SPECIAL PROJECT PROGRESS REPORT

Reporting year	2020/21
Project Title:	Mining 5th generation reanalysis data for changes in the global energy cycle and for estimation of forecast uncertainty growth with generative adversarial networks
Computer Project Account:	spatlh00
Principal Investigator(s):	Leopold Haimberger
Affiliation:	University of Vienna
Name of ECMWF scientist(s) collaborating to the project (if applicable)	Hans Hersbach, M. Balmaseda
Start date of the project:	1.1.2021
Expected end date:	31.12.2023

Computer resources allocated/used for the current year and the previous one

		Previous year		Current year	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	-	-	10000	0.0
Data storage capacity	(Gbytes)	-	-	1000	32

Summary of project objectives

The special project focuses on detecting and estimating changes in the coupled oceanic/atmospheric water and energy cycles. Its second focus, also employing reanalysis data, are novel methods for describing forecast uncertainty growth.

Summary of problems encountered (if any)

Summary of results of the current year (from July of previous year to June of current year)

Within a new Austrian Science funds project (start in Jan 2020) we are continuing our work on calculating high accuracy budgets from ERA5. The budgets are now available from 1985-2020. The working group participated in writing a review paper (von Schuckmann et al. 2021) and published results on the ERA5 energy and moisture budgets in some detail (Mayer J. et al. 2021a). Another paper on the quality of indirectly estimated surface energy fluxes from ERA5 (Mayer J et al. 2021b) will be submitted in June 2021. Work on Arctic budgets led to a comprehensive study on the Arctic hydrological cycle, from indirect P-E estimates over Arctic river catchments to water storage estimates from GRACE and freshwater flow through Arctic ocean gateways (Winkelbauer et al. 2021, submitted).

