E., Caumont, O., and Mahfouf, J.-F.: Operational implementation of the 1D+3DVar assimilation method of radar reflectivity data in the AROME model, *Mon. Weather Rev.*, 142, 1852–1873, doi:10.1175/MWR-D-13-00230.1, 2014.
- Wulfmeyer, V., Behrendt, A., Bauer, H. S., Kottmeier, C., Corsmeier, U., Blyth, A., Craig, G., Schumann, U., Hagen, M., Crewell, S., DiGirolamo, P., Flamant, C., Miller, M., Montani, A., Mobbs, S., Richard, E., Rotach, M., Arpagaus, M., Russchenberg, H., Schlusser, P., König, M., Gartner, V., Steinacker, R., Dorninger, M., Turner, D., Weckwerth, T., Hense, A., and Simmer, C.: The convective and orographically-induced precipitation study (COPS): a research project for improving quantitative precipitation forecasting in low-mountain regions, *B. Am. Meteorol. Soc.*, 89, 1477–1486, 2008.

Table A1. Contingency table.

	observed yes	observed no
forecast yes	hits	false alarms
forecast no	misses	correct rejections

$N = \text{Total}$

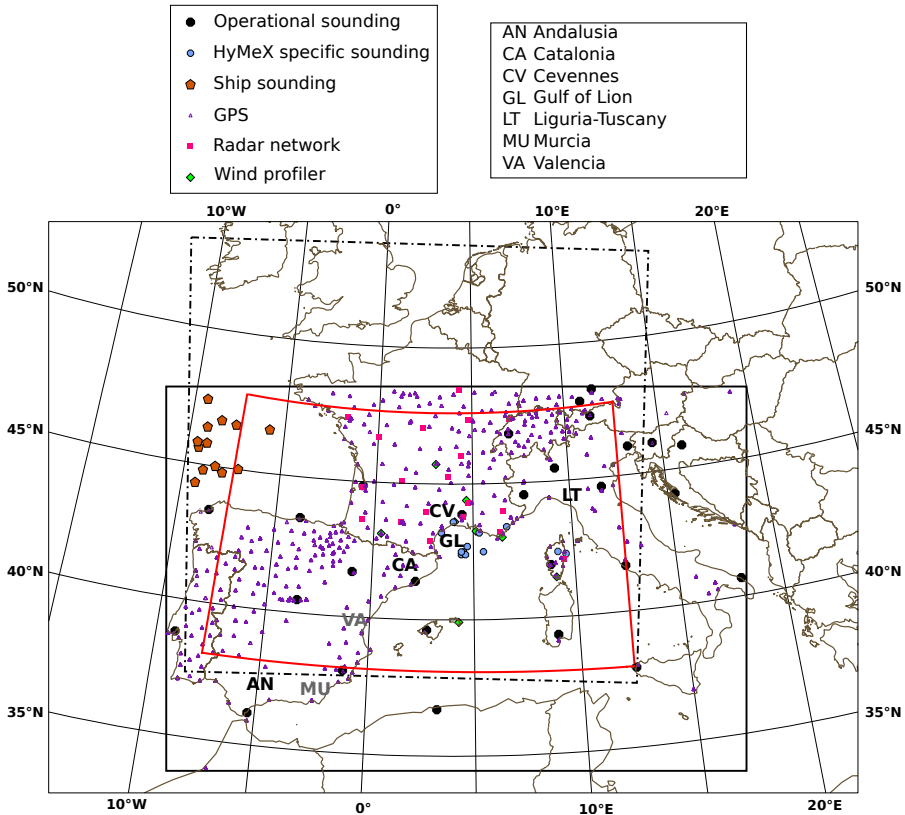


Figure 1. Location of various observation types assimilated in AROME-WMED model represented with symbols and domains of AROME-WMED (solid line box) and AROME-France (dashed dotted line box). **The red rectangle represents the verification area.**

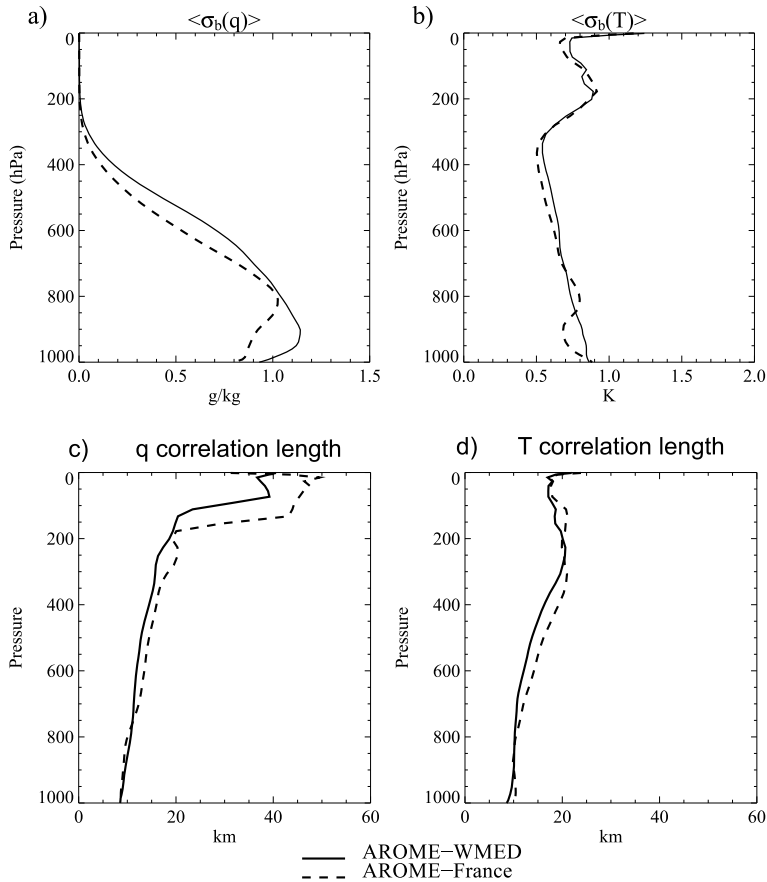


Figure 2. Standard deviation (a and b) and correlation length (c and d) of the AROME-WMED model (solid line) and of the AROME-France model (dashed line) for the humidity (left-hand panels) and the temperature (right-hand panels).

