Supplement of

A robust data cleaning procedure for eddy covariance flux measurements

Domenico Vitale[1], Gerardo Fratini[2], Massimo Bilancia[3], Giacomo Nicolini[1], Simone Sabbatini[1], and Dario Papale[1,4]

[1]Department for Innovation in Biological, Agro-Food and Forest Systems (DIBAF), University of Tuscia, via San Camillo de Lellis, 01100 Viterbo, Italy
[2]LI-COR Biosciences Inc., Lincoln, Nebraska 68504, USA
[3]Ionian Department of Law, Economics and Environment, University of Bari Aldo Moro, Via Lago Maggiore angolo Via Ancona, 74121 Taranto, Italy
[4]Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), 01100 Viterbo, Italy

Correspondence: domvit@unitus.it, domvit@pec.it

Appendix A: High-frequency eddy covariance time series characteristics

In order to define a simulation model suitable for the representation of raw, high-frequency eddy covariance (EC) data, we preliminarily established their main statistical properties based on observed data. In particular, two aspects were taken into consideration: the temporal dependence of each time series and the correlation between vertical wind speed and other atmospheric scalars. The assessment was performed by means of the sample autocorrelation function (ACF) and of the sample cross-correlation function (CCF), respectively.

The ACF provides the correlation of a signal with a delayed copy of itself as a function of the time lag h. Consider a sequence of data y_t with $t = 1, .., n$, the ACF can be defined as

$$r_h = \frac{\sum_{t=h+1}^{n}(y_t - \overline{y})(y_{t-h} - \overline{y})}{\sum_{t=h+1}^{n}(y_t - \overline{y})} \qquad \text{for} \qquad h = 1, 2, ... \tag{A1}$$

where $\overline{y}$ is the sample mean.

The CCF provides a measure of the linear dependence between two signals as a function of the displacement of one relative to the other. Consider two signals x_t and y_t, the CCF at lag h can be defined by $\rho_h(x_t, y_t) = \text{cor}(x_t, y_{t-h}) = \text{cor}(x_{t+h}, y_t)$.

Illustrative examples of high-frequency time series from Fi-Sii and their respective ACFs and CCFs (after time series alignment by means of the maximum covariance procedure as implemented in the EddyPro® software, LI-COR Biosciences, 2019) are depicted in Figure A1. Panels a, b and c represent ideal situations where CCFs exhibit a dominant peak at lag 0. Typically, this happens when the ACF decreases exponentially towards 0 as the time lag increase. Conversely, panels d-i refer to a selection of raw data where the ACF for scalar variables decays towards 0 very slowly, a symptom of very strong persistence, meaning that past values have a long-lasting impact on following values. In such cases, either the dominant peak of the CCFs at lag 0 does not represent their absolute maximum (or minimum) (panels d and e) or, when the underlying correlation is of very low magnitude, the CCF presents a shape of dubious interpretation (panels f-i). Time series characterized by high temporal dependence exhibit a markedly unstable long-run behavior. The nature of such a non-stationarity cannot always be imputed solely to the presence of deterministic trend components. As we will describe in more detail in Appendix B, a particular class of non-stationary stochastic processes, known as difference-stationary processes, characterized by the presence of stochastic trends provides a more appropriate modelization framework.

Similar considerations hold for data collected at the other EC sites considered in this work (see Section 2.5 of the manuscript). To summarize the main characteristics of the ACF of EC raw data, Figure A2 shows the variability of r_h at lag $h = 1$ for vertical wind speed (w), sonic temperature (TSonic), carbon dioxide (CO_2) and water vapour (H_2O) collected at 10 EC sites. Such estimates are often larger than 0.95 and close to unity, indicative of a high degree of serial dependence. Lower values (around 0.90) were found for w, at BE-Lon and Fi-Sii sites, and for CO_2 at SE-Htm. No significant difference between day- and night-time was observed suggesting that, for the EC sites examined here, serial dependence is not influenced by the different ecophysiological processes at play. Figure A3 shows instead the variability of ρ_h at lag $h = 0$ between w and each scalar time series (i.e. the contemporaneous linear association entering in flux computation). Values range between -0.8 and 0.8, although most correlation estimates are less than 0.25, in absolute value.
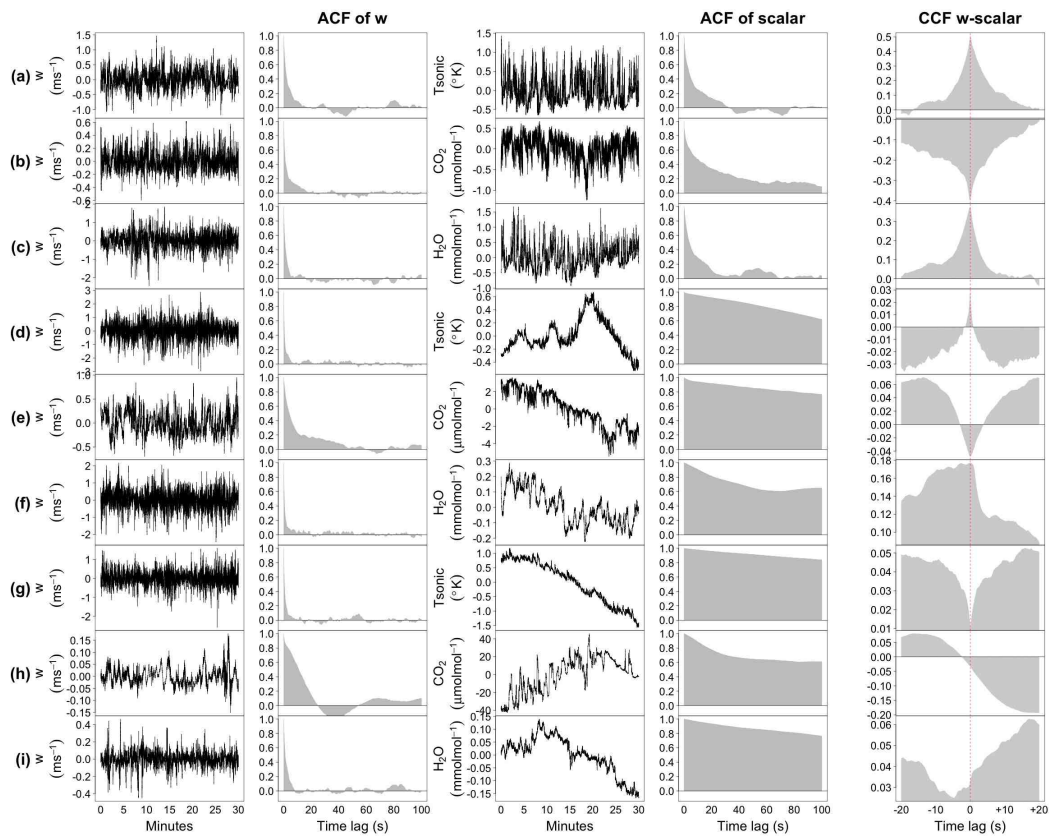
Figure A1. Illustrative examples of high-frequency time series from Fi-Sii site with their respective ACFs and CCFs. From left to right: vertical wind speed (w) and its ACF; atmospheric scalars and their ACFs; CCFs between w and scalars.
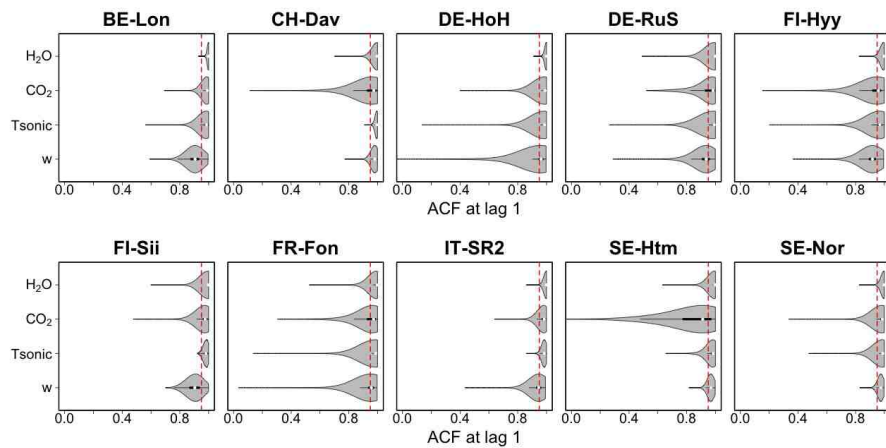
Figure A2. Distribution of the ACF at lag 1 for high-frequency EC time-series collected in 10 sites. Dashed red lines are in correspondence of 0.95.
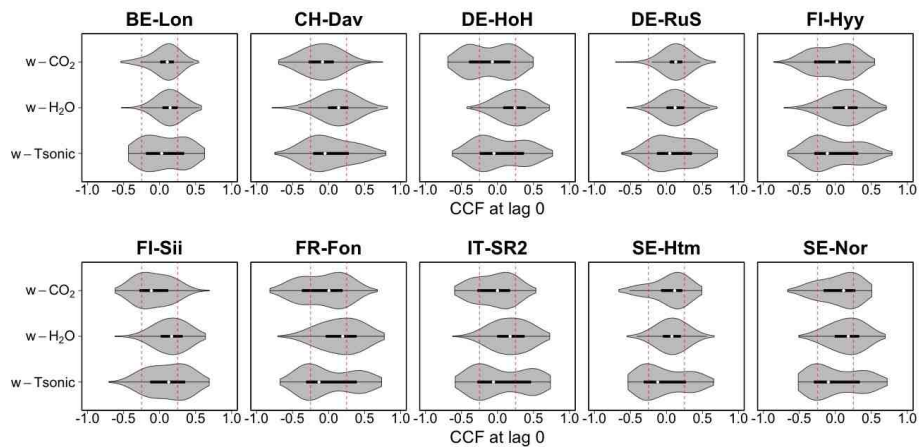
Figure A3. Distribution of the and CCF at lag 0 for high-frequency EC time-series collected in 10 sites. Dashed red lines are in correspondence of ± 0.25.

