Nudging technique for scale bridging in air quality/climate atmospheric composition modelling

A. Maurizi¹, F. Russo¹, M. D’Isidoro², and F. Tampieri¹

¹Istituto di Scienze dell’Atmosfera e del Clima, Consiglio Nazionale delle Ricerche, 40129 Bologna, Italy
²Italian National Agency for New Technologies, Energy and Sustainable Economic Development ENEA, Air Pollution Unit, via Martiri di Monte Sole, 4 40129 Bologna, Italy

Correspondence to: A. Maurizi (a.maurizi@isac.cnr.it)

Received: 18 May 2011 – Published in Atmos. Chem. Phys. Discuss.: 20 June 2011
Revised: 31 January 2012 – Accepted: 6 April 2012 – Published: 24 April 2012

Abstract. The interaction between air quality and climate involves dynamical scales that cover a very wide range. Bridging these scales in numerical simulations is fundamental in studies devoted to megacity/hot-spot impacts on larger scales. A technique based on nudging is proposed as a bridging method that can couple different models at different scales.

Here, nudging is used to force low resolution chemical composition models with a run of a high resolution model on a critical area. A one-year numerical experiment focused on the Po Valley hot spot is performed using the BOLCHEM model to assess the method.

The results show that the model response is stable to perturbation induced by the nudging and that, taking the high resolution run as a reference, performances of the nudged run increase with respect to the non-forced run. The effect outside the forcing area depends on transport and is significant in a relevant number of events although it becomes weak on seasonal or yearly basis.

1 Introduction

Processes determining the atmospheric composition cover the whole spectrum of dynamical scales, ranging from the global, relevant for climate and large scale transport episodes, to molecular, where dissipation of energy takes place along with the basic transformation processes (chemical reactions and removal).

Studying the interaction between climate and anthropogenic activities, especially those concentrated in megacities/hot-spots, requires the description of a wide variety of scales, from those typical of anthropogenic emissions, to the global scale. The simultaneous explicit description of these processes is limited by computer resources. Therefore, the need to allow different scales to interact calls for a variety of approaches to bridge them. This connection is already well established in the direction of global-to-regional (or regional-to-local), since it is the result of using output from large scale models to feed boundary conditions into limited area models. This one-way nesting approach has long been used in atmospheric modelling including meteorological applications, operational weather forecast and, more recently, atmospheric composition studies and forecast applications like the GEMS/MACC projects (e.g. Huijnen et al., 2010). This coupling is implemented between different models (i.e. when boundary conditions are provided by an external model) or within the same model when nesting (one-way) is needed for zooming purposes.

The exchange of information in the reverse direction is highly problematic though of great interest. In fact, uncertainties in large scale simulations critically depend on the contribution from the most polluted areas. Here the highly inhomogeneous distribution of sources and the intrinsic non-linearity of the processes involved may give rise to significant differences between coarse and fine resolution simulations.

Different methods have been used so far to implement high resolution information on specific location resulting from detailed anthropogenic emission inventories and refined resolution simulations. In the two-way nesting approach the “parent” run domain includes one or more sub-domains where the same model is run (children processes) at a resolution finer
than the parent. The fine scale simulation communicates with the coarser one taking the parent results at the children sub-domains boundaries as boundary conditions, and feeding back the parent with high resolution fields resulting from the children simulations. Two-way nesting has been used to bridge very different scales, from global (Krol et al., 2005) to urban (Wolke et al., 2008). The benefit of two-way nesting in atmospheric composition applications is not completely assessed (Misenis and Zhang, 2010) although some studies have shown more accurate predictions of chemical species (Gego et al., 2005; Fast et al., 2006).

Other models adopt stretched grids (Struzewska and Kaminski, 2008; de Meij et al., 2009), where a single grid with variable horizontal resolution permits focusing on a selected area. This approach has the advantage of no duplicate runs with different resolutions and the information can propagate in a natural way across the scales. On the other hand, with this approach refining resolution in the area of interest also implies refinement in other areas unnecessarily increasing the computational cost. Using it to focus on different spots simultaneously can become very expensive.

Both approaches, two-way nesting and grid stretching, have common problems when seen from the present perspective. They are designed to run one model at different resolutions. However when looking at the air quality/climate interactions, processes accounting for different scales are so different that they cannot be dealt with using one model only.

Coupling different models can be performed by forcing a low resolution run with results from a high resolution simulation using a data assimilation approach. Here, we adopt nudging as the simplest technique of performing the bridging. The most interesting feature of this technique is that it is suitable, for example, to force global models using regional or local models, without a direct dynamic two-way interaction.