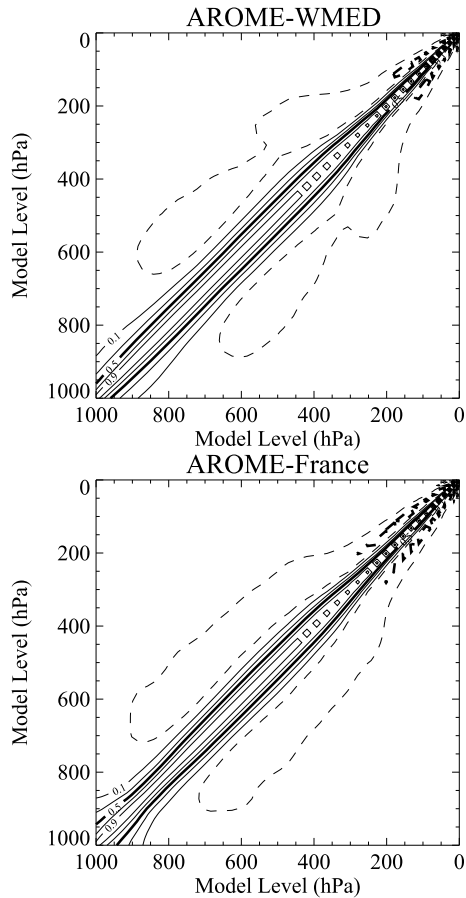


Figure 3. Vertical correlations of temperature forecast errors for AROME-WMED (top panel) and AROME-France (lower panel). Isolines plotted every 0.2-plain line: positive values, dashed line: negative values.

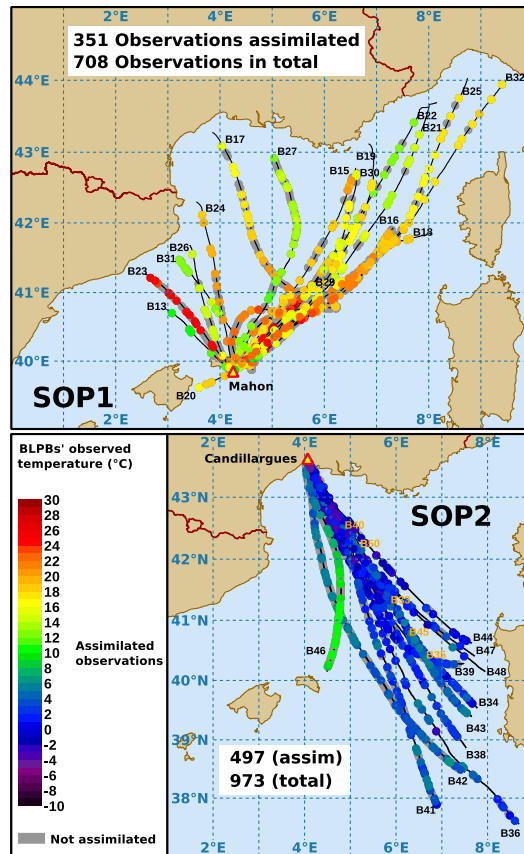


Figure 4. Trajectories of the 18 assimilated Boundary Layer Pressurized Balloons (BLPBs) during SOP1 (top) and of the 16 assimilated BLPBs launched during SOP2 (bottom).

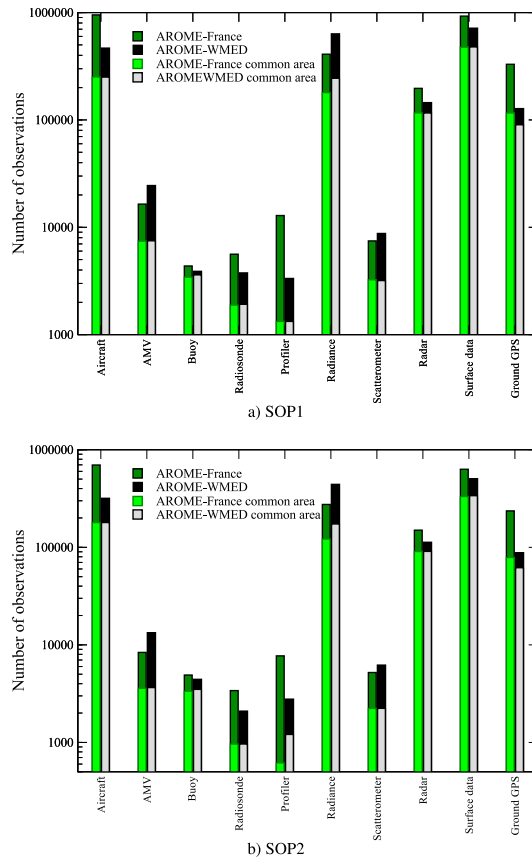


Figure 5. Total number of assimilated observations by each model **(a)** during the SOP1 period (5 September 00:00 UTC to 5 November 2012 21:00 UTC) and **(b)** for the SOP2 (1 February 00:00 UTC to 15 March 2013 21:00 UTC). y axis uses a logarithmic scale. Light-coloured bars represent the amount of data present in the common area of both models (39° N to 47.5° N and from 8° W to 11° E).

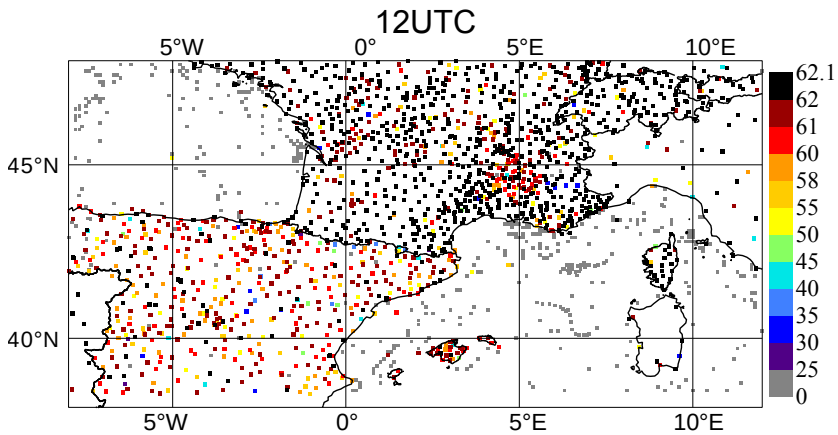
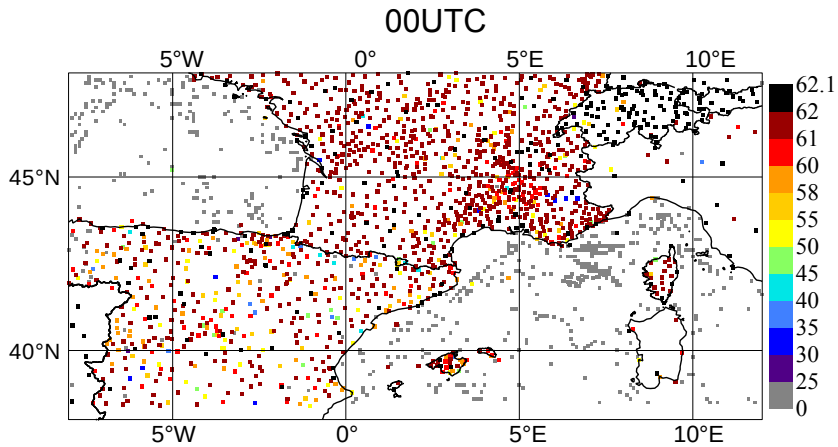


