EDDY COVARIANCE METHOD

DATA PROCESSING PRINCIPLES AND PROCEDURES

Gerardo Fratini
Principal Scientist, R&D
LI-COR Environmental

July 2022
FROM THIS
• Introduction to the Eddy Covariance data processing

• Eddy Covariance raw datasets

• Description of main processing steps
Several instruments from different manufacturers

Weatherized Analyzers for Field Measurements:

- **LI-7500A**
  - Open Path CO₂/H₂O Analyzer
  - Measures CO₂ and H₂O in situ; it is the most widely used open path CO₂/H₂O analyzer worldwide.

- **LI-7200**
  - Enclosed CO₂/H₂O Analyzer
  - Combines the advantages of open and closed path analyzers.

- **LI-7700**
  - Open Path CH₄ Analyzer
  - Designed for in situ measurements through the extremes of the environment.
In theory flux computation is simple

\[
F_c = \rho \cdot w'c' + \int_{z=0}^{h_m} \frac{\partial \tilde{c}}{\partial t} dz
\]
Fluxes are not measured, they are estimated!

• Fast, precise instruments measure high-frequency data of wind, gas concentration, temperature, ...

• Slow, accurate instruments measure low-frequency data of biomet variables

• Raw measurements and initial flux estimates need to be corrected for instrumental limitations, setup imperfections and less-than-ideal turbulence conditions

• Fluxes must be accompanied by quality flags and other means of quality evaluation
Sources of bad data that cannot be corrected

Preventable

• Measurements are not done inside the boundary layer of interest

• Terrain is not horizontal and uniform: advection

• Distortion of the air flow field

Not preventable

• Air flow is not fully turbulent (advection fluxes, storage)

• Turbulence is not stationary
EC ≈ Endless Correction

We “correct” for:

• Instrumental limitations:
  o Native gas measurement is density (not concentration)
  o Finite time response
  o Finite measuring volume and presence of sampling line
  o Instrumental drifts
  o Transducers’ shadowing
  o Technology-specific quirks (LASER, NDIR...)
  o Short-term malfunctions, power-downs

• Deployment limitation:
  o Measurement height
  o Leveling of the instruments
  o Lack of instruments’ co-location
Raw Eddy Covariance datasets – Fast EC data

- High frequency (1-20 Hz) data of 3D wind speed, gas concentrations, temperature, pressure, diagnostics...
**Raw Eddy Covariance datasets – Fast EC data**

<table>
<thead>
<tr>
<th>Time</th>
<th>CO2 (mmol/m³)</th>
<th>H2O (mmol/m³)</th>
<th>Total_P (kPa)</th>
<th>U (m/s)</th>
<th>V (m/s)</th>
<th>W (m/s)</th>
<th>T (°C)</th>
<th>CO2_dry (umol/mol)</th>
<th>H2O_dry (nmol/mol)</th>
<th>Cell_T (°C)</th>
<th>AGC</th>
<th>Flow_Rate (lpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00:00</td>
<td>620.54</td>
<td>98.0508</td>
<td>2.9473</td>
<td>0.0150</td>
<td>5.6496</td>
<td>2.9542</td>
<td>19.4576</td>
<td>312.599</td>
<td>16.1208</td>
<td>21.4958</td>
<td>71.23</td>
<td>16.8569</td>
</tr>
<tr>
<td>12:15:00</td>
<td>625.332</td>
<td>95.8301</td>
<td>0.78082</td>
<td>5.66779</td>
<td>3.16065</td>
<td>19.6566</td>
<td>312.608</td>
<td>16.2469</td>
<td>21.612</td>
<td>81.23</td>
<td>16.8569</td>
<td></td>
</tr>
</tbody>
</table>

**Timestamp, often not used but worth storing**

Gas molar densities are “native” gas measurements that one should always store.

3D wind components, sonic temperature, diagnostics.

Gas concentrations only meaningful in closed/enclosed path instruments.

Diagnostic information (AGC, RSSI) can be used to filter out individual measurements flagged for poor quality.

