



*Supplement of*

## **FluxHourly: global long-term hourly 9 km terrestrial water-energy-carbon fluxes**

**Qianqian Han et al.**

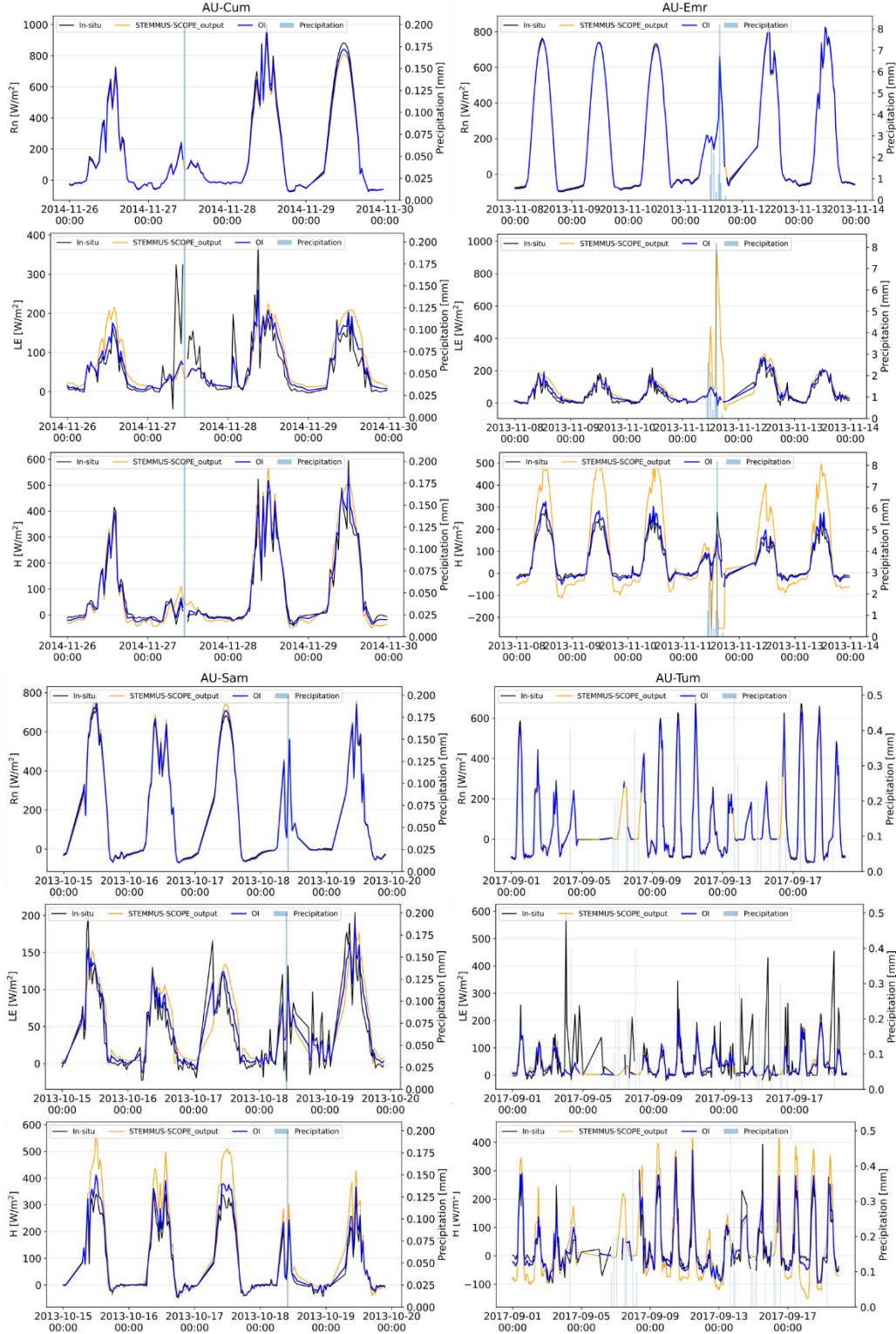
*Correspondence to:* Bob Su (z.su@utwente.nl)

The copyright of individual parts of the supplement might differ from the article licence.



- b) Remove abnormal H values: If  $H < -250$ , then  $H = -250$ ; if  $R_{in} > 10$  &  $H > R_n$ , then  $H = R_n - G$ ,  $LE = 0$ ; if  $R_{in} \leq 10$  &  $H > 100$ , then  $H = R_n - G$ ,  $LE = 0$ .
- c) Remove abnormal LE values: If  $R_{in} > 10$  &  $LE > 1.26 R_n$ , then  $LE = R_n - G - H$ ; if  $R_{in} \leq 10$  &  $LE > 100$ , then  $LE = R_n$ ,  $H = -G$ .

(6) Perform optimal interpolation for  $R_n$ ,  $LE$ ,  $H$ , and keep STEMMUS-SCOPE output for  $G$ , SIF685, SIF740 and GPP.



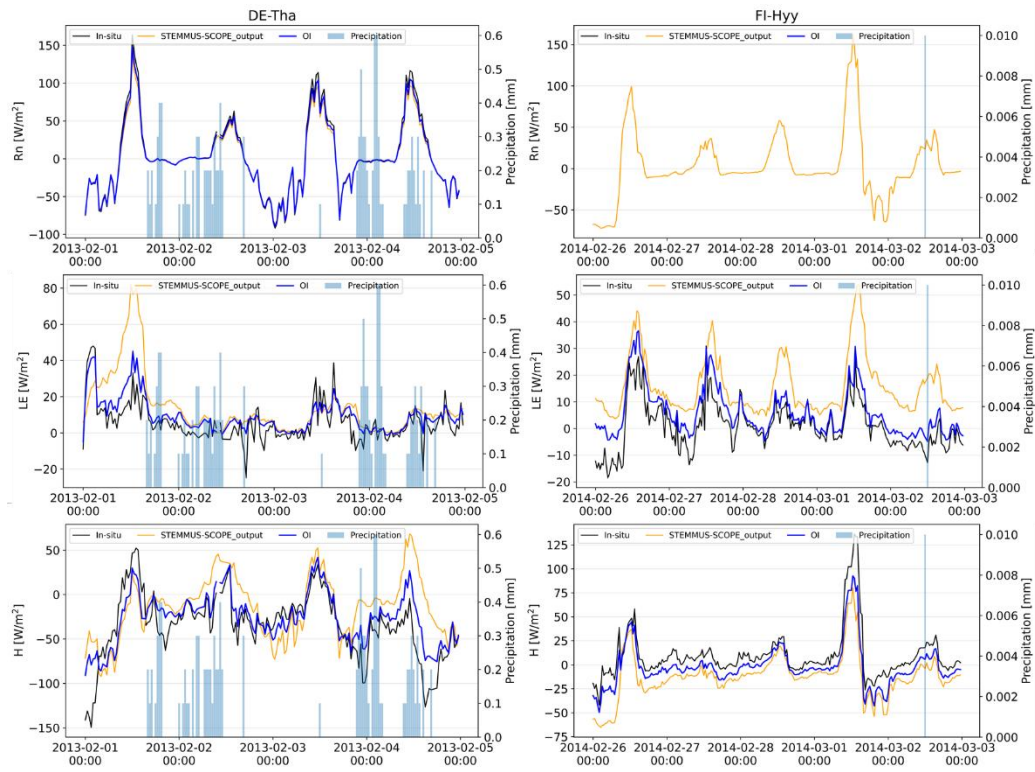


Figure S2. Comparison of in-situ measurements, STEMMUS-SCOPE model outputs, and OI results for net radiation (Rn), latent heat flux (LE), and sensible heat flux (H) at the AU-Cum, AU-Emr, AU-Sam, AU-Tum, DE-Tha and FI-Hyy sites. The blue bars indicate precipitation.

- (7) Apply the Savitzky–Golay (S-G) filter on the seven (optimally interpolated) variables, then use the optimal interpolated values to subtract the S-G filtered result to get residuals.
- (8) Calculate the standard deviation (std) based on the residuals data in two ways: a) daily std; b) use residuals in all stations in the same IGBP type and in the same hemisphere to calculate std for the same timestep in all years (Fig S1).
- (9) Split data into 4 parts: training, testing\_random, testing\_time, testing\_space. Station JP-SMF in training, testing\_random and testing\_time does not have in-situ Rn, station US-PFa in testing\_space does not have in-situ Rn. These two stations were filtered out.
- (10) Convert local time to UTC time.
- (11) Resample from half-hourly to hourly.

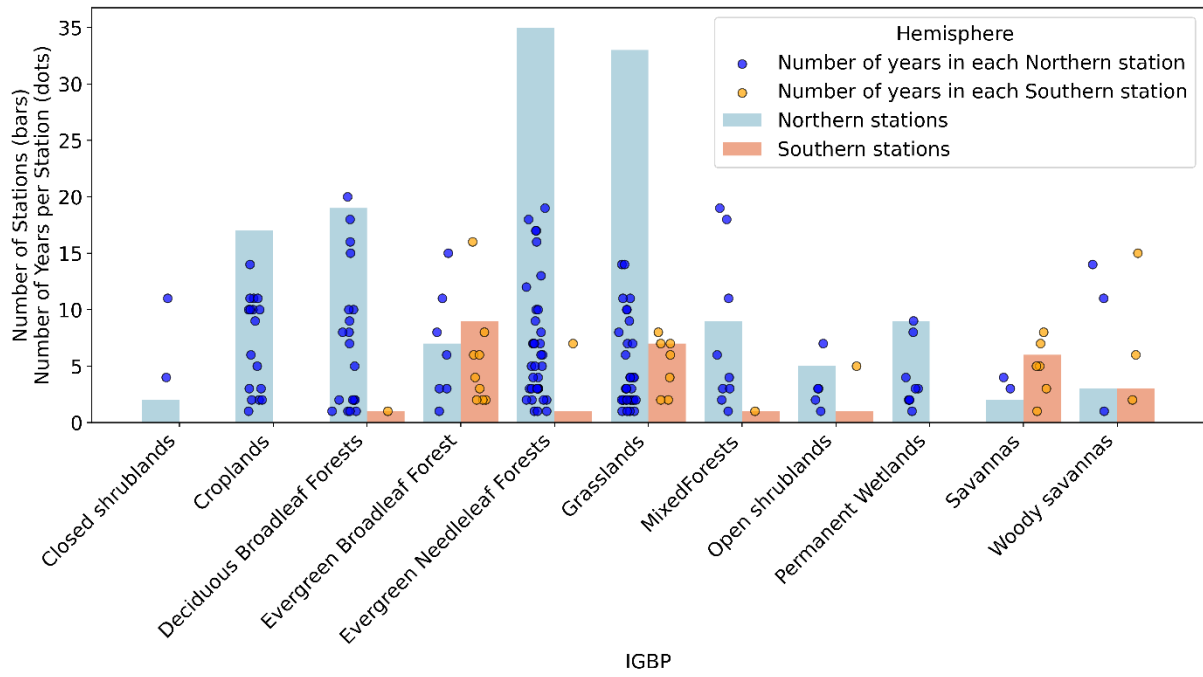
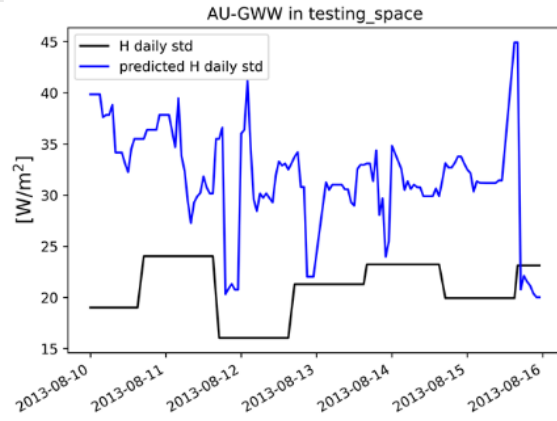
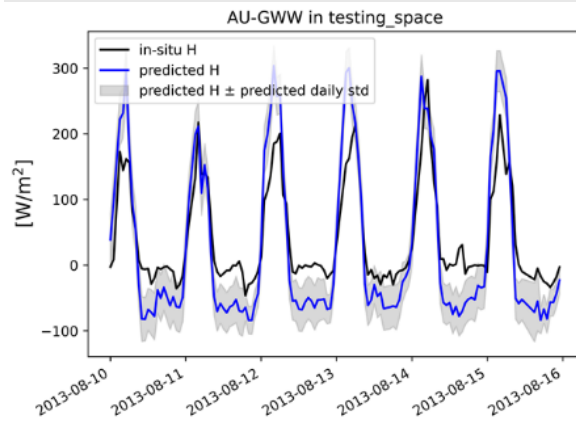
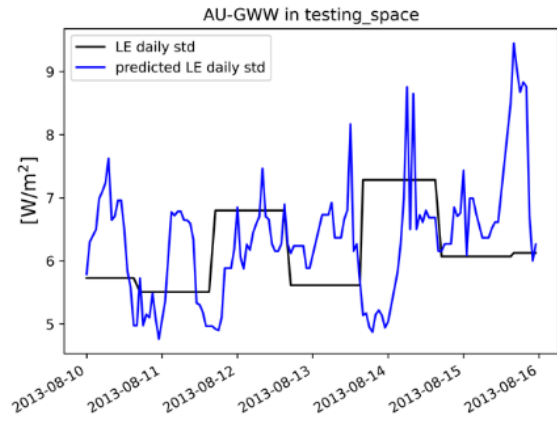
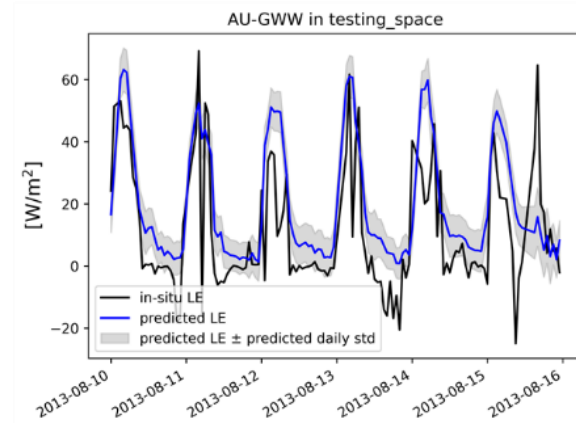
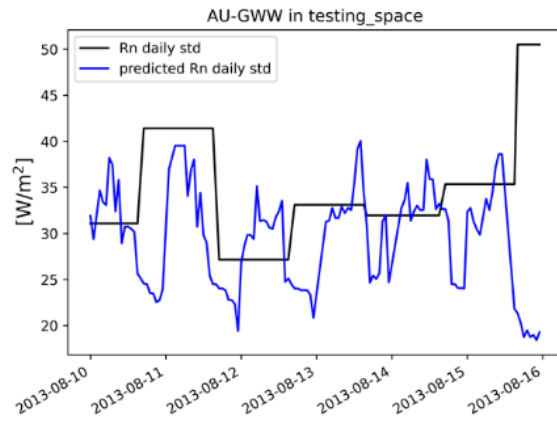
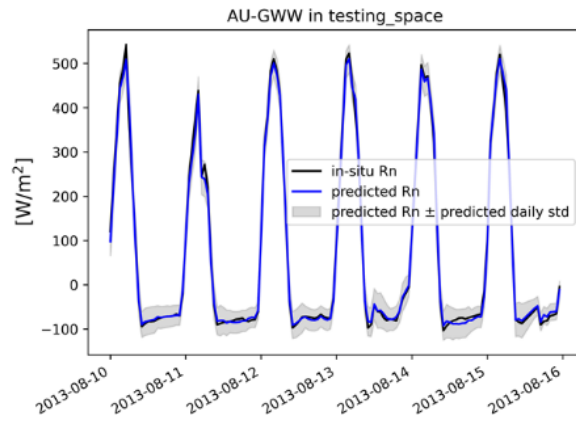


