EUCLIPSE

EU Cloud Intercomparison, Process Study & Evaluation Project

Grant agreement no. 244067

Deliverable D4.3 Report on a study identifying the utility of NWP based methods for identifying and narrowing sources of divergent behaviour in cloud-climate feedbacks in ESMs.

Responsible partner: ECMWF

Delivery date: 36 months

Report on a study identifying the utility of NWP based methods for identifying and narrowing sources of divergent behaviour in cloud-climate feedbacks in ESMs

M.J. Rodwell, ECMWF, 3 July 2014

This study focuses on the use of Numerical Weather Prediction (NWP) to assess climate models. In particular, the assessment of errors in the parameterisation of "fast" aspects of the physics; which have an impact on weather forecasts and also represent a major source of uncertainty in climate sensitivity. The study is in two parts. The first is an assessment of diagnostic techniques themselves; to understand their strengths and weaknesses in the evaluation of climate model errors. The second part uses one of these techniques to evaluate various formulations of the parameterisation of shallow convection.

1. Comparison of diagnostic methods

In Numerical Weather Prediction, the forecast is routinely assessed against verifying observations and this provides a powerful framework for investigating model errors. When attempting to predict the climate decades into the future, such verifying observations are obviously not available. However, a lot of the uncertainties in climate models are associated with 'fast' aspects of the physics of the atmosphere (convection, clouds, etc.) which are also important in NWP. This has led previous studies to suggest that a lot can be learned about climate models by running them in 'NWP mode'. Here we compare a couple of the most promising NWP approaches to climate model assessment.

The first (and most widely feasible) NWP approach is to simply initialise a climate model with a set of initial conditions (analyses) produced by one of the world's leading NWP centres, and to inspect the systematic errors that evolve. This approach is sometimes called the "Transpose-AMIP" method (Philips et al., 2004). At forecast lead-times of a few days or more, this approach can usefully assess the climate of the model. While it is important to recognise this benefit of the Transpose-AMIP method, this study is primarily concerned with a somewhat different issue; that of diagnosing deficiencies and sensitivities within individual parameterised processes (such as stratocumulus clouds for example) that are the root-causes for model climate deficiencies. At such lead-times (of a few days or more), interactions that take place within the atmosphere (between planetary waves and convection in the tropics for example) mean that it can be difficult to identify the root-causes themselves. In addition, these interactions give rise to chaos which generally leads to reduced statistical significance of any results - necessitating longer experimental periods. By examining shorter lead-times, one reduces the time over-which interactions can obscure the underlying model problems, but a second diagnostic barrier can arise. This barrier is associated with the fact that the analyses themselves are not perfect, and so the systematic 'errors' one obtains will actually conflate two sets of errors - those in the analyses and those in the climate model. At short lead-times (less than a few days), forecast errors are smaller and so these analysis errors cannot necessarily be neglected. The ability, or otherwise, of the Transpose-AMIP approach to identify model deficiencies thus lies in our ability to find a forecast lead-time sufficiently long to be able to neglect analysis error, but sufficiently short to avoid complicating interactions and loss of significance. This search has been one component of this work topic.

The second NWP approach to model assessment aims to implicitly account for the errors in the analysis, so that weather forecasts made with the climate model can be assessed at very short lead-times - thus minimising the complicating effect of interactions. To understand how this approach can account for the errors in the analysis, one needs to understand how the analyses are produced within the forecast centre's data assimilation system. While data assimilation systems vary in many details, a key aspect remains the same - new observations are 'optimally' combined with a short "background" or "first-guess" forecast (initialised from a previous analysis) in a way that is consistent with estimated errors in the observations and background. A key diagnostic of the data assimilation is the "analysis increment" which is simply the difference between the new analysis and the previous background forecast. The key component to this second diagnostic approach is that the background forecast is actually made with the climate model - including at the climate model's resolution (consistent linearized and adjoint models are also employed in the incremental 4D variational assimilation scheme used at ECMWF). By doing this, the magnitude of the increments will be a good estimate of the inconsistency between the climate model and the observations. Furthermore, the tendencies to the state vector that individual process parameterisations (and the resolved dynamics) produce within the background forecast can be compared with the increments, so that the inconsistencies may point to key parameterisation problems. This approach is sometimes called the "Initial Tendencies" method. It was proposed by Klinker and Sardeshmukh (1992) and refined by Rodwell and Palmer (2007). It relies on the ability to perform data assimilation with the background forecasts made using the climate model. This is perhaps the key limitation of the approach. If such data assimilation is possible, then the approach is generally easier to perform than Transpose-AMIP since it avoids issues associated with interpolation of analyses to the climate model's own grid. In order to assess the ability of the Initial Tendencies method to identify model deficiencies, artificial perturbations have been introduced into the climate model, and the method tasked with identifying these perturbations.