In the analysis presented here it is assumed that the high resolution run provides results nearest to the truth. Although this is not always straightforward for the state-of-the-art models and typical resolutions (Valari and Menut, 2008), this assumption relies on the consideration that models that can deal with fine scale are typically calibrated to perform well in areas with strong emissions at fine scale resolution. Moreover, improvements with resolution are expected to become more relevant in the future as a result of research efforts to improve processes parameterisations and emissions description.

The aim of this paper is to evaluate the effectiveness of such approach. An experiment will be performed using the BOLCHEM model (Mircea et al., 2008) at two different resolutions to simulate a two-model configuration. It must be kept in mind that the real application to the two-model configuration would also present the difficulty of exchanging information between two different chemical mechanisms and two possibly different vertical coordinate systems. Using the same model at two resolutions allows us to isolate the effect of exchanging information, highlighting the effect of non-linearities only.

2 Bridging the scales: the nudging technique

The basic idea of the nudging technique is to force the model run towards available measurements (the “truth”) with some relaxation term in the governing equations. The nudging method is widely used in many fields of atmospheric modelling (Haase et al., 2000; Davolio and Buzzi, 2004) and air quality simulations (Kim et al., 2010) to assimilate data into model runs. In the present application of the nudging concept, the concentration fields in a given region of a low resolution (LRn) run are forced towards the values obtained from a high resolution (HR) run which uses a non-forced low resolution (LR) run as boundary conditions.

The technique consists of adding a term to the concentration tendency equation in the low resolution model. This term acts to force the computed concentrations towards the high resolution concentrations obtained from a fine scale model:

$$\frac{dC_{LRn}}{dt} = \frac{1}{\tau} [C_{HR} - C_{LRn}]$$

(1)

where $C_{HR}$ is the concentration obtained from the HR run remapped onto the LR grid, $\tau$ is the relaxation time, which regulates how fast the actual fields are relaxed to the high resolution fields. For integration time step $\Delta t \ll \tau$, the numerical solution is:

$$C_{LRn}(x,y,z,t+\Delta t) = C_{LRn}(x,y,z,t) + \frac{\Delta t}{\tau} [C_{HR}(x,y,z,t) - C_{LRn}(x,y,z,t)]$$

(2)

Alternatively, the full exponential solution of Eq. (1) is used if the above condition is not fulfilled.

Nudging is applied at each point of a given domain of the LRn run. The forcing is determined by the value of the HR run remapped onto the LR grid. This makes the forcing grid coincide with that of the forced run implying no need of any weighting function, in contrast to what is usually assumed when applying nudging to station data. Actually, this corresponds to a “top-hat” weighting function in for each point of the HR remapped.

The rationale of this approach lies in the fact that non-linearities in dynamics and chemistry lead to concentration fields, obtained from the HR run and averaged over the LR grid, which differ significantly from those directly obtained from the LR run. Furthermore, the actual forcing term, HRr, is obtained from HR by remapping the HR results onto the LR grid. It is known that HRr is likely to perform better than HR when compared to data (Valari and Menut, 2008) because HR runs are more prone to inaccuracies in the wind direction when dealing with the plume-receptor problem. This is...
known in meteorological applications as the double-penalty problem (Mass et al., 2002).

In the present paper the nudging technique is applied to concentration fields only. The same kind of forcing could in principle be applied to the meteorological fields as well. This requires adding the nudging term in the dynamical equations. However, this problem is far more complex because it would produce an unbalance in the dynamical fields that could require damping for the sake of numerical stability. This would complicate the picture and make results less clear. On the other hand, data assimilation in dynamical systems is a complex and evolving problem that is beyond the scope of the present study. Because of the extremely complex picture that could arise as a consequence of forcing meteorology, this option is not considered here.

3 Experimental setup

The nudging procedure is applied to couple two BOLCHEM runs at two different resolutions over two different domains. Three different types of model runs were performed for the whole year 2005:

1. the non-forced LR run (0.5° × 0.5°, Δt = 400 s) over Europe (6°W–44°E, 30°N–57°N), driven by boundary conditions from ECMWF for meteorology and by climatological values for chemistry. The anthropogenic emissions prepared by INERIS for the CityZen project http://www.cityzen-project.eu/ are used.

2. the HR run (0.1° × 0.1°, Δt = 90 s) that covers the Po Valley (7.95°E–12.83°E, 44.4°N–45.98°N) driven by boundary conditions (both for meteorology and chemistry) taken from LR run; and

3. several LRn runs using different parameters and configurations, over the same domain and with the same resolution as the LR run applying the nudging in the above defined Po Valley region to all model species at all levels.

