Gap-filling soil respiration data

-a highlight of time-series analysis methods

NIBIO
NORWEGIAN INSTITUTE OF BIOECONOMY RESEARCH

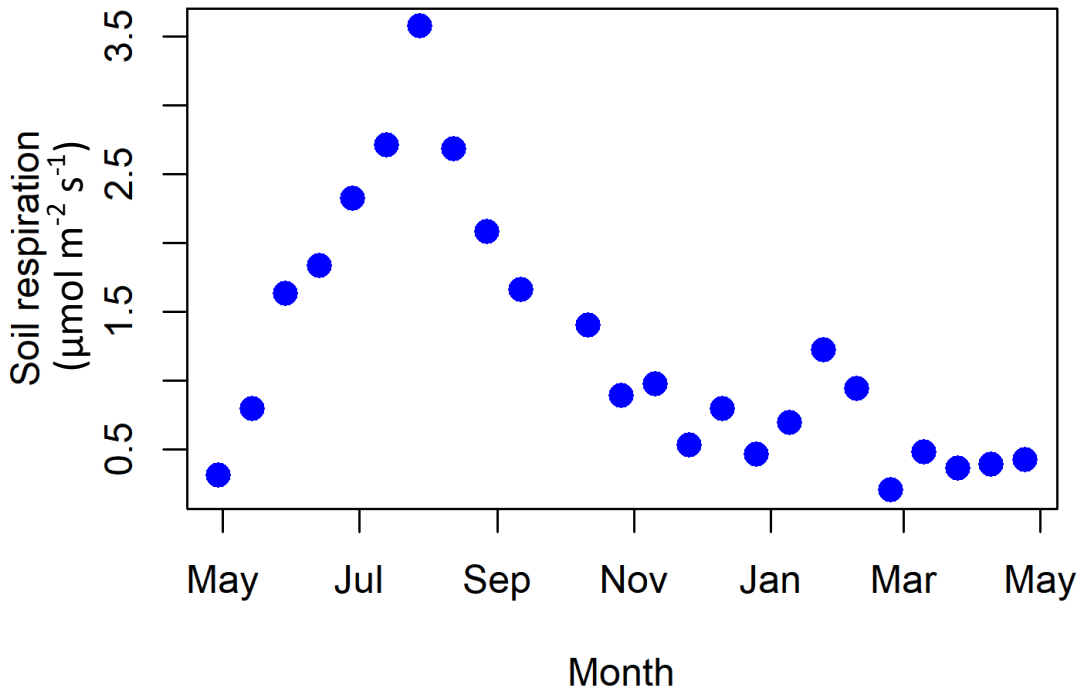
Junbin Zhao, Holger Lange, Helge Meissner

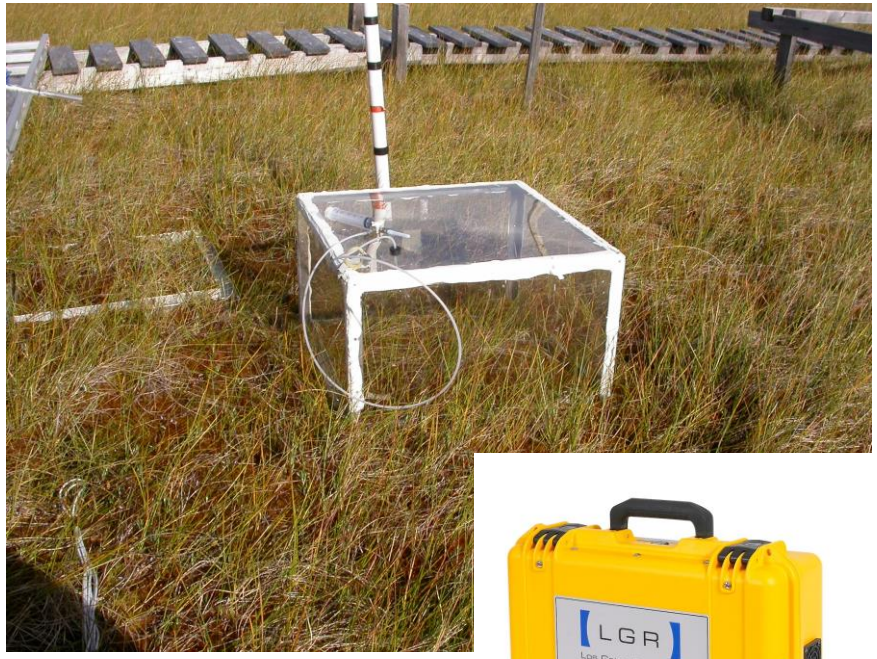
Thanks to: Manon Sueur (ENSAIA, Nancy, France)

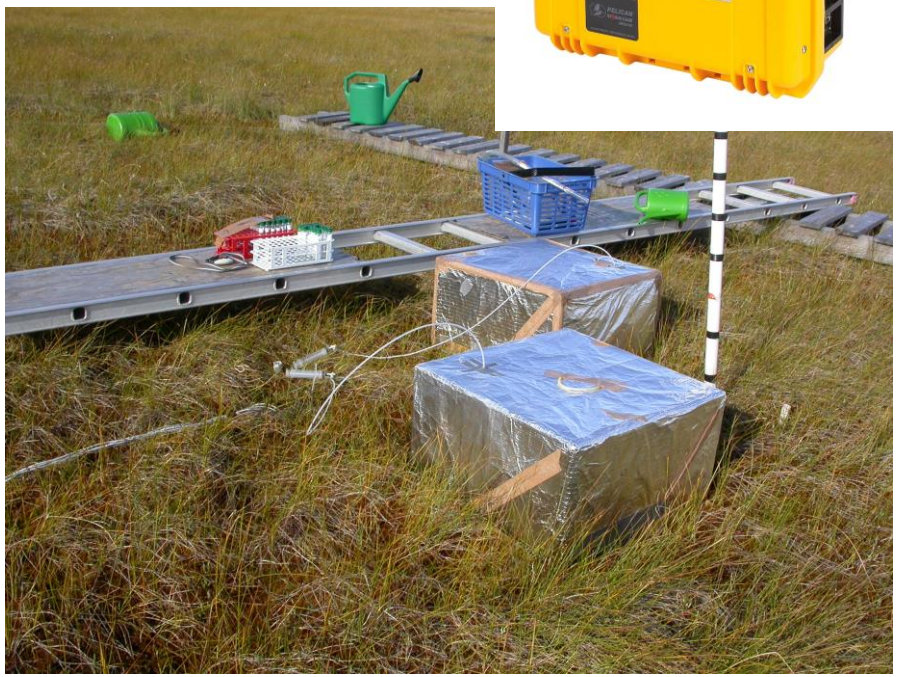
Norwegian Research Council grant no. 255061

Published as Zhao et al. 2020, Agricultural and Forest Meteorology
The methods for gap-filling are included in a R package available from Github.