Figure S3. Stations statistics over each IGBP land cover class and each hemisphere

Because there is both diurnal and seasonal changes in fluxes, we first applied the Savitzky–Golay (S-G) filter on the optimal interpolation result to remove such patterns. Then we used the difference between the optimal interpolation result and the S-G filtered result to get residuals, from which we calculated std in two ways. This std can be understood as data fluctuation. We first used this std to train a Random Forest algorithm, then we used this trained Random Forest to upscale std to gridded scale for each grid and timestep. We call this std uncertainty and calculated it in two ways as follows.

- Daily std from the residuals (using 48 values with half-hourly data).
- Std from the same time step of residuals from different stations within one IGBP vegetation type.



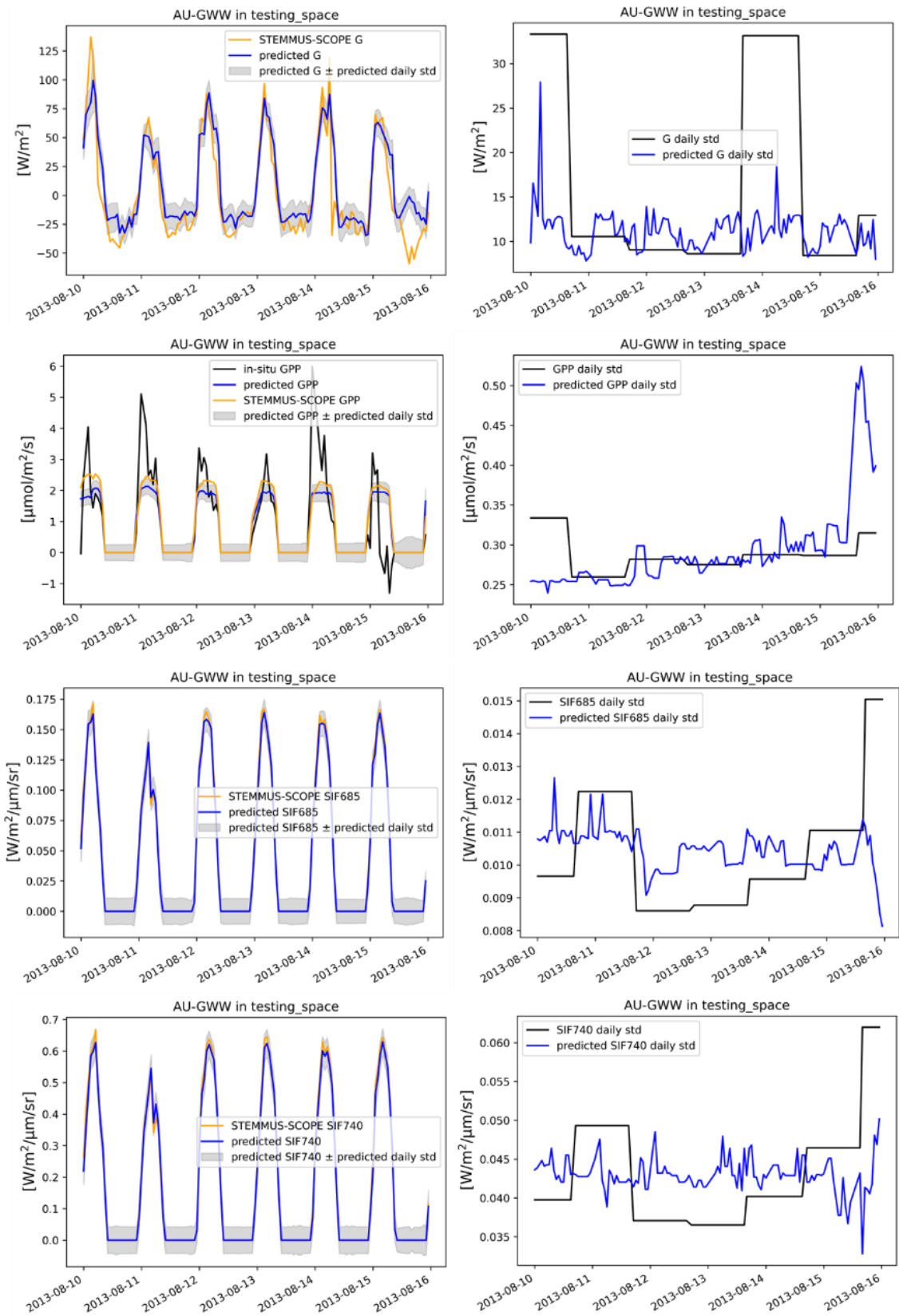
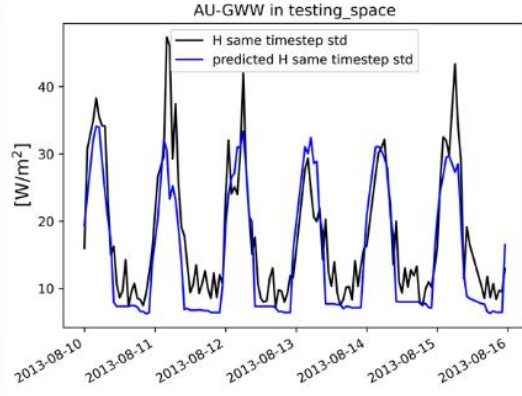
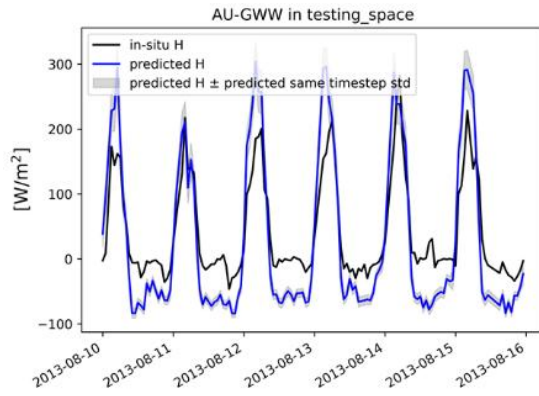
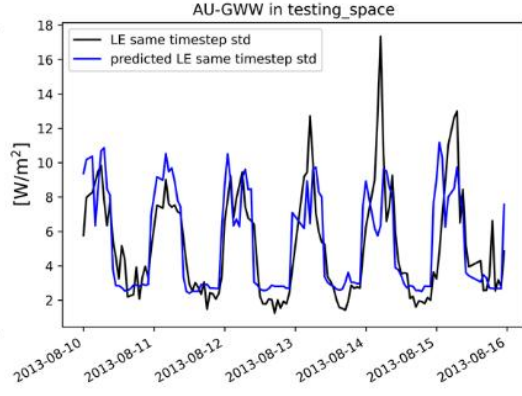
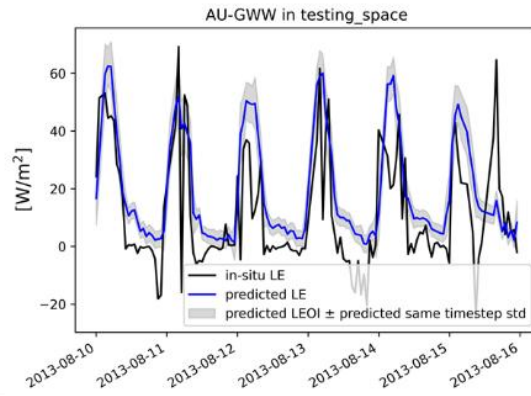
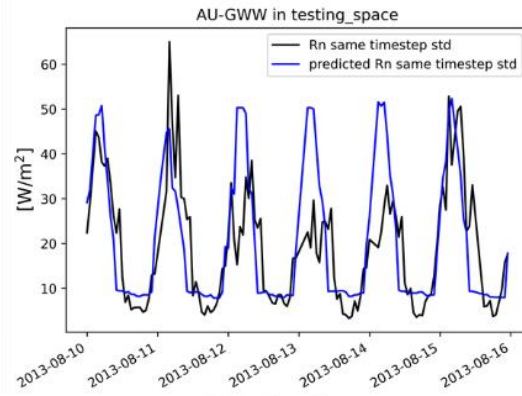
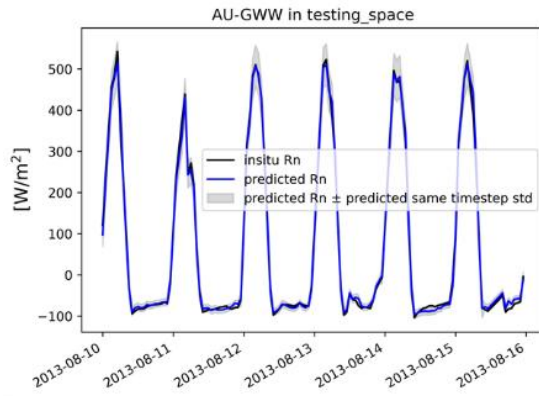


Figure S4. In-situ fluxes, predicted fluxes, and predicted daily std on 7 variables in AU-GWW station (right figures are calculated std and predicted std)



