SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Incorporating land-surface model uncertainty into the IFS
Computer Project Account:	spgbweis
Start Year - End Year :	2014 - 2016
Principal Investigator(s)	Dr. Antje Weisheimer
Affiliation/Address:	University of Oxford Atmospheric, Oceanic and Planetary Physics Clarendon Lab Dept. of Physics University of Oxford Oxford OX1 3PU Email: Antje.Weisheimer@physics.ox.ac.uk
Other Researchers (Name/Affiliation):	Tim Palmer (Oxford) Dave MacLeod (Oxford) Florian Pappenberger (ECMWF) Sarah-Jane Lock (ECMWF) Gianpaolo Balsamo (ECMWF)

The following should cover the entire project duration.

Summary of project objectives

(10 lines max)

The land surface is a key source of seasonal predictability and it has been shown that land-atmosphere coupling contributes to the development of seasonal heatwaves. It is highly heterogeneous in space and time, which introduces uncertainty when the climate system is simulated on the typical grid resolution used in seasonal forecasting. Uncertain model parameter values are also a source of uncertainty in simulation.

The aim of this project was to improve the representation of uncertainty in the land surface component of the IFS (HTESSEL). This will be explored by running experiments that perturb uncertain aspects of land surface processes.

Summary of problems encountered

(If you encountered any problems of a more technical nature, please describe them here.)

No problems encountered.

Experience with the Special Project framework

(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

We have found the experience very productive; the access to supercomputer resources has enabled experiments to be carried out which would have otherwise been quite limited. In practice, carrying out these experiments would have been impossible without help from user support at ECMWF, which was invaluable.

Summary of results

(This section should comprise up to 10 pages and can be replaced by a short summary plus an existing scientific report on the project.)

The work on this special project examined the impact of various perturbation strategies to explicitly represent land surface uncertainties. Two peer-reviewed papers have resulted from this work so far: one which analyses the impact of perturbing key soil hydrological parameters on the representation of the 2003 European warm summer in seasonal forecasts and a second looking at the impact of uncertainty in hydrological parameters on representation of soil moisture memory. A third paper is planned, which will summarize, compare and contrast the overall impact of all hydrology perturbation schemes investigated. An investigation into the impact of stochastic perturbation of the land-atmosphere thermal coupling parameter was also carried out as part of a three-month placement of Dave MacLeod at ECMWF in Spring 2016: results showed little impact and no plan is in place to write up this work.

Abstracts from the two published papers are reproduced below. A summary of the results from the planned third paper follows. Finally, the results on thermal coupling perturbation are described.

List of publications/reports from the project with complete references

Paper #1: MacLeod DA, Cloke HL, Pappenberger F, Weisheimer A. 2016. Improved seasonal prediction of the hot summer of 2003 over Europe through better representation of uncertainty in the land surface. *Q. J. R. Meteorol. Soc.* 142:79–90 (also published as ECMWF technical memorandum #761)

Methods to represent uncertainties in weather and climate models explicitly have been developed and refined over the past decade and have reduced biases and improved forecast skill when implemented in the atmospheric component of models. These methods have not yet been applied to the land-surface component of models. Since the land surface is strongly coupled to the atmospheric state at certain times and in certain places (such as the European summer of 2003), improvements in the representation of land-surface uncertainty may potentially lead to improvements in atmospheric forecasts for such events.

Here we analyse seasonal retrospective forecasts for 1981–2012 performed with the European Centre for Medium-Range Weather Forecasts (ECMWF) coupled ensemble forecast model. We consider two methods of incorporating uncertainty into the land-surface model (H-TESSEL): stochastic perturbation of tendencies and static perturbation of key soil parameters.

We find that the perturbed parameter approach improves the forecast of extreme air temperature for summer 2003 considerably, through better representation of negative soil-moisture anomalies and upward sensible heat flux. Averaged across all the reforecasts, the perturbed parameter experiment shows relatively little impact on the mean bias, suggesting perturbations of at least this magnitude can be applied to the land surface without any degradation of model climate. There is also little impact on skill averaged across all reforecasts and some evidence of overdispersion for soil moisture.

The stochastic tendency experiments show a large overdispersion for the soil temperature fields, indicating that the perturbation here is too strong. There is also some indication that the forecast of the 2003 warm event is improved for the stochastic experiments; however, the improvement is not as large as observed for the perturbed parameter experiment.

Paper #2: MacLeod DA, Cloke H, Pappenberger F, Weisheimer A. 2016. Evaluating uncertainty in estimates of soil moisture memory with a reverse ensemble approach. *Hydrol. Earth Syst. Sci.* 20:2737–2743

Soil moisture memory is a key component of seasonal predictability. However, uncertainty in current memory estimates is not clear and it is not obvious to what extent these are dependent on model uncertainties. To address this question, we perform a global sensitivity analysis of memory to key hydraulic parameters, using an uncoupled version of the H-TESSEL land surface model.

Results show significant dependency of estimates of memory and its uncertainty on these parameters, suggesting that operational seasonal forecasting models using deterministic hydraulic parameter values are likely to display a narrower range of memory than exists in reality. Explicitly incorporating hydraulic parameter uncertainty into models may then give improvements in forecast skill and reliability, as has been shown elsewhere in the literature. Our results also show significant differences with previous estimates of memory uncertainty, warning against placing too much confidence in a single quantification of uncertainty.

Representing uncertainty in soil hydrology: a comparison of the impact of three methods on seasonal climate forecasts (paper in preparation)

Following on from the work described in paper #1 above, perturbation of soil hydrology has been implemented in CY41R1. Seasonal hindcast experiments were carried out, following a standard setup: initialised on 1-May for 1981-2013 and integrating forecasts for four months.

