DYCO: A Python package to dynamically detect and compensate for time lags in ecosystem time series

Lukas Hörtnagl[1]

1 ETH Zurich, Department of Environmental Systems Science, ETH Zurich, CH-8092, Switzerland

Summary

In ecosystem research, the eddy covariance (EC) method is widely used to quantify the biosphere-atmosphere exchange of greenhouse gases (GHGs) and energy (Aubinet et al., 2012; Baldocchi et al., 1988). The raw ecosystem flux (i.e., net exchange) is calculated by the co-variance between the turbulent vertical wind component measured by a sonic anemometer and the entity of interest, e.g., CO_2, measured by a gas analyzer. Due to the application of two different instruments, wind and gas are not recorded at exactly the same time, resulting in a time lag between the two time series. For the calculation of ecosystem fluxes this time delay has to be quantified and corrected for, otherwise fluxes are systematically biased. Time lags for each averaging interval can be estimated by finding the maximum absolute covariance between the two turbulent time series at different time steps in a pre-defined time window of physically possible time-lags (e.g., McMillen, 1988; Moncrieff et al., 1997). Lag detection works well when processing fluxes for compounds with high signal-to-noise ratio (SNR), which is typically the case for e.g. CO_2. In contrast, for compounds with low SNR (e.g., N_2O, CH_4) the cross-covariance function with the turbulent wind component yields noisier results and calculated fluxes are biased towards larger absolute flux values (Langford et al., 2015), which in turn renders the accurate calculation of yearly ecosystem GHG budgets more difficult and results may be inaccurate.

One method to adequately calculate fluxes for compounds with low SNR is to first calculate the time lag for a *reference* compound with high SNR (e.g., CO_2) and then to apply the same time lag to the *target* compound of interest (e.g., N_2O), with both compounds being recorded by the same analyzer (Nemitz et al., 2018). DYCO uses this method by facilitating the dynamic lag-detection between the turbulent wind data and a *reference* compound and the subsequent application of found *reference* time lags to one or more *target* compounds.

Statement of need

Lag detection between the turbulent departures of measured wind and the scalar of interest is a central step in the calculation of EC ecosystem fluxes. If the covariance maximization fails to detect a clear peak in the covariance function between the wind and the scalar, current flux calculation software can apply a constant default (nominal) time lag to the respective scalar. However, both the detection of clear covariance peaks in a pre-defined time window and the definition of a reliable default time lag is challenging for compounds that are often characterized by low SNR, such as N_2O. In addition, the application of one static default time lag may produce inaccurate results when systematic time shifts are present in the raw data.

DYCO is meant to assist current flux processing software in the calculation of fluxes for compounds with low SNR. In the context of current flux processing schemes, the unique features offered as part of the DYCO package include:

- the dynamic application of progressively smaller time windows during lag search for a *reference* compound (e.g., CO_2),
- the calculation of default time lags on a daily scale for the *reference* compound,
- the application of daily default *reference* time lags to one or more *target* compounds (e.g., N_2O)
- the dynamic normalization of time lags across raw data files,
- the automatic correction of systematic time shifts in *target* raw data time series, e.g., due to failed synchronization of instrument clocks, and
- the application of instantaneous *reference* time lags, calculated from lag-normalized files, to one or more *target* compounds.