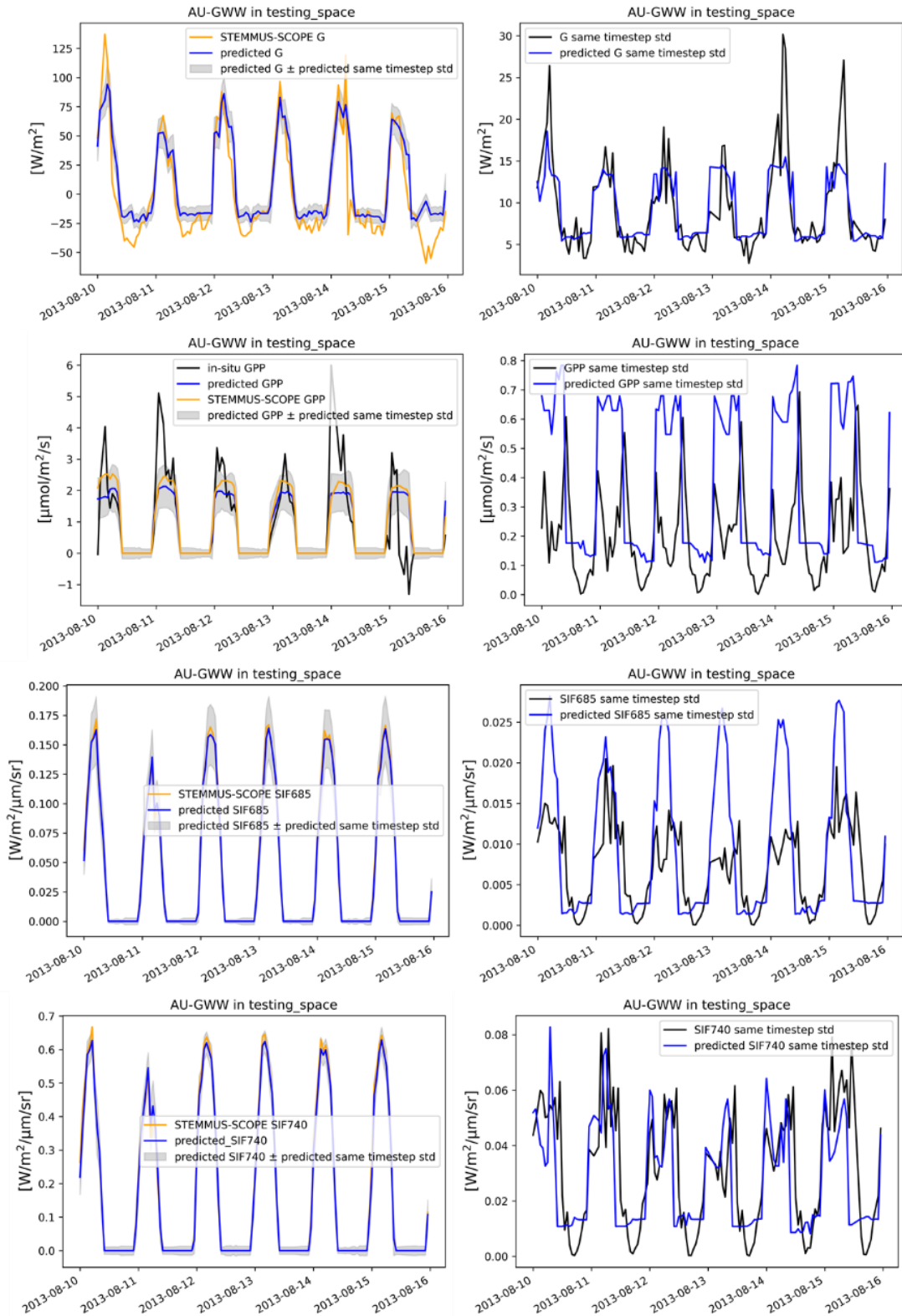


Figure S5. In-situ fluxes, predicted fluxes, and predicted same timestep std on 7 variables in AU-GWW station (right figures are calculated std and predicted std)

## S2 Comparison on testing\_space set on monthly & hourly scale

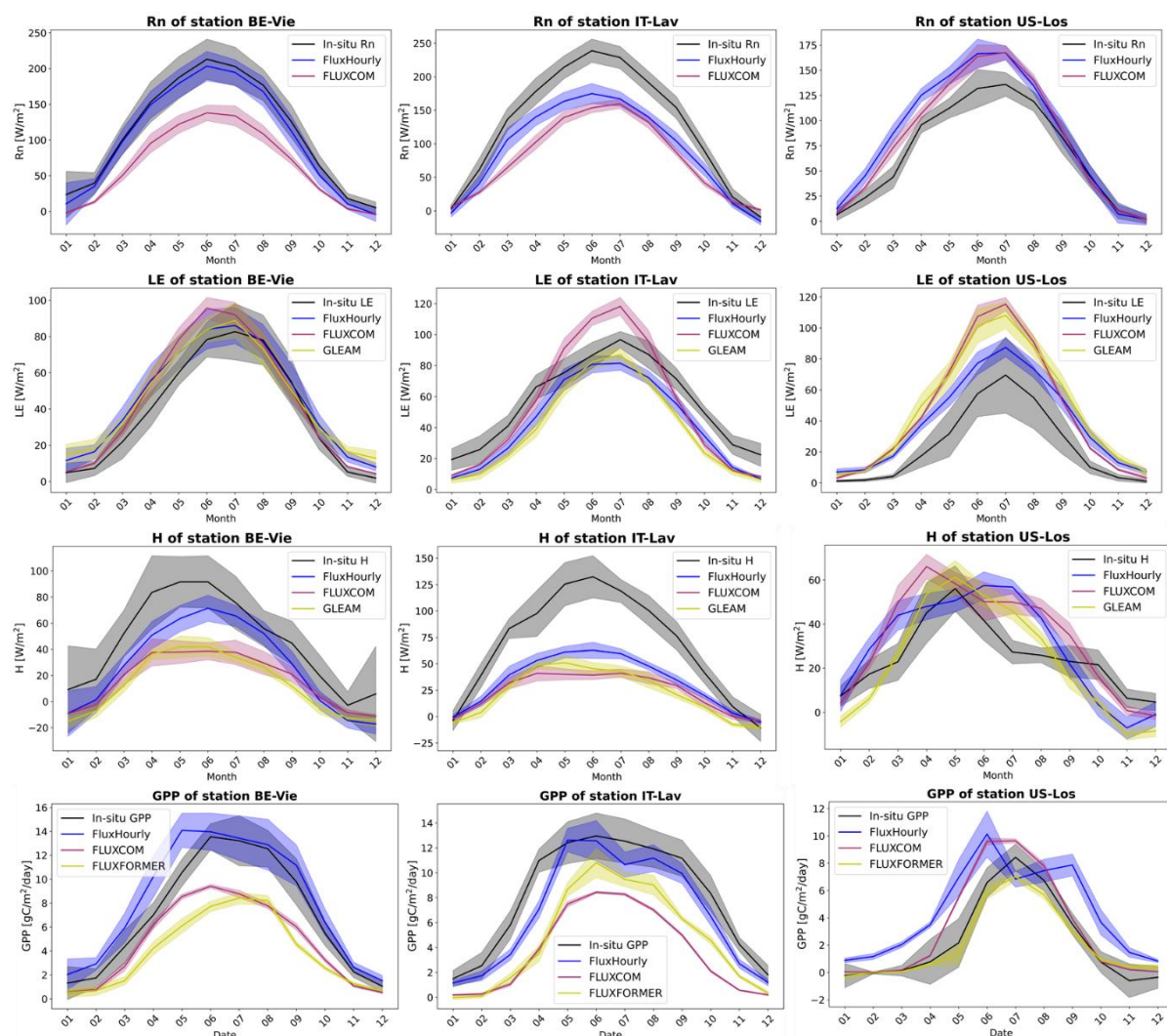
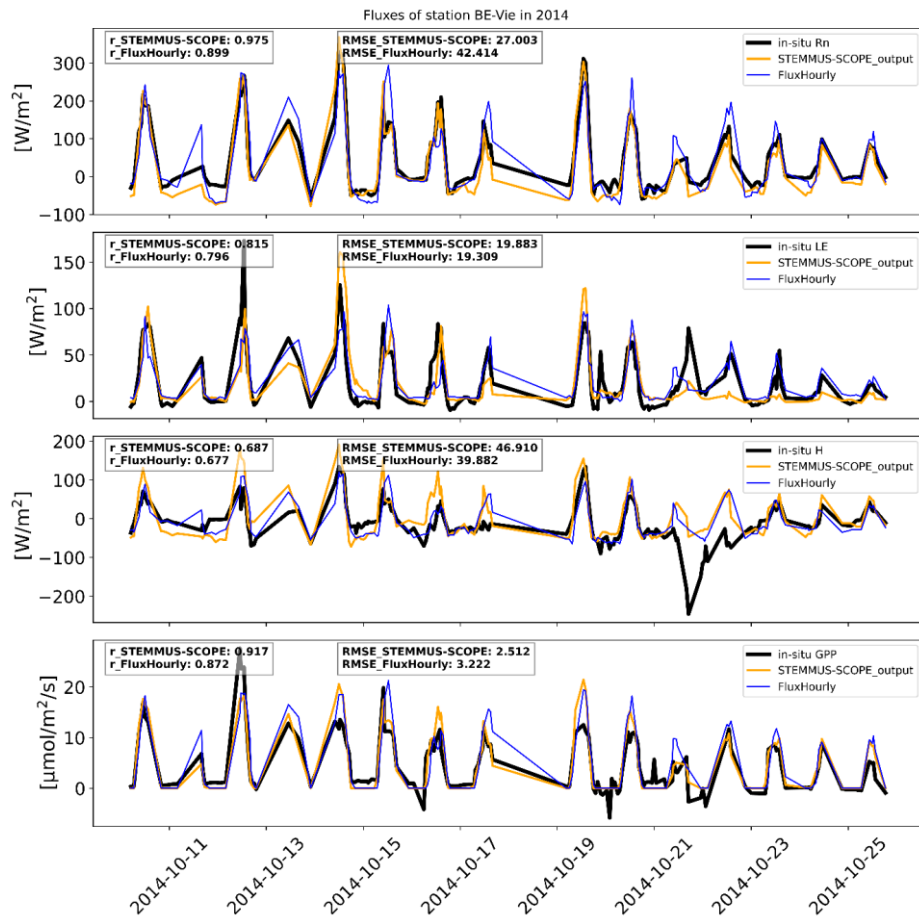
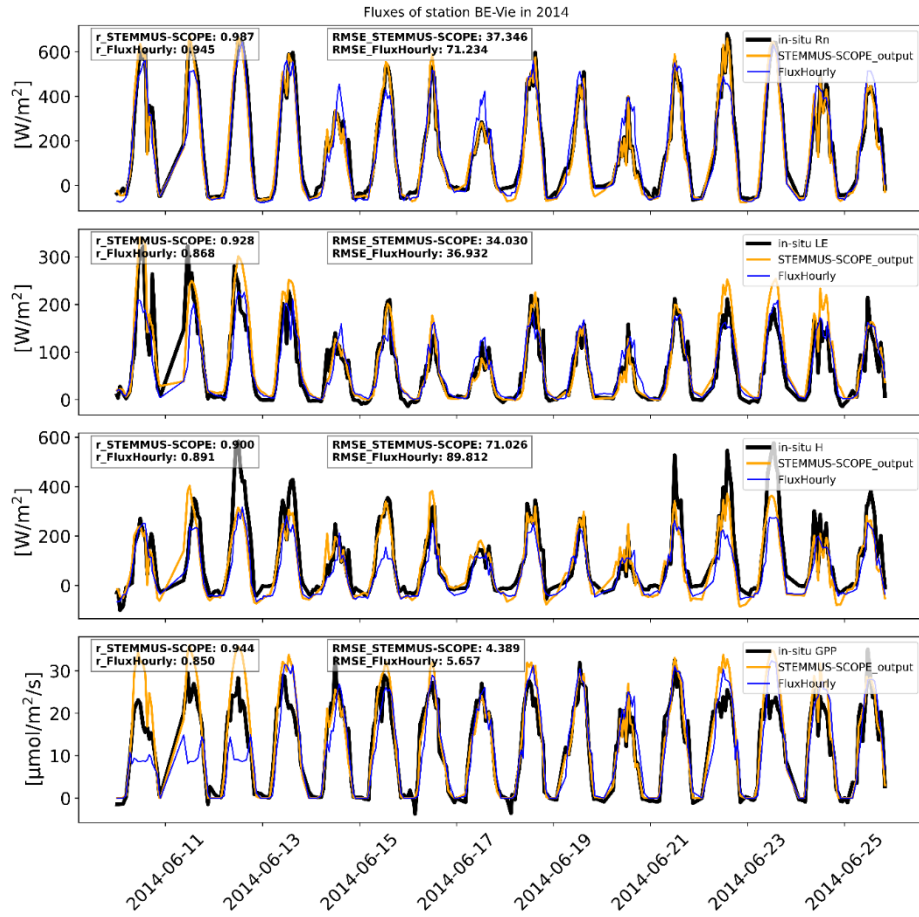


