Subject: Scientific visit at ECMWF and University of Oxford (ECMWF/University of Oxford, Reading/Oxford, UK, 29/02/2016-04/03/2016)

From: François Massonnet

To: BSC-ES

Copy: -

Introduction

- Objective: I have been invited for a 4-day scientific visit at ECMWF in the team of Magdalena Balmaseda to discuss, among others: ocean and sea ice data assimilation, the use of new products of sea ice thickness for model evaluation, and the use of the sea ice model LIM3 in the ECMWF seasonal forecast system. I took advantage of my trip to spend the last day of the week at University of Oxford, in the teams of Tim Palmer and Laure Zanna. Topics covered there were essentially related to reproducibility of EC-Earth on different platforms and other computer science issues. I also attended two brilliant seminars by Kerry Emanuel (worldwide expert on hurricanes) who was visiting at the same time.
- Funding: ECMWF
- Agenda: -
- Attendants: me.

Wherever this is possible, I put in [**CAPITAL BOLD]** the names of BSC members that could be interested in the topic discussed.

Feedback on presentations

I gave two presentations on Monday 29 Feb 2016: one about the importance of observational products uncertainty for seasonal forecast verification and one about reproducibility of EC-Earth.

[**OMAR, PACO**]. The talk about observational uncertainty was extremely well received. People like the framework presented and see it as a promising way to perform joint model- and obs-evaluation in one shot. According to the framework introduced, the two analyses cannot be decoupled (because observations and models share a common signal), and that's precisely what people liked. They also see in the study a very neat connection between data assimilation and forecast verification. Indeed, the unifying model used in the study (truth-plus-noise model) is also the one at the heart of data assimilation, used to derive optimal analyses in both variational methods (3/4D VAR) and Kalman filters. An interesting comment came from Tony McNally, an expert in satellite observations. He said that it would be really easy to create an new observation that would fool my analysis. Create for example a new observation equal to the first forecast of my ensemble: this new observation will correlate a lot with all forecasts we have, but still be non-sense since that's not a proper observation. That shows that correlation is not a sufficient metric to claim that some observation is good, but well a necessary one. I also had a long discussion with

experts in observations who explained to me the difference between all datasets used. I have everything clear in mind now and I can start writing up the paper.

[**REPRODUCIBILITY TEAM**]. Regarding reproducibility, I showed our results from last year and our current strategy. One striking figure that we had suggests that Antarctic sea ice extent is much more different on ECMWF's platform than on Ithaca or MareNostrum3. The audience was quite skeptical about the reasons that could explain this. People are thinking about a bug, rather than hardware/software uncertainty, to justify the difference. Indeed, the NEMO code is known to contain several "uninitialized arrays": these are arrays that are declared but not filled with initial data. When processed in the code for the first time, the compiler has to insert some default values. Some compilers put NaN, some put 0, some put -1e24. If that's what happens, then indeed the origin of differences would be rather human- than machine-induced. This is also why we must absolutely re-run the experiment.

Summary of activities at ECMWF

ECMWF has undergone substantial changes in the past weeks. Florence Rabier is now the head of the center since January, and a re-organization of the center's structure has been undertaken to account for Copernicus activities. There are several departments, one of which is "Research". The Research department is made up of several sections, and I had most interactions with the Predictability section and Data Assimilation section.

The ocean reanalyses that we use extensively at BSC (ORAS4, ORAS5, ORAP5) are produced using NEMO. ORAS4 is run at ORCA1, while ORAS5 and ORAP5 are run at ORCA025. ORAS4 does not have a prognostic sea ice model while the two others use LIM2. The data assimilation method is "3D-VAR FGAT", which in short is a variational method where the initial guess is adjusted as to minimize the squared misfit to observations over the assimilation window. All in all, assimilation windows are very short (1 day) compared to what I used to do (15 days), but ECMWF works in an operational framework (I'm more on the climate side). All ocean reanalyses are available from the MyOcean portal.

