ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

A comparison of methods for gap-filling sensible and latent heat fluxes in different climatic conditions

Erkka Rinne (erkka.rinne@fmi.fi), Henriikka Vekuri, Mika Aurela & Juha-Pekka Tuovinen

Eddy covariance **flux measurements need to be gap-filled** when utilising the data for the calculation of annual balances. The measurement technique itself is prone to errors and technical failures may also lead to gaps of various lengths. Gap-filling of the flux time series is typically based on estimating statistically representative values based on various environmental variables through linear regression, lookup tables or machine learning methods.

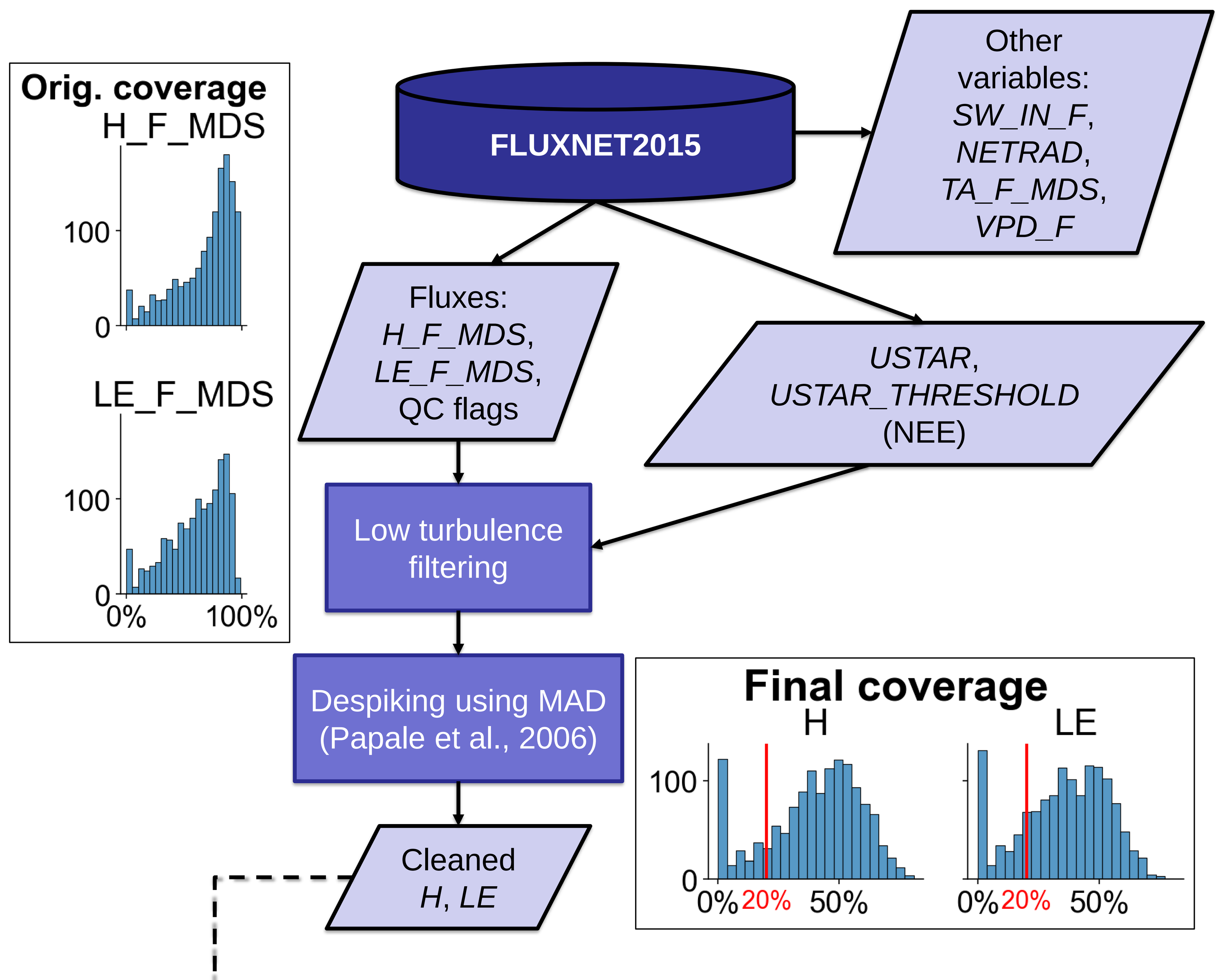
A large number of methods for the imputation of energy fluxes have been applied and compared in recent literature (Zhu et al., 2022; Mahabbati, 2022; Khan, Jeon, and Jeong, 2021; Foltýnová, Fischer, and McGloin, 2019). Both latent and sensible heat fluxes are strongly driven by the incoming solar radiation, and it is usually used as an independent variable in gap-filling methods. Vekuri et al. (2023) showed that marginal distribution sampling (MDS), **a widely used method for gap-filling carbon dioxide fluxes, creates a systematic bias in higher latitudes**, where the distribution of incoming radiation is highly skewed.

