Nowcasting with the AROME Model: First Results from the High-Resolution AROME Airport

SUSANNA HAGELIN AND LUDOVIC AUGER

CNRM-GAME/GMAP, Météo-France, and CNRS, Toulouse, France

PASCAL BROVELLI AND OLIVIER DUPONT

DPrévi-PI, Météo-France, Toulouse, France

(Manuscript received 5 July 2013, in final form 11 March 2014)

ABSTRACT

Applications of Research to Operations at Mesoscale (AROME) Airport is the high-resolution version of the French operational numerical weather prediction model AROME. The purpose of AROME Airport is to provide rapidly updated forecasts with a 500-m resolution for nowcasting over an airport and to help with the prediction of wake vortices in order to increase the efficiency of the airport. Here, the model is evaluated for the area around Paris Charles de Gaulle Airport (CDG), France. AROME Airport and other configurations of the AROME model are compared to observations during two 4-week periods, when additional observations are available for the area around CDG (May–June 2011 and September–October 2012). The root-mean-square error and the bias are calculated for the 2-m temperature, 10-m wind speed, and the vertical distribution of the wind speed using the available measurements. The performance of AROME Airport is compared to the operational configuration of the AROME model both for synoptic hours and in a more realistic setting using the forecasts from AROME Airport starting every hour. It is shown that the forecasts from AROME Airport are an improvement with respect to the operational when comparing runs from all forecast hours, particularly for the wind speed, but when comparing the synoptic hours, the results are less clear. The sensitivity of AROME Airport to its data assimilation and initial and lateral boundary conditions is also discussed.

1. Introduction

The air traffic over Europe has increased dramatically during the last few decades and, despite a recent decrease in air traffic, the amount of traffic is very likely to continue to increase. It has become increasingly difficult to manage the higher demand for air space and to manage the traffic flow at the busiest European airports (EUROCONTROL 2012). The Single European Sky Air Traffic Management Research project (SESAR; http://www.sesarju.eu) was created to reform the air traffic management of the European skies, and one part of the SESAR project is dedicated to increasing the capacity of existing airports. This study is a part of the 12th work package of the SESAR project, subproject 12.2.2 "Runway Wake Vortex Detection, Prediction and Decision Support Tools," which aims to detect and forecast the wake vortices created by airplanes during landing and takeoff.

The capacity of an airport can be increased if a more flexible approach to the safety distance between aircraft taking off and landing is adopted. These separation times are necessary to ensure that an aircraft avoids the wake vortices created by the landing or takeoff of the previous aircraft using the runway. However, the strength and duration of the wake vortices vary not only by the size and weight of the aircraft but also by the meteorological conditions. During weather favorable to wake vortices dissipation, the separation time between consecutive aircraft can be decreased. Likewise, if the weather conditions are such that the wake vortices are more persistent than otherwise, the separation times can be increased (Gerz et al. 2005).

Historically, nowcasting has been mostly focused on severe weather conditions and early systems used the extrapolation of radar or satellite images (Browning et al. 1982; Smith et al. 1982). Later generations of

Corresponding author address: Susanna Hagelin, CNRM-GAME/GMAP, Météo-France, and CNRS, 42 Av. G. Coriolis, 31057 Toulouse, CEDEX 1, France.
E-mail: susanna.hagelin@meteo.fr

DOI: 10.1175/WAF-D-13-00083.1

nowcasting systems became increasingly reliable and new applications have been found as models for numerical weather prediction (NWP) became more complex and computing power increased while progress was being made in mesoscale data assimilation and short-term modeling (Mass 2012).

The use of nowcasting systems for airports, whether based on NWP models or observations, or both, now ranges from lightning detection (see, e.g., Brovelli et al. 2005; Li and Lau 2008); to nowcasting of winter weather, where knowing the expected icing conditions and snowfall helps with the planning of deicing of airplanes and clearing of runways (Rasmussen et al. 2001); to fog and visibility (Fabbian et al. 2007; Gultepe et al. 2006); and to more general systems (Huang et al. 2012; Isaacs et al. 2014).

The Applications of Research to Operations at Mesoscale (AROME) nowcasting system (AROME NWC; Auger et al. 2014, manuscript submitted to *Quart. J. Roy. Meteor. Soc.*) is based on the operational mesoscale weather forecasting model of Météo-France (Seity et al. 2011) and AROME Airport is developed using AROME NWC as a base.

The purpose of AROME Airport is to provide a rapidly updated high-resolution weather forecast at a dedicated location. AROME Airport is also meant to provide the input data for a wake vortex prediction model that can simulate the actual wake vortices. It is important that the data from AROME Airport are as accurate as possible in order to provide this wake vortex prediction model with the best possible startup conditions available.

In order for the wake vortex prediction model to simulate the formation, dissipation, and movements of the wake vortices, this model, in addition to the standard meteorological parameters (temperature, wind speed, pressure, etc.), also needs to be provided with the eddy dissipation rate (EDR), which is directly related to the turbulent kinetic energy (TKE).

The aim of this paper is to present the capabilities of the high-resolution configuration of the AROME model, AROME Airport, run with a grid size of 500 m on a small domain around Paris Charles de Gaulle Airport (CDG), and to verify the accuracy of the short-term forecasts from AROME Airport for wind speed and temperature as well as studying the TKE of the AROME Airport system at different resolutions. The results from AROME Airport will be compared to both observations and the forecasts, as well as analyses from the operational AROME model.

During May and early June 2011, Thales Air Systems organized a first measurement campaign, XP0, at CDG as a part of the SESAR project. During this campaign observations of wind speed were taken using several wind profilers: a sodar, a lidar, and two ultrahigh-frequency (UHF) profilers. We use these data both as a means to verify the wind speeds forecasted by the model and, in the data assimilation of the model, to see whether additional data can increase the accuracy of the forecasts and, if so, to estimate their relative impacts on the forecasts. During a later observation campaign, XP1, during September and October 2012, data from two wind speed profilers (a lidar and a UHF profiler) were available.

This paper is organized in the following way. Section 2 describes the model configuration of AROME Airport and the measurement campaigns XP0 and XP1. Section 3 discusses the TKE of the AROME model for different resolutions. Section 4 shows the results of the verification of forecasts from AROME Airport with respect to the wind speed and the temperature. Section 5 discusses the results and analyzes the sensibility of the data assimilation system of AROME Airport and the impact of the initialization data and the lateral boundary conditions. The conclusions are presented in section 6.

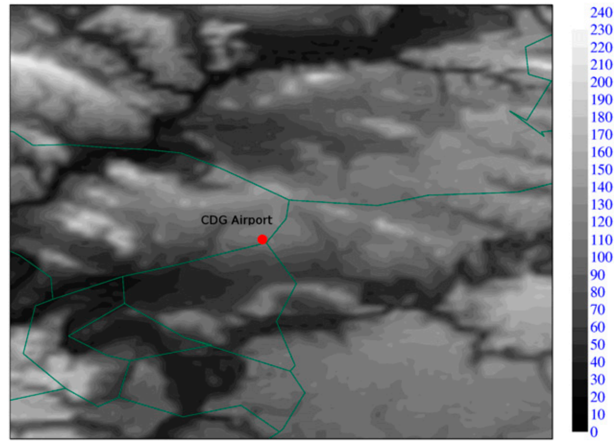

FIG. 1. The orography of the 100 km × 100 km area covered by the inner domain around CDG. The height scale on the rhs is given in m ASL. The green lines indicate the borders between departments.

