On the Prediction of Stratospheric Balloon Trajectories: Improving Winds with Mesoscale Simulations

VALÉRIAN JEWTOUKOFF, a RIWAL PLOUGONVEN, AND ALBERT HERTZOG
Laboratoire de Météorologie Dynamique, CNRS, Ecole Polytechnique, Palaiseau, France

CHRIS SNYDER AND GLEN ROMINE
National Center for Atmospheric Research, b Boulder, Colorado

(Manuscript received 27 May 2015, in final form 13 May 2016)

ABSTRACT
Safety compliance issues for operational studies of the atmosphere with balloons require quantifying risks associated with descent and developing strategies to reduce the uncertainties at the location of the touchdown point. Trajectory forecasts are typically computed from weather forecasts produced by an operational center, for example, the European Centre for Medium-Range Weather Forecasts. This study uses past experiments to investigate strategies for improving these forecasts. Trajectories for open stratospheric balloon (OSB) short-term flights are computed using mesoscale simulations with the Weather and Research Forecasting (WRF) Model initialized with ECMWF operational forecasts and are assimilated with radio soundings using the Data Assimilation Research Testbed (DART) ensemble Kalman filter, for three case studies during the Strapolé 2009 campaign in Sweden. The results are very variable: in one case, the error in the final simulated position is reduced by 90% relative to the forecast using the ECMWF winds, while in another case the forecast is hardly improved. Nonetheless, they reveal the main source of forecasting error: during the ceiling phase, errors due to unresolved inertia–gravity waves accumulate as the balloon continuously experiences one phase of a wave for a few hours, whereas they essentially average out during the ascent and descent phases, when the balloon rapidly samples through whole wave packets. This sensitivity to wind during the ceiling phase raises issues regarding the feasibility of such forecasts and the observations that would be needed. The ensemble spread is also analyzed, and it is noted that the initial ensemble perturbations should probably be improved in the future for better forecasts.

1. Introduction
Several types of balloons are used to provide upper-air observations: radiosondes, superpressure balloons, and zero-pressure balloons (Pettifer 2009). For several reasons, the ability of forecasting balloon trajectories can be of great importance: safety is the primary concern, as scientific balloon activities are viable only if they can be performed without putting in danger the overflown population. But other reasons are directly associated with the scientific activities and include, for instance, the efficient planning of measurements, with the prospect of combining several measurements to sample the same air mass (match technique), making both measurements more valuable (e.g., Rex et al. 1997, 1998; Cirisan et al. 2014; Inai et al. 2013).

Producing a balloon trajectory forecast \([x(t), y(t), z(t)]\) will involve two components:
1) a numerical weather prediction (NWP) model to forecast the evolution of the environment (wind, temperature, and infrared radiation) and
2) a model describing the balloon’s behavior and displacement within this environment (flight physics model).

Uncertainty in forecasts of balloon trajectories will come from both components, in different proportions
depending on the balloon type and meteorological situation. Nonetheless, horizontal displacements of the balloon at a given time in a trajectory (i.e., $dx/dt$ and $dy/dt$) will be mostly dependent on the horizontal wind (hence, on the forecast from the NWP model, provided that an accurate advection scheme is used), whereas vertical displacements (i.e., $dz/dt$) will be more dependent on the understanding of the flight physics.

The balloons we consider in this study are open stratospheric balloons (OSBs, which are also called zero-pressure balloons) because they carry the largest payloads and hence impose the most stringent safety requirements. While numerous studies have addressed the physics of such balloon flights (e.g., Kreith and Kreider 1974; Romero et al. 1981; Carlson and Horn 1983; Franco and Cathey 2004; Dai et al. 2012), our objective in the present study is to investigate the uncertainty tied to the forecast of the environment (point 1 above). Since the 1970s, the Centre National d’Études Spatiales (CNES; French space agency) has been using these OSBs, which can fly for a few hours (3–12 typically) with loads up to 1 t, to perform scientific missions devoted to atmospheric sciences or to astronomy. OSB flights generally consist of three distinct phases: an ascent, which typically lasts 1.5–2 h; a ceiling, during which the altitude (between 20 and 35 km for the cases considered) varies slowly; and a quick descent under parachute (less than an hour). A long-standing concern related to this kind of flight is the risk associated with the final touchdown location, typically of the order of several tens of kilometers. This uncertainty seriously reduces the places on Earth where such OSB flights can occur, as well as the flight opportunities during a flight campaign, since the larger the uncertainty, the wider the landing polygon.

The time scale for such balloon flights is associated with the horizontal travel of several tens of kilometers. Common operational practice consists of forecasting balloon trajectories from forecasts produced with a global NWP model, for example, from the European Centre for Medium-Range Weather Forecasts (ECMWF; Lebeau and Sanfourche 2011). Given the small scale of the trajectory, it is possible however that a mesoscale meteorological model would be suitable. A great deal of effort has been devoted in the last decade to improving atmospheric flow predictions using high-resolution mesoscale models (Coniglio et al. 2010) and assimilating surface and radiosondes (Faccani et al. 2009; Benjamin et al. 2010; Wheatley et al. 2012; Choi et al. 2015), radar (Snyder and Zhang 2003; Zhang et al. 2004; Caya et al. 2005; Aksoy et al. 2009, 2010), and satellite data (Hoffman et al. 1990; Bauer et al. 2011; Liu et al. 2012). Using winds provided by a mesoscale model to forecast balloon trajectories would thus benefit from a higher resolution and allow for the assimilation of observations that are not currently used in global forecasts.

