GRACE-FO Science Team Meeting 2025

WROCŁAW UNIVERSITY OF ENVIRONMENTAL AND LIFE SCIENCES

Validating OS LISFLOOD with SLR: Three Decades of Hydrological Variability

Filip Gałdyn[1], Laura Jensen[3], Radosław Zajdel[2,1], Krzysztof Sośnica[1], Henryk Dobslaw[3]

[1] Wrocław University of Environmental and Life Sciences, Institute of Geodesy and Geoinformatics, Wrocław, Poland
[2] Research Institute of Geodesy, Topography and Cartography, Geodetic Observatory Pecný, Czechia
[3] GFZ Helmholtz Centre for Geosciences, Potsdam, Germany

Research Objectives

1. Evaluate hydrological model performance
Assess and compare the ability of OS LISFLOOD and LSDM to reproduce TWS variability in the world’s largest river basins, including seasonal, interannual, and long-term signals.

2. Analyze spatial and climatic patterns
Investigate how model performance varies across latitudinal and hydroclimatic zones and identify systematic strengths and weaknesses of each model.

3. Validate and extend reference datasets
Examine the consistency and applicability of satellite-based combinations (SLR+GRACE, SLR+DORIS) as independent validation datasets in the pre-GRACE era

OS LISFLOOD

Development and Purpose

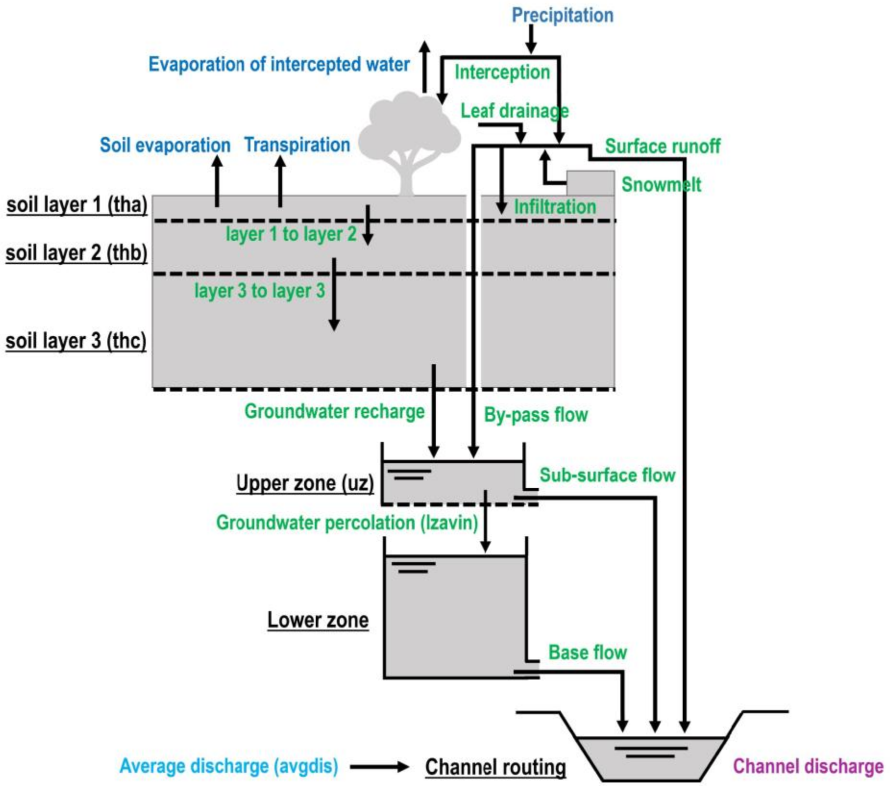
- Developed by the JRC (De Roo et al., 2000; Van Der Knijff et al. 2010)
- Operational core of EFAS/GloFAS
- Continuously maintained, open-source, available on GitHub

Key Features

- Full hydrological cycle
- High resolution: 0.05° (~5 km) daily data (1960–now)
- Meteorological forcing – ERA5 reanalysis (ECMWF)
- Includes anthropogenic water use (LSDM not)

Applications

- Simulation of Terrestrial Water Storage (TWS) variability
- Superior to LSDM (Dill, 2008) in capturing interannual signals (Jensen et al., 2025)



Overview of the main processes included in OS LISFLOOD. The scheme is adapted from https://ec-jrc.github.io/lisflood-model/2_01_stdLISFLOOD_overview/ (last visited 24/9/2025).

External validation data

SLR only (10x10)
(Gałdyn et al. 2024)

- data from 1995 to December 2023 based on 8 geodetic satellites
- splitting and re-stacking NEQ
- published on ICGEM

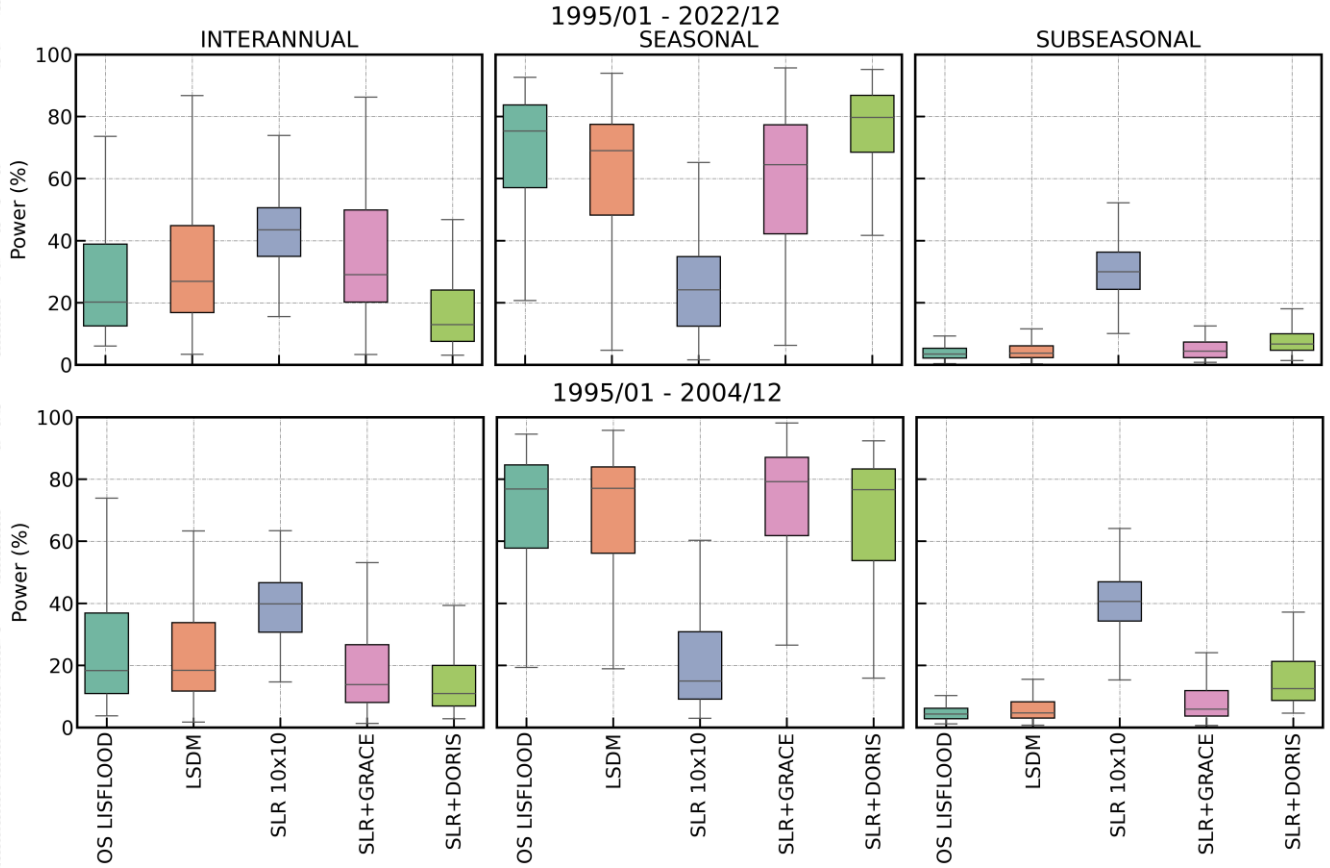
SLR + GRACE (60x60)
(based on Gałdyn et al. 2024)