Appendix B: Stationary and non stationary time series models

A sequence of random variables $\{x_t\}$ indexed by time $t \in Z$, is called a *stochastic process*. A generic observed time series represents one possible realization of a stochastic process. Having only a single realization, in order to make statistical inferences about the structure of a stochastic process we must usually make some simplifying assumptions about that structure. The most important of such assumptions is that of stationarity, which ensures that the process is in statistical equilibrium, meaning that the laws of probability that govern the behavior of the process do not change over time.

A strictly stationary time series is one for which the probabilistic behavior of every collection of values $\{x_{(t_1)}, x_{(t_2)}, ..., x_{(t_k)}\}$ is identical to that of the time lagged set $\{x_{(t_1+h)}, x_{(t_2+h)}, ..., x_{(t_k+h)}\}$. This means that if a time series is strictly stationary, then all of the multivariate distribution functions for subsets of variables must agree with their counterparts in the shifted set for all values of the lag parameter h. Since this version of stationarity is too strong for most applications, a milder version that imposes conditions only on the first two moments of the series is often used. In particular, a time series x_t is said to be weakly stationary if (i) the mean value is constant and does not depend on time index t, and (ii) $\gamma(h) = cov(x_{t+h}, x_t)$ depends on t and h only through their difference $|t - h|$. In the following, unless specified otherwise, the term stationary will refer to weak stationarity.

Stationary processes

An important stationary process is defined by a sequence ε_t of independent, identically distributed (IID) random variables with mean 0 and variance σ_ε^2. Such a sequence is referred to as white noise and is indicated by the notation $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Its importance stems not from the fact that it is an interesting model itself, but from the fact that many useful processes can be constructed from white noise. The first-order moving average process, denoted MA(1) and defined as:

$$x_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ and μ and θ could be any constants, is an example of a stationary process constructed from white noise. An MA(1) process has mean $E[x_t] = \mu$, variance $Var[x_t] = \gamma(0) = \sigma_\varepsilon^2/(1 + \theta^2)$, while the autocovariance function is $\gamma(h) = \theta\sigma_\varepsilon^2$ for $h = 1$ and $\gamma(h) = 0$ for $h > 1$. An MA(1) process has therefore no correlation beyond lag 1. Its ACF at lag 1 is equal to $\frac{\gamma(h)}{\gamma(0)} = \frac{\theta}{1+\theta^2}$. As a consequence the largest value that lag-one ACF can attain is 0.5 when $\theta = +1$ and the smallest value is -0.5, which occurs when $\theta = -1$. A realization of MA(1) process with $\theta = 0.9$ and its sample ACF are depicted in panel a of Figure B1.

A widely used model to describe time-varying processes is the first-order autoregression, abbreviated AR(1), defined as

$$x_t = \mu + \phi x_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. An AR(1) process has mean $E[x_t] = \mu/(1 - \phi)$, variance $Var[x_t] = \gamma(0) = \sigma_\varepsilon^2/(1 - \phi^2)$, while the autocovariance function is $\gamma(h) = \frac{\sigma_\varepsilon^2 \phi^h}{1-\phi^2}$. An AR(1) process is stationary only when $|\phi| < 1$. The ACF of a stationary AR(1) process is equal to $\frac{\gamma(h)}{\gamma(0)} = \phi^h$. Therefore for a stationary AR(1), the magnitude of the ACF decreases exponentially with rate ϕ

as the number of lags h increases. If $0 < \phi < 1$, all correlations are positive; If $-1 < \phi < 0$ the ACF at lag 1 is negative and the signs of successive autocorrelations alternate from positive to negative, with their magnitudes decreasing exponentially. Panels b and c of Figure B1 depict realizations of two stationary AR(1) processes with ϕ equal to 0.9 and 0.99, respectively.