Qualitatively, one expects that the stratospheric flow will be rather large scale and very well approximated by balanced dynamics in the extratropics (Hoskins et al. 1985). This implies that satellite observations of temperature, which are assimilated in global NWP models, constrain the wind field well. Hence, forecasts of the stratospheric flow from global NWP models are generally considered as being reliable (Simmons et al. 2005). Quantitatively, the errors in the analyzed winds in the lower stratosphere have been studied using independent wind observations gathered during long-duration (i.e., several days) superpressure balloon flights (e.g., Hertzog et al. 2003). In the mid- and high latitudes, these studies have actually reported rather small errors (Bocca et al. 2008), in contrast to tropical latitudes (Podglajen et al. 2014). For these reasons, our a priori expectation was that errors in forecast trajectories would not predominantly come from the ceiling phase of the OSB flights. Yet, our results below will reveal the contrary, that is, that errors in the forecasts for the three cases studied come mainly from the ceiling phase in the lower stratosphere.

The improvements that can be gained from assimilation of additional observations has been examined in the ocean to predict drifter trajectories (Özgökmen et al. 2000; Piterbarg 2001; Özgökmen et al. 2001) and to show encouraging results for midterm forecasting (typically 5–10-day lead times). In the atmosphere, Keil et al. (2001) have compared balloonborne observations with analyses and forecasts for midterm prediction of long-duration balloon trajectories. They showed that better initial conditions resulted in better forecasts for 2–5-day lead times with a reduction of 50% on average of the distance between the simulated and real balloons in the Northern Hemisphere. But this study did not focus on shorter-time forecasts (several hours), which has to be considered for OSB flights. At these time scales, mesoscale processes could have an influence on the balloon...
trajectories, but, to our knowledge, studies addressing the impact of such processes on the accuracy of short-term predictions of atmospheric balloon trajectory are still lacking.

The aim of the present work is thus to explore possible improvements of trajectory forecasts for OSBs using a mesoscale model and data assimilation. As previously mentioned, our focus is entirely on reducing the uncertainty associated with the wind field forecast. A straightforward and clean way to isolate this uncertainty from that mentioned, our focus is entirely on reducing the uncertainty associated with the wind field forecast. A straight-forward and clean way to isolate this uncertainty from that tied to flight physics is to work on hindcasts of past flights, for which the balloon altitude \( \zeta(t) \) is known. This is nearly equivalent to assuming that we have a perfect flight physics model and hence a corresponding perfect prediction of the balloon altitude.

The article is organized as follows. In section 2, we present the Strapolété campaign, and the numerical and experimental setup of the numerical simulation and data assimilation. In section 3, results from those experiments are shown for the case studies. In the last section, we give our conclusions about the improvement of predictability for balloon trajectory forecasts and discuss the limitations of our analysis.

### 2. Data and methodology

#### a. The Strapolété summer campaign

From 2 August to 16 September 2009, CNES released nine OSBs from the Esrange Space Center launch base (67.894°N, 21.107°E) near Kiruna in Sweden (see flight information in Table 1) (Payan et al. 2010). OSBs are open (zero pressure) balloons that have one or several openings in their envelopes. These balloons are used for short-lived experiments (\( \sim 1-10 \) h) and fly up to the upper stratosphere with significant payloads up to 1 ton. OSBs are subject to a complex thermodynamic/radiative equilibrium. Their volume varies as the visible and infrared radiative fluxes change, since these fluxes control the gas temperature inside the balloon. In addition to this fairly complex flight physics, the OSB flight is not passive unlike a radiosonde: the operator has some control on the altitude through the possibility of gas discharge or ballast release. This further complicates the prediction of their trajectories (Alexander and de La Torre 2003). In the Strapolété campaign, CNES used balloons with volumes ranging from 35,000 to 400,000 m\(^3\) that carried scientific payloads between 100 kg and 1 ton. The three-dimensional positions of the balloon are recorded every 10 s by a GPS receiver in the flight chain, and the horizontal velocities of the wind along the trajectory are inferred from the successive balloon positions: we thus assume that the balloon is a perfect tracer of the horizontal components of the wind. The ambient temperature and pressure are furthermore measured with the same time resolution during the flights. Note that the OSB observations were not assimilated within the ECMWF model suite during the campaign. The Strapolété campaign was designed to study stratospheric processes and chemical composition. Once in the midstratosphere, the balloons thus drifted for a few hours at altitudes between 29 and 36 km. OSB-altitude time series and OSB trajectories are shown in Fig. 1 for the three Strapolété balloons used in this study (see also Table 1).

During the Strapolété campaign, CNES also released daily radiosondes from the launching site typically 6–10 h before the OSBs flights to evaluate the meteorological conditions. Those radiosondes constitute the uniquely available observations to provide detailed information on the atmospheric state before each OSB launch. Measurements in these radiosondes are done every 2 s from the surface up to about 35 km. Just like the OSBs, the radiosondes were not assimilated in the ECMWF system. Information about the Strapolété radiosondes is summarized in Table 1.

#### b. Computation of balloon trajectories

In this study, we will perform numerical simulations of the balloon trajectories in various dynamical fields as described in the next section. Here, we present the advection model that was used to calculate these trajectories. As previously explained, the complex OSB flight physics and the balloon pilot actions on the balloon require the use of a flight physics model to thoroughly simulate the evolution of the balloon altitude during the flight. This kind of modeling is in itself a subject of much research, which, for instance, has to account for heat exchanges through radiation, conduction, and

### Table 1. Strapolété balloon flights and corresponding radiosondes. RRTMG denotes Rapid Radiative Transfer Model for general circulation models.