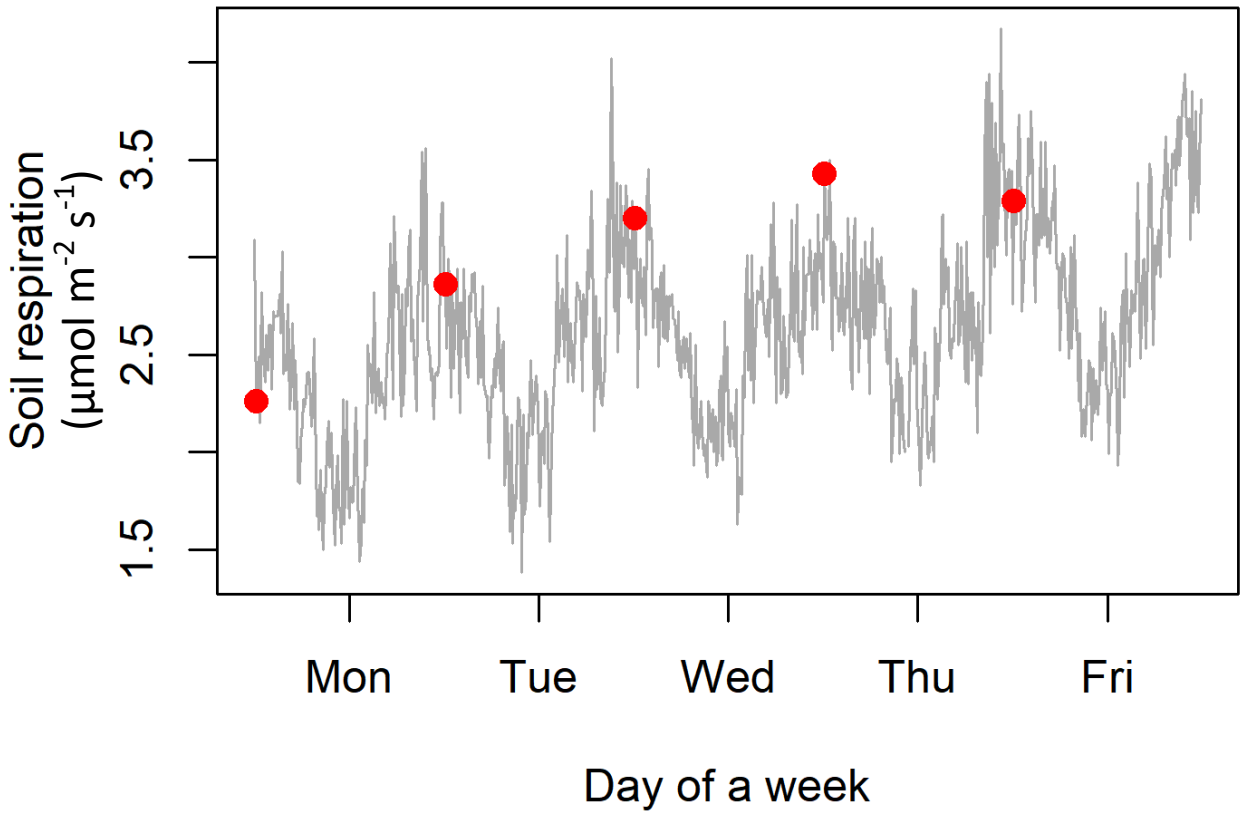
Closed chambers are widely used for soil respiration measurements; however, manually measured fluxes have **low-temporal resolution**







Compared to the continuous measurements,
low-resolution flux could lead to biased budgets for soil respiration.

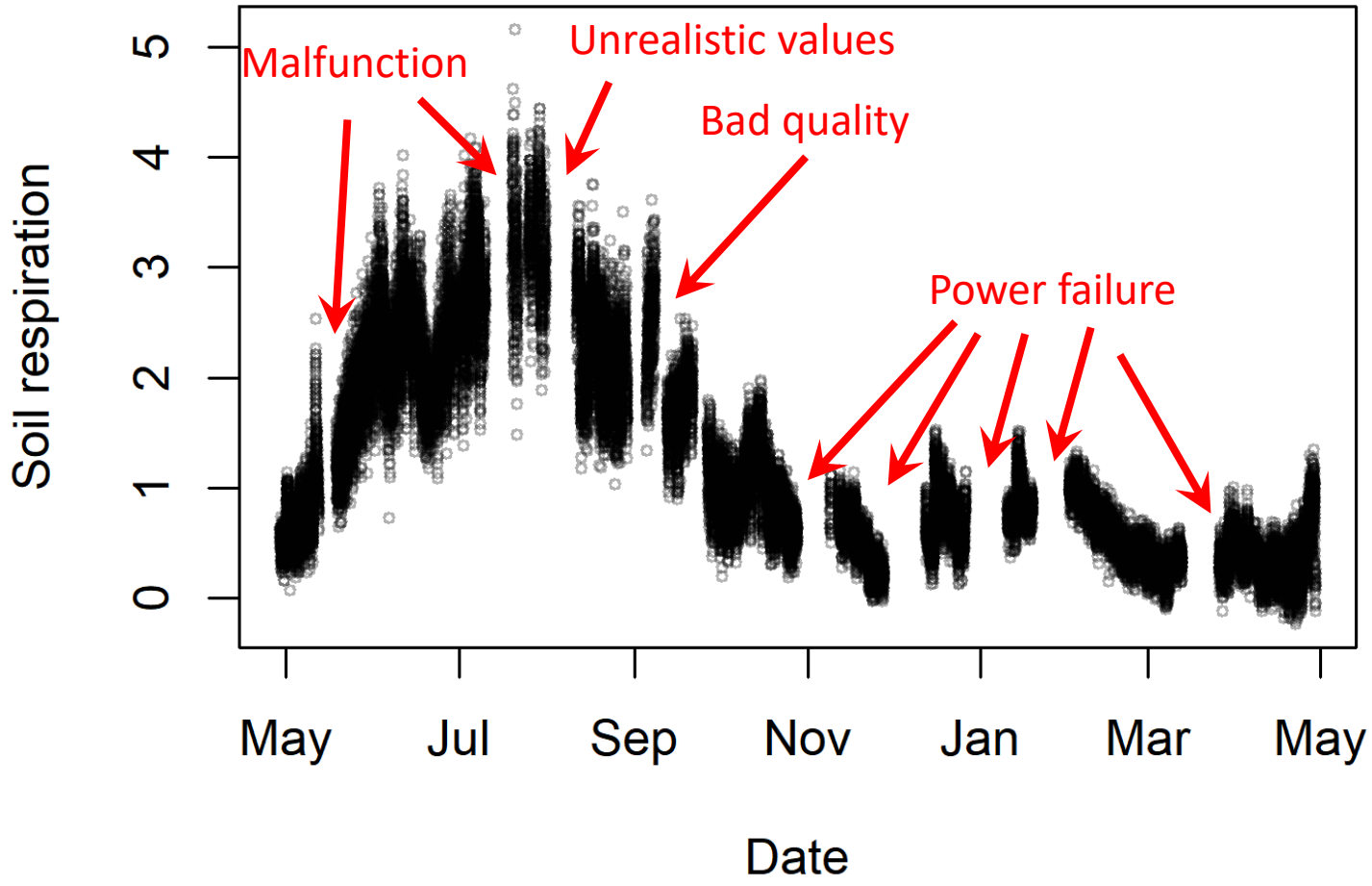


Forced diffusion chamber (Eosense)

Photo: Helge Meissner

Forced diffusion chambers measure soil respiration continuously (e.g. every 10 mins)

The continuous measurements are still associated with many data gaps that need to be filled with reasonable values.



Filling gaps in respiration time series: four methods

1. Process-oriented: non-linear regression to explanatory variables
2. Machine learning: artificial neural networks
3. Data-oriented: Singular spectrum analysis
4. Utilizing multiple measurements: Imputation by expectation-maximization

The four methods are included in a R package available from Github.

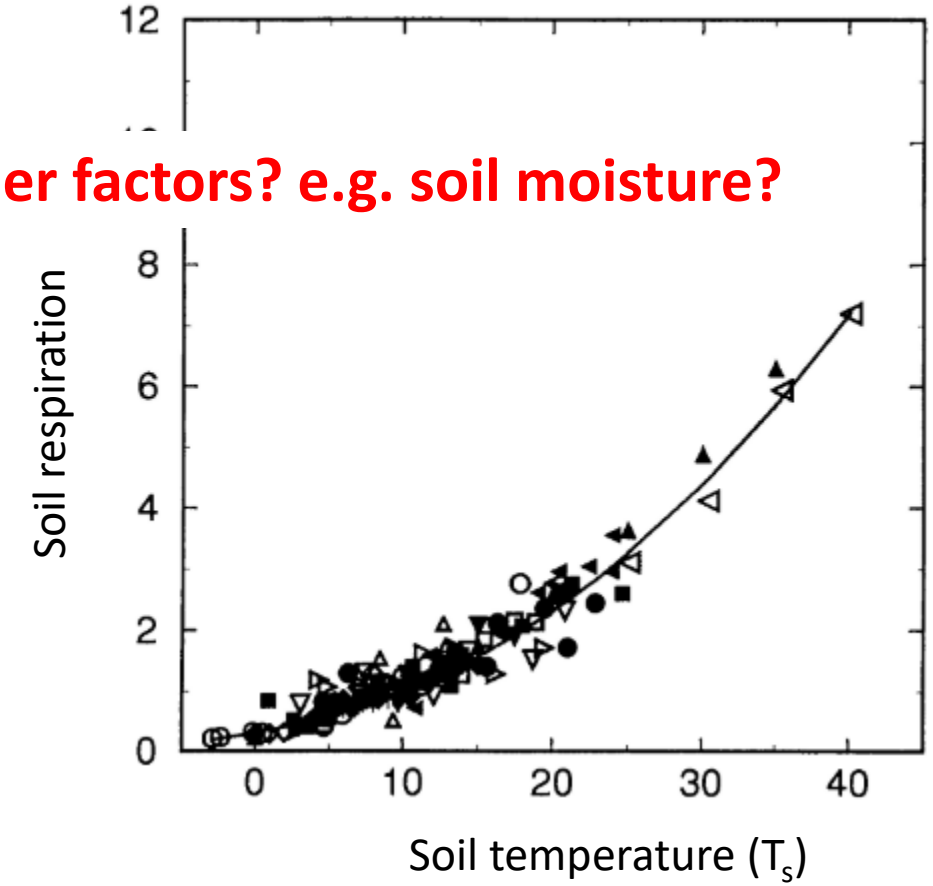
Method #1: Non-linear least squares (NLS)

