Project report: Project report of bu0801

New Project title: MiKliP2 Modul A : Module A - Determination of initial conditions and initialization

(old title: Atmospheric and Oceanic Data Assimilation Plus ENsembles Generation; AODA-PENG)

Project lead: Johanna Baehr (IfM Hamburg), Andreas Hense (Meteorologisches Institut Bonn), Reporting period: 01.01.2014 – 31.10.2015

WP numbers follow the joint MiKlip II proposal (as used in the joint DKRZ proposal for 2016).

WP 3 (project number 801)

WP3.1 Ensemble generation

A new ensemble generation approach (Romanova and Hense, 2015), Anomaly Transform method (AT), considering i) orthogonal conditions on perturbation patterns; ii) total energy norm; iii) conserved perturbed initial fields only in the ocean; iv) constrained disturbance amplitudes to the observational data set; was developed in the earlier stage of the MIKLIP I project. The evaluation of the hindcast on independent data observational data sets and different reanalysis products was conduction on the following evaluation metrics and variables: i) estimated energy perturbation growth – globally and vertically averaged; ii) exponential fit on time growth, growth rate; iii) box plots: lower and upper whiskers and the outliers; iv) partitioning the total energy into kinetic and potential parts; v) RMS of the perturbation patterns; vi) anomaly correlation coefficients (ACC); vii) Ensemble Spread Skill Score (ESS); viii) Beta-scores in time (ARH, beta-score). All scores are calculated on the post-processing WIZARD and MIKLIP server. The evaluation of the hindcast showed that reasonable spreads were produced during the whole integration period and are getting optimal in most of the regions at approximately around the lead year five.

However, a study was needed to assess the compatibility of the AT technique toward other ensemble generation methods. An additional hindcast, which starts from ten-day lagged ocean states in January 2000 and integrates forward for ten years, was performed and the AT initialized hindcast was appended over the time to cover a common time span. The new and the extended one, required integration of 50 model years, which consumed approximately 90 000 CPUh and occupied a disc space of about 21 TB. Most of the compressed data, later on, were transferred to the archive. Together with the prototype system, which begins the ensembles from one day lagged atmosphere and ocean and the hindcast created by using Ensemble Kalman Filter (EnKF) (Brune et al. 2015), four different ensemble generation techniques were available for the analysis. From the other side, four different reanalysis and namely GFDL Ocean Data Assimilation Experiment, GECCO2 Ocean Reanalysis, NCEP R2 and MERRA Modern-ERA Retrospective Analysis were used for the comparison. Fig. 1 shows the spatial distribution of the correlation coefficients between hindcasts and reanalysis of monthly anomalies for the hindcast produced only with ocean disturbances (AT and Ocean lagged) and the mean of the four reanalysis for the lead year 1 and 5. Other evaluation scores were calculated such as cumulative frequency distribution (probability estimates), integrated squared difference of the cumulative frequency distribution, spread skill score and reliability diagrams. The data processing and evaluation was performed on the Wizard server.

Another method for ensemble generation used in the operational forecast is the Bred Vector (BV) technique (Toth and Kalnay, 1993). The BV has to be tested also in a long-range, inter-annual, seasonal and decadal forecast. Here we implemented a new version which does not depend on a long control run. Rather the disturbances are initiated at the beginning of a selected period (one year at the moment), extracted and rescaled at the end of that period and added back to the beginning (Fig. 2). The iteration is started with random disturbances. Targeting decadal prediction it is searched only for the slowest modes of the ocean physical processes. The metric used to scale the disturbed fields is the weighted total energy as a sum of zonal and meridional flow components of the kinetic energy and the available potential energy. A rescaling coefficient of the initial fields between every iterative step is defined as a ratio between the total energy growth at the first and the final state. The weighted total energy norm is also used to monitor the growths rates of the fastest growing error modes. Breeding routines are externally designed and are able to be implemented in different coupled climate models through shell scripts. The routines do not depend on the model resolution and the time period of the iterative looping could be defined by the user. The breeding routines, up to date, are implemented on T31L31/GR30L40 with arbitrary initial conditions. The method and the breeding application are still in a testing phase. Nine Bred Vectors are calculated using 12 months looping period over 10 iterative steps. Investigated are the rescaling coefficients and the change of global mean temperature between the iterations aiming to justify the saturation level of the error growth. Fig. 3 shows the most sensitive regions in the ocean, in terms of total energy change, responsible for inter-annual to decadal variability.