A full account of this comparison of diagnostic methods is presented in Klocke and Rodwell (2014). It aims to address an important yet under-represented topic in the climate literature until now. Here we discuss one figure from this study. Figure 1 shows differences in outgoing long-wave radiation (-ΔOLR) between a set of control forecasts and, for each column, a differently perturbed set of forecasts. The control forecasts use the ECMWF model run at a 'climate resolution' of ~80km in the horizontal, and initialised from "native" analyses that are produced using the same background model. The perturbed forecasts differ from the control forecasts in their initial conditions and/or in the model used to make the forecast. Results are averaged over forecasts initialised at 12UTC each day during April and May 2011. Each row shows a different forecast lead-time.

In the right-hand column in Figure 1, the perturbed forecasts involve a change to the convection scheme (a decrease in convective entrainment coefficient). This change is also used in the model used to produce the initialising analyses. This column therefore represents the Initial Tendency methodology. Statistically significant differences in OLR are apparent in the Inter-Tropical Convergence Zone at lead-times of 6h and 24h. The lower convective entrainment rate leads to higher-reaching convection, an increase in high-cloud cover and consequently to reduced OLR. These differences demonstrate that the Initial Tendency approach can 'find' the signal of the artificially-introduced physics perturbation in a climatologically important radiation field (as well as in the initial tendencies themselves). After 24h, the signal and its significance are lost due to interactions and the growth of chaos.

In the left-hand column in Figure 1, the perturbed forecasts involve a different set of initial conditions (analyses from the UK Met. Office) but use the same ECMWF control model for the forecasts. This experimental set-up is to investigate the "spin-up" issues inherent in the Transpose-AMIP methodology. Higher temperatures in the Met Office analyses lead to strong, and statistically significant, initial cooling of the Earth system via the OLR. This spin-up issue is present for lead-times up to at least 2 days - beyond the lead-time where the Initial Tendencies (right-hand column) were able to identify the model perturbation and thus suggests there is no lead-time at which the Transpose-AMIP methodology would be able to identify the forecast model perturbation. However, the design for these experiments is somewhat crude. For example, we interpolated Met Office initial conditions from only 16 pressure levels, and there are inconsistencies associated with the initialisation of the surface. Although such difficulties are, to some degree, inherent in the transpose-AMIP methodology, it is interesting to see how well the transpose-AMIP approach would perform when interpolation issues are kept to a minimum. This is explored in the final set of perturbed forecasts.

In the middle column in Figure 1, the perturbed forecasts involve the entrainment perturbation to the forecast model, but are initialised from the control analyses. This experimental set-up represents the Transpose-AMIP methodology with interpolation issues minimised as much as possible. Since both sets of forecasts are initialised from the same set of analyses, we might expect the difference in OLR to be small at a lead time of 6h. From the results in the right-hand column, we would also expect the difference in OLR signal to be statistically insignificant after day 2. Both of these predicted results are evident in the middle column. However, even with this near-perfect transpose-AMIP experimental design, we see that the true signal of the model perturbation (i.e. reduced OLR due to higher convection) does not emerge at intermediate lead times (such as at 24 h). Hence, even with the best possible initialisation of Transpose-AMIP style experiments, it does not seem possible to be able to identify our introduced model perturbation. It is possible that Transpose-AMIP style assessments of more strongly differing models might still be possible, but these results highlight the fundamental importance of initialisation and spin-up.