Figure S6. Monthly Rn, LE, H, GPP from in-situ, FluxHourly, FLUXCOM, GLEAM, FLUXFORMER at station BE-Vie, IT-Lav, US-Los



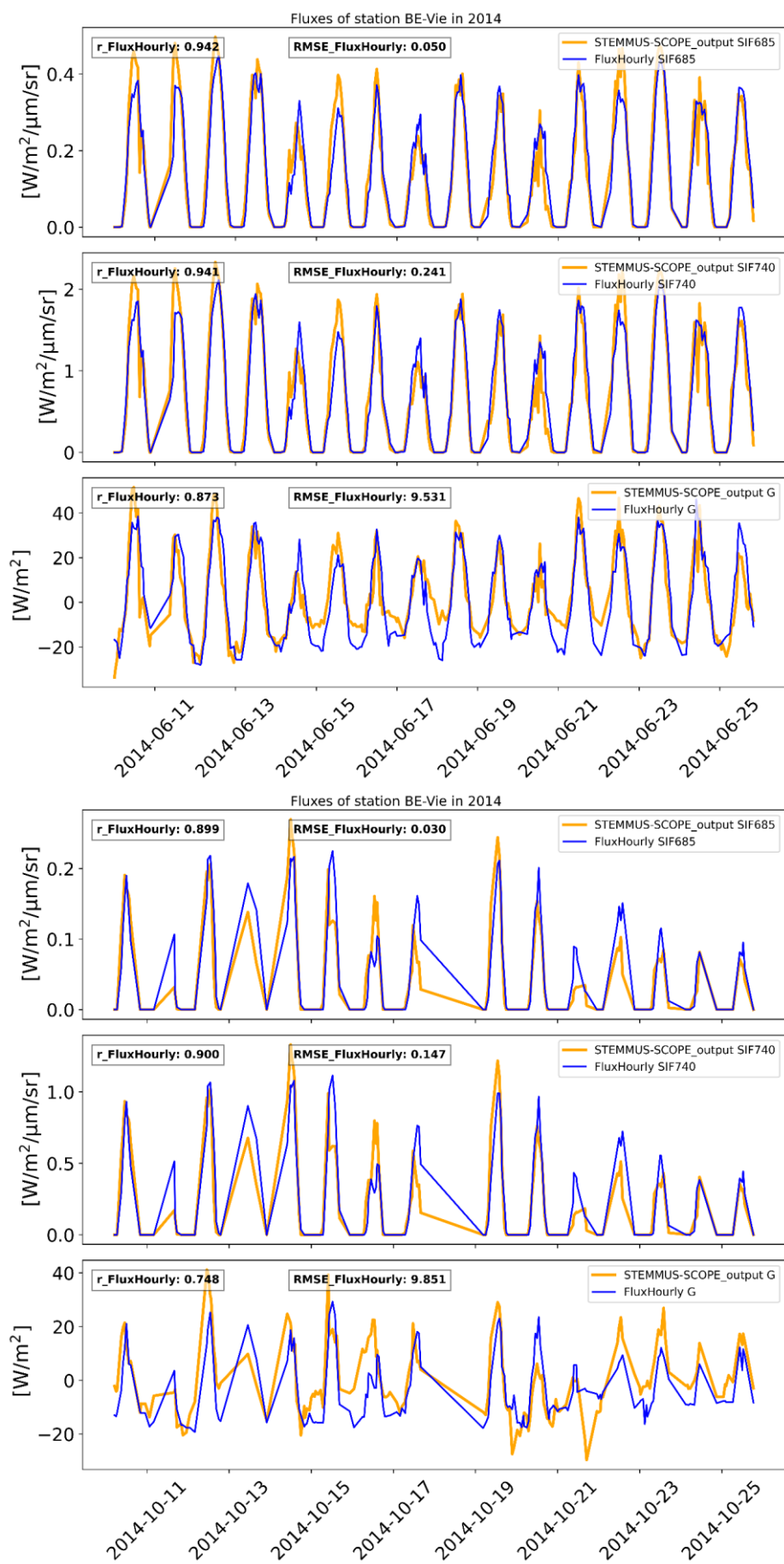
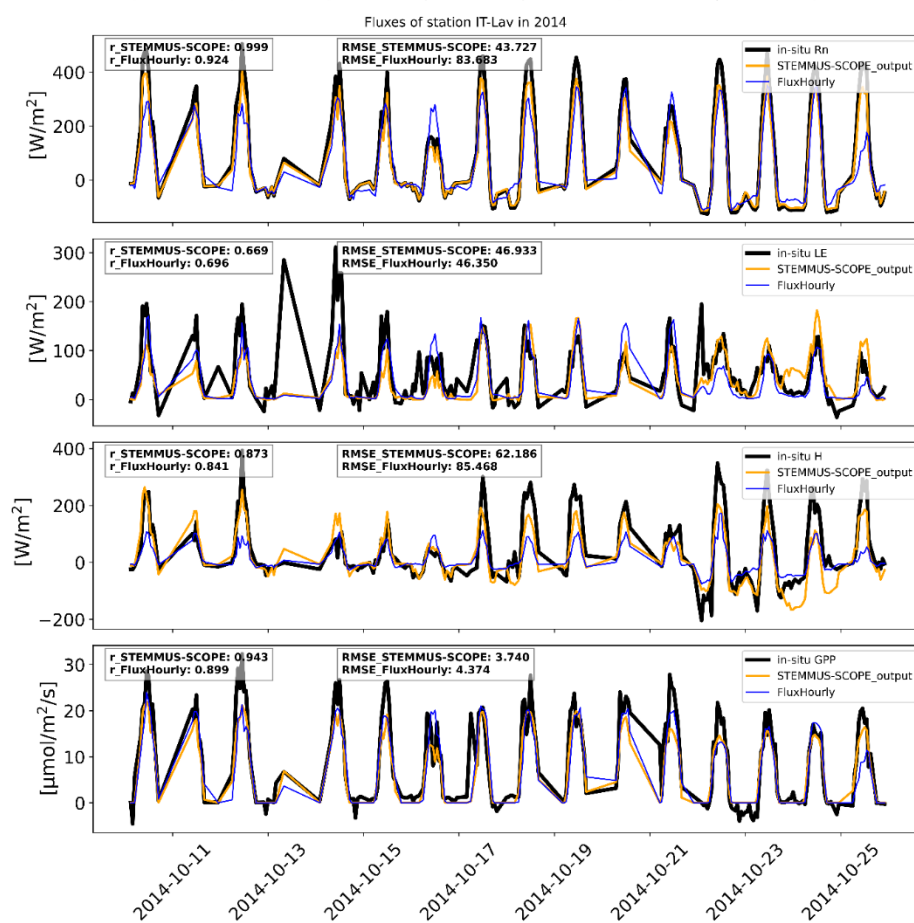
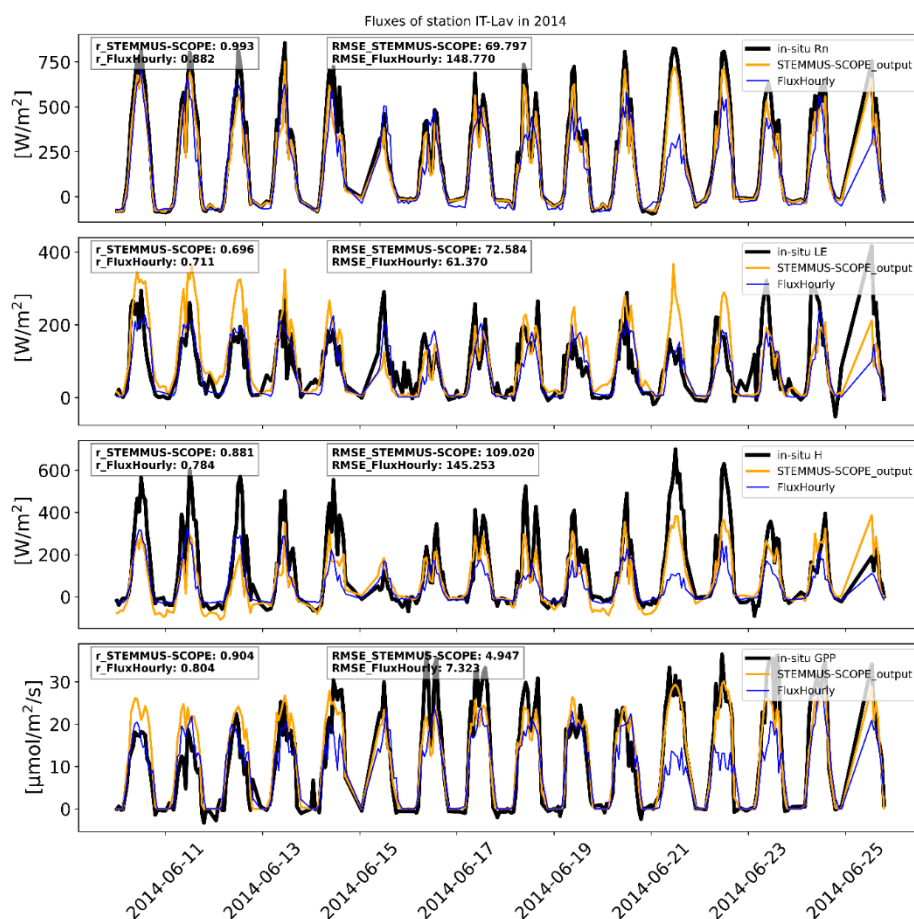


Figure S7. Rn, LE, H, G, GPP, SIF685, SIF740 at BE-Vie, June and Oct (time is in local time)