2. Model configuration

The basis for AROME Airport is the operational AROME model with a resolution of 2.5 km, which is run 4 times a day (Seity et al. 2011). AROME Airport consists of two models. The first is a model with the same resolution as the operational model, which covers most of northern France, an area of 600 km × 600 km, hereafter called the Paris domain (not shown). As mentioned in the introduction, this part of the model is based on AROME NWC. The second, smaller, domain is centered around the airport (see Fig. 1) with a resolution of 500 m and covering an area of approximately 100 km × 100 km. The Paris domain has the same vertical resolution

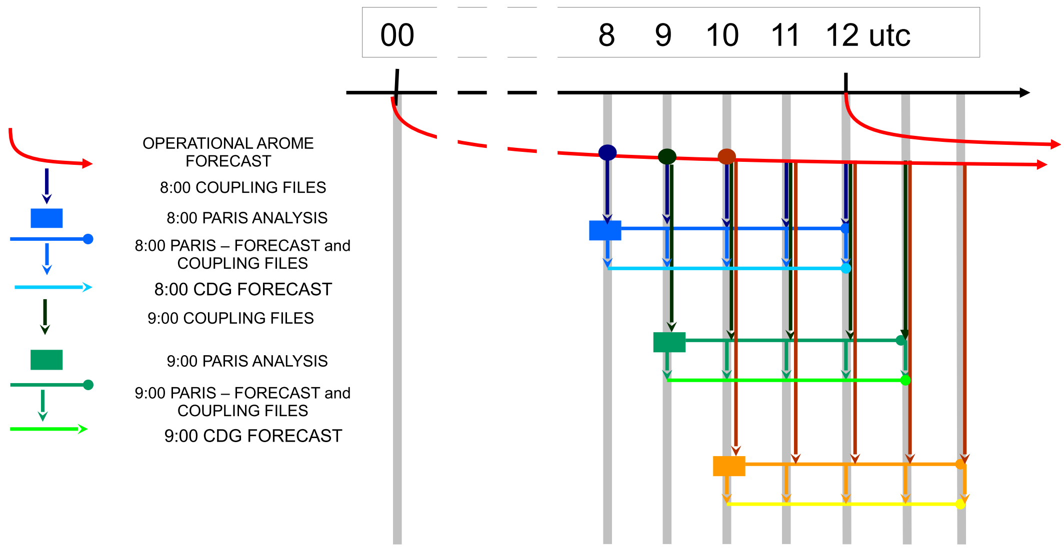

FIG. 2. Schematics of the different components of the AROME Airport system and how they interact. Vertical arrows show the couplings between the model components and horizontal arrows the forecasts.

and time step as the operational model, 60 vertical levels and a 60-s time step, whereas the high-resolution model is run with 113 vertical levels and a time step of 20 s. At times, outside of the periods presented in this paper, when there were strong winds in the upper troposphere, the time step had to be modified and a new time step of 10 s was necessary in order to avoid model crashes.

The lower boundary conditions of AROME Airport come from the Surface Externalisée (SURFEX), which uses an orography averaged at 1-km resolution and a land-use database, also at 1-km resolution.

AROME Airport starts by taking the best available forecast from the operational AROME, which is used for the initialization of the forecast on the Paris domain and to provide lateral boundary conditions. The data assimilation is also performed on the Paris domain where additional data, which might not have been available for the operational model, are assimilated. The forecasts from the Paris domain are, in their turn, used as the initialization data and boundary conditions for the high-resolution domain centered on the airport. The schematics of AROME Airport are presented in Fig. 2. With this model setup there is no cycling, so each hourly run is independent from the last one. This makes for a much simpler configuration than would otherwise have been the case. As discussed in section 5b, this also means that AROME Airport does not take advantage of its own forecast because each run is independent.

AROME Airport provides a 5-h forecast every hour. Between the first and the second forecast hour there are outputs every 5 min, whereas from forecast hour two, outputs are provided for every hour. During operational conditions AROME Airport will start with a cutoff time of 15 min, leaving sufficient time for additional observations to be used in the data assimilation, and within 45 min after the cutoff time all forecasts will be available. A complete run of AROME Airport for every hour is scheduled to be available 1 h after its initial time; for example, the forecasts from the 1300 UTC run will be available at 1400 UTC.

A number of different configurations of AROME Airport have been tested, particularly changes in the parameters for the innermost model. However, to go through them all in detail would take too much space without providing much relevant information. Also for some of the different configurations the differences in the results were indeed very small. Therefore, only the results from the best-performing version will be discussed in this paper unless something else is explicitly stated.

Measurement campaigns

As mentioned in the introduction, the performance of AROME Airport is verified for two 4-week periods where extra data from the two measurement campaigns in early summer (XP0) and autumn (XP1) were available. For each measurement campaign a continuous 4-week period is selected: 9 May–5 June 2011 (XP0) and 22 September–19 October 2012 (XP1). During both XP0 and XP1, several wind speed profilers were deployed near the runways of CDG, primarily to detect the wake vortex turbulence but they also provided measurements of the vertical profiles of the wind speed. The wind speed profiles from some of these instruments are assimilated by the AROME Airport system, but not by the operational AROME model, and are used in section 4c to assess the impact of these wind speed profiles in the data assimilation and how well the AROME model describes the vertical distribution of the wind speed. The wind speed profilers used in this study during XP0 include the following:

- one Windcube-70 lidar developed by Leosphere and Onera (the French aerospace laboratory), which measures the wind speed from 100 to 2000 m with a 50-m vertical resolution;
- one PCS 2000 sodar from Meteorologische Messtecknik GmbH (METEK), which measures the wind speed between 20 and 460 m with 10-m vertical resolution; and
- two UHF profilers, which measure the wind speed between 150 and up to 4000 m; both of the UHF profilers were continually operated in both high (150-m resolution and a top height of 4 km) and low mode (75-m vertical resolution and a top height of 3 km).

During XP1, only the lidar and one of the UHF profilers were available.

3. Assessment of the turbulent kinetic energy of the AROME Airport configuration

Because one of the purposes of AROME Airport is to provide a wake vortices prediction system with turbulence-related fields, the behavior of the model with regard to the turbulence will be assessed in this section.

TKE in fluid dynamics is the mean kinetic energy per unit mass associated with eddies in turbulent flow, which is associated with the fluctuating part of the motion. The concept of decomposing the fluid motion into a mean and fluctuating part requires the use of an averaging operator to separate each term. There exist two slightly different approaches for giving a relevant mathematical background to the partition between a mean and a fluctuating part. In the first method, the mean operator is a mathematical expectation in an ideal probabilistic space.

Applying this method to the Navier–Stokes equations leads to the Reynolds-averaged Navier–Stokes equations (RANS). These equations have been described in many reference books and articles (see, e.g., Reynolds 1895; Tennekes and Lumley 1972; Libby 1996). This first point of view is coherent with the atmospheric TKE measurements (Wyngaard and Coté 1971), where TKE is considered an intrinsic physical field independent of any numerical simulation context.

In the second approach, the governing fluid equations are obtained by applying a space and time averaging directly to the Navier–Stokes equations, with the use of a kernel function convoluted to each term of the original equations. Its length scale is equal to the grid size of the model and its time scale is equal to its time step. This approach, which is used in many numerical weather prediction models such as AROME, is generally referred to as large eddy simulation (Smagorinsky 1963; Deardorff 1970). With these equations the TKE is a prognostic variable that represents the part of the kinetic energy that is due to the subgrid-scale motion and its value decreases when the grid size of the model decreases, contrary to the TKE derived from RANS theory.

The AROME and AROME Airport models use the following one-dimensional equation for TKE:

$$\frac{de}{dt} = \alpha e^{1/2}\left(\frac{\partial U}{\partial z}\right)^2 - \beta \frac{g}{\theta_v} e^{1/2}\frac{\partial \theta}{\partial z} - \gamma e^{3/2}, \qquad (1)$$

where e is the TKE; θ_v the virtual potential temperature; θ the potential temperature; U the horizontal velocity; g gravitational constant; z is height; t is time; and α, β, and γ are three suitable coefficients. Notice that the three terms represented on the right-hand side of the equation are the shear production, the buoyancy production, and the dissipation, respectively; the turbulent transport term is taken into account in the shear production and the buoyancy term, while the pressure correlation term is neglected. This TKE equation is derived from concepts that were first used in atmospheric boundary layer large-eddy simulation (LES) models (Moeng 1984).

The TKE parameterization is important inside the atmospheric boundary layer where the highest values of TKE in the whole atmosphere are to be found. The 2.5-km horizontal resolution of the AROME model is too large to explicitly represent the motion of the eddies in the boundary layer, thus making it necessary to use a shallow convection model. The AROME and AROME Airport shallow convection model [eddy diffusivity Kain–Fritsch (EDKF)] described in Pergaud et al. (2009) provides the TKE equation with a source term, which is indeed the main contribution to the evolution of the TKE. EDKF requires, among other assumptions, that the grid size of the model not be too small.

The behavior of AROME Airport was tested with various horizontal resolutions, and the evolution of the vertical TKE profile was assessed. Figure 3 presents vertical profiles of kinetic energy (a resolved part and a turbulent part; i.e., TKE) for two different horizontal resolution configurations (dx = 500 and 2500 m). It is averaged over a period of 15 days, from 1 to 15 August 2012. The values are issued from 5-h forecasts from the 1200 UTC runs so that free convective boundary layer conditions are sampled. The resolved and the turbulent kinetic energy are plotted for both resolutions.