Three types of perturbation were tested:

- PP: perturbed parameter. The parameters perturbed are the Van-Genuchten alpha parameter and the saturated soil conductivity. These are soil-type dependent, and are used in the equations for hydraulic conductivity in the soil, describing the soil moisture release curve and the conductivity of the soil at saturation. Each parameter is perturbed from its default value by [-80,-40,0,40,80]%, giving a 25-member ensemble overall.
- ST: stochastic tendencies. Here the soil moisture tendencies output by the land surface parameterization subroutine are perturbed stochastically. This follows the atmospheric stochastic scheme, where the tendencies to temperature, humidity and wind fields from the atmospheric physics parameterizations are perturbed, using an autoregressive AR1 process with 3 scales (the spectral pattern generator, SPG). The scales of the SPG for SPPT are chosen to represent typical short, medium and long temporal/spatial scales, with more weight given to the short/small scales. The default of the SPG used in SPPT is to weight the small/medium/long scales with a standard deviation of 0.52/0.18/0.06. For the stochastic perturbation of HTESSEL tendencies, the moisture fields were perturbed for each vertical soil level equally keeping the overall standard deviation constant (i.e. 0.32/0.32/0.32).
- iSP: stochastic parameters. Here the two parameters perturbed in the PP experiment were perturbed using stochastic perturbations generated by independent realizations of the SPG. That is, the perturbation of the parameters is independent. Three SP experiments were carried out, using different weights of the SPG: iSP-1 (0.32/0.32/0.32), iSP-2 (0.06/0.18/0.52) and iSP-3, which uses only one scale which is unused in the other experiment and default SPPT. This scale has a decorrelation length of 2000km and time of 365 days, roughly approximating a static perturbation across 4 months.

The PP experiment is identical to that which produces an improvement in the 2003 warm summer described above, except it is implemented in a more recent cycle of the model. The ST experiment is similar to the ST-2 experiment described in the paper, though only the moisture tendency is perturbed since perturbation of the temperature tendencies led to a large increase in the spread of the soil temperature fields. The SP experiments have not been published.

Overall the SP experiments have the largest impact on the mean state seen in Figure 1. Though the parameter perturbation is symmetric they lead overall to a large increase in runoff. This is likely due to the nonlinear relationship between the soil parameters and subsurface runoff. For the SP experiments this leads to a drying of the soil, and consequently a reduction in latent heat flux due to lower availability of soil moisture. Decreases in latent heat flux are balanced by increases in sensible heat flux, leading to an increase in summertime soil temperatures by around 0.5K (Figure 2). Mean state impacts for the ST and PP experiments are lower, though PP also leads to a slight increase in 2m air temperature in some regions, despite no significant impact on the mean soil moisture.

Note that the water budget closure term (dW, calculated as the difference between observed change in soil moisture and the net flux into the soil) is not identically zero. This is partly because snow is artificially capped at 10m. However, it has not been possible to fully close the water budget in the control run, even for grid points that have no snow. Imbalanced surface water budgets have previously been reported for HTESSEL (*Kauffeldt et al. 2015*) and work is on-going with Bart van den Hurk (KNMI) and Souhail Boussetta (ECMWF) to investigate this.

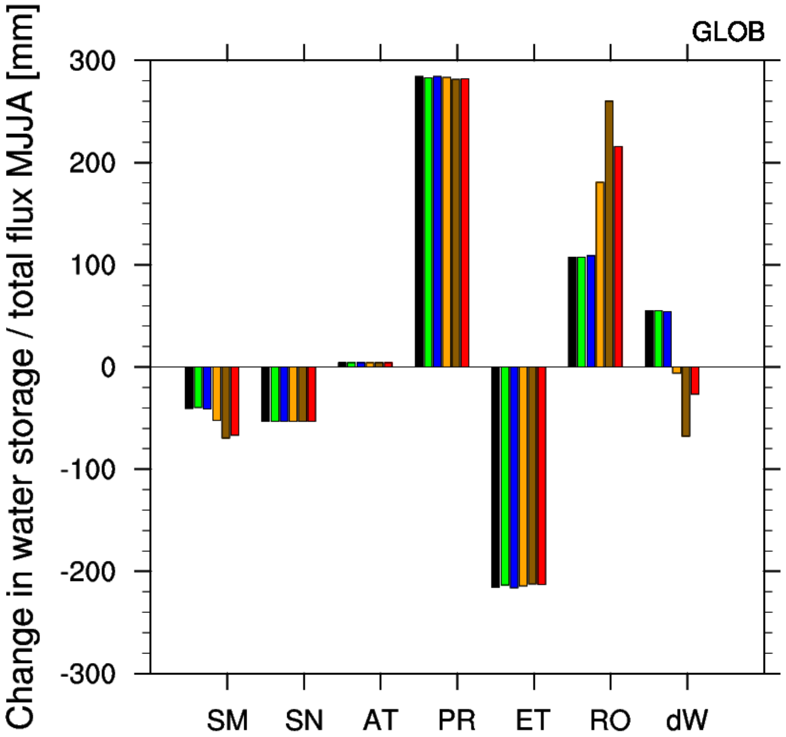

Figure 1: Global water budget for all experiments: control, PP, ST, iSP-1, iSP-2, iSP-3 (black, green, blue, orange, brown, red). Change in total storage total column soil moisture (SM), snow (SN) is plotted, along with total flux of atmospheric water vapour (AT), precipitation (PR), evapotranspiration (ET) and runoff (RO). The final term dW indicates water closure, which is non-zero, partially because snow depth is capped at 10m. However closure has not been achieved for non-snow points, see discussion above.

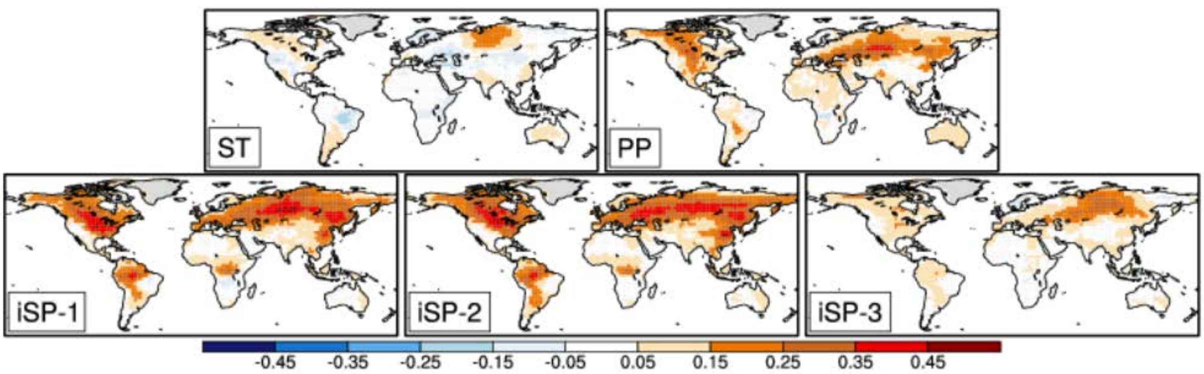

Figure 2: Mean JJA 2m air temperature difference from control in each hydrology perturbation experiment