`DYCO` aims to complement current flux processing schemes. `DYCO` uses EC raw data files as input and produces final lag-removed files that can be directly used in current flux calculation software.
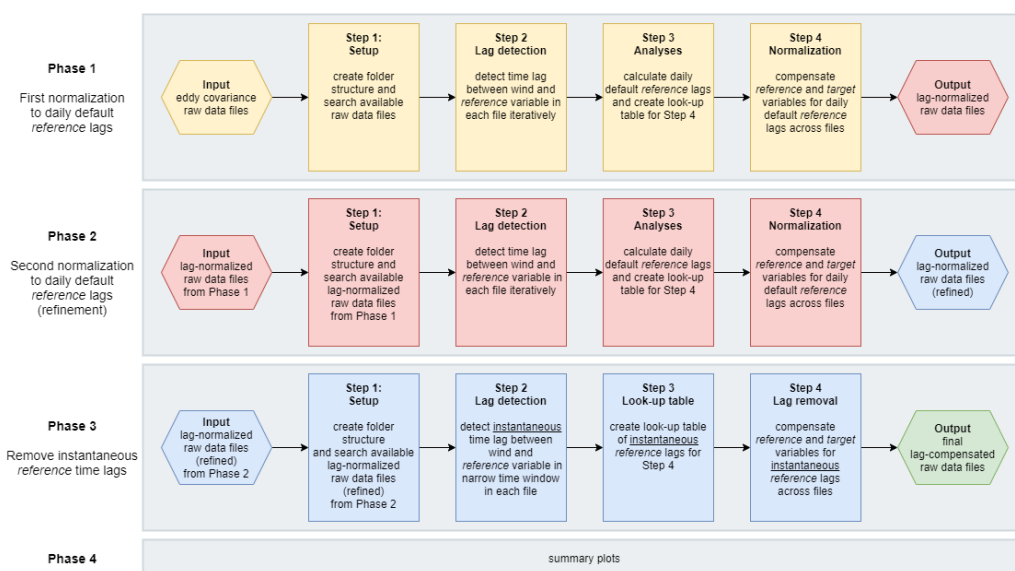
Processing chain

Figure 1. *The DYCO processing chain.*

`DYCO` uses eddy covariance raw data files as input and produces lag-compensated raw data files as output.

The full `DYCO` processing chain comprises four phases and several iterations during which *reference* lags are iteratively refined in progressively smaller search windows (Figure 1). Generally, the *reference* lag search is facilitated by prior normalization of default (nominal) time lags across files. This is achieved by compensating *reference* and *target* time series data for daily default *reference* lags, calculated from high-quality *reference* lags available around the respective date (Figure 2). Due to this normalization, *reference* lags fall into a specific, pre-defined, and therefore known time range, which in turn allows the application of increasingly narrower time windows during lag search. This approach has the advantage that *reference* lags can be calculated from a *reference* signal that shows clear peaks in the cross-covariance analysis with the wind data and thus yields unambiguous time lags due to its high SNR. In case the lag search failed to detect a clear time delay for the *reference* variable (e.g., during the night), the respective daily default *reference* lag is used instead. *Reference* lags can then be used to compensate *target* variables with low SNR for detected *reference* time delays.

Phase 1 uses eddy covariance raw data files as input and removes the daily default *reference* lag from *reference* and *target* variables, producing lag-normalized raw data files. This is achieved by first calculating the time lag between the turbulent fluctuations of the vertical wind and the turbulent measurements of the *reference* variable (e.g., atmospheric CO_2) by covariance maximization. The time window for this lag search is iteratively narrowed down (default 3 iterations). Detected time lags are analyzed and daily default *reference* lags are calculated from a selection of high-quality covariance peaks found around the respective date. The *reference* variable and *target* variables are then compensated for found daily default *reference* lags by shifting time series data accordingly. At the same time DYCO removes systematic shifts from the data, i.e., after normalization all lags are found in the same time range (Figure 2). Finally, the shifted data are saved to new files.

Phase 2 comprises the exact same processing steps as Phase 1, with the difference that instead of the original eddy covariance raw data files the lag-normalized files produced during Phase 1 are used as input. In essence, Phase 2 refines *reference* results from Phase 1. Since the *reference* time series data were already compensated for daily default *reference* lags during the previous phase, Phase 2 can apply a much narrower time window during *reference* lag search.

Phase 3 detects the final instantaneous lags between the turbulent wind and the *reference* variable, using the lag-normalized files from Phase 2 as input. Due to the compensation for daily default time lags in preceding Phases, instantaneous lags for each time period are now found close to zero. Therefore, a narrow time window of +/- 100 records is applied for the final lag search. Lags found within the acceptance limit of +/- 50 records are selected as the final instantaneous *reference* lag for the respective time period. For time periods with lags outside the acceptance limits, the final instantaneous *reference* lag is set to zero. Final instantaneous *reference* time lags are then applied to the *target* variables and final lag-compensated files are produced.