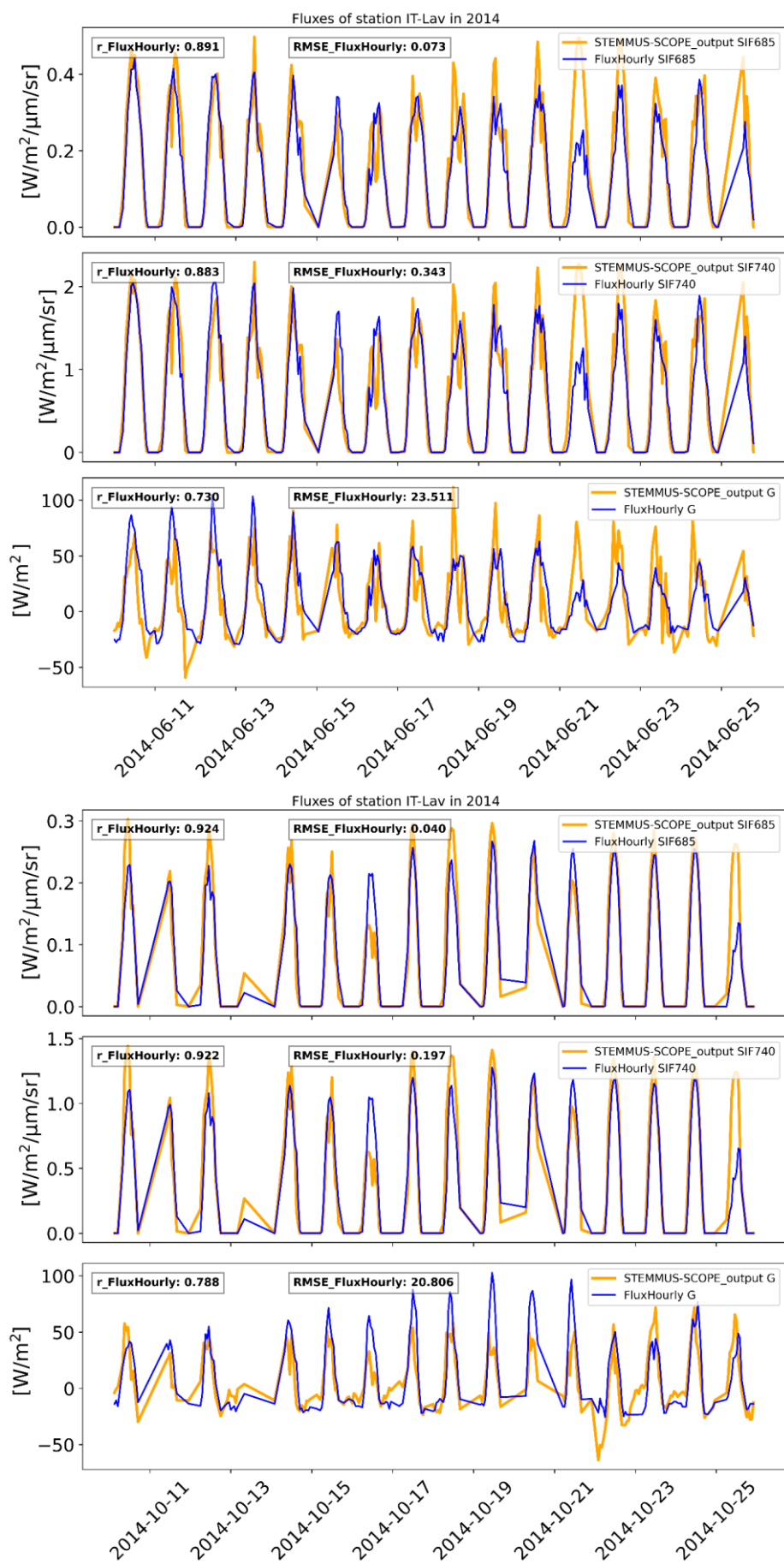
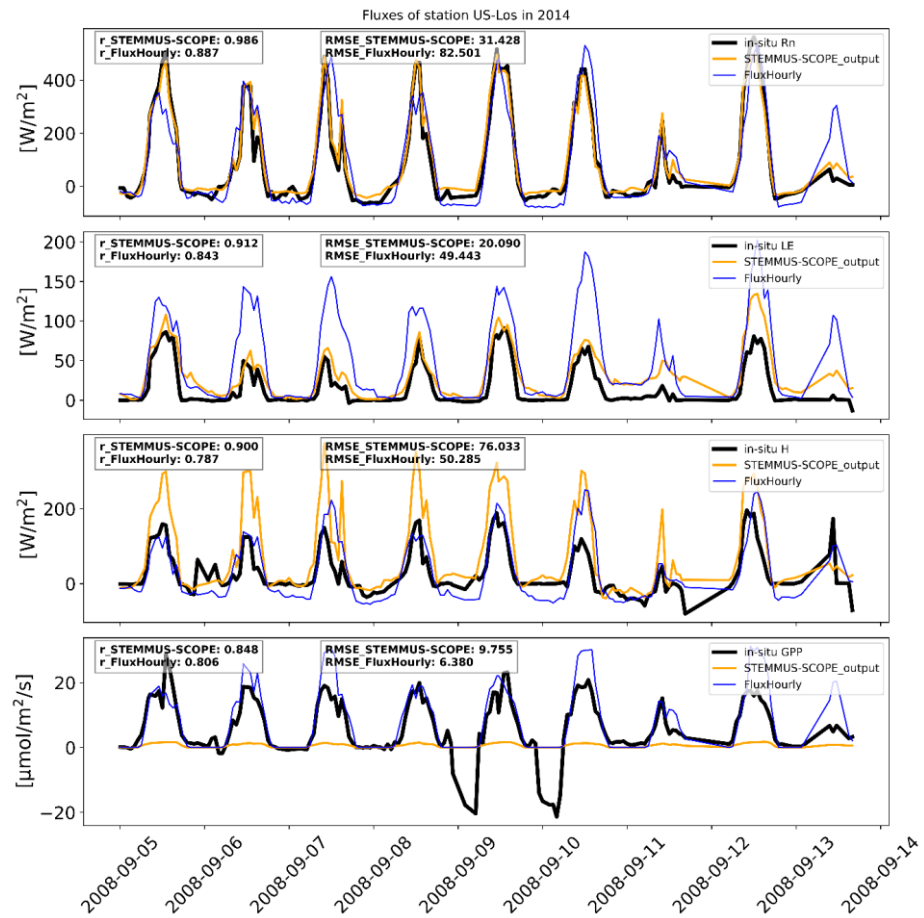
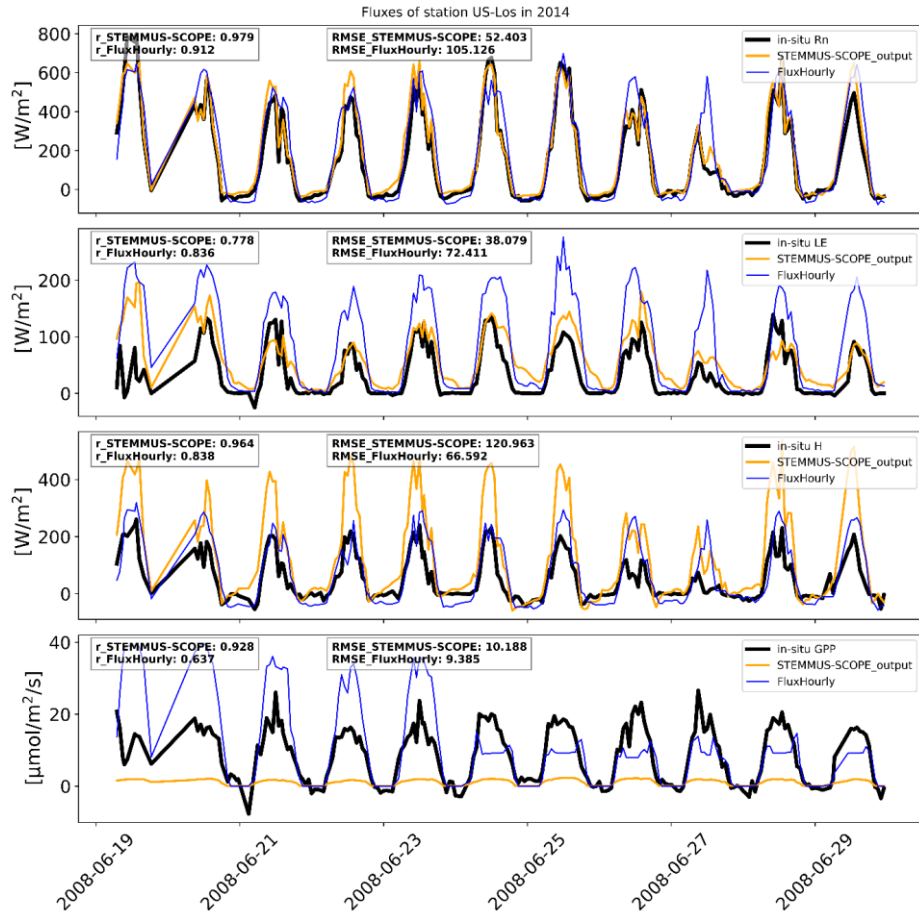


Figure S8.  $R_n$ , LE, H, G, GPP, SIF685, SIF740 at IT-Lav, June and Oct (time is in local time).



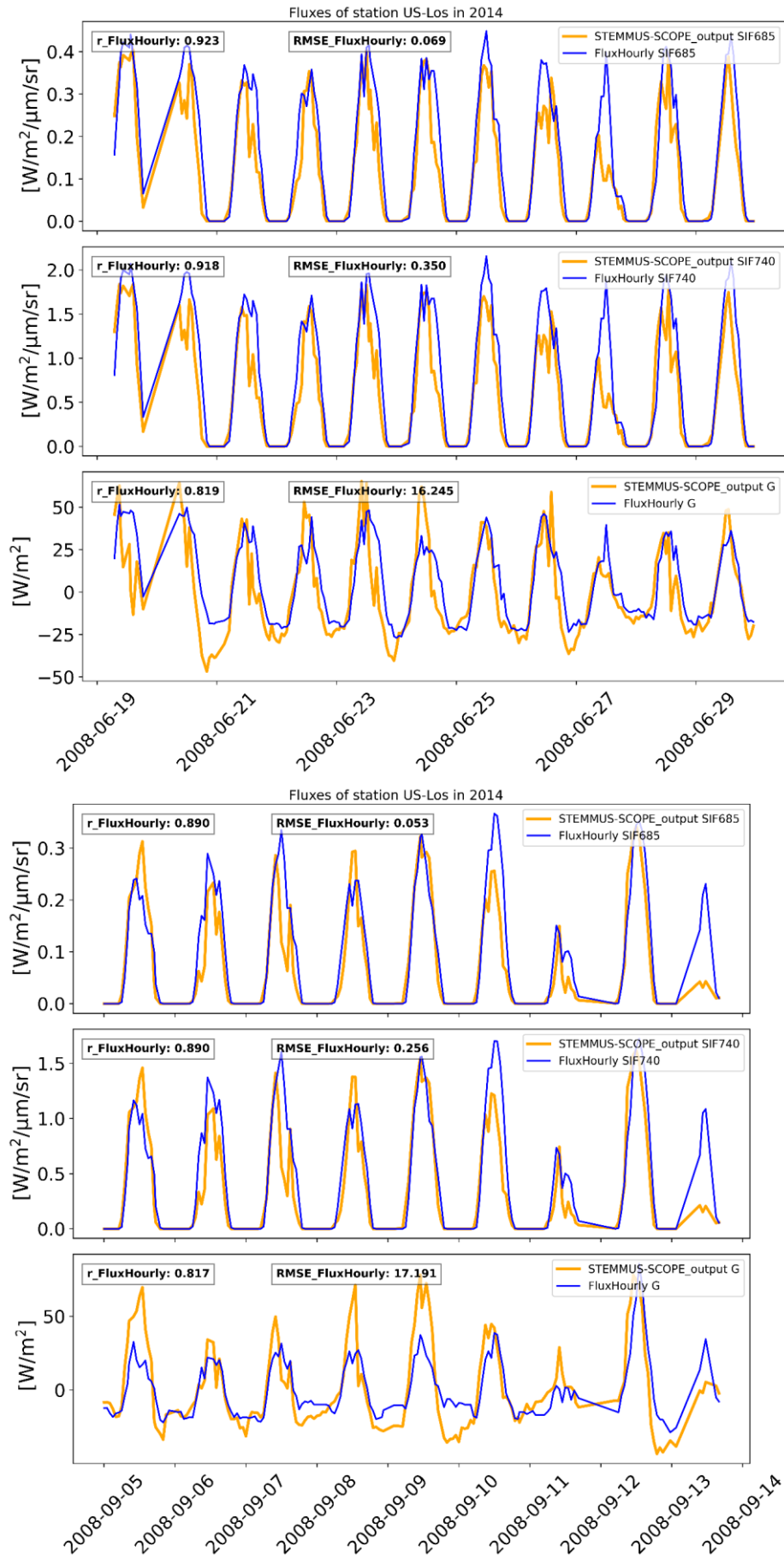
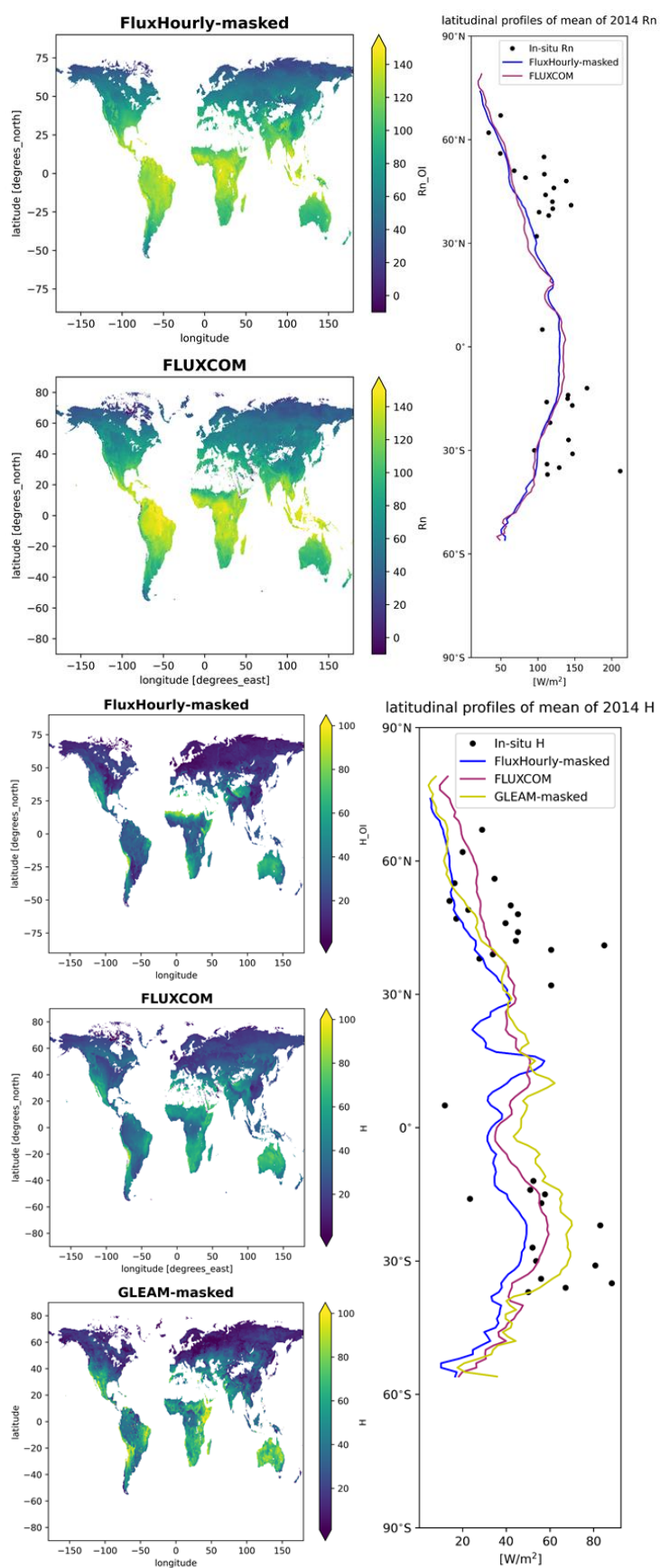


Figure S9.  $R_n$ , LE, H, G, GPP, SIF685, SIF740 at US-Los, June and Sep (time is in local time).