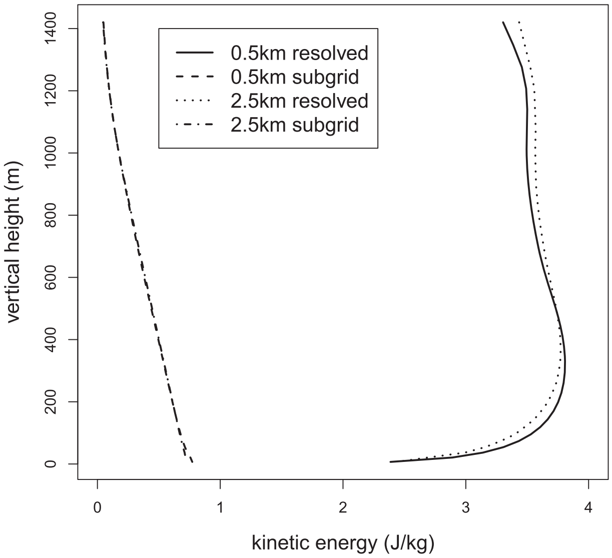

FIG. 3. Profiles of the resolved and subgrid kinetic energy as a function of vertical height for two different horizontal resolutions: 0.5 and 2.5 km.

Resolved and subgrid kinetic energy results present small differences between the two resolutions. The vertical profile of resolved energy increases greatly from zero at the ground up to a maximum of 3.8 J kg^{-1} at approximately 300 m, then it slightly decreases. A second local maximum is found at the top of the boundary layer. These resolved profiles have values that are in agreement with the theory. The TKE profiles are also the same between the two resolutions. Their values linearly decrease with height, approaching zero at the top of the boundary layer. To sum up, the AROME models behave as if the resolved and subgrid parts of the boundary layer wind are the same at these two resolutions, whereas the averaging box size that discriminates between subgrid and resolved motion is different. (It is supposed here that the difference in the orographic resolution has a weak impact on the total kinetic energy.)

According to spectrum studies of the convective atmospheric boundary layer, the classical length scale for horizontal motion ranges from tens of meters inside the surface layer up to 1 km when convection is fully developed (Stull 1988). Consequently, with a horizontal 500-m resolution, a mesoscale model is expected to partly resolve these motions, which lies in contradiction with the previous results. However, as shown in Skamarock (2004) and Ricard et al. (2013), the dynamical cores used for numerical weather models do not resolve motions at the horizontal resolution, but rather at $5 \times \Delta x$. Consequently, these results are coherent with the diagnostics carried out in Fig. 3 for both horizontal resolutions (dx = 500 and 2500 m), and the effective horizontal resolution remains much larger than the largest eddies of the boundary layer. This allows the EDKF parameterization to be used, without violating the minimum grid-size assumption.

The sensitivity of the TKE scheme [Eq. (1)] has also been tested with different settings of EDKF: the buoyant and shear production were modified by changing the value of the multiplicative coefficient according to values identified in the literature. Although with new coefficients for the dx = 500-m configuration the TKE values decreased a bit and hence were more in agreement with what is expected (not shown), these new settings led to other problems, such as worse wind scores, and the same configuration was kept for dx = 500 m as for the dx = 2500-m resolution for further model runs.

4. Verification of AROME Airport for wind speed and temperature

In this section the forecasts from AROME Airport are validated using observations at screen level as well as using the available wind speed profiles from the two measurement campaigns, XP0 and XP1, and radiosonde data. To have a first indication of the capabilities of the AROME Airport system with respect to the operational AROME model, initially the comparisons are made for the forecasts starting at the synoptic hours in order to compare the forecasts for the same hours.

In section 4e, AROME Airport is evaluated during more realistic conditions employing the forecasts for all hours of the day, and also looking at the differences between using the most recent forecasts from the operational AROME model and using the forecasts that realistically are available at the same time as the forecasts from AROME Airport. Depending on the hour of the day, the realistically available forecasts from the operational model are from somewhere between 2 and 9 h before whereas in the idealized case the best available forecast is never older than 5 h.

Another effect to take into consideration when looking at these results is that the two periods studied are characterized by different weather conditions. During the first campaign (XP0 in early summer), the weather was warmer and with a more marked diurnal variation in the temperature than during the autumn campaign XP1. The average wind speed was more similar during both campaigns, but the day-to-day variations were larger during XP1.

a. 10-m wind speed

The data from all available weather stations within the CDG domain are used to validate the model data. Each

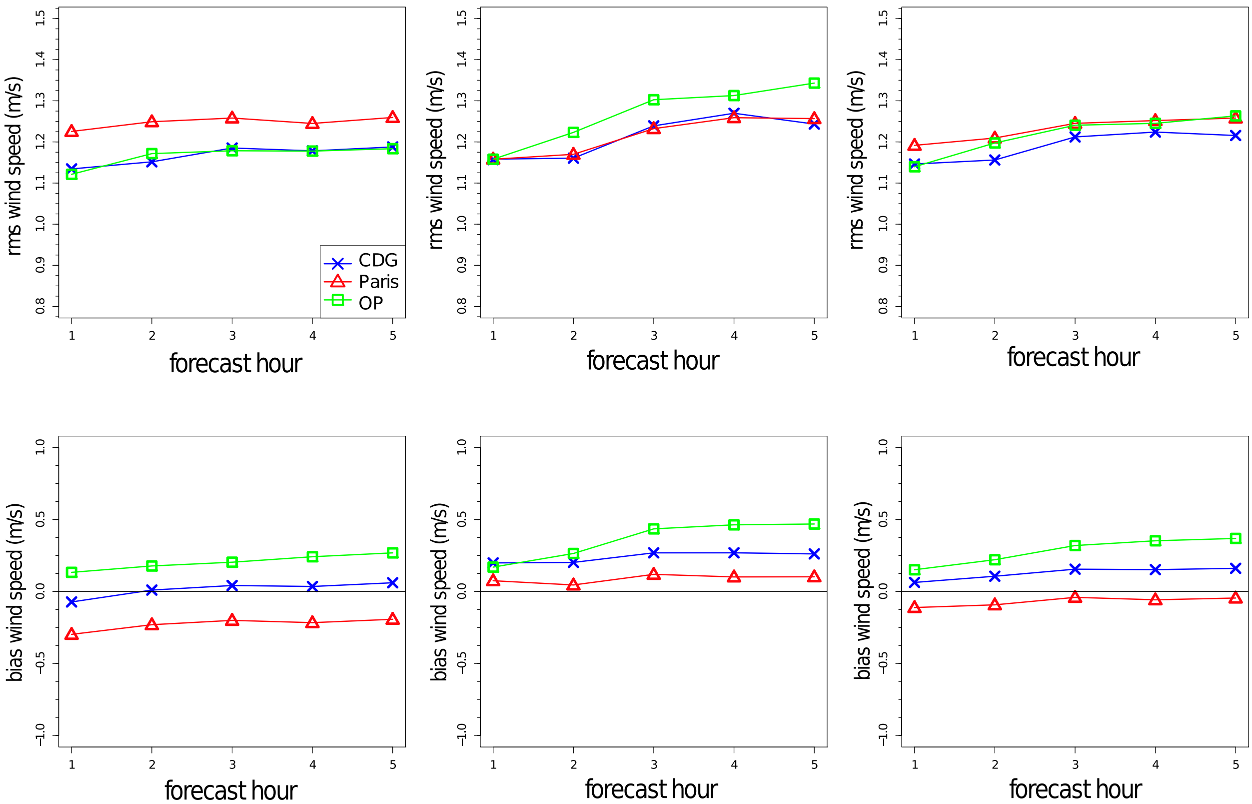

FIG. 4. (top) RMSE and (bottom) bias calculated from forecasts starting at the synoptic hours for the 10-m wind speed during 4 weeks in (left) May–June and (center) September–October 2011, and (right) for both periods combined for the simulations by the operational AROME system (green squares), and for AROME Airport for the Paris domain (red triangles) and the high-resolution CDG domain (blue crosses).

hour this dataset is compared to the output of the model from the grid point closest to each station. Both the data from the Paris and the CDG domains of AROME Airport are compared to the output from the operational AROME runs, at the corresponding forecast hours. To compare AROME Airport with the operational AROME at the same forecast hours, at first this comparison is made only at the synoptic hours.

Figure 4 shows the root-mean-square error (RMSE) and the bias for the 10-m wind speed for two 4-week periods during May–June 2011 (left) and September–October 2012 (center), as well as for both of the periods combined (right). During the summer period it is interesting to note that while the forecast over the Paris domain is systematically worse than the operational AROME simulation, the high-resolution forecasts over the CDG domain have RMSE values that are indistinguishable from the operational model despite the fact that the initialization and boundary conditions for the CDG domain come from the Paris domain. During the autumn, both the Paris and CDG domains perform better than the operational model, except for the first forecast hour. Calculating the RMSEs for both 4-week periods together shows that all three models have a comparable performance. For the first forecast hour there is a slightly lower RMSE for the operational model than for the others, but for the following forecast hours the CDG domain shows the lowest RMSE values.