For all the runs, the number of levels used is 33 for meteorology, up to σ = 0.01 and 16 for chemistry, up to σ = 0.5, where σ is the ratio of the air pressure at the considered level divided by the pressure at the surface. First vertical level is at about 20 m (40 m thickness), stretched according to a cubic law.

The forcing term HRt is obtained by remapping the HR results over the LR grid conserving the mass according to Jones (1999). Given that the number of levels is the same for LR and HR, the mass of each species remapped is conserved provided that for each LR grid cell the average thickness of the grid does not differ too much from the LR thickness. This particular procedure can be applied here because we are dealing with the same model and we use the same levels for different horizontal resolutions. Given the methodological nature of the present study, it is intended only to exclude spurious effects and cannot be kept valid for more general applications.

The relaxation time τ in Eq. (1) was set to 1200 s, so that in steady conditions and for a passive tracer the LR run would converge to the HR run before the updated forcing value was used. This relaxation time is used as a reference and is varied for sensitivity experiments as described in the next section. The value of CHRt is updated every hour.

4 Results

The simulations were performed according to the configuration described in the previous section. The time interval at which boundary conditions are updated was fixed to one hour. Then different LRn runs were performed in different configurations for sensitivity analysis.

Model results were analysed in different regions of the domain: the core region (frame 0), where the forcing is applied, plus three additional frames surrounding the core:

1. frame 1 (region limited internally by the core bounds and externally by the box: 6.7°E–14.3°E, 43.3°N–47.0°N);

2. frame 2 (region limited internally by the frame 1 bounds and externally by the box: 6.0°E–15.1°E, 42.8°N–47.5°N);

3. frame 3 (region limited internally by the frame 2 bounds and externally by the box: 5.4°E–15.8°E, 42.2°N–47.9°N).

Frames were selected to keep the number of model grid points as uniform as possible across them.

The model orography in the focus area is presented in Fig. 1 along with a graphical representation of the different areas involved in the model analysis. Frame 3 extends up to the central Italy to the south and to Marseilles on the west segment. Figure 2 shows in the same area an example of the HR run and the forcing term CHRt − CLR for O3 on 2005-06-17, 12:00 UT at the surface. This highlights the fine resolution used in the area and the strength of the forcing.

Seasonal statistics were computed for ground level concentrations (first model level) of carbon monoxide (CO), nitrogen dioxide (NO2), ozone (O3) and particulate matter >10 µg (PM10) in order to verify the stability and quantify the strength of the effect of nudging and its spread outside the forcing area. The standard deviations of the differences CLRn−CLR for winter and summer for the four species are reported in Fig. 3. The effect of nudging is evident in the core region for the four species with some difference and with some seasonality. Some spread is observed in external frames, with PM10 qualitatively showing an influence outside

www.atmos-chem-phys.net/12/3677/2012/
the core that is larger than any other species. Looking at seasonality, it can be observed the strongest differences between summer and winter for CO and O$_3$. This corresponds to the seasonality in concentration: O$_3$ concentration is larger in summer due to photochemical processes while CO concentration can be large at the surface level in winter due to the reduced boundary layer height. NO$_2$ is a mixture of both because it is partly emitted like CO and partly a product of photochemical reactions. PM$_{10}$ is observed to present high values in both summer and winter with a relevant influence outside the forcing region also in the season where it presents the largest concentration.

These results are summarised and quantified in Fig. 4 that shows the Taylor diagrams computed using the whole year data, and for the 4 different regions. The Taylor diagram (Taylor, 2001) provides concise two-dimensional plots of statistical properties which show how well the simulated patterns match a reference. In this case in Fig. 4, the comparison between low resolution simulations (with and without nudging) and the reference (HR run) is performed. The axes refer to the standard deviation of the LR run normalised to the standard deviation of the reference. The correlation coefficient between a given field and the reference is represented by its azimuthal position, while the centred pattern root mean square difference between a run and the reference (in units of standard deviation) is proportional to their distance. Therefore each point plotted on the diagram defines how close a simulation is to the reference in terms of the above statistical properties.

The Taylor diagrams on Fig. 4 show a clear improvement within the forcing region that demonstrates the stability of the system to perturbations introduced by the forcing term. The improvement progressively decreases, while remaining positive, as we consider outer regions and is similar for the four species. It can be observed that in frames 2 and 3, LR performs better than in the core. This can be rather fortuitous. However, it is possible to give an interpretation considering that the majority of emissions in the area are concentrated in the Po Valley, i.e., within the core and that the Po Valley is a semi-enclosed basin. It can be argued that most of the pollutants, on the average, remain bounded in this region making concentrations there much higher than outside. Small absolute errors are likely to be associated with small concentrations in external frames. It is evident that LR concentration of O$_3$ do not seem to significantly change its behaviour from frame to frame. Being O$_3$ a secondary and long lived (when not destroyed by titration in presence of high NO levels) specie and, therefore, less connected to local emissions it is more likely to be exported making the behaviour in the different frames less variable. Looking at the improvements introduced by nudging, no noticeable differences among species can be observed besides PM$_{10}$ always showing the largest improvement.