- data from 1995 to December 2023
- different modeling approaches for each degree range, including fitting annual/semiannual signals, stochastic pulses and extrapolation backwards

SLR + DORIS (60x60)
(Löcher et al. 2025)

- data from 1984 to December 2023.
- use of empirical orthogonal functions (EOFs) from GRACE and GRACE-FO
- combination SLR and DORIS observations

Model band characteristics

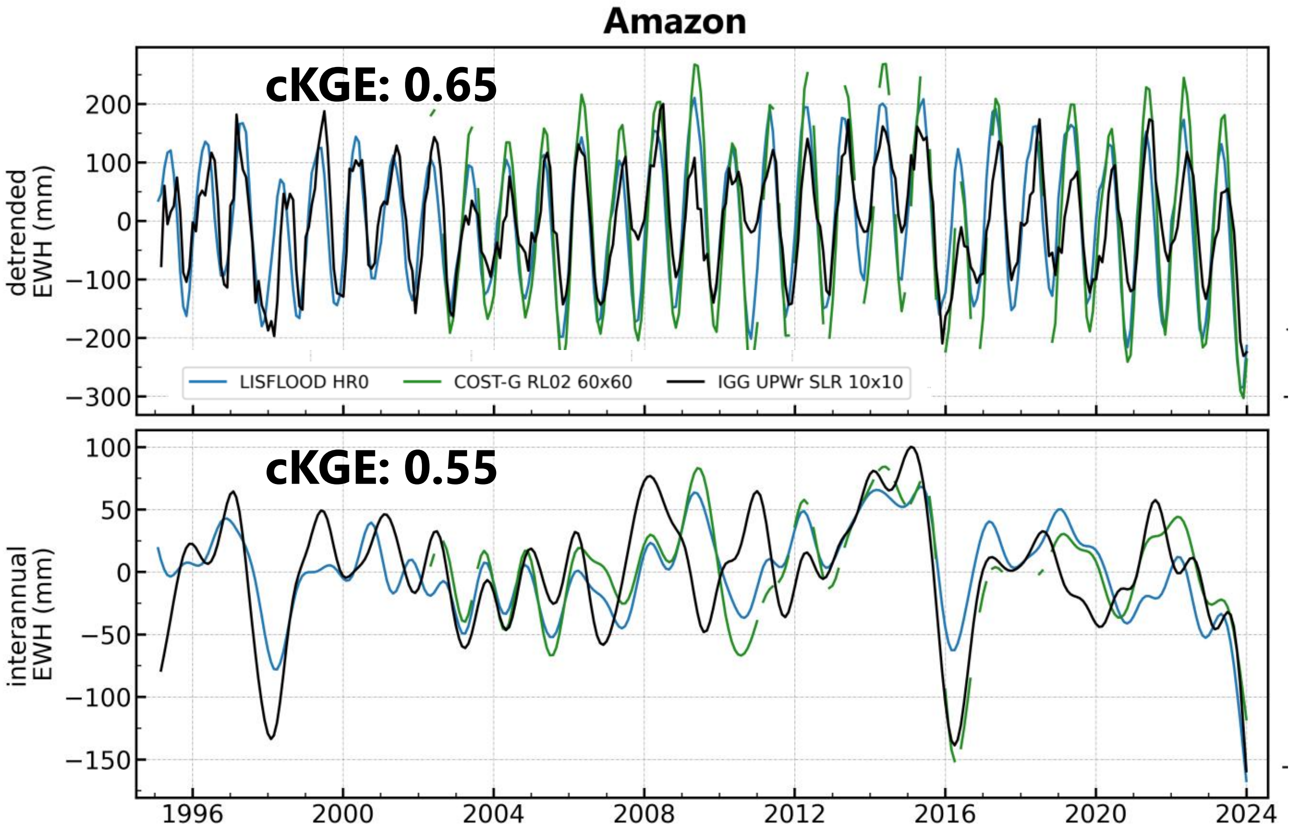


Share of signal power (%) by temporal band for monthly TWS anomalies in 100 largest river basins. Boxplots summarize results for five datasets: OS LISFLOOD, LSDM, SLR 10×10 (SLR-only with, Gaussian 300km filter applied), SLR+GRACE (with DDK3 filter), and SLR+DORIS (with Gaussian 300 km).