<table>
<thead>
<tr>
<th>Case</th>
<th>Date</th>
<th>Launch time (UTC)</th>
<th>Duration</th>
<th>Max altitude (km)</th>
<th>Launch time (UTC)</th>
<th>Max altitude (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16 Aug 2009</td>
<td>1620</td>
<td>7 h, 13 min</td>
<td>30.142</td>
<td>0552</td>
<td>34.3</td>
</tr>
<tr>
<td>2</td>
<td>2 Aug 2009</td>
<td>1800</td>
<td>4 h, 00 min</td>
<td>29.56</td>
<td>1349</td>
<td>36.33</td>
</tr>
<tr>
<td>3</td>
<td>14 Aug 2009</td>
<td>0926</td>
<td>5 h, 58 min</td>
<td>33.5</td>
<td>0156</td>
<td>36.277</td>
</tr>
</tbody>
</table>

RRTMG denotes Rapid Radiative Transfer Model for general circulation models.
convection between the balloon envelop, the lifting gas, and the environment, as well as the complex interactions of these exchanges with the shape of the balloon and the resulting lifting force. Including such a model in our trajectory computations would unavoidably generate uncertainties in the balloon altitude, which would translate to further uncertainties in the simulated balloon horizontal positions, given the predominant vertical shear of the winds in the stratosphere. Hence, we have adopted another approach in this study in which we focus on the hindcasts of balloon flights: we have used the real balloon pressure \(P_{rb}(t)\) in the trajectory simulations. Specifically, we have interpolated at each time step \(t_i\) along the simulated trajectories the dynamical fields at \(P_{rb}(t_i)\) while letting the forecast zonal and meridional wind components advect the balloon horizontally. This enables us to focus on only the trajectory uncertainties associated with the prediction of the atmospheric dynamics. In operations, one has to use a flight physics model, which adds uncertainties to those coming from the atmospheric forecast, which is discussed in this study. Here, since we use the real balloon altitudes in our simulations, we are unable to assess the impact of an imperfect balloon-altitude simulation on the quality of the trajectory forecast.

The advection model we use thus integrates the kinematic equations

\[
\begin{align*}
\frac{dx_{sb}}{dt} &= u(x_{sb}, y_{sb}, P_{rb}, t), \\
\frac{dy_{sb}}{dt} &= v(x_{sb}, y_{sb}, P_{rb}, t),
\end{align*}
\]

where \((x_{sb}, y_{sb})\) refers to the horizontal coordinates of the simulated balloon, and \((u, v)\) are the zonal and meridional wind components in either the ECMWF control forecast or the various mesoscale model simulations. The interpolation of the dynamical fields is done linearly in time and with cubic spline in the three spatial directions. The integration is performed with a fourth-order Runge–Kutta scheme, and we have checked that our integration time step (10 s) is small enough so that our results are not sensitive to the chosen integration scheme. The dynamical fields we used are available every 6 h on 15 pressure levels (1000, 925, 850, 700, 500, 300, 250, 200, 100, 50, 10, 7, 5, 3, 2 hPa), with a horizontal resolution of 0.25° × 0.25° (i.e., 20–30 km) for the ECMWF forecast. The WRF fields are made available every hour with a horizontal grid spacing of 10 × 10 km² on model levels.

In Eq. (1), we assume that the balloon is a perfect tracer of the horizontal wind. This is not the case, as the balloon inertia will induce some delay before the

![Fig. 1. Altitude of the OSB for (a) case 1 (16 Aug 2009), (b) case 2 (2 Aug 2009), and (c) case 3 (14 Aug 2009) as a function of time. Red triangles denote the start (upward pointing) and end of the ceiling (downward pointing).](image)
balloon adjusts to wind changes. Vial et al. (2001) have studied the response of spherical balloons to such changes. We use here their formulas for the adjustment time scale and for the geographical separation between a real balloon and one without inertia (see their appendix A). In doing so, we thus consider that the OSB is nearly spherical, which is likely valid at least at first order in the stratosphere but more questionable when the balloon is not fully inflated, that is, during the ascent and final descent phases. With a typical balloon radius of 50 m, the adjustment time scale for an initial difference between the balloon velocity and that of the air of 1 m s\(^{-1}\) is \(\sim 400\) s, which is shorter than the time scale of most of the wave motions in the stratosphere. It is recalled that this time scale scales as the inverse of the initial velocity difference, and thus it reduces if the initial velocity difference increases. Integrating the velocity difference in time yields a geographical separation between the real balloon and the one simulated with Eq. (1) of less than 2500 m after 10 h of simulations even if the initial velocity difference is as large as 5 m s\(^{-1}\). Such separations are much smaller than those resulting from errors in the wind forecasts, which are discussed below, and we can therefore safely simulate the real balloon trajectory with Eq. (1).

c. Numerical modeling and data assimilation setup

The numerical simulations in this study are done using the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2008). We use the Advanced Research version of WRF (ARW) dynamical core. The WRF Model is a nonhydrostatic compressible mesoscale model well adapted for mesoscale prediction. It uses a staggered horizontal Arakawa C grid (Arakawa and Lamb 1977) and a vertical stretched terrain-following grid. The dynamical core integrates the momentum and conservation equations using a flux-divergence fifth-order scheme. Time integration uses a third-order Runge–Kutta split–explicit scheme to handle the acoustic waves on a small time step (Wicker and Skamarock 2002). The model is well suited for data assimilation experiments and has been widely used with the Data Assimilation Research Testbed (DART; Anderson et al. 2009) package for convective prediction purposes using radar data assimilation (Aksoy et al. 2009, 2010).