We assessed the performance of MDS in predicting sensible (H) and latent heat (LE) fluxes and also compared against a machine learning algorithm (XGB) as well as simple linear regression (LR). Measurement data is from the Northern hemisphere sites in the FLUXNET2015 dataset (Pastorello et al., 2020).

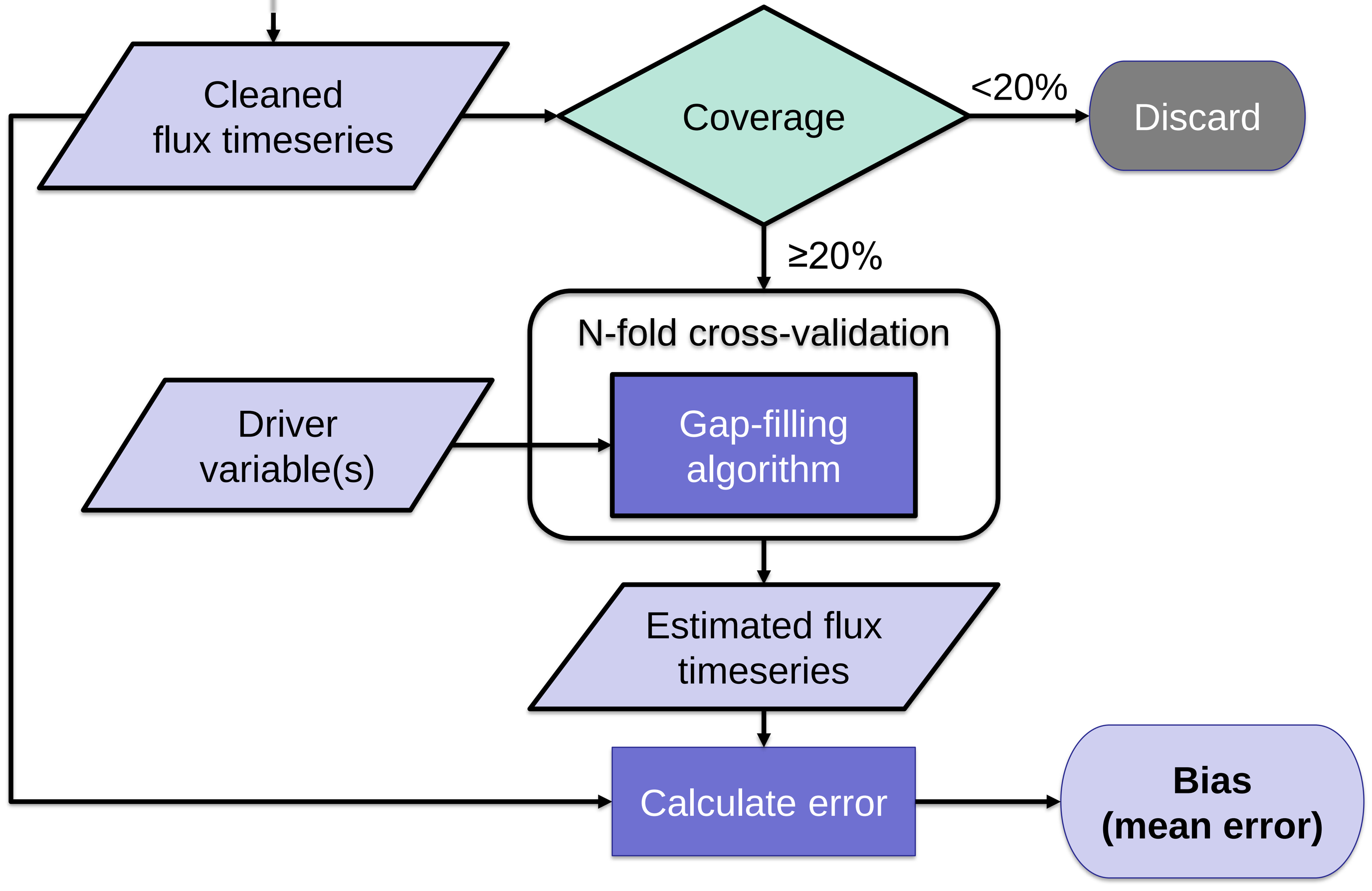
Gap-filling methods used

Method	Description	Cross-validation
MDS	Marginal distribution sampling lookup table (Reichstein et al., 2005). Drivers: SW_IN, VPD and TA	Each point