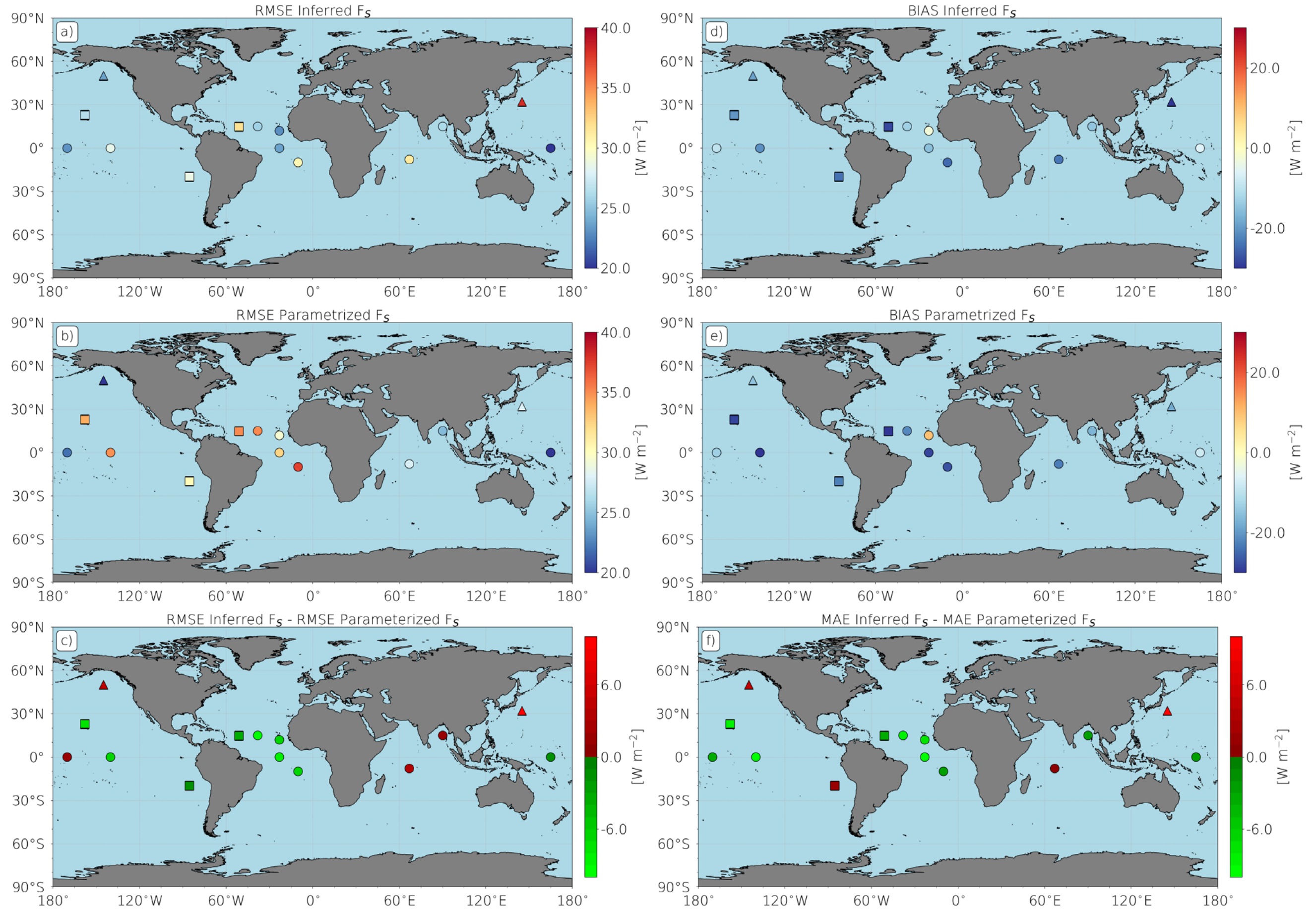

Fig. 1:Differences between estimated surface energy balance (Fs) at ocean buoy sites minus the observed values (actually fluxes calculated with bulk formulae) at the buoys. Fs estimates are (a,c) diagnostic from CERES TOA fluxes and ERA5 energy flux divergence, (b,d) from ERA5 parameterizations. Panels c), f) show that RMSE and MAE deviations from the buoy observations are slightly smaller for the diagnostic estimates overall.

Fig. 1 summarizes a comparison of indirectly and parameterized net surface energy fluxes at selected buoy sites. Correlations on a monthly timescale are quite high, with indirectly estimated surface energy fluxes outperforming the parameterized ones. Still, there are buoys where parameterized fluxes perform better, and there is quite a strong bias of estimated fluxes vs. the fluxes from buoys, which need to be better understood.

Fig. 2 depicts the major water flows through the Arctic climate system. Fluxes have been made consistent using a variational method, but imbalances are generally quite low, thus also the uncertainties are usually in the sub-5% range, except for the storage terms. Analysis has shown, however, that the data sets must be chosen really carefully, and results can strongly depend on choices of which pixel to use for getting estimates of major river outflows or which ocean pixels to include to estimate Greenland ice loss from GRACE.

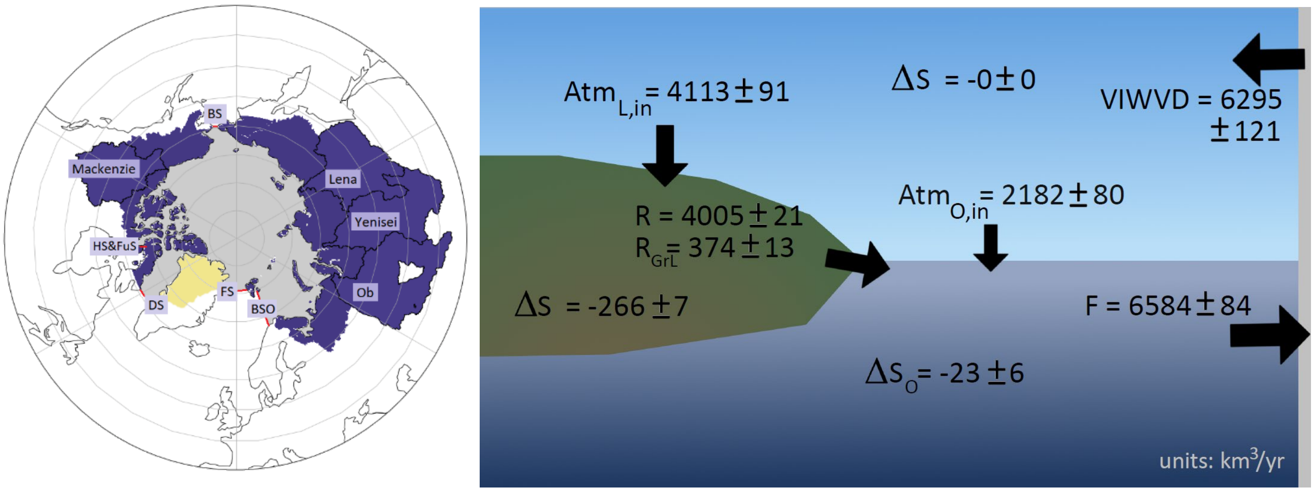

Fig. 2:Study region and water budget estimates from Winkelbauer et al. (2021). DS=storage in ocean, land surface (including glaciers) and atmosphere. VIWVD=Vertically integrated water vapor flux divergence averaged over Arctic region, $Atm_{L,in}$=P-E over the Arctic land mass, $Atm_{O,in}$=P-E over ocean, R, R_{Grl} are runoff estimates for the whole region and Greenland, respectively, F=water volume flux through ocean gateways estimated from ocean reanalyses and moorings. Uncertainties estimated from ensemble of calculations based on different reanalyses and different observation data sets. Further details can be found in the manuscript.