Figure 6. Number of days for which 2 m temperature data were available for the period between 5 September and 6 November 2012 at 00:00 UTC ([top](#)) and 12:00 UTC ([bottom](#)) in the HyMeX database.

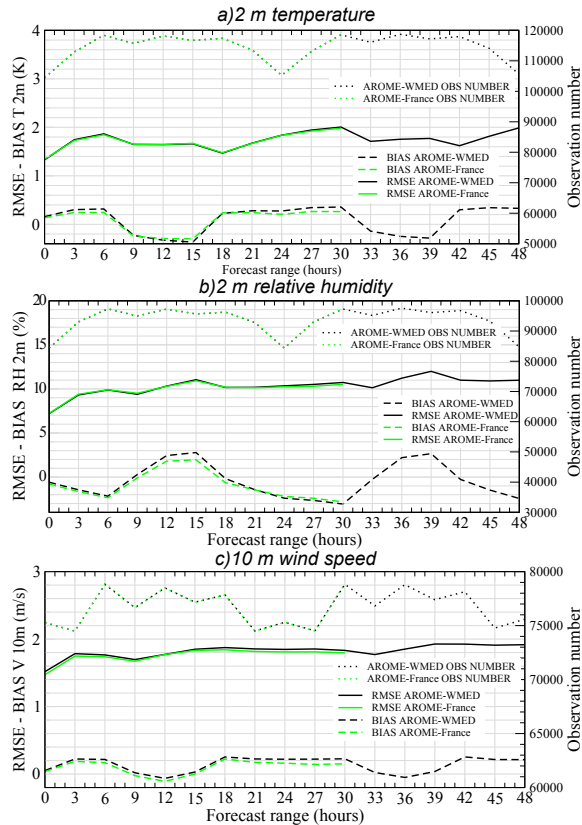


Figure 7. Bias (dashed lines) and **Root Mean Square Error** (solid lines) computed with 2 m temperature **(a)**, 2 m relative humidity **(b)** and 10 m wind speed **(c)** with respect to the forecast range for the AROME-WMED model (black) and the AROME-France model (green) for SOP1 from 5 September to 6 November 2012. Dotted lines denote the number of observations used for the comparison (right y axis).

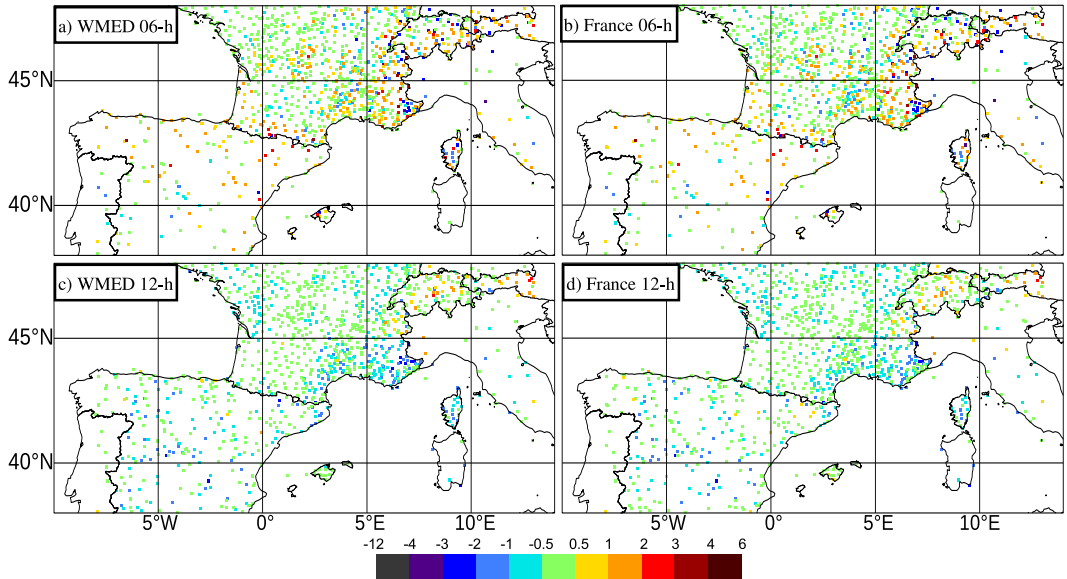


Figure 8. 2 m temperature bias (model–observation, in °C) computed with surface data for the 6 and 12 h forecast ranges for the SOP1 period ranging from 7 September to 5 November 2012, for AROME-WMED (left) and for AROME-France (right).