The bias is positive for the operational model of the same magnitude for both of the periods. The Paris domain has a negative bias during the summer period and a much smaller positive bias during the autumn period. The high-resolution CDG domain has almost no bias during the summer period, whereas during the autumn there is a small positive bias. The net effect is that the smallest bias is found over the Paris domain whereas the overall largest bias is still found in the operational model. Almost all of the differences found in the bias between the different configurations are statistically significant whereas none of the differences in the RMSE values are. For further details please, see the appendix (Table A1).

b. Wind direction

The differences in the RMSE and the bias of the wind direction between AROME Airport and the operational

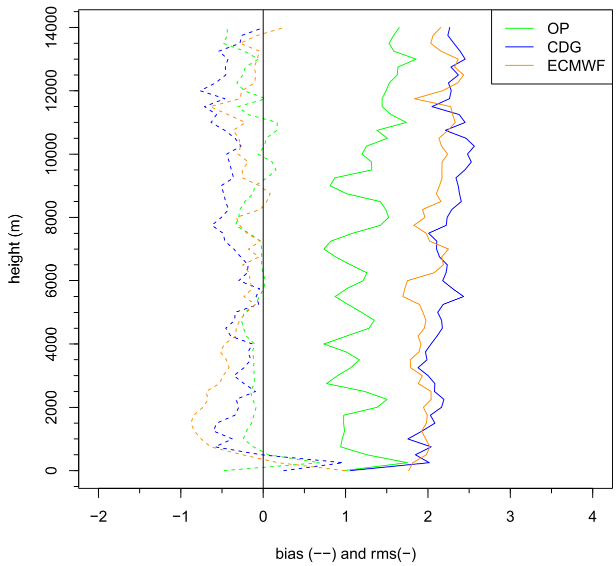

FIG. 5. RMSE (solid lines) and bias (dashed lines) of the analyses from ECMWF (orange), the analyses of the operational AROME model (green), and the 1-h forecasts from AROME Airport over the CDG domain (blue) compared to the radiosonde data from the Trappes station.

AROME are small and therefore these figures are not shown here. During XP0 the RMSE is almost constant for all models and all forecast hours, at around 25°, and during XP1 it is 22°–23°. The bias also changes very little with the forecast hour; during XP0 it does not exceed 3° for any model, whereas during XP1 the maximum bias is 6° for the operational AROME, at the 4-h forecast. The maximum bias for AROME Airport (both domains), meanwhile, is found for the fifth forecast hour and is slightly above 5°.

c. *Wind speed profiles*

Vertical profiles of the wind speed are available at 0000 and 1200 UTC from the radiosondes at the Trappes, France, station, situated near the western border of the CDG domain. To have an independent reference for the capability of AROME to forecast the vertical profile of the wind speed, the wind speed data from the Trappes radiosondes are also compared with the analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF). It is worth mentioning that these radiosonde data are assimilated by both models.

Because the radiosonde profiles are only available twice daily, the comparison with the model data is necessarily performed for a smaller statistical sample. Nonetheless, this comparison still provides a useful insight into the performance of AROME Airport in the free atmosphere. Figure 5 shows the vertical profiles of the bias and the RMSE of three models: the analyses from the ECMWF (using 0.125° resolution), the analyses from the operational AROME, and the forecasts from the high-resolution AROME Airport for XP0 and XP1. The profiles from AROME Airport are calculated from the 1-h forecasts valid for the same time because no analyses are saved from the 500-m resolution. The analyses from the coarser Paris domain are saved. However, because the RMSE and bias from the Paris domain are very similar to the CDG domain, these results are not shown here in order to make Fig. 5 clearer. It is obvious from these figures that the best scores come from the analyses of the operational AROME model. However, it is encouraging that the 1-h forecasts from AROME Airport show scores that are comparable to the analyses of the ECMWF.

Looking at the wind speed profile using the data from the wind profilers obtained during XP0 and XP1, it is possible to get a more detailed look at the performance of the AROME model nearer to the ground. The different profilers during the measurement campaigns have different resolutions and maximum measurement heights, as described in section 2. An instrument-by-instrument comparison shows little difference in the RMSE and the bias for the different models. To get a more concise view of the performance of the model with respect to the profiler data, the RMSE and the bias for each of the profilers are combined. The result of this analysis is found in Fig. 6, which shows the results for the CDG domain (crosses), the Paris domain (triangles), and the operational AROME model (squares) for the same forecast hours starting from the following synoptic hours: 0000, 0600, 1200, and 1800 UTC. The difference between the three versions of AROME is small, but the bias is slightly better for AROME Airport on the small domain. The bias of the Paris domain is practically identical to the bias for the operational model, particularly for the first 3 h. The differences in the RMSE values are also small, where the AROME CDG has slightly smaller RMSEs for the first forecast hour and to a lesser extent for the second and third hours whereas the operational model is somewhat better for forecast hours four and five.

Also taking advantage of the large amount of simulations that are produced daily from the AROME Airport configurations, the same analysis as above is repeated but using all available forecasts for the 24-h runs from the CDG and Paris domains (not shown). Looking at these data, there is a small improvement in the AROME simulation at 500 m (CDG) with respect to the Paris domain in the first kilometer above the ground. Higher up, there is still a small difference (albeit very small) in favor of the small-scale model where the RMSE and average difference between the model and

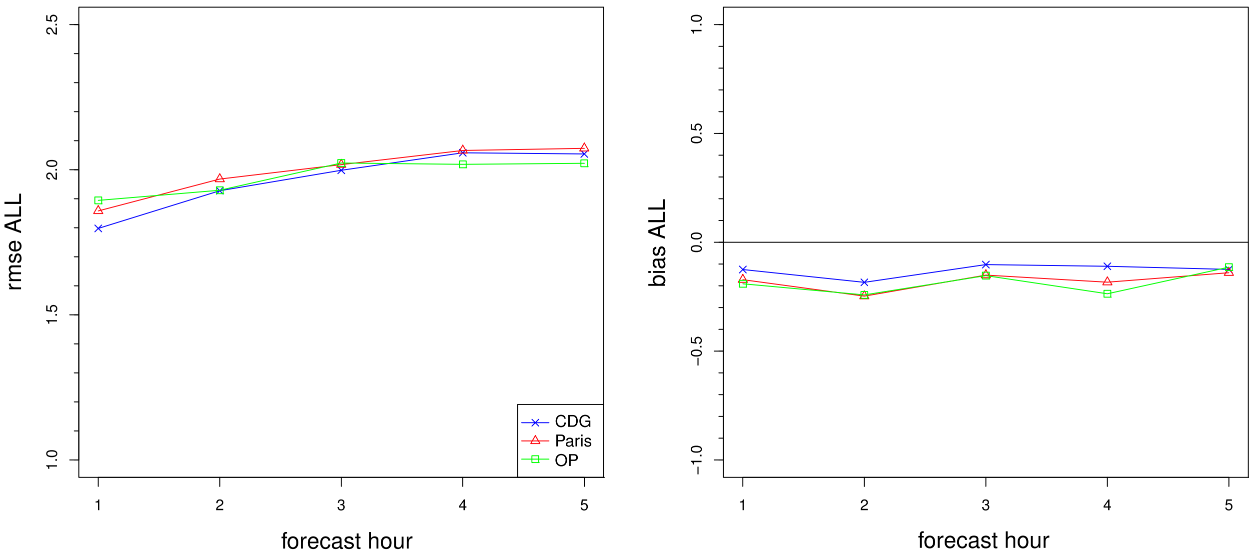

FIG. 6. (left) RMSE and (right) bias of the wind speed profiles for both XP0 and XP1 for three versions of the AROME model: the operational AROME (green lines), the Paris domain from AROME Airport (red lines), and the high-resolution CDG domain from AROME Airport (blue lines).

the observations are slightly smaller for the forecasts from the CDG domain.

d. 2-m temperature

The temperature at 2 m is compared with the forecasts in the same way as the 10-m wind speed. This comparison is also at first made for the forecasts originating from all four synoptic hours. The RMSE and bias for these forecasts for the AROME Airport and the operational AROME from the simulations of the 4 weeks during XP0 (left) and the similarly long XP1 (center), as well as for both periods combined (right), are shown in Fig. 7. Looking at the RMSE in Fig. 7 (top) it is obvious that the operational model is predicting the 2-m temperature better than AROME Airport for both the Paris domain and the high-resolution CDG domain, regardless of the period studied.

The bias behaves differently during the two periods. During XP0, clearly the best model is the high-resolution CDG domain, which starts with an almost neutral bias for the first 2 h and thereafter gets slightly higher. With the exception of the first forecast hours for AROME Airport, all three models show a positive bias. However, during XP1 all models show a negative bias with respect to the observations and the worst model is the CDG domain. Adding up all the data gives a very small bias and the overall best model is the operational AROME. A summary of the forecast hours when the differences found in the bias and the RMSE are statistically significant is shown in the appendix (Table A2).