- Most commonly used method!

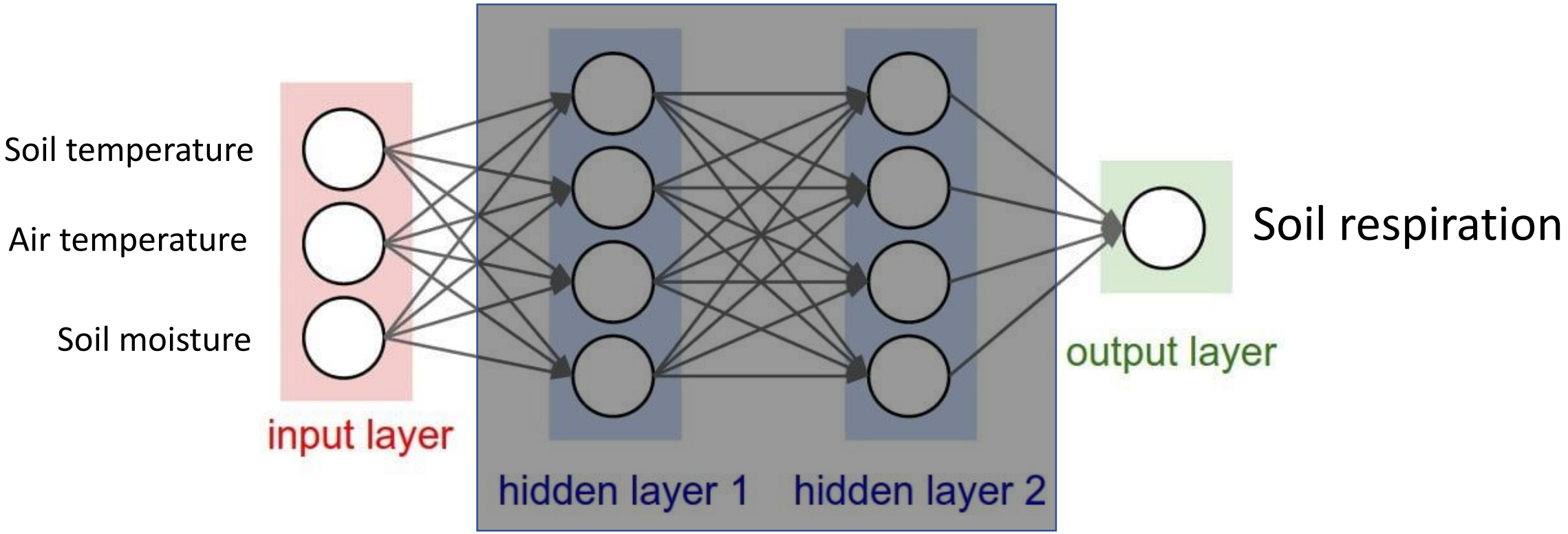
$$R_s = R_{10} \cdot \mathrm{e}^{E_0\left(\frac{1}{T_{ref}-T_0}-\frac{1}{T_s-T_0}\right)}$$

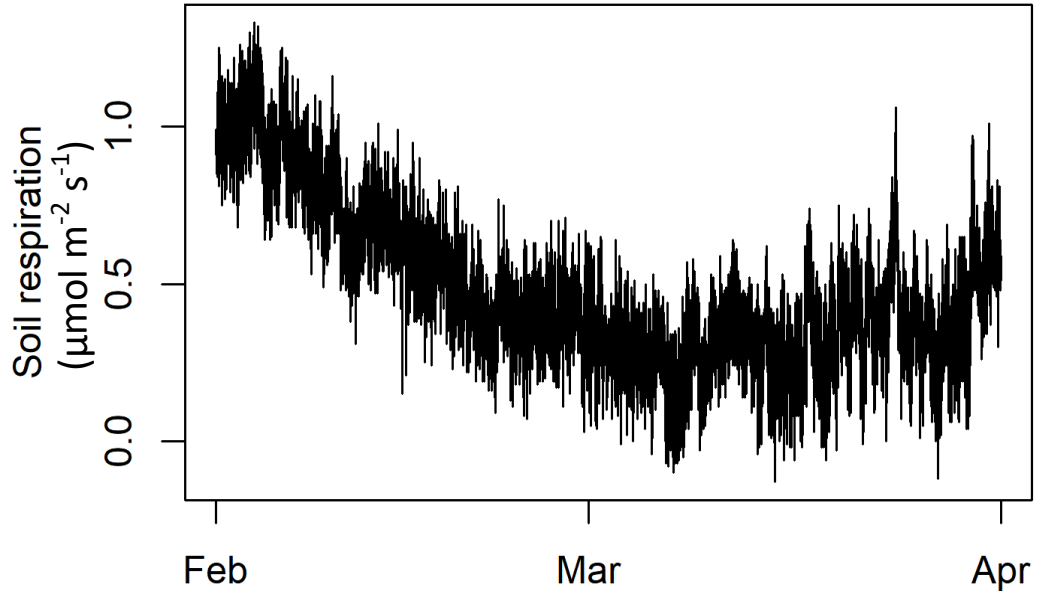
(Lloyd J., Taylor, 1994)

Other factors? e.g. soil moisture?