We use a single 1000 km \(\times\) 1000 km domain centered on Kiruna in our experiments, with a horizontal grid spacing of 10 km. Extending the domain in the vertical to include a significant portion of the stratosphere requires some manual adaptation of the code (S. Evan 2014, personal communication): manually prescribing the eta levels and imposing the mean constant temperature in the stratosphere (here \(T = 270\) K based on the mean temperature calculated in the ECMWF forecasts). Nonetheless, the depth of the domain of our experiments remains constrained by the maximum height of the ensemble forecasts used for the initial and lateral boundary conditions. Our domain uses 120 vertical levels from the surface up to the pressure level of 5 hPa (\(\leq 40\) km) to resolve accurately the flow in the lower and midstratosphere. To avoid wave reflections at the model top, a sponge layer is applied to the damp vertical velocity within the 5 km below the model lid. We use physical schemes and parameterizations that are used in NCAR’s Antarctic Mesoscale Prediction System (AMPS; Powers et al. 2012). The model outputs the dynamical fields at an hourly frequency. A summary of our WRF implementation is shown in Table 2.

The atmospheric initial and boundary conditions are provided by the ECMWF operational ensemble perturbations and forecasts from the Ensemble Prediction System (EPS; Molteni et al. 1996). The ECMWF ensemble consists of an unperturbed control forecast and 50 members generated by perturbing the operational analysis. The perturbations are constructed from singular vectors using a forward adjoint linear of the ECMWF model (Molteni et al. 1996) to get a maximal growth for lead times of 48–72 h. The forecasts are available every 6 h starting from each analysis and lasting for 10 days. The perturbed analysis and forecasts have a horizontal grid spacing of 0.25° longitude and latitude corresponding to 15–30 km. The perturbed analysis has 62 vertical model levels from the surface up to 0.01 hPa and is used as our ensemble initial conditions, while the forecasts have 16 pressure levels extending up to 5 hPa and we use them for our lateral boundary conditions during time integration. Therefore, the minimum pressure allowed by the forecasts constrains the model height we use.

We use an ensemble Kalman filter (EnKF) system to assimilate the available data. Readers are referred to Evensen (2003) for a comprehensive explanation of the EnKF technique. In the present study, DART (Anderson et al. 2009) is used to perform data assimilation of the

<table>
<thead>
<tr>
<th>Table 2. Model physics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model physics</td>
</tr>
<tr>
<td>WRF single-moment 5-class microphysics scheme</td>
</tr>
<tr>
<td>Yonsei University planetary boundary layer scheme</td>
</tr>
<tr>
<td>MM5-derived surface layer scheme</td>
</tr>
<tr>
<td>Noah land surface model</td>
</tr>
<tr>
<td>RRTMG for longwave radiation</td>
</tr>
<tr>
<td>RRTMG for shortwave radiation</td>
</tr>
</tbody>
</table>
radiosondes and part of the OSB observations in the WRF ensemble. DART uses the ensemble adjustment filter method (Anderson 2001) to control the effects of an underestimated analysis-error covariance resulting from the use of a finite number of members in an ensemble. For simplicity we use the ECMWF ensemble, and hence our WRF ensemble comprises 50 members. Limitations of this choice will be described in section 4. We use Gaspari and Cohn’s (1999) horizontal localization correlation function with a half-width localization radius of 300 km. This value is within typical ranges used for this kind of observation (Otkin 2012). The observation error variance has been determined in previous studies from a comparison between NCEP and radiosondes, with typical values of 3 m s\(^{-2}\) at the upper-tropospheric jets and a constant value of 2 m s\(^{-2}\) throughout the stratosphere. We assimilate all the available data within an assimilation window of 3 h. The boundary conditions are updated after each assimilation. We have chosen to assimilate the horizontal wind components and temperature from the nearest radiosondes. Assimilation of temperature with the wind components is necessary to enforce consistency at large scale (i.e., through thermal wind balance). Although moisture is available from our radiosonde measurements, we restrict our study to the assimilation of winds and temperature because moisture does not play a significant role in the stratosphere. Finally, we use an adaptative inflation technique (Anderson 2009) to maintain the ensemble spread.

Since the assimilation of data will be relevant only at length scales comparable to or larger than the grid spacing, we need to smooth the vertical profiles (OSB and radiosondes) prior to assimilation. We apply a local regression using a weighted linear least squares and a second-order polynomial so that it filters out the vertical wavelengths smaller than \(~500\) m. Then, the profiles are undersampled so that the ratio of observations to model levels is roughly 4.

d. Description of experimental design

We conducted three experiments using WRF and DART to investigate the impact of the assimilation of very few localized observations on the forecast of Lagrangian balloon trajectories. See Fig. 2 for schematics of all the experiments. EPS forecasts provide initialization and boundary conditions for all three experiments. The experiments are as follow:

1) WRF-Ens consists of an ensemble control forecast, started 24 h before the first available radiosonde (prior to the OSB launch) and propagated for 48 h without additional information.

2) DART-RS is similar to WRF-Ens, with the additional assimilation of a radiosonde after 24 h of simulation. The radiosonde is assimilated at its midascent time (i.e., the time evolution during the radiosonde ascent is not taken into account). The WRF ensemble is restarted from this updated analysis until the OSB touches the ground.