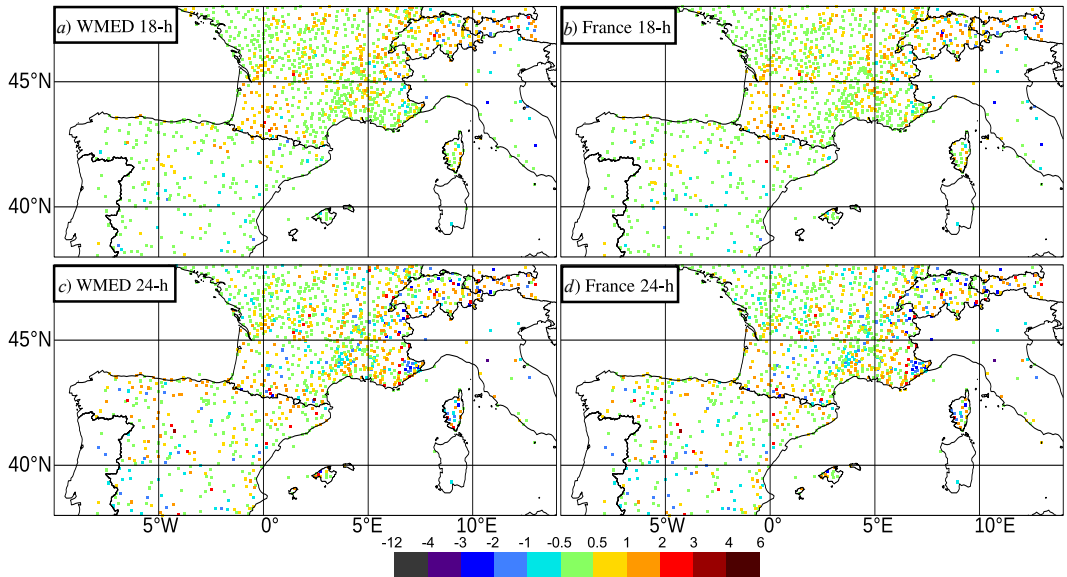


Figure 9. 2 m temperature bias (model–observation, in °C) computed with surface data for the 18 and 24 h forecast ranges for the SOP1 period ranging from 7 September to 5 November 2012, for AROME-WMED (left) and for AROME-France (right).

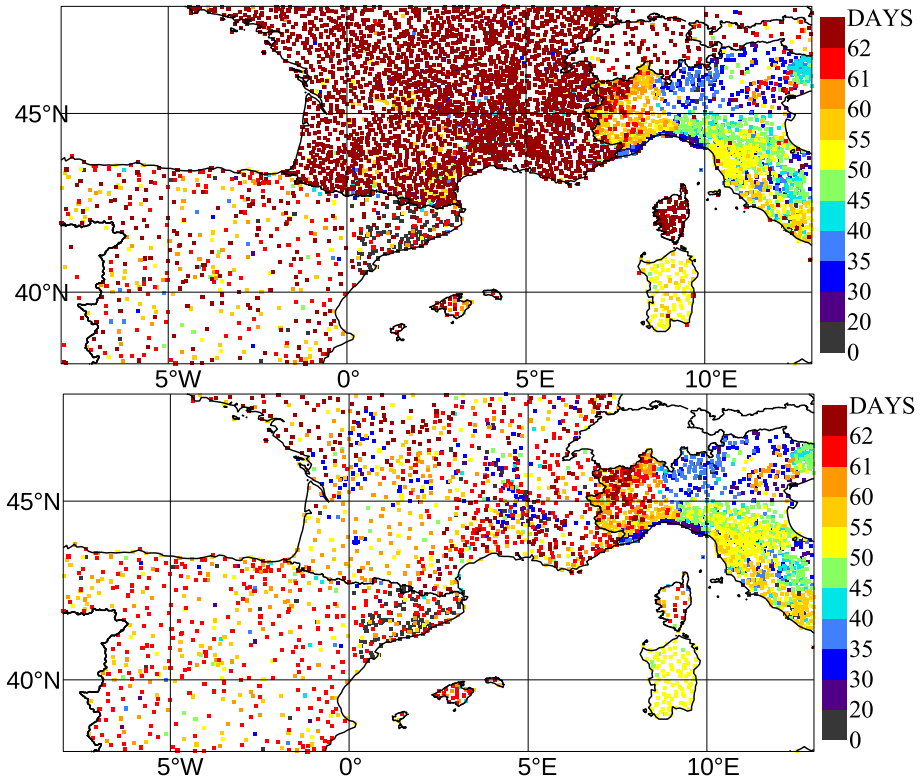


Figure 10. Number of available 24 h accumulated precipitation observations from 06:00 to 06:00 UTC next day (top) and from 00:00 to 00:00 UTC next day (bottom) for the period ranging from 5 September to 5 November 2012 (62 days) in the HyMeX database on 1 January 2014.

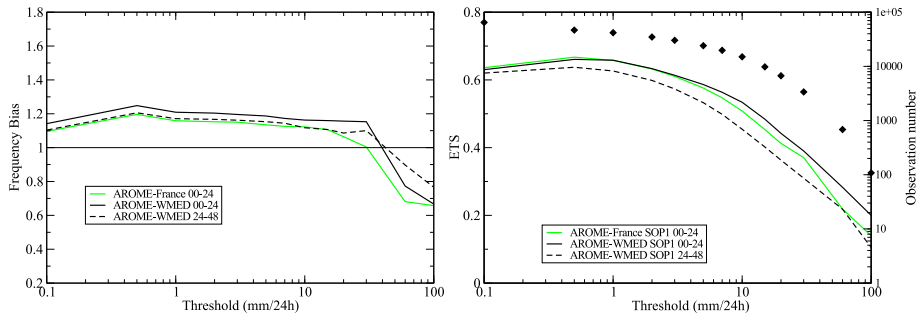


Figure 11. Frequency bias (left hand-side) and Equitable Threat Score (ETS, right hand-side) for the 0–24 h accumulated precipitation simulated by AROME-France (green line) and AROME-WMED (black line) and for the 24–48 h accumulated precipitation predicted by AROME-WMED (dashed black line) during the SOP1 period from 5 September to 6 November 2012. Logarithm scale on x axis. Diamonds represent the number of observations used for the comparison.

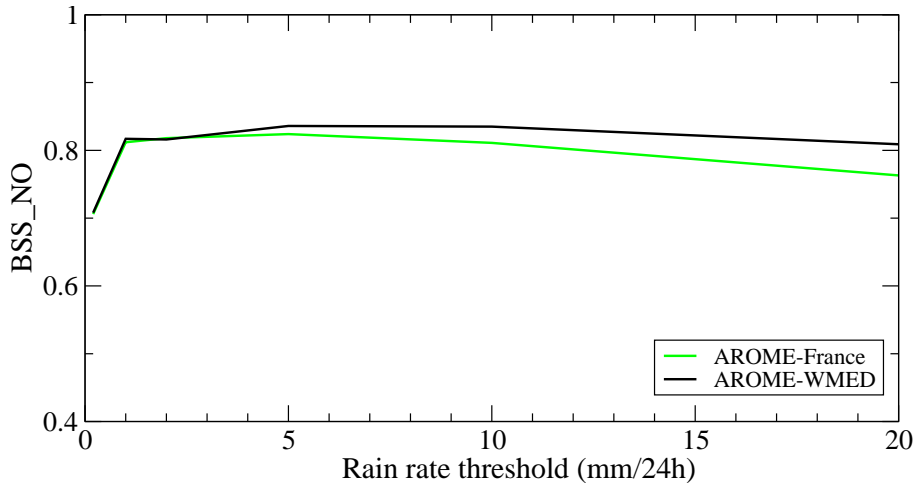


Figure 12. Brier skill score computed in a neighbouring distance of 54 km for AROME-France (green) and AROME-WMED (black) computed with raingauge data from France as a function of 6–30 h rain rate threshold during the SOP1 period from 5 September to 5 November 2012. [In this case, there is no significant difference.](#)