### S3 Inter comparison with existing flux datasets on spatial pattern



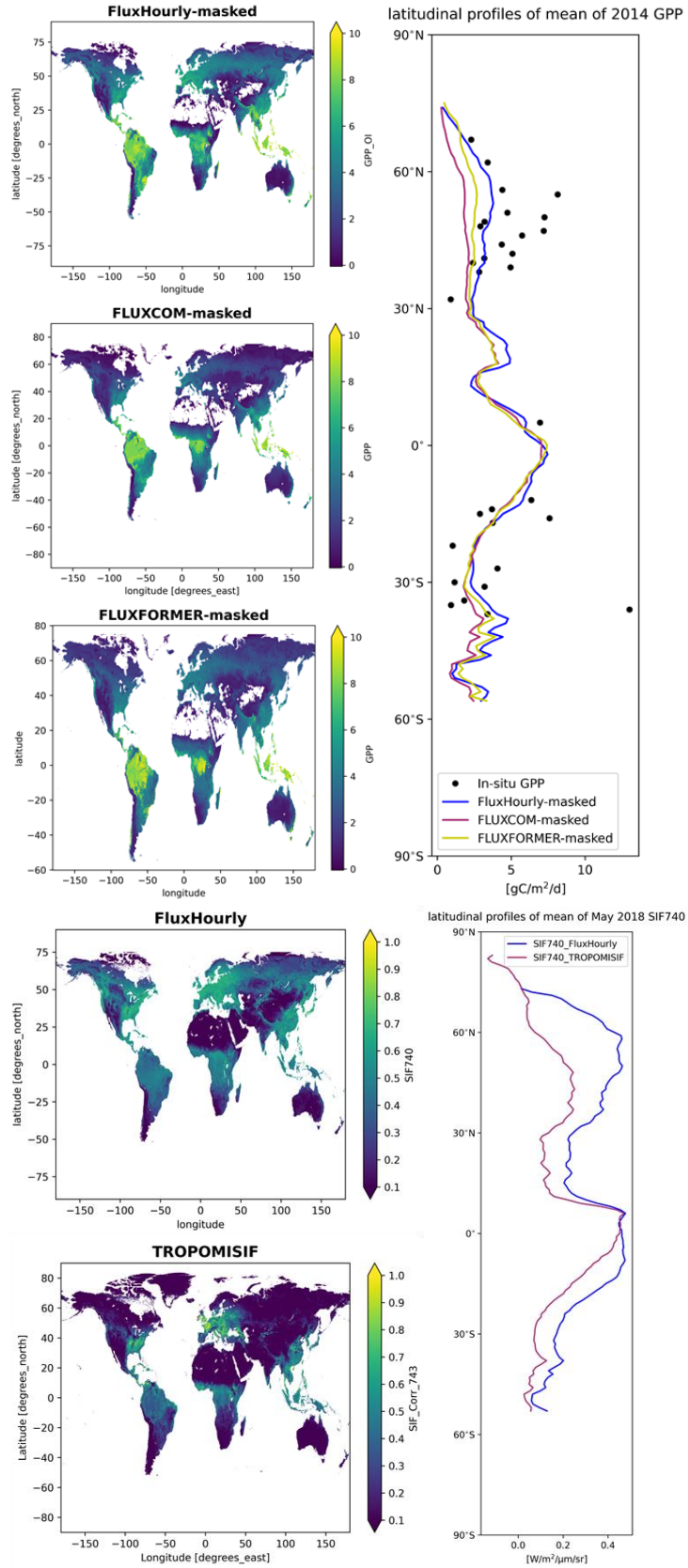


Figure S10. Annual mean of predicted hourly 9 km Rn, H, GPP, SIF740 by RF\_OI, and FLUXCOM, GLEAM, FLUXFORMER (only for GPP) in 2014 and TROPOMISIF (only for SIF740 in May 2018). Note: GPP is masked by both FLUXCOM and FLUXFORMER as both datasets has missing values. Rn and H is masked by FLUXCOM.

## S4 Diurnal patterns and seasonal changes: Rn, H, G

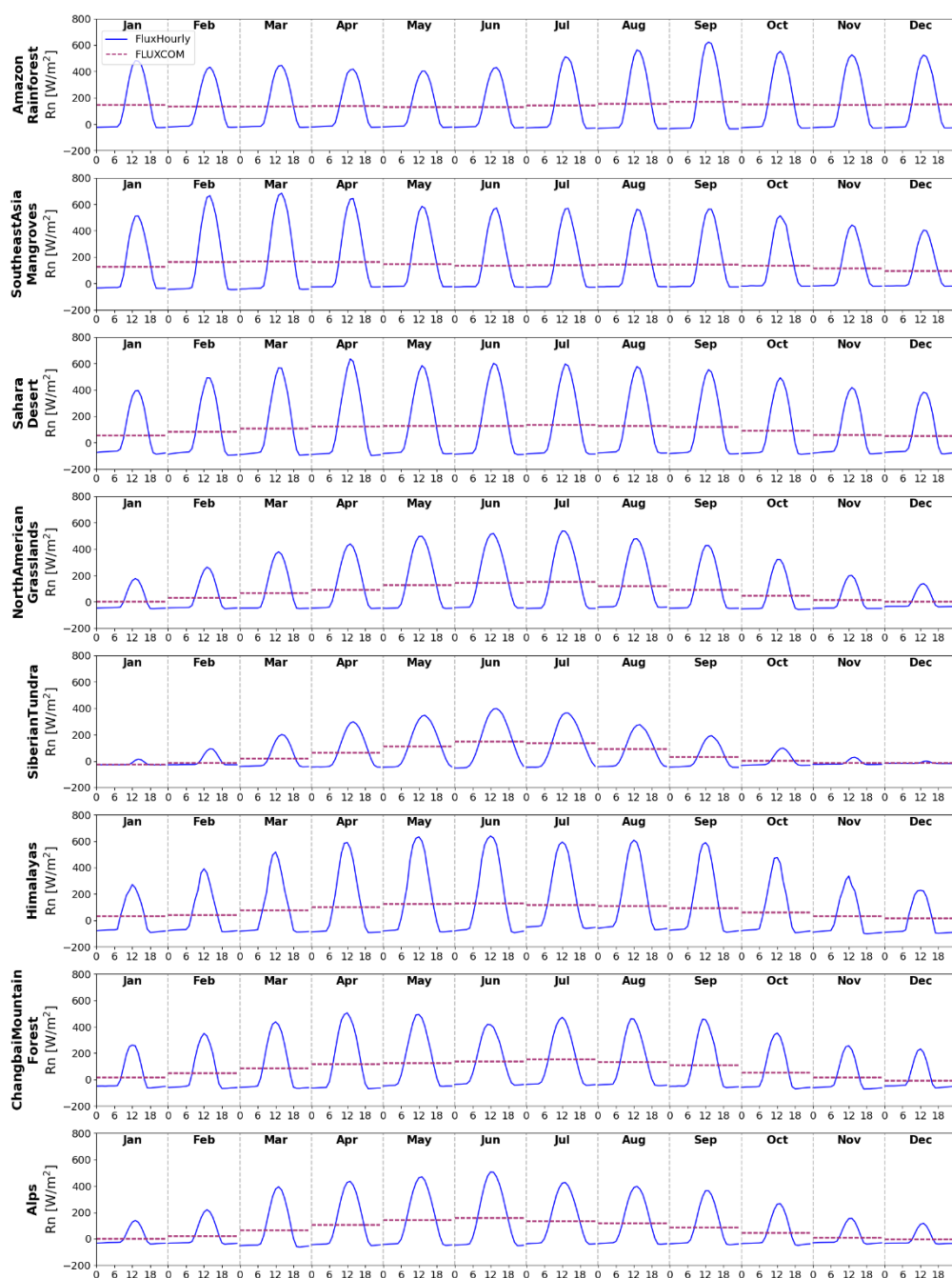


Figure S11. Diurnal cycles of  $R_n$  for 8 regions for each month of the year, where each panel refers to a region. The data is converted to local time zone.

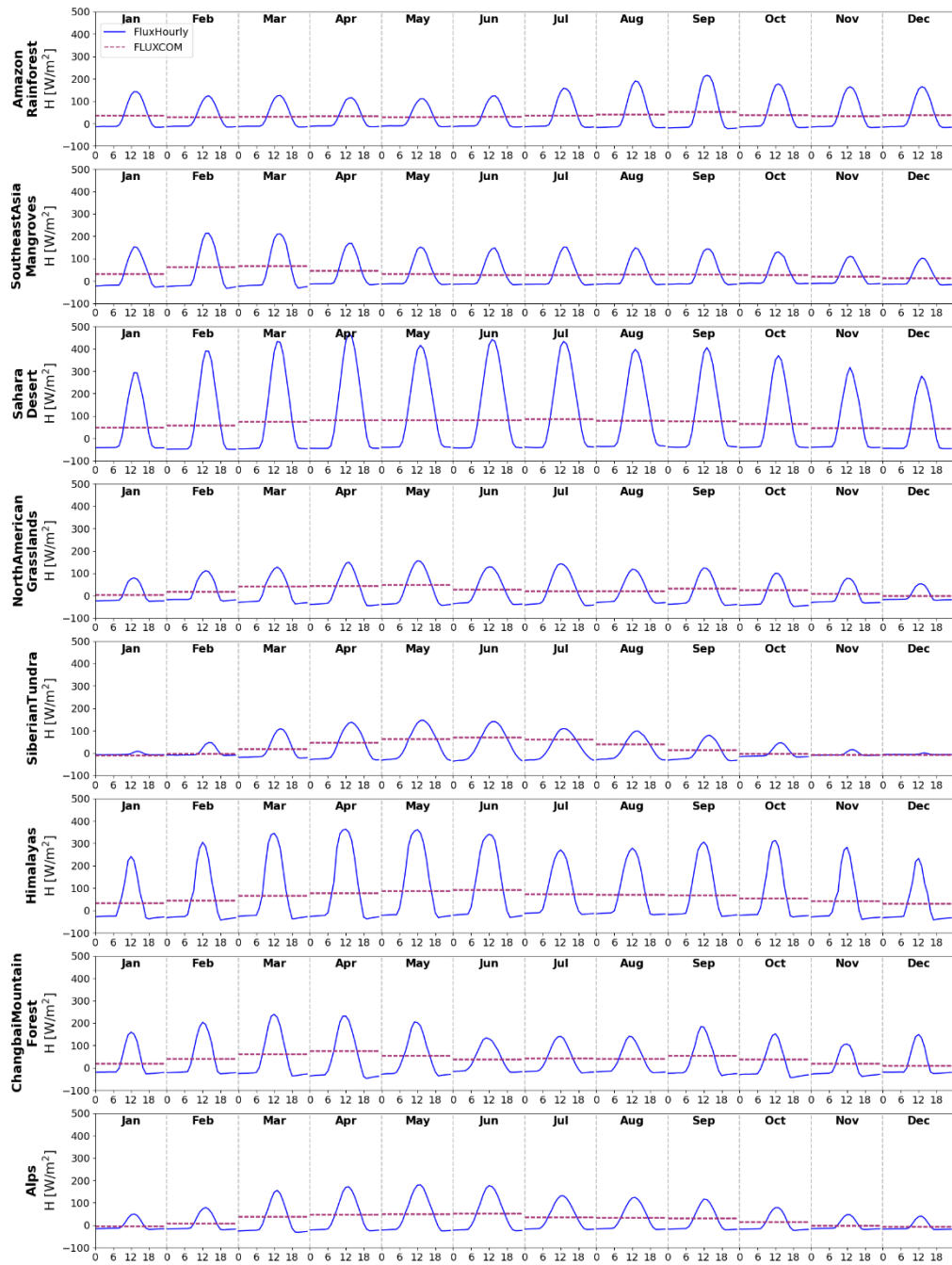


Figure S12. Diurnal cycles of  $H$  for 8 regions for each month of the year, where each panel refers to a region. The data is converted to local time zone.