The reason for the apparently better bias in the 2-m temperature during XP0 for the high-resolution domain is obvious if the hour-by-hour bias is studied instead. Figure 8 (top) shows the hourly bias during XP0 (May–June). The forecasts from AROME Airport (triangles, Paris; crosses, CDG) are 1-h forecasts, whereas the data from the operational model (squares) correspond to the most recent forecast for the corresponding hour, or the analysis (marked in red). The bias from the operational model is close to zero during most of the day, but as the afternoon progresses the bias becomes more and more positive, with the exception of the 1800 UTC analysis, and the net result is thus a positive bias. The biases for the Paris and CDG domains are similar, even though the absolute value of the bias from the Paris forecasts is slightly higher than for the CDG forecasts. In the early morning the bias decreases by almost 1°C between 0400 and 0600 UTC.[1] Afterward, the bias slowly approaches zero and in the evening the bias is positive again. The net effect of a strong negative bias in the morning and an equally strong positive bias in the evening is close to zero. During XP1 (Fig. 8, bottom) the hourly bias does not show the same drastic decrease during the morning hours and the sharp increase in the evening hours is also absent.

It is important to note that these comparisons, and also those in section 4a, are not showing the full potential of

[1] The values are 0.89°C for the Paris domain and 0.84°C for the CDG domain.

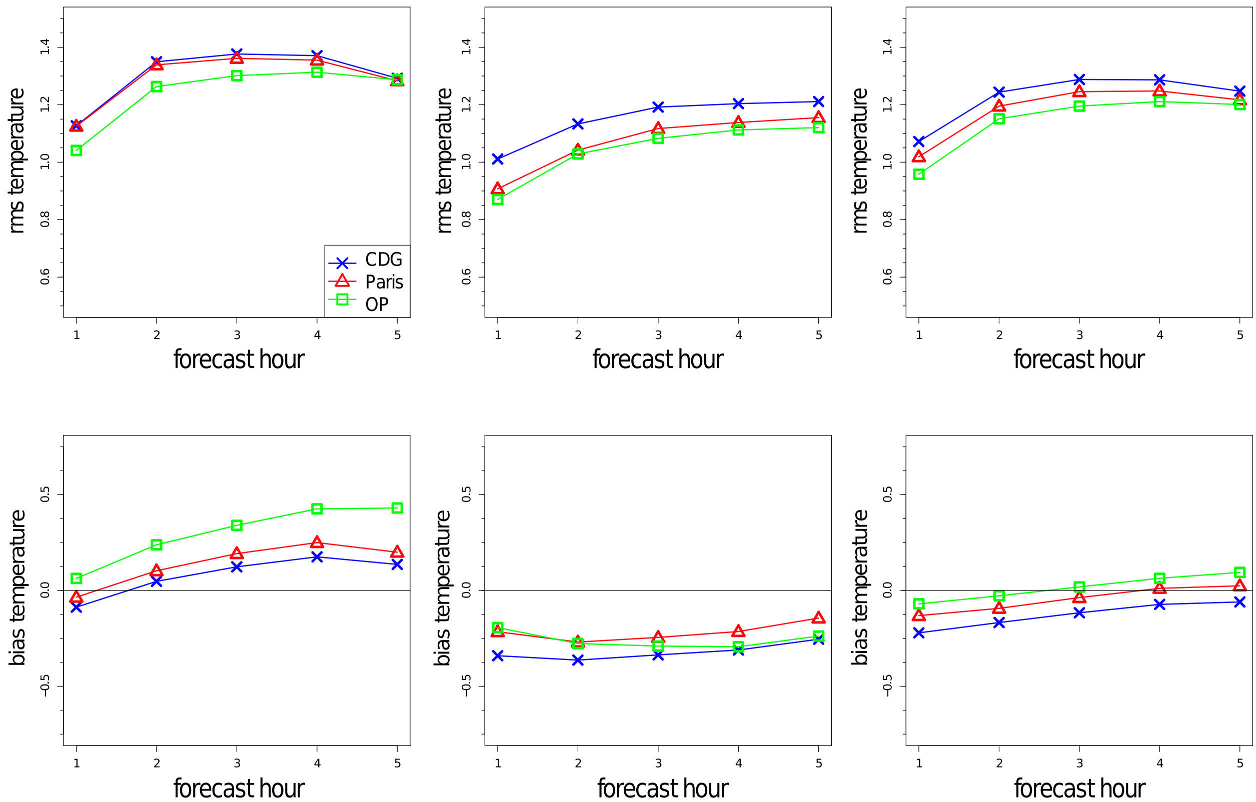

FIG. 7. (top) RMSE and (bottom) bias calculated from forecasts starting at the synoptic hour for the 2-m temperature during the 4 weeks in (left) May–June and (center) September–October 2012, and (right) for both periods combined for the simulations by the operational AROME (green squares), and for AROME Airport for the Paris domain (red triangles) and for the high-resolution CDG domain (blue crosses).

the AROME Airport system because the idea behind AROME Airport is to have a rapid forecast available at all hours of the day and the operational AROME is only reinitialized every 6 h. In the next section, the scores for AROME Airport are presented using the forecast from all hours of the day and night in order to give a more detailed view of the system's capabilities.

e. Validation of AROME Airport in an operational setting

The least expensive way to forecast on a very short time scale is to use the method of forecast by persistence, that is, presuming that the weather will not change from what has already been observed. For very short time scales, such as 1–2 h, this can be an accurate method. Figure 9 shows the RMSEs for the 2-m temperature and the 10-m wind speed for XP0 and XP1 calculated for all hours, where the worsening performance of the operational model here is represented by the green squares. Solid lines represent the most recent forecast, though in a real-time system for at least half of the time these forecasts would not be available yet. Dashed lines represent the most recent forecast that is actually available in real time at the time considered.

The performance of the forecast made by the persistence technique (circles) varies with the period and parameter studied. Looking at the 2-m temperature (Fig. 9, top) during the May–June period (XP0) the forecast from the persistence method clearly provides the worst forecasts whereas the best performance comes from AROME Airport. Both AROME Airport domains have almost equal scores for all forecast hours; however, the Paris domain is marginally better. During September–October (XP1), for the first forecast hour, the forecast by the persistence method provides a forecast that is better than either the CDG domain or the operational model but worse than the forecast from the 2.5-km Paris domain, which is the best model with regard to the RMSE values during this period. Looking at the RMSE just 1 h later, the forecast by the persistence method is worse than that of the worst-performing model (the operational AROME in a real-time setting) and afterward the forecast made by the persistence method continues to deteriorate. Overall, the best model for the 2-m temperature is the Paris domain, the worst-performing numerical model is the operational AROME model using the realistically available forecasts, and all model configurations outperform the forecast made by the

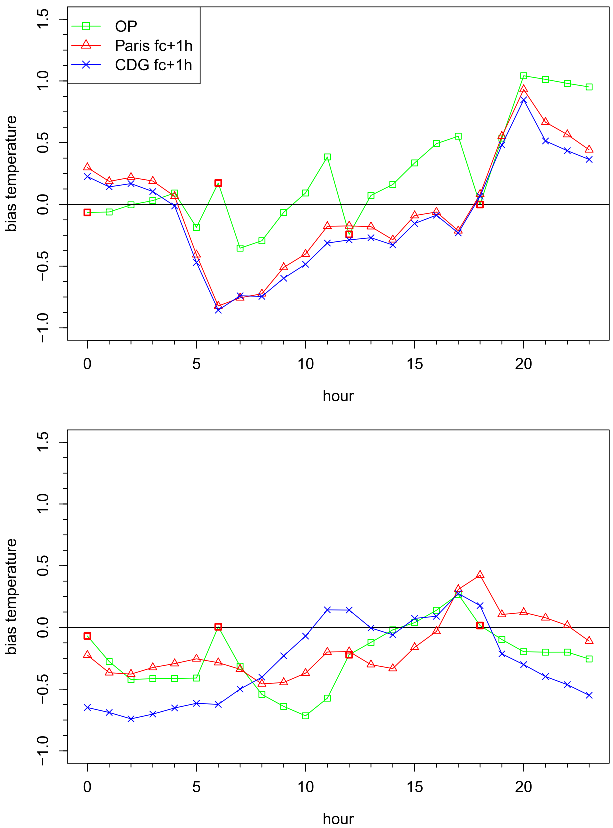

FIG. 8. Hourly bias of the 2-m temperature during (top) XP0 and (bottom) XP1 for the operational AROME (green squares), Paris domain (red triangles), and CDG domain (blue crosses) simulations. The red squares indicate the synoptic hours, where the data from the operational model correspond to the analysis and not to a forecast.

persistence method. The statistical significance between the differences in the RMSE values for the different AROME configurations for the 2-m temperature and the 10-m wind speed is presented in the appendix (Table A3).

