Manuscript number: HESS-2025-365

Title: Learning Evaporative Fraction with Memory

Zhao et al. developed a machine learning model to predict the daily evaporative fraction (EF) across 67 eddy-covariance flux sites with different plant functional types (PFTs) worldwide. A Long Short-Term Memory (LSTM) unit is used to examine the effect of previous climatic conditions (e.g., precipitation, air temperature) – analogous to a memory effect – on EF predictions. The manuscript highlights that memory effects – ranging from 7 to 365 days – largely contribute to EF dynamic predictions, particularly across forest sites, likely due to deep roots and water use strategies compared to grassland and savanna sites. More specifically, the results show that air temperature and precipitation, rather than net radiation, emerged as the most influential drivers of EF predictions. The study provides new insights into how PFTs, soil properties, and the climate characteristics of the previous year (i) impacted EF dynamics in the following year (i + 1). This has the potential to improve the understanding of water and carbon fluxes under less frequent but more intense precipitation events due to climate change. Below are my comments:

Major comments:

- Your results suggest that air temperature is the most influential driver of EF predictions for forest sites (Figures 4f, 4h, and 4i), while precipitation is the dominant factor in savannas, grasslands, and shrubland sites. While the role of precipitation is clear (lines 256-275), the role of air temperature requires a more detailed explanation in the context of the memory effect on EF. For example, while Chen et al. (2018) highlighted that air temperature dominates the half-hourly ET in evergreen needleleaf forests, their analysis did not consider the memory effect.
- Lines 203-205: You state that "*we observed a consistent bias in the model's predictions at higher EF values for some sites, compared to lower ones. A closer examination reveals that higher EF values typically occur at forested sites, while lower EF values are found at non-forested sites*". However, EF also varies depending on precipitation seasonality (dry or wet season) or growing season in each site. Additionally, Miguez-Macho and Fan (2021) show that transpiration during the wet

season mainly depends on recent infiltration (topsoil groundwater). Conversely, transpiration is mainly supplied by water from past infiltrations stored in deep soil during the dry season, suggesting that the effect of deep root on evaporative fraction (EF) changes between seasons. In this sense, the memory effect on EF predictions could be higher during the dry season – EF would exhibit a higher dependence on previous precipitation events – than during the wet season. Thus, how does memory effect on EF prediction changes during the dry vs wet seasons or growing/non-growing season? Similarly, how does model accuracy vary as a function of dry down length?

- Line 36. Authors state that "*a slowly declining EF during periods of droughts is indicative of a deep root system*". Is this the only factor modulating EF dynamics? What about the role of shallow water table depth (WTD)? For example, Chen et al. (2024) found that enhanced vegetation index (EVI) anomalies during drought years in the Amazon rainforest depend on both root deep and WTD. Thus, how could EF predictions improve by including WTD data (e.g., from Fan et al., (2013))?
- In Lines 235-245, you state that EF depends on precipitation distribution throughout the year. It would be interesting to quantify how the memory effect on EF varies across regions with contrasting precipitation seasonality, characterized by metrics such as the Seasonality Index (SI; Teegavarapu, 2019) and perform a similar analysis as in Figure 6. This would allow you to assess the contribution of root depth/plant functional types in comparison to precipitation distribution.
- Lines 301-305. You state that "*the maximum duration of the precipitation reveals that the legacy effect of winter precipitation, occurring 365 days prior, continues to impact water use strategies, as shown by our analysis. This observation aligns with the previous studies and experimental findings. For instance, antecedent winter precipitation from one or two years prior has been shown to influence growth in the current period*". However, this may not be clear for tropical regions and regions where most of the precipitation occurs in the summer (e.g., southwestern US).

Minor comments:

- Add a space between text and all citations (e.g., Line 31).

- How do you define drydown? For example, Fu et al., (2022) define dry-downs like *"episodes with no rain for several consecutive days during which soil moisture (SM) shows a short-term "pulse" rise after rain and then decays until the next rain event".*
- In the manuscript you used "dry-down" and drydown expressions. Please standardize the term.
- Line 95. What is the source of soil moisture data?
- Line 94. Describe the root depth and evapotranspiration data used in the Figure 6.
- Line 115. Does MODIS extract over the full available period? Did you apply any quality control?
- Line 120. From which soil depth are soil properties retrieved?
- Line 131. A descriptor of precipitation seasonality such as the Seasonality Index (SI; Teegavarapu, 2019), can provide additional information about its distribution throughout the year.
- Lines 166-167. You state that "*whole datasets were separated by site, with one site from each PFT randomly selected for testing and validation, and the remaining sites used for training*". How many EBF sites are included in your analysis? Based on Figure S1, I assume that only one is included.
- Figure 4j. Why is the evergreen broadleaf forest (EBF) not included in this panel, but it shows in Figure 4g? The same question applies to figure 5.
- Figure 6: How do you rank rooting depth? Are the RD values in panel B the mean values for all sites within each soil sand fraction level? Why did you define the aridity index as ET/PET instead of the more traditional definition of P/PET?.
- Figure 3. I suggest replacing at least one grassland or savanna site with an EBF site.
- Figure 4. Please indicate the number of sites included for each PFT.
- Table S1. Please indicate the sites selected for your analysis.
- Figures 4-6. In line 105 you mention that only sites with > 10 years of records were included in your analysis. So, the note "*sites with more than 10 years of observations are summarized here*" is not needed in figures descriptions may be redundant.
- Please check some minor grammar issues, for example:
    - Line 9. plant "water stress"?
    - Line 17 and 275. "precipitation" and "VPD"

- Line 18. Add comma after “(OSH”)

References

Chen, S., Stark, S.C., Nobre, A.D., Cuartas, L.A., De Jesus Amore, D., Restrepo-Coupe, N., Smith, M.N., Chitra-Tarak, R., Ko, H., Nelson, B.W., Saleska, S.R., 2024. Amazon forest biogeography predicts resilience and vulnerability to drought. Nature 631, 111–117. https://doi.org/10.1038/s41586-024-07568-w

Chen, Y., Xue, Y., Hu, Y., 2018. How multiple factors control evapotranspiration in North America evergreen needleleaf forests. Science of The Total Environment 622–623, 1217–1224. https://doi.org/10.1016/j.scitotenv.2017.12.038

Fan, Y., Li, H., Miguez-Macho, G., 2013. Global Patterns of Groundwater Table Depth. Science 339, 940–943. https://doi.org/10.1126/science.1229881

Fu, Z., Ciais, P., Feldman, A.F., Gentine, P., Makowski, D., Prentice, I.C., Stoy, P.C., Bastos, A., Wigneron, J.-P., 2022. Critical soil moisture thresholds of plant water stress in terrestrial ecosystems. Sci. Adv. 8, eabq7827. https://doi.org/10.1126/sciadv.abq7827

Miguez-Macho, G., Fan, Y., 2021. Spatiotemporal origin of soil water taken up by vegetation. Nature 598, 624–628. https://doi.org/10.1038/s41586-021-03958-6

Teegavarapu, R.S.V., 2019. Changes and Trends in Precipitation Extremes and Characteristics, in: Trends and Changes in Hydroclimatic Variables. Elsevier, pp. 91–148. https://doi.org/10.1016/B978-0-12-810985-4.00002-5