Supplementary information to “Global long-term hourly 9 km terrestrial water-energy-carbon fluxes (FluxHourly), by Qianqian Han, Yijian Zeng, Yunfei Wang, Fakhereh (Sarah) Alidoost, Francesco Nattino, Yang Liu, Bob Su

Section 1 Protocol for data processing of fluxes samples (training & testing data)

We have developed the following protocol for data processing of eddy-covariance fluxes.

(1) We first plot a figure with LE/Rn and Pre/Rn to detect outlier stations (Abramowitz et al. 2024). We aggregated half-hourly original values into annual values to calculate LE/Rn and Pre/Rn. From Figure S1, 5 stations shows abnormal values, then we checked one by one the variables. The precipitation at station AU-Rob has problems, so this station was removed. AU-Ctr and AU-Cow just have high precipitation. FR-Lq1 and FR-Lq2 has abnormal Rn.

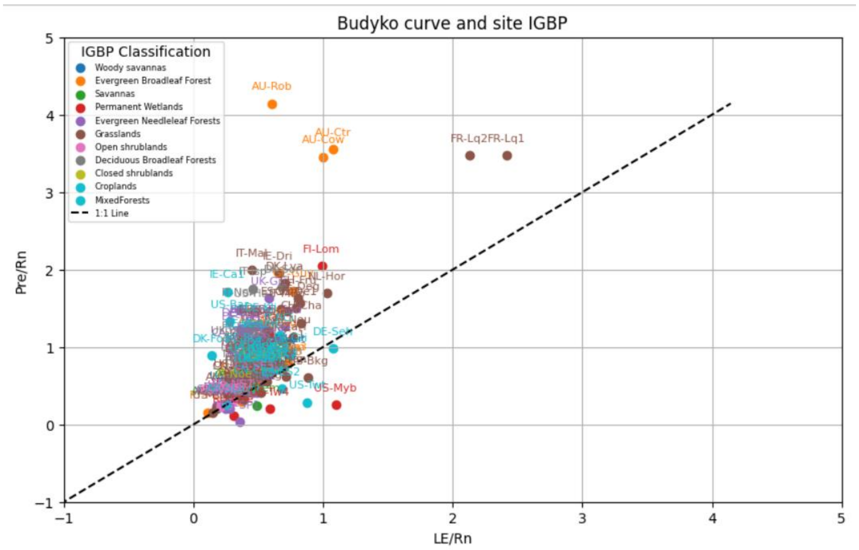


Figure S1. Distribution of site IGBP vegetation types on a Budyko style plot, using in-situ fluxes for all 170 sites.

(2) Use quality control of LE and H to keep only measurements, removing interpolated values.

(3) When Precipitation>0 & LE > (Rn – H), then we reject in-situ LE, H, implying that the measured LE has been contaminated by precipitation event.

(4) For Rn, only IT-LMa, FR-Lq1, FR-Lq2, IT-CA1, US-SP2 has strange values. These outliers were removed manually.

(5) Limit STEMMUS-SCOPE output LE and H based on the rule devised by (Wang et al. 2025).
If G > 300, then G = 300; if G < -200, then G = -200.
If H < -250, then H = -250; if Rin > 10 & H > Rn, then H = Rn – G, LE = 0; if Rin <= 10 & H > 100, then H = Rn – G, LE = 0.
If Rin > 10 & LE > 1.26 Rn, then LE = Rn – G – H; if Rin <= 10 & LE > 100, then LE = Rn, H = - G.

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

(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.

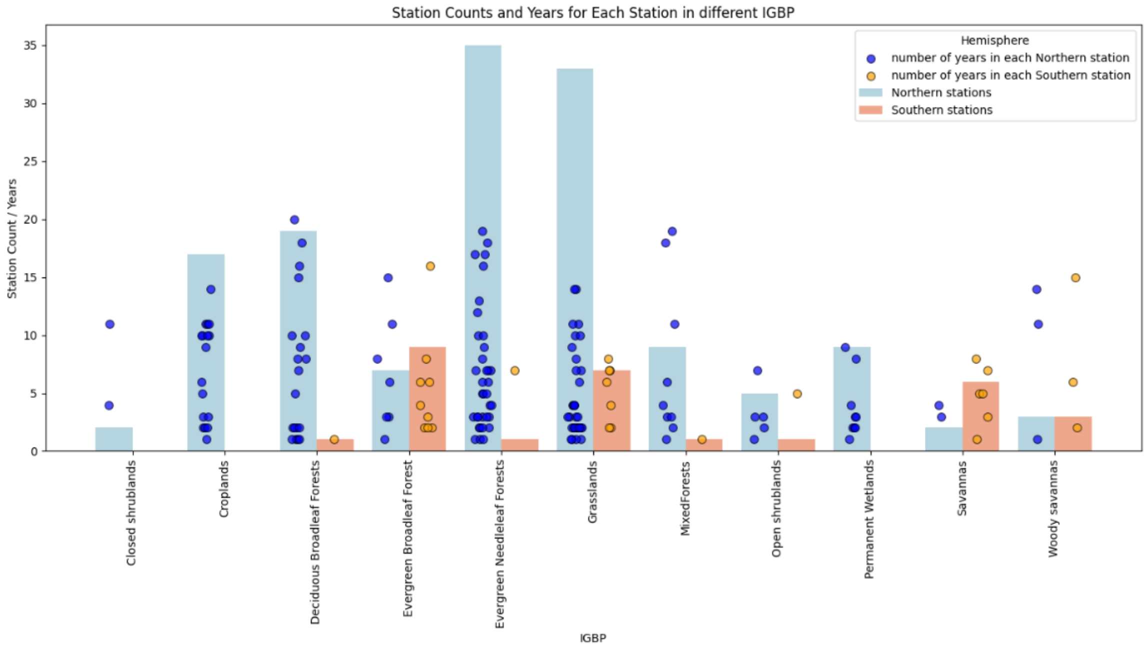
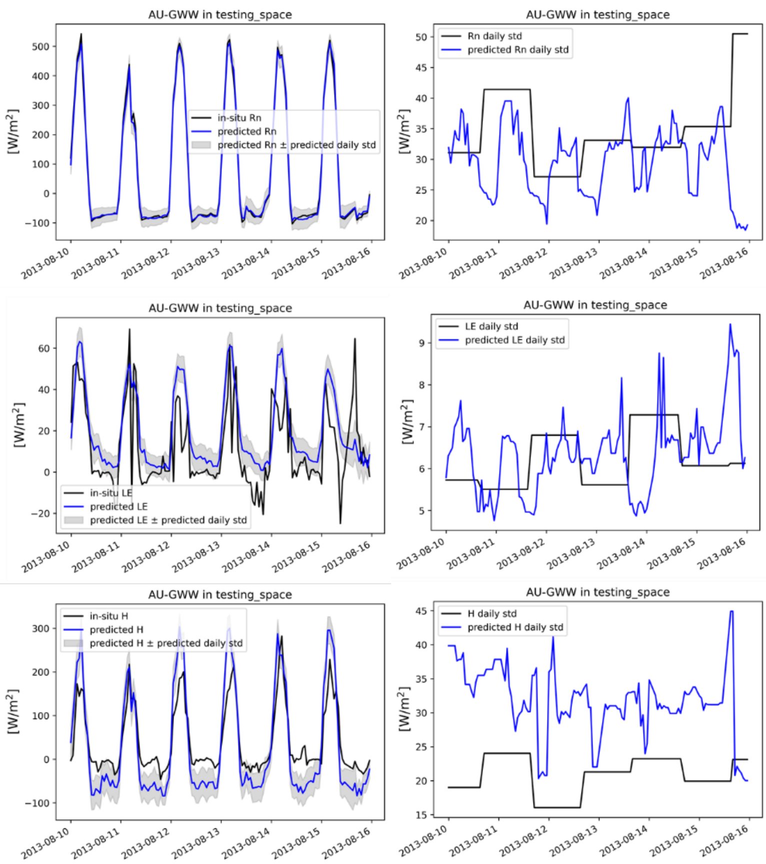


Figure S2. Stations statistics over each IGBP land cover class and each hemisphere

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.



AU-GWW in testing_space
in-situ Rn
predicted Rn
predicted Rn ± predicted daily std
Rn daily std
predicted Rn daily std
in-situ LE
predicted LE
predicted LE ± predicted daily std
LE daily std
predicted LE daily std
in-situ H
predicted H
predicted H ± predicted daily std
H daily std
predicted H daily std
[W/m2]
2013-08-10
2013-08-11
2013-08-12
2013-08-13
2013-08-14
2013-08-15
2013-08-16

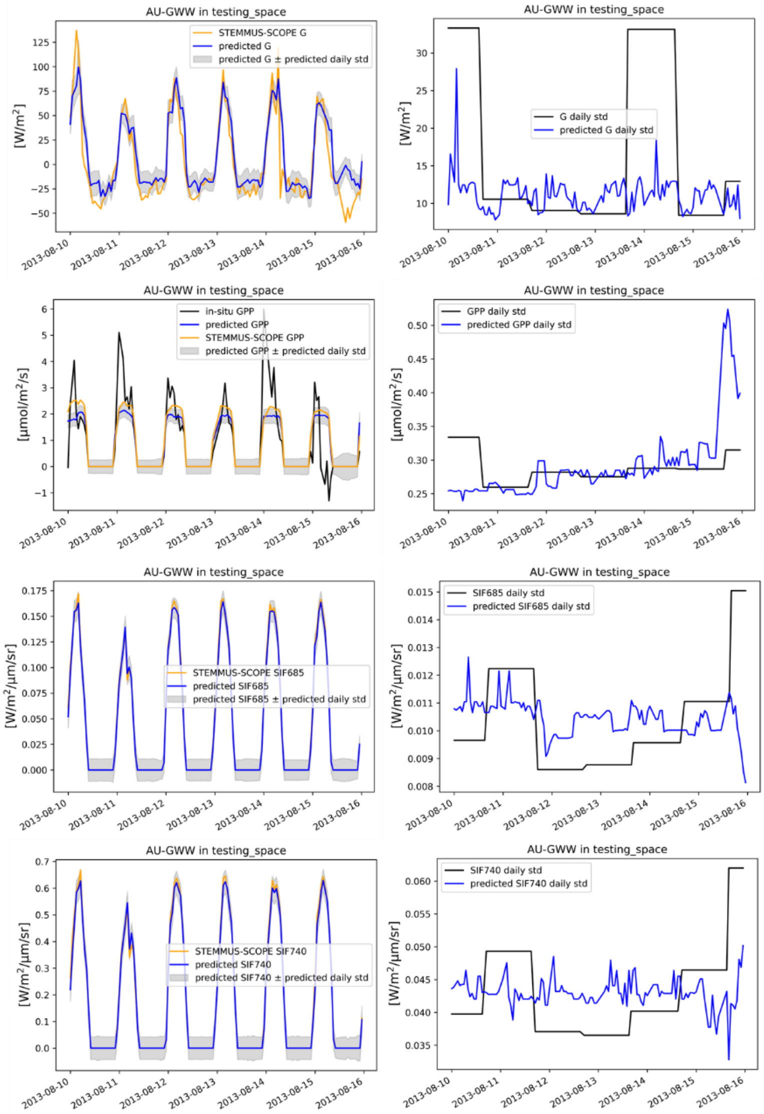
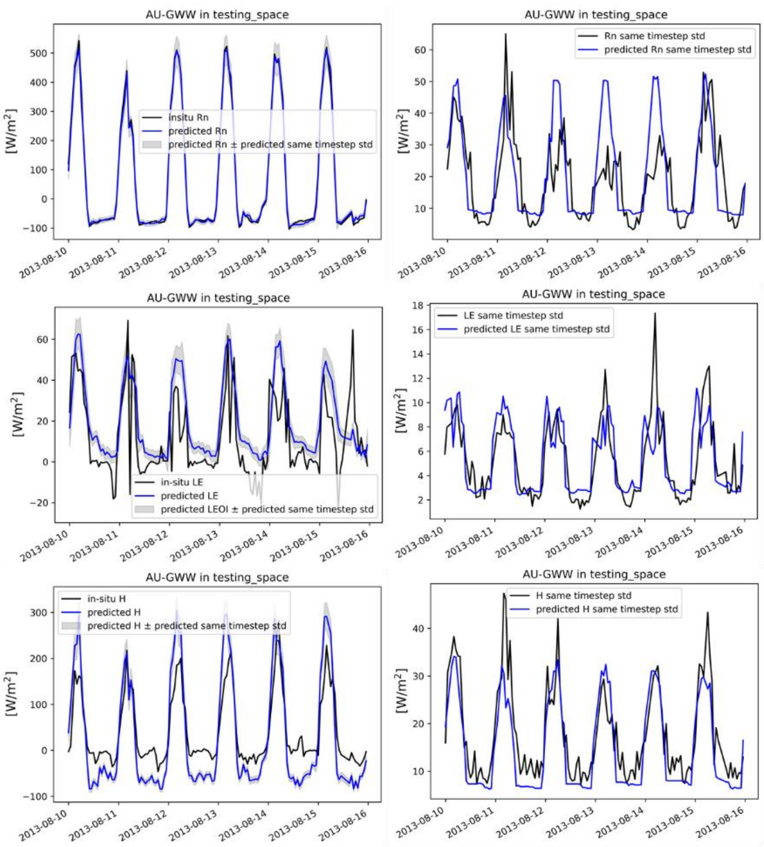