The main conclusions of this study are that, although the Transpose-AMIP methodology can clearly be useful for comparing the overall performance of (climate) models, it may not be able to readily identify the reasons for the difference in this performance (and thus not be able to say much about the formulation of cloud parameterisations for example). The Initial Tendency approach is more able to identify errors in model formulation (but requires the use of a data assimilation system).

There are other, more widely applicable, approaches to climate model assessment which go some way to following the Initial Tendency methodology. For example, the calculation of the "nudging" required to force a climate model to follow a set of analyses produced elsewhere could be assessed. The nudging takes the place of the analysis increments in the Initial Tendency approach (e.g. Jeuken et al., 1996; Mapes and Bacmeister, 2012). It is also worth experimenting with less complex data assimilation systems, which do not involve so many "forward models" (which project model fields onto satellite observations). Simplified approaches such as these may represent a smaller first step which could be sufficient, or may give further impetus towards the development of full data assimilation systems. The difficulty that the development of a full data assimilation system presents should also encourage those who already possess a full data assimilation system to apply it to the climate question.

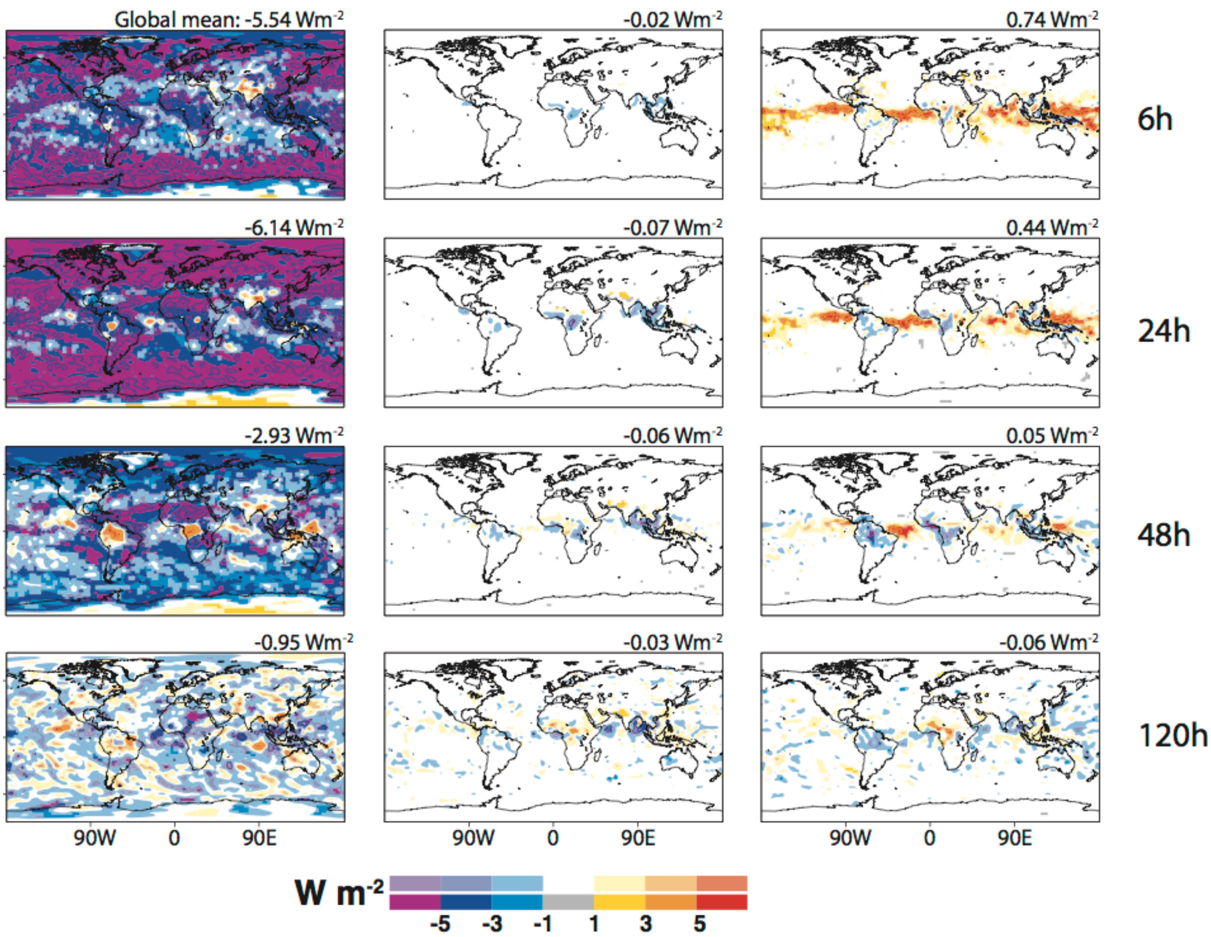

Figure 1. Differences in top-of-atmosphere outgoing long-wave radiation (−ΔOLR) between a perturbed set of forecasts and the control set of forecasts (perturbed minus control). The perturbed forecasts are (left) control model initialised from the UK Met Office analyses, (middle) perturbed model initialised from the control analyses, and (right) perturbed model initialised from its own (native) analyses. The lead-time ranges are (from top to bottom) 0–6 h, 0–24 h, 24–48 h and 96–120 h. Saturated colours indicate statistically significant differences at the 5% level. Note that -ΔOLR is used (rather the +ΔOLR) as positive values then represent a positive change in atmospheric heating.

2. Diagnosing different representations of shallow convection using the Initial Tendencies approach

Two approaches to the representation of shallow convection have been compared. The full comparison is documented in Bechtold et al., (2014), but here we focus on the evaluation using the Initial Tendencies method.

The first approach is that used in the operational ECMWF model. It is discussed in more detail in Bechtold et al. (2014) and references therein. In summary, the planetary boundary layer (PBL) includes the interacting processes of dry diffusion, cumulus mass flux, clouds and radiation. The model distinguishes between a stable PBL, a dry convective PBL and a cloudy PBL. The cloudy boundary layer is further classified into either a well-mixed PBL with stratocumulus clouds, a convective so called ‘decoupled’ layer with cumulus clouds (the decoupling criterion is the inversion strength), and a purely convective PBL with cumulus clouds. One weakness of this approach is the threshold-dependent switching which makes it difficult to enable smooth transitions between regimes. Nevertheless, there are important strengths in the current operational framework; for example it permits a consistent treatment of shallow and deep convection which is important for representing the diurnal cycle.

