

A APPENDIX

A.1 FLUX TOWER PICTURE

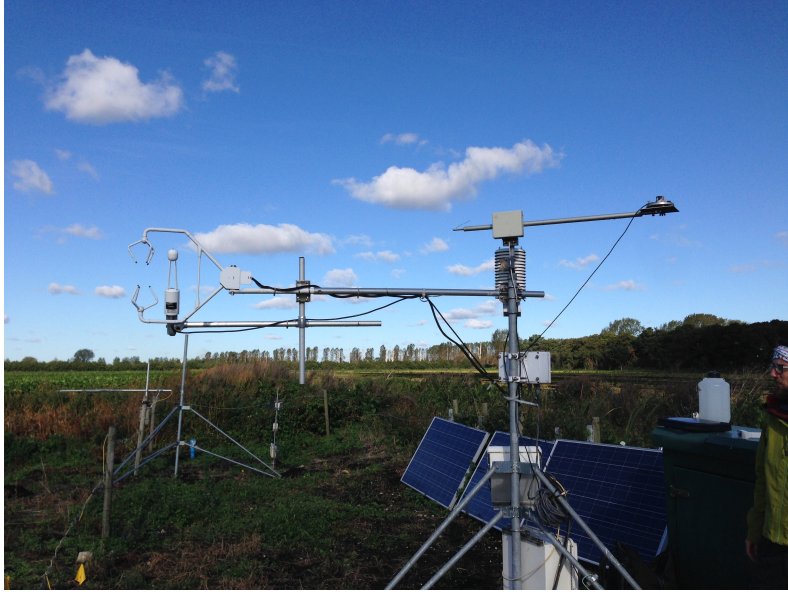


Figure 3: Flux tower in East Anglia, UK, composed of a sonic anemometer, an infrared analyser on the arm of the tower at the left of the picture

A.2 TEMPERATURE MODEL

A key component in expressing the dynamics of NEE as a function of time White & Luo (2008); Weng (2011) is the differential daily temperature model, which is described in the JULES, the Joint UK Landscape Simulator Williams & Clark (2014), as the equation:

$$T_{air,t} = T_{air,r} + \frac{\Delta T_{air,t}}{2} \cos(2\pi \frac{t - t_{T_{max}}}{t_{day}}) \quad (17)$$

where $T_{air,r}$ and $\Delta T_{air,t}$ are the temperature and diurnal temperature ranges respectively, t_{day} is the length of a day and $t_{T_{max}}$ is the time of day where the temperature is highest. We assume $t_{T_{max}}$ occurs 0.15 of the length of the day after local noon time as:

$$t_{T_{max}} = \frac{t_{up} + t_{down}}{2} + 0.15(t_{up} - t_{down}) \quad (18)$$

with t_{up} , t_{down} are the local sunrise and the sunset times.

A.3 RADIATION MODEL

Another key component in expressing the dynamics of NEE as a function of time is the differential radiation model which is described in JULES as being based on the downward longwave R_{sw}^{down} , shortwave R_{sw}^{down} and the diffuse radiation R_{diff} . The downward longwave solar radiation is defined as:

$$R_{lw,t}^{down} = R_{lw,0}^{down} (4 \frac{T_{air,t}}{T_{air,r}} - 3) \quad (19)$$

where $R_{lw,0}^{down}$ is the downward longwave solar radiation. Furthermore, the downward shortwave solar radiation is defined as:

$$R_{sw,t}^{down} = R_{sw}^{0,down} R_{norm,t}^{\odot} \quad (20)$$

where $R_{norm,t}^\odot$ is the solar radiation normalisation factor depending on the time t and $R_{sw}^{0,down}$ is the downward shortwave solar radiation. The diffuse radiation is thus defined as:

$$R_{diff,t} = R_{diff}^0 R_{norm,t}^\odot. \quad (21)$$

The constants $R_{lw,0}^{down}$, $R_{sw}^{0,down}$, R_{diff}^0 are fixed based on geo-scientific studies. The normalisation solar radiation $R_{norm,t}^\odot$ comes from Huntingford et al. (2010).

All in all, we consider the global radiation as:

$$R_{g,t} = R_{lw,t}^{down} + R_{sw,t}^{down} + R_{diff,t} \quad (22)$$

A.4 DRIFT OF THE STOCHASTIC DIFFERENTIAL EQUATION

Here are the drift details of the Wiener process. Firstly, we use the equation 17

$$\frac{dT_{air,t}}{dt} = \pi \frac{\Delta T_{air,t}}{t_{day}} \sin(2\pi \frac{t - t_{T_{max}}}{t_{day}}). \quad (23)$$

Moreover, we consider the Arrhenius-type law of ecosystem respiration in equations 2 and 3

$$\frac{d}{dT_{air}} R_{eco,t}(T_{air,t}) = \frac{E_0}{(T_{air,t} - T_0)^2} R_{eco,t} \quad (24)$$

where $r_{night/day}$ is a generic base respiration at reference temperature T_{ref} .