Figure S3. In-situ fluxes, predicted fluxes, and predicted daily std on 7 variables in AU-GWW station (right figures are calculated std and predicted std)



AU-GWW in testing_space
insitu Rn
predicted Rn
predicted Rn ± predicted same timestep std
[W/m²]
AU-GWW in testing_space
Rn same timestep std
predicted Rn same timestep std
[W/m²]
AU-GWW in testing_space
in-situ LE
predicted LE
predicted LEOI ± predicted same timestep std
[W/m²]
AU-GWW in testing_space
LE same timestep std
predicted LE same timestep std
[W/m²]
AU-GWW in testing_space
in-situ H
predicted H
predicted H ± predicted same timestep std
[W/m²]
AU-GWW in testing_space
H same timestep std
predicted H same timestep std
[W/m²]
2013-08-10
2013-08-11
2013-08-12
2013-08-13
2013-08-14
2013-08-15
2013-08-16

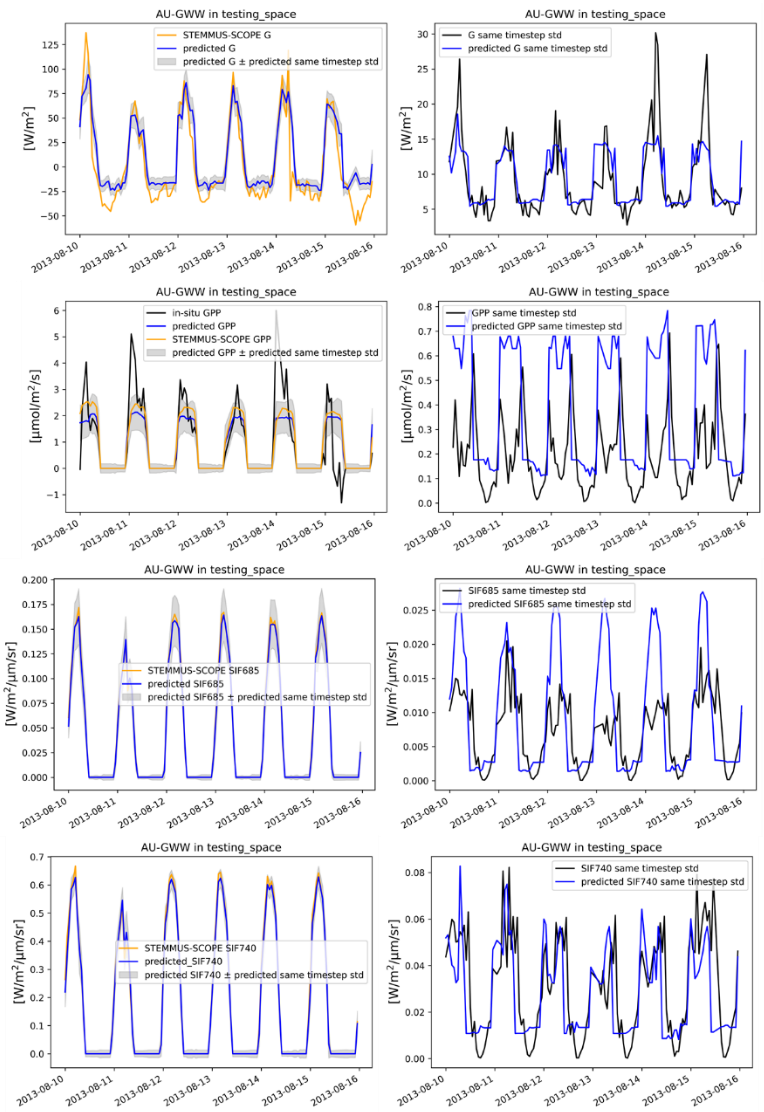


Figure S4. In-situ fluxes, predicted fluxes, and predicted same timestep std on 7 variables in AU-GWW station (right figures are calculated std and predicted std)

Section 2. Comparison on testing_space set on monthly & hourly scale

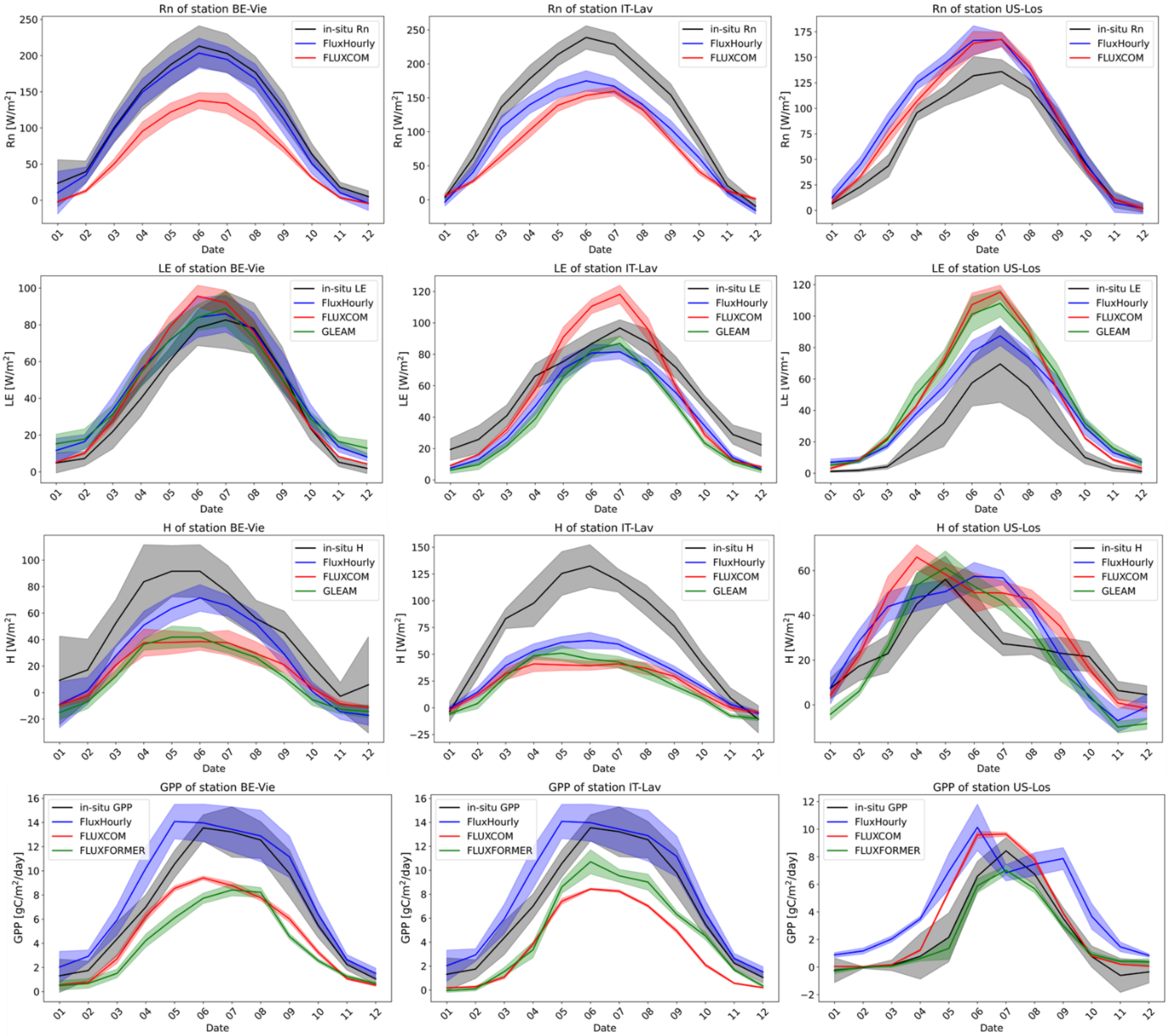
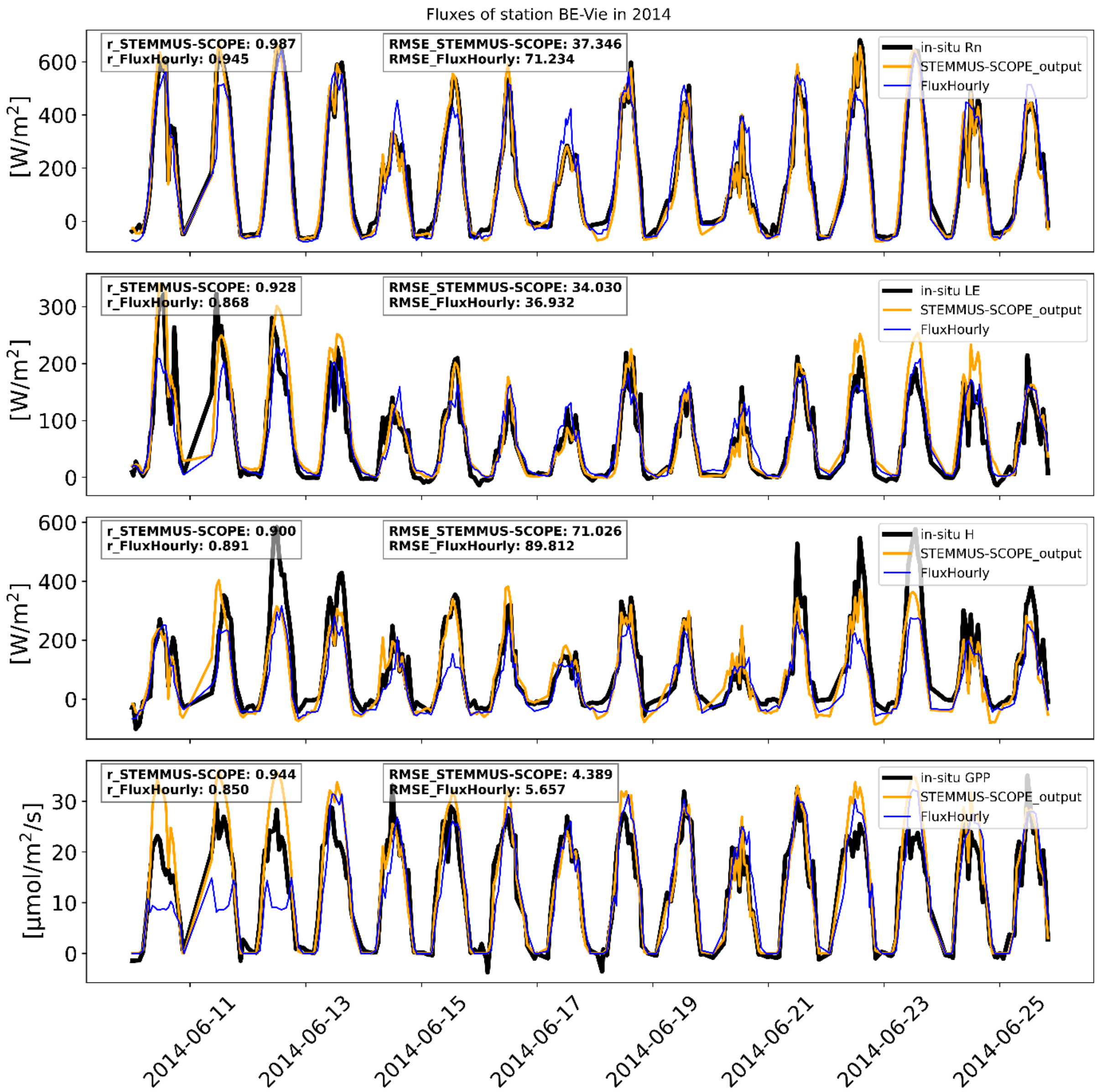
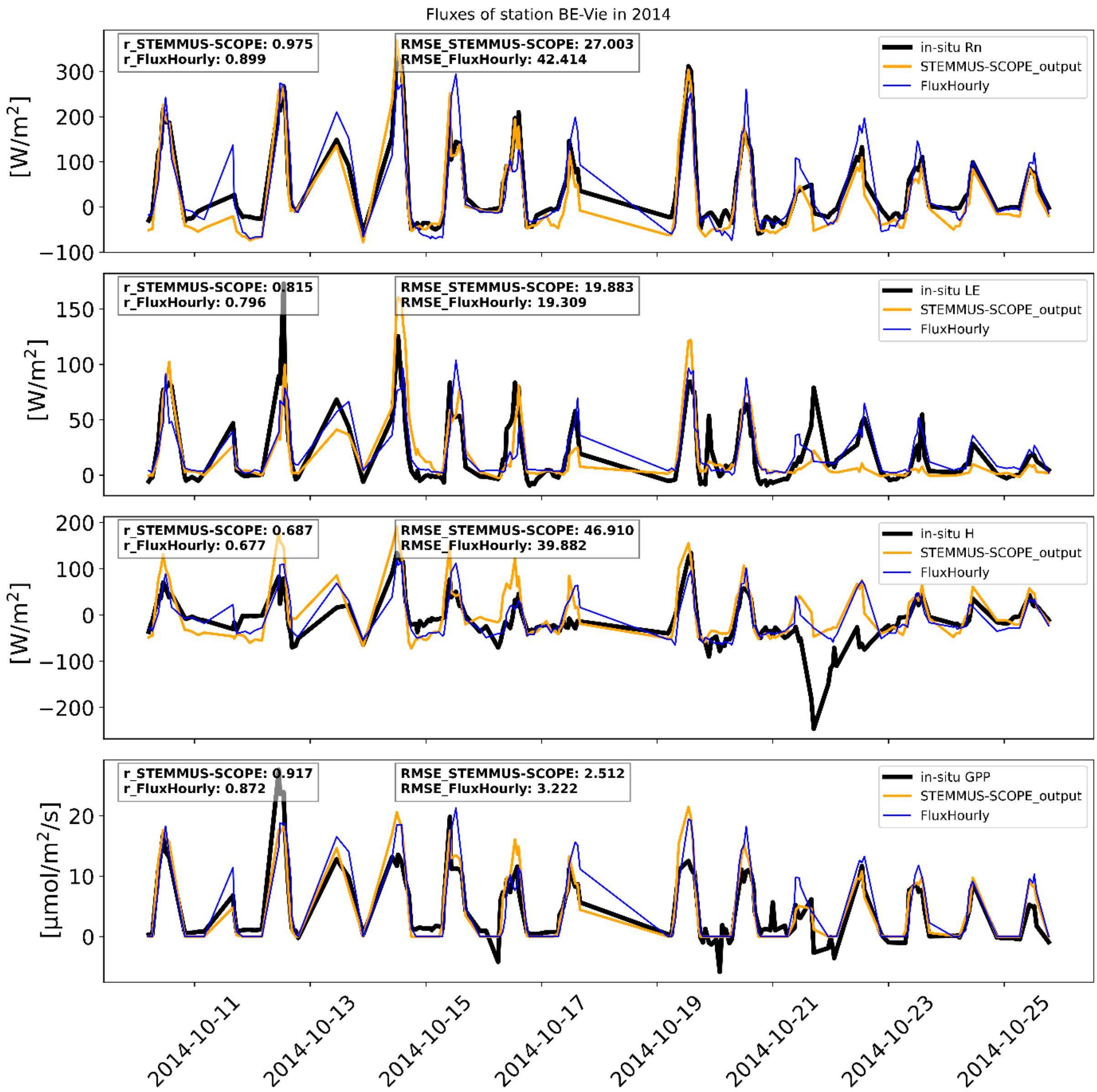