The second approach takes a different view-point and seeks more consistency within the PBL by integrating the well-mixed and decoupled regimes into a single scheme that includes two ascents: a dry plume stopping at cloud base, and a more buoyant 'moist' parcel that reaches the cloud top. Such an approach, called "DUAL-M", was pursued by Neggers et al. (2009). It showed encouraging results in the ECMWF model, with more realistic trade cumulus cloud structures and lower and more realistic cloud top heights. However, problems with the DUAL-M scheme remained, notably the underestimation of continental shallow clouds, leading to a warm bias over the continents and some lack of stabilisation in non-surface driven convection as encountered in frontal clouds. The scheme is therefore implemented in the ECMWF model (as an option) with the original shallow convection scheme taking-over if parcels are ascending from elevated layers.

The data assimilation and weather forecasts used in this comparison have been performed at a horizontal resolution of ~40 km. For each model configuration, consistent initial conditions are generated using the same model and a 6h assimilation window. Forecasts are initialized from these analyses every 6h (at 00, 06, 12 and 18 UTC) daily between 30 December 2011 and 2 February 2012. Initial tendencies from physical processes (and the dynamics) are accumulated over the first 6h to obtain a more detailed insight into the importance of the individual processes.

Figure 2 shows these initial temperature tendencies averaged over the 4 control forecasts per day for January 2012, together with the corresponding analysis increments and mean analysed evolution over January. Also shown is the zonal-mean analysis temperature. Saturated colours indicate statistically significant values. Heating by convection in the tropical troposphere is largely balanced by dynamical cooling (in the ascending branch of the Hadley Circulation), radiative cooling and net (evaporative) cooling in the cloud scheme. Convective heating is also strong in the wintertime middle latitude storm track regions where it tends to be balanced by radiative and cloud evaporative cooling. The cloud scheme produces net condensational heating in the extratropical mid troposphere, where it is largely balanced by radiative cooling. The zonal-mean radiative tendency (cooling) is rather uniform with latitude, but has a peak cooling rate near 800 hPa in the storm tracks (not evident in the zonal mean plots shown). In the PBL, diffusive heat transport is strong. It is compensated by dynamical cooling and cooling in the sub-convective-cloud-layer through the evaporation of rain.