The RMSE values for the 10-m wind speed during XP0 (Fig. 9, bottom left) show that the persistence method is outperformed by all models. During XP1 (Fig. 9, bottom center) the forecast by the persistence method is worse than both versions of AROME Airport but better than the operational AROME for the first forecast hour; thereafter, all AROME configurations have lower RMSE values than the forecasts by the persistence method. Looking at only the AROME models, the wind speed is most accurately described (as shown in the RMSE values) by AROME Airport with the high-resolution CDG domain during XP0, and the worst model is AROME Airport on the larger Paris domain. During XP1 the situation is different and now AROME Airport for both the Paris and the CDG domains have the lowest RMSE values and the worst of the models is the operational AROME calculated using the realistically available hours. Calculating the RMSEs of the wind speed for both periods together shows that there are small differences in the performance of the models both with regard to the forecast hour and between the models themselves (Fig. 9, bottom right); however, the high-resolution model does have the smallest RMSE values, and as expected, the worst-performing model is again the operational AROME calculated using the realistically available forecasts. The forecast by persistence method is more reliable for the wind speed than for the temperature, but it is still better to rely on the model forecasts.

5. Discussion

a. *Sensitivity of the data assimilation system*

To better understand why the operational model appears to provide a more accurate description of the 2-m temperature, if the same forecast hours are compared with one another, further tests on the data assimilation system were performed looking at

(i) increasing the background error covariances (**B** matrix) used in the data assimilation by a multiplicative coefficient (i.e., configuration i),
(ii) introducing a stricter cutoff limit so that observations that arrive too late are not included in the data assimilation (i.e., configuration ii), and
(iii) removing all observations of the temperature and the humidity at 2 m from the data assimilation (i.e., configuration iii).

To better quantify the effect of the data assimilation, the scores are calculated on the larger Paris domain, where approximately 500 observations are available for each hour. The RMSE and the bias of the 2-m temperature for the Paris domain are shown in Fig. 10. These scores are calculated for the synoptic hours during a test week during XP1 (8–14 October 2012). In Fig. 10, there are two curves representing the operational AROME model, which, just as in section 4e, depends on whether the data come from forecasts originating from the synoptic hour in question or whether they originate from the most recent forecasts that would be available in real time. The forecasts that are available in the real-time setting are also the forecasts that are used as initial and lateral boundary conditions by AROME Airport; these forecasts can be up to 9 h old. With this in mind, it is not surprising that in Fig. 10 both the lowest RMSE values and the smallest bias come from the operational AROME using the most recent forecasts (green solid lines) and

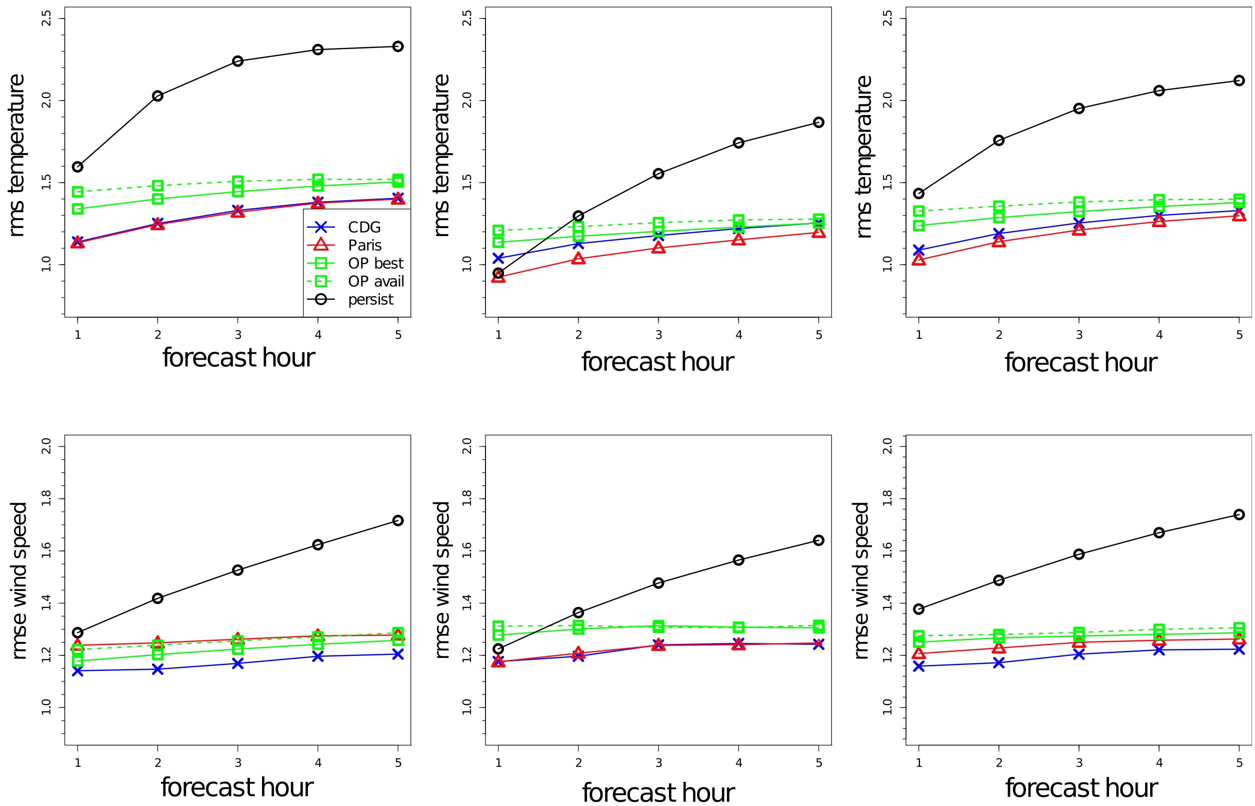

FIG. 9. RMSE for the (top) 2-m temperature and (bottom) 10-m wind speed during (left) XP0 and (center) XP1, and (right) for both periods combined for the AROME models compared to the forecast by persistence method (black circles). Green squares with solid lines are for the operational model where the forecasts are the best match for the hour, whereas the dashed lines are for the same scores from the forecasts that are actually available at the given time. For AROME Airport the red triangles are the forecasts from the Paris domain and the blue crosses come from the high-resolution CDG domain.

the worst from the operational AROME with the forecast that would have been available in real time (green dashed lines). The data from AROME Airport with the lowest RMSE and bias come from the configuration with the multiplicative coefficient in front of the **B** matrix (configuration i), which has a small but positive influence on the scores with respect to the standard configuration (triangles), whereas the introduction of a stricter cutoff limit (crosses) marginally worsens the scores. The worst-performing AROME Airport configuration is configuration iii (blue diamonds), which does not assimilate the 2-m temperature or the relative humidity. This configuration does not present any improvement with respect to the operational AROME data, which are available in a real-time setting.

These data show that the AROME Airport system can provide an improvement in the short-range forecasts for the 2-m temperature when compared to the operational AROME forecasts available at the same time. This comparison also shows that using screen-level observations in the data assimilation has a large positive influence on the forecasts, something that has also been shown by Brousseau et al. (2013), though the impact on the forecast skill of both a specifically tuned **B** matrix and of a long cutoff, while positive, seems less important than in the case of forecast lead times beyond nowcasting ranges.

b. *Initialization files and boundary conditions*

In an attempt to verify the hypothesis that the forecasts from AROME Airport would be better if the model was using more recent initial conditions, boundary conditions, and surface analyses, a test was performed where AROME Airport was run with unrealistic startup conditions, that is, starting AROME Airport with files from the operational AROME that in a real-time system would not be available yet. A number of test simulations using combinations of the initial startup conditions (guesses), surface analyses, and coupling files from different starting times were tested. Using the difference in the 2-m temperature between the operational AROME and AROME Airport on the Paris