The atmospheric reanalyses are 4D-VAR, and followa similar approach. The ocean and atmosphere are first run separately and innovations are recorded. Then, a linear system is solved in order to determine the change in initial conditions that would minimize the cost function over the assimilation window. For this, there is a need to compute the adjoint and tangent operators of the models. This is quite a painful work, but when in place, it's quite efficient. (In short, it involves writing two additional routines for every routine of the code). In a second run, the ocean and atmosphere are forwarded in coupled mode and exchange information. A new analysis is made, initial conditions are updated and the model runs a third time coupled. The assimilation is weakly coupled, because no information is transferred from the atmosphere to the ocean apart from physical adjustments during the run.

There is finally a long, century reanalysis that has been produced: ERA-20C. It is already available from ECMWF portal. In addition, a century-long *coupled* reanalysis is currently taking place. The reanalysis is conducted by chunks of 10-years, in parallel, otherwise it would be too long to do. Expected time of completion is 12 months.

ECMWF is finally about to upgrade it's operational forecast system for better resolution. The new forecasts are currently being verified and the whole system should switch to higher resolution rather soon (a matter of months).

Topics discussed with members of Predictability and Data assimilation teams

[**OMAR, NEVEN**] Steffen Tietsche. Steffen currently works on the use of SMOS thin ice data. Given temperature, salinity and thickness, sea ice emits or scatters long waves in a certain part of the spectrum. The algorithm uses prior information from temperature and salinity, obtained from reanalyses, to infer what sea ice thickness must be in order to match the wavelength measured by SMOS. The data is rather uncertain because temperature and salinity are themselves not well constrained. Furthermore, the algorithm is no longer sensitive at above 1 m of ice thickness. This data is publicly available, and now assessed by Steffen. We could think to use it for model assessment and data assimilation.

[**ELEFTHERIA, VIRGINIE**]. Sarah Keeley. The ECMWF team is willing to use LIM3 in the forecast system, but for this some patches have to be fixed. One recurrent question is how to start a LIM3 simulation from a LIM2 restart. I faced similar problems during my PhD so I was able to give her some advices and idea. Sarah has now the routines I gave her and will give them a try soon. More generally, if we want to restart ORAP5 from LIM2 sea ice initial conditions but with LIM3, this is also possible and I can help people here.

[**VIRGINIE, ELEFTHERIA, CHLOÉ, VALENTINA**]. Kristian Mogensen. Kristian is in charge of running the coupled model and interacting with EC-Earth. He now has our namelists, but apparently they use different configuration files from ours. This is because they are synced with the NOCS (Southampton) for ORCA1 and with DRAKKAR (Grenoble) for ORCA025. Kristian is now aware of our results at ORCA1 and ORCA025, but they don’t seem to be willing to change their configuration.

Magdalena Balmaseda. Magdalena was wondering about the role of sea ice salinity in the global ocean circulation. She made some tests by varying the parameter controlling brine rejection during ice formation, and saw a very large sensitivity of the AMOC. She would be interested in performing more sensitivity experiments, perhaps with the new NEMO3.6.

[**ELEFTHERIA, VIRGINIE, CHLOÉ, VALENTINA**]. Hao Zuo. Hao is the data assimilation guy of the team. He is maintaining the ORA reanalyses. We don't have to hesitate to contact him whenever we have a question on the ORA suite. The ORA will not use the new bathymetry files released by SHACONEMO to ensure continuity with previous versions. So, it will be in our hands to make the scripts that transform ORA restarts in to readable EC-Earth restarts.

Summary of discussions at University of Oxford

[**ORIOL TINTÓ**] Peter Düben is working with Tim Palmer and Stephan Juricke on the speedup of openIFS/NEMO codes by reducing precision from double to single. The big argument is to say that the system is chaotic anyway, so that we don't loose much information by truncating precision. Speedup is quite impressive (30-40 %, if I remember correctly) and has been also tested in NEMO. It is however not sufficient to just change precision level in the namelist, because in many parts of the code precision has been hardcoded.