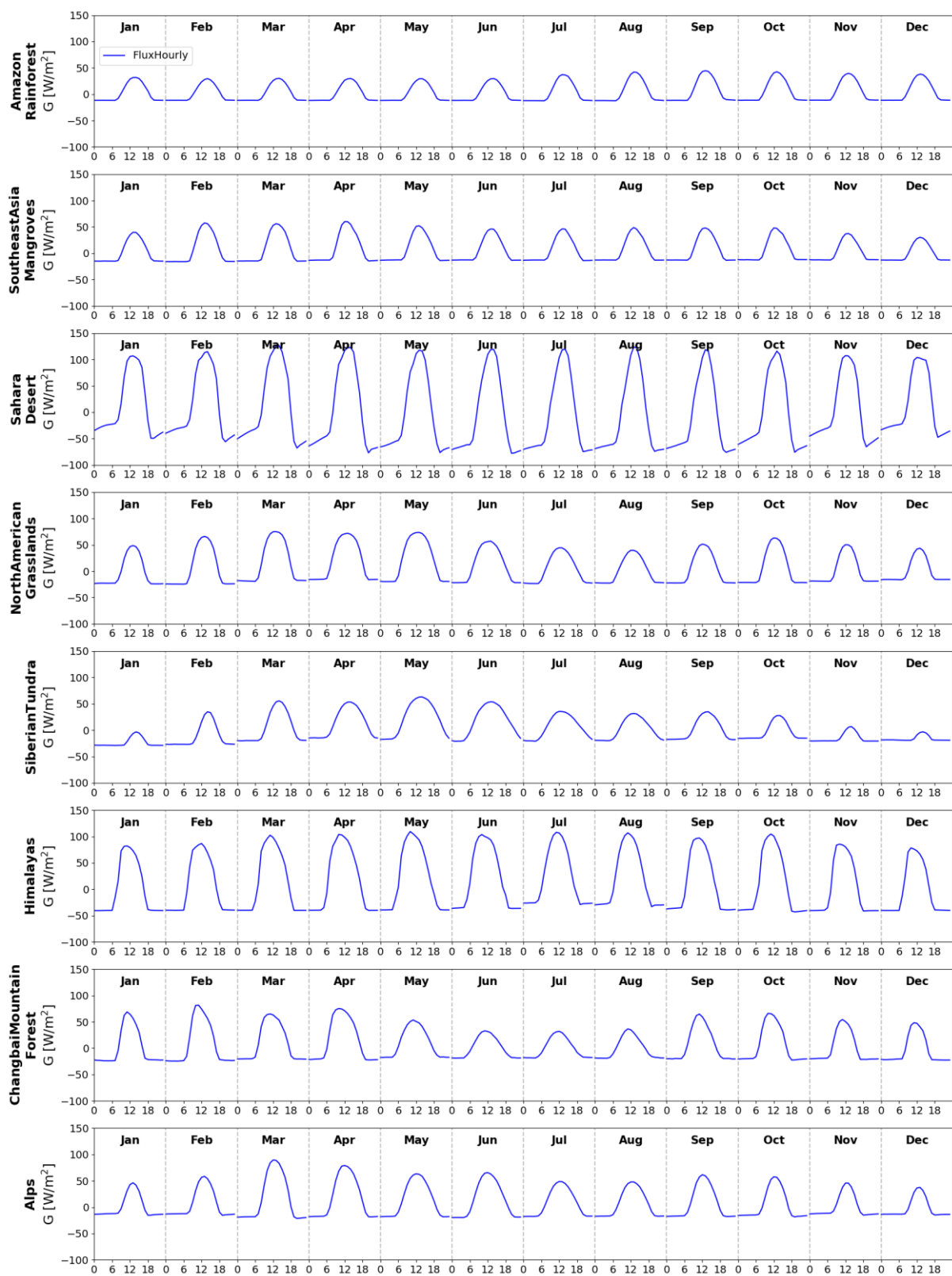
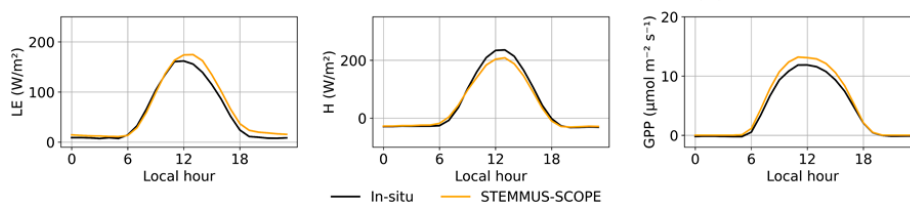


Figure S13. Diurnal cycles of  $G$  for 8 regions for each month of the year, where each panel refers to a region. The data is converted to local time zone.