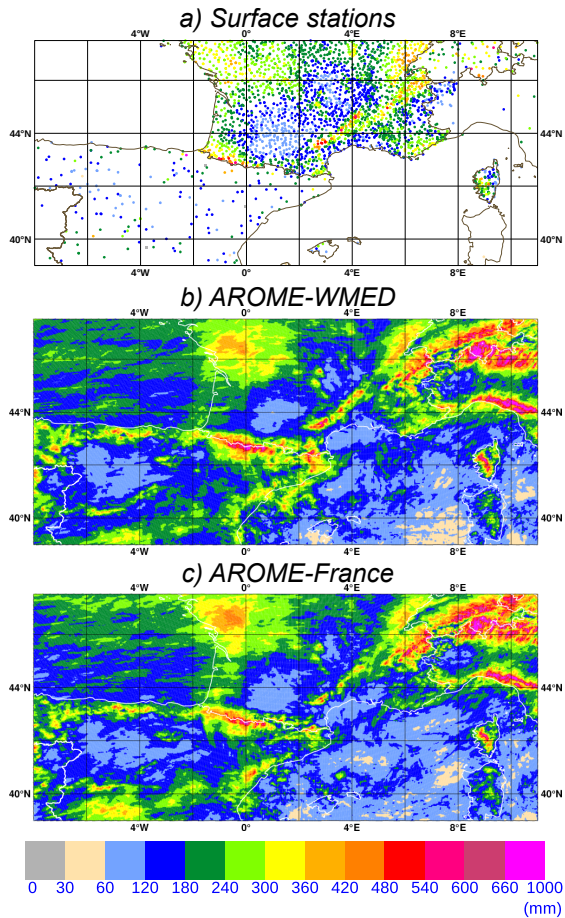


Figure 13. Accumulated precipitation (mm) during SOP1 period ranging from 5 September to 5 November 2012 measured by surface stations (Accumulation between 06:00 and 06:00 UTC next day, **(a)** predicted by AROME-WMED **(b)** and by AROME-France **(c)**).

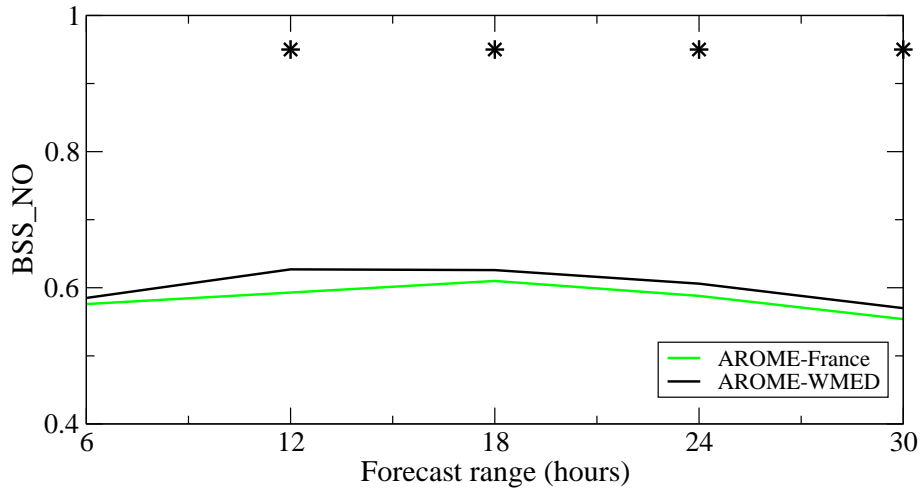


Figure 14. Brier skill score computed in a neighbouring distance of 54 km for AROME-France (green) and AROME-WMED (black) for simulated $10.8\mu\text{m}$ SEVIRI brightness temperature lower than 240 K during the SOP1 period from 5 September to 5 November 2012. Stars represent the statistical significance of the difference between the two models, which is 90% for all forecast ranges above 12 hours.

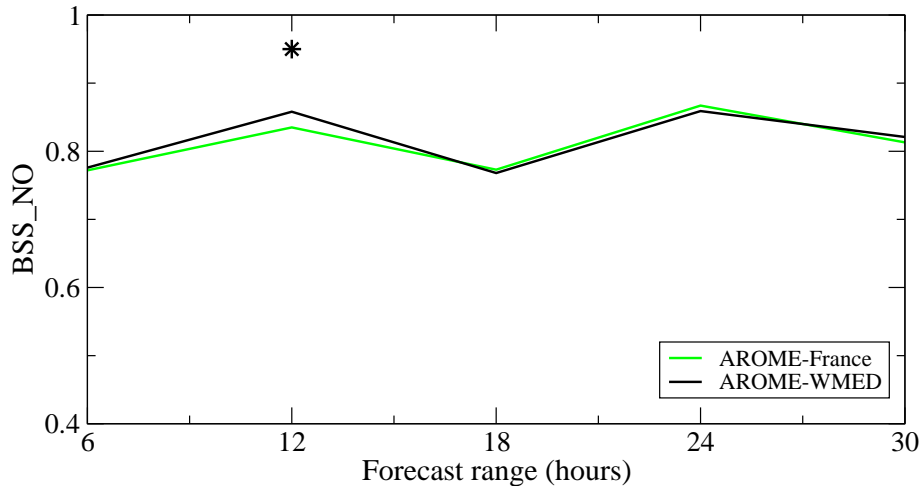


Figure 15. Brier skill score computed in a neighbouring distance of 54 km for AROME-France (green) and AROME-WMED (black) for 1 h wind gust exceeding 40 km h^{-1} during 6 h as a function of the neighbouring during the SOP1 period ranging from 5 September to 5 November 2012. Stars indicate where the statistical significance of the difference between the two models exceeds 90%.