Comparison to 2014 request

Consumed resources for breeding in 2015 are as follows: CPUh for BV testing phase around 3500 CPUh, and BV 10 member ensemble around 5500 CPUh.

WP3.2 Ensemble Kalman Filter

Work package 2 of AODA-PENG (MiKlip I) is concerned with the implementation of an EnKF scheme into the ocean component of the coupled Earth system model MPI-ESM to initialize decadal predictions. It is carried out at IfM Hamburg. We are using the "LR" setup of MPI-ESM, that is ECHAM6 T63L47 and MPIOM GR15 with 40 levels. The Ensemble Kalman Filter (EnKF) for MPIOM has been implemented using the Parallel Data Assimilation Framework (PDAF, Nerger and Hiller 2013). Implementation work was concluded with one to three year assimilation tests (8 ensemble members, global variant of EnKF) of the observation representativeness error, in 2013. As a result observational errors of 1 K for temperature and 1 psu for salinity were found to be applied along with the global variant of the EnKF. In the next step we realized two assimilation experiments with 8 members each in MPI-ESM-LR on a monthly basis from 1996 to 2010 with EnKF assimilation in MPIOM but a free atmosphere in ECHAM6. In the first experiment EN3 observations of sub-surface temperature and salinity as well as HadISST sea surface data were used, in the second experiment only EN3 observations of temperature and salinity below 50 m depth were used for assimilation. We also performed a free control simulation with 8 members by extending the already existing MiKlip baseline-1 10-year hindcasts initialized in 1996 for 5 years.
We evaluated the three experiments in terms of correlation with observed temperature, upper ocean heat content and sea surface height. The results are threefold: firstly, correlation with observations generally improves with both assimilation experiments over the free run, secondly, assimilation of observations only below 50 m depths generally leads to better correlations than assimilation of data over the whole water column, and thirdly, correlation improvements are largest

near the ocean's surface and decrease rapidly with depth.

Following the experiments with the free atmosphere we transferred the baseline-1 setup for atmospheric nudging to our assimilation system and prepared an assimilation experiment with MPI-ESM-LR (8 members) running from 1958 to 2014 on a monthly basis with oceanic EnKF assimilation and atmospheric nudging to ERA40/ERAInterim re-analysis data. Compared to the baseline-1 assimilation this weakly coupled coupled EnKF/MPI-ESM assimilation shows almost the same quality in terms of correlation with observations of 2m air temperature.

Concluding the work in 2014 and 2015, we initialized and simulated 10-year hindcasts (8 members) from our weakly coupled assimilation on a yearly basis from 1961 to 2014. The correlations of the hindcast lead year time series with observations were compared to those from the baseline-1 hindcasts. In lead year one our hindcast shows the same pattern in correlation with observations as the baseline-1 hindcasts, although the correlation is generally slightly weaker (Fig. 4a,c). For the averaged lead year 2 to 5 time series our hindcasts show the same correlation quality as baseline-1 hindcasts over the continents, but an improved correlation quality over the tropical and sub-tropical Pacific Ocean as well as over the North Atlantic Current in the central North Atlantic Ocean (Fig. 4b,d).

Comparison to 2014 request

The simulation of the 10-year hindcasts from 1958 to 2014 had to be partly shifted to 2015. We therefore used all but a small part of the allocated resources for 2014 in the requested way, however, the last part of the hindcasts were accounted to 2015.