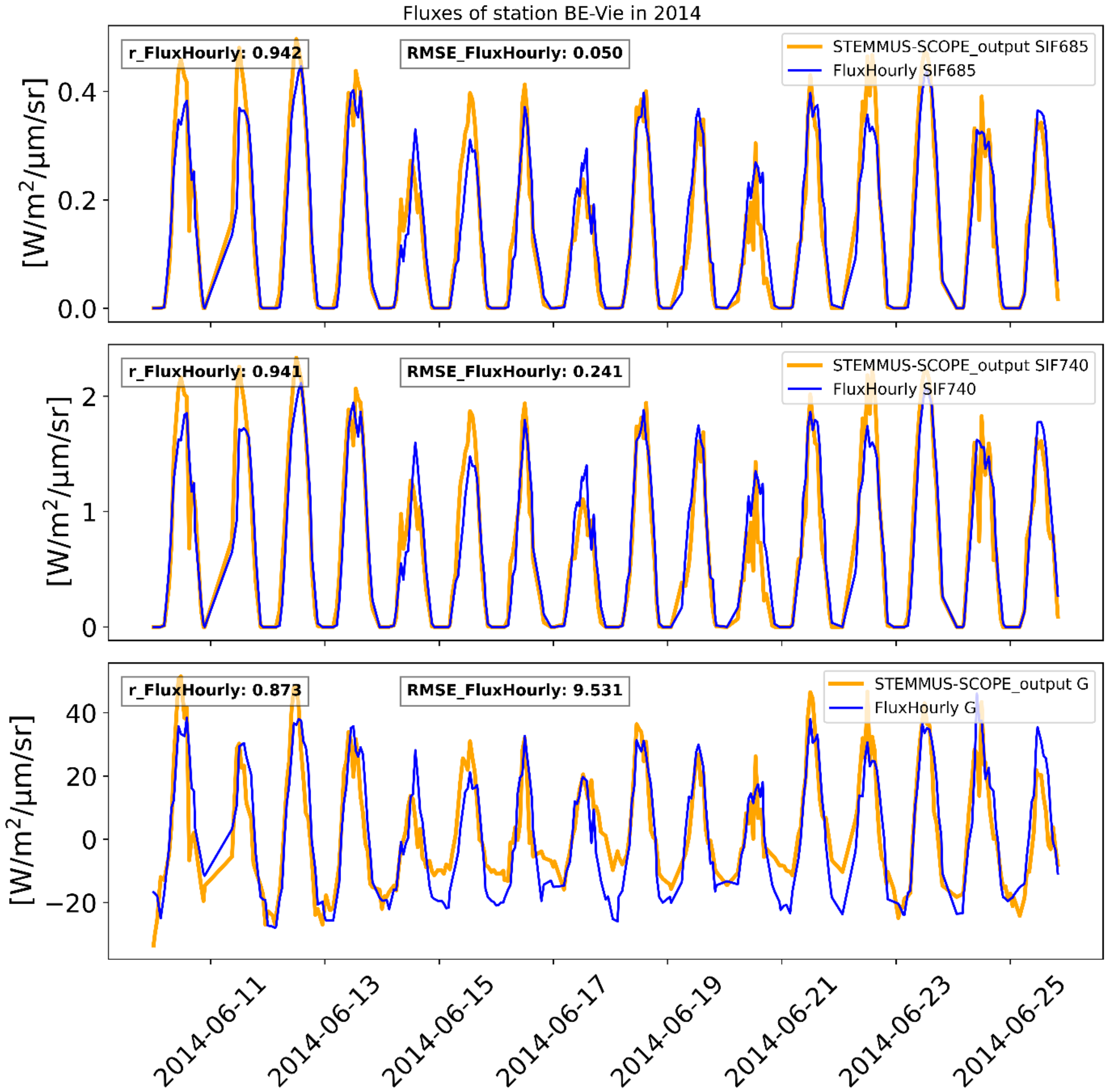
Figure S5. Monthly Rn, LE, H, GPP from in-situ, FluxHourly, FLUXCOM, GLEAM, FLUXFORMER in station BE-Vie, IT-Lav, US-Los



Fluxes of station BE-Vie in 2014
r_STEMMUS-SCOPE: 0.987
r_FluxHourly: 0.945
RMSE_STEMMUS-SCOPE: 37.346
RMSE_FluxHourly: 71.234
in-situ Rn
STEMMUS-SCOPE_output
FluxHourly
[W/m^2]
r_STEMMUS-SCOPE: 0.928
r_FluxHourly: 0.868
RMSE_STEMMUS-SCOPE: 34.030
RMSE_FluxHourly: 36.932
in-situ LE
STEMMUS-SCOPE_output
FluxHourly
[W/m^2]
r_STEMMUS-SCOPE: 0.900
r_FluxHourly: 0.891
RMSE_STEMMUS-SCOPE: 71.026
RMSE_FluxHourly: 89.812
in-situ H
STEMMUS-SCOPE_output
FluxHourly
[W/m^2]
r_STEMMUS-SCOPE: 0.944
r_FluxHourly: 0.850
RMSE_STEMMUS-SCOPE: 4.389
RMSE_FluxHourly: 5.657
in-situ GPP
STEMMUS-SCOPE_output
FluxHourly
[µmol/m^2/s]
2014-06-11
2014-06-13
2014-06-15
2014-06-17
2014-06-19
2014-06-21
2014-06-23
2014-06-25



Fluxes of station BE-Vie in 2014
r_STEMMUS-SCOPE: 0.975
r_FluxHourly: 0.899
RMSE_STEMMUS-SCOPE: 27.003
RMSE_FluxHourly: 42.414
in-situ Rn
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.815
r_FluxHourly: 0.796
RMSE_STEMMUS-SCOPE: 19.883
RMSE_FluxHourly: 19.309
in-situ LE
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.687
r_FluxHourly: 0.677
RMSE_STEMMUS-SCOPE: 46.910
RMSE_FluxHourly: 39.882
in-situ H
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.917
r_FluxHourly: 0.872
RMSE_STEMMUS-SCOPE: 2.512
RMSE_FluxHourly: 3.222
in-situ GPP
STEMMUS-SCOPE_output
FluxHourly
[μmol/m2/s]
2014-10-11
2014-10-13
2014-10-15
2014-10-17
2014-10-19
2014-10-21
2014-10-23
2014-10-25



Fluxes of station BE-Vie in 2014
r_FluxHourly: 0.942
RMSE_FluxHourly: 0.050
STEMMUS-SCOPE_output SIF685
FluxHourly SIF685
[W/m2/μm/sr]
r_FluxHourly: 0.941
RMSE_FluxHourly: 0.241
STEMMUS-SCOPE_output SIF740
FluxHourly SIF740
[W/m2/μm/sr]
r_FluxHourly: 0.873
RMSE_FluxHourly: 9.531
STEMMUS-SCOPE_output G
FluxHourly G
[W/m2/μm/sr]
2014-06-11
2014-06-13
2014-06-15
2014-06-17
2014-06-19
2014-06-21
2014-06-23
2014-06-25