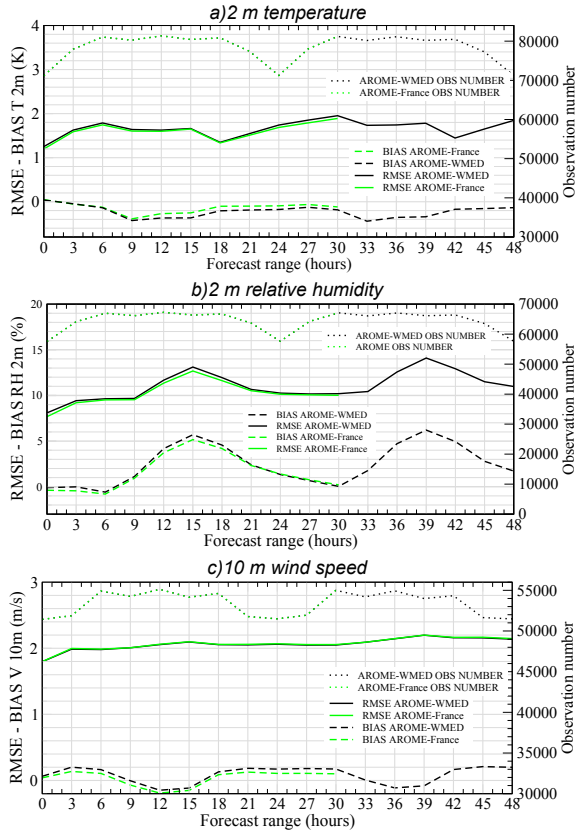


Figure 16. Bias (dashed lines) and root mean square errors (RMSE, solid lines) computed with 2 m temperature observations (a), 2 m relative humidity (b) and 10 m wind (c) with respect to the forecast range for the AROME-WMED model (black) and the AROME-France model (green) for SOP2 from 1 February to 15 March 2013. Dotted lines denote the number of observations used for the comparison (right y axis).

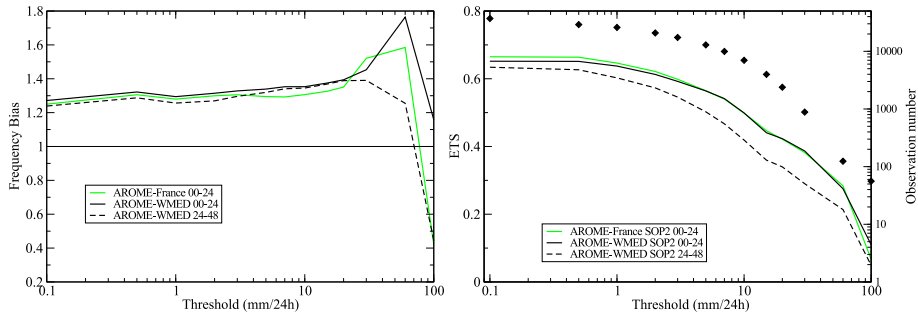


Figure 17. Frequency bias (left hand-side) and Equitable thread score (ETS, right hand-side) for the 0–24 h accumulated precipitation for AROME-France (green line) and AROME-WMED (black line) and for the 24–48 h accumulated precipitation for AROME-WMED (dashed black line) for the SOP2 period from 1 February to 15 March 2013. Logarithm scale on x axis. Diamonds represent the number of observations used for the comparison (logarithm scale).

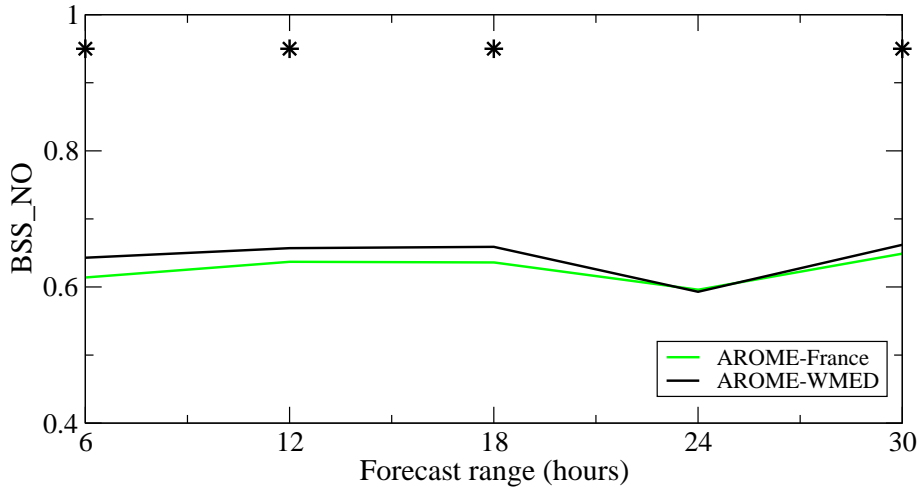


Figure 18. Brier skill score for simulated $10.8\ \mu\text{m}$ SEVIRI brightness temperature lower than 240 K as a function of forecast range for SOP2 period (1 February 2013–15 March 2013). Stars indicate where the statistical significance of the difference between the two models exceeds 90%.

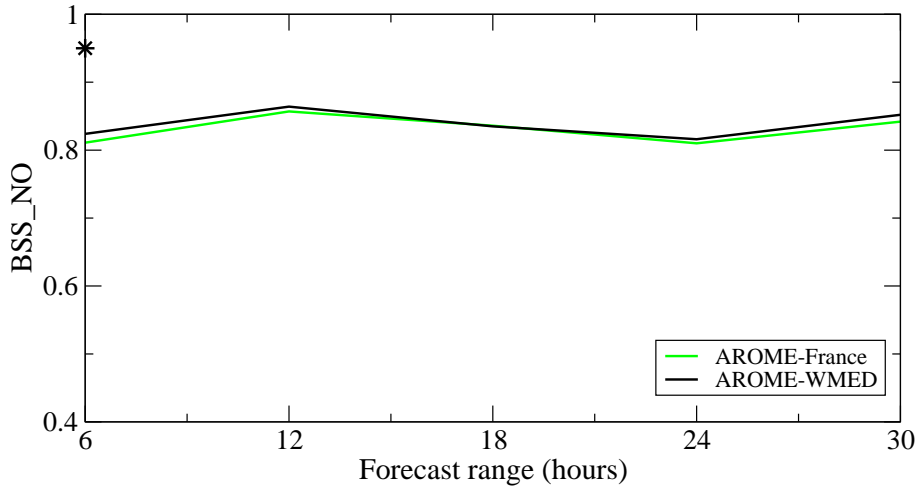


Figure 19. Brier skill score for 1 h wind gust exceeding 40 km h^{-1} during the last hour as a function of forecast range during the SOP2 (1 February 2013–15 March 2013). Stars indicate where the statistical significance of the difference between the two models exceeds 90%.