Then, we use the global radiation equation 22

$$\frac{dR_{g,t}}{dt} = \frac{d}{dt} (R_{lw,t}^{down} + R_{sw,t}^{down} + R_{diff,t}) \quad (25)$$

Where each component of the sum is computed from the equations 19, 20 and 21, we deduce

$$\begin{aligned} \frac{d}{dt} R_{lw,t}^{down} &= \frac{R_{lw,0}^{down}}{T_{air,0}} \frac{dT_{air,t}}{dt} \\ \frac{d}{dt} R_{sw}^{down} &= R_{sw}^{0,down} \frac{d}{dt} R_{norm,t}^\odot \\ \frac{d}{dt} R_{diff,t} &= R_{diff}^0 \frac{d}{dt} R_{norm,t}^\odot \end{aligned} \quad (26)$$

where $\frac{d}{dt} R_{norm,t}^\odot$ is deduced from the IMOGEN routine `sunny` from JULES Huntingford et al. (2010).

Moreover, we deduce from the daytime equation model 3

$$\frac{d}{dR_g} GPP_t(R_{g,t}) = \frac{\alpha\beta^2}{(\alpha R_{g,t} + \beta)^2} \quad (27)$$

All things considered, we have the analytic expression for

$$\mu_t = \frac{dNEE_t}{dt} = \frac{d}{dT_{air}} R_{eco,t}(T_{air,t}) \frac{dT_{air,t}}{dt} - \frac{d}{dR_g} GPP_t(R_{g,t}) \frac{dR_{g,t}}{dt} \quad (28)$$

A.5 DIFFUSION COEFFICIENT, NOISE OF MEASUREMENTS

We consider the diffusion coefficient for both nighttime and daytime $\sigma_{night}, \sigma_{day}$ that can be computed from the histograms of the error between the NEE values in flux data and the day/night Physics Models defined in section 3.1. We confirmed the Gaussian assumption using normality tests White & Luo (2008). These coefficients will be learned implicitly by the PIAE.

A.6 PIAE AND AE NEURAL NETWORK CONFIGURATIONS

A.7 DAYTIME MODEL RESULTS

A.8 PERFORMANCE ON PREDICTIONS ON NIGHTTIME AND DAYTIME NEE PARAMETERS BOTH PIAE AND AE

```

702
703
704
705 PIAE_SDE(
706     (encoder): Sequential(
707         (0): Linear(in_features=18, out_features=16, bias=True)
708         (1): ReLU()
709         (2): Linear(in_features=16, out_features=16, bias=True)
710         (3): ReLU()
711         (4): Linear(in_features=16, out_features=32, bias=True)
712     )
713
714     (nee_decoder): Sequential(
715         (0): Linear(in_features=32, out_features=16, bias=True)
716         (1): ReLU()
717         (2): Linear(in_features=16, out_features=16, bias=True)
718         (3): ReLU()
719         (4): Linear(in_features=16, out_features=1, bias=True)
720     )
721
722     (fc_mu): Sequential(
723         (0): Linear(in_features=32, out_features=4, bias=True)
724         (1): ReLU()
725         (2): Linear(in_features=4, out_features=1, bias=True)
726     )
727
728     (fc_logvar): Sequential(
729         (0): Linear(in_features=32, out_features=4, bias=True)
730         (1): ReLU()
731         (2): Linear(in_features=4, out_features=1, bias=True)
732     )
733
734     (temp_derivative_decoder): Sequential(
735         (0): Linear(in_features=32, out_features=16, bias=True)
736         (1): ReLU()
737         (2): Linear(in_features=16, out_features=16, bias=True)
738         (3): ReLU()
739         (4): Linear(in_features=16, out_features=1, bias=True)
740     )
741
742     (k_decoder): Sequential(
743         (0): Linear(in_features=32, out_features=16, bias=True)
744         (1): LeakyReLU(negative_slope=0.01)
745         (2): Linear(in_features=16, out_features=16, bias=True)
746         (3): LeakyReLU(negative_slope=0.01)
747         (4): Linear(in_features=16, out_features=2, bias=True)
748         (5): LeakyReLU(negative_slope=0.01)
749     )
750 )
751 )
752

```

Figure 4: Neural Network Configuration for PIAE Model

```
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770 AE(
771     (encoder): Sequential(
772         (0): Linear(in_features=18, out_features=16, bias=True)
773         (1): ReLU()
774         (2): Linear(in_features=16, out_features=32, bias=True)
775         (3): ReLU()
776     )
777     (nee_decoder): Sequential(
778         (0): Linear(in_features=32, out_features=16, bias=True)
779         (1): ReLU()
780         (2): Linear(in_features=16, out_features=16, bias=True)
781         (3): ReLU()
782         (4): Linear(in_features=16, out_features=1, bias=True)
783     )
784     (k_decoder): Sequential(
785         (0): Linear(in_features=32, out_features=16, bias=True)
786         (1): ReLU()
787         (2): Linear(in_features=16, out_features=16, bias=True)
788         (3): ReLU()
789         (4): Linear(in_features=16, out_features=2, bias=True)
790     )
791 )
792
793
794
795
```