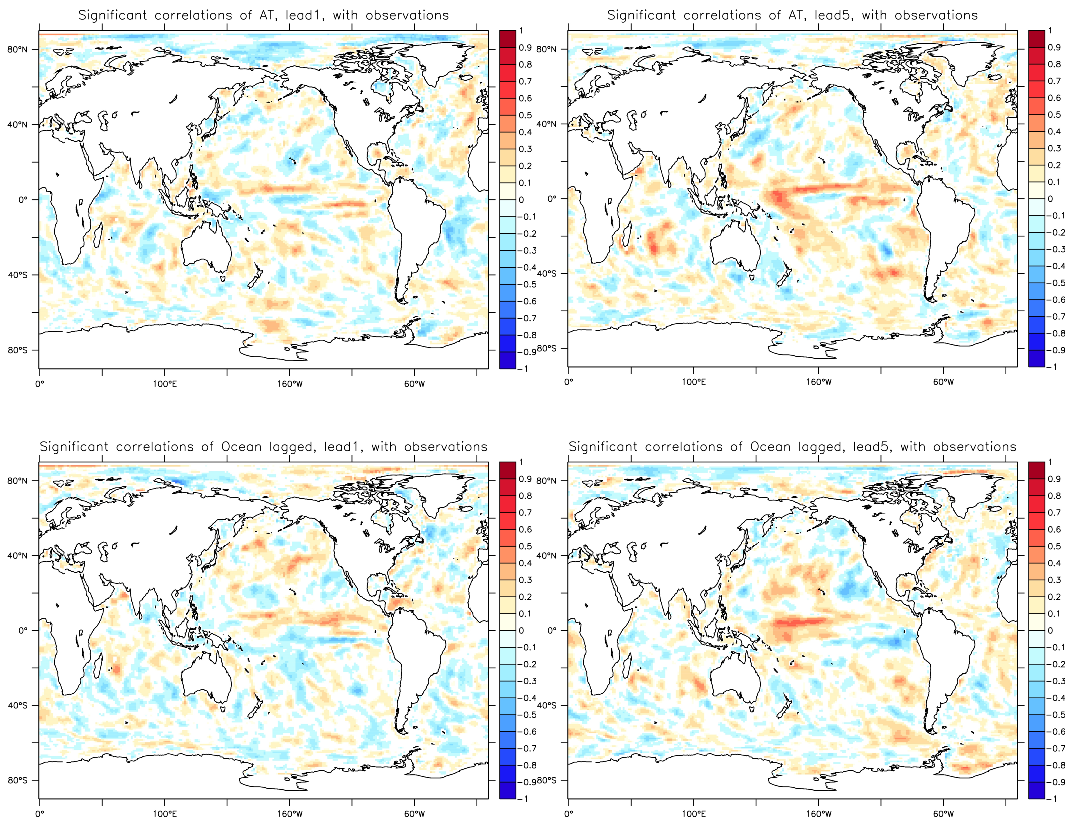

Fig. 1 Spatial correlation patterns for the hindcast produced with AT method (upper panel) and Lagged ocean (lower pannel) for the lead years 1 and 5. The plot shows overall increase in the correlation coefficients with the lead year.

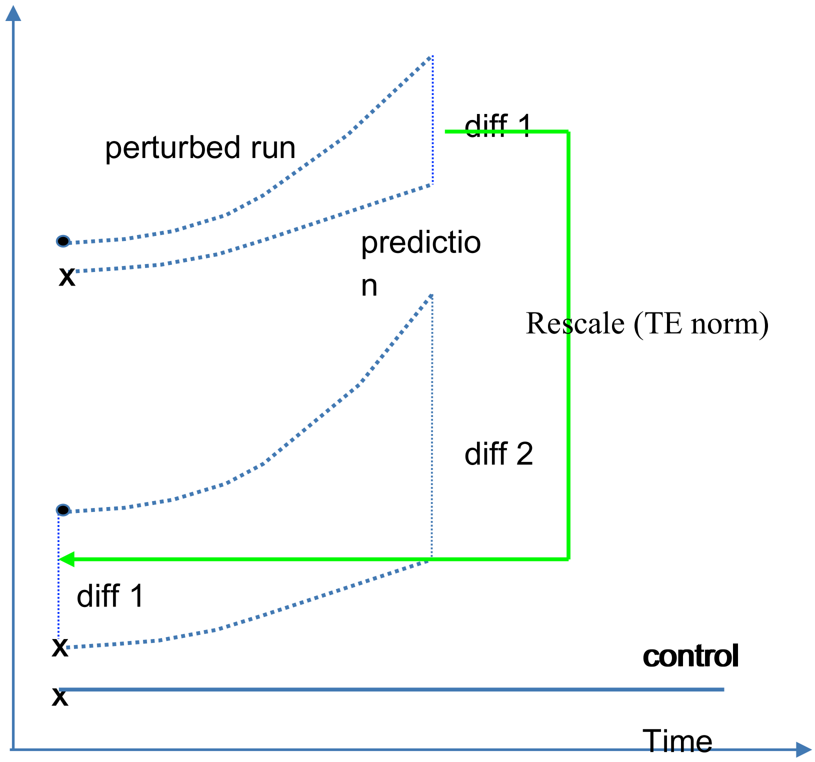

Fig. 2 Schematic figure of breeding, discussed issues: a) time period of looping period which is appropriate for decadal forecast; and b) number of iterations necessary to reach saturation of the growing modes.