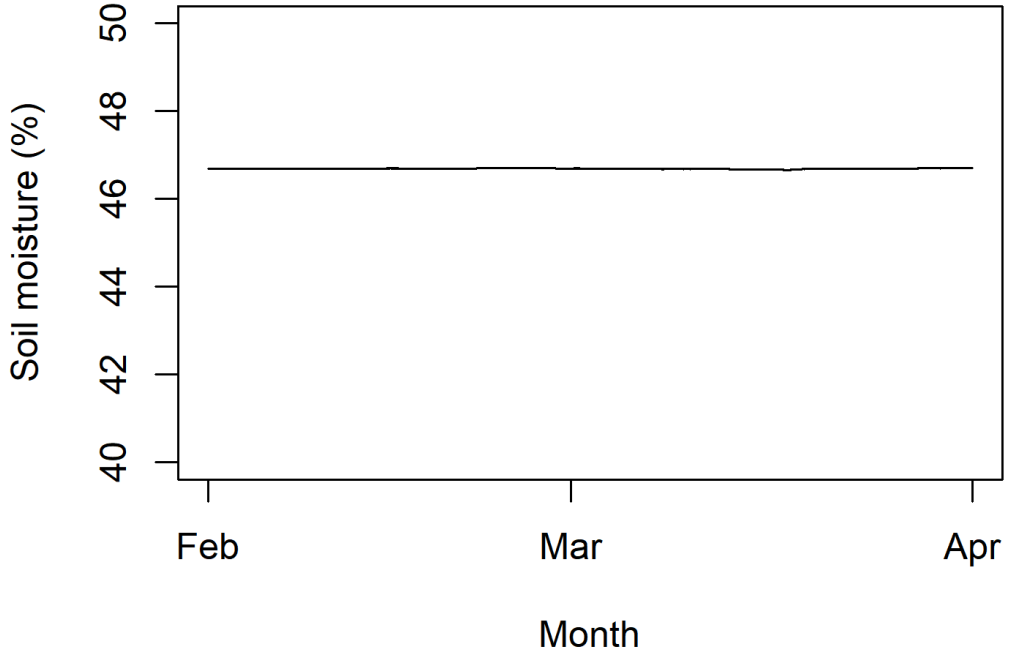


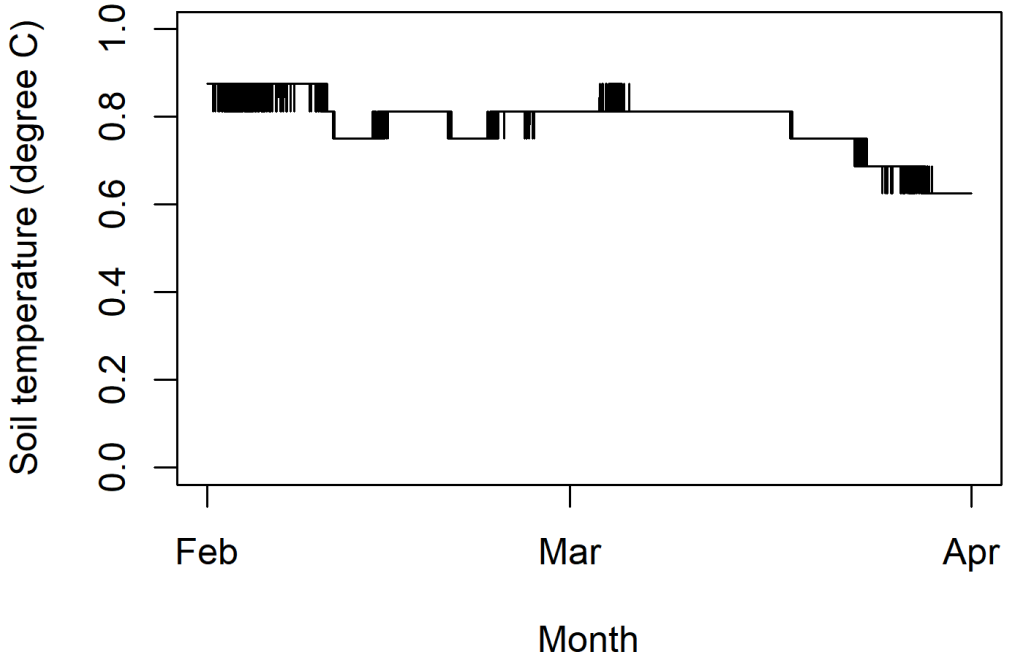
Methods #2: Artificial neural network (ANN)