Further extensions of MA and AR processes can be taken into account to model more complex time series dynamics. For example, MA or AR processes could be of higher order or combined together to form the so-called ARMA models (the reader is referred to Hamilton, 1994; Brockwell and Davis, 2002; Cryer and Chan, 2008; Shumway and Stoffer, 2017). All the above processes constitute a valid choice when data are stationary.

Nonstationary processes

When observed time series exhibit deviations from stationarity, it is required to take into account the nature of trends affecting data to avoid erroneous interpretations of the analysis results. Models useful in describing trending time series fall into two categories: *trend-stationary* (TS) and *difference-stationary* (DS) processes. A stochastic process is TS if an underlying deterministic (not necessarily linear) trend can be removed, resulting in a stationary process. Conversely, a DS process requires differencing between consecutive values to achieve stationarity. In both DS and TS processes, the mean can be increasing or decreasing over time and for this reason it can be sometime difficult to distinguish between them. However, they are different in many aspects. For example, in the presence of a shock, TS processes are mean-reverting (i.e. after the shock, the time series will converge again towards the growing mean, which was not affected by the shock) while DS processes have a permanent impact on the mean (i.e. no convergence over time).

The simplest example of a TS process is given by the linear trend model

$$x_t = \alpha + \beta \cdot t + e_t \tag{B1}$$

where e_t is, for convenience, a WN process. A TS process is clearly non-stationary since its mean change with time t. As said before, an implication of models as in (B1) is that the slope β applies for all time t. We should therefore have good reasons for assuming such a model, not just because the series looks somewhat linear over a given time period observed. The ACF of a TS process decays very slowly towards zero (see panel d of Figure B1).

A markedly unstable long-run behavior is obtained by considering the class of DS processes. The simplest specification of a DS process is the *random walk* (RW). A random walk is defined as a process where the current value of a variable is composed of the past value plus an error term defined as a white noise. A RW can be seen as an AR(1) process with $\phi = 1$ and is represented as follows:

$$x_t = x_{t-1} + \epsilon_t \tag{B2}$$

where ϵ_t is WN. Note that we may rewrite (B2) as a cumulative sum of white noise variates. That is, $x_t = x_0 + \sum_t \epsilon_t$, where x_0 represents the initial condition while $\sum_t \epsilon_t$ is representative of the stochastic trend component. The implication of a process of this type is that the best prediction of x for next period is the current value or, in other words, the process does not allow to predict the change $x_t - x_{t-1}$. If ϵ_t has a symmetric distribution around zero, then conditional on x_{t-1}, x_t has 50-50 chance to

go up or down, implying that x_t would go up or down at random. It can be shown that the mean of a random walk process is constant, but its variance is not. By means of recursive substitutions, in fact, we get

$$Var[x_t] = Var[x_{t-1} + \epsilon_t] = Var[x_{t-1}] + \sigma_\epsilon^2$$
$$= Var[x_{t-2} + \epsilon_{t-1}] + \sigma_\epsilon^2 = Var[x_{t-2}] + 2\sigma_\epsilon^2$$

5 ...

$$= Var[x_0] + t \cdot \sigma_\epsilon^2$$

Assume that x_0 is given, then $Var[x_t] = t \cdot \sigma_\epsilon^2$, which means that the variance of the process is not constant but increases without bound as time t increases and therefore the process is not stationary.

Another specification of a DS process is given by the RW with *drift* having the following form:

10 $\qquad x_t = \delta + x_{t-1} + \epsilon_t.$ $\hfill$ (B3)

For $\delta > 0$ the process will show an upward trend. Note in fact that a drift acts like a deterministic linear trend since we may rewrite (B3) as

$$x_t = x_0 + \delta \cdot t + \sum_t \epsilon_t,$$ $\hfill$ (B4)

where $x_0 + \delta \cdot t$ is the deterministic linear trend and $\sum_t \epsilon_t$ is the stochastic trend. In case of a RW with drift, therefore, not only

15 its variance is not constant but also its mean function changes with time. A realization of RW and RW with drift with their sample ACF are depicted in panels e and f of Figure B1. Both RW and RW with drift exhibits extremely high autocorrelation that does not decrease very rapidly as the lag increases, although the ACF tends to decay more slowly in presence of a drift.

On the choice of the detrending procedure

In light of the above considerations, if the ACF of the time series values either cuts off or dies down fairly quickly, then the time

20 series values can be considered stationary. On the other hand, if the ACF of the time series values either cuts off or dies down extremely slowly, then it should be considered non-stationary. In this last case, it would be appropriate to correctly identify the nature of non-stationarity, meaning distinguish between TS and DS processes. Usually this task is achieved through application of the so-called unit root tests, although, it is recognized that they have low power in finite samples leading often to ambiguous results. In this work, we do not investigate further. However, the possibility that EC data are affected by deterministic and/or

25 stochastic trends will be not neglected and will be taken into account during data simulation.

Despite being aware that further and in depth investigations are required, in the following we provide pros and cons different methods of detrending have on either TS and DS processes. To this end, consider first the case when the data were really generated by a RW with drift process as in (B4) and the detrending procedure consists in (erroneously) subtracting a linear trend from x_t:

