

Interactive comment on “Reaching the lower stratosphere: validating an extended vertical grid for COSMO” by J. Eckstein et al.

Anonymous Referee #1

Received and published: 19 February 2015

GENERAL COMMENTS

The paper describes an extension into the lower stratosphere for the COSMO model and shows two boundary forced limited-area simulations for a selected European area, with particular focus on simulating the lower stratospheric region in northern latitudes. One simulation used ERA-Interim, the other used NCEP re-analysis data. The vertical extent is increased from 22km to 33km. The starting height of the newly designed sponge layer is chosen to be at 28km. Typically a large sponge starting just above 11km is used in standard COSMO simulations and hence does not allow to simulate the lower stratosphere. Simulation results are presented for a 11-months simulation. The simulated fields are compared to regridded boundary forcing data and to radiosonde observations, revealing biases in both. The results are shown to be strongly forced

C51

by the boundary data of the domain. The authors claim added value with an example of orographically induced lee waves that are absent in the lower resolution global boundary data.

The paper is concise and clear but I don't think so far any of the authors' claims are presented beyond doubt.

a) It is pointed out that COSMO runs stably for the entire 11 months period suggesting that it is a suitable model for investigating the lower stratosphere. If the study really would address the stability and accuracy of the COSMO model as opposed to the accuracy of the boundary forcing data, one would need to show for example the sensitivity to different coupling intervals, e.g. the boundary forcing is used every 6 hourly, how do the results change if it is used only every 12 hours? For the scatter plots at which output times are the modeled values compared with observed values? What is the actual variability of the temperature and humidity fields compared to the boundary forcing within the 6 hourly periods?

b) The new sponge is insufficiently tested and described. It is smaller in its vertical extent than the previous one. Why? Moreover, little is said about what the sponge layer does. There are three aspects to a damping layer, its vertical extent, its scale-selective/or not absorbing mechanism, and its time-scale. How active is this layer really? How much does the boundary forcing matter? There are no cross sections in this paper of for example gravity wave momentum flux or vertical velocity to judge any of these aspects. How sure are you that there is no artificial reflection of this sponge or that this sponge has any impact at all? In global models, even slight changes to the sponge have dramatic impacts on the overall circulation. I assume that this is alleviated by the strong boundary forcing and the relatively small domain but this is not said.

c) The paper claims added value by showing a case study with strong gravity waves signatures absent in the boundary forcing that are interpreted as orographically induced lee waves. Again, a vertical cross section of vertical velocity and gravity wave

C52

momentum fluxes would allow to judge on these. How sure are you that these are orographically induced ? One should also show a comparison of the forcing orography in a line along the wind direction (E-W), comparing ERA-Interim and COSMO. How different are the actual slopes ? Moreover, the gravity waves shown have a large horizontal extent (Fig.17). A quick look at NOAA 19 satellite pictures from that date do not reveal any such structures but they do show some (much) smaller-scale gravity-wave activity. So if the gravity waves seen in Fig.17 are matching the scales selected by the COSMO orography filtering/resolution, and do not represent reality, how can the claim be upheld that it provides added value ? Afterall, the lower resolution forcing data equally only shows what it can realistically represent and the effect of gravity waves are parameterized at these scales. The problem with much higher resolutions such as used in COSMO, however, is that these may show resolved gravity waves but at the wrong scales with unrealistic momentum fluxes. Unless the authors can believably make the case that the waves seen in the simulation are a better representation of the truth, I suggest to revise the conclusions. Notably, I do not disagree in principle that a limited-area model could be a useful tool to discover the UTLS region, but the case has not been made here.

d) Not much is said about model physics in the extended region. Which physical parameterizations operate in the extended region and how ? Is this perhaps a reason why the tropospheric variability is larger than the lower stratospheric one ? In this context, how is advection effecting the humidity bias for example across the tropopause ? Conservation of moist species is particularly sensitive in these regions due to sharp vertical gradients and small absolute values. What is the “added value” of COSMO in this context ?

e) The paper does show a lower stratospheric / upper tropospheric humidity bias comparing model and observation, the bias appears to be already in the forcing data and/or in the observations. There is also a known dry bias for radiosondes which could explain the differences found for stations 11 and 23 ? The 4 percent at least seem to match the

C53

findings in Wang et al 2013 (Radiation Dry Bias Correction of Vaisala RS92 Humidity Data and Its Impacts on Historical Radiosonde Data) in terms of amplitude and sign.

SOME DETAILED COMMENTS

para 2.3, T42 (quadratic grid, ?full?) corresponds to 2.8 degrees (128 points on a latitude), T255 (linear grid, reduced) corresponds to 0.7 degrees (512 points per latitude, reducing towards the pole to keep average distance constant), this should be perhaps spelled out given the focus on higher latitudes.

For the US standard atmosphere 2.7hPa is ~40km, this could be clarified, e.g. in the context of an assumed colder basic state ?

para 2.4 The way the sentence is phrased “model runs stably for eleven months” implies that it aborted after due to stability. Perhaps “eleven months” can be deleted here as it is later explained why 11 months.

before 8076 “of of”

para 4.1 explain what you mean with “free running level”

para 4.2 In the context of the main comments above “the model is able to simulate mean temperature well in all heights” seems a premature statement to me. I think the main point is there is no difference to the regridded boundary forcing data.

para 4.2 “the variability in higher altitudes is lower”, so could it be that the boundary forcing is “less disturbed” by COSMO model physics ?

para 4.3.2 “only larger scale fluctuations ... can be captured by the model”, does this point towards inadequate physics and/or resolution and or dynamics in the lower stratosphere and/or too strong boundary forcing ?

para 4.3.2 which radiosondes do these stations use ? See comment e) above.

para 5, p. 494 The case has not been made sufficiently that these indeed are oro-

C54

graphically induced lee waves and if these are only model lee waves as opposed to observed lee waves.

para 5, "COSMO is almost four times as good" implies that increasing resolution is always good. I would say "The COSMO simulations used 3.5 times higher horizontal resolution compared to ERA-Interim" and back this up with a comparison plot of resolved orographic slopes.

para 6, "The measurements of temperatures are well reproduced by the model for all stations and heights" This should be qualified further in the context of the quality of the boundary forcing and interval. Likewise for the sentence "The error in heights above 11km is even smaller ..." considering my previous comments.

Figs 7+12 The scatter plot data used needs to be explained further.

Figs 6 + 13 It is difficult to see details. Perhaps a enlarged smaller time interval would help to see variability within a series of a few forcing intervals ?

Interactive comment on Geosci. Model Dev. Discuss., 8, 483, 2015.