3) DART-all is similar to DART-RS, with the additional assimilation of the OSB observations (wind and temperature) during the ascent. This is equivalent to assimilating a single radiosonde \(~1\) h before the balloon reaches its ceiling. In contrast with experiment DART-RS, this time the assimilation takes place at launch time to constrain the low-level flow. In operational conditions, this type of setup is equivalent to a nowcasting situation. See Fig. 2 for the schematics of all the experiments.

In DART-RS and DART-all, the model is run for a typical period of \(~24\) h to allow the perturbations to grow and optimize the ensemble spread before encountering the first observation to be assimilated. Since those perturbations are constructed to grow the most unstable dynamical modes for lead times of
72 h, this value of ~24 h was found to be a good compromise between predictability, computational cost, and ensemble spread. We verified that the ensemble spread after the 24-h lead time was enough by comparing the error between the WRF ensemble and the radiosonde to the ensemble spread (10 m s$^{-1}$ for the $u$ component and 2–5 m s$^{-1}$ for the $v$ component between 7 and 10 km).

3. Case studies and results

In this section, three case studies are presented. See Table 1 for information about the selected case studies. Case 1 was chosen because the root-mean-square error (RMSE) between the balloon and ECMWF analysis velocities was the highest on this case. The other two cases were chosen because they correspond to the shortest time intervals between the radiosonde and the OSB launch time. For each of the three cases, the WRF-Ens, DART-RS, and DART-all forecasts are presented and compared to the observations and the ECMWF control forecast. The purpose of these comparisons is 1) to validate the assimilation process, 2) to evaluate the improvement in the modeled winds, and 3) to assess the impacts for the trajectories. We focus on the ensemble means in this section. All trajectories are integrated starting at the launch of the OSB except in the DART-all experiment, in which integration starts right after the OSB assimilation.

a. Case 1: 16 August 2009

The synoptic situation from 0000 UTC 15 August to 0000 UTC 17 August in the ECMWF operational analyses consists of a geopotential shortwave low at midaltitude moving eastward across the Scandinavian Mountains, with westerly winds stronger approximately 500–1000 km south of the launch site location but relatively weak above the launch site (Fig. 3a). At the upper levels, the launch site is located in a zone with very weak winds (Fig. 3b).

The vertical profiles of the horizontal velocity interpolated along the radiosonde (RS) position in the ECMWF are shown in Fig. 4 (black dashed curves). The ECMWF profile does not compare well with the observations, in particular with the absence of the jet between 8 and 12 km in the zonal wind. On the other hand, the WRF-Ens simulation (blue curves) accurately forecasts the zonal velocity ($u$), except near the lower (~6 km) and upper flanks (~12–14 km) of the tropospheric jet, where local errors of 5 and 2 to 3 m s$^{-1}$, respectively, are observed. In contrast, WRF-Ens displays significant errors in the meridional velocities ($v$): a 2 to ~3 m s$^{-1}$ low bias over most of the profile, with even higher errors in the mid- and upper troposphere (4–10 km). The DART-RS analysis (red curves) efficiently reduces these errors in the troposphere (corrections in the range of 3–5 m s$^{-1}$) and significantly improves the representation of the jet structure in the simulation, less so in the stratosphere (between 18 and 25 km), where the errors in the WRF-Ens simulations were smaller. The root-mean-square error with respect to the radiosonde is slightly reduced in the $u$ component and significantly reduced in the zonal velocity $v$ with a difference of almost 1 m s$^{-1}$ (see Table 3). Now, the critical region to predict the Lagrangian OSBs’ trajectories is between 25 and 33 km on
average, as the balloon spends most the time flying at these altitudes (ceiling phase). The red curves show a better fit to the radiosonde compared to WRF-Ens on average in this region, although the 5-km vertical wavelengths waves are damped by the ensemble mean. In Fig. 5, we have represented the vertical profiles in the DART-RS (red) and DART-all (green) analyses, along with the OSB profiles (black). The DART-all analysis provides a correction of several meters per second, in particular at the jet located at 10 km.

The trajectories integrated into the ECMWF—the WRF-Ens, DART-RS, and DART-all experiments—are represented along with the observed trajectory in Fig. 6. The trajectory in the ECMWF control forecast (dashed black line) sees mainly zero $u$ and positive $v$ wind and drifts mostly northward during the integration period. This indicates that the ECMWF forecasts have substantial errors, in particular in the stratosphere. At the end of the time integration, the distance between the OSB and the ECMWF trajectory is 90 km. In contrast, the trajectories obtained with the ensemble from WRF all capture reasonably the eastward drift in the troposphere and the subsequent southeastward motion during the ceiling phase, although small-scale details remain not well represented in the simulation. The landing points are found to be 30 km away in WRF-Ens, 15 km away in DART-RS (improved winds in the stratosphere), and only 5 km away in DART-all (improved description of the tropospheric winds). Overall, the error has been reduced by a factor of 3 due to using a mesoscale model and by a further factor of 6 due to assimilating the initial balloon ascent. We note however that all of the experiments miss the observed initial inflection in the trajectory during the ceiling phase. We have evidence (not shown) that this is because the ensemble mean significantly damps the waves with short vertical wavelengths between 28 and 30 km of altitude (this will be discussed in section 4).

### b. Case 2: 2 August 2009

The meteorological condition for case 2 is characterized by a complex flow in the troposphere with the launch site located near a saddle point between a low to the southeast and a broad anticyclone to the north (Fig. 7a). In the stratosphere, a simpler flow consisting of a broad region of strong easterly winds is found (Fig. 7b).