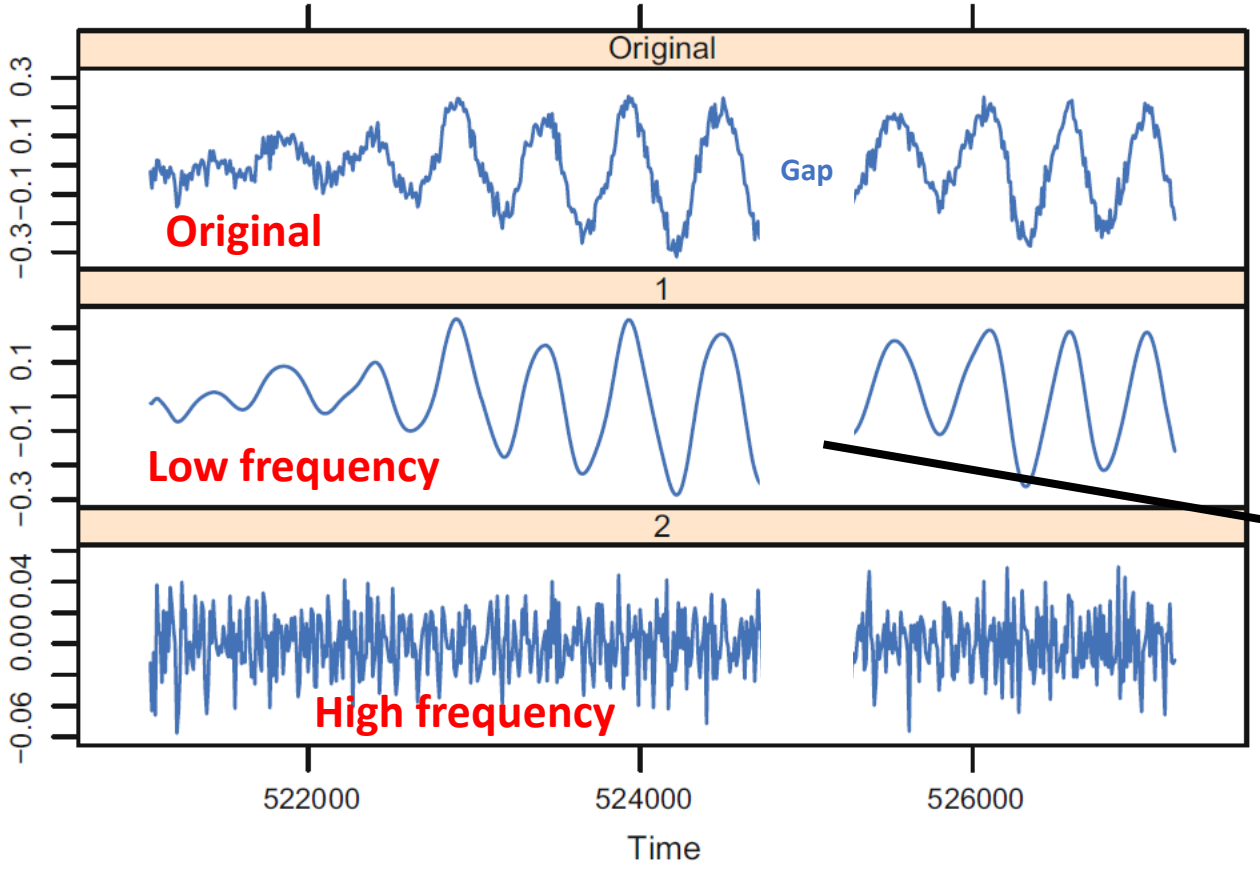


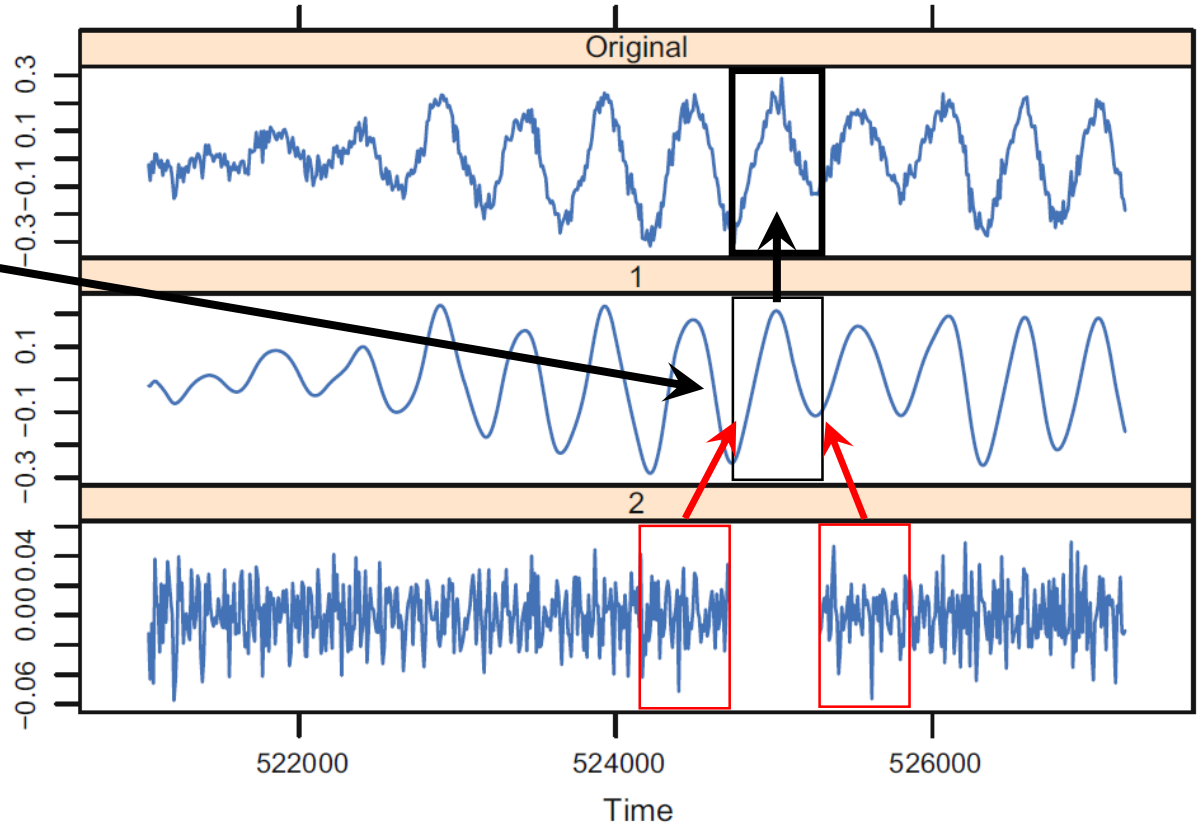
Winter flux:
bad relationships between environmental factors and soil respiration





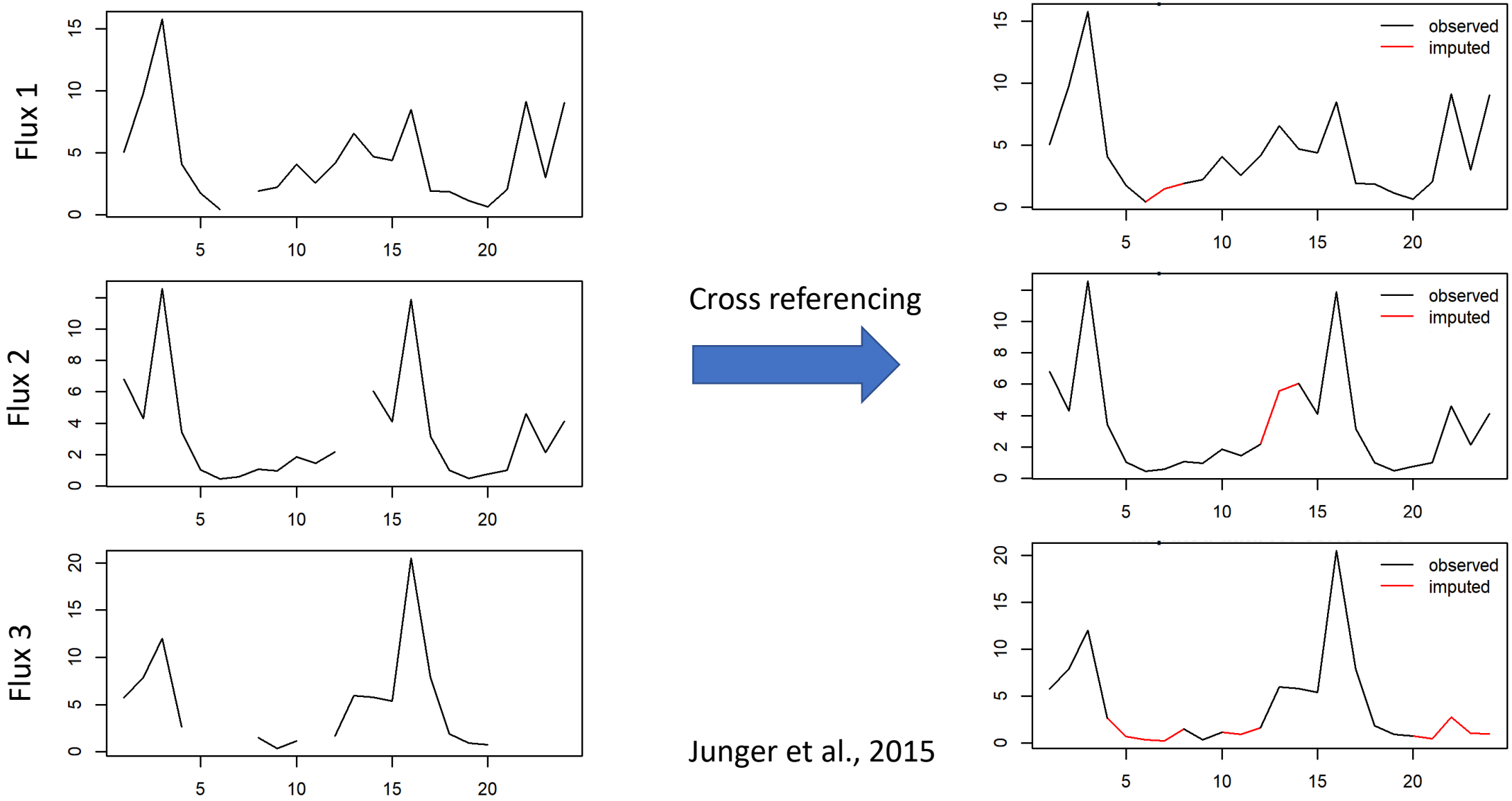
Methods #3: singular spectrum analysis (SSA)





(Graphs from Golyandina et al. 2018)

Methods #4: Expectation-maximization (EM)