Improvements can be quantitatively defined using the centred pattern RMS difference (Taylor, 2001) of LR with respect to HR (say $A$) and the same quantity for LRn ($B$). An Index of Improvement (IM) can be defined as: $IM = (A - B)/A \cdot 100$. For example in Fig. 4a for CO we have $A \sim 0.8$ $B \sim 0.2$, therefore we have $IM = 75$, which corresponds to an improvement of 75 %. Figure 5 shows the frequency functions for daily values of IM, for the different species. It can be observed that the most frequent value of IM in the core region lies between 60 and 70. We can also see that in the outer frames the most frequent values lie between 0 and 10 (as for CO) with a long positive tail (in particular for frame 1). This means that on average the effect in the outer frames is small, but in a number of cases can be as large as 40 (PM$_{10}$ and CO).

To illustrate the improvement due to nudging in the outer frames, we selected one of the cases that contribute to the long tails observed in Fig. 5. From the examination of the time series of daily averaged IM, the case study of 2005-06-17 has been selected, in which large values of IM occur in outer frames. The values of IM for the different species for this case study are reported in Fig. 6 as function of the frame number assumed as an indicator of the distance from the core. The figure shows, especially in the case of PM$_{10}$ but also for CO and O$_3$, that the improvement is large in the core and propagates significantly up to frame 3.

Advection is the main mechanism responsible for the spread of the information from the core to outer frames. In order to assess this aspect, the improvement in 4 different sectors of each frame: north, east, south and west sectors, was analysed. Figure 7 shows four panels representing the decay of IM with the distance in each of the sectors. We notice that the largest value of IM in the outer frames occur in the south and west sectors, suggesting a strong influence of transport in the propagation of the effect of nudging. Figure 8...
illustrates the occurrence of transport towards the Tyrrenian sea by representing the average differences $C_{\text{LRn}} - C_{\text{LR}}$ for PM$_{10}$ and the wind field at 850 hPa for the present case study. It is worth noting that the wind at 10 m (not reported) is consistent with the picture given at 850 hPa. The propagation of the influence of nudging is clearly correlated with the wind field. This explains partially the reduced effects on longer time periods, since the Po Valley is frequently characterised by low winds conditions. Windy regions could present a larger influence of nudging outside the forcing area.

4.1 Sensitivity tests

The influence of $\tau$ on the nudging procedure was tested by comparing IM in the different frames and for different species for different $\tau$. It is expected that the effects vanish as $\tau \to \infty$. Selected values, other than 1200 s, are: 0, 2400 s, 3600 s. The statistics were computed for the case study, to highlight effects in outer regions. It is found that from $\tau = 0$ s to $\tau = 3600$ s the differences in IM are very small in the core (less than 10%) and negligible in the frames. In order to test the consistency of the model, $\tau = 360000$ s was also tested and it is found that as expected the IM values are really small also in the core (less than 10%). It can be concluded that the influence of $\tau$ on IM is strongly non-linear with very small variations in the range of values between the $\Delta \tau$ and the nudging forcing upgrade interval (kept to one hour in this study). The upgrade interval could also be changed to test sensitivity. However, the upgrade of the forcing that requires exchange information between two runs can be considered similar to the boundary condition information exchange and it is not expected to change very much with respect to one hour.
Further sensitivity tests were performed in order to evaluate the effect of dealing with species in different ways. This point is very important when applying the technique to couple different models with different chemical schemes. In the frame of the present study where only one model is available, different criteria of forcing the sets of species were tested.

The first test deals with forcing only the emitted species as they are common to any chemical scheme. The species forced are: CO, NH$_3$, NO, NO$_2$, SO$_2$, HCHO, CCHO, RCHO, MEK, ETHE, OLE1, OLE2, ALK1 of the SAPRC90 chemical mechanism (Carter, 1990). Table 1 compares values of IM in the core to the reference case for O$_3$, a purely secondary compound and for PM$_{10}$, a mixture of primary (emitted) and secondary components. It is observed that the effect becomes negligible for O$_3$ even in the core. Also, effects on PM$_{10}$ are reduced, but remain important. This difference is explained by the above described nature of the two observed species.