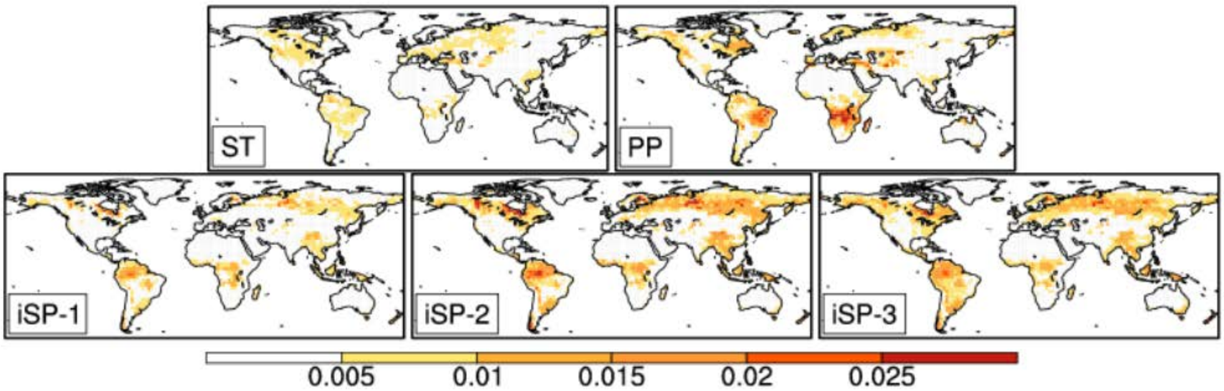

Figure 3: Average JJA soil moisture standard deviation, difference from control for each experiment

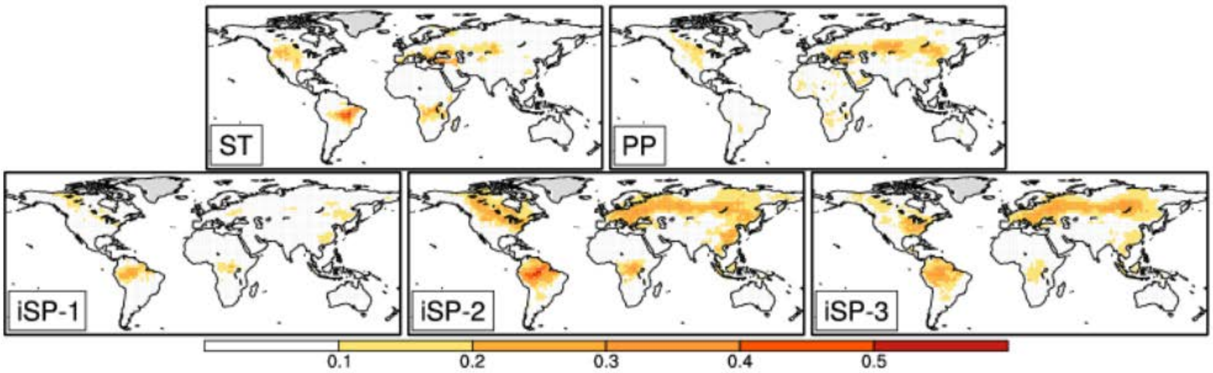

Figure 4: As figure 3, for JJA 2m air temperature

Impacts on soil moisture ensemble spread and 2mT spread are shown in Figures 3 and 4. Here the SP experiments show impacts in roughly the same locations as impacts on the mean state. Interestingly iSP-2 and iSP-3 show the largest impact on 2m temperature ensemble spread, whilst iSP-1 and iSP-2 show the largest impact on mean state. That is, the SPG scale weightings used in iSP-1 have a large impact on the bias but small impact on spread, whilst the iSP-3 experiment has a larger impact on spread and lower impact on bias. PP has impacts on spread of similar magnitude to the SP experiments, however the pattern of impact is quite different. That is; PP shows large impact over Brazil where SP does not, whilst SP increases spread over the boreal forests where PP has little impact. This is somewhat puzzling as the experiments perturb the same parameters. Indeed, iSP-3 is in theory an almost identical setup to PP, the only difference being in the spatial component to the perturbation of iSP-3. Though this spatial perturbation has a large decorrelation length in iSP-3 and is evenly distributed globally, we must presume that the different impacts on spread seen for iSP-3 arise somehow from this difference, though it is not immediately obvious how.

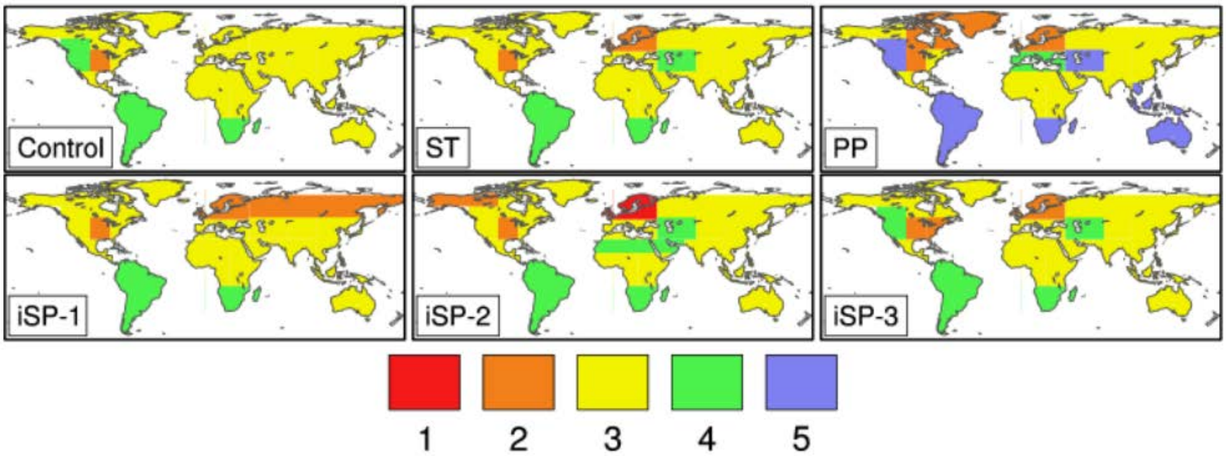

Figure 5: Reliability category of JJA soil moisture falling in the highest quintile, for all hydrology perturbation experiments