<table>
<thead>
<tr>
<th>Table 3: RMSE with respect to the radiosonde (upper part) and with respect to the OSB vertical profile (lower part) for the three cases.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Experiment</strong></td>
</tr>
<tr>
<td>WRF-Ens</td>
</tr>
<tr>
<td>DART-RS</td>
</tr>
<tr>
<td>DART-RS</td>
</tr>
<tr>
<td>DART-all</td>
</tr>
</tbody>
</table>
The DART-RS analysis produces large corrections as shown in Fig. 8 (up to 8 m s$^{-1}$ at an altitude of ~23 km) and efficiently corrects the velocity profile: the RMSE is reduced by about 40% in the $u$ and 60% in the $v$ components. Corrections are present in both the troposphere and stratosphere, but they are larger in the latter, where the assimilation reduces the northward and westward components of the wind. Observations of the OSB unfortunately start only at an altitude of 11 km. Their assimilation again is successful at correcting stratospheric winds that have grown too large in DART-RS (RMSE reduced by 10%–25%).

From the trajectories shown in Fig. 9, it is clear that most of the error comes from the motion of the balloon during the ceiling phase of the flight. The modeled wind in the four cases (ECMWF, WRF-Ens, DART-RS, and DART-all) does not vary significantly during this phase, resulting in a straight trajectory, with a corresponding error in location linearly increasing. The orientation and amplitude of the wind are improved by going from the ECMWF to the WRF-Ens, and further improved by the DART-RS analysis. Somewhat surprisingly, no significant improvement is obtained by the DART-all analysis. In total, the error in the landing point has been reduced by a factor of 2, but it remains 40 km away from the real landing point.

There is an important qualitative difference between all of the modeled trajectories and the real trajectory: the latter is curved during the ceiling phase, changing from northwestward to southwestward. Inspection of the velocity profile from the OSB reveals that there are shortscale fluctuations of the wind between the heights of 27 and 30 km. A hodograph of the observed wind shows clear portions of ellipses (Fig. 10), which is a signature of an inertia–gravity wave, which is absent from the models. The amplitude is significant but not especially large (1–2 m s$^{-1}$). The importance of such motions and the challenge that they represent for trajectory forecasts will be further discussed in section 4.
c. Case 3: 14 August 2009

On 14 August 2009, the flow at midaltitude displays a large synoptic cyclone located northwest of the Scandinavian Mountains, and the launch site is embedded in north-northeastward winds (Fig. 11a). In the stratosphere, the flow shows a jet exit region with relatively weak winds over the launch site (Fig. 11b).

As for previous cases, large corrections in the stratospheric winds result from the DART-RS analysis. Data assimilation improves winds that generally had a large southerly component (Fig. 12). On the other hand, the DART-all analysis has difficulty producing convincing corrections, especially on \( y \). There are strong fluctuations in the observed velocities between 20 and 33 km. The modeled profile at best follows a smoothed approximation to the observed wind (Fig. 12).

The trajectories represented in Fig. 13 clearly show that the mesoscale forecasts are rather poor in this case and that they hardly improved relative to the simulations with the ECMWF analyses. The altitude during the ceiling phase hardly changes. It is high, close to the model top, so that the model is not very reliable at those levels.

4. Discussion

a. Improvements and limitations of balloon trajectories with mesoscale simulations

On its own, the use of a mesoscale model (WRF) already has significant potential for improving the trajectory forecasts. This is illustrated by case 1, where the final position of the simulated balloon trajectory in WRF-Ens is 70 km closer to the real balloon than the
trajectory performed with ECMWF operational analyses. There are at least two reasons that could explain the benefit of the WRF-Ens simulations over those with the ECMWF operational analyses in general:

1) Using a mesoscale model first allows us to tackle the limitations in terms of model resolution and frequency of forecast outputs (Walmsley and Mailhot 1983). The ECMWF model has a typical horizontal resolution on the order of $\sim 15$ km, meaning that it is not able to simulate correctly the forcing that the balloon experiences at scales shorter than a few tens of kilometers (Kahl and Samson 1986). In our case, the slight increase of resolution in our WRF simulations ($10$ km) may be enough to resolve parts of the complex orographic flow near the Scandinavian Mountains, although the vertical resolution in our WRF configuration is too coarse to simulate short vertical waves as evidenced in case 2. Kuo et al. (1985) and Knudsen and Carver (1994) also showed that the sparse time resolution of the forecast fields ($6$ h in our use of ECMWF operational analyses) is also likely to affect the computed trajectories because such sampling may be too infrequent to identify the local accelerations of the flow. Although the increase in temporal resolution is likely to have positive impacts, we have not examined further the sensitivity of the simulated trajectories to time resolution and thus cannot state whether there is an optimal resolution. To study the sensitivity of the trajectories to horizontal grid spacing, an additional experiment is done with a resolution of $\Delta x = 2$ km in case 2. We use a nested domain of dimension $151 \times 151$, centered on the launch site. One-way nesting is used. The number of vertical levels remains unchanged, and the model physics is the same as for the 10-km simulations with the convective scheme turned off. The 2-km experiment follows the same successive cycles of assimilation as the 10-km grid. For that experiment, assimilation of the radiosonde profiles is done on both the coarse and nested grids. The 2-km experiment produced trajectory forecasts almost indistinguishable from those in the 10-km experiment for case 2 (not shown). Based on these results, we can consider 10 km to be a reasonable choice in terms of horizontal grid spacing.