Figure 5: Neural Network Configuration for AE Model

Table 3: Tree Configuration using Python Sci-kit Learn Library for the RF and XGB models

	Tree Model	Number of Trees	Max Depth	Max Leaves
RF	100	None	None	None
XGB	100	None	None	None

Table 4: Results for NEE prediction on Day time data and model experiments. The metrics MMD, Wasstn (Wassstein Distance), KL (Kullback Leibler Divergence), and MAE (Mean Absolute Error) are expressed as the lower the better. R2 (score) is expressed as higher the better.

Approach	MMD	Wasstn	KL	MAE	R2
PIAE	0.027 ± 0.004	0.135 ± 0.01	0.042 ± 0.003	1.521 ± 0.09	0.868 ± 0.03
AE	0.026 ± 0.003	0.190 ± 0.02	0.071 ± 0.005	1.452 ± 0.08	0.877 ± 0.03
RF	0.050 ± 0.005	0.350 ± 0.03	0.149 ± 0.010	1.593 ± 0.10	0.857 ± 0.04
XGB	0.033 ± 0.004	0.234 ± 0.02	0.976 ± 0.020	1.532 ± 0.09	0.863 ± 0.03

Table 5: Performance on predictions of nighttime model parameters for both PIAE and AE models

	MMD	Wasstn	KL	MAE	R2
E_0 (PIAE)	0.172	3.388	1.237	5.629	0.941
E_0 (AE)	0.151	1.519	0.835	3.075	0.957
$r_{night/day}$ (PIAE)	0.124	0.304	0.99	0.437	0.942
$r_{night/day}$ (AE)	0.033	0.122	0.731	0.253	0.971

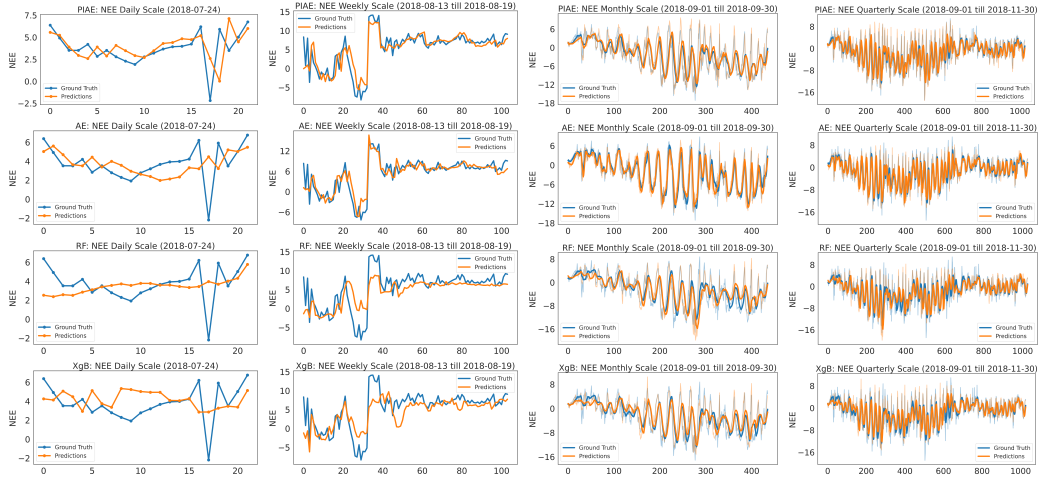


Figure 6: Day time model results on test data across different time scales for each approach in the experimentation. Row 1 represents results based on the PIAE model, Row 2 represents results from the AE model, Row 3 represents results from the RF model, and Row 4 represents results from the XGB model respectively. The sequences illustrated in the graphs are randomly sampled from the test dataset and are kept consistent for each approach for fair validation. The actual timestamps of the sequences are mentioned at the top of each graph.

Table 6: Performance on predictions of daytime model parameters for both PIAE and AE models

	MMD	Wasstn	KL	MAE	R2
E_0 (PIAE)	0.155	1.44	0.590	3.407	0.97
E_0 (AE)	0.118	1.145	0.465	2.236	0.982
$r_{night/day}$ (PIAE)	0.018	0.047	0.071	0.22	0.978
$r_{night/day}$ (AE)	0.037	0.091	0.067	0.247	0.977
α (PIAE)	0.0007	0.0079	1.139	0.0163	0.859
α (AE)	0.0196	0.0251	0.959	0.048	-1.117
β (PIAE)	0.285	3.502	0.594	6.182	0.984
β (AE)	0.239	2.212	0.348	3.387	0.997