While these processes are largely in balance with each other over the season, there is a small evolution term (note the smaller contour interval, and the lack of statistical significance) that represents the annual cycle and the average of synoptic variations. The other notable term in the budget is the analysis increment (but again note the reduced contour interval). This is statistically significant and largely reflects a correction for systematic errors in the modelled processes. The tropical zonal-mean increments show a heating/cooling dipole of O(0.06 K/6h = 0.25 K/day). This is related to errors in the zonal-mean Hadley circulation which the model tends to weaken (not shown). Heating increments in the extratropical upper troposphere are thought to result from a drift in humidities (Leroy and Rodwell, 2014) - that are poorly constrained by observations - that then leads to an imbalance between radiative cooling and dynamical warming.

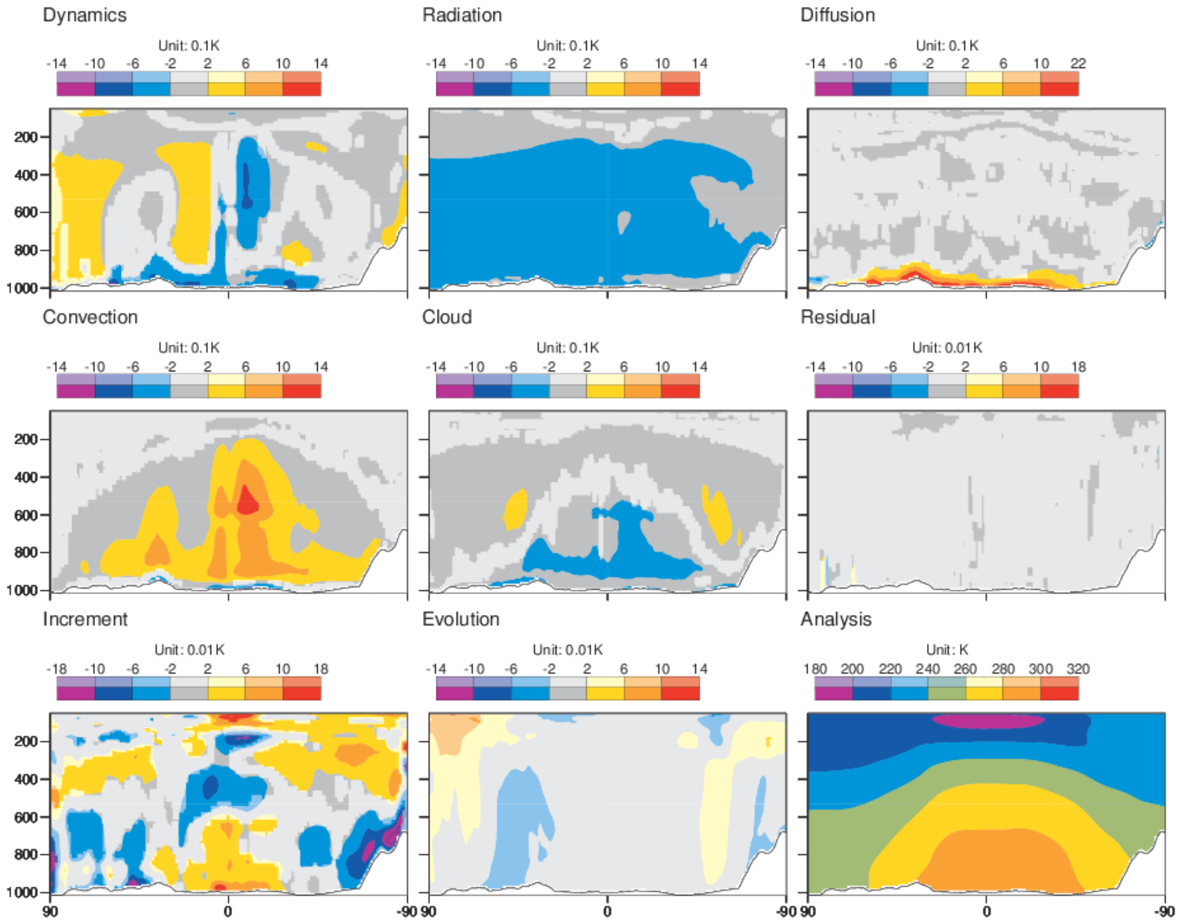

Figure 2. Zonal-mean initial temperature tendencies (accumulated over the first 6h, and averaged over forecasts initialised every 6h during January 2012) together with the mean analysis increment, mean analysed evolution over January and the residual that represents other numerics within the model and closes the budget. Also shown is the zonal-mean analysis temperature. Values significantly different from zero at the 5% level using a Student's T-test are shaded with more saturated colours.

When the DUAL-M scheme used instead of the operational PBL scheme, cloud evaporative cooling is reduced in the PBL and convective heating is also reduced (in compensation). Diffusion heating tendencies strengthen near the surface and there is less radiative cooling at the top of the PBL. Much of these changes (not shown) compensate each other but there is a small change in the overall balance. This is reflected in the change to the zonal-mean analysis increments shown in Figure 3. For the tropics, note that these changes in mean increments act to reduce the increments shown in Figure 2, and thus indicate that the DUAL-M scheme is beneficial in this respect. These and other results demonstrate that the 'dual mass flux' approach is promising. It is thought that this can be realized within the current framework without the inconsistency of involving the additional statistical cloud scheme within the present formulation of DUAL-M.

