Project: **854**

Project title: **Erdsystemmodellevaluierung (DLR-Institut für Physik der Atmosphäre)**

Principal investigator: **Axel Lauer**

Report period: **2025-11-01 to 2026-10-31**

1. ESMValTool development

New versions of the Earth System Model Evaluation Tool (ESMValTool) are developed and tested within project 854 before being released. Within the reporting period, ESMValTool v2.14.0 (March 2026) and v2.15.0 (July 2026) have been released. New features in the two versions include, for instance, improved support for COREDX-CMIP5 and newly added support for CORDEX-CMIP6 models, implementation of new CMIP7-REF recipes, support for intake-esgf and CMIP7 data. New diagnostics such as calculation of Lamb weather types, hour of maximum precipitation and a set of sanity and consistency checks tailored to model development have been implemented, new observational products such as ESA CCI snow, EN4 and new parameters from ESA CCI aerosol are now available.

The CMIP7-REF diagnostics are documented in:

Hoffman, F. M., Hassler, B., Swaminathan, R., Lewis, J., Andela, B., Collier, N., Hegedűs, D., Lee, J., Pascoe, C., Pflüger, M., Stockhause, M., Ullrich, P., Xu, M., Bock, L., Chun, F., Gier, B. K., Kelley, D. I., Lauer, A., Lenhardt, J., Schlund, M., Sreeush, M. G., Weigel, K., Blockley, E., Beadling, R., Beucher, R., Dugassa, D. D., Lembo, V., Lu, J., Brands, S., Tjiputra, J., Malinina, E., Medeiros, B., Scoccimarro, E., Walton, J., Kershaw, P., Lanfer Marquez, A., Roberts, M. J., O'Rourke, E., Dingley, B., Turner, B., Hewitt, H., and Dunne, J. P.: Rapid Evaluation Framework for the CMIP7 Assessment Fast Track, Geosci. Model Dev., 19, 7415-7455, doi: 10.5194/gmd-19-7415-2026, 2026.

Some examples of the new sanity and physical consistency checks are given in the next section. The entire set of diagnostics is documented in:

Lauer, A., Schlund, M., Bock, L., Hassler, B., Behrens, G., Gier, B., Lindenlaub, L., Lorenz, S., Malles, J.-H., Müller, W. A., van Pham, T., Weigel, K., Zeng, G., and Eyring, V.: An ESMValTool-based framework for sanity checks, physical consistency and climate fidelity during model development - ICONEval v1.0, Geosci. Model Dev., 19, 8731-8754, doi: 10.5194/gmd-19-8731-2026, 2026.

2. Analysis and evaluation with ESMValTool

Continuous evaluation and performance monitoring of Earth System Models (ESMs) is critical, particularly for emerging hybrid machine-learning-enhanced models, to detect unrealistic behavior, insufficient climate reproduction, or violations of physical laws early in the development process. For this, ICONEval has been developed, an open-source evaluation framework built on ESMValTool that complements existing community diagnostics with rapid, reproducible, and physically informed assessments suitable for use during model development. ICONEval enables efficient parallel processing of ESMValTool recipes and automatically generates HTML summary reports for streamlined performance monitoring. Its diagnostics are organized into three complementary categories: (1) sanity checks of global mean values against observational bounds, (2) physical consistency checks (e.g., conservation of total air mass, realistic water-vapor–temperature dependence), and (3) climate fidelity diagnostics across atmospheric, oceanic, and land components. The framework is demonstrated on a historical ICON-XPP simulation and, while focused on the ICON model, its underlying ESMValTool diagnostics remain applicable to other models compliant with the CMOR output standard.

With sanity checks, the global representation of selected variables in an ESM simulation, such as the temporal evolution of global means, minima, and maxima can be assessed. "Reasonable" upper and lower limits for global means are defined using the global minimum and maximum monthly values across all available years of multiple observational and reanalysis datasets. Categories implemented include radiation

and energy, moisture and precipitation, clouds, and temperature and wind stress.

Physical consistency checks ensure that the model's equations, parameterizations, and numerical methods correctly represent known fundamental laws of physics. This is of importance as simplified representations of subgrid processes such as newly developed data driven parameterizations of convection or cloud microphysics might violate some physical constraints. Complying with basic physical consistency checks can be seen as a necessary even though not sufficient precondition to allow for realistic future projections. The currently implemented consistency checks are focused on atmospheric parameters and include the categories mass conservation, temperature dependence of water vapor, cloud ice fraction, and tropospheric lapse rates. As an example, the tropospheric lapse rate diagnostic is illustrated with an example ICON-XPP run in Figure 1. On average, lapse rates in the tropical troposphere above the boundary layer are expected to be smaller than in mid- or high-latitudes. In the example shown, the peak of the lapse rate PDFs are around 5 K km^{-1} in the Tropics and 6.5 K km^{-1} in mid-latitudes.