MDS_DT	Modified version of MDS taking into account radiation skewness	Each point
XGB	Gradient boosted decision trees using the eXtreme Gradient Boosting library and a whole year for training (see Vekuri et al., 2023). Drivers: SW_IN, VPD and TA	100 fold
LR	Localised linear regression against incoming radiation, training window at least 20 samples and up to 28 days wide.	20 fold

Data preparation

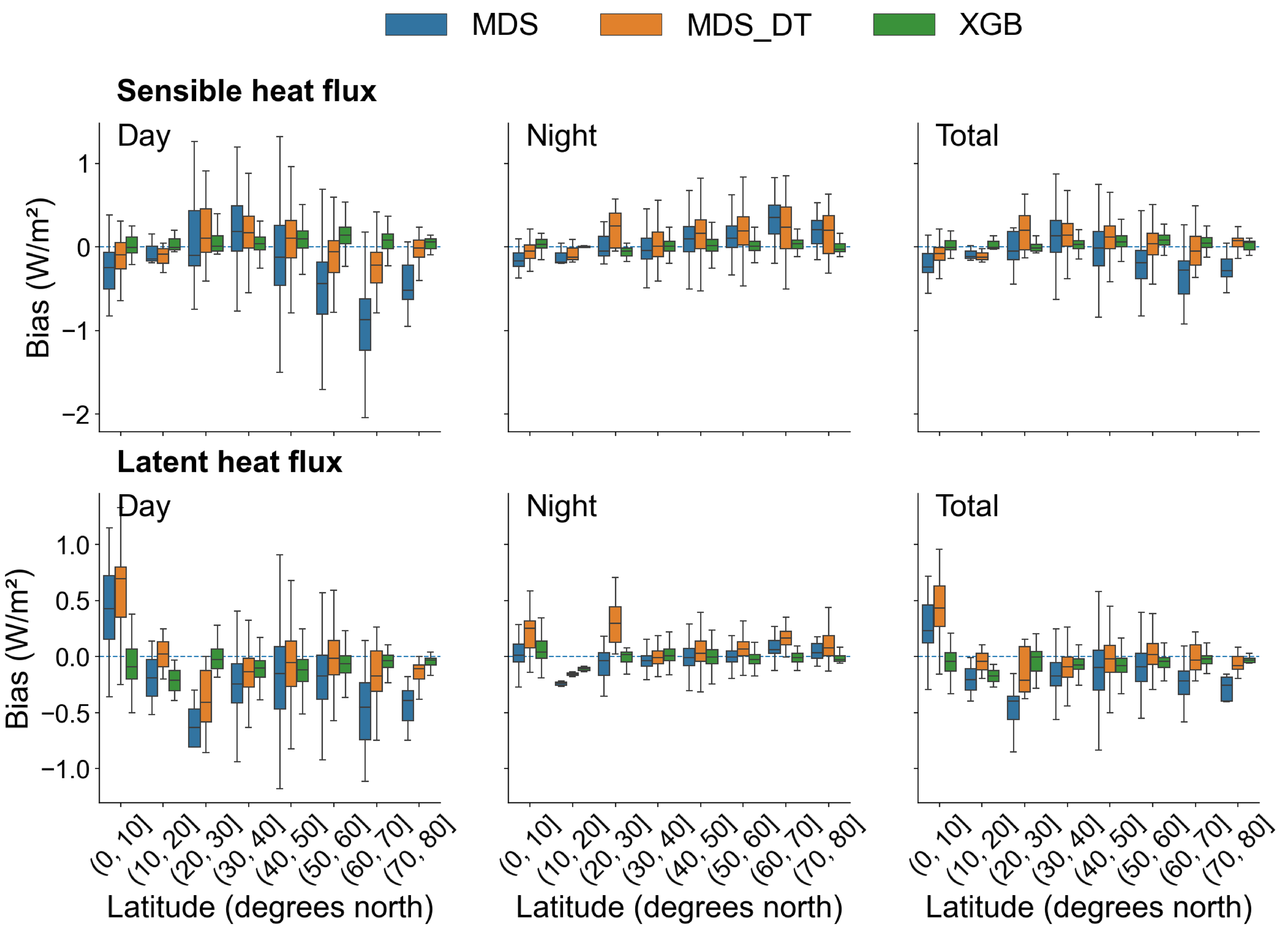


Estimation of gap-filling bias
(single site-year, flux and method)