30 $\qquad x_t - \delta \cdot t = x_0 + \sum_t \epsilon_t.$ $\hfill$ (B5)
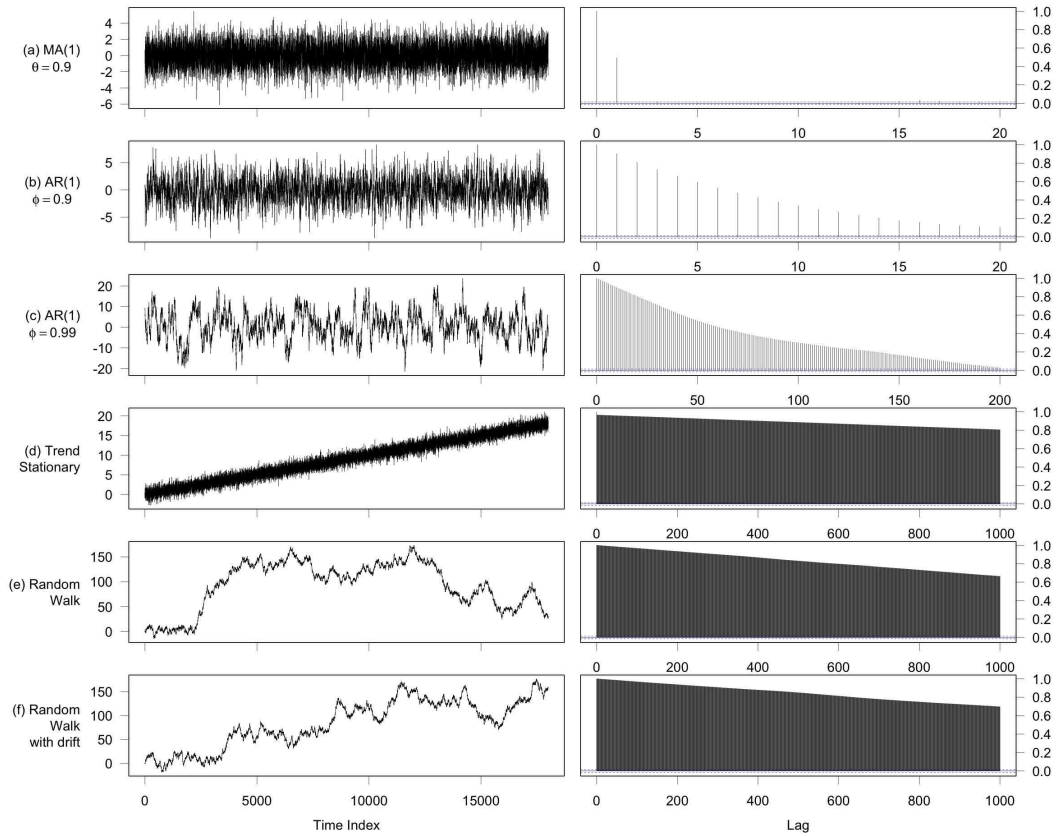
Figure B1. Realizations of stochastic processes (left panels) and corresponding autocorrelation function (ACF, right panels).

Being $Var[x_0 + \sum_t \epsilon_t] = t \cdot \sigma_\epsilon^2$, the transformed data is not stationary. Thus, subtracting a deterministic trend from a RW with drift process is not sufficient to produce stationary time series. The correct treatment for these processes consists in taking the difference between successive values, i.e. $x_t - x_{t-1} = \delta + \epsilon_t$, so differencing can get rid of trends which are built out of the summation of persistent random shocks.

5   Consider instead the case when data were generated by a TS process as in (B1) and the detrending procedure consists in (erroneously) taking difference between successive values of x_t:

$$x_t - x_{t-1} = \beta + e_t - e_{t-1}. \tag{B6}$$

Now the deterministic trend goes away and transformed data are stationary, although a unit root has been introduced in the MA representation. MA processes with unit root ($\theta = \pm 1$) are not-invertible and subject to potential analytical difficulties due to 10   the fact that standard estimation procedures leading asymptotically normal estimates are not readily available (see Plosser and Schwert, 1977, for more details).

Appendix C: Simulating two AR(1) processes with known covariance structure

In the following, we describe the modelling approach used to simulate a bivariate system of stationary time series having a known and a priori fixed correlation. For a more comprehensive and detailed exposition in the context of multivariate time series analysis, the reader is referred to the authoritative textbook of Lütkepohl (2005).

According to the main properties of EC time series outlined in Appendix A, time series were simulated from two AR(1) processes (see Appendix B for details), $x_t = \phi_x x_{t-1} + \varepsilon_{x,t}$ and $y_t = \phi_y y_{t-1} + \varepsilon_{y,t}$, representative of the vertical wind speed velocity and scalar atmospheric concentration variable, respectively. To ensure that simulated AR(1) processes have a pre-fixed correlation structure, i.e. $\mathrm{cor}(x_t, y_t) = \rho$, we considered a bivariate system of AR(1) processes[1], given by

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_x & 0 \\ 0 & \phi_y \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \sim MVN \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_{\varepsilon_x}^2 & \sigma_{\varepsilon_x \varepsilon_y} \\ \sigma_{\varepsilon_y \varepsilon_x} & \sigma_{\varepsilon_y}^2 \end{bmatrix} \tag{C1}$$

Such a system can be also seen as a first-order Vector AutoRegressive (VAR) model, and can be rewritten as $Z_t = A Z_{t-1} + U_t$, where $Z_t = (x_t, y_t)'$ and $U_t = (\varepsilon_{x,t}, \varepsilon_{y,t})'$ are (2×1) random vectors, while A is a fixed (2×2) coefficient matrix subject to parameters restriction with elements a_{12}, a_{21} equal to 0. This means that neither x_t and y_t depends on y_{t-1} and x_{t-1}, respectively. In such a way, it is allowed only a contemporaneous causal relationship between variables which is due to the correlation between $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$. Once defined the AR(1) coefficients, the target is therefore generate $\varepsilon_{x,t}, \varepsilon_{y,t}$ with a correlation structure ensuring that $\mathrm{cor}(x_t, y_t)$ is of pre-fixed amount.

To this end, denoting by Σ_Z and Σ_U the variance-covariance matrix of Z_t and U_t, respectively, then it can be checked (see Lütkepohl, 2005, Ch. 2) that

$$\Sigma_Z = A \Sigma_Z A' + \Sigma_U \tag{C2}$$

which can be solved as:

$$vec(\Sigma_Z) = [I - A \otimes A']^{-1} vec(\Sigma_U), \tag{C3}$$

where vec stands for vectorization and $\otimes$ for the Kronecker product and I is the identity matrix.

---

[1] Similar procedures are provide by:

mpiktas (https://stats.stackexchange.com/users/2116/mpiktas), How to simulate two correlated AR(1) time series?, URL (version: 2013-10-03): https://stats.stackexchange.com/q/71831

Rob Hyndman (https://stats.stackexchange.com/users/159/rob-hyndman), Generate two correlated ARMA(1,1) processes, URL (version: 2012-07-24): https://stats.stackexchange.com/q/32908

javlacalle (https://stats.stackexchange.com/users/48766/javlacalle), How do you simulate two correlated AR(p) time series?, URL (version: 2015-02-05): https://stats.stackexchange.com/q/136318

javlacalle (https://stats.stackexchange.com/users/48766/javlacalle), How do you generate correlated ARMA(1,1) models?, URL (version: 2017-04-13): https://stats.stackexchange.com/q/136502

For a bivariate system of time series, (C3) is as follow

$$
\begin{bmatrix} \sigma_x^2 \\ \sigma_{xy} \\ \sigma_{yx} \\ \sigma_y^2 \end{bmatrix} = \left(\begin{bmatrix} 1-\phi_x^2 & 0 & 0 & 0 \\ 0 & 1-\phi_x\phi_y & 0 & 0 \\ 0 & 0 & 1-\phi_x\phi_y & 0 \\ 0 & 0 & 0 & 1-\phi_y^2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 \\ \sigma_{\varepsilon_x\varepsilon_y} \\ \sigma_{\varepsilon_y\varepsilon_x} \\ 1 \end{bmatrix}, \tag{C4}
$$

where $\sigma_{xy}, \sigma_{yx}, \sigma_{\varepsilon_x\varepsilon_y}, \sigma_{\varepsilon_y\varepsilon_x}$ denoting covariances while the error terms $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$ are assumed, for convenience, to have unit variances. Note that from (C4), $\sigma_x^2 = 1/(1-\phi_x^2)$ and $\sigma_y^2 = 1/(1-\phi_y^2)$ are the analytical expressions of the variance of