## S5 Diurnal cycles for each IGBP class, global mean and trend

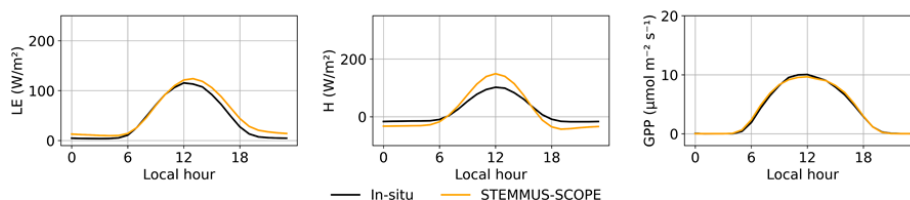
### Closed shrublands | 2005

Number of stations in 2005=2 / total number of stations in this igbp=2



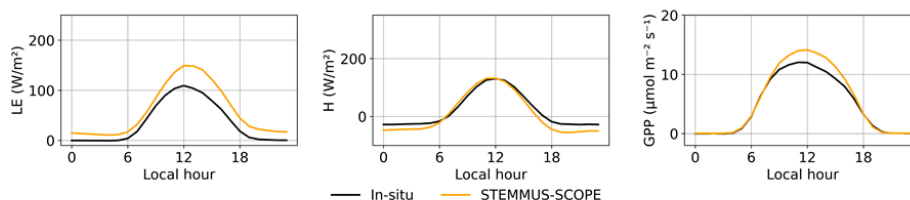
### Croplands | 2005

Number of stations in 2005=14 / total number of stations in this igbp=17



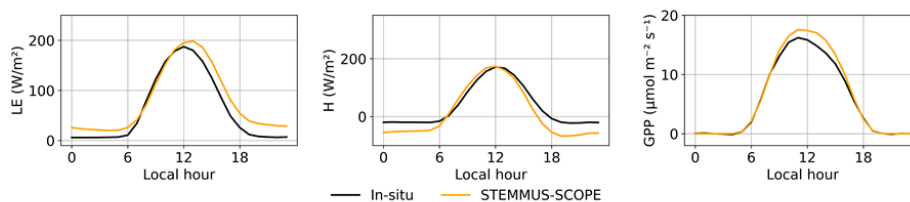
### Deciduous Broadleaf Forests | 2004

Number of stations in 2004=12 / total number of stations in this igbp=20



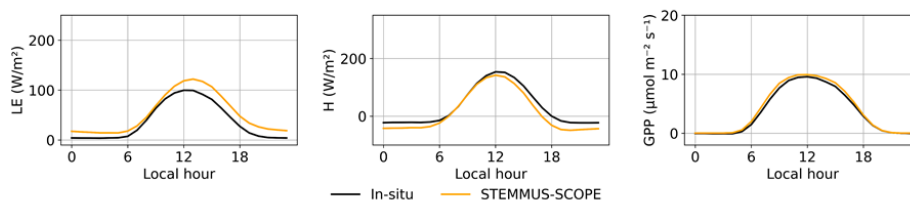
### Evergreen Broadleaf Forest | 2003

Number of stations in 2003=8 / total number of stations in this igbp=16



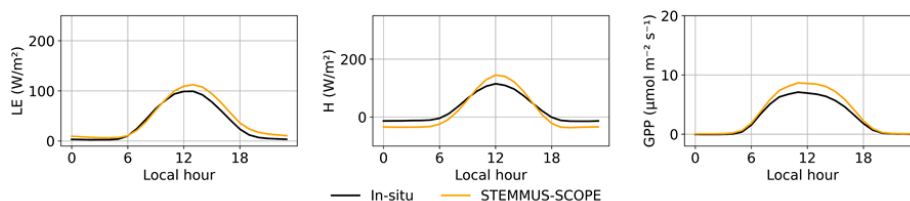
### Evergreen Needleleaf Forests | 2004

Number of stations in 2004=22 / total number of stations in this igbp=36



### Grasslands | 2005

Number of stations in 2005=19 / total number of stations in this igbp=40



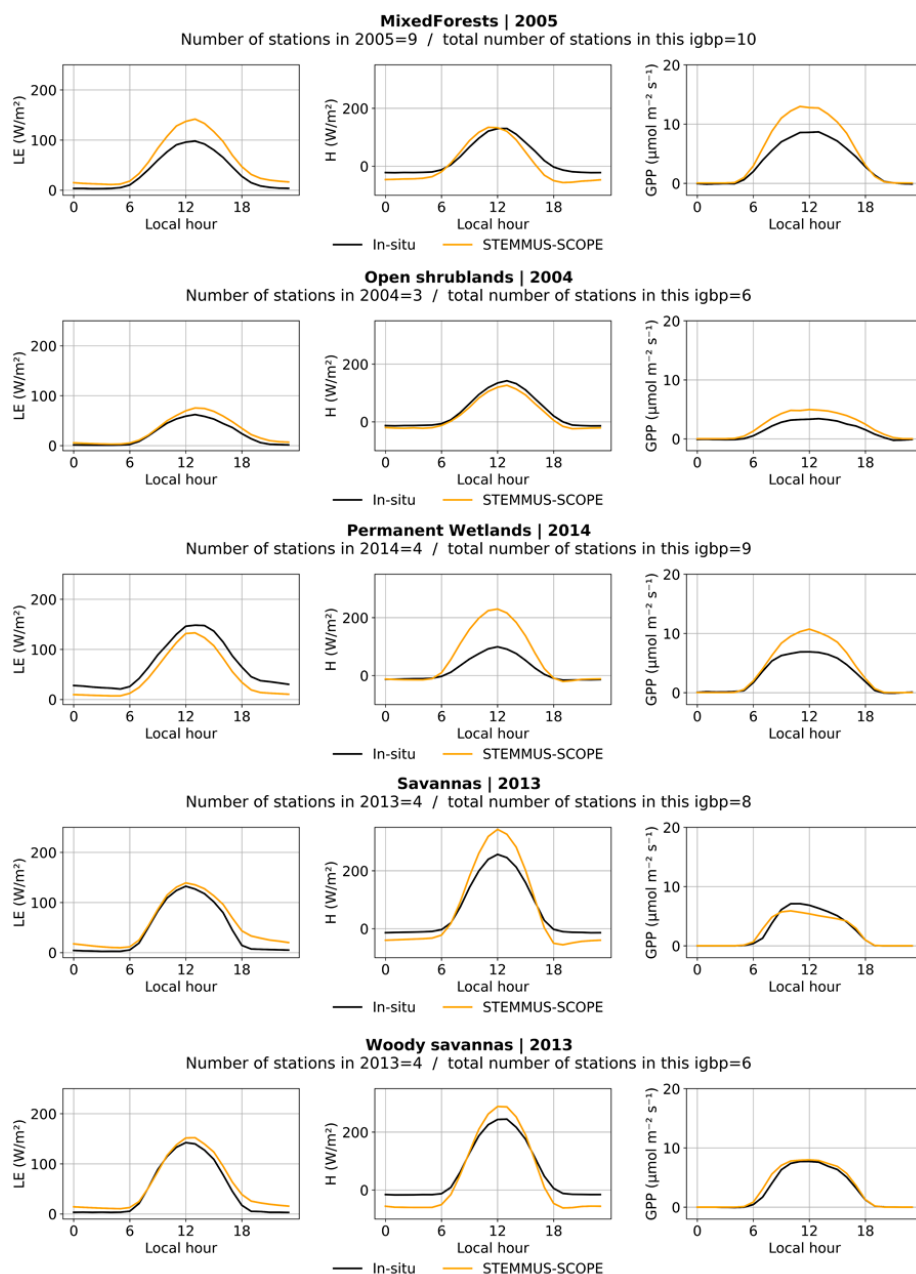
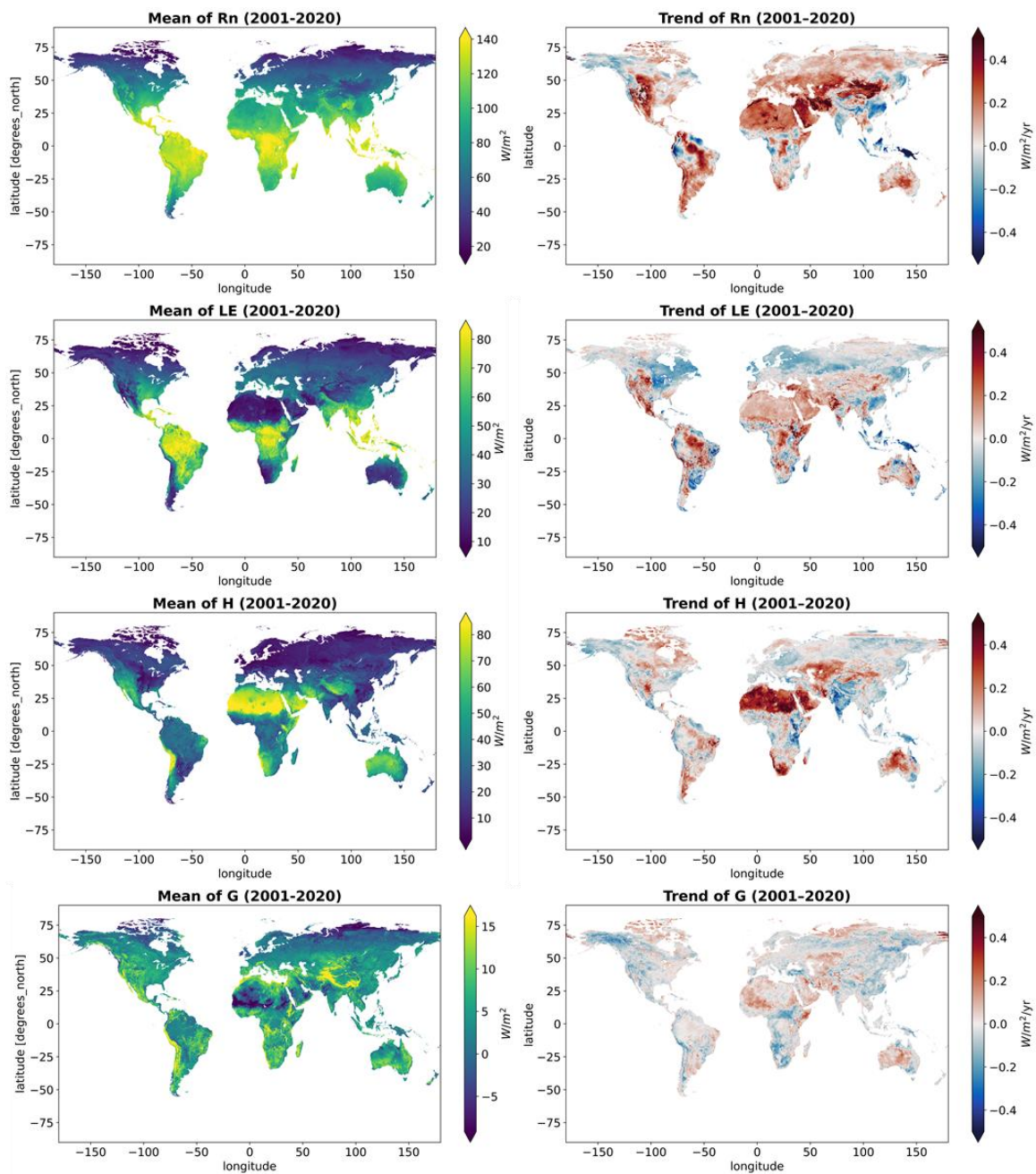


Figure S14. Diurnal cycles of LE, H, GPP for each IGBP class between In-situ measurements and STEMMUS-SCOPE output.



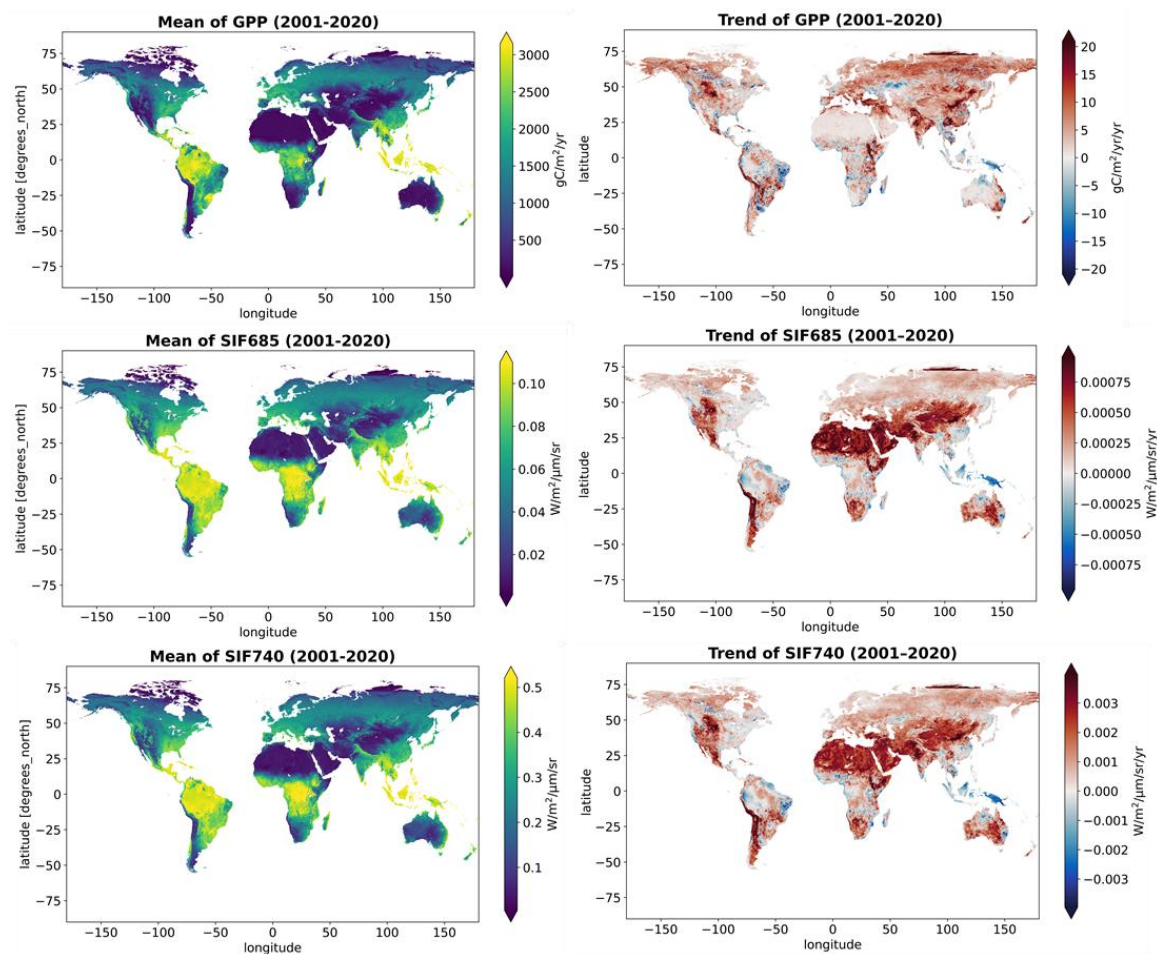


Figure S15. Global mean and yearly trends for Rn, LE, H, G, GPP, SIF685, SIF740 during 2001-2020.

#### S6 Testing on site level

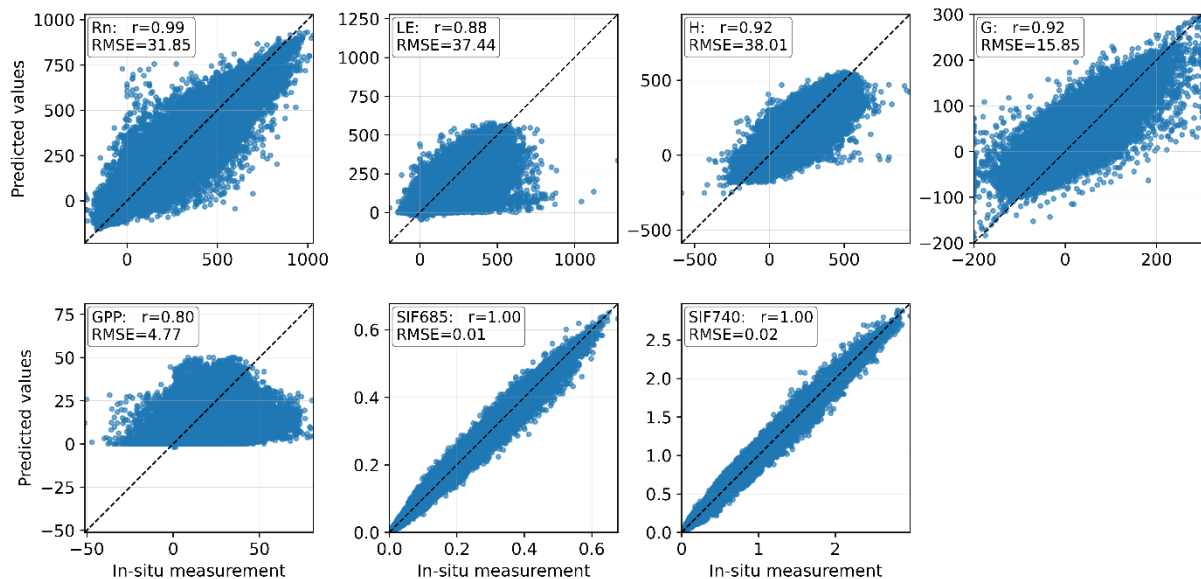


Figure S16. Comparison between predicted and in-situ measured fluxes for all variables on testing\_random set.

#### S7 Köppen–Geiger (KG) climate classification full name

Af – tropical rainforest climate,  
Am – tropical monsoon climate,

Aw – tropical wet and dry or savanna climate,  
BSh – hot semi-arid climate,  
BSk – cold semi-arid climate,  
BWh – hot desert climate,  
BWk – cold desert climate,  
Cfa – humid subtropical climate,  
Cfb – temperate oceanic climate,  
Cfc – subpolar oceanic climate  
Csa – hot-summer Mediterranean climate,  
Csb – warm-summer Mediterranean climate,  
Csc – cold-summer Mediterranean climate  
Cwa – monsoon-influenced humid subtropical climate,  
Cwb – subtropical highland climate,  
Cwc – cold subtropical highland climate  
Dfa – hot-summer humid continental climate,  
Dfb – warm-summer humid continental climate,  
Dfc – subarctic climate,  
Dfd – extremely cold subarctic climate,  
Dsa – mediterranean-influenced hot-summer humid continental climate,  
Dsb – mediterranean-influenced warm-summer humid continental climate,  
Dsc – mediterranean-influenced subarctic climate,  
Dwa – monsoon-influenced hot-summer humid continental climate,  
Dwb – monsoon-influenced warm-summer humid continental climate,  
Dwc – monsoon-influenced subarctic climate,  
Dwd – monsoon-influenced extremely cold subarctic climate.  
EF – ice cap climate.  
ET – tundra climate.

## References

- Abramowitz, G., Ukkola, A., Hobeichi, S., Cranko Page, J., Lipson, M., De Kauwe, M.G., Green, S., Brenner, C., Frame, J., & Nearing, G.: On the predictability of turbulent fluxes from land: PLUMBER2 MIP experimental description and preliminary results, *Biogeosciences*, 21, 5517-5538, <https://doi.org/10.5194/bg-21-5517-2024>, 2024.
- Wang, Y., Zeng, Y., Alidoost, F., Schilperoort, B., Song, Z., Yu, D., Tang, E., Han, Q., Liu, Z., & Peng, X.: A physically consistent dataset of water-energy-carbon fluxes across the Soil-Plant-Atmosphere Continuum, *Sci. Data*, 12, 1146, <https://doi.org/10.1038/s41597-025-05386-x>, 2025.