Supplementary Information for the manuscript: Assessment of prediction skill of SEAS5 forecast using ERA5 soil moisture data and relation to crop production

Padmavathi Bevara and Strobach Ehud

Figures S1-S3

Figures S1, S2, and S3 present the first six normalized principal components (PCs) of ERA5 soil moisture anomalies and the corresponding EOF-projected SEAS5 anomalies for lead times 0–6 months. These figures demonstrate that the temporal evolution of the leading PCs in SEAS5 closely matches that of ERA5 across all lead times, which capture the dominant modes of short-term hydroclimatic variability. The strong agreement between ERA5 and projected SEAS5 PCs indicates that the reconstruction procedure successfully isolates physically meaningful, large-scale soil moisture signals from model noise. The first few PCs primarily represent high-frequency, synoptic-scale fluctuations and land–atmosphere coupling processes, hence their strong temporal coherence and amplitude agreement across datasets.

Figure S1 presents the PCs for the swvl1 (0–7 cm) soil layer. The PCs in this layer primarily represent high-frequency atmospheric fluctuations and surface-driven processes, and the close match between the datasets demonstrates that the reconstruction effectively filters noise while preserving physically meaningful variability.

Figure S2 shows the PCs for the swvl2 (7–28 cm) soil layer. In this layer, PC1, PC2, PC4, and PC68 exhibit particularly good correspondence between ERA5 and SEAS5 at shorter lead times. This reflects the fact that swvl2 retains both high-frequency atmospheric influences from the surface and lower-frequency memory contributions from deeper layers, resulting in a mixed temporal signature that SEAS5 is able to capture well through EOF projection. The close PC alignment also suggests that the reconstruction preserves intermediate-depth soil moisture processes such as evapotranspiration feedbacks, infiltration timing, and root-zone buffering.

Figure S3 illustrates the PCs for swvl3 (28–100 cm), where PC1, PC2, PC3, and PC5 show a strong match between ERA5 and reconstructed SEAS5. In deeper layers,

the temporal variability becomes smoother and dominated by lower-frequency, longer-memory components, as soil moisture integrates hydroclimatic anomalies over weeks to months. The reduced amplitude and slower oscillations in PCs at this depth reflect reduced sensitivity to short-lived atmospheric forcing and increased dependence on accumulated wetting and drying cycles. The strong SEAS5–ERA5 agreement in these lower-frequency modes demonstrates that the EOF reconstruction method effectively captures the persistent hydrological memory that characterizes subsurface soil layers.

Taken together, Figures S1 to S3 illustrate a systematic transition in the temporal characteristics of soil moisture variability with depth. In **swvl1**, PCs represent fast atmospheric fluctuations; in **swvl2**, intermediate memory processes emerge; and in **swvl3**, the PCs reflect long-term hydrological storage and multi-month persistence. The EOF–PC reconstruction leverages this structure by projecting SEAS5 anomalies onto the physically dominant ERA5 modes, thereby recovering meaningful variability even when raw SEAS5 fields are noisy or misaligned. This analysis demonstrates that principal-component comparison provides a powerful diagnostic tool for evaluating model fidelity, identifying predictable modes, and guiding reconstruction of soil moisture fields for improved seasonal forecast performance.