Study sites in Norway

Løten

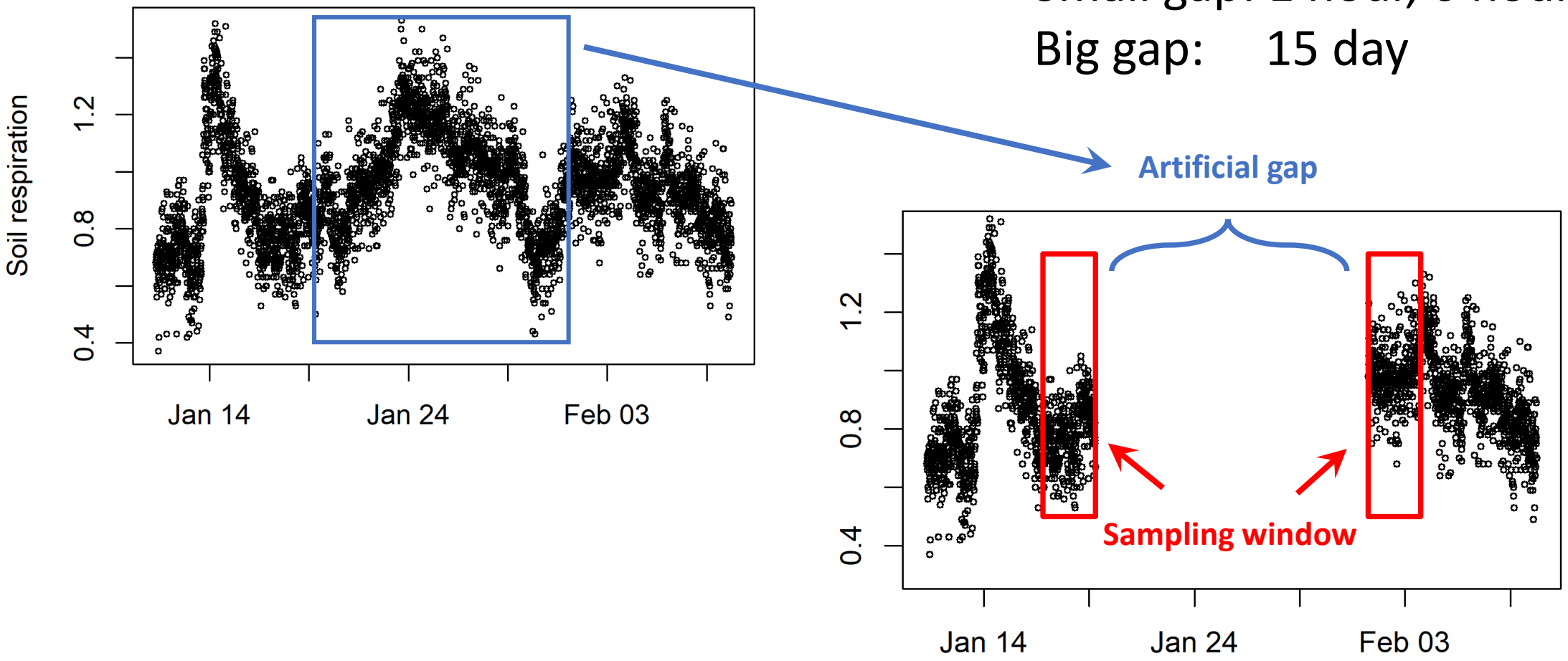
Hurdal

Photos: Helge Meissner

Add artificial gaps to be filled

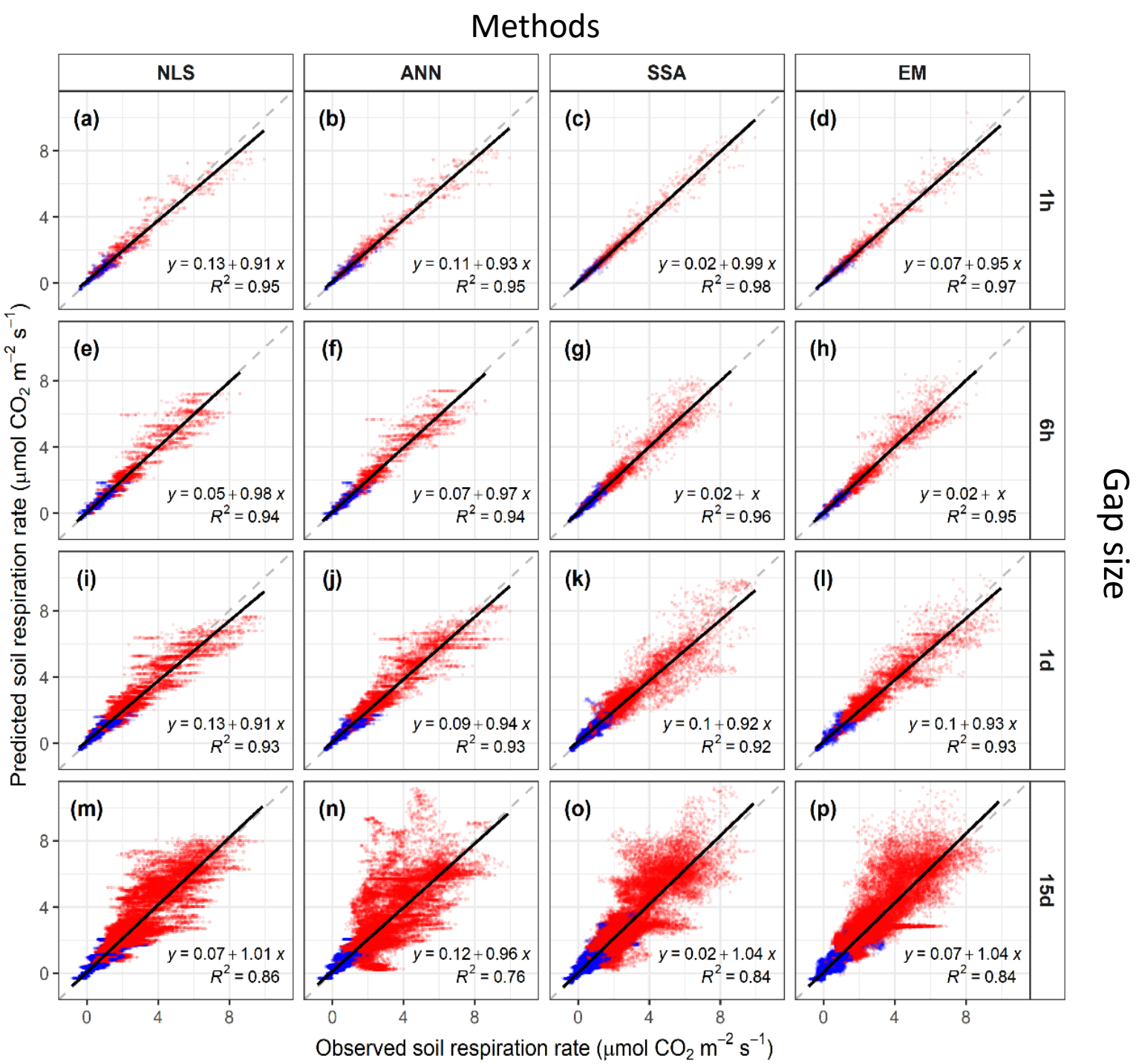
Small gap: 1 hour, 6 hour, 1 day

Big gap: 15 day

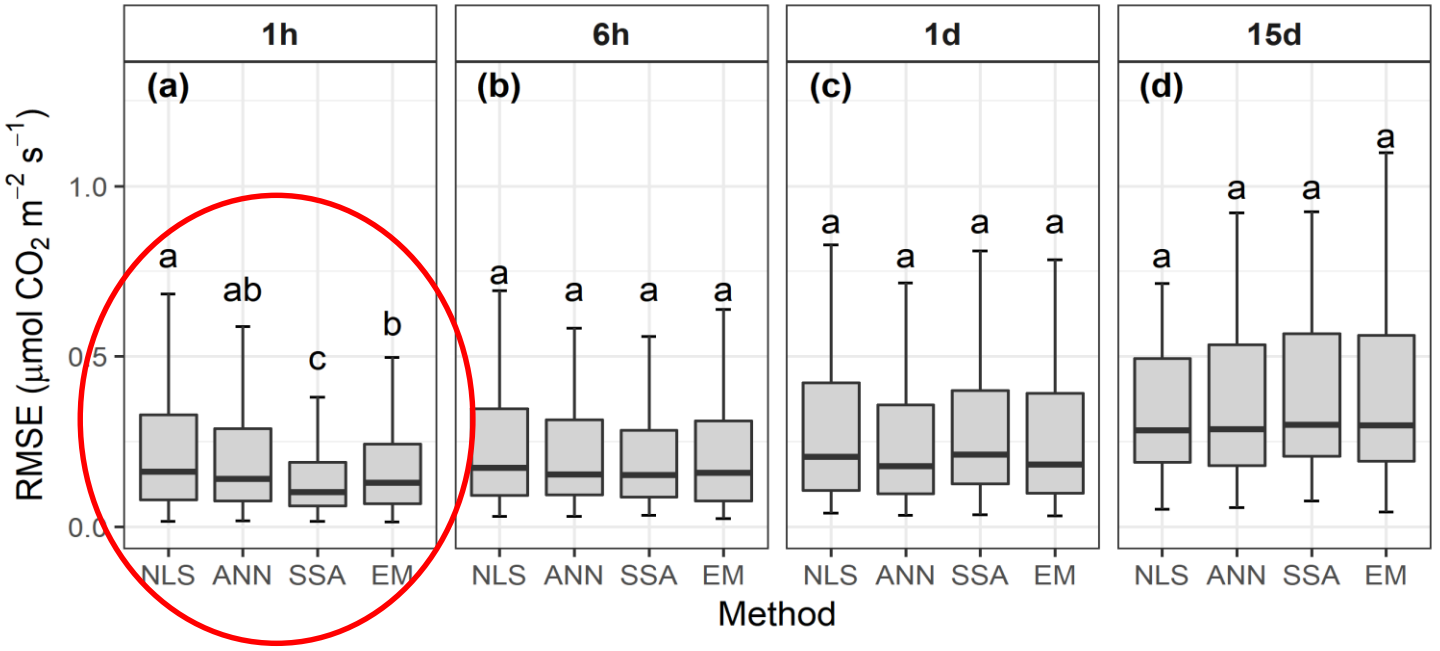


- All the methods perform well in gaps ≤ 1 day with slope close to 1 and $R^2 > 0.9$.