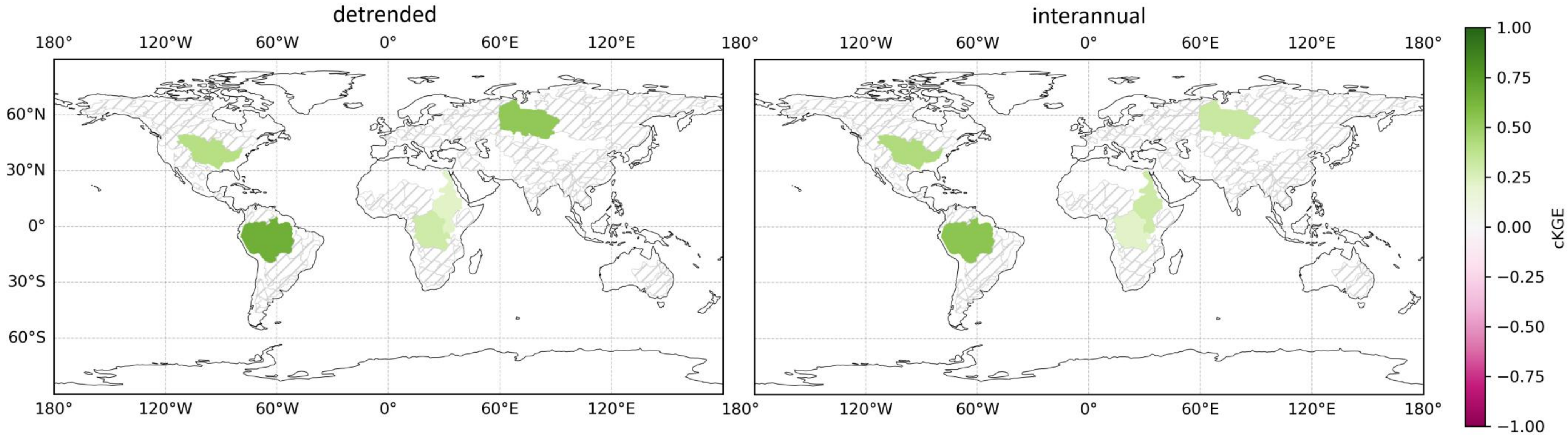
- OS LISFLOOD and LSDM are predominantly seasonal, with limited subseasonal variability.
- The SLR-only solution (10×10) exhibits increased subseasonal power and reduced seasonal dominance due to draconitic/orbital aliasing and low spatial resolution.
- SLR+GRACE, SLR+DORIS improve signal-to-noise ratio by rebalancing the spectrum towards the seasonal band, reducing subseasonal variability, and enhancing spatial resolution.
- Patterns stable across both periods, confirming the robustness of the conclusions regardless of the record length.

OS LISFLOOD performance vs SLR only (1995-2024)

- SLR lower spatial resolution limits amplitude accuracy, but the comparison is focused on 5 the largest river basins to mitigate these effects.
- In Amazon river basin variablity is comparable to OS LISFLOOD especially in pre GRACE era for detrended signal and interannual

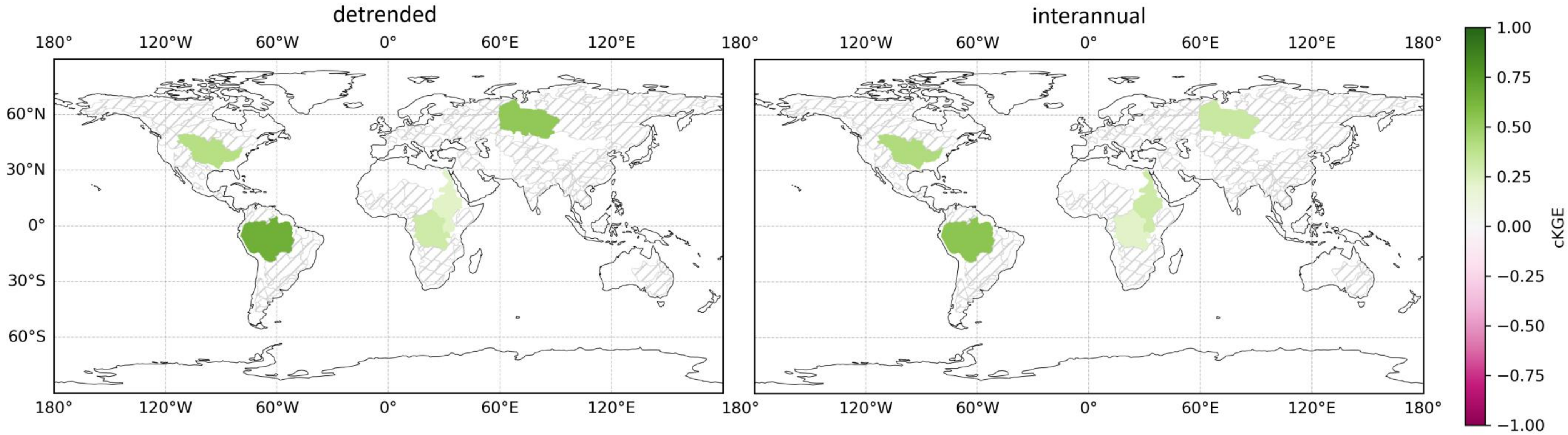


OS LISFLOOD performance vs SLR only (1995-2024)



- The centered modified Kling-Gupta-Efficiency (cKGE) shows positive values for **all basins**, indicating good agreement with SLR data, with slightly lower performance for interannual signals

OS LISFLOOD performance vs SLR only (1995-2024)