Cell T and P are essential for calculating concentrations from molar densities, or for applying WPL correction.
Raw Eddy Covariance datasets – Slow biomet data

- Low frequency (<1 Hz) data of ambient T, P, RH, radiations, precipitation, ...
Raw Eddy Covariance datasets – Metadata

- Site location, instrument models and their actual deployment (height above aground, separations, acquisition frequency..)

![Diagram showing LI-7700, Gill R3, and LI-7200 instruments](image)
Eddy Covariance Software

- EddyPro®
- EdiRe
- TK3
- EddyUH
- eddy4R
- EasyFlux® DL/PC

- Released and maintained by LI-COR®
- 7000+ users
- 1000+ peer-reviewed citations
- Used by and developed with ICOS, AmeriFlux
INTRODUCTION TO EDDY COVARIANCE DATA PROCESSING

Data processing workflow

**INPUTS**
- Raw dataset
  - High freq EC data
  - Low freq biomet
  - Metadata

**COMPUTATIONS**
- Statistical screening of raw time series
  - 9 statistical tests
  - Despiking and data exclusion

  **Time series correction**
  - Angle of attack correction
  - Axis rotation for tilt correction
  - Time delay compensation
  - Cross-wind and humidity correction of $T_s$
  - Detrending

- Flux correction
  - Correction of frequency losses
  - Compensation of density fluctuations (WPL)
  - (Off-season uptake correction)
  - (Spectroscopic correction)

- Quality parameters and other calculations
  - Micrometeorological tests
  - Footprint models
  - Turbulence calculations
  - Flux random uncertainty

**OUTPUTS**
- Time series quality flags
- Rotation angles
- Scalar time delays
- Uncorrected fluxes
- Spectra and cospectra
- Spectral correction factors
- WPL terms
- Corrected fluxes
- Flux quality flags
- Footprint estimation
- Turbulence parameters
- ...
A parenthesis on fluxes on-the-fly

• Sometimes the EC logging system allows for on-the-fly flux computation, on a 30-min basis

• Good for checking status of acquisition

• Without a complete use of metadata, on-the-fly fluxes are likely to be not accurate

• In some conditions, on-the-fly fluxes cannot be accurate (e.g. closed path systems with long sampling lines, complex topographies)

• In any case, storing raw data is essential for future reprocessing or re-evaluation of calculated fluxes
First step

Quality screening of high frequency time series and removal of implausible values
Detecting anomalies in raw time series (e.g., Vickers and Mahrt, 1997)

• Attempt to detect different problems in the time series

• Provide good/bad flags for each test, for each time series, for each raw file

• Can be used to clean up raw time series

• Complement micrometeorological flags for QA/QC of fluxes
Spike test

Good time series
Drop-outs

Good time series

STATISTICAL SCREENING OF RAW TIME SERIES
Amplitude resolution

Good time series

STATISTICAL SCREENING OF RAW TIME SERIES
Discontinuities

Good time series

STATISTICAL SCREENING OF RAW TIME SERIES

LI-COR Biosciences
Other 5 statistical tests

• Variables shall not get physically implausible values

• Skewness and Kurtosis shall not deviate too much from normal distribution

• Angle of attack shall not be very large for too many samples

• Time delays shall be close enough to expectations

• Time series shall not be instationary
Second step

Correction of time series for instrumental and setup limitations
Angle-of-attack correction

• Corrects for the flow distortion induced by the anemometer (shadowing)

• Available only for vertical-mounting Gill anemometers (R3, WindMaster)

• If not applied, all fluxes are underestimated by some 10-15%!