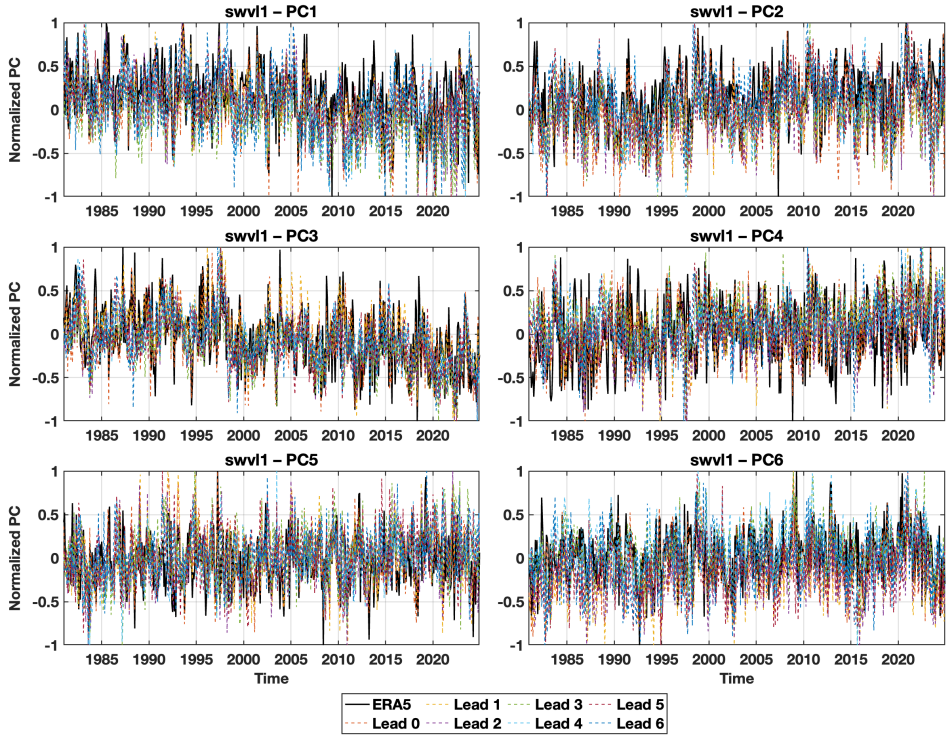

Fig. S1 Normalized PCs of ERA5 and reconstructed SEAS5 soil moisture anomalies for the first soil layer (swvl1). Figure S1 shows the first six PCs derived from ERA5 and their SEAS5 reconstructions.

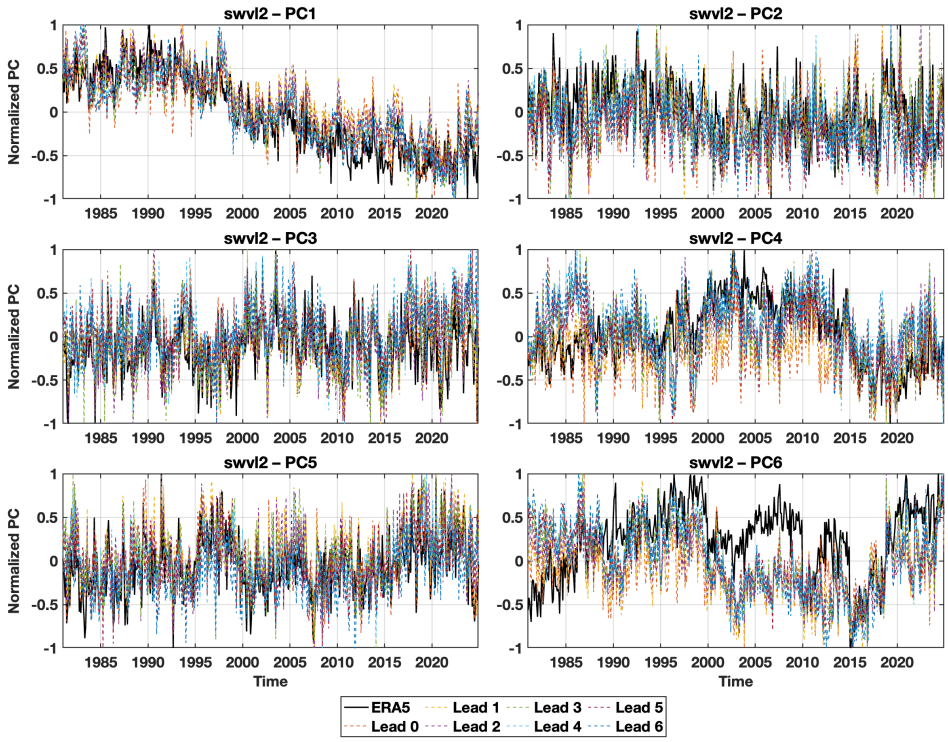

Fig. S2 Normalized PCs of ERA5 and reconstructed SEAS5 soil-moisture anomalies for the second soil layer (swvl2). Figure S2 displays the first six PCs for this layer.