Another test on species manipulation was performed simulating the exchange of grouped species like Volatile Organic Compounds (VOC) and PM$_{10}$. Rather than exchanging the concentration of single components of those two groups from HRr to LRn, the total mass of VOC and PM$_{10}$ is exchanged and then distributed to the single components of LRn run proportional to the ratio between the concentration of the single component and the total mass in LRn. The results (not reported here) show a very small departure from the reference case which means that in the present configuration, exchanging the single species or their total mass, with the mechanism described above, do not produce relevant change. It must be kept in mind that the present configuration with one model only is not well suited for this test because exactly the same
Fig. 4. Taylor diagrams of LR run vs. HR run ... of LR with re-spect to HR (sayA) and the same quantity for LRn (B). An Index of Improvement (IM) can be defined as: IM=

concentrations in external frames. It is evident that LR concentrations in external frames. It is evident that LR

gion making concentrations there much higher than outside. 
gion making concentrations there much higher than outside.

the pollutants, on the average, remain bounded in this re-

Fig. 5. Frequency distribution of the Index of Improvement for each species and for each frame. The dataset shown here includes the entire year of data.

Fig. 6. Daily IM as a function of the frame index for the four selected compounds, in the case study 2005-06-17. The Frame index 0 refers to the core region.

5 Conclusions

A new method for allowing different models with different spatial resolutions to interact in two way has been investigated. The method is based on the usual boundary condition exchange to provide forcing in one way and on the nudging technique for coupling from high- to low-resolution. To test the method, numerical experiments were performed using the BOLCHEM model at different spatial resolutions, to simulate coupling of different models. The focus of the experiments was on the Po Valley hot-spot.

Results show a large improvement in the forcing region, highlighting the stability of the method. As far as the influence outside the forcing region is concerned, small values of IM are observed averaging over yearly or even seasonal time periods. Nevertheless, it is observed that a relevant fraction of cases present fairly large improvement outside the forcing area.

Analysing one test case has shown that influence of nudging can extend as far as 200 km with an improvement of 10 to 20 % for some species due to advection. It can be argued that in regions that are more windy (on yearly or seasonal basis) than the Po Valley, the spread would become relevant also at these time scales.

Based on the test case, some sensitivity studies were also performed and they shown that the method is quite

component of grouped VOC and PM$_{10}$ are present in both HR and LRn runs. However this case is important as it defines a method for exchanging information between really different chemical mechanisms.
Maurizi et al.: Bridging the scales in atmospheric composition modelling

Other test cases were performed to simulate the problem of coupling different models with different chemical mechanisms. Results show that forcing emitted species only does not produce much improvement. However, a method is presented for coupling different schemes based on the exchange of information between high- and low-resolution runs of only the grouped VOC and PM$_{10}$ species, redistributed according to the actual chemical mechanism. Although the results do not differ much from the reference case, because here a unique model is used, this constitutes the basis for testing the method in the coupling of actually different models.

The main limitation arising from the present study is that the influence outside the forcing region, though very strong in a relevant number of cases, seems to be weak on long time scale average (season to year). However, as this effect is largely related to advection, it can be argued that the Po Valley represent a restrictive test for the method. A more windy area (e.g., the BeNeLux) is expected to display larger effects.

The application to the “real world case” could be the coupling of a global model, with resolutions of the order of 1° and a regional model, focused on hot spots with resolution up to 0.1°, possibly with different chemical schemes and also different vertical coordinate systems. This last point would require a special attention in the identification of dynamically equivalent layers, in particular the boundary layer, taking care of exchanging information not by a simple vertical

Fig. 7. Daily statistics for the case study 2005-06-17. From left to right, from top to bottom: North (a), South (b), East (c) and West (d) sides of the three frames (value for core is unique).

Fig. 8. Contour plot of the average difference of C$_{LRn}$−C$_{LR}$ for PM$_{10}$ for the case study 2011-06-17 at 07:00 UTC and the wind field at 850 hPa.

insensitive to the relaxation time of the forcing term, at least in a range of values comparable to the time step and to the forcing upgrade time interval. It is necessary to increase the relaxation time by two orders of magnitude to obtain an almost complete reduction of the effects.
interpolation but matching the equivalent layers between the two models. An additional problem that should be investigated is how to deal with different stability conditions of the boundary layer as seen by the low- and high-resolution runs.

Acknowledgements. The research leading to these results was funded by the European Union’s Seventh Framework Programme (FP 2007-2013) within the project CityZEN (megaCITY – Zoom for the Environment) and the project MACC (Monitoring Atmospheric Composition and Climate). Most of the software used for the production of this article is free software. The authors would like to thank all the free software community and in particular the Debian Project (http://www.debian.org).

References