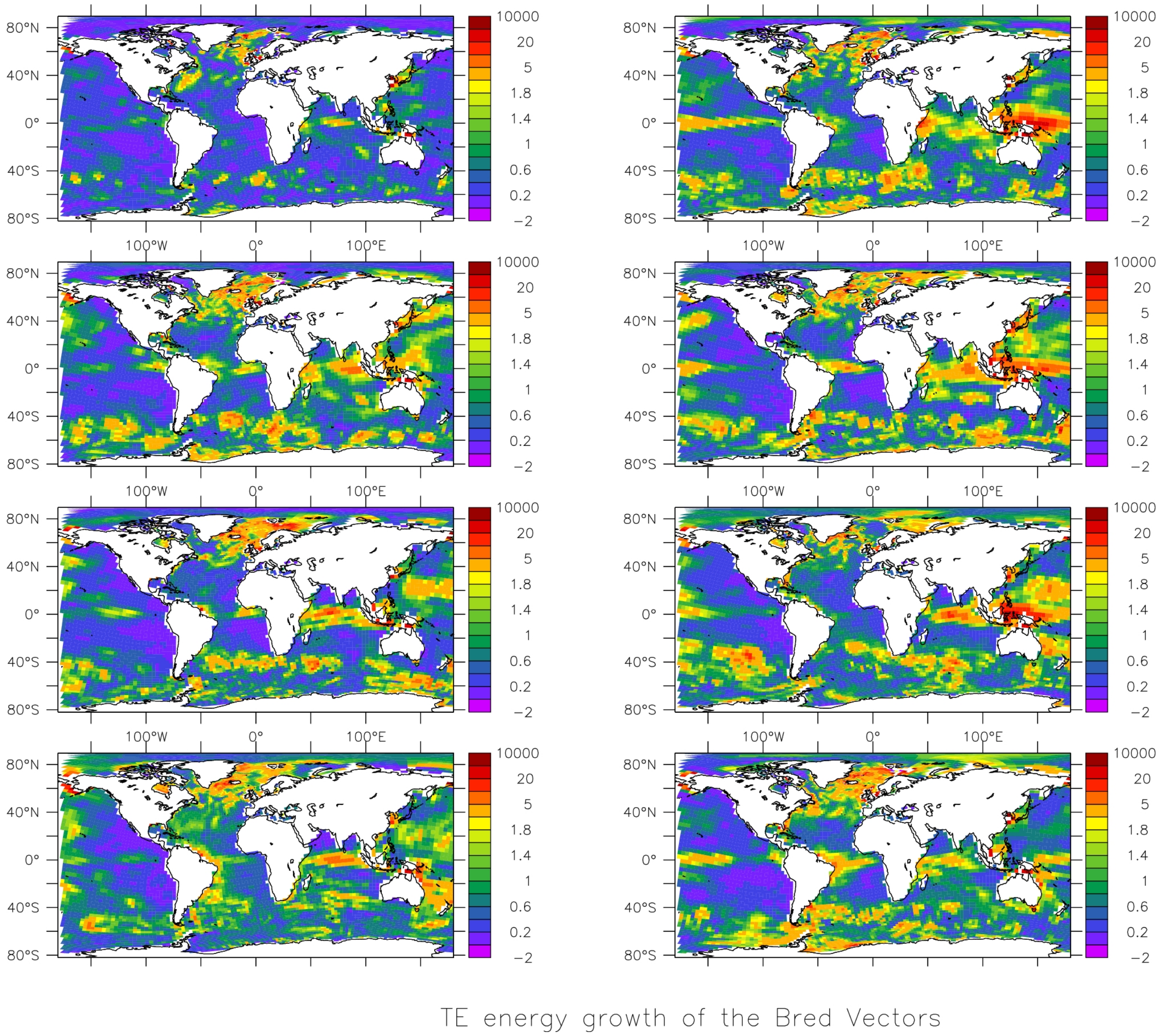

Fig. 3 Spatial pattern of ensemble mean perturbation growth rate (averaged with depth) at final iteration for eight Bred Vectors. The unit of energy growth is dimensionless.