S1 Figure S4

The statistical significance of the ERA5–SEAS5 correlations for both the original and EOF-reconstructed soil moisture datasets was assessed using the Fisher z-test together with an F-test for improvement in correlation strength. Figure S4 illustrates the spatial distribution of significant correlations at the p ¡ 0.05 level (95% confidence). The EOF-reconstructed fields show a substantially larger fraction of significant grid points compared to the original SEAS5 output, indicating that reconstruction effectively enhances the physically meaningful large scale signal while suppressing high-frequency noise. The F-test confirms that the increase in correlation from the original to the reconstructed dataset is statistically significant, demonstrating that the improvement reflects genuine enhancement of model reanalysis agreement rather than sampling variability. This strengthened significance is most pronounced across central and eastern Europe, where soil moisture memory is inherently higher, while regions that remain insignificant such as arid areas of North Africa align with known hydrological and modeling limitations. Overall, Figure S4 shows that EOF reconstruction produces a statistically robust improvement in ERA5–SEAS5 correspondence, validating its use for downstream applications such as soil-moisture predictability assessment and crop-yield forecasting.

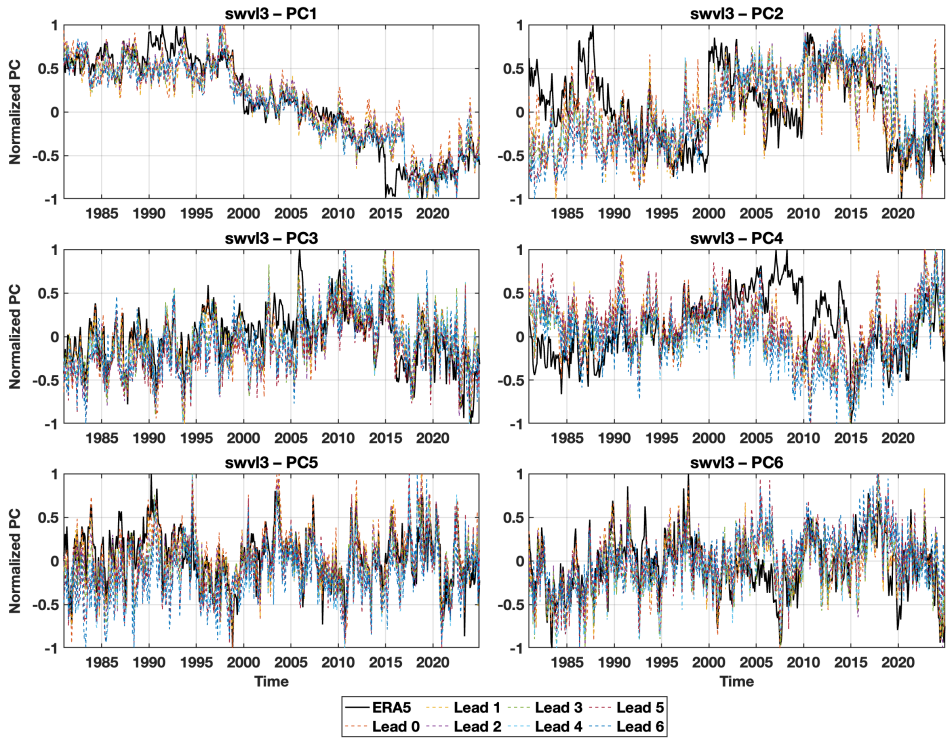

Fig. S3 Normalized PCs of ERA5 and reconstructed SEAS5 soil-moisture anomalies for the third soil layer (swvl3). Figure S3 presents the first six PCs corresponding to this deeper layer.

S2 Figure S5

Figure S5 shows the pixel-wise correlation between reconstructed ERA5 soil-moisture anomalies (depth-weighted average of swvl1–swvl4) and observed winter-wheat and maize yield anomalies across Europe. The reconstruction is based on principal components (PCs) 2–10; PC1 is excluded because it exhibits discontinuous and non-physical behaviour that would otherwise introduce spurious spatial patterns into the reconstructed fields. Using only the physically meaningful modes yields clearer and more spatially coherent soil moisture–yield correlations than those obtained from the raw ERA5 anomalies, with the most evident improvements over southeastern and central Europe where rain-fed crop production is highly sensitive to soil-moisture variability. Overall, the depth-weighted reconstruction using significant PCs provides a modest but consistent enhancement in the soil moisture–yield relationship for both winter wheat and maize across the European domain.