Kochendorfer et al., 2011
Nakai and Shimoyama, 2012
Nakai and Shimoyama, 2012
Axis rotations for tilt correction

• Correct anemometer misalignment with respect to local flow streamlines

• If not applied, all fluxes are likely to be strongly overestimated

• Different correction approaches:
  
  • Double rotations, triple rotations
Axis rotations for tilt correction – Double rotation

- Assumes that average vertical wind component is zero on a 30-min basis (no vertical advection fluxes)

Step 1: Nullify \( \nu \) component of average wind vector

\[
\begin{align*}
    u_1 &= u_m \cos \gamma + v_m \sin \gamma \\
    v_1 &= -u_m \sin \gamma + v_m \cos \gamma \\
    w_1 &= w_m
\end{align*}
\]

\[\gamma = \tan^{-1}\left(\frac{v_m}{u_m}\right)\]

Step 2: Nullify \( \omega \) component of average wind vector

\[
\begin{align*}
    u_2 &= u_1 \cos \alpha + w_1 \sin \alpha \\
    v_2 &= v_1 \ (= 0) \\
    w_2 &= -u_1 \sin \alpha + w_1 \cos \alpha
\end{align*}
\]

\[\alpha = \tan^{-1}\left(\frac{w_1}{u_1}\right)\]
Axis rotations for tilt correction – Sector-wise planar fit

- Assumes that average vertical wind component is zero on the long term (weeks), not on a 30-min basis
- Requires pre-processing raw data to assess rotation matrices

Griessbaum and Schmidt, 2009
Time delay compensation

• Time delays (lags) arise due to the air sampling systems, signal phase-shift, instruments separation

• If not compensated, lead to systematic flux underestimation

• The covariance maximization procedure is used to calculate time lags automatically.
Covariance maximization

- We impose a range of (plausible) artificial lags between the time series of \( w \) and gas concentration, and determine the time lag that maximizes the covariance.
The special case of water vapor in closed-path systems

• Being a sticky gas, H₂O travelling time in the sampling line depends on RH (and secondarily on T).

• Covariance maximization must be adapted to account for such dependency.

Runkle et al. (2012, BLM)

Nordbo et al. (2012, Tellus B)
**Time lag optimizer**

- A pre-processing step for an in-situ assessment of the dependency of H₂O time lag on RH
Detrending

• Calculation of turbulent fluctuations (elimination of non turbulent trends)

\[ H = \rho c_p \overline{w'T'} \]
\[ T' = T - \overline{T} \]
Detrending

• Calculation of turbulent fluctuations (elimination of non turbulent trends)

\[ H = \rho c_p \, \overline{w'T'} \quad T' = T - \overline{T} \]
Detrending

- Long-term (30-min) motions are to be expected, or should they be interpreted as non-turbulent trends?
OPEN-PATH AND CLOSED-PATH INSTRUMENTS WITHOUT ADEQUATE PROTECTION AGAINST POLLUTANTS ARE SUBJECT TO CONTAMINATION

- If the optical path is obstructed by pollutants, concentration measurement can be affected (usually shows up as a drift in time)

**Instrumental drift**

---

**CORRECTION OF TIME SERIES**

**Fratini, internal LI-COR report**
Instrumental drift

- For dynamics with clear patterns and “reference values” (e.g. atmospheric CO$_2$), drifts can easily be recognized. For H2O it’s much less obvious.

- Effects of contamination depends on pollutants size, effective refractive index, etc.

- RSSI (Signal Strength) is a proxy, but not always a reliable one.

_Fratini, internal LI-COR report_
CORRECTION OF TIME SERIES

Instrumental drift

• It is commonly believed that errors in mean concentrations do not bias resulting fluxes.

• After all:  
  \[ F = \bar{\rho}_a \cdot \bar{w}' \bar{c}' \]

• In reality, on account of the non-linearity of calibration curves, this is not true:

Fratini et al. 2014, Biogeosciences
Instrumental drift

- Flux error depends directly on concentration bias
- For water vapor, flux error depends strongly also on the concentration at which the bias occurs

_Fratini et al. 2014, Biogeosciences_
Instrumental drift

- A procedure is available in EddyPro® to correct concentration biases
- LE fluxes can improve as much as 8%

Fratini et al. 2014, Biogeosciences
Third step

Correction of preliminary flux estimates
Compensation of air density fluctuations (WPL term)

\[ c = d \frac{RT}{P(1 - \chi_w)} = \frac{d}{d_a} \]