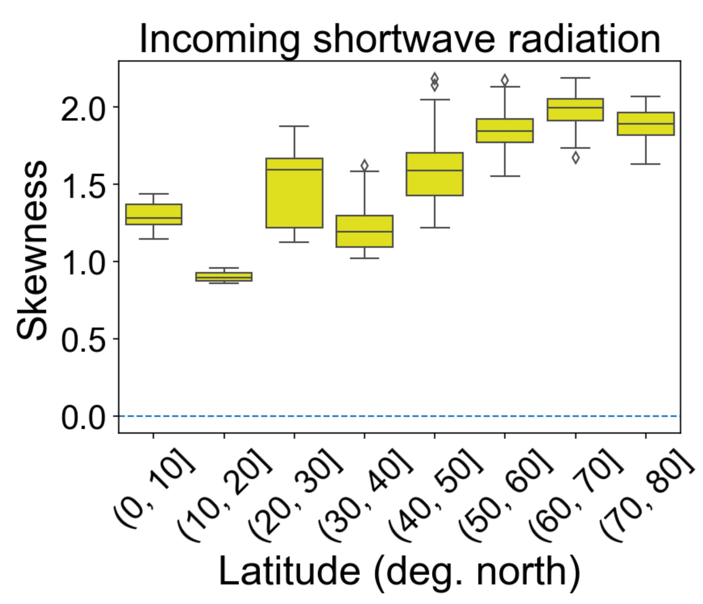


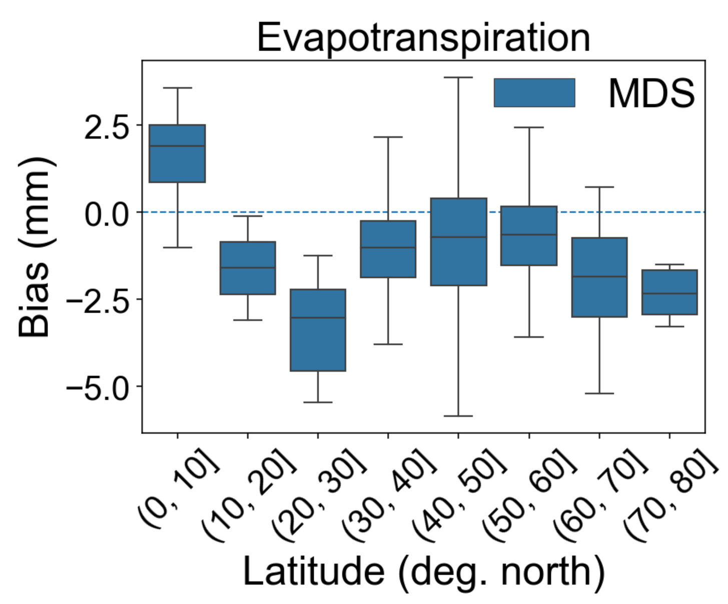
Results

- MDS produces a systematic negative error increasing towards higher latitudes for both sensible and latent heat flux, more pronounced during the daytime.
- The modified version, MDS_DT reduces the systematic error and results in closer to zero biases overall.
- XGB machine learning algorithm performs best of the studied methods.