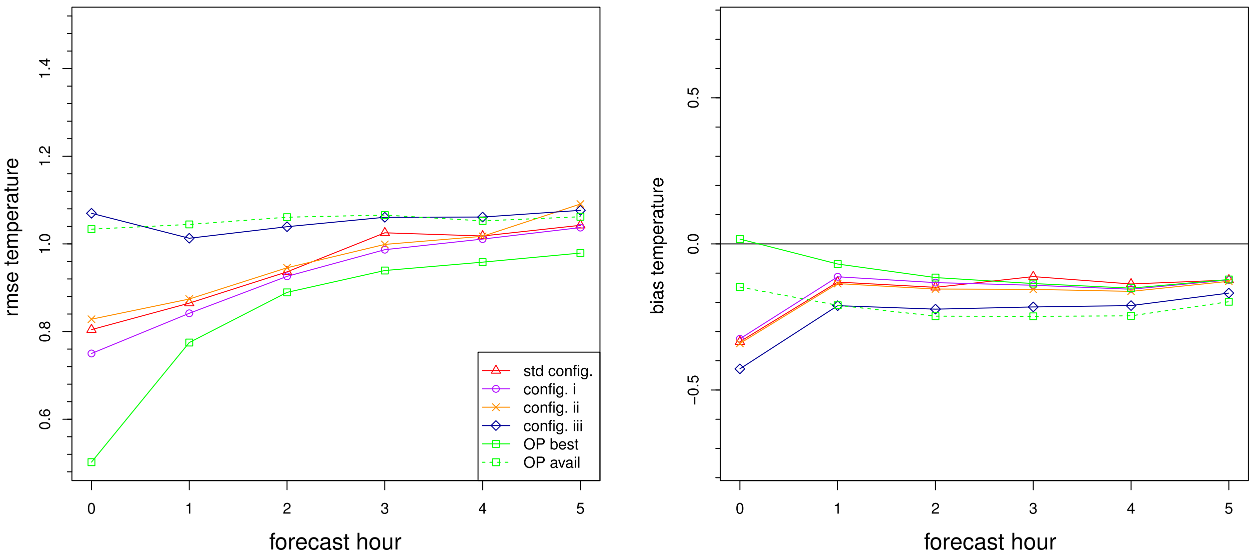

FIG. 10. (right) RMSE and (left) bias for the 2-m temperature with differences in the data assimilation calculated for the larger Paris domain. Green-squared lines are for the operational model where the solid lines are the forecasts from each synoptic hour, whereas the dashed lines the same scores from the forecasts that are actually available at the given time. For AROME Airport the red triangles are for the standard configuration, purple circles are for configuration i, orange crosses show configuration ii, and navy diamonds are for configuration iii.

domain as an indication of the impact of the different configurations, there is always a benefit in using an up-to-date surface analysis from the operational model. These tests show that AROME Airport would certainly benefit from having its own surface analysis. However, adding an extra step for the surface analysis would also result in a much more complex configuration. Using synchronous files for the initialization conditions and lateral boundary conditions (i.e., using forecasts from the same AROME run as in the initialization and for the lateral boundary conditions) also improved the performance of AROME Airport. However, due to the time constraint of a future operational AROME Airport system, it is necessary to use asynchronous data, because the data analysis also showed that it is better to use the most recently available forecast for the initialization and the lateral boundary conditions.

6. Conclusions

This study presents the AROME Airport system, including its configuration and capabilities. AROME Airport is a high-resolution NWP model that is intended to be used as an aid for nowcasting purposes. The output of this model is also meant to be used as initialization data for a wake vortices prediction model to forecast the wake vortices of airplanes during landings and takeoffs. For this purpose AROME Airport provides forecasts of both the classic meteorological parameters, such as wind speed and temperature, as well as the turbulence parameters TKE and EDR.

This study has focused on the verification of the high-resolution AROME model, with a resolution of 500 m, surrounding Paris Charles de Gaulle Airport. This domain is very flat, so the higher resolution is not expected to provide any dramatic improvement in the forecast with respect to the operational version of the AROME model.

Looking at the TKE of AROME Airport at different resolutions ($dx = 500$ and 2500 m) in a free convective boundary layer showed that the difference in the resolved energy (TKE) between the different resolutions was very small and the subgrid kinetic energy was very similar for both resolutions. This means that the effective resolution when using the 500-m grid is larger than the largest eddies in the boundary layer and that the shallow convection parameterization can be used.

The scores (root-mean-square error and bias) for the forecasts from AROME Airport have been calculated for the wind speed and temperature during two observational campaigns at CDG, during May–June 2011 (XP0) and September–October 2012 (XP1). During these periods several vertical profilers measuring the wind speed were available in addition to the standard observations.

Calculating the scores for AROME Airport and comparing the results to those of the operational AROME for the synoptic hours, the best scores for the

10-m wind speed are found for the high-resolution model of AROME Airport during XP0 whereas during XP1 there is practically no difference between the high-resolution CDG domain and the Paris domain and both of them are better than the operational model. Combining the data from the two periods demonstrates that the CDG domains have slightly lower RMSE values than do the other models. Comparing the models with the data from the wind speed profiles, the scores are very similar but with a very small improvement for the high-resolution model. Looking at the 2-m temperature for the forecasts starting at the synoptic hours, the best RMSE values are obtained using the operational model. The bias is inconclusive and shows a positive bias during XP0 whereas during XP1 it is negative, which leads to an overall bias that is almost neutral. The differences between the different model configurations are small and it is no surprise that a test of the statistical significance of these scores shows no statistical significance for the RMSE of the 2-m temperature and only for a few hours between the CDG domain and the other two domains for the 10-m wind speed. The differences in the bias were found to be more statistically significant, especially for the 2-m temperature.

However, because AROME Airport is designed to run every hour, calculating the same scores for the wind speed and the temperature in the boundary layer and comparing the results to what realistically is available from the operational model, as well as the forecast made by the persistence method, gives a much more realistic view of the capabilities of AROME Airport. During XP0 the high-resolution AROME Airport simulation provides the lowest RMSE values for the 10-m wind speed whereas during XP1 both the high-resolution CDG domain and the Paris domain are comparable. Combining the two periods shows that overall there are small differences in the RMSEs for all models, but the high-resolution CDG domain provides the lowest RMSE values and the forecast by persistence method is outperformed by all of the AROME configurations. The 2-m temperature is equally well forecasted over the Paris and the CDG domains during XP0 whereas during XP1 the Paris domain outperforms the others, even though for the first forecast hour the forecast by persistence method has an almost equally low RMSE. Combining the data for both periods shows that the overall best model is the Paris domain and again the forecast by persistence method is outperformed by all model configurations. The statistical significance was only calculated for the RMSEs for these runs and interpreting the results is complicated. The only clear conclusion is that the results for the 2-m temperature were found to be more statistically significant than the results for the 10-m wind speed.

The sensitivity of the data assimilation and its impact on the 2-m temperature was also shown. The data assimilation improves the results from AROME Airport on the Paris domain, which is the domain where the data assimilation of AROME Airport is performed. It was also shown that the screen-level observations and radar observations can be more important for short-term forecasts than for longer-term forecasts, which are more driven by the large-scale initialization.

Having tested the impact of the initialization files and the lateral boundary conditions, it was found that the forecasts from AROME Airport are improved if the initialization files and the files used as boundary conditions are synchronous; however, it is also beneficial to use the latest available forecasts and analyses for these tasks. Because AROME Airport is supposed to be run in an operational environment, it is not possible to benefit from synchronous initialization conditions.

A future version of AROME Airport might benefit from a cycled configuration in order for it to take advantage of its own forecasts. AROME Airport would also benefit from a $\mathbf{B}$ matrix that is recalculated and adjusted specifically for AROME Airport. This would optimize the data assimilation. However, the statistical sample provided here is yet too small to provide an accurate sample for this task.

All in all, AROME Airport has shown some encouraging results for the Paris CDG area at high resolution. Despite the fact that the high-resolution domain is rather small (100 km $\times$ 100 km) and the orography here is very smooth, the results for the 10-m wind speed are particularly encouraging. It would also be useful to continue this study in order to verify the model for a longer period than the 8 weeks used here to assess the ability of the model during all seasons.

AROME Airport has already been run in a semi-operational test during XP1 and further tests are planned for September 2014. Tests using AROME Airport in real time are scheduled for 2015.

Acknowledgments. This work has been supported by Thales Air Systems.

APPENDIX

Statistical Significance of the Differences between the Model Configurations

The statistical significance of the differences found in the scores for the screen-level observations between the different configurations of AROME Airport and the operational AROME model in section 4 is

TABLE A1. Summary of the forecast hours for which the differences in the scores between the different AROME configurations are statistically significant for the 10-m wind speed for the synoptic hours in Fig. 4. The first column shows the results for the high-resolution CDG domain and the larger Paris domain, the second column is between the CDG domain and the operational model, and the third column is for the Paris domain and the operational model.

	CDG–Paris (h)	CDG–OP (h)	Paris–OP (h)
RMSE			
XP0	—	—	—
XP1	—	—	—
Both	—	—	—
Bias			
XP0	1–5	1–5	1–5
XP1	1–5	3–5	2–5
Both	1–5	1–5	1–5

TABLE A2. Summary of the forecast hours for which the differences between the scores for different AROME configurations are statistically significant for the 2-m temperature for the synoptic hours in Fig. 7. The first column shows the high-resolution CDG domain and the larger Paris domain, the second column is between the CDG domain and the operational model, and the third column is for the Paris domain and the operational model.