- Compensate for the dependency of gas molar density measurements on fluctuations of air temperature, pressure and humidity
- This is not a correction: it is an actual term in the mass balance equation
Compensation of air density fluctuations (WPL term)

\[ d = c \frac{P(1 - \chi_w)}{RT} \]

\[ w' \downarrow \]
\[ T' \downarrow \]
\[ d' \uparrow \]

\[ w' \cdot d' < 0 \]

Apparent CO\textsubscript{2} uptake!
Compensation of air density fluctuations (WPL term)

\[ d = c \frac{P(1 - \chi_w)}{RT} \]

\[ w' \uparrow \]
\[ T' \uparrow \]
\[ d' \downarrow \]

\[ w' \cdot d' < 0 \]

Apparent CO\textsubscript{2} uptake!
Compensation of air density fluctuations (WPL term)

• Must be applied for all gas analyzers that measure gas molar densities

• Several approaches available:
  
  • Transform densities to mixing ratios at high frequency
    
    □ enclosed-path analyzers with fast T and P measurements
  
  • Neglect T and P fluctuations, compensate for humidity
    
    □ Closed-path analyzers with long sampling lines
  
  • WPL approach to add T, P and humidity terms to preliminary fluxes
    
    □ Open-path analyzers
Compensation of spectral attenuations

- Spectra attenuations arise due to:
  - Instrumental limitations (response time, volume averaging)
  - Instruments deployment (sampling lines, instrument separations, height above the ground)
  - Site characteristics (surface roughness)
  - Processing choices (finite flux averaging interval)
Compensation of spectral attenuations

- True cospectrum
- Measured cospectrum
- Filter (Transfer function)

[Graph showing cospectra and frequency with log scale on both axes, with a distinction between low and high frequency losses.]
Strategies for spectral attenuations

• **Analytic** (Moncrieff et al. (1997); Massman (2000)):

  • **PROS**: Simple, robust, physically based
  
  • **CONS**: some important sources of attenuation are not described (filters, water vapor interactions with surfaces..)

• **In-situ** (Ibrom et al. (2007); Runkle et al. (2012), Fratini et al. (2012)):

  • **PROS**: don’t make (too many) assumptions on attenuation sources, assess attenuations based on collected data

  • **CONS**: more complex, require data pre-processing, prone to errors
Other corrections (instrument-specific)

• Correction of sensible heat fluxes for effects of humidity on air temperature estimated with sonic temperature

• Cross-wind correction of sonic temperature (often included in anemometer firmware)

• «Self-heating correction» for some open-path CO2/H2O analyzer models in cold environments

• Spectroscopic correction for LASER-based analyzers (e.g. LI-7700)

• Band-broadening correction for NDIR-based analyzers
Micrometeorological QC tests: stationarity

Foken et al. 2004, Handbook of Micrometeorology

\[
(x'_w')_i = \frac{1}{N-1} \left[ \sum_j x_j w_j - \frac{1}{N} \sum_j x_j \sum_j w_j \right]
\]

Covariance calculated on single sub-period \( i \)

\[
x'_w' = \frac{1}{M} \sum_i (x'_w')_i
\]

Average covariance calculated using the \( M \) sub-periods

\[
(x'_w')_o = \frac{1}{M(N-1)} \left[ \sum_i \sum_j x_j w_j \right] - \frac{1}{MN} \sum_i \sum_j x_j \sum_j w_j \]

Covariance calculated on the full averaging period

\[
RN_{cov} = \left| \frac{(x'_w') - (x'_w')_o}{(x'_w')_o} \right|
\]

Relative difference between the two covariance estimates, should not exceed a threshold (30% or 100%)
In atmospheric turbulence the ratio between variance of a quantity and its turbulent flux is constant or function of the stability parameter.