The Copernicus service contract C3S 311c Lot2 (early upper air data), devoted to developing a data portal that allows to extract sub-daily upper air data together with bias adjustments and metadata information, is in its final phase. The service currently runs on a CDS development server but should become operational very soon.

We have studied the use of both data-only and differential equations-based machine learning for solving meteorological forecasting problems. In particular, we have trained a probabilistic conditional generative adversarial network against geopotential, temperature, wind and specific humidity data with the goal to predict these quantities for the next 4 days. Owing to the probabilistic nature of the model, we can also use the model for estimating the uncertainty in its prediction. This provides a potential alternative for ensemble forecasting. We are currently working at a calibration of the model against ECMWF ensemble data.

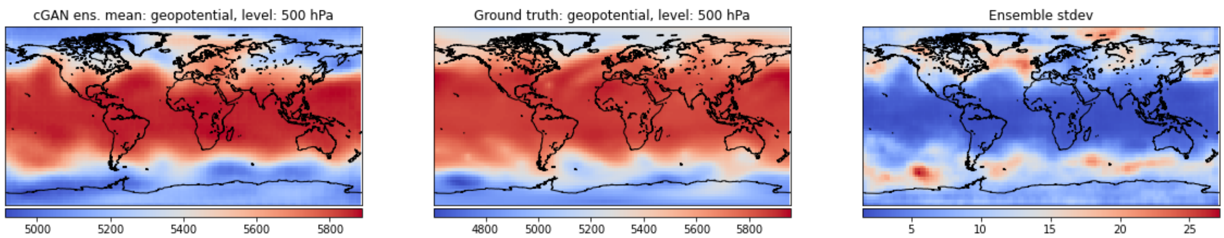

Fig. 3: Day-2 500 hPa geopotential forecast using a probabilistic conditional generative adversarial network. Left: ML prediction, middle: Ground truth, right: Estimated uncertainty in model forecast. The model correctly predicts that its forecast over the North Atlantic is uncertain, thus correctly identifying its own shortcomings.

We have also implemented and tested a machine learning model for solving the shallow-water equations on the sphere. This model does not yet take advantage of ECWMF model data, but has been shown to correctly reproduce standard (analytical) benchmarks for the shallow-water equations. The combination with model data will be pursued in the near future.

List of publications/reports from the project with complete references

Bihlo, A., 2021: A generative adversarial model for global ensemble weather forecasting. *In preparation*

Marquet, P. (2015). On the computation of moist-air specific thermal enthalpy. *Quarterly Journal of the Royal Meteorological Society*, *141*(686), 67-84.

Mayer, J., M. Mayer and L. Haimberger, 2021: Consistency and Homogeneity of Atmospheric Energy, Moisture, and Mass Budgets in ERA5. J. Climate **34**, 3955-3974, https://doi.org/10.1175/JCLI-D-20-0676.1

Mayer, J., Mayer, M., Haimberger, L., 2021: Comparison of Surface Energy Fluxes from Global to Local Scale. *To be submitted*

von Schuckmann, K., Cheng, L., Palmer, M. D., Tassone, C., Aich, V., Adusumilli, S., Beltrami, H., Boyer, T., Cuesta-Valero, F. J., Desbruyères, D., Domingues, C., García-García, A., Gentine, P., Gilson, J., Gorfer, M., Haimberger, L., Ishii, M., Johnson, G. C., Killik, R., King, B. A., Kirchengast, G., Kolodziejczyk, N., Lyman, J., Marzeion, B., Mayer, M., Monier, M., Monselesan, D. P., Purkey, S., Roemmich, D., Schweiger, A., Seneviratne, S. I., Shepherd, A., Slater, D. A., Steiner, A. K., Straneo, F., Timmermans, M.-L., and Wijffels, S. E., 2020: Heat stored in the Earth system: Where does the energy go? The GCOS Earth heat inventory team, Earth Syst. Sci. Data **12**, 2013-2041, https://doi.org/10.5194/essd-12-2013-2020.

Winkelbauer, S, Mayer, M, Seitner, V, Zsoter, E, Zuo, H & Haimberger, L 2021, 'Diagnostic evaluation of river discharge into the Arctic Ocean and its impact on oceanic volume transports', HESS, https://doi.org/10.5194/hess-2021-318

Summary of plans for the continuation of the project

The next focus for energy budget evaluations will be the quantification of changes in the Western Boundary Currents. We also plan to compare indirectly estimated surface fluxes with those from the Mosaic expedition. We also strive to improve the energy budget evaluation methods following the ideas of Marquet et al. (2015), and also to be fully consistent with the updated depiction of the energy budget within the IFS (to be included in chapter 12 of part IV of the documentation for CY47r3).

A bid for continuation of the Copernicus C3S-311c Lot2 service has been submitted with CNR as lead contractor.

We plan to continue the development of differential equations-based machine learning models for global weather forecasting.