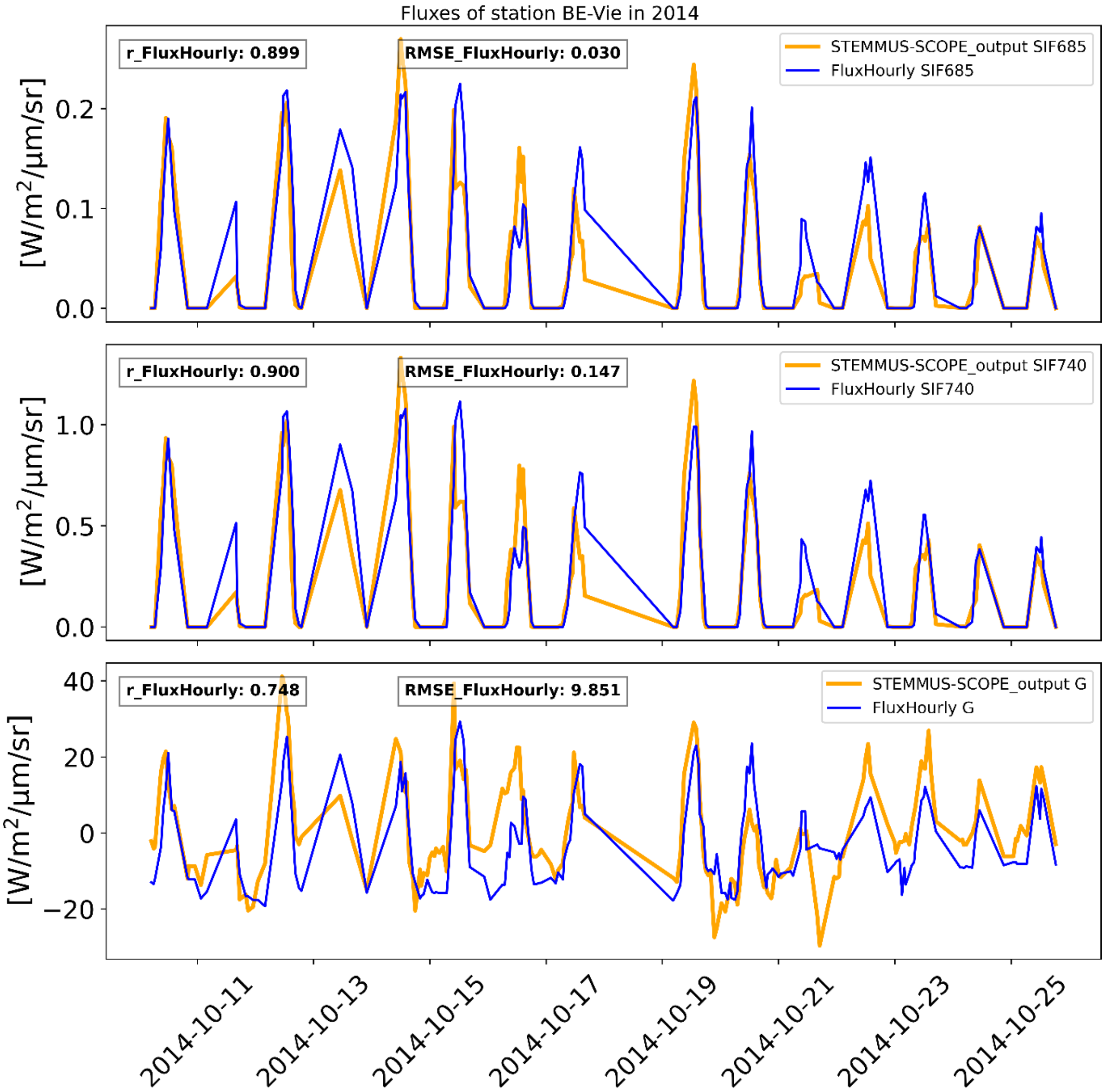
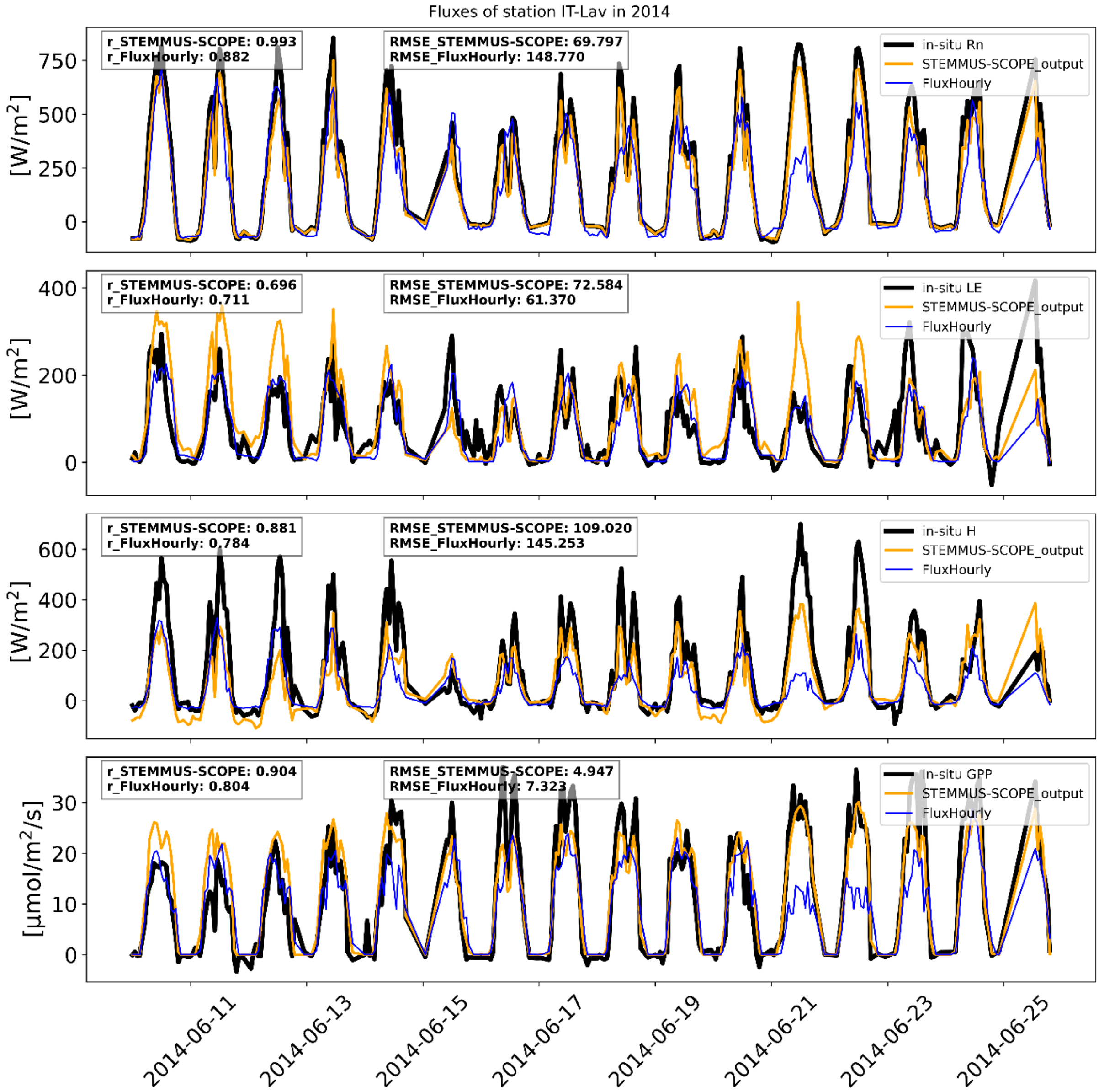
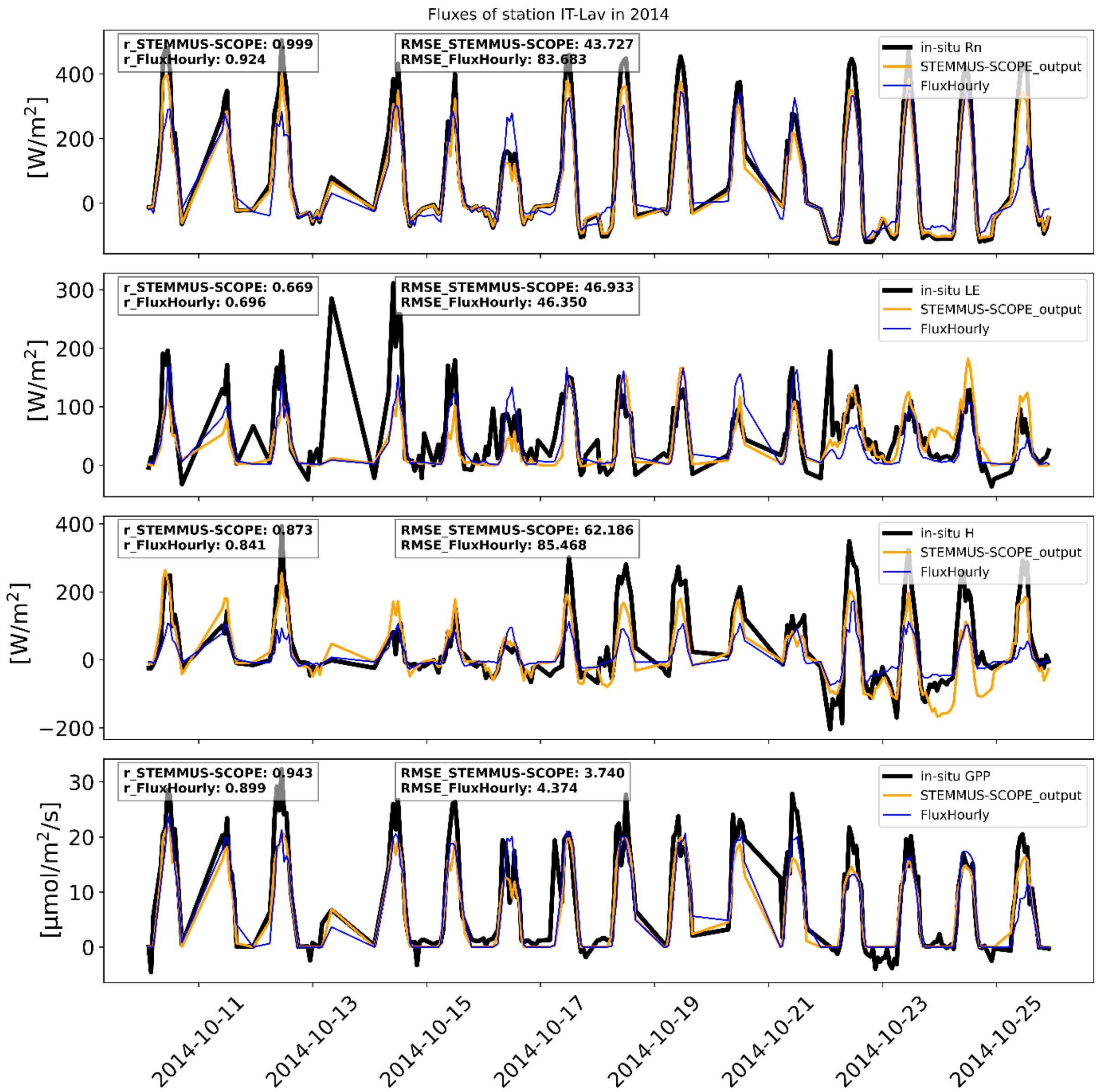
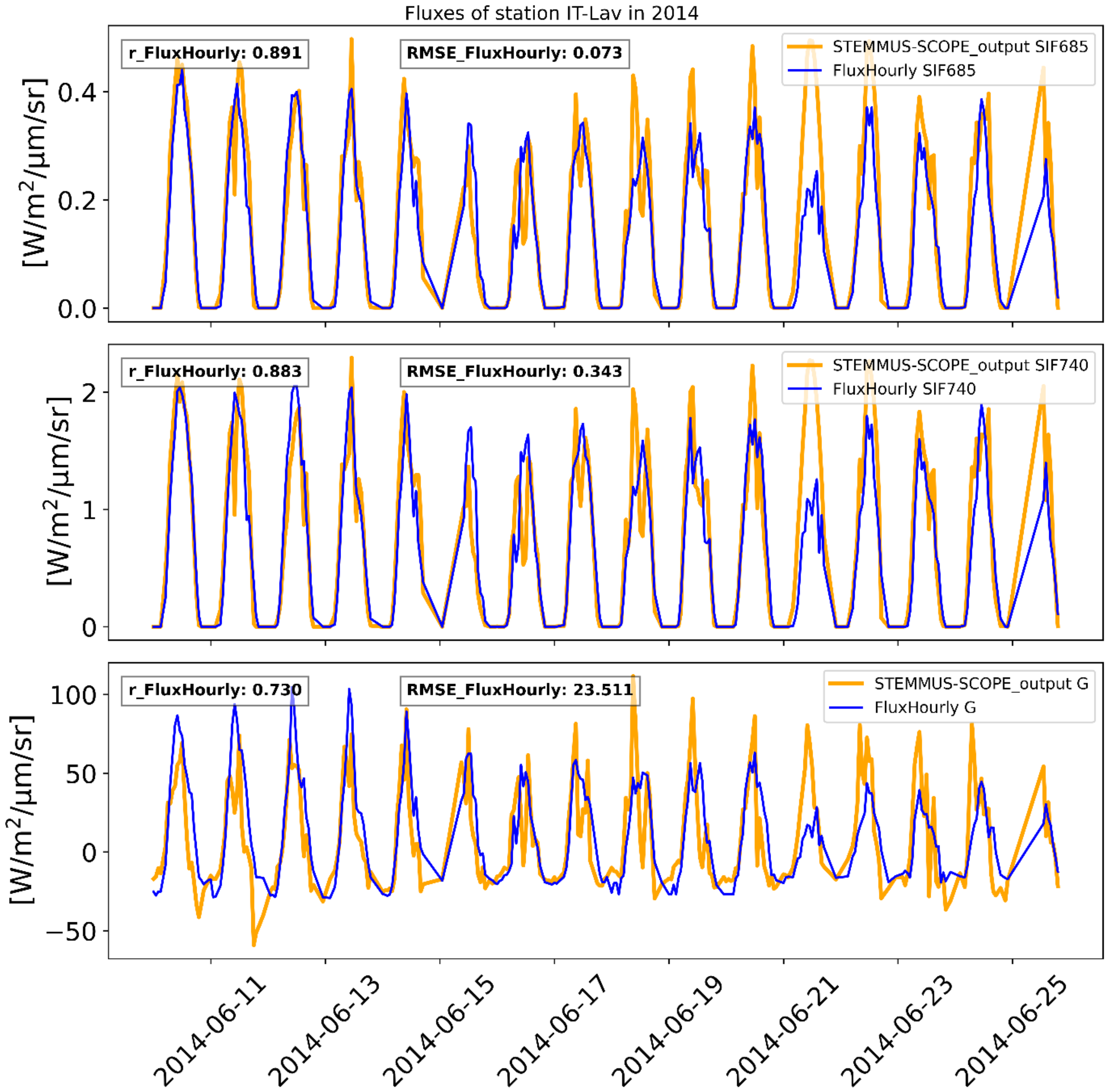