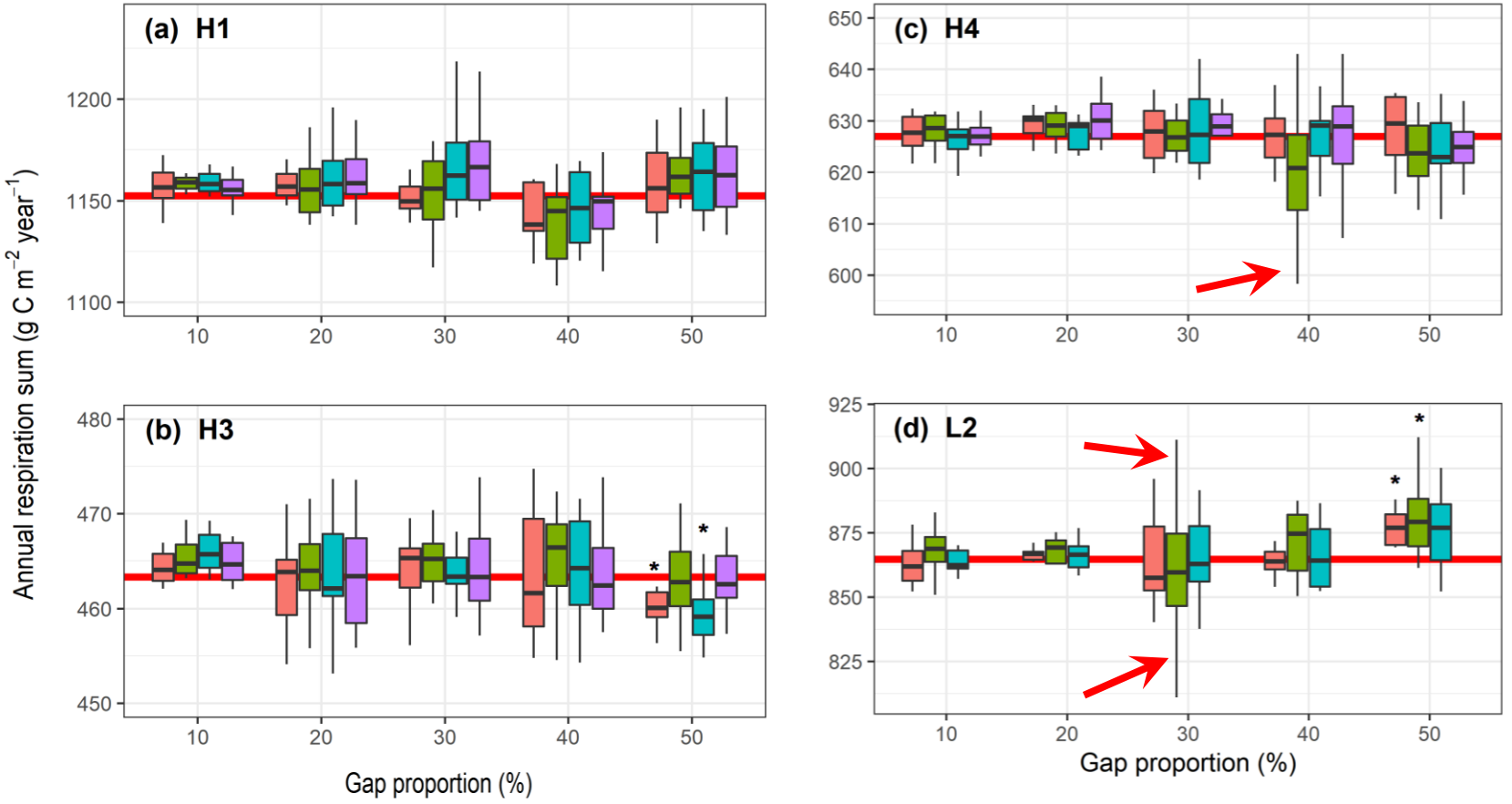
- Predictions for the 15-day gaps are not systematically biased, but are **low in R^2**, especially **ANN**.

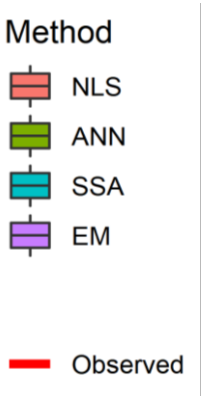


Data predicted by all the methods showed similar root mean square error (RMSE), **except** in 1 hour gaps where **SSA** and **EM** had better performance.



Annual Budget





Hurdal site: H1, H3, H4

Løten site: L2

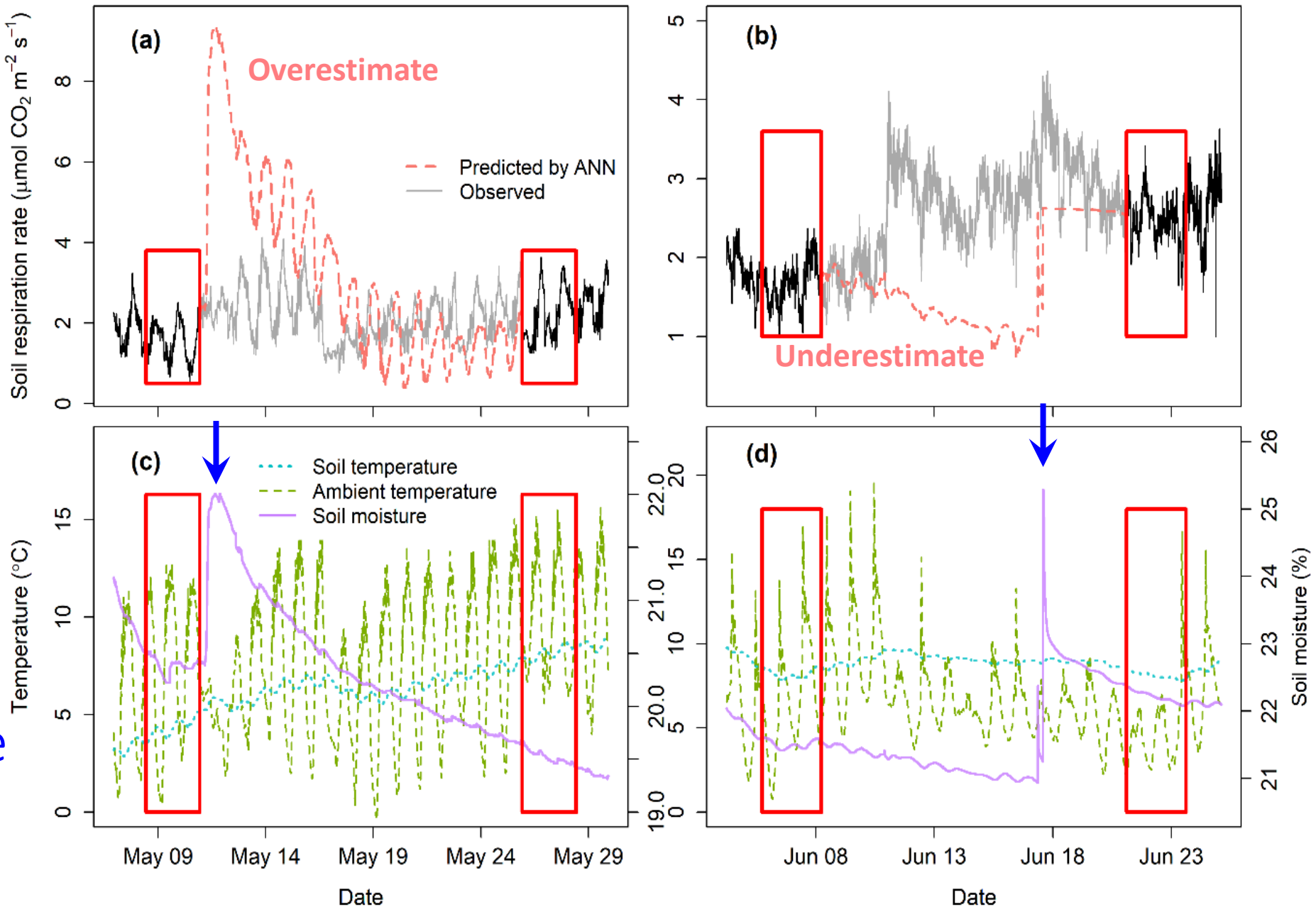
Why is ANN worst?