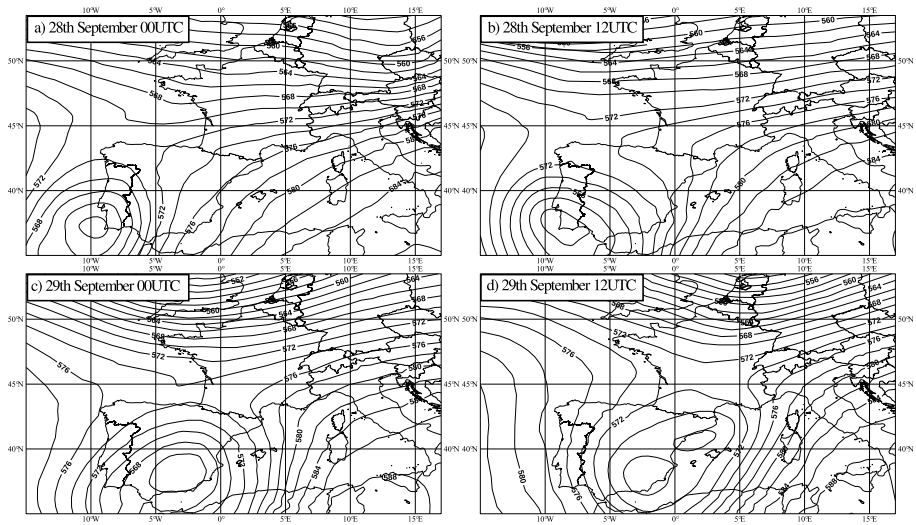


Figure 20. Geopotential height at 500 hPa (in dam) from operational ARPEGE analyses.

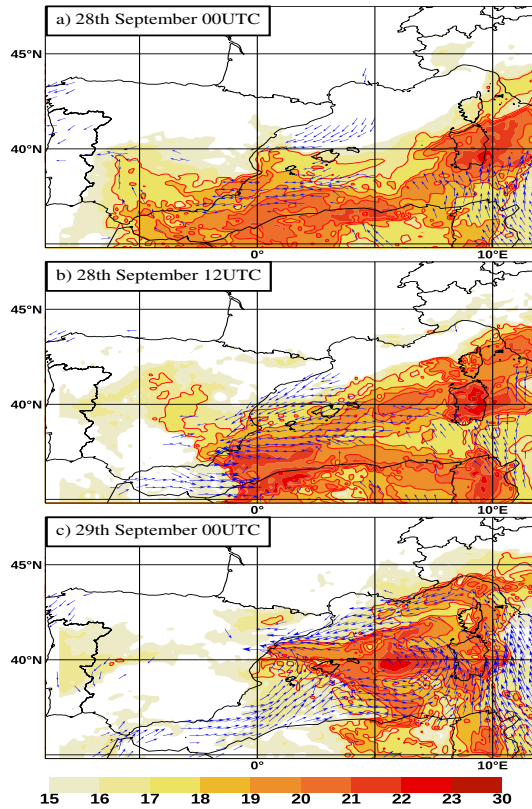


Figure 21. Wind at 925 hPa (blue arrows in excess of 20 m s^{-1}) and wet bulb temperature ($^{\circ}\text{C}$, colour scale) at 925 hPa from AROME-WMED analyses.

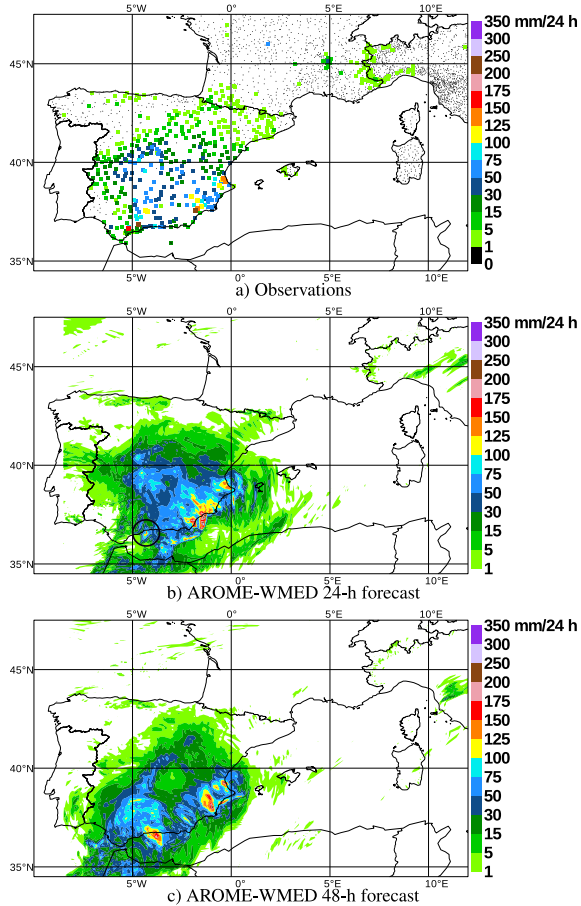


Figure 22. 24 h accumulated precipitation (mm) from 28 September 2012 at 00:00 UTC to 29 September at 00:00 UTC: observations (a), 24 h AROME-WMED forecast (b) and 48 h AROME-WMED forecast (c). Black circle indicates the Andalusia precipitation maximum in the 24 h forecast.