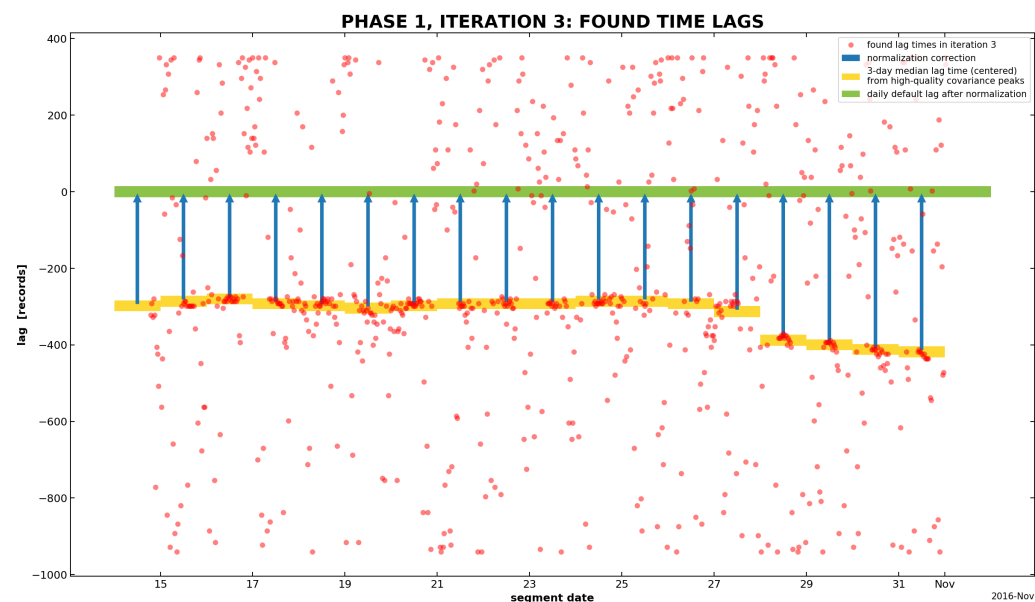**Phase 4** finalizes the processing chain by producing additional plots.

Figure 2. *Example showing the normalization of default reference time lags across raw data files recorded at the forest site* Davos *in Switzerland. Shown are found instantaneous time lags (red) between turbulent wind data and turbulent CO_2 mixing ratios, calculated daily reference default lags (yellow bars), normalization correction (blue arrows) and the daily default reference lag after normalization (green bar). Negative lag values mean that the CO_2 signal was lagged behind the wind data, e.g., -400 means that the instantaneous CO_2 records arrived 400 records (corresponds to 20s in this example) later at the analyzer than the wind*

data. Daily default reference lags were calculated as the 3-day median time lag from a selection of high-quality time lags, i.e., when cross-covariance analyses yielded a clear covariance peak. The normalization correction is applied dynamically to shift the CO_2 data so that the default time lag is found close to zero across files. Note the systematic shift in time lags starting after 27 Oct 2016.

Real-world examples

The ICOS Class 1 site Davos (CH-Dav), a subalpine forest ecosystem station in the east of Switzerland, provides one of the longest continuous time series (24 years and running) of ecosystem fluxes globally. Since 2016, measurements of the strong GHG N_2O are recorded by a closed-path gas analyzer that also records CO_2. To calculate fluxes using the EC method, wind data from the sonic anemometer is combined with instantaneous gas measurements from the gas analyzer. However, the air sampled by the gas analyzer needs a certain amount of time to travel from the tube inlet to the measurement cell in the analyzer and is thus lagged behind the wind signal. The lag between the two signals needs to be compensated for by detecting and then removing the time lag at which the cross-covariance between the turbulent wind and the turbulent gas signal reaches the maximum absolute value. This works generally well when using CO_2 (high SNR) but is challenging for N_2O (low SNR). Using covariance maximization to search for the lag between wind and N_2O mostly fails to accurately detect time lags between the two variables (noisy cross-correlation function), resulting in relatively noisy fluxes. However, since N_2O has similar adsorption / desorption characteristics as CO_2, it is valid to assume that both compounds need approximately the same time to travel through the tube to the analyzer, i.e., the time lag for both compounds in relation to the wind is similar. Therefore, `DYCO` can be applied (i) to calculate time lags across files for CO_2 (*reference* compound), and then (ii) to remove found CO_2 time delays from the N_2O signal (*target* compound). The lag-compensated files produced by `DYCO` can then be used to calculate N_2O fluxes. Since `DYCO` normalizes time lags across files and compensates the N_2O signal for instantaneous CO_2 lags, the *true* lag between wind and N_2O can be found close to zero, which in turn facilitates the application of a small time window for the final lag search during flux calculations.