2) The mesoscale and the global models do not have the same biases and limitations (e.g., the WRF is non-hydrostatic; models have different parameterizations, numerics and/or solvers, though this is likely not of importance for the present case studies). Comparison of the two models is not the purpose and is outside the scope of the present paper. Here we will only emphasize the importance of having a top as high as possible for the prediction of the flow in the stratosphere. Indeed, we used model tops located approximately at 33 km in our preliminary experiences.

![Figure 9](attachment:image9.png)  
*Fig. 9.* As in Fig. 6, but for case 2.

![Figure 10](attachment:image10.png)  
*Fig. 10.* Hodograph for case 2 in the height interval 27–30 km, from the velocity disturbances obtained by filtering out the scales larger than 5 km and shorter than 300 m. Numbers indicate the altitude in kilometers.
(compared to ≤40 km now), causing the ceiling of the OSBs to be in the sponge layer. This proved to be detrimental to the accuracy of the forecast trajectories with an increase of the error by 20% relative to ECMWF for the location of the final position for case 3. Like previously, the 2-km experiment with assimilation of radiosonde profiles did not improve significantly the trajectory forecasts.

However, this benefit of using a mesoscale model to forecast balloon trajectories is not systematic as demonstrated in cases 2 and 3, for which the forecast final landing positions in WRF simulations are only 10–15 km closer to the real landing point than those performed in the ECMWF operational analyses. These cases suggest that the main limitation with the use of WRF is that the initial and boundary conditions provided by the ECMWF are not always sufficient to represent the real atmospheric state accurately (Knudsen et al. 2001). Hence, the use of additional available observational data is justified.

Using standard values for the observational errors, assimilation of only one single radiosonde produces significant local horizontal and vertical corrections on the wind velocity with typical values of ~1 m s\(^{-1}\) or more, and the root-mean-square error is significantly reduced. The resulting simulated trajectories are 10–40 km better for the final position than those calculated in WRF-Ens. Assimilation of the OSB’s ascent also produces significant corrections, and the trajectories are generally improved as we see in cases 1 and 2, although this is not systematic (case 3).

To better identify the trajectory parts that are more prone to forecast errors, we show in Fig. 14 integrations.
of the simulated balloon trajectory that start from the end of the ascent phase and from the end of the ceiling phase. In these integrations, the simulated balloon position is assumed to be that of the real balloon at the initial time of the integration. Figure 14 reveals that most of the error adds up in the ceiling part of the trajectory (left panels). Although this may seem logical if one considers that the ceiling accounts for 50%–70% of the whole trajectory, this may also look paradoxical, as one may expect the stratospheric flow to be dominated by a large-scale balance.

As argued in the introduction, there are several reasons to expect that the stratospheric winds are well known. The circulation is dominated by the large scale, which is close to the geostrophic balance in the extratropics and thus is well constrained by observations of the temperature field. Uncertainties in the midlatitudes have been quantified and are rather weak for the large-scale winds (a standard deviation of the errors on the order of 1 m s⁻¹; Bocca et al. 2008). Such waves typically have vertical wavelengths of 1–4 km (Sato 1994; Guest et al. 2000) and are hence poorly resolved in the ECMWF forecasts. The absence of such waves in the forecast wind field is not a major concern for the ascent phase of an OSB (or for trajectories of long-duration balloons) because the balloon samples the different phases of a wave packet and hence the wave effects cancel out. For the ceiling phase of an OSB however, the problem is very different because the time scale of this phase is typically a fraction of that of inertial waves. OSBs may thus experience only part of the whole wave oscillation, so that the integrated wave-induced displacement is no longer zero. This is in strong contrast with longer-duration balloon flights, for which the effect of wave-induced disturbances on the trajectory should vanish (Keil et al. 2001; Hertzog et al. 2008). For typical inertial waves with horizontal wind disturbances of a few meters per second, such displacement can be of several tens of kilometers and thus it explains (at least partly) the differences between the observed OSB flight and those simulated with ECMWF forecasts. Accurately forecasting such mesoscale components of the flow (i.e., right amplitude and phase) will remain challenging unless significantly more observations are available. The descent phase, on the other hand, which is crucial during balloon operations, greatly benefits from the successive assimilations of available data (Fig. 14, right panel). Indeed, the previously noted limitations associated with the ceiling phase mostly vanish during the descent phase, as the balloon experiences one or several vertical wavelengths of the gravity waves that are present.

**b. On the optimality of the initial perturbations**

In the present study, we have obtained contrasted success in forecasting stratospheric balloon trajectories. This suggests that predictability of such trajectories may vary with the structure of the stratospheric flow. Estimating this predictability is important during balloon
operations, and the spread of the ensemble forecasts may provide information relevant to this issue. In Fig. 15, we have presented the balloon trajectory spread in our WRF simulations. Figure 15 shows that case 1 is the most predictable case: the OSB trajectory is embedded in the DART-all ensemble trajectories and the root-mean-square distance (RMSD) of the final landing point is 12 km. In contrast, the real balloon trajectory is not well captured by the ensemble trajectories in case 2 (Fig. 15b), although the dominant westward component is simulated correctly. The RMSD for this case is 70% larger than for case 1, suggesting a weaker predictability. Case 3 (Fig. 15c) is representative of a bifurcation situation where the spread is almost distributed uniformly in all directions. The RMSD for this case is the highest with a value of approximately 70 km. Moreover, Fig. 16 clearly shows that the error grows mostly at the ceiling. Those results and the fact that several profiles suggest that the corrections were not as effective in the stratosphere as in the troposphere support that the ECMWF ensemble is not sufficient to predict stratospheric balloon trajectories. It is indeed challenging to construct an initial set of perturbations that reasonably samples the probability density function of possible atmospheric states (Snyder and Zhang 2003; Zhang et al. 2004). Methods for quickly generating the perturbations
used for short-term convective/mesoscale prediction, such as additive noise (Tong and Xue 2005; Dowell and Wicker 2009), may be appropriate here. Improving the ensemble, with an emphasis on the spread in the lower stratosphere, is a path for future investigation. The forecast trajectories are also sensitive to the tuning of the EnKF system, such as the localization radius, and the inflation. Also, we have used adaptative covariance inflation described by Anderson (2009), but inflation has given mixed results for convective studies (Snyder and Zhang 2003; Dowell et al. 2004; Tong and Xue 2005). Since the largest portion of the OSBs’ flights remain in the stratosphere, this is not clear how this would impact their trajectories. Further studies would be needed to explore the sensitivity to those parameters.