Previously the PP experiment showed improvement for representation of the 2003 warm European summer over the CY36R4 control. This result was not replicated in CY41R1, as the simulation of this event in the control improved markedly. However, for the CY41R1 experiments the reliability of soil moisture forecasts improved markedly for the PP and SP experiment (Figures 5 and 6). Here, the reliability categories (defined by *Weisheimer and Palmer, 2014*) for JJA soil moisture falling in the highest quintile improved for many regions. An example of this for West North America is shown in Figure 6 for all experiments. The points on the reliability diagram for the experiments lie directly on the diagonal "perfect forecast" line, for PP and iSP-3, significantly improved over the control. The improvement here is shown to come mainly from higher probability forecasts, which are under-confident in the control and much improved in the experiments. Similar improvements were seen for lower quintile soil moisture, though do not bring any systematic improvements to 2m air temperature reliability.

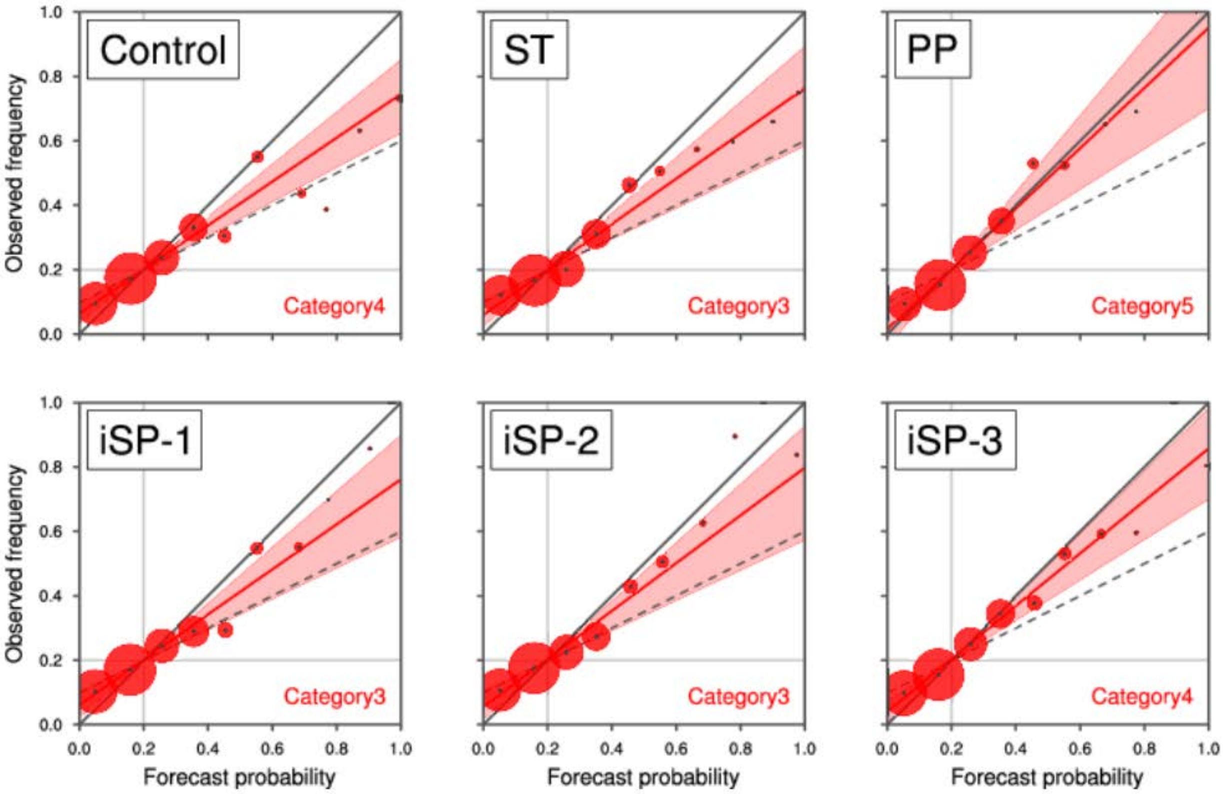

Figure 6: Reliability diagrams for JJA highest quintile soil moisture over West North America, for all hydrology perturbation experiments

Perturbation to the land-atmosphere coupling parameter in extended-range forecasts

From April–June 2016 Dave MacLeod worked temporarily at ECMWF in order to investigate the impact of stochastic perturbation of land-atmosphere coupling. This work has been in conjunction with Sarah-Jane Lock and Gianpaolo Balsamo, and involves the application of the new Stochastically Perturbed Parameters (SPP) scheme to the Lambda parameter, which controls the thermal coupling between the skin temperature and surface. There is uncertainty in the parameter, which has a large influence on the diurnal temperature cycle at the surface.

Experiments were carried out with IFS CY42R1, with 1 month EPS runs initialized every 16 days from 1-Dec 2013 to 26-Nov 2014, with 20 perturbed members. Perturbation was made to the thermal coupling parameter using the SPP framework. Initial experiments were carried out, using a perturbation distribution parameter of 0.3. Extensive analysis of these runs was carried out, but it was not possible to find any impact of the perturbation. Following this, additional experiments were carried out, one which increased the perturbation parameter to 0.5 and a second which increased the parameter to 0.5 and used long time and spatial scales for the pattern which were essentially infinite, for the 1 month run.