In Fig. S5, panels (a) and (c) present the spatial (grid-cell) correlations between ERA5 soil-moisture anomalies and winter-wheat and maize yield anomalies, respectively. Panels (b) and (d) show area-averaged time series extracted from the boxed region indicated in the corresponding maps. Soil-moisture time series are computed for

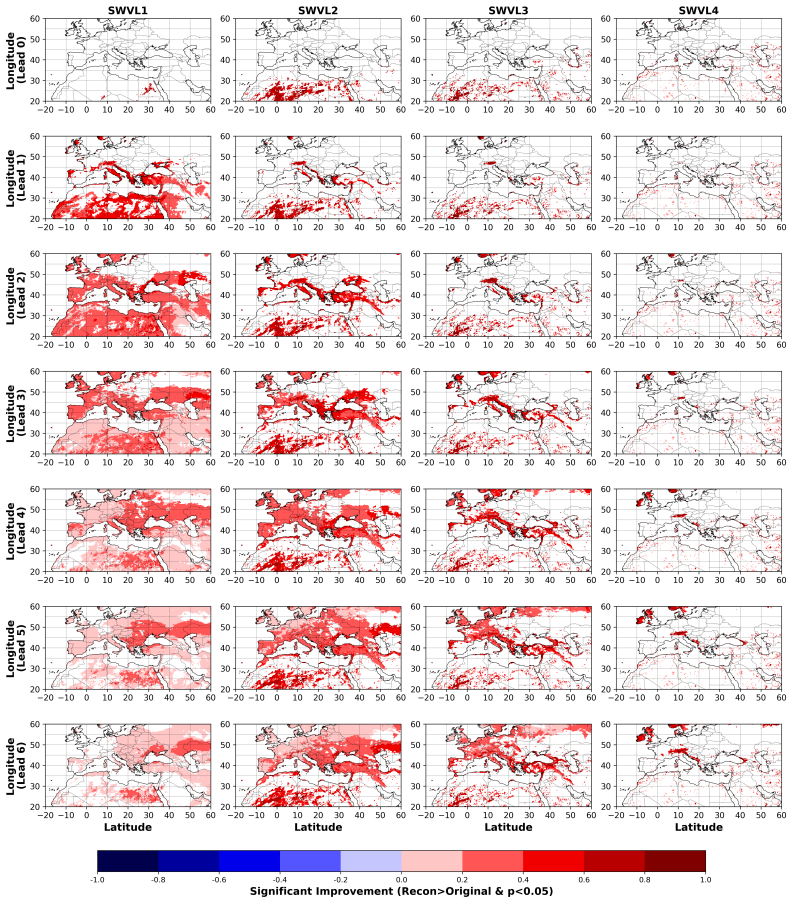

Fig. S4 Significance of correlation between original soilmoisture anomalies of ERA5 and SEA5 data and reconstructed soilmoisture anomaly of ERA5 and EOF projected SEAS5 soilmoisture anomalies with P avlues($p < 0.05$)

the relevant seasonal windows, whereas yield is available at annual resolution; therefore, we compare the seasonal soil-moisture anomalies with annually averaged yield anomalies. All time series shown are detrended and normalized (see Methods) to facilitate comparison. Although strong spatial correlations are obtained in several regions, the agreement between the paired area-mean time series is weaker, likely reflecting differences in temporal aggregation (seasonal windows versus annual yield) and potential lags between soil moisture conditions and harvested yield. This mismatch merits further investigation in future work.

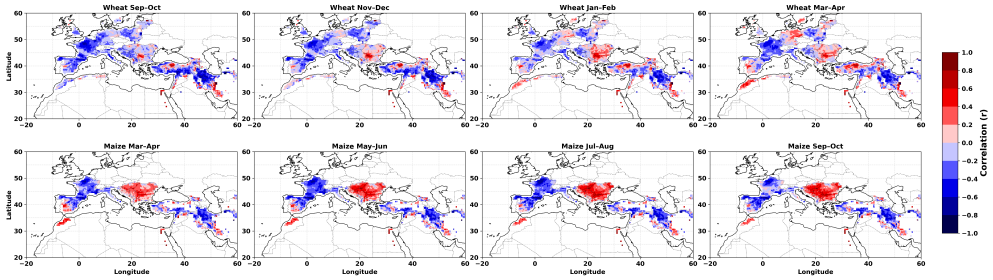

Fig. S5 Pannels 'a' and 'c' represents the pixel-wise correlation between the reconstructed ERA5 soil moisture (depth-weighted average of 4 layers) anomaly derived from the 2–10 principal components and the crop-yield anomalies (Winter wheat and Maize), Pannels 'b' and 'd' represents the normalized timeseries of soil moisture and crop yield (Winter wheat and Maize) for the rectangular region for the period of 1981 to 2016