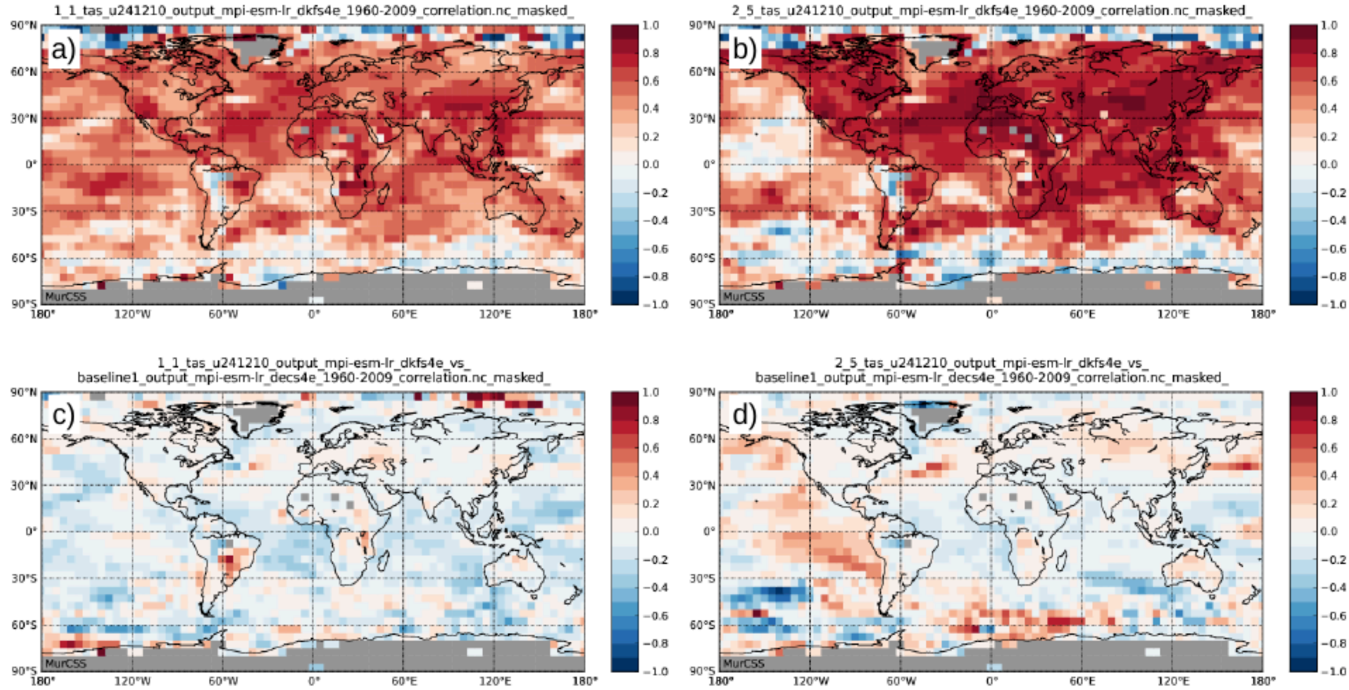

Fig 4: Comparison of correlations of simulations on a yearly basis - hindcasts initialized by coupled assimilation 1960-2014 (above), difference between hindcasts initialized by coupled assimilation and hindcasts initialized by baseline-1 1960-2014 (below), lead year one (left) and average over lead years 2-5 (right) with observations of 2m air temperature from HadCRUT4 (Morice et al. 2012).

References:

Brune, S., L. Nerger, J. Baehr, 2015: Assimilation of oceanic observations in a global coupled Earth system model with the SEIK filter, Ocean Modelling, accepted.

Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, L., Zheng, F., 2014. Seasonal-to-decadal predictions with the ensemble Kalman filter and the Norwegian Earth System Model: a twin experiment. Tellus A 66.