5. Summary

We have investigated the prediction of trajectories for three open stratospheric balloons during the Strapolété summer campaign in Sweden in 2009. Our purpose was to explore the uncertainty tied to the forecasts of the horizontal winds. Our approach has consisted of isolating this from other uncertainties by assuming a perfect forecast of the balloon altitudes: in other words, we have used the real balloon altitudes as a given time series and calculated trajectories from the simulated wind fields interpolated at these altitudes (this essentially amounts to assuming a perfect flight physics model).

The model used for the simulations was the WRF, initiated and forced on the boundaries by fields from the ECMWF. Improvements relative to the forecasts of trajectories performed with the ECMWF fields can come from two aspects: the finer resolution (spatial and higher output frequency) and the assimilation of additional data (radiosondes launched specifically for the campaign). Unfortunately, only one radiosonde was launched before each OSB flight during the Strapolété campaign. Hence, we have also carried out experiments with the assimilation of the measurements collected during the ascent phase of the OSB as a limit case, to probe how much of an improvement could be obtained with this fresh information for the subsequent phases of the flight. Namely, we have examined the impact of using 1) ensemble simulations with the mesoscale model WRF, 2) assimilation of a single available radiosonde in WRF using DART, and 3) assimilation of both the radiosonde and additional data from the initial ascending phase of the OSB.

Case 1 is illustrative of a very encouraging scenario and underlines the potential of using mesoscale simulations with data assimilation of radiosondes for trajectory forecasts. Using the WRF Model reduced the error in the landing point from 90 down to 30 km. The assimilation of the radiosonde further reduces the error by half. However, success is not systematic in our experiments, and cases 2 and 3 display only modest improvement from the use of WRF and the assimilation of the
radiosonde. These failures help clearly identify the limitations and directions for future efforts:

1) Significant errors were found during the ceiling phase of the flights for simulations with model tops too low (we made test simulations with a model top at approximately 35–37 km, for which the ceiling was partly located in the sponge layer). Use of a deeper domain solved this problem, but it is possible that further improvement will be obtained from pushing the model top yet further.

2) The quality of the forecast will crucially depend on improved knowledge of the initial atmospheric state and hence should benefit from enhanced observations prior to a launch.

One important finding of the present study is that, even with a perfect flight physics (i.e., without errors on the vertical position of the balloon), very large errors are present in the forecast trajectories (the errors can be comparable to the total distance flown). In other words, regardless of the further difficulties from the description of flight physics, the uncertainty originating from the forecast errors in winds constitute a major difficulty for balloon trajectory.

The second major outcome is that error mainly originates from the ceiling phase—that is, when the balloon flies in the lower stratosphere—with limited variations of altitude. This may come as a bit of a surprise, given that lower stratospheric winds are dominated by large-scale, balanced dynamics. However, it is consistent with the poor representation of mesoscale motions in NWP models, in particular inertia–gravity waves, which can have amplitudes of several meters per second. The sampling of such waves during a fraction of the wave period leads to errors of several tens of kilometers in the balloon’s forecast position. This constitutes a significant challenge: having an accurate description of inertia–gravity waves in NWP models requires significant vertical resolution but also very likely significant assimilation from detailed observations. Indeed, the requirement, for trajectory calculations, would be to describe accurately all the characteristics of the wave (amplitude and wavelengths), including in particular the phase. Although studies demonstrating the ability of numerical weather prediction models to simulate a satisfactory gravity wave field (though with underestimated amplitudes) have accumulated in recent years (Wu and Eckermann 2008; Shutts and Vosper 2011; Plougonven et al. 2013; Jewtoukoff et al. 2015), these have mainly focused on time-averaged fields. At present, it is not clear whether having a reliable nowcast of inertia–gravity waves in the lower stratosphere, with an accurate phase, is even feasible. Further and more systematic research will be required for this investigation.

Acknowledgments. This work was initiated while V. Jewtoukoff and R. Plougonven were visiting the Meso-scale and Microscale Meteorology (MMM) division at the National Center for Atmospheric Research (NCAR). Financial support was provided for the two first authors by the NCAR/MMM Visiting Scientist Program. The authors are grateful to Alexis Doerenbecher (Météo-France), Philippe Cocquerez, and Stéphane Louvel (both of CNES) for their helpful comments on this research. V. Jewtoukoff would also like to thank Maria Frediani (University of Connecticut) for the discussions on topics related to this research. The WRF simulations and data assimilation experiments were conducted using NCAR’s Yellowstone and CISL facilities, the NCAR Data Assimilation Research Testbed, and ARW models. The Strapoléité project was funded by the French Agence Nationale de la Recherche (ANR-BLAN08-1-31627, ANR-13-BS06-0011-01), the
REFERENCES