5 AR(1) processes involved (see Appendix B).

Since, $\text{cor}(x_t, y_t) = \rho_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2\sigma_y^2}}$ by substituting the terms from (C4), we have

$$
\rho_{xy} = \frac{\sigma_{\varepsilon_x,\varepsilon_y}\sqrt{(1-\phi_x^2)(1-\phi_y^2)}}{1-\phi_x\phi_y} \tag{C5}
$$

from which it follows that to generate two AR(1) processes with a priori specified correlation ρ, we need to generate $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$ with unit variance and correlation given by

10 $$
\rho_{\varepsilon_x,\varepsilon_y} = \rho_{xy}\frac{1-\phi_x\phi_y}{\sqrt{(1-\phi_x^2)(1-\phi_y^2)}} \tag{C6}
$$

In the following we provide the R code to simulate the bivariate time series of AR(1) processes of length $n = 18000$ with $\phi_x = \phi_y = 0.95$ and pre-fixed correlation $\rho_{xy} = 0.25$:

```
rho_inn_f <- function(rho_xy, phi_x, phi_y) {
            rho_xy * (1 - phi_x * phi_y)/sqrt((1 - phi_x^2)*(1 - phi_y^2))
            }
phi_x <- 0.95 ## ar1 coefficient of x
phi_y <- 0.95 ## ar1 coefficient of y
burn.in <- 1000 ## number of observations of the burn-in period
n <- 18000 ## sample size
rho_xy <- 0.25 ## pre-fixed correlation between x and y
rho_inn <- rho_inn_f(rho_xy, phi_x, phi_y)
innovations <- mvrnorm(n + burn.in, mu=c(0,0), Sigma=cbind(c(1, rho_inn), c(rho_inn, 1)))
x <- arima.sim(list(ar=phi_x), n, start.innov=innovations[1:burn.in, 1],
            innov=innovations[(burn.in + 1) : (burn.in + n),1])
y <- arima.sim(list(ar=phi_y), n, start.innov=innovations[1:burn.in, 2],
            innov=innovations[(burn.in + 1) : (burn.in + n), 2])
```

Appendix D: Details of the sequential data cleaning applied to flux data measured at 10 EC sites

This Appendix contains the detail of the data cleaning procedure applied to H, LE and NEE flux variables measured at the 10 EC sites part of the ICOS network (www.icos-ri.eu) as described in Section 2.5 of the main paper:

- ✓ BE–Lon (Lonzee, Belgium, cropland; Moureaux et al., 2006);

- ✓ CH–Dav (Davos, Confoederatio Helvetica, evergreen needleleaf forest; Zielis et al., 2014);

- ✓ DE–HoH (Hohes Holz, Deutschland, alluvial forest; Wollschläger et al., 2017);

- ✓ DE–RuS (Selhausen Juelich, Deutschland, cropland; Mauder et al., 2013);

- ✓ FI–Hyy (Hyytiala, Finland, evergreen needleleaf forest; Suni et al., 2003);

- ✓ FI–Sii (Siikaneva, Finland, wet grassland; Haapanala et al., 2006);

- ✓ FR–Fon (Fontainebleau, France, deciduous broadleaf forest; Delpierre et al., 2015);

- ✓ IT–SR2 (San Rossore 2, Italy, evergreen needleleaf forest; Matteucci et al., 2015);

- ✓ SE–Htm (Hyltemossa, Sweden, evergreen needleleaf forest; van Meeningen et al., 2017);