Error:

ANN: -11.3 to 16.0%

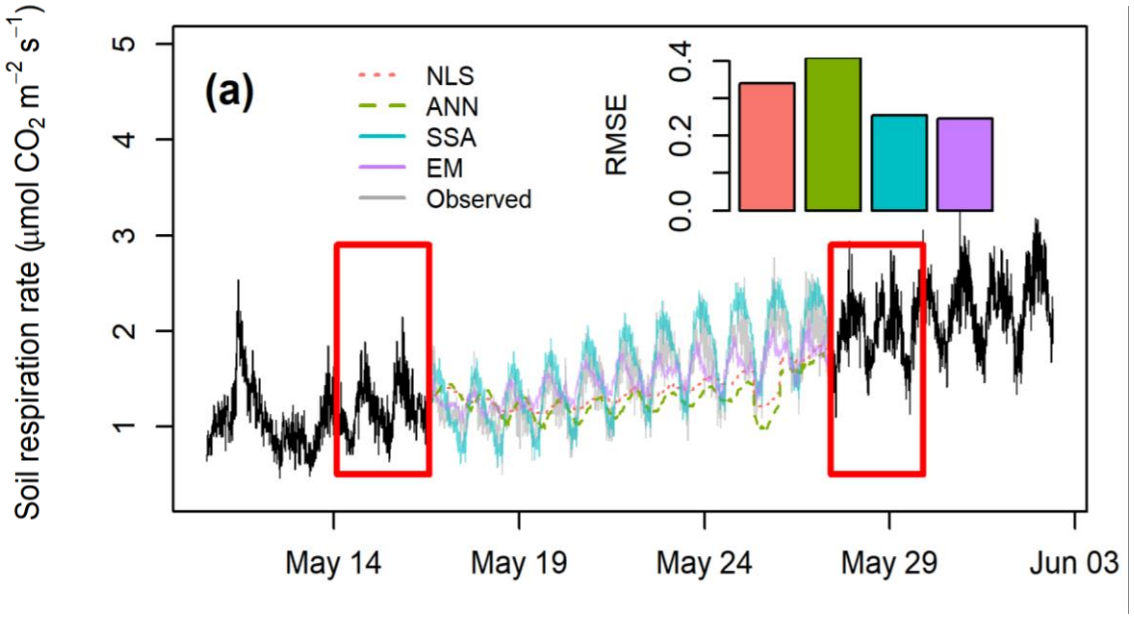
Others: -3.7 to 5.8%

Rain events reduce the performance of ANN!

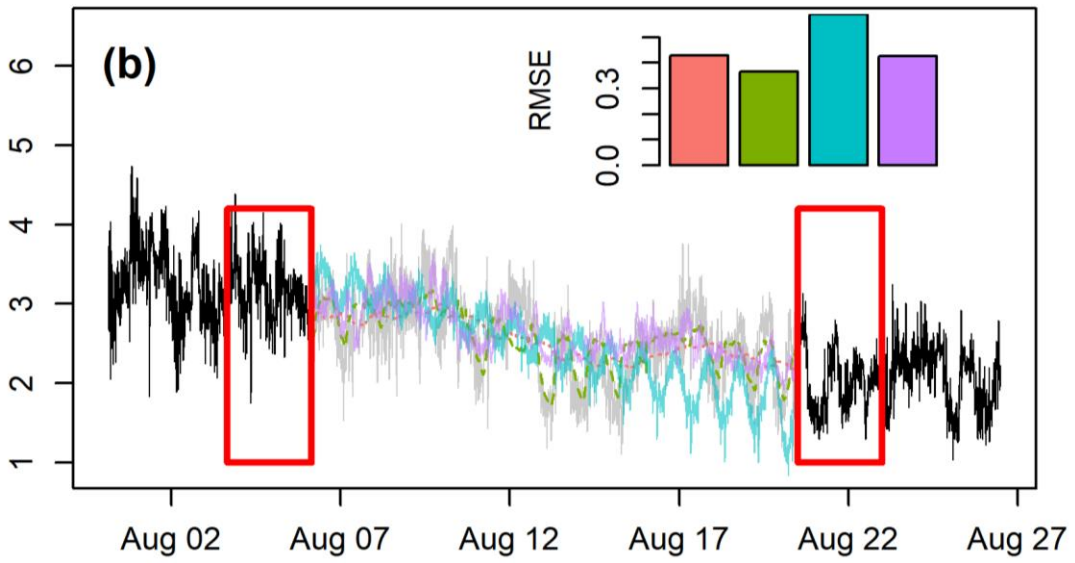


SSA for large gaps: Good & bad?

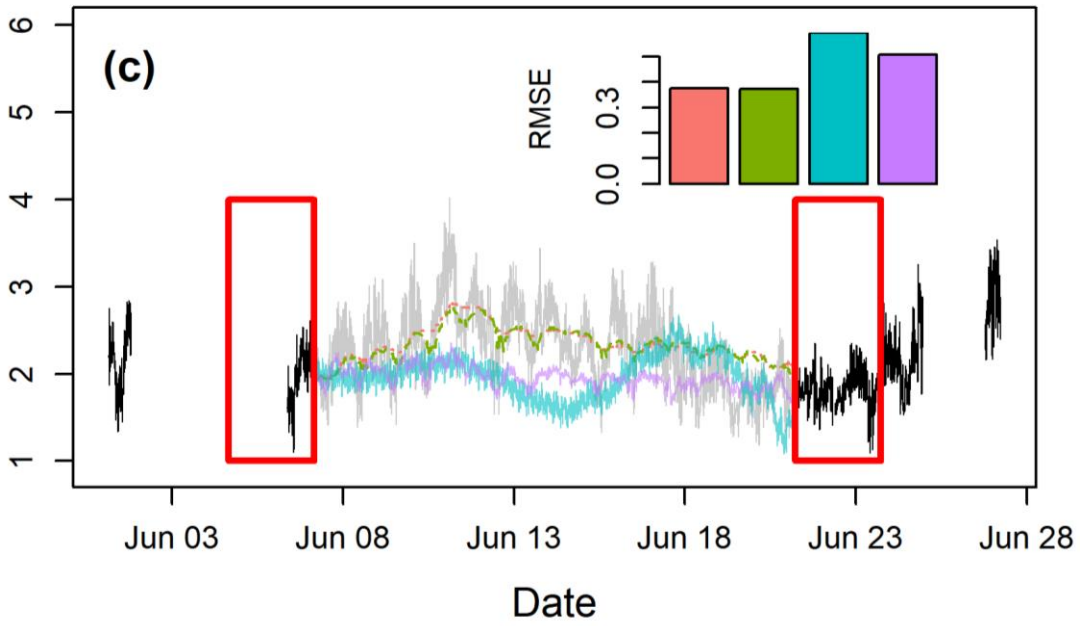
Case 1: good diurnal pattern



Case 2: irregular diurnal pattern



Case 3: missing data in sampling window



Conclusion

- NLS
  - Based on well-acknowledged model - most commonly used!
  - Low variance

- ANN:
  - Multiple input variables
  - Great potential to improve

- SSA:
  - Great for small gaps, especially in winter
  - No other variables needed

- EM:
  - Require reference dataset
  - No environmental variable needed

SSA and **EM** have great potentials to be used on other flux data (e.g., eddy covariance, CH_4 flux) that needs to be tested!

References

- J. Lloyd, J.A. Taylor. On the temperature dependence of soil respiration. Funct. Ecol., 8 (1994), pp. 315-323
- Golyandina, N., Korobeynikov, A. and Zhigljavsky, A., 2018. Singular spectrum analysis with R.
- W.L. Junger, A. Ponce de Leon. Imputation of missing data in time series for air pollutants. Atmos. Environ., 102 (2015), pp. 96-104