Figure S6. Rn, LE, H, G, GPP, SIF685, SIF740 in BE-Vie, June and Oct (time is in local time)


Fluxes of station IT-Lav in 2014
r_STEMMUS-SCOPE: 0.993
r_FluxHourly: 0.882
RMSE_STEMMUS-SCOPE: 69.797
RMSE_FluxHourly: 148.770
in-situ Rn
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.696
r_FluxHourly: 0.711
RMSE_STEMMUS-SCOPE: 72.584
RMSE_FluxHourly: 61.370
in-situ LE
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.881
r_FluxHourly: 0.784
RMSE_STEMMUS-SCOPE: 109.020
RMSE_FluxHourly: 145.253
in-situ H
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.904
r_FluxHourly: 0.804
RMSE_STEMMUS-SCOPE: 4.947
RMSE_FluxHourly: 7.323
in-situ GPP
STEMMUS-SCOPE_output
FluxHourly
[μmol/m2/s]
2014-06-11
2014-06-13
2014-06-15
2014-06-17
2014-06-19
2014-06-21
2014-06-23
2014-06-25


Fluxes of station IT-Lav in 2014
r_STEMMUS-SCOPE: 0.999
r_FluxHourly: 0.924
RMSE_STEMMUS-SCOPE: 43.727
RMSE_FluxHourly: 83.683
in-situ Rn
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.669
r_FluxHourly: 0.696
RMSE_STEMMUS-SCOPE: 46.933
RMSE_FluxHourly: 46.350
in-situ LE
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.873
r_FluxHourly: 0.841
RMSE_STEMMUS-SCOPE: 62.186
RMSE_FluxHourly: 85.468
in-situ H
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.943
r_FluxHourly: 0.899
RMSE_STEMMUS-SCOPE: 3.740
RMSE_FluxHourly: 4.374
in-situ GPP
STEMMUS-SCOPE_output
FluxHourly
[μmol/m2/s]
2014-10-11
2014-10-13
2014-10-15
2014-10-17
2014-10-19
2014-10-21
2014-10-23
2014-10-25



Fluxes of station IT-Lav in 2014
r_FluxHourly: 0.891
RMSE_FluxHourly: 0.073
STEMMUS-SCOPE_output SIF685
FluxHourly SIF685
[W/m²/μm/sr]
0.0
0.2
0.4
r_FluxHourly: 0.883
RMSE_FluxHourly: 0.343
STEMMUS-SCOPE_output SIF740
FluxHourly SIF740
[W/m²/μm/sr]
0
1
2
r_FluxHourly: 0.730
RMSE_FluxHourly: 23.511
STEMMUS-SCOPE_output G
FluxHourly G
[W/m²/μm/sr]
−50
0
50
100
2014-06-11
2014-06-13
2014-06-15
2014-06-17
2014-06-19
2014-06-21
2014-06-23
2014-06-25

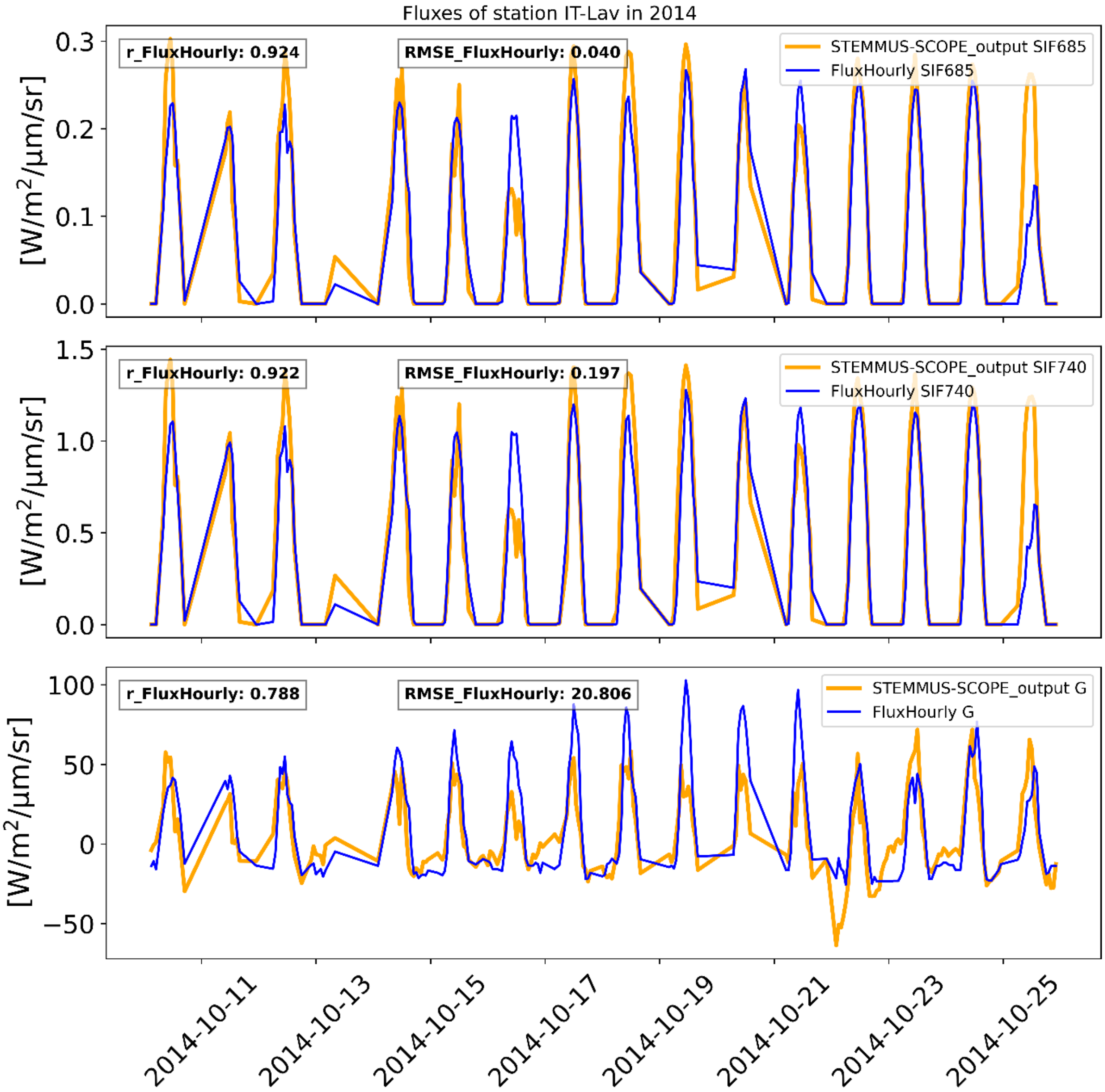
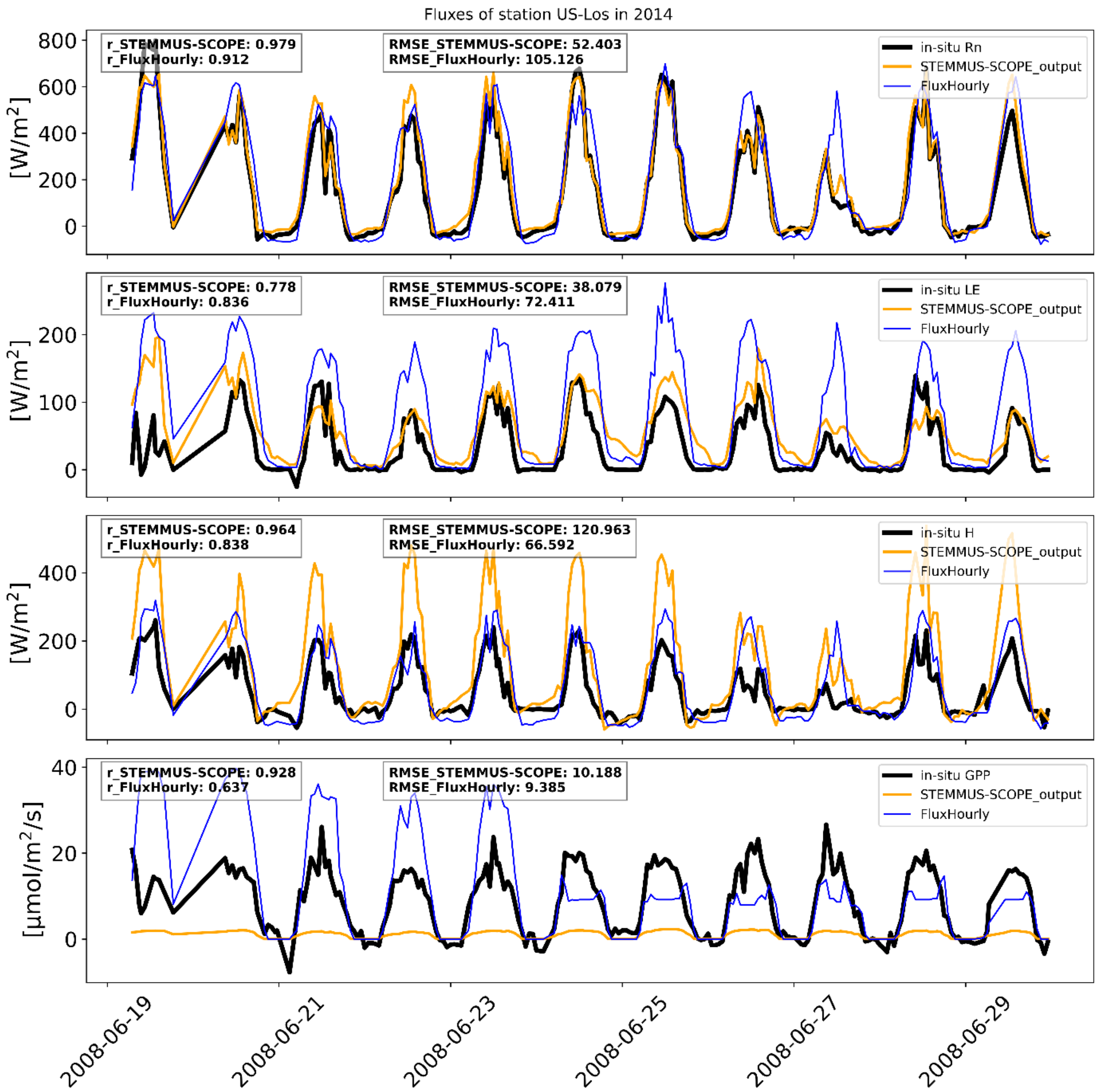
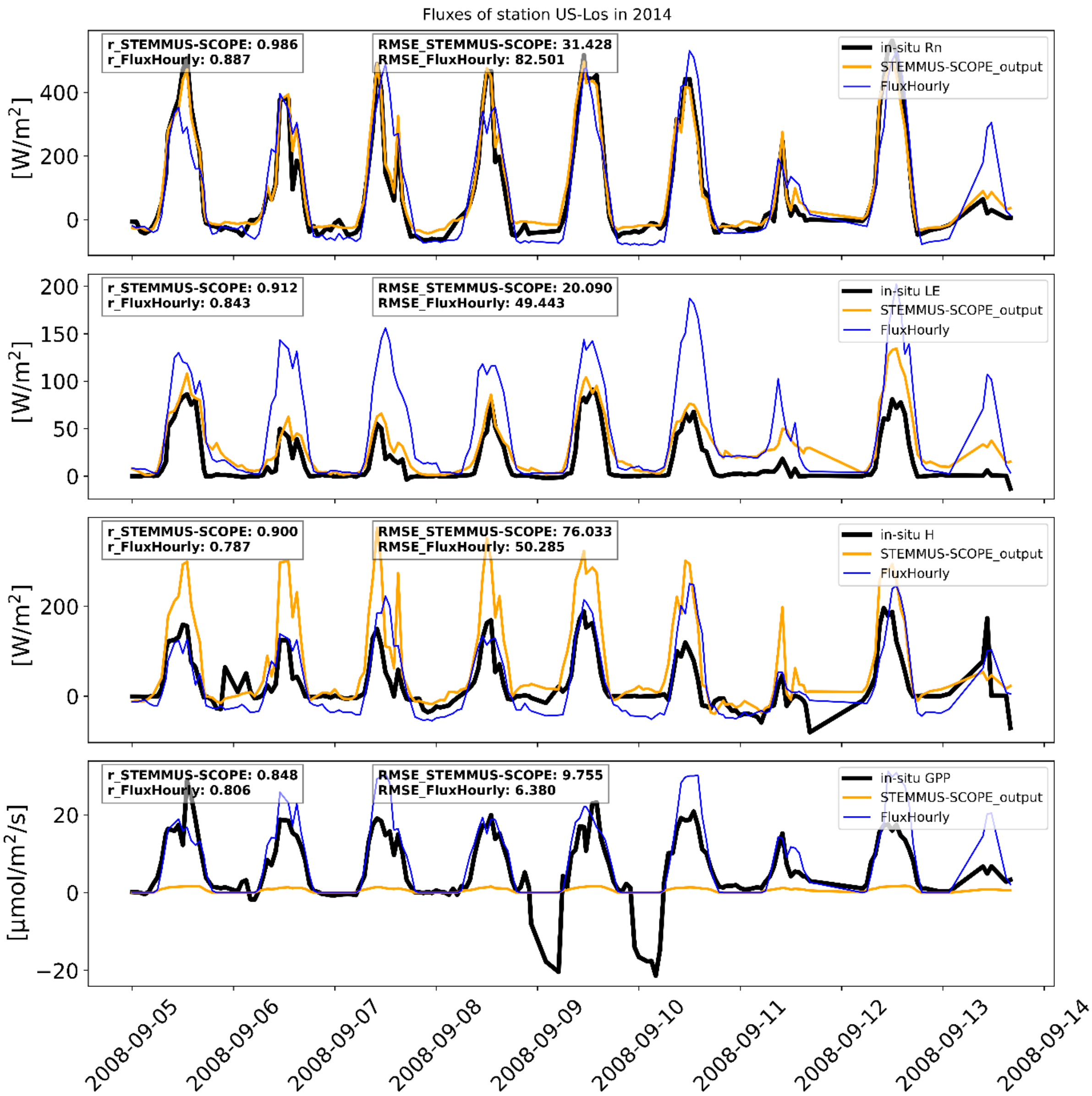