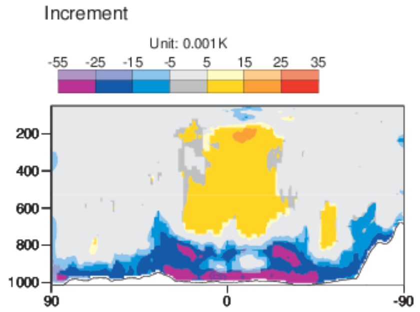

Figure 3. Zonal-mean difference in temperature increments resulting from replacing the operational PBL scheme with the DUAL-M scheme.

3. References

Bechtold P., I. Sandu, D. Klocke, N. Semane, M. Ahlgrimm, A. Beljaars, R. Forbes, M. Rodwell. 2014: The role of shallow convection in ECMWFs integrated forecasting system. *ECMWF Technical Memorandum*, **725**, pp.

Jeuken A.B.M., P.C. Siegmung, L.C. Heijboer, J. Feichter and L. Bengtsson. 1996: On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *J. Geophys. Res.,* **101**, 16939–16950.

Klinker E. and P.D. Sardeshmukh. 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. *J. Atmos. Sci.*, **49**, 608–627.

Klocke D. and M.J. Rodwell. 2014. A comparison of two numerical weather prediction methods for iagnosing fast-physics errors in climate models. *Q. J. R. Meteorol. Soc*. **140,** 517–524.

Leroy S.S. and M.J. Rodwell. 2014: Leveraging highly accurate data in diagnosing errors in atmospheric models. *Bull. Am. Meteorol. Soc.* Early Online Release. Doi: 10.1175/BAMS-D-12-00143.1.

Mapes B.E. and J.T. Bacmeister. 2012. Diagnosis of tropical biases and the MJO from patterns in MERRA's analysis tendency fields. *J. Climate* **25**, 6202–6214.

Neggers, R., 2009: A dual mass flux framework for boundary layer convection. Part II: Clouds. *J. Atmos. Sci.*, **66**, 1488–1506.

Phillips T.J., G.L. Potter, D.L. Williamson, R.T. Cederwall, J.S Boyle, M. Fiorino, J.J. Hnilo, J.G Olson, S. Xie and J.J Yio. 2004. Evaluating parameterizations in general circulation models: climate simulation meets weather prediction. *Bull. Am. Meteorol. Soc.,* **85**, 1903–1915.

Rodwell M.J. and T.N. Palmer. 2007: Using numerical weather prediction to assess climate models. *Q. J. R. Meteorol. Soc.,* **133**, 129–146.