In order to ensure that the basic characteristics of the atmosphere, the ocean and the land surface are represented correctly in a model, the simulations are compared to observational data records. Diagnostics implemented for this basic check include calculation of climatologies, annual and diurnal cycles, geographical distributions and time series of different dataset aggregations. The diagnostics are flexible enough to allow their use for many different variables without the need to adjust anything in the code.

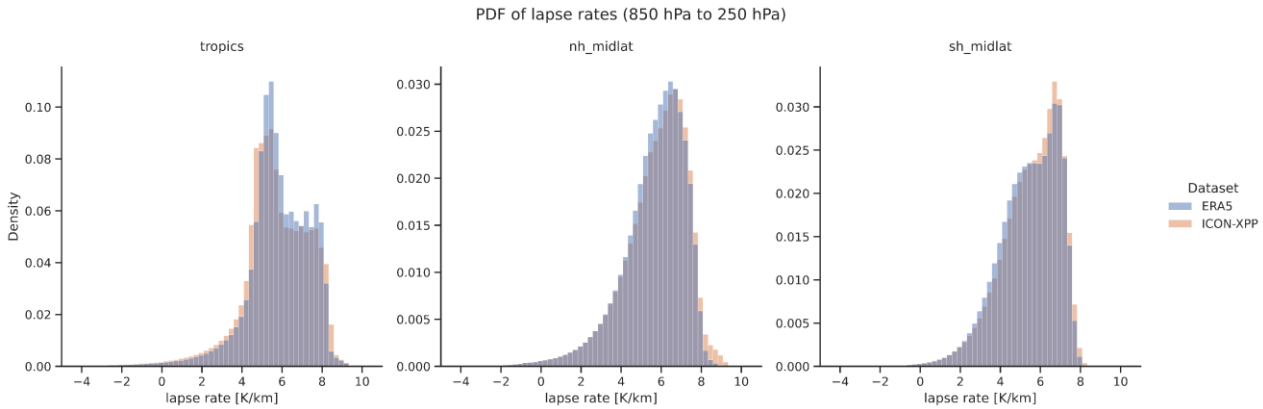


Figure 1 Probability density functions of the 3D lapse rates from ERA5 (blue) and ICON-XPP (orange) in the altitude range 850 to 250 hPa calculated from monthly mean temperatures over the time period 1995-2014. From left to right: Tropics (30°S-30°N), Northern Hemisphere mid-latitudes (40-60°N) and Southern Hemisphere mid-latitudes (60-40°S). From Lauer et al. (2026).

Publications in 2026 related to project 854

Bock, L., A. Lauer, and J. Runge: Quantifying the Causal Effect of Cloud-Controlling Factors on Marine Stratocumulus Clouds, J. Atmos. Sci., 83, 255-264, doi: 10.1175/JAS-D-25-0153.1, 2026.

Hoffman, F. M., Hassler, B., Swaminathan, R., Lewis, J., Andela, B., Collier, N., Hegedűs, D., Lee, J., Pascoe, C., Pflüger, M., Stockhause, M., Ullrich, P., Xu, M., Bock, L., Chun, F., Gier, B. K., Kelley, D. I., Lauer, A., Lenhardt, J., Schlund, M., Sreeush, M. G., Weigel, K., Blockley, E., Beadling, R., Beucher, R., Dugassa, D. D., Lembo, V., Lu, J., Brands, S., Tjiputra, J., Malinina, E., Medeiros, B., Scoccimarro, E., Walton, J., Kershaw, P., Lanfer Marquez, A., Roberts, M. J., O'Rourke, E., Dingley, B., Turner, B., Hewitt, H., and Dunne, J. P.: Rapid Evaluation Framework for the CMIP7 Assessment Fast Track, Geosci. Model Dev., 19, 7415-7455, doi: 10.5194/gmd-19-7415-2026, 2026.

Lauer, A., Schlund, M., Bock, L., Hassler, B., Behrens, G., Gier, B., Lindenlaub, L., Lorenz, S., Malles, J.-H., Müller, W. A., van Pham, T., Weigel, K., Zeng, G., and Eyring, V.: An ESMValTool-based framework for sanity checks, physical consistency and climate fidelity during model development - ICONEval v1.0, Geosci. Model Dev., 19, 8731-8754, doi: 10.5194/gmd-19-8731-2026, 2026.