	CDG–Paris (h)	CDG–OP (h)	Paris–OP (h)
RMSE			
XP0	—	—	—
XP1	1	1	—
Both	—	1–3	—
Bias			
XP0	—	1–5	1–5
XP1	1	1	—
Both	1	1–2	—

presented in this section. The statistical significance is calculated using the Student's t test with a confidence interval of 95%. These calculations are performed for the results presented in Figs. 4, 7, and 9.

Tables A1 and A2 show for which forecast hours the differences in the RMSE and the bias between the different AROME configurations were statistically significant for the 10-m wind speed (Table A1) and the 2-m temperature (Table A2) for the synoptic hours. For the 10-m wind speed none of the differences found in the RMSE values are statistically significant whereas the differences in the bias are statistically significant for almost all of the forecast hours. For the 2-m temperature the only differences in the RMSEs found to be statistically significant were between the CDG domain and the operational model for the first forecast hour during XP1 and for the first 3 h when the RMSE is calculated for both XP0 and XP1 combined. The difference in the bias is statistically significant between both domains of AROME Airport and the operational AROME model during XP0. During XP1 only the difference in the bias for the first forecast hour between the CDG domain and the Paris domain and the difference between the CDG domain and the operational model are statistically significant. The statistical significance for the differences between the model configurations when both of the studied periods are combined is only present for the first forecast hour between the Paris and CDG domains and for the first 2 h between the CDG domain and the operational model.

Table A3 shows the statistical significance of the differences in the RMSE values for the forecasts for all 24 h for both the 2-m temperature and the 10-m wind speed. For the 2-m temperature almost all of the differences found in the first forecast hour are statistically significant and for the forecast hours that follow, fewer and fewer of the differences found between the different configurations are statistically significant. For the 10-m wind speed only three of the differences between the different model configurations are statistically significant. These

TABLE A3. Summary of the forecast hours for which the differences between the RMSE values for the different AROME configurations are statistically significant for the 2-m temperature and the 10-m wind speed for all 24 h in Fig. 9. CDG is the high-resolution domain, Paris is the Paris domain, OP refers to the most recent forecasts from the operational model, and OPre indicates the forecasts from the operational model that are available in a real-time setting.

	CDG–Paris (h)	CDG–OP (h)	CDG–OPre (h)	Paris–OP (h)	Paris–OPre (h)	OP–OPre (h)
2-m temp						
XP0	—	1–3	1–5	1–5	1–5	1
XP1	1	—	1	1–2	1–3	—
Both	1	1–2	1–4	1–4	1–5	1
10-m wind						
XP0	—	—	—	—	—	—
XP1	—	—	—	—	1	—
Both	—	1–2	1–2	—	—	—

are the differences between the CDG domain and the operational model (both the most recent forecasts and the realistically available forecasts) for the first 2 h and between the Paris domain and the older, realistically available, forecast from the operational model for the first hour.

REFERENCES

Brousseau, P., G. Desroziers, F. Bouttier, and B. Chapnik, 2013: *A posteriori* diagnostics of the impact of observations on the AROME-France convective-scale data assimilation system. *Quart. J. Roy. Meteor.,* **140,** 982–994, doi:10.1002/qj.2179.

Brovelli, P., S. Sénési, E. Arbogast, P. Cau, S. Cazabat, M. Bouzom, and R. Reynaud, 2005: Nowcasting thunderstorms with SIGOONS: A significant weather object oriented nowcasting system. *World Weather Research Program Symposium on Nowcasting and Very Short Range Forecasting,* Toulouse, France, WMO, 7.08.

Browning, K. A., C. G. Collier, P. R. Larke, P. Menmuir, and R. G. Owens, 1982: On the forecasting of frontal rain using a weather radar network. *Mon. Wea. Rev.,* **110,** 534–552, doi:10.1175/1520-0493(1982)110<0534:OTFOFR>2.0.CO;2.

Deardorff, J. W., 1970: Convective velocity and temperature scales for the unstable planetary boundary layer and for Rayleigh convection. *J. Atmos. Sci.,* **27,** 1211–1213, doi:10.1175/1520-0469(1970)027<1211:CVATSF>2.0.CO;2.

EUROCONTROL, cited 2012: Performance review report: An assessment of air traffic management in Europe during the calendar year 2011. [Available online at http://www.eurocontrol.int/prc.]

Fabbian, D., R. de Gear, and S. Lellyet, 2007: Application of artificial neural network forecasts to predict fog at Canberra International Airport. *Wea. Forecasting,* **22,** 372–381, doi:10.1175/WAF980.1.

Gerz, T., F. Holzäpfel, W. Bryant, F. Köpp, M. Frech, A. Tafferner, and G. Winkelmans, 2005: Research towards a wake-vortex advisory system for optimal aircraft spacing. *C. R. Phys.,* **6,** 501–523, doi:10.1016/j.crhy.2005.06.002.

Gultepe, I., M. D. Müller, and Z. Biybei, 2006: A new visibility parameterization for warm-fog applications in numerical weather prediction models. *J. Appl. Meteor. Climatol.,* **45,** 1469–1480, doi:10.1175/JAM2423.1.

Huang, L. X., G. A. Isaacs, and G. Sheng, 2012: Integrating NWP forecasts and observation data to improve nowcasting accuracy. *Wea. Forecasting,* **27,** 938–953, doi: http://dx.doi.org/10.1175/WAF-D-11-00125.1.

Isaacs, G. A., and Coauthors, 2014: The Canadian Airport Nowcasting System (CAN-Now). *Meteor. Appl.,* **21,** 30–49, doi:10.1002/met.1342.

Li, P. W., and D. S. Lau, 2008: Development of a lightning nowcasting system for Hong Kong International Airport. Preprints, *13th Conf. on Aviation, Range, and Aerospace Meteorology,* New Orleans, LA, Amer. Meteor. Soc., P1.13. [Available online at https://ams.confex.com/ams/pdfpapers/130793.pdf.]

Libby, P. A., 1996: *Introduction to Turbulence*. Taylor and Francis, 341 pp.

Mass, C., 2012: Nowcasting: The promise of new technologies of communication, modelling and observation. *Bull. Amer. Meteor. Soc.,* **93,** 797–809, doi:10.1175/BAMS-D-11-00153.1.

Moeng, C. H., 1984: A large-eddy-simulation model for the study of planetary boundary-layer turbulence. *J. Atmos. Sci.,* **41,** 2052–2062, doi:10.1175/1520-0469(1984)041<2052:ALESMF>2.0.CO;2.

Pergaud, J., V. Mason, S. Malardel, and F. Couvreux, 2009: A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction. *Bound.-Layer Meteor.,* **132,** 83–106, doi:10.1007/s10546-009-9388-0.

Rasmussen, R., and Coauthors, 2001: Weather Support to Deicing Decision Making (WSDDM): A winter weather nowcasting system. *Bull. Amer. Meteor. Soc.,* **82,** 579–595, doi:10.1175/1520-0477(2001)082<0579:WSTDDM>2.3.CO;2.

Reynolds, O., 1895: On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Philos. Trans. Roy. Soc. London,* **186A,** 123–164, doi:10.1098/rsta.1895.0004.

Ricard, D., C. Lac, S. Riette, R. Legrand, and A. Mary, 2013: Kinetic energy spectra characteristics of two convection-permitting limited-area models AROME and Meso-NH. *Quart. J. Roy. Meteor. Soc.,* **139,** 1327–1341, doi:10.1002/qj.2025.

Seity, Y., P. Brosseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Mason, 2011: The AROME-France convective-scale operational model. *Mon. Wea. Rev.,* **139,** 976–991, doi:10.1175/2010MWR3425.1.

Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.,* **132,** 3019–3032, doi:10.1175/MWR2830.1.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Mon. Wea. Rev.,* **91,** 99–164, doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Smith, W. L., V. E. Suomi, F. X. Zhou, and W. P. Menzel, 1982: Nowcasting applications of geostationary satellite atmospheric sounding data. *Nowcasting,* K. Browning, Ed., Academic Press, 123–135.

Stull, R., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic, 666 pp.

Tennekes, H., and J. L. Lumley, 1972: *A First Course in Turbulence*. The MIT Press, 300 pp.

Wyngaard, J. C., and O. R. Coté, 1971: The budgets of turbulent kinetic energy and temperature variance in the atmospheric surface layer. *J. Atmos. Sci.,* **28,** 190–201, doi:10.1175/1520-0469(1971)028<0190:TBOTKE>2.0.CO;2.