- ✓ SE–Nor (Norunda, Sweden, boreal forest; Lagergren et al., 2005).
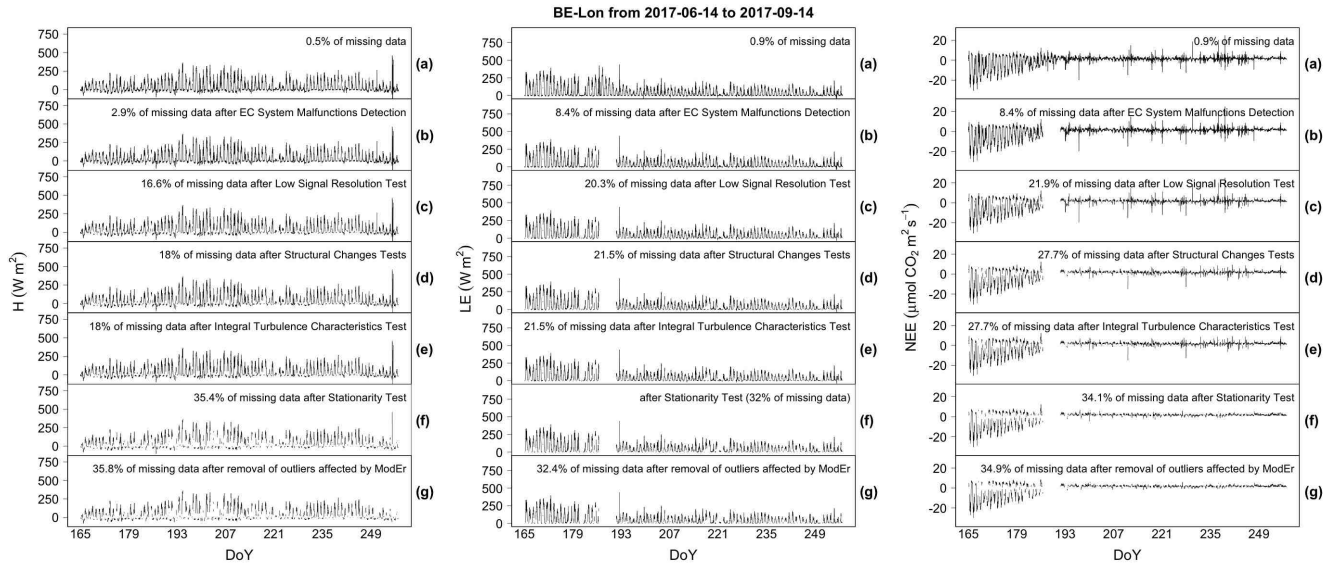
Figure D1. Sequential data cleaning procedure applied to H, LE and NEE fluxes at BE-Lon site.
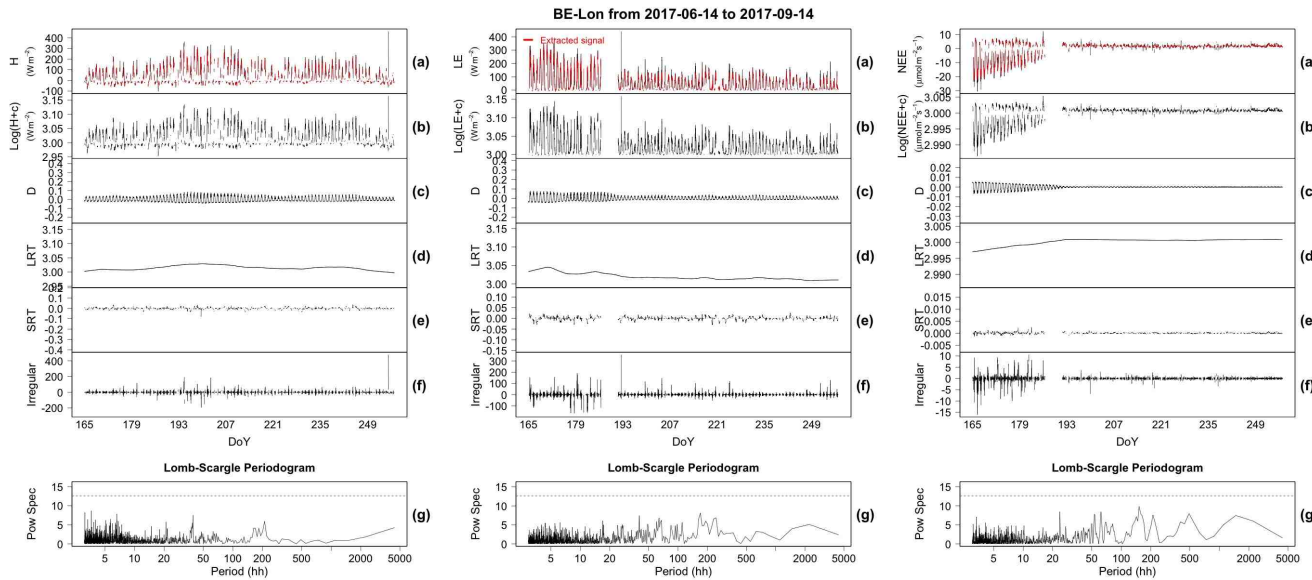
Figure D2. STL decomposition (panels a-f) of H, LE and NEE fluxes at BE-Lon site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
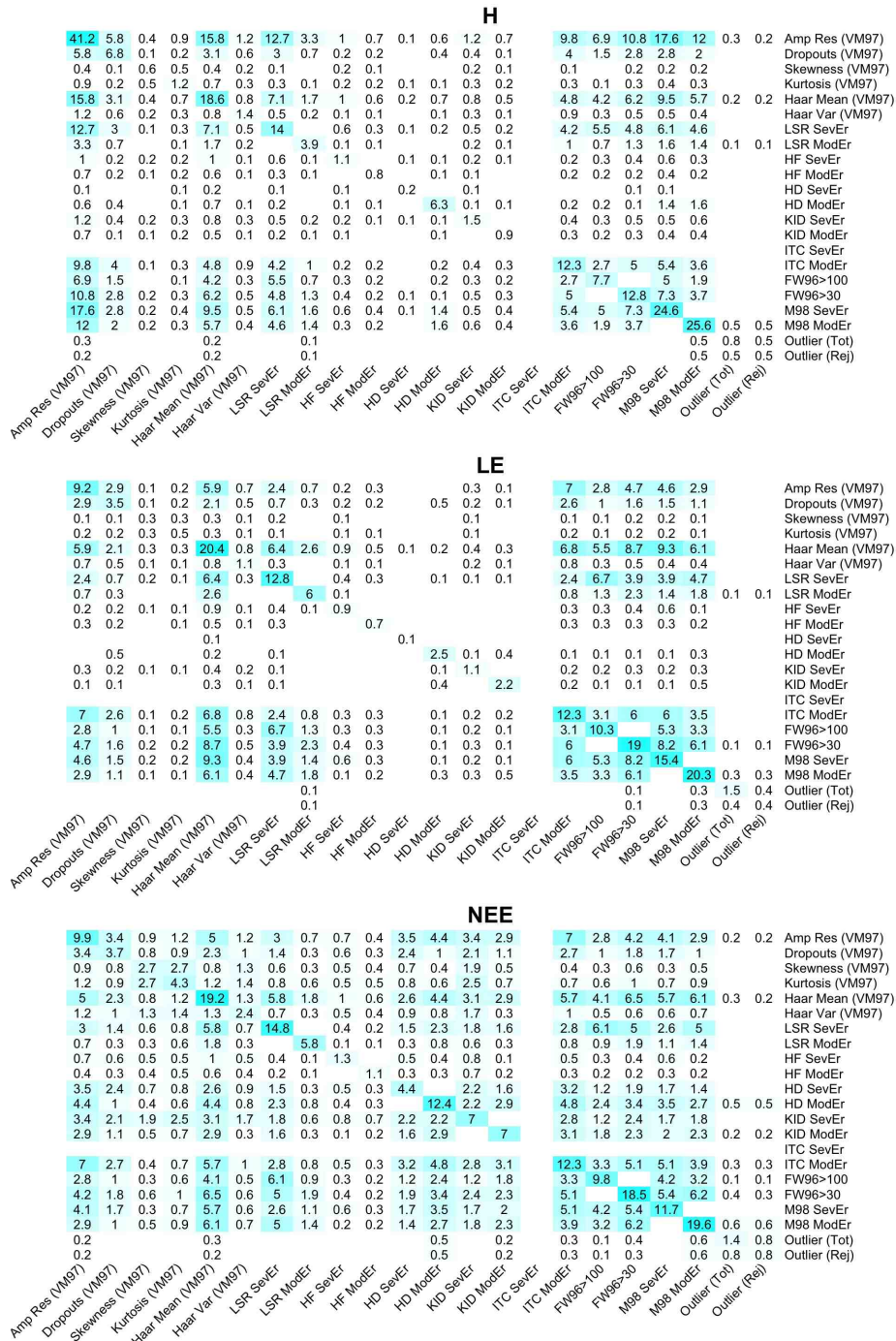