[**VIRGINIE, NEVEN**] Stephan Juricke is working on stochastic parameterizations in sea ice. He's now more on the ocean side, but plans to get back to sea ice soon. His modifications to the NEMO code are not intrusive and could be tested in Barcelona, if we want to. He plans to perturb P*, the albedo of ice and drag coefficients. It is not clear however how to perturb albedo: before the albedo routine, in the middle of it or after it?

Seminars of Kerry Emanuel at University of Oxford

Kerry Emanuel is a recognized scientist in the domain of atmospheric convection, radiation, and cyclogenesis. He gave first a technical seminar presenting a simple convective-radiative model of the atmosphere. The model is quite idealized but complex enough to test how various parameters (SST, resolution, domain size) affect cumulo-nimbus formation and their clustering. In the much more general second talk, he had an overview about hurricanes their response to climate change and their role in the climate system. He made the nice analogy between a hurricane and a Carnot engine. In fact, hurricanes are extremely efficient at converting latent heat into kinetic energy! He also showed analyses that suggest the critical role of hurricanes in storing heat in the Earth System. After a hurricane has passed, waters are mixed (due to stirring) and the sun can then input a substantial amount of heat in the ocean. He estimated the power input from hurricanes and other storms to be ~1.4 PW, which is just huge in comparison to typical values for e.g. poleward heat transport (~2-3 PW).

Areas of future collaboration

Investigation of the role of sea ice salinity on global ocean circulation. Magdalena is puzzled by the result she obtained and would like to test how NEMO oceanic circulation (e.g., AMOC) is sensitive to the parameterization of salinity rejection in LIM. I can launch experiments at ORCA1 with nn_icesal taking different values.

Ice-ocean covariances in ORAS4 and ensemble simulations. Steffen had explored relationships between sea ice thickness variability and SSS variability. He noticed interesting correlations in the ORAS4 reanalysis, but that could be good to see if we have the same in our free ensemble runs. Also, how this all depends on resolution might be an important topic for PRIMAVERA.

Assimilation with the ensemble Kalman Filter (EnKF). The current version of 3D-VAR at ECMWF already involves members, and they are thinking to test or implement the EnKF for sea ice and ocean estimation. However, at ECMWF, scientific objectives and strategic objectives are sometimes orthogonal, so this will depend a lot on the freedom that Steffen has to do this. But they are definitely interested in keeping collaboration with us (EC-Earth, BSC, UCL) for those important aspects of polar predictions.

LIM3 in the operational system. Again, this is more a scientific possibility than a strategic decision at this stage. But they will be testing LIM3 in the coming months; on the long term, it could become their reference model for future reanalyses. I'm keen in helping Sarah for restartability issues.

Important information for BSC members

[**CLIMATE PREDICTION**] There's a general feeling that BSC does not contribute/interact enough on the EC-Earth development portal. Those of us using EC-Earth should systematically subscribe to the development portal, and we should be alert when a new thread is posted.

[**CLIMATE PREDICTION**] ECMWF will not use the extended ORCA1 grid (the one used by BSC and SMHI, and EC-Earth in general), to ensure continuity in their reanalyses.

[**ALL**] ERA5 will be the next big reanalysis released by ECMWF. Hopefully end of next year (2017). This will be a high-resolution atmospheric reanalysis with members. In parallel, a century-long reanalysis will be issued

[**NEVEN, VIRGINIE**] Stephan Juricke is highly interested in coming over in Barcelona. We should try to get funding for him (for a month or so).

[**CLIMATE PREDICTION**] There is a climate database called SPHINX. It is a kind of pre-PRIMAVERA set of simulations exploring the impact of resolution in ensemble runs. https://eudat.eu/communities/data-sphinx-data-storage-and-preservation-of-high-resolution-climate-experiments