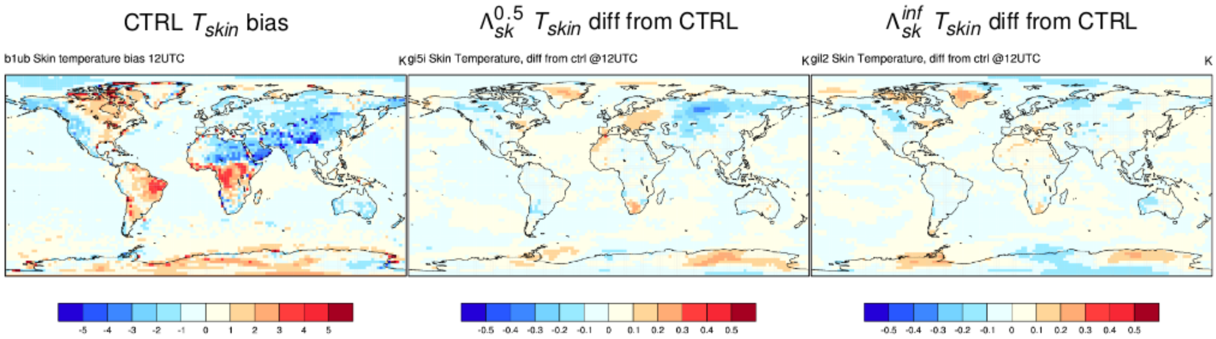

Figure 8: Control skin temperature bias for lead month 1 and impact of the two thermal coupling perturbation experiments expressed as the difference in the mean skin temperature

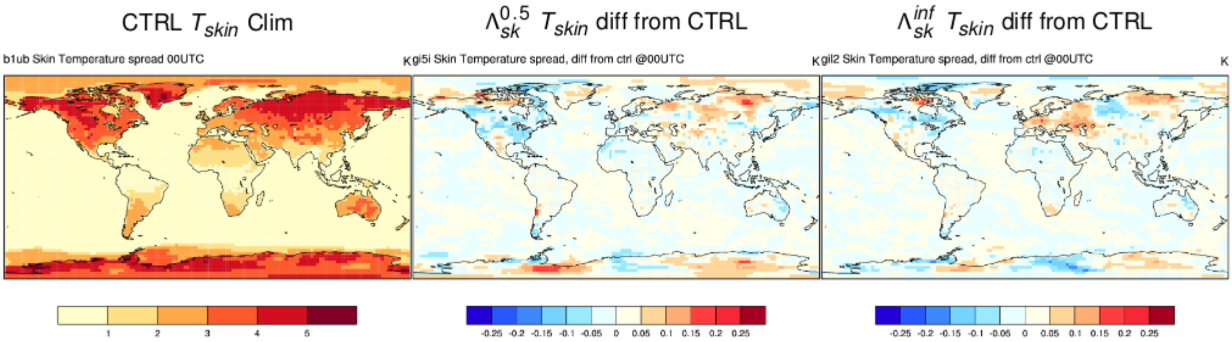

Figure 9: As figure 8, for ensemble spread of skin temperature

These experiments had a very minor impact on the mean state, with a localised slight decrease in the skin temperature in North Asia (Figure 8), along with a slight increase in spread in the same region (Figure 9). However in general the impact on the mean state was below 0.1K, with no impacts in the Tropics or the Southern Hemisphere, whilst differences in spread are less than 2% of the control spread and not significant. Average impact on spread across all start dates as a percentage of control spread is shown in Figure 10 for various land regions, with the impact on skin temperature RMSE (as a percentage of control RMSE) is shown in figure 11. The impacts are not large, with no systematic sign of the impact in spread, and a mild tendency for increasing the RMSE, of the order of 2%.

Further analysis (not shown) indicated that the perturbation was insensitive to the diurnal cycle, with no differing impact at 00UTC and 12UTC. Additionally, no impact on sensible heat fluxes over the first day was found. Looking across all start dates individually (Figure 12) reveals that the impact on ensemble spread shows much larger variability across start dates than the average impact of the perturbation.

Overall, perturbation experiments for the coupling did not provide any evidence that perturbation to the thermal coupling parameter would provide a useful method for increasing the near-surface spread. Given that the coupling parameter of 0.5 is not particularly small, further experiments might consider alternative routes for addressing the under-dispersion near the surface.

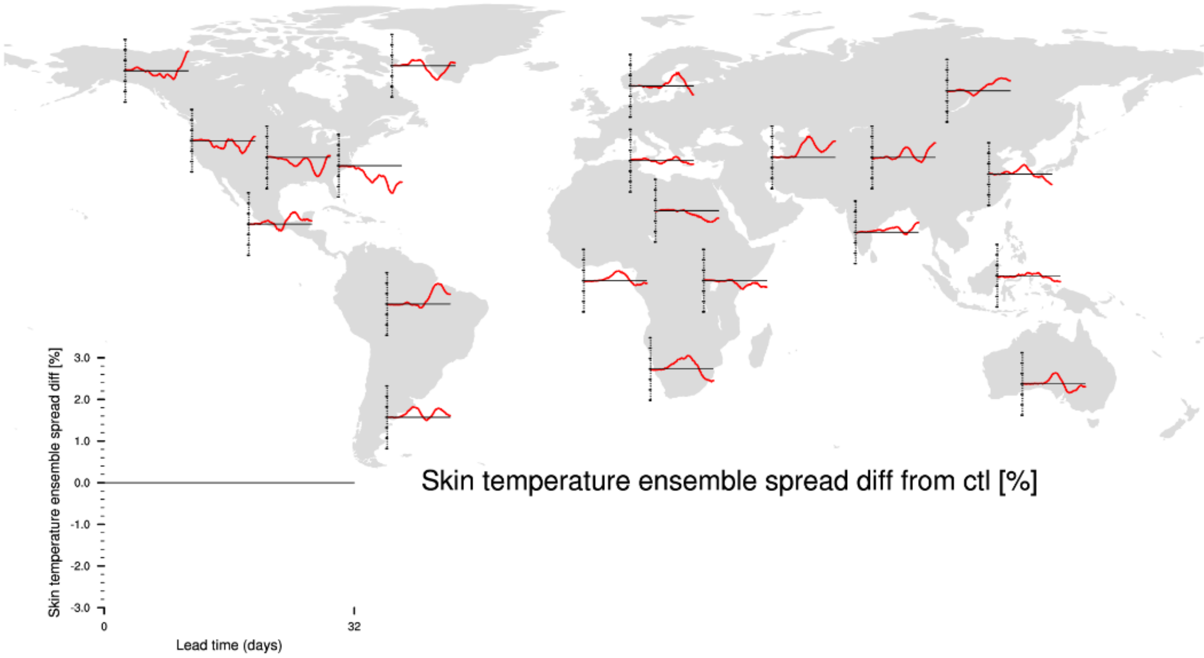

Figure 10: Average skin temperature ensemble spread as a function of lead time for standard land regions, expressed as a percentage of control spread