Figure D3. Percentages of H, LE and NEE fluxes at BE-Lon site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
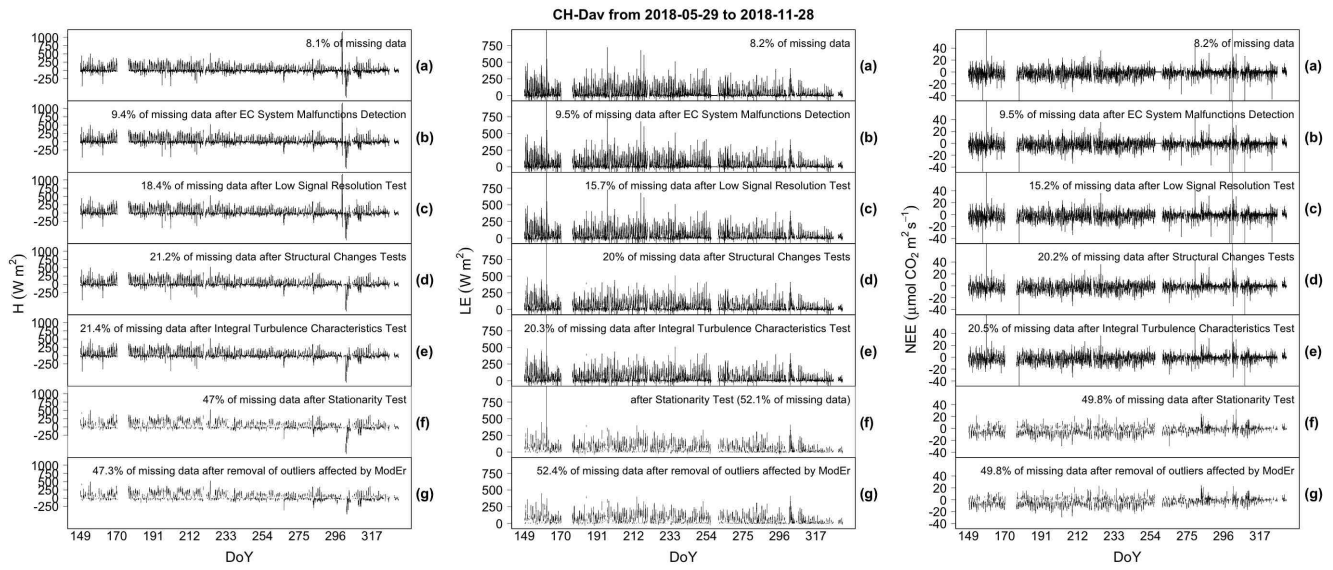
Figure D4. Sequential data cleaning procedure applied to H, LE and NEE fluxes at CH-Dav site.
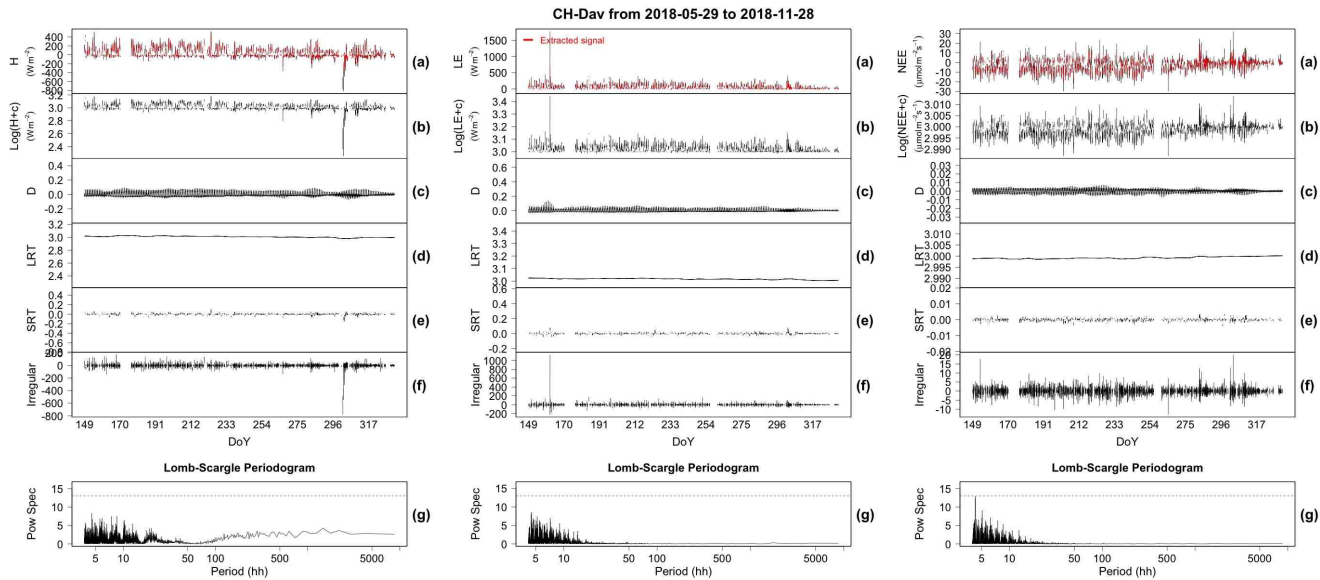
Figure D5. STL decomposition (panels a-f) of H, LE and NEE fluxes at CH-Dav site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
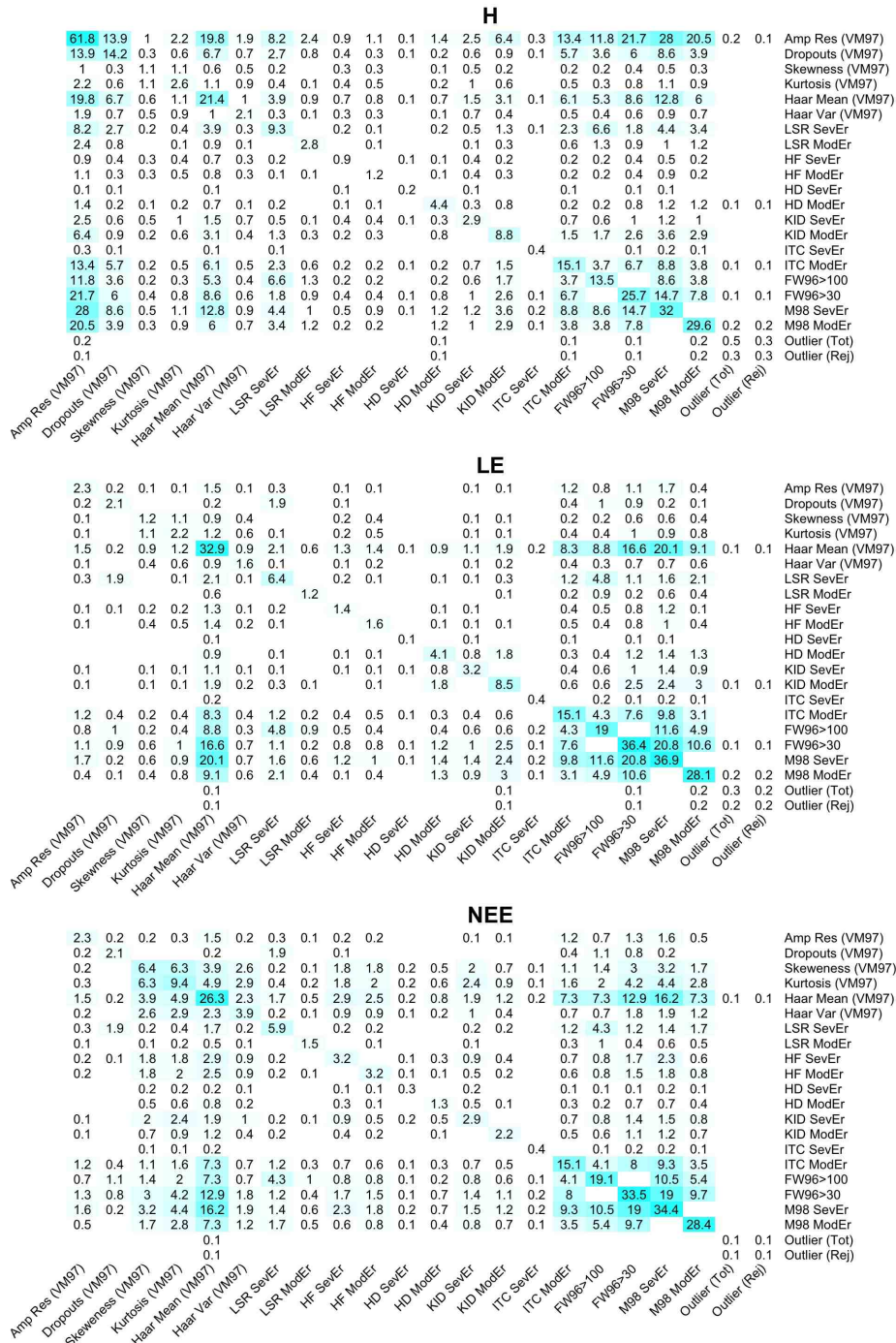
Figure D6. Percentages of H, LE and NEE fluxes at CH-Dav site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