Foken et al. 2004, *Handbook of Micrometeorology*

Relative difference between the two estimates of turbulence parameters must be below a given threshold.
Micrometeorological QC tests combination

- Stationarity and developed turbulence tests provide an individual “flag” (one per test)
- The two flags are combined in a single general flag that can be used to filter the data

<table>
<thead>
<tr>
<th>a</th>
<th>range</th>
<th>b</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-15%</td>
<td>1</td>
<td>0-15%</td>
</tr>
<tr>
<td>2</td>
<td>16-30%</td>
<td>2</td>
<td>16-30%</td>
</tr>
<tr>
<td>3</td>
<td>31-50%</td>
<td>3</td>
<td>31-50%</td>
</tr>
<tr>
<td>4</td>
<td>51-75%</td>
<td>4</td>
<td>51-75%</td>
</tr>
<tr>
<td>5</td>
<td>76-100%</td>
<td>5</td>
<td>76-100%</td>
</tr>
<tr>
<td>6</td>
<td>101-250%</td>
<td>6</td>
<td>101-250%</td>
</tr>
<tr>
<td>7</td>
<td>251-500%</td>
<td>7</td>
<td>251-500%</td>
</tr>
<tr>
<td>8</td>
<td>501-1000%</td>
<td>8</td>
<td>501-1000%</td>
</tr>
<tr>
<td>9</td>
<td>&gt;1000%</td>
<td>9</td>
<td>&gt;1000%</td>
</tr>
</tbody>
</table>

a: State-state test according to Equation 9.6.
b: Integral turbulence characteristics according to Equation 9.9.

Foken et al. 2004, Handbook of Micromet
How does EddyPro use biomet data?

• Fluxes (e.g. over 30 min intervals) are calculated and corrected based on:
  
  • **Covariances** calculated from fast measurement, acquired at $f > 5$Hz
  
  • **Mean quantities**, averaged over the 30 min interval

For example, CO₂ flux:

$$F = \rho_a \cdot w' c'$$

Mean value over 30 mins, calculated starting either from fast or slow measurements

Covariance over 30 mins, calculated starting from fast measurements
How does EddyPro use biomet data?

• Average air temperature and RH are involved in various flux equations, for example:

air density: \( \rho_d = \frac{P}{R_d T_a} - \rho_w \frac{m_d}{m_w} \)

WPL term: \( F = F_o + \mu \frac{E}{\rho_d} \frac{\rho_c}{1 + \mu \frac{\rho_w}{\rho_d}} + \frac{H}{\rho_a C_p T_a} \rho_c + P_{term} \)

H correction: \( H = \rho C_p \overline{w'T_a'} + \rho C_p \frac{-0.51 T_a E}{\rho_a} \)
How does EddyPro use biomet data?

- Global ($R_g$) and longwave incoming ($R_l$) radiations and PAR are needed in the off-season uptake correction of the LI-7500(A), Burba et al. 2008:

<table>
<thead>
<tr>
<th>Parameters for multiple regression* between $T_s-T_a$ and $T_a$, $R_g$, $R_l$, $U$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offset</td>
</tr>
<tr>
<td>Daytime</td>
</tr>
<tr>
<td>$T_s^{\text{bot}}-T_a$</td>
</tr>
<tr>
<td>$T_s^{\text{top}}-T_a$</td>
</tr>
<tr>
<td>$T_s^{\text{par}}-T_a$</td>
</tr>
<tr>
<td>Night-time</td>
</tr>
<tr>
<td>$T_s^{\text{bot}}-T_a$</td>
</tr>
<tr>
<td>$T_s^{\text{top}}-T_a$</td>
</tr>
<tr>
<td>$T_s^{\text{par}}-T_a$</td>
</tr>
</tbody>
</table>

*Multiple regression: daytime $T_s-T_a = \text{offset} + \text{parameter } 1 \times T_a + \text{parameter } 2 \times R_g + \text{parameter } 4 \times U$; night-time $T_s-T_a = \text{offset} + \text{parameter } 1 \times T_a + \text{parameter } 3 \times R_l + \text{parameter } 4 \times U$. All temperatures are in °C.