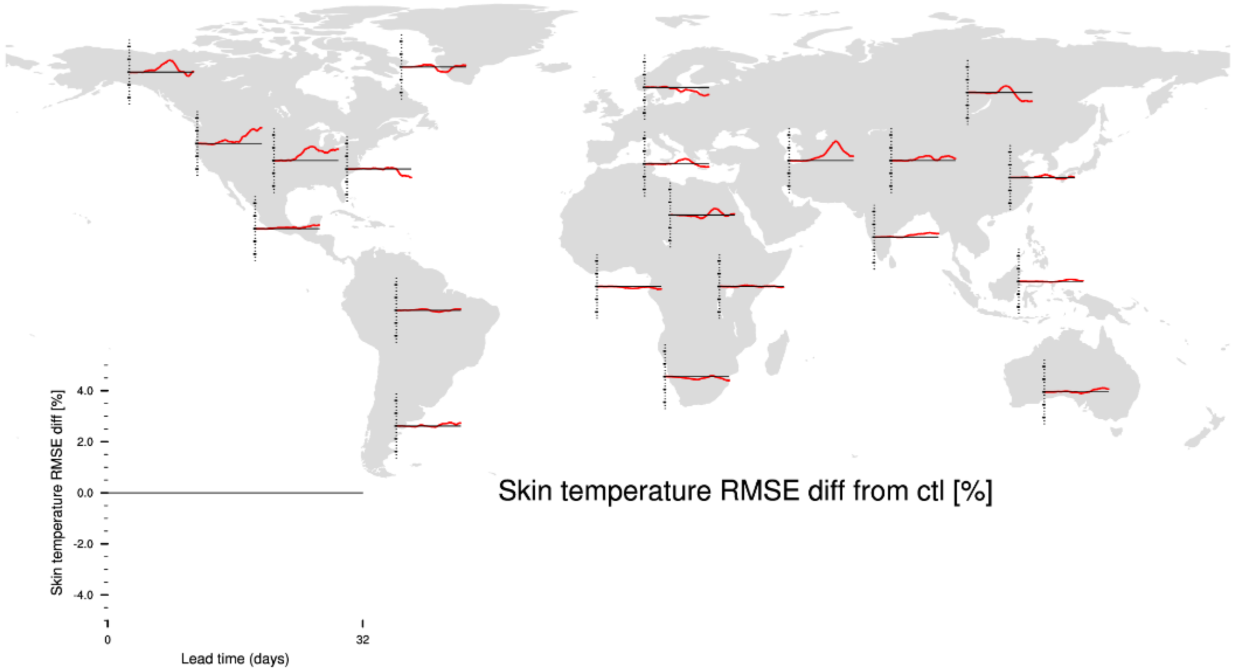

Figure 11: As figure 10, for skin temperature RMSE

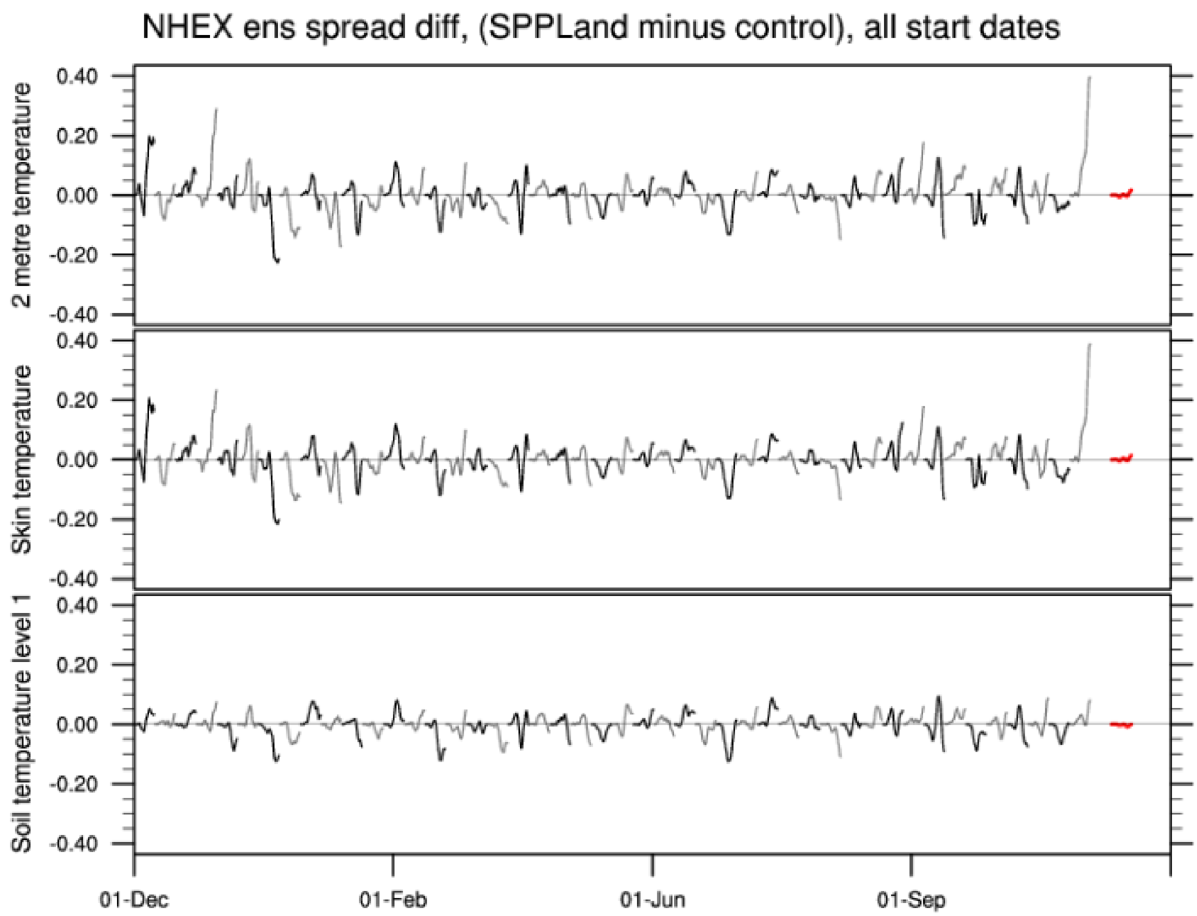

Figure 12: Ensemble spread in Northern Hemisphere Extratropics expressed as the difference from control, for variables 2m air temperature, skin temperature and top level soil temperature (top, middle and bottom). All start dates are plotted (NB black and grey are used alternately for visual clarity and do not have any further significance). The average difference over all start dates is plotted as the final red line.