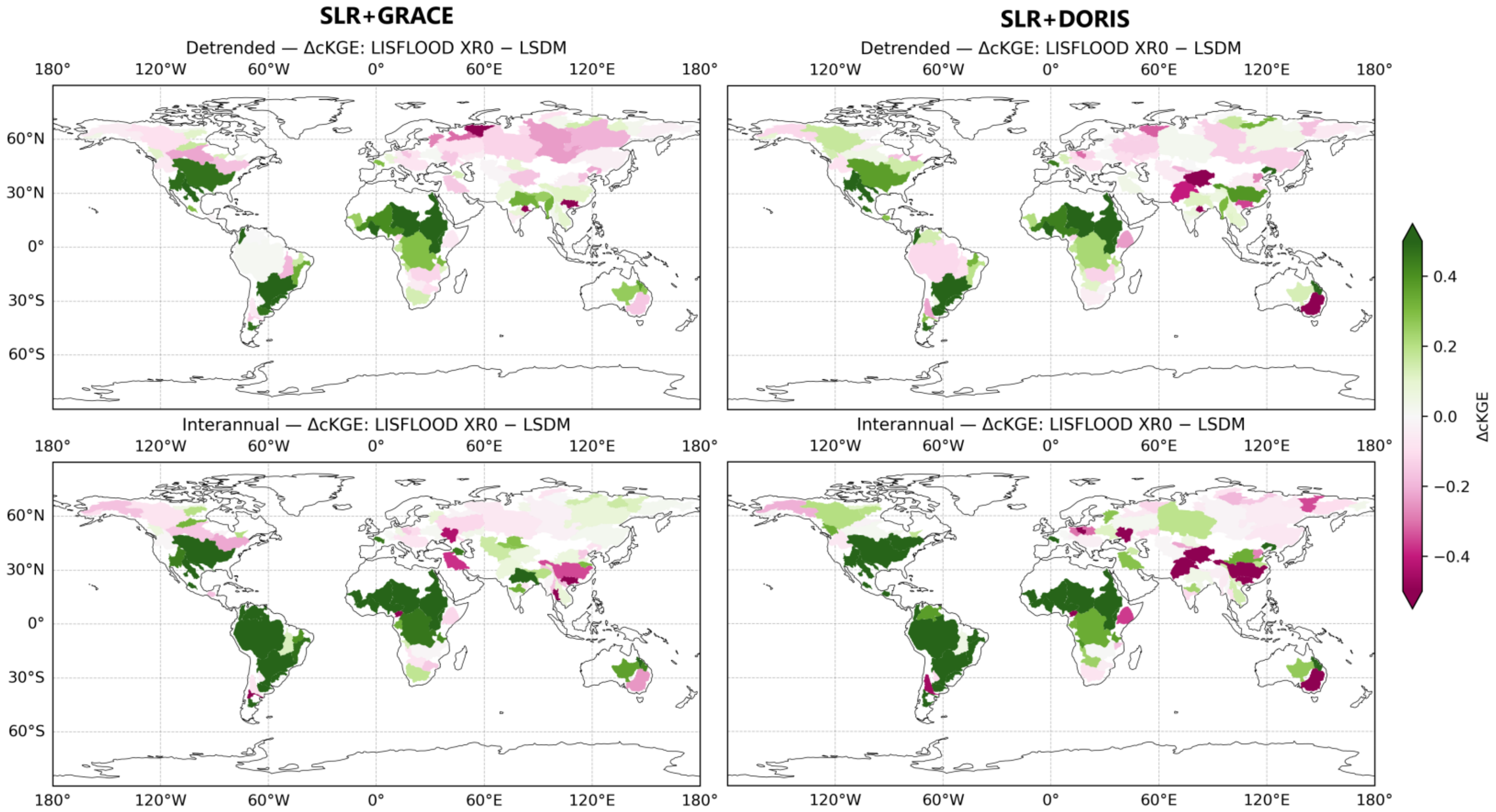


- The centered modified Kling-Gupta-Efficiency (cKGE) shows positive values for **all basins**, indicating good agreement with SLR data, with slightly lower performance for interannual signals

SLR-only is consistent with OS LISFLOOD, but for global validation, a model with higher spatial resolution <u>is needed</u>.

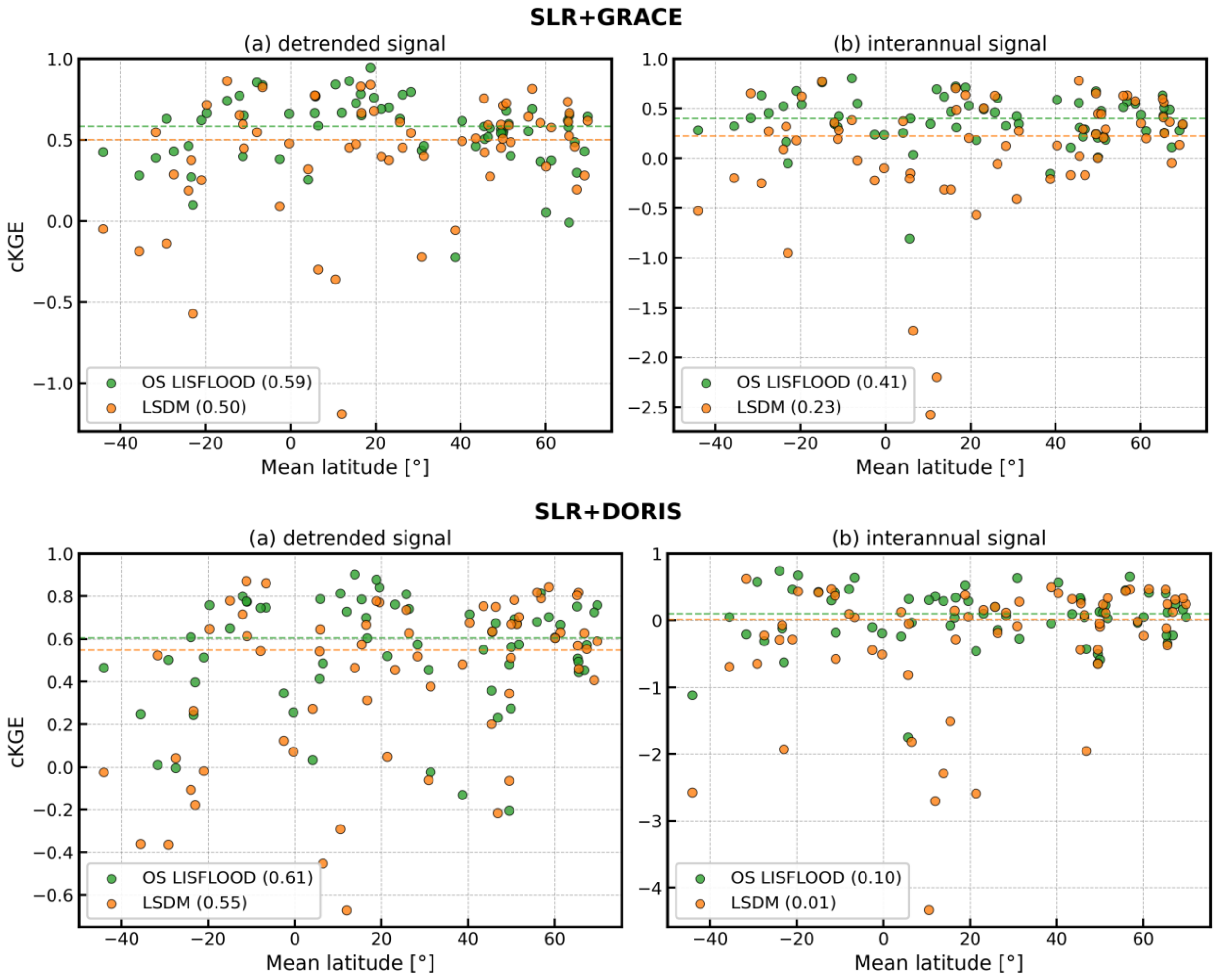
OS LISFLOOD vs LSDM (1995-2022)

- OS LISFLOOD outperforms LSDM in tropical and subtropical basins for detrended signals; LSDM remains competitive in Asia and high latitudes for interannual signals.
- Both SLR+GRACE and SLR+DORIS show consistent patterns, confirming OS LISFLOOD's robustness compared to LSDM



OS LISFLOOD vs LSDM (1995-2022)

- OS LISFLOOD explains more variability in detrended signals (cKGE 0.59) than LSDM (cKGE 0.50).
- For interannual signals, OS LISFLOOD maintains a cKGE of 0.41, nearly double that of LSDM (cKGE 0.23).