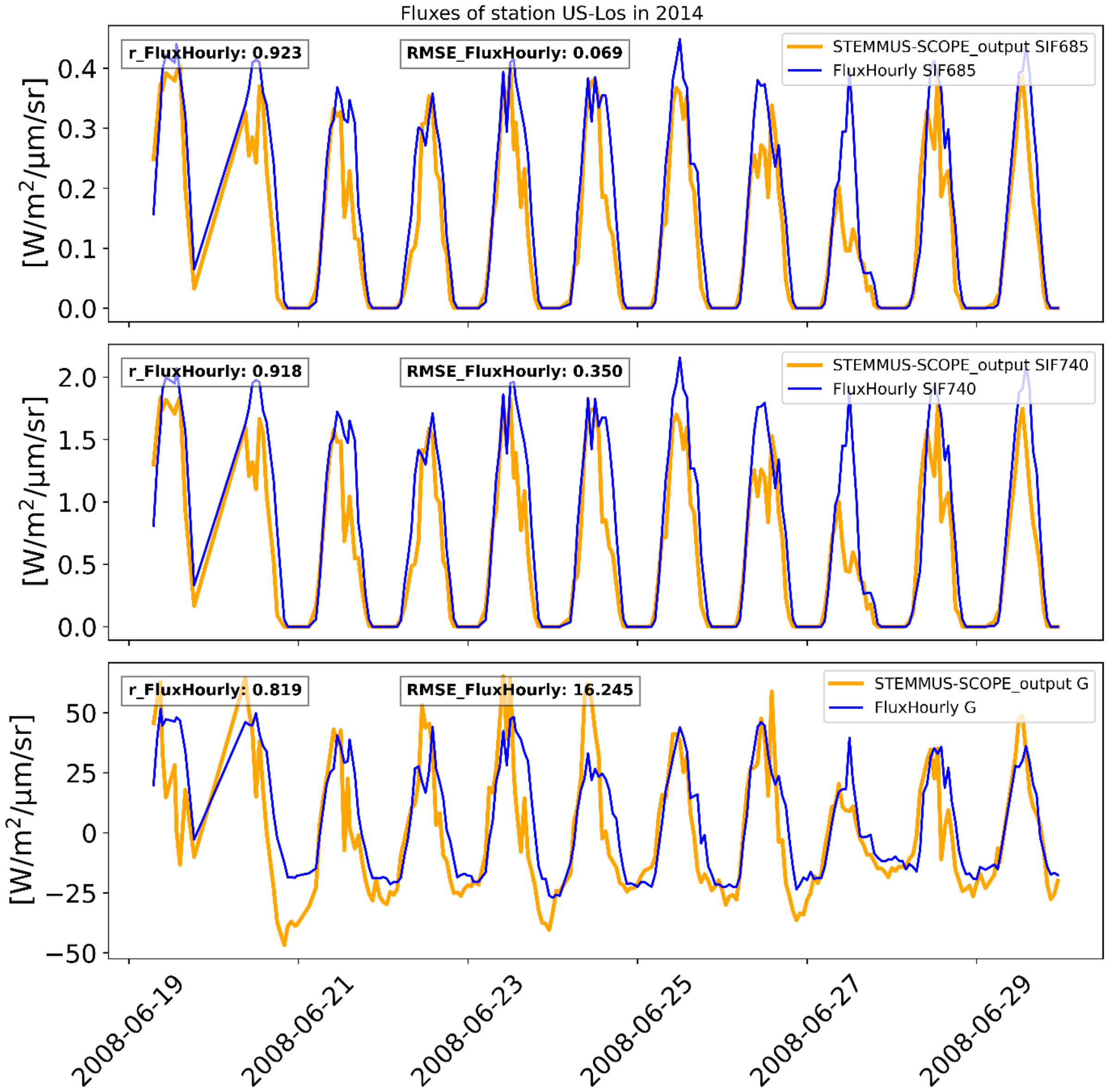
Figure S7. Rn, LE, H, G, GPP, SIF685, SIF740 in IT-Lav, June and Oct (time is in local time).



Fluxes of station US-Los in 2014
r_STEMMUS-SCOPE: 0.979
r_FluxHourly: 0.912
RMSE_STEMMUS-SCOPE: 52.403
RMSE_FluxHourly: 105.126
in-situ Rn
STEMMUS-SCOPE_output
FluxHourly
[W/m^2]
r_STEMMUS-SCOPE: 0.778
r_FluxHourly: 0.836
RMSE_STEMMUS-SCOPE: 38.079
RMSE_FluxHourly: 72.411
in-situ LE
STEMMUS-SCOPE_output
FluxHourly
[W/m^2]
r_STEMMUS-SCOPE: 0.964
r_FluxHourly: 0.838
RMSE_STEMMUS-SCOPE: 120.963
RMSE_FluxHourly: 66.592
in-situ H
STEMMUS-SCOPE_output
FluxHourly
[W/m^2]
r_STEMMUS-SCOPE: 0.928
r_FluxHourly: 0.637
RMSE_STEMMUS-SCOPE: 10.188
RMSE_FluxHourly: 9.385
in-situ GPP
STEMMUS-SCOPE_output
FluxHourly
[μmol/m^2/s]
2008-06-19
2008-06-21
2008-06-23
2008-06-25
2008-06-27
2008-06-29



Fluxes of station US-Los in 2014
r_STEMMUS-SCOPE: 0.986
r_FluxHourly: 0.887
RMSE_STEMMUS-SCOPE: 31.428
RMSE_FluxHourly: 82.501
in-situ Rn
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.912
r_FluxHourly: 0.843
RMSE_STEMMUS-SCOPE: 20.090
RMSE_FluxHourly: 49.443
in-situ LE
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.900
r_FluxHourly: 0.787
RMSE_STEMMUS-SCOPE: 76.033
RMSE_FluxHourly: 50.285
in-situ H
STEMMUS-SCOPE_output
FluxHourly
[W/m2]
r_STEMMUS-SCOPE: 0.848
r_FluxHourly: 0.806
RMSE_STEMMUS-SCOPE: 9.755
RMSE_FluxHourly: 6.380
in-situ GPP
STEMMUS-SCOPE_output
FluxHourly
[μmol/m2/s]
2008-09-05
2008-09-06
2008-09-07
2008-09-08
2008-09-09
2008-09-10
2008-09-11
2008-09-12
2008-09-13
2008-09-14


Fluxes of station US-Los in 2014
r_FluxHourly: 0.923
RMSE_FluxHourly: 0.069
STEMMUS-SCOPE_output SIF685
FluxHourly SIF685
r_FluxHourly: 0.918
RMSE_FluxHourly: 0.350
STEMMUS-SCOPE_output SIF740
FluxHourly SIF740
r_FluxHourly: 0.819
RMSE_FluxHourly: 16.245
STEMMUS-SCOPE_output G
FluxHourly G
[W/m²/µm/sr]
2008-06-19
2008-06-21
2008-06-23
2008-06-25
2008-06-27
2008-06-29