Karspeck, A. R., Yeager, S., Danabasoglu, G., Hoar, T., Collins, N., Raeder, K., Anderson, J., Tribbia, J., May 2013. An Ensemble Adjustment Kalman Filter for the CCSM4 Ocean Component. J. Climate 26 (19), 7392–7413.

Morice, C. P. et al. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. Journal of Geophysical Research: Atmospheres, 117 (D8)

Nerger, L., Hiller, W., 2013. Software for ensemble-based data assimilation systems - Implementation strategies and scalability. Computers and Geosciences 55, 110 – 118

Romanova, V. and A. Hense, 2015, Anomaly Transform Methods Based on Total Energy and Ocean Heat Content Norms for Generating Ocean Dynamic Disturbances for Ensemble Climate Forecasts. ClimDyn. DOI: 10.1007/s00382-015-2567-4

Toth, Z. and E. Kalnay., 1993, Ensemble forecasting at NMC: the generation of perturbations. Bull. Amer. Meteor. Soc., 74, 2317-2330.

Project: 756
Project title: MiKlip PastLand
Project leader: Stefan Hagemann
Report period: 31.10.2013 - 31.10.2015

Optimum parameter and state estimation of the land and biosphere

For the 2014/2015 period, PastLand aimed to investigate the suitability of Transformation of Algorithms in Fortran (TAF) as superior method for soil moisture initialization. Furthermore, a soil moisture initialization experiment was conducted to learn about the life time of soil moisture perturbations in an Earth System Model (ESM).

Soil moisture initialization experiment

In order to evaluate whether the initialization of soil moisture has the potential to improve the prediction skill of coupled climate models at seasonal to decadal time scales, an elaborated AMIP-type experiment was conducted. The experiment design considered soil moisture initialization in different seasons and years, and yields information about the life-time (memory) of extreme yet realistic soil moisture perturbations. Our analyses were focused on root zone soil moisture (RSM) as it comprises the part of the soil that directly interacts with the atmosphere via bare soil evaporation and transpiration. We found that RSM memory differs not only spatially but also depends on the time of initialization. Long memory up to one year is evident mostly for dry soil moisture regimes, after heavy precipitation periods or prior to snow covered conditions. Short memory below two weeks prevails in wet soil moisture regimes and prior to distinct precipitation periods or snow melt. Furthermore, RSM perturbations affect other land surface states, e.g. soil temperature and leaf carbon content, and even induce anomalies with specific memory in these variables. Especially for deep layer soil temperature these anomalies can last up to several years. As long as RSM memory is evident, we found that anomalies occur periodically in other land surface states whenever climate conditions allow for interactions between that state and RSM. Additionally, anomaly recurrence is visible for RSM itself. This recurrence is related to the thickness of the soil layer below the root zone and can affect RSM for several years. From our findings we conclude that soil moisture initialization has the potential to improve the predictive skill of climate models on seasonal scales and beyond. However, a sophisticated, multi-layered soil hydrology scheme is necessary, to allow for the interactions between RSM and the deep soil layer reservoir.

Further information about this study are published in Stacke and Hagemann (2015).

Revision of planned simulations

While being a superior method which does not violate the water and energy balances, TAF has the shortcoming that it cannot utilize the most recent observation data for the initialization of predictions but enforces a certain time lag. Thus, a comparison of an ideal prediction ensemble (which is not possible with TAF) with the realistic prediction ensemble (which can be done with TAF) was planned to learn about the decrease in prediction skill due to the enforced time lag. However, this experiment became infeasible due to severe cuts in our requested computing time.

Instead, we utilized our granted resources for different simulations dedicated to improve process representation on the land surface. Furthermore, an ensemble simulation using optimized soil parameters derived from different satellite data sets is currently prepared and will be conducted soon. Please note, that this ensemble is still part of PastLand I which was extended until the end of the 2015.

References

Stacke, T. and S. Hagemann (2015). Life time of soil moisture perturbations in a coupled land-atmosphere simulation. *Earth Syst. Dynam. Discuss.* 6(2), pp. 1743–1788. ISSN: 2190-4995. DOI: 10.5194/esdd-6-1743-2015 (cit. on p. 1).