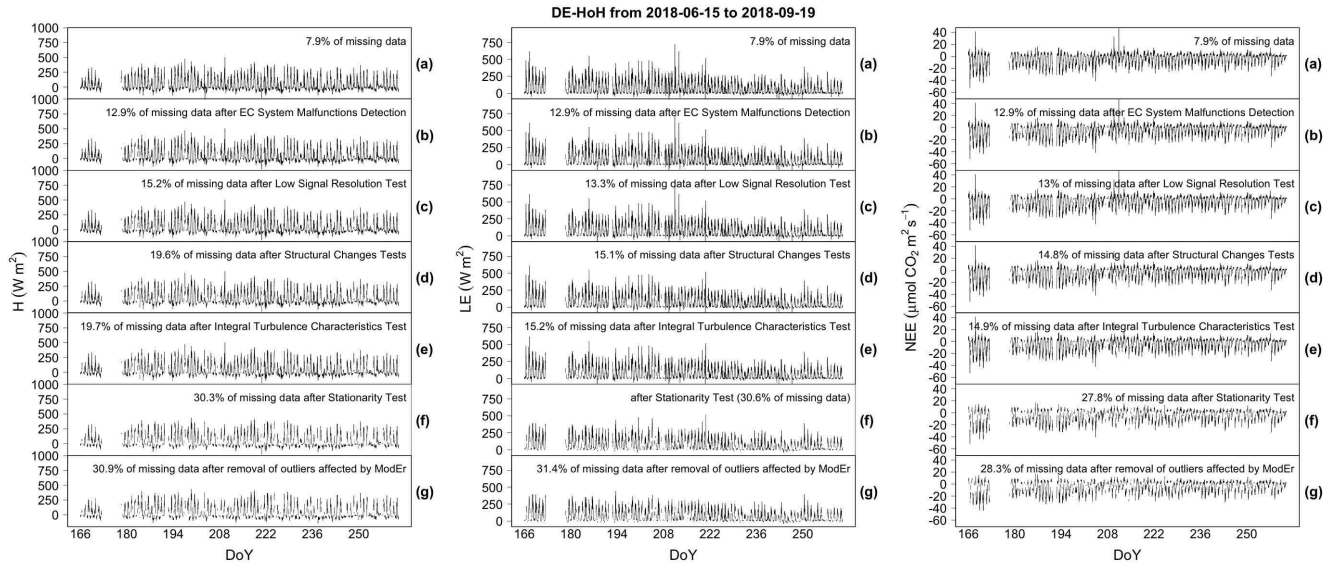
Figure D7. Sequential data cleaning procedure applied to H, LE and NEE fluxes at DE-HoH site.
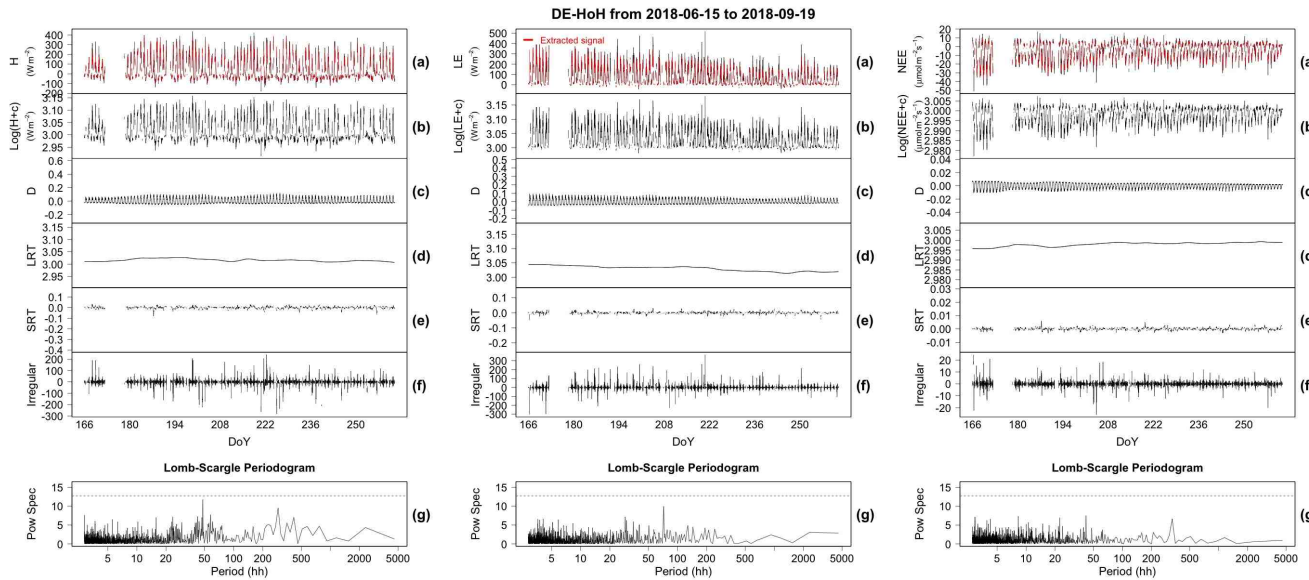
Figure D8. STL decomposition (panels a-f) of H, LE and NEE fluxes at DE-HoH site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.

Figure D9. Percentages of H, LE and NEE fluxes at DE-HoH site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