References cited above

Kauffeldt, A., S. Halldin, F. Pappenberger, F. Wetterhall, C.-Y. Xu, and H. L. Cloke. 2015. Imbalanced land surface water budgets in a numerical weather prediction system. *Geophys. Res. Lett.*, **42**, 4411–4417. doi: 10.1002/2015GL064230.

Weisheimer A and Palmer TN 2014. On the reliability of seasonal climate forecasts. *J. R. Soc. Interface*, **11** 20131162; DOI: 10.1098/rsif.2013.1162.

Work from the special project is also reported in: Leutbecher M et al, 2017. Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. Accepted at QJRMS (also published as ECMWF technical memorandum #785)

Future plans

Our current plans are to write up and publish the work on the newer set of experiments with CY41R1, described above. No special projects focusing on the land surface uncertainty are active or planned, however the stochastic hydrology scheme has been implemented in EC-Earth. Simulations are planned to examine the impact of this and other schemes on climate timescales. This work is being carried out as part of the PRIMAVERA project by Kristian Strømmen (University of Oxford). Some of these simulations will be carried out at ECMWF, under the special project spgbtpsp.

Description of additional experiments performed based on the long seasonal hindcasts

Several seasonal hindcast sensitivity experiments were run in late 2016, based on the long hindcast control runs described in the papers above. The setup of these experiments is described below, along with some initial results; analysis of these experiments is on-going.

SST denial experiment

To test the impact of the anomalies in the Indian and West Pacific Ocean region, experiments were run with the SST forcing in the region shown in Figure 1 replaced by daily climatology. A smoothing is applied at the edges to avoid non-physical steps in the applied forcing field and any resulting spurious model behaviour. This experiment gives an indication of the sensitivity of the seasonal forecast to anomalies in this region. The run follows the setup described in paper #1 above, with a reduced ensemble size of 25 members. The experiment was run for 1st November start dates, to look at Northern Hemisphere winter and compare with the results of the published papers.

Initial results from the November start date are shown in Figure 2. A drop in skill is seen for correlation skill of 2m air temperature over the SST denial region, as is expected. However, the drop in skill is not uniform across the region, and latitude band across the central Indian Ocean retains significant skill. Figure 2 also shows impact on 500hPa geopotential height, revealing that the removal of the anomalies leads to reduction in skill across a larger domain extending westward from the region, across Africa, the Atlantic and South America.

An additional experiment for the 1st February start date was run, using the same denial box as the 1st November run. This was run in order to study the impact on predictability of the East African long rains, which will be used as part of the NERC/DFID project ForPAc.