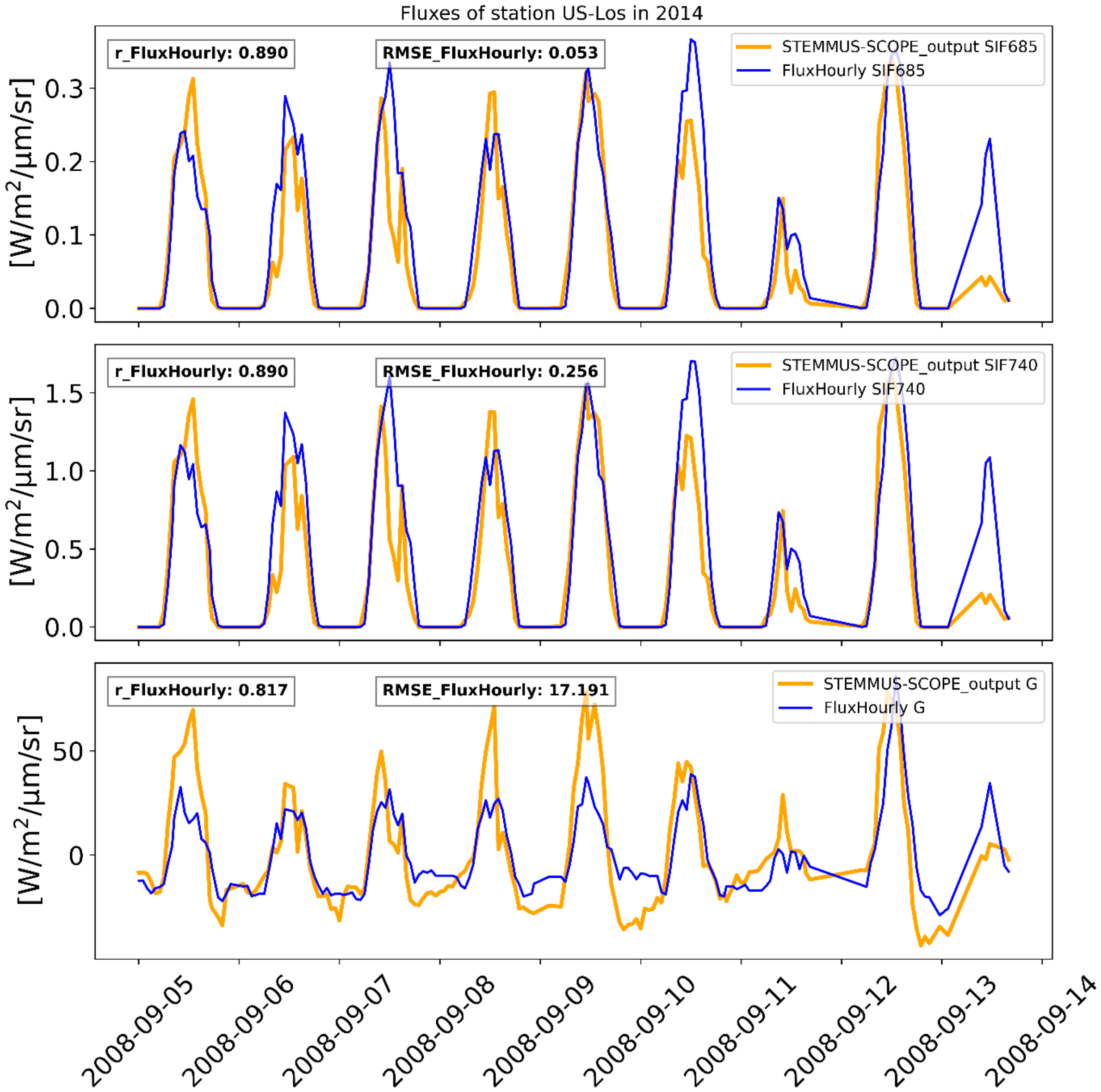
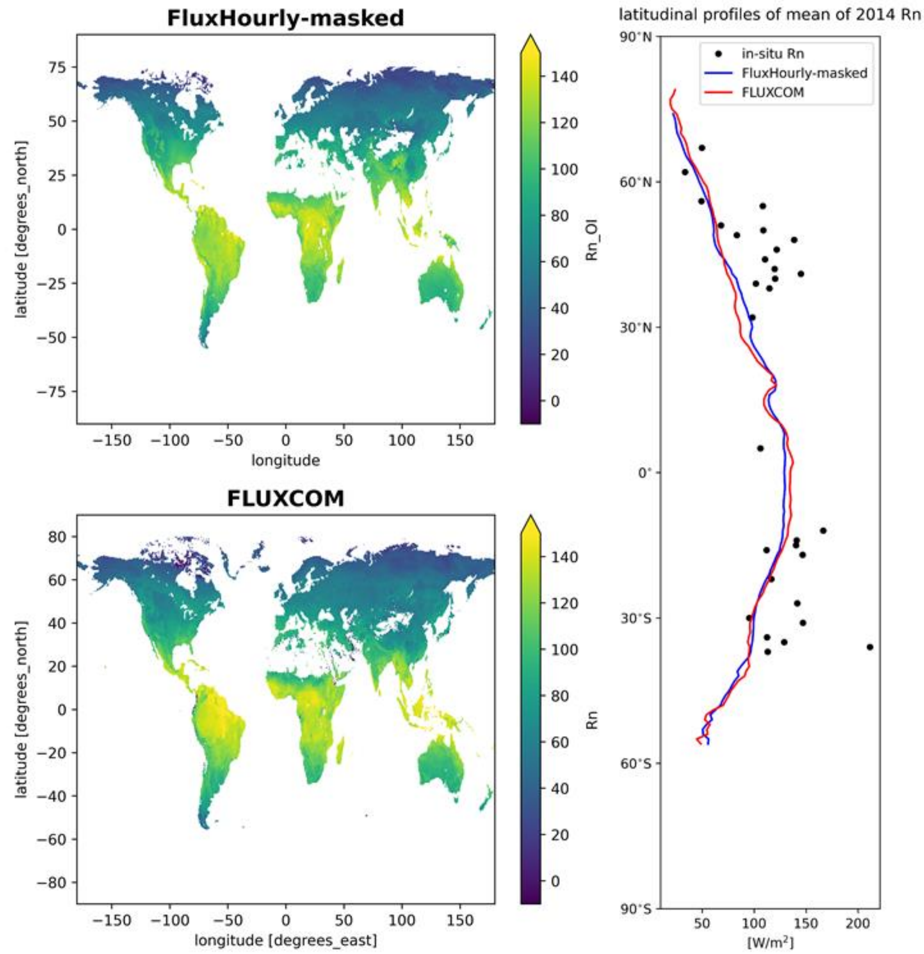
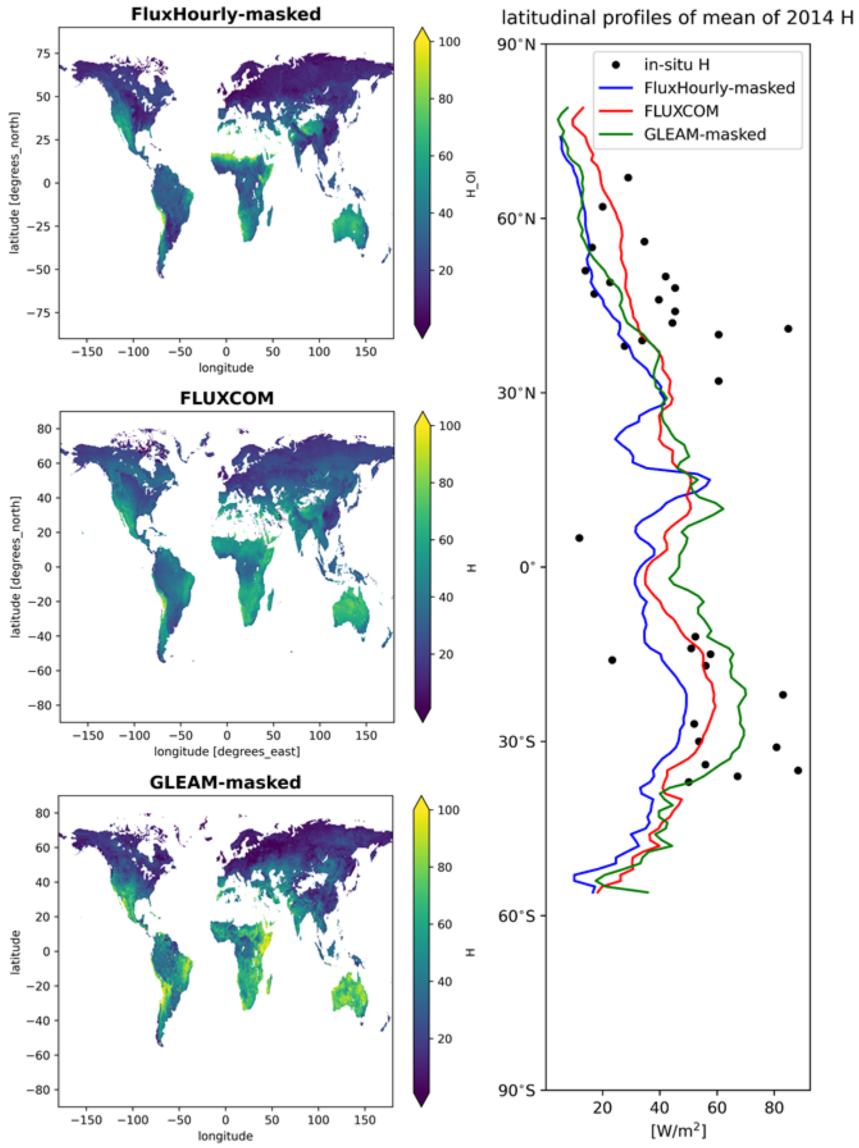


Figure S8. Rn, LE, H, G, GPP, SIF685, SIF740 in US-Los, June and Sep (time is in local time).

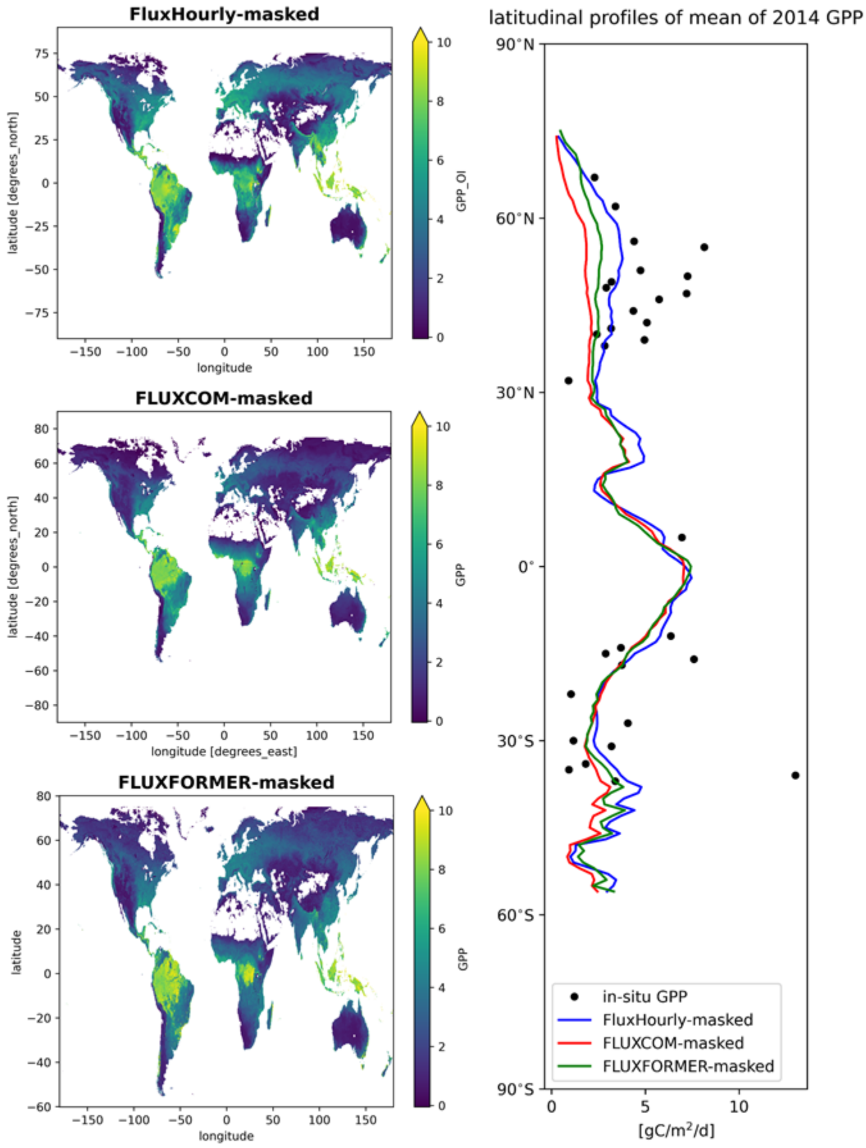
Section 3. Inter comparison with existing flux datasets on spatial pattern



FluxHourly-masked
latitude [degrees_north]
longitude
Rn_OI
FLUXCOM
latitude [degrees_north]
longitude [degrees_east]
Rn
latitudinal profiles of mean of 2014 Rn
in-situ Rn
FluxHourly-masked
FLUXCOM
[W/m2]



FluxHourly-masked
FLUXCOM
GLEAM-masked
latitude [degrees_north]
latitude
longitude
longitude [degrees_east]
H_OI
H
latitudinal profiles of mean of 2014 H
in-situ H
FluxHourly-masked
FLUXCOM
GLEAM-masked
[W/m2]


FluxHourly-masked
FLUXCOM-masked
FLUXFORMER-masked
latitude [degrees_north]
latitude
longitude
longitude [degrees_east]
GPP_OI
GPP
latitudinal profiles of mean of 2014 GPP
90°N
60°N
30°N
0°
30°S
60°S
90°S
in-situ GPP
FluxHourly-masked
FLUXCOM-masked
FLUXFORMER-masked
[gC/m²/d]