Another application example are managed grasslands where the biosphere-atmosphere exchange of N_2O is often characterized by sporadic high-emission events (e.g., Hörtnagl et al., 2018; Merbold et al., 2014). While high N_2O quantities can be emitted during and after management events such as fertilizer application and ploughing, fluxes in between those events typically remain low and often below the limit-of-detection of the applied analyzer. In this case, calculating N_2O fluxes works well during the high-emission periods (high SNR) but is challenging during the rest of the year (low SNR). Here, `DYCO` can be used to first calculate time lags for a *reference* gas measured in the same analyzer (e.g., CO_2, CO, CH_4) and then remove *reference* time lags from the N_2O data.

Acknowledgements

References

Aubinet, M., Vesala, T., & Papale, D. (Eds.). (2012). *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*. Springer Netherlands. https://doi.org/10.1007/

978-94-007-2351-1

Baldocchi, D. D., Hincks, B. B., & Meyers, T. P. (1988). Measuring Biosphere-Atmosphere Exchanges of Biologically Related Gases with Micrometeorological Methods. *Ecology*, *69*(5), 1331–1340. https://doi.org/10.2307/1941631

Hörtnagl, L., Barthel, M., Buchmann, N., Eugster, W., Butterbach-Bahl, K., Díaz-Pinés, E., Zeeman, M., Klumpp, K., Kiese, R., Bahn, M., Hammerle, A., Lu, H., Ladreiter-Knauss, T., Burri, S., & Merbold, L. (2018). Greenhouse gas fluxes over managed grasslands in Central Europe. *Global Change Biology*, *24*(5), 1843–1872. https://doi.org/10.1111/gcb.14079

Langford, B., Acton, W., Ammann, C., Valach, A., & Nemitz, E. (2015). Eddy-covariance data with low signal-to-noise ratio: Time-lag determination, uncertainties and limit of detection. *Atmospheric Measurement Techniques*, *8*(10), 4197–4213. https://doi.org/10.5194/amt-8-4197-2015

McMillen, R. T. (1988). An eddy correlation technique with extended applicability to non-simple terrain. *Boundary-Layer Meteorology*, *43*(3), 231–245. https://doi.org/10.1007/BF00128405

Merbold, L., Eugster, W., Stieger, J., Zahniser, M., Nelson, D., & Buchmann, N. (2014). Greenhouse gas budget (CO_2, CH_4 and N_2O) of intensively managed grassland following restoration. *Global Change Biology*, *20*(6), 1913–1928. https://doi.org/10.1111/gcb.12518

Moncrieff, J. B., Massheder, J. M., Bruin, H. de, Elbers, J., Friborg, T., Heusinkveld, B., Kabat, P., Scott, S., Soegaard, H., & Verhoef, A. (1997). A system to measure surface fluxes of momentum, sensible heat, water vapour and carbon dioxide. *Journal of Hydrology*, *188-189*, 589–611. https://doi.org/10.1016/S0022-1694(96)03194-0

Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., Burba, G. G., Dengel, S., Gielen, B., Grelle, A., Heinesch, B., Herbst, M., Hörtnagl, L., Klemedtsson, L., Lindroth, A., Lohila, A., McDermitt, D. K., Meier, P., Merbold, L., Nelson, D., Nicolini, G., … Zahniser, M. (2018). Standardisation of eddy-covariance flux measurements of methane and nitrous oxide. *International Agrophysics*, *32*(4), 517–549. https://doi.org/10.1515/intag-2017-0042