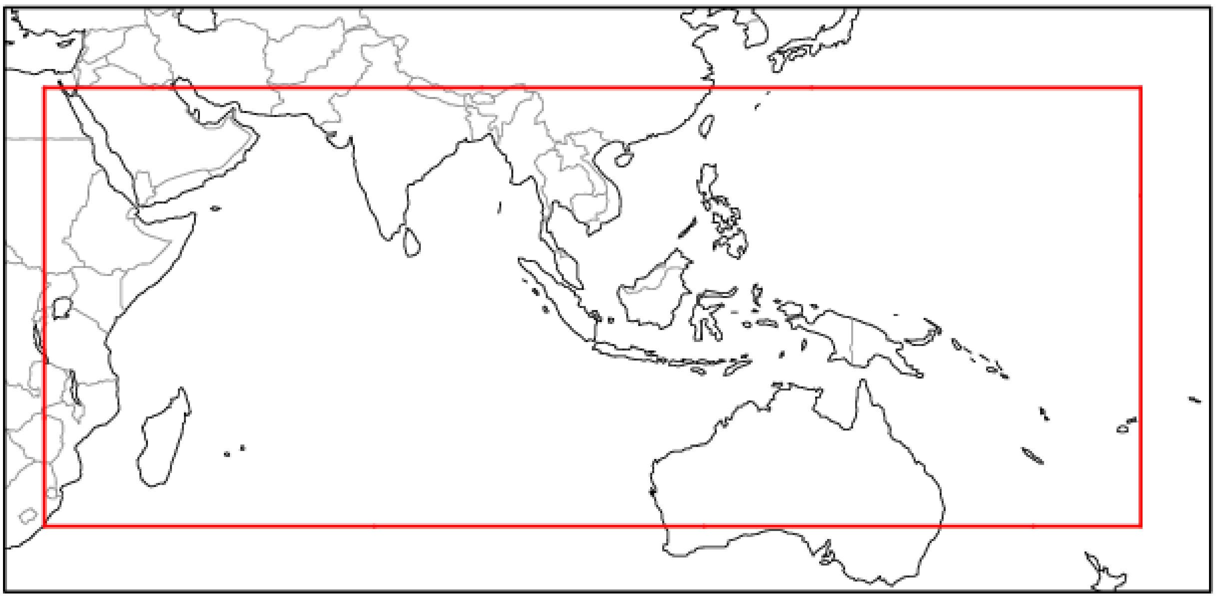

Figure 2: Region used to remove SST anomalies for the SST denial experiments

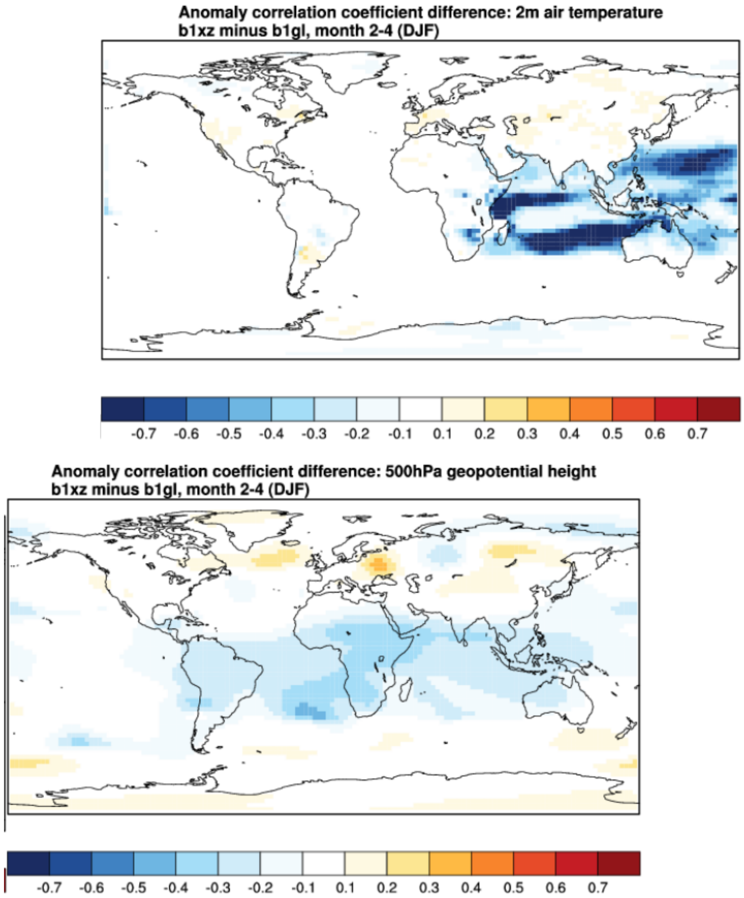

Figure 3: Impact of removing SST anomalies in the box shown in figure 1: DJF ensemble mean anomaly correlation coefficient for 2m air temperature (left) and 500hPa height (right).

Scrambled Initial Conditions

For this experiment, the 1st November start date was re-run following the control hindcasts, except for each year the initial condition is chosen from a year at random. This experiment is able to reveal the aspects of sensitivity of the forecast to initial conditions in a long-term context. Additionally, since the initial condition years are chosen randomly without replacement, it will also be possible to reorder the hindcast and study the sensitivity of the predictions to the SST forcing. Analysis of this experiment is planned, with an initial result shown in Figure 3. This shows the impact of the initial condition on the DJF 2m air temperature ensemble mean correlation, indicating some importance of initial states for the forecast skill, over India, Asia and the Sahel. It is hypothesised that this predictable component arises from the initial conditions of the land surface, rather than the atmosphere.

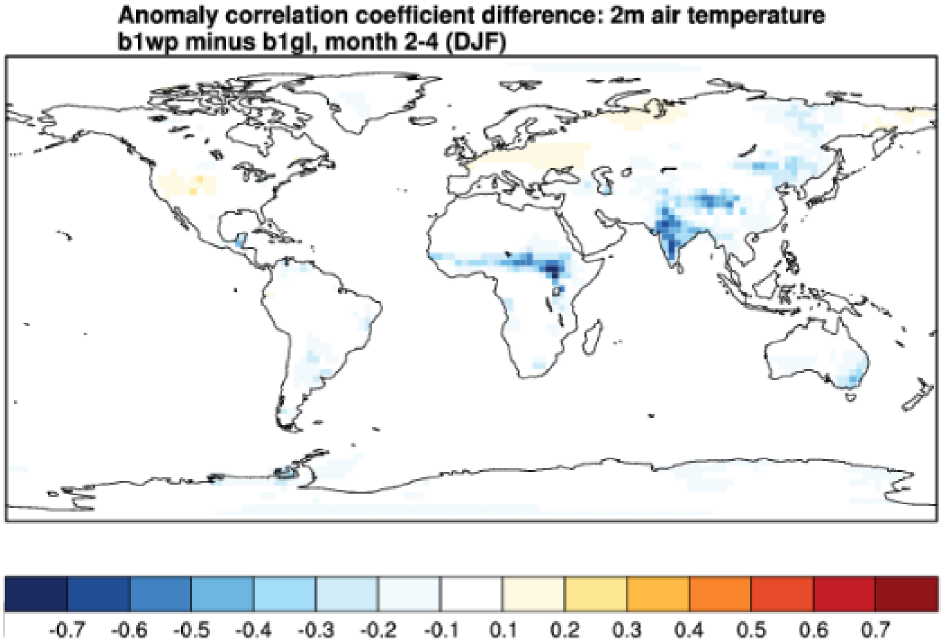

Figure 4: Impact of the scrambled initial condition experiment on ensemble mean correlation of DJF 2m air temperature

Tropical Relaxation

For this experiment a band in the tropics (20N-20S, 0-360E) is 'relaxed' to the ERA20C reanalysis. That is, throughout the model simulation all atmospheric fields in this region are pushed back toward the reanalysis if they start to diverge. This experiment has only been carried out for the 1st November start date, for a 15-member ensemble. A smaller ensemble is sufficient as the relaxation reduces the potential for dispersion. Analysis of this experiment will be carried out soon.

References cited above

Müller W, Appenzeller C, Schär C. 2005. Probabilistic seasonal prediction of the winter North Atlantic Oscillation and its impact on near surface temperature. *Clim. Dyn.* **24**: 213–226, doi: 10.1007/s00382-004-0492-z.

Shi W, Schaller N, MacLeod DA, Palmer TN, Weisheimer A. 2015. Impact of hindcast length on estimates of seasonal climate predictability. *Geophys. Res. Lett.* 42:1554–1559

Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new Special Project.)

In this project, the seasonal hindcasts of the 20th Century were based on atmosphere-only simulations with prescribed SSTs and sea ice initialised with the atmospheric reanalysis ERA-20C. The next step is to perform a similar set of hindcasts but with a fully coupled ocean-atmosphere model. The availability of ECMWF's first coupled reanalysis of the 20th Century CERA-20C will make this possible. This work is being carried out as part of the 2017-2018 special project spgbweis "Coupled seasonal forecasts of the 20th Century (CSF-20C)".

Analysis of the experimental data produced above will continue. Particularly of note is that Bart van den Hurk (KNMI) has extended the scrambled initial condition run in Spring 2017 whilst working as a visitor to ECMWF. Bart has carried out experiments based on the long-term seasonal hindcast, where only the land surface initial conditions are randomised, to assess the sensitivity of the predictability to land initial state, and examine any decadal-scale variability of the relationship.

The 1st February SST denial experiment as well as the control seasonal hindcasts will be used as part of the NERC/DFID project ForPAc, which is working toward improved forecast based action for humanitarian planning in East Africa (Oxford lead Tim Palmer, with Dave MacLeod). The experiments here will be used to establish baseline reliability and economic value of seasonal forecasts for the two rainy seasons across the region.