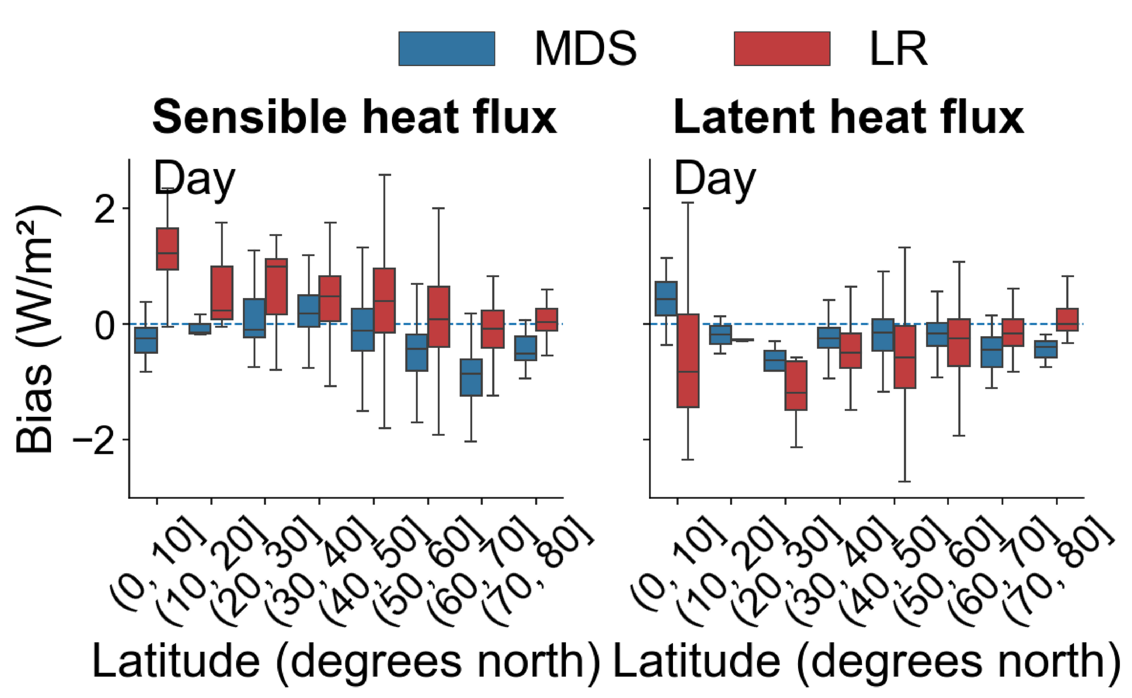


- Distribution of incoming radiation is right-skewed
- MDS lookup method finds more samples in the lower range of the radiation distribution causing a negative heat flux bias (cf. Vekuri et al., 2023)





- Simple linear regression against incoming radiation works during daytime, no similar systematic error in high latitudes
- Net radiation could work even better, especially during nights



Conclusions

- Assuming the same systematic error for all gaps leads to a total annual energy imbalance of few megajoules.
- Evapotranspiration imbalance as much as few millimetres.
- Networks (FLUXNET, ICOS) should update the processing pipelines to avoid systematic under- or overestimation of the heat fluxes or evapotranspiration during gap-filling.

References

Foltýnová, L., M. Fischer, and R.P. McGloin, ‘Recommendations for Gap-Filling Eddy Covariance Latent Heat Flux Measurements Using Marginal Distribution Sampling’, Theoretical and Applied Climatology, Vol. 139, No. 1–2, September 11, 2019, pp. 677–688.
Khan, M.S., S.B. Jeon, and M.-H. Jeong, ‘Gap-Filling Eddy Covariance Latent Heat Flux: Inter-Comparison of Four Machine Learning Model Predictions and Uncertainties in Forest Ecosystem’, Remote Sensing, Vol. 13, No. 24, January 2021, p. 4976.
Mahabbati, A., ‘Investigating the Application of Machine Learning Models to Improve the Eddy Covariance Data Gap- Filling’, The University of Western Australia, 2022.
Papale, D., M. Reichstein, M. Aubinet, E. Canfora, C. Bernhofer, W. Kutsch, B. Longdoz, et al., ‘Towards a Standardized Processing of Net Ecosystem Exchange Measured with Eddy Covariance Technique: Algorithms and Uncertainty Estimation’, Biogeosciences, Vol. 3, No. 4, November 27, 2006, pp. 571–583.
Pastorello, G., C. Trotta, E. Canfora, H. Chu, D. Christianson, Y.-W. Cheah, C. Poindexter, et al., ‘The FLUXNET2015 Dataset and the ONEFlux Processing Pipeline for Eddy Covariance Data’, Scientific Data, Vol. 7, No. 1, July 9, 2020, p. 225.
Reichstein, M., E. Falge, D. Baldocchi, D. Papale, M. Aubinet, P. Berbigier, C. Bernhofer, et al., ‘On the Separation of Net Ecosystem Exchange into Assimilation and Ecosystem Respiration: Review and Improved Algorithm’, Global Change Biology, Vol. 11, No. 9, 2005, pp. 1424–1439.
Vekuri, H., J.-P. Tuovinen, L. Kulmala, D. Papale, P. Kolari, M. Aurela, T. Laurila, J. Liski, and A. Lohila, ‘A Widely-Used Eddy Covariance Gap-Filling Method Creates Systematic Bias in Carbon Balance Estimates’, Scientific Reports, 2023.
Zhu, S., R. Clement, J. McCalmont, C.A. Davies, and T. Hill, ‘Stable Gap-Filling for Longer Eddy Covariance Data Gaps: A Globally Validated Machine-Learning Approach for Carbon Dioxide, Water, and Energy Fluxes’, Agricultural and Forest Meteorology, Vol. 314, March 1, 2022, p. 108777.