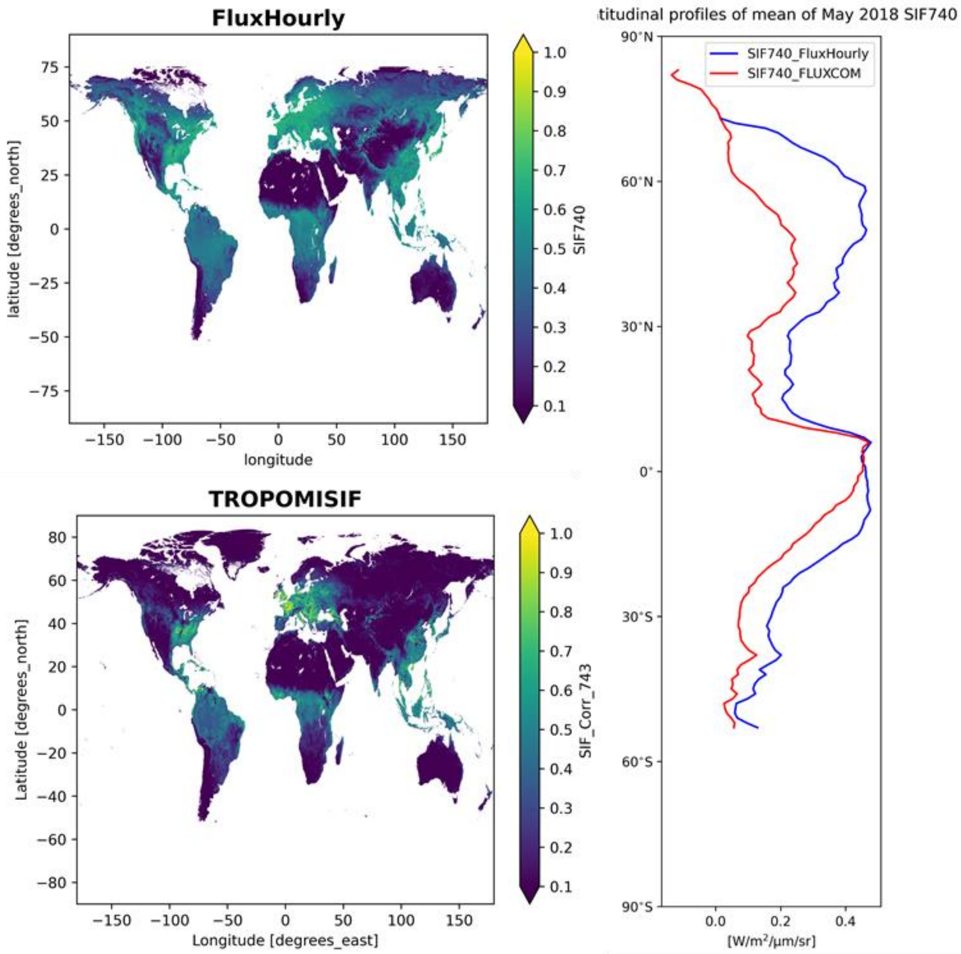


Figure S9. Annual mean of predicted hourly 9 km Rn, H, GPP, SIF740 by RF_OI, and FLUXCOM, GLEAM, FLUXFORMER (only for GPP) in 2014 and TROPOMISIF (only for SIF740 in May 2018). Note: GPP is masked by both FLUXCOM and FLUXFORMER as both datasets has missing values. Rn and H is masked by FLUXCOM.

Section 4. Diurnal patterns and seasonal changes: Rn, H, G

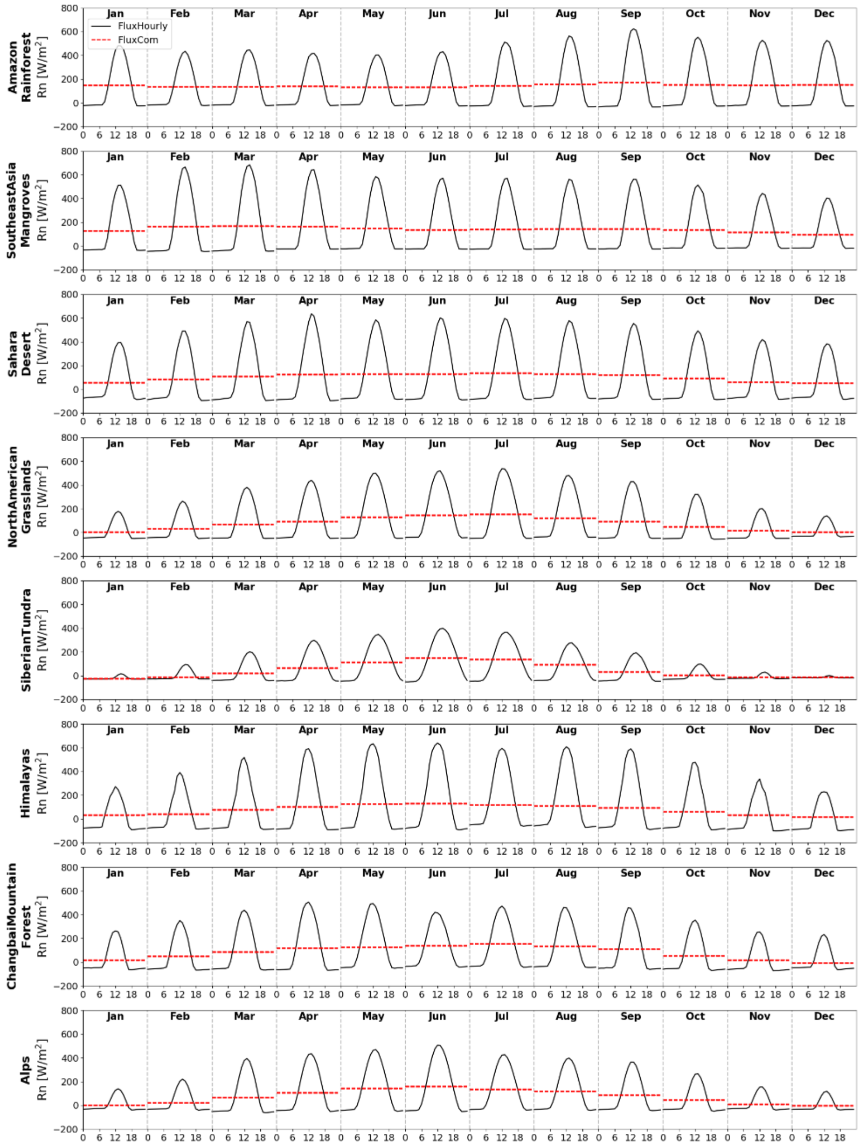


Figure S10. Diurnal cycles of Rn for 8 regions for each month of the year, where each panel refers to a region. The data is converted to local time zone.

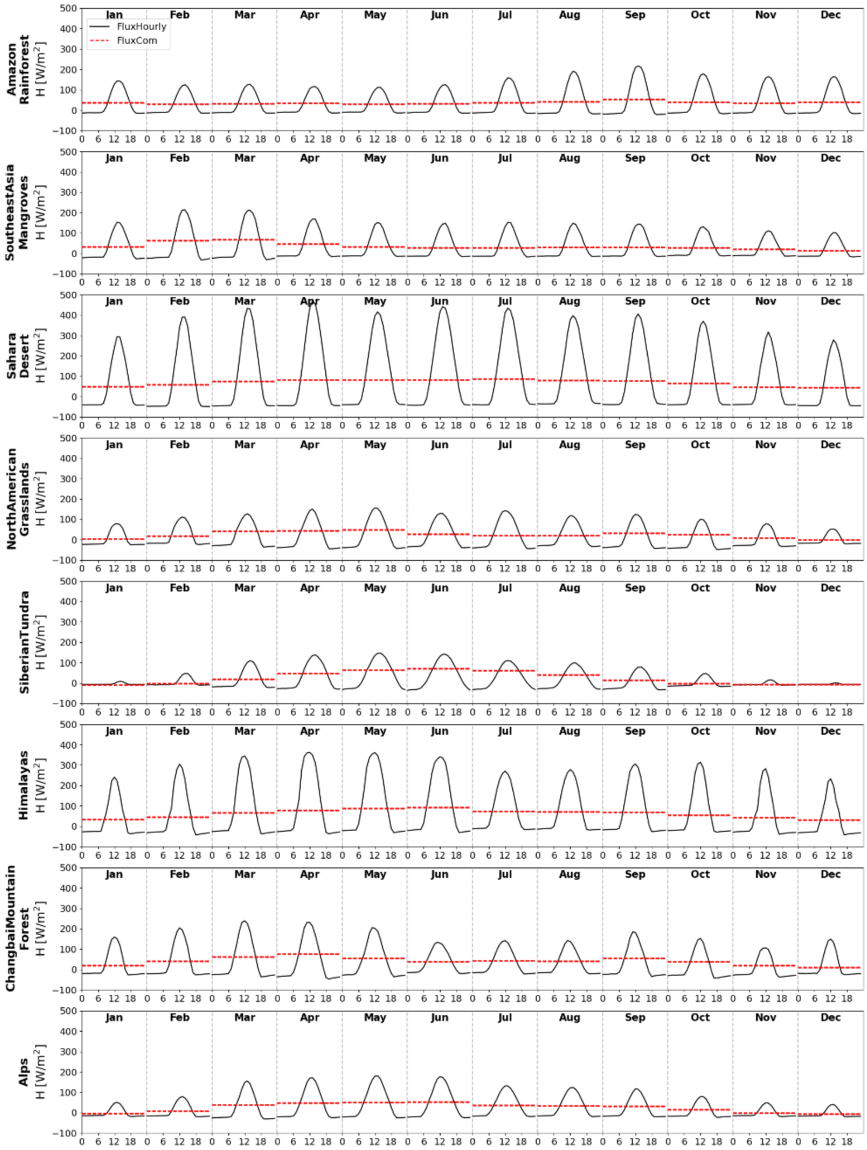


Figure S11. Diurnal cycles of H for 8 regions for each month of the year, where each panel refers to a region. The data is converted to local time zone.

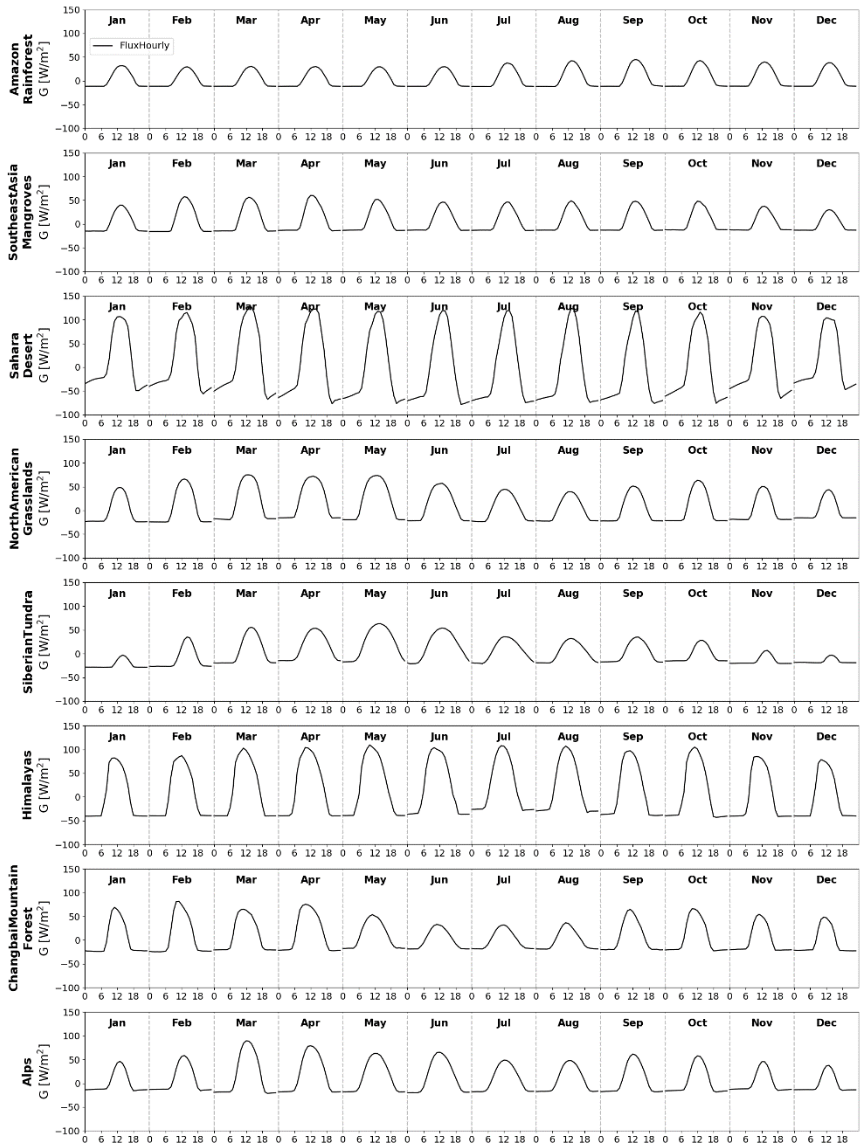


Figure S12. Diurnal cycles of G for 8 regions for each month of the year, where each panel refers to a region. The data is converted to local time zone.

Section 5. 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, 2024

Wang, Y., Zeng, Y., Alidoost, F., Schilperoort, B., Song, Z., Yu, D., Tang, E., Han, Q., Liu, Z., Peng, X., Zhang, C., Retsios, B., Girgin, S., Lü, X., Zuo, Q., Cai, H., Yu, Q., Van der Tol, C., & Su, Z. A physically consistent dataset of water-energy-carbon fluxes across the Soil-Plant-Atmosphere Continuum. Sci. Data2025