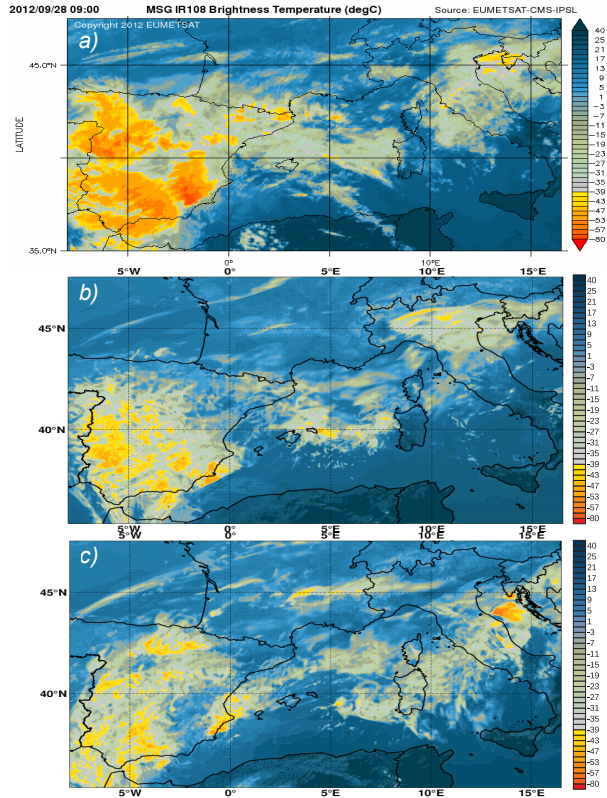


Figure 23. Brightness temperatures ($^{\circ}\text{C}$) observed with SEVIRI on 28 September 2012 09:00 UTC (a) and simulated with AROME-WMED 9 h forecast from 28 September 2012 00:00 UTC (b) and with 33 h forecast from 27 September 2012 00:00 UTC (c).

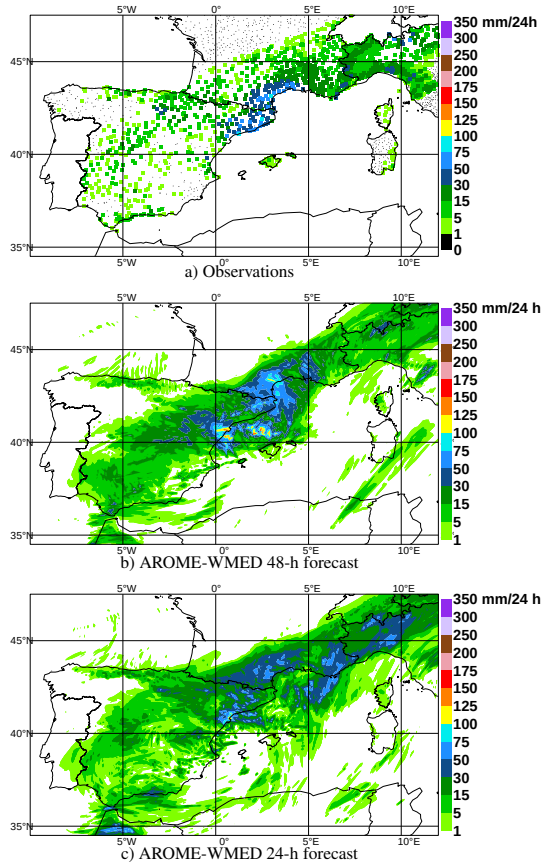


Figure 24. 24 h accumulated precipitation (mm) from 29 September 2012 at 00:00 UTC to 30 September at 00:00 UTC: observations (a), 48 h AROME-WMED forecast (b) and 24 h AROME-WMED forecast (c).

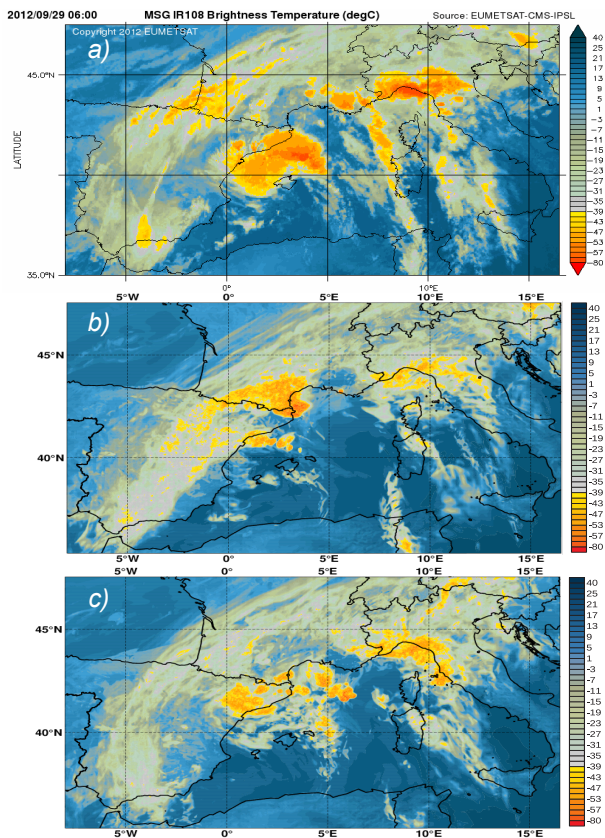


Figure 25. Brightness temperatures ($^{\circ}\text{C}$) observed with SEVIRI on 29 September 2012 at 06:00 UTC (a) and simulated with AROME-WMED 6 h forecast from 29 September 2012 00:00 UTC (b) and with 30 h forecast from 28 September 2012 00:00 UTC (c).

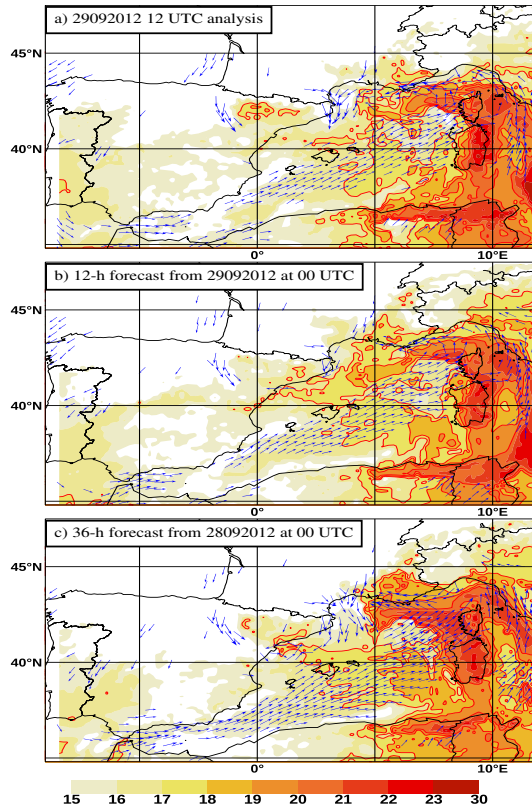


Figure 26. Wind at 925 hPa (blue arrows in excess of 20 m s^{-1}) and wet bulb temperature ($^{\circ}\text{C}$, colour scale) at 925 hPa from AROME-WMED analyses (a), 12 h (b) and 36 h (c) forecast ranges.