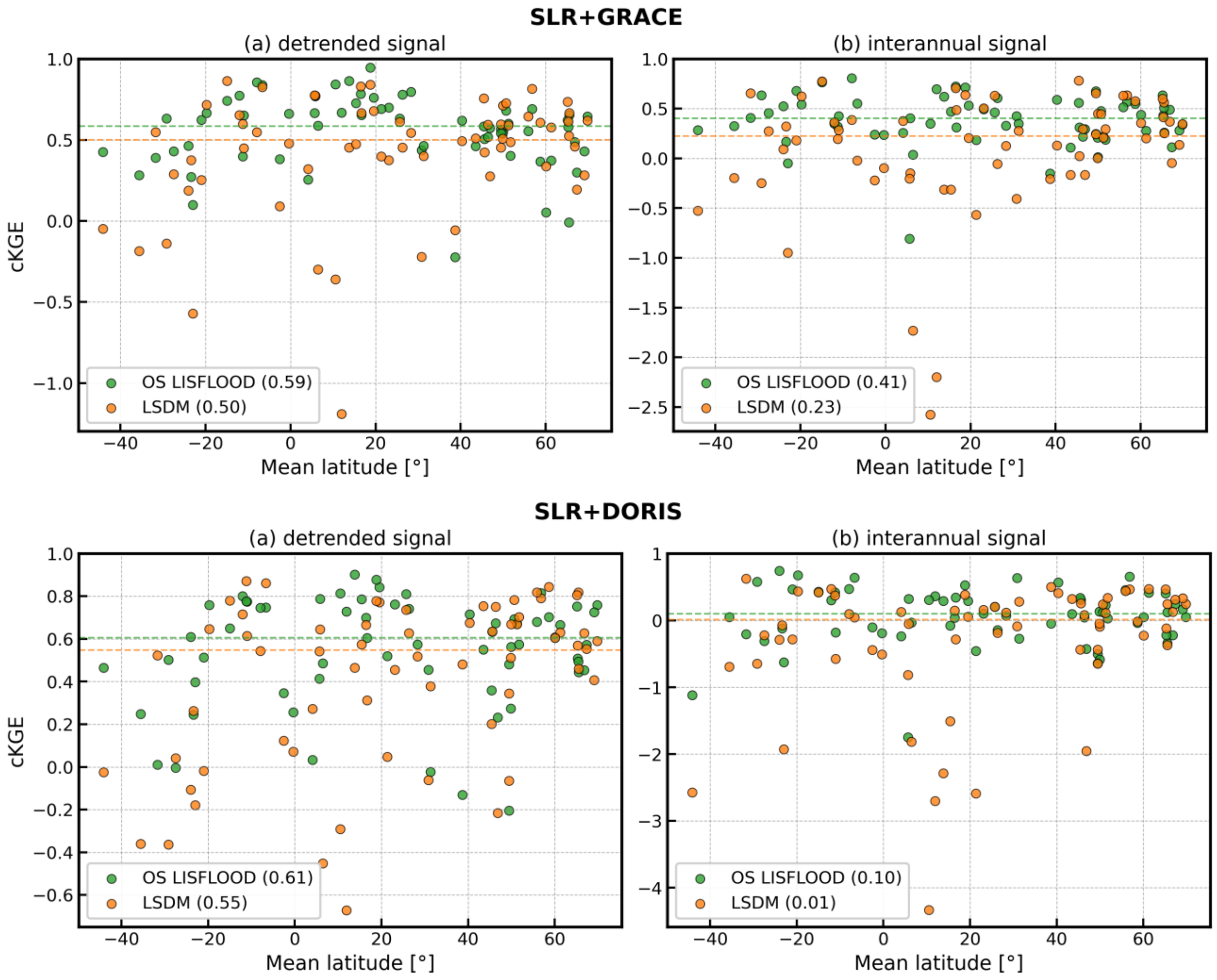


Scatter plot of cKGE for the (a) detrended signal and (b) interannual signal of the 100 largest river basins (sorted by mean latitude), evaluated against SLR+GRACE and SLR+DORIS. Green dots denote OS LISFLOOD and orange dots denote LSDM. Horizontal dashed lines indicate the median cKGE values across all basins (given in parentheses in the legend).

OS LISFLOOD vs LSDM (1995-2022)

- OS LISFLOOD explains more variability in detrended signals (cKGE 0.59) than LSDM (cKGE 0.50).
- For interannual signals, OS LISFLOOD maintains a cKGE of 0.41, nearly double that of LSDM (cKGE 0.23).

OS LISFLOOD is more consistent with gravimetric estimates, especially in large low-latitude basins.

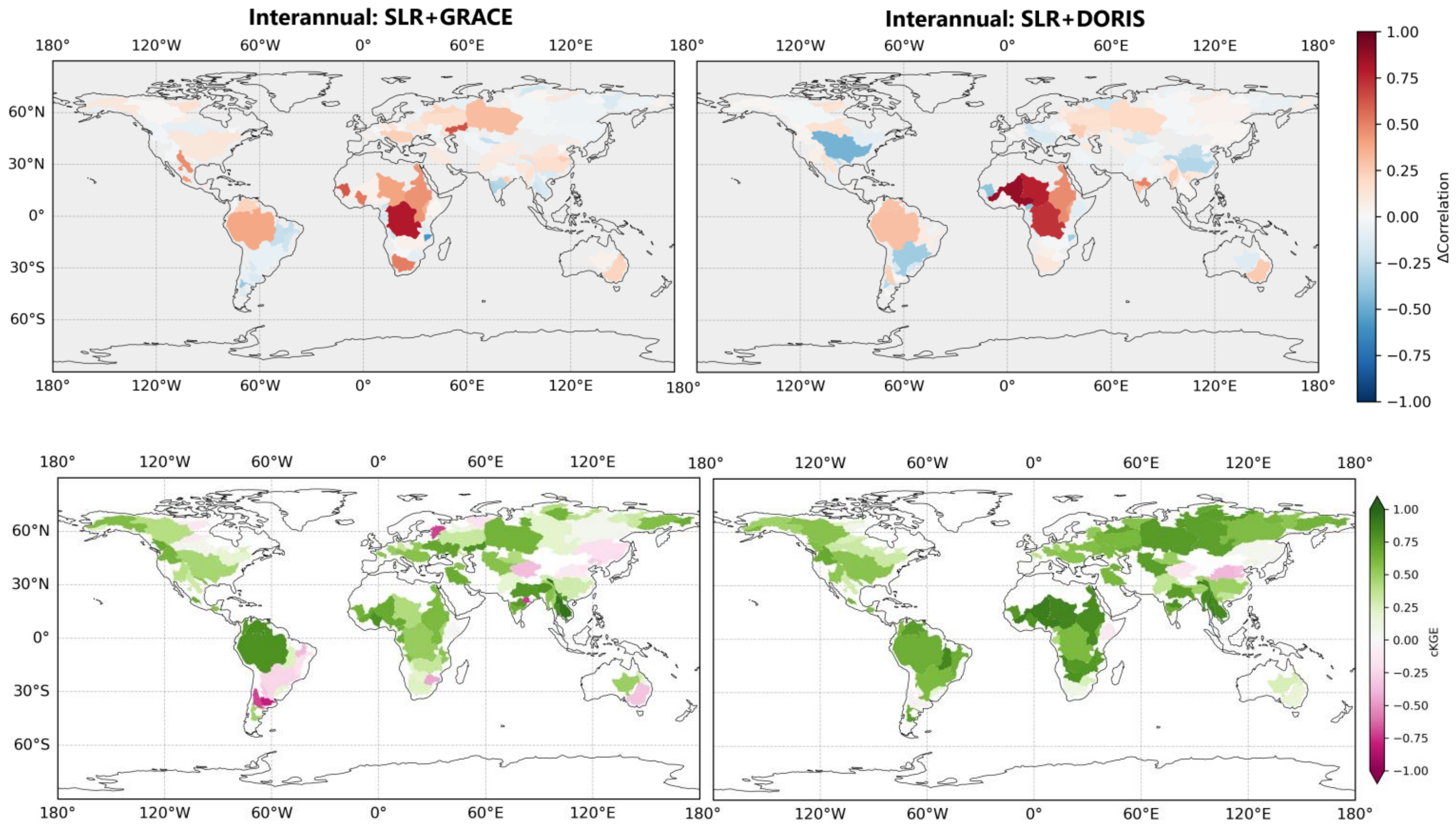


Scatter plot of cKGE for the (a) detrended signal and (b) interannual signal of the 100 largest river basins (sorted by mean latitude), evaluated against SLR+GRACE and SLR+DORIS. Green dots denote OS LISFLOOD and orange dots denote LSDM. Horizontal dashed lines indicate the median cKGE values across all basins (given in parentheses in the legend).

OS LISFLOOD vs LSDM (1995-2004)

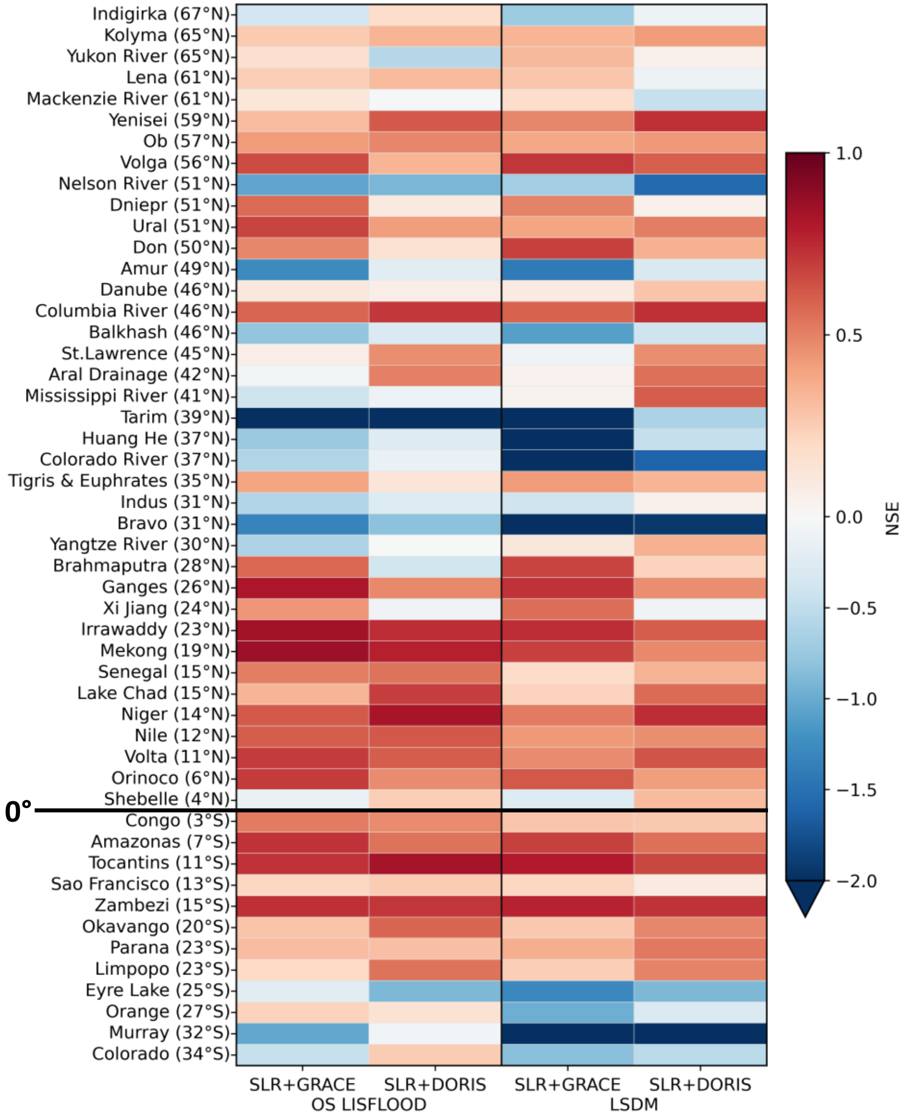
- OS LISFLOOD exhibits higher correlations in many tropical and subtropical basins (e.g., Amazon, Congo, Niger) when compared to LSDM, while mid- to high-latitude regions (e.g., Ob, Yenisei) show less advantage or negative correlations.

- Basin-averaged cKGE values indicate that OS LISFLOOD provides better consistency, especially in large tropical basins. Both models, however, exhibit limitations in cold and arid regions, highlighting the need for further improvements.



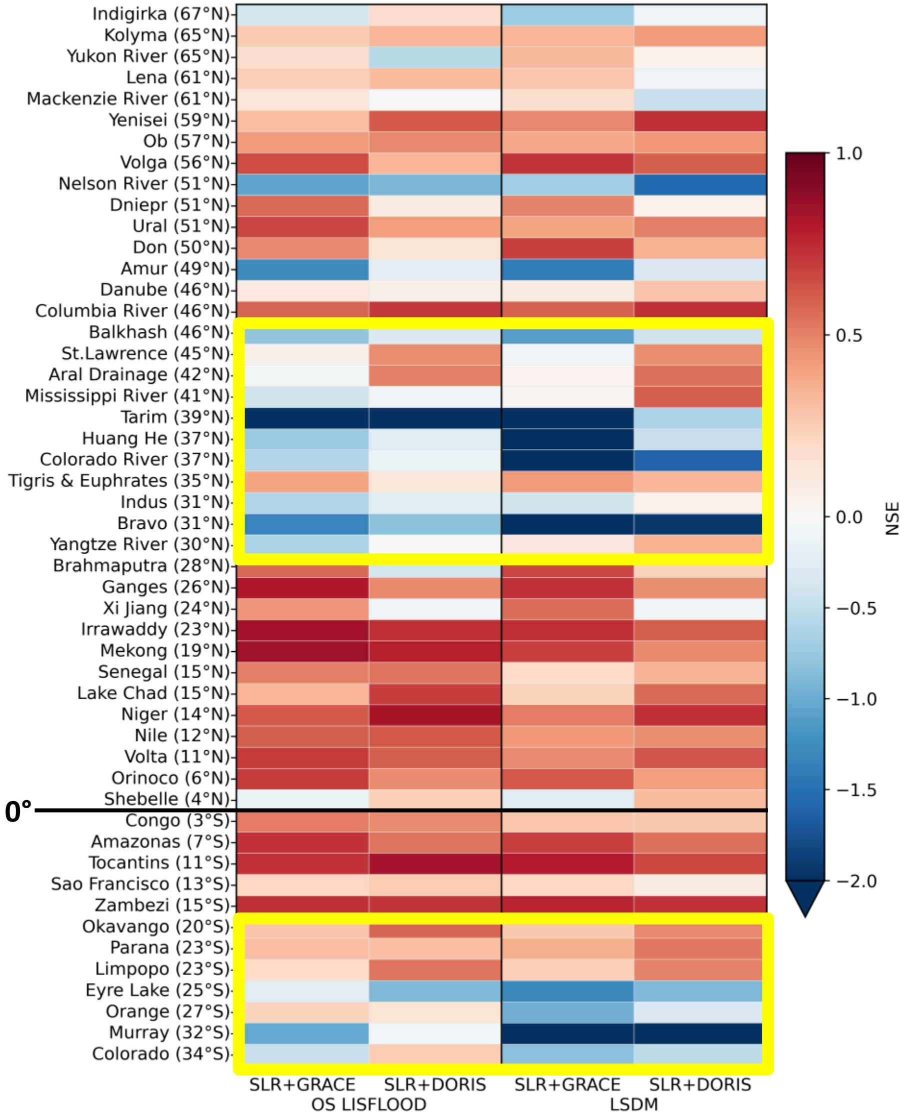
OS LISFLOOD vs LSDM (1995-2004)

- OS LISFLOOD outperforms LSDM in most tropical and subtropical basins, showing positive Nash–Sutcliffe Efficiency (NSE) when validated with SLR+GRACE and SLR+DORIS data.



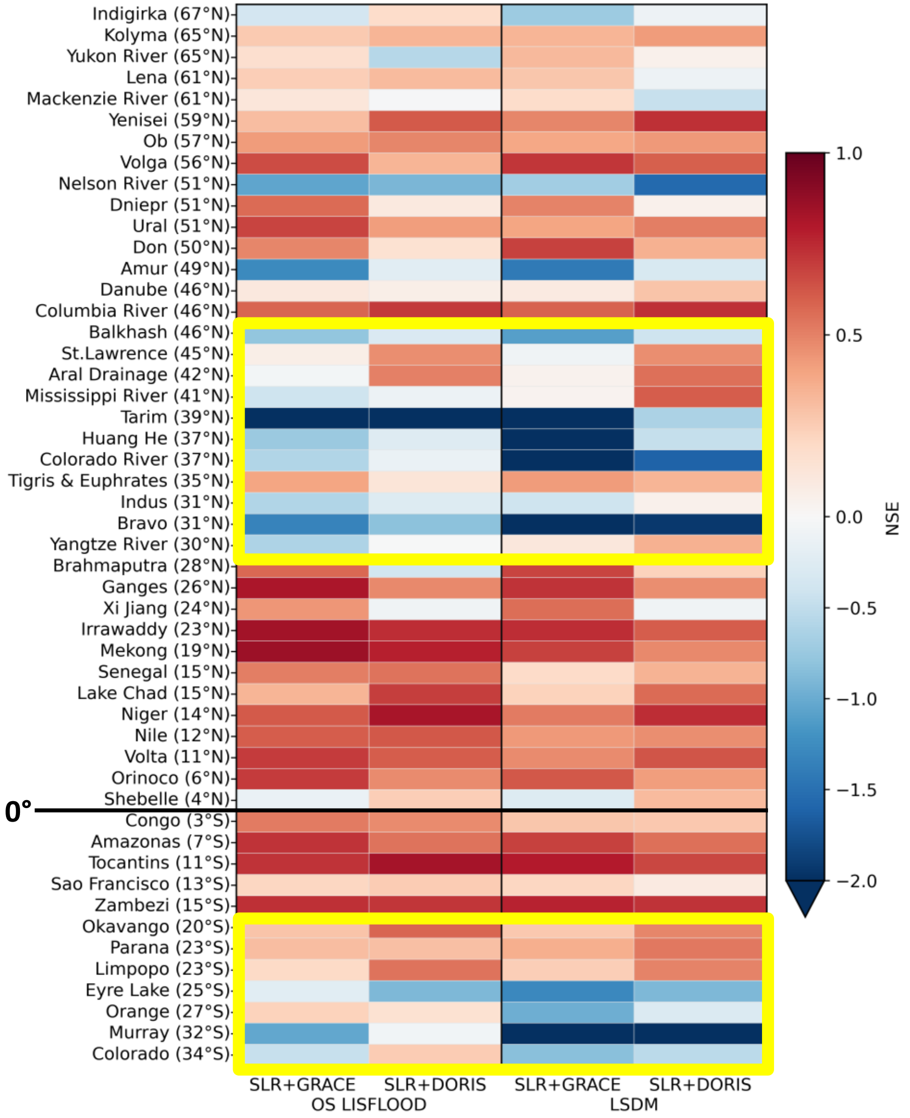
OS LISFLOOD vs LSDM (1995-2004)

- OS LISFLOOD outperforms LSDM in most tropical and subtropical basins, showing positive Nash–Sutcliffe Efficiency (NSE) when validated with SLR+GRACE and SLR+DORIS data.
- Both OS LISFLOOD and LSDM models show negative NSE in high-latitude and dry basins, indicating difficulties in capturing hydrological variability in cold and water-limited regions.

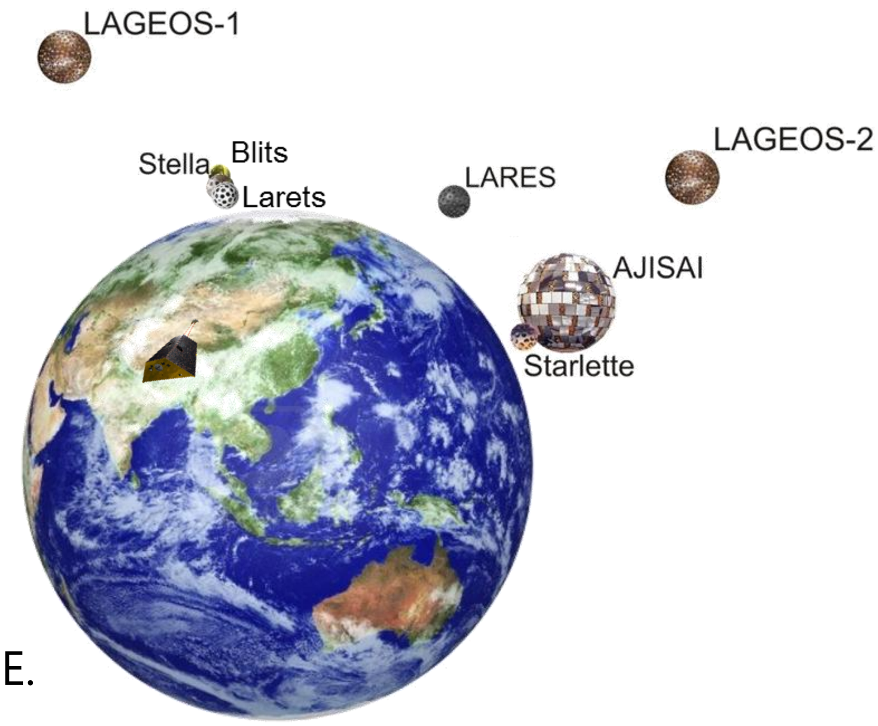


OS LISFLOOD vs LSDM (1995-2004)

- OS LISFLOOD outperforms LSDM in most tropical and subtropical basins, showing positive Nash–Sutcliffe Efficiency (NSE) when validated with SLR+GRACE and SLR+DORIS data.
- Both OS LISFLOOD and LSDM models show negative NSE in high-latitude and dry basins, indicating difficulties in capturing hydrological variability in cold and water-limited regions.
- The consistency between SLR+GRACE and SLR+DORIS models highlights the robustness of the OS LISFLOOD model, with improved performance in some regions compared to LSDM during the 1995–2004 period.



Conclusions



1. OS LISFLOOD outperforms LSDM
OS LISFLOOD better captures large-scale TWS variability, especially seasonal and interannual signals in tropical and subtropical basins.

2. Robust performance in the pre-GRACE era
Strong agreement with SLR+GRACE and SLR+DORIS confirms that OS LISFLOOD reliably reproduces interannual TWS variability before GRACE.

3. Geodetic and hydrological relevance
The improved consistency of OS LISFLOOD with independent gravimetric estimates highlights its value for geodetic applications and long-term water cycle reconstructions.

4. Regional limitations of both models (1995-2004)
At high latitudes and in arid regions, both models struggle to reproduce TWS variability, often yielding negative NSE values and underestimating amplitudes of interannual changes.

Thank you for your attention

Filip Gałdyn

filip.galdyn@upwr.edu.pl

Gałdyn, F., Sośnica, K., Zajdel, R, Mayer, U., Jäggi., A. (2024). Long-term ice mass changes in Greenland and Antarctica derived from satellite laser ranging. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2024.113994

Gałdyn, F., Jensen, L., Zajdel, R., Sośnica, K., Dobslaw, H., (2025). OS LISFLOOD hydrological model validation with SLR observations over nearly three decades (in preparation)

This study was supported by the National Science Centre, Poland (NCN), under Grant No. UMO-2024/53/N/ST10/00522, and by the Foundation for Polish Science (FNP)

References

De Roo, A. P. J., Wesseling, C. G., & Van Deursen, W. P. A. (2000). Physically based river basin modelling within a GIS: the LISFLOOD model. Hydrological Processes, 14(11–12), 1981–1992. https://doi.org/10.1002/1099-1085(20000815/30)14:11/12<1981::AID-HYP49>3.0.CO;2-F

Dill, R. (2008). Hydrological model LSDM for operational Earth rotation and gravity field variations: Vol. 08/09 (D. GeoForschungsZentrum (ed.)). https://doi.org/10.2312/GFZ.b103-08095

Gałdyn, F., Sośnica, K., Zajdel, R., Meyer, U., & Jäggi, A. (2024). Long-term ice mass changes in Greenland and Antarctica derived from satellite laser ranging. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2024.113994

Jensen, L., Dill, R., Balidakis, K., Grimaldi, S., Salamon, P., & Dobslaw, H. (2025). Global 0.05° water storage simulations with the OS LISFLOOD hydrological model for geodetic applications. *Geophysical Journal International*, *241*(3), 1840–1852. https://doi.org/10.1093/gji/ggaf129

Löcher, A., Kusche, J., & Nie, Y. (2025). A 40-year record of the Earth's time-variable gravity field from SLR and DORIS. Advances in Space Research. https://doi.org/10.1016/j.asr.2025.05.089

Van Der Knijff, J. M., Younis, J., & De Roo, A. P. J. (2010). LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. International Journal of Geographical Information Science, 24(2), 189–212. https://doi.org/10.1080/13658810802549154

Evaluation Metrics

For the quantitative evaluation of model–observation agreement we use the cKGE, i.e., the bias-insensitive variant of modified Kling-Gupta-Efficiency (KGE', Gupta et al., 2009; Kling et al., 2012). Prior to comparison, each time series is demeaned to remove constant offsets; thus, the mean-bias term $\beta = \mu_s/\mu_o$ is effectively set to 1. The resulting metric combines correlation and variability components,

$$\mathrm{cKGE} = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2}$$

where r is the Pearson correlation coefficient between simulated (s_i) and observed (o_i) anomalies,

$$r = \frac{\sum (s_i - \mu_s)(o_i - \mu_o)}{\sqrt{\sum (s_i - \mu_s)^2 \sum (o_i - \mu_o)^2}},$$

and α quantifies the variability ratio, $\alpha = \frac{\sigma_s}{\sigma_o}$.

With demeaning, the variability term of KGE' based on coefficients of variation (γ) reduces to the standard-deviation ratio (α), making the adopted formulation equivalent to KGE' with $\beta = 1$. The metric attains its optimum at $\mathrm{cKGE} = 1$, indicating perfect correlation and matched variability.

A widely used performance metric in hydrology is NSE (Nash & Sutcliffe, 1970) that measures the predictive skill of a model relative to the mean of observations. NSE is defined as:

$$NSE = 1 - \frac{\sum_i (o_i - s_i)^2}{\sum_i (o_i - \bar{o})^2},$$

where *oi* are the observed values, *si* the simulated values, and *o* An NSE of 1 corresponds to perfect agreement, values between 0 and 1 indicate that the model outperforms the mean of observations, while negative values imply that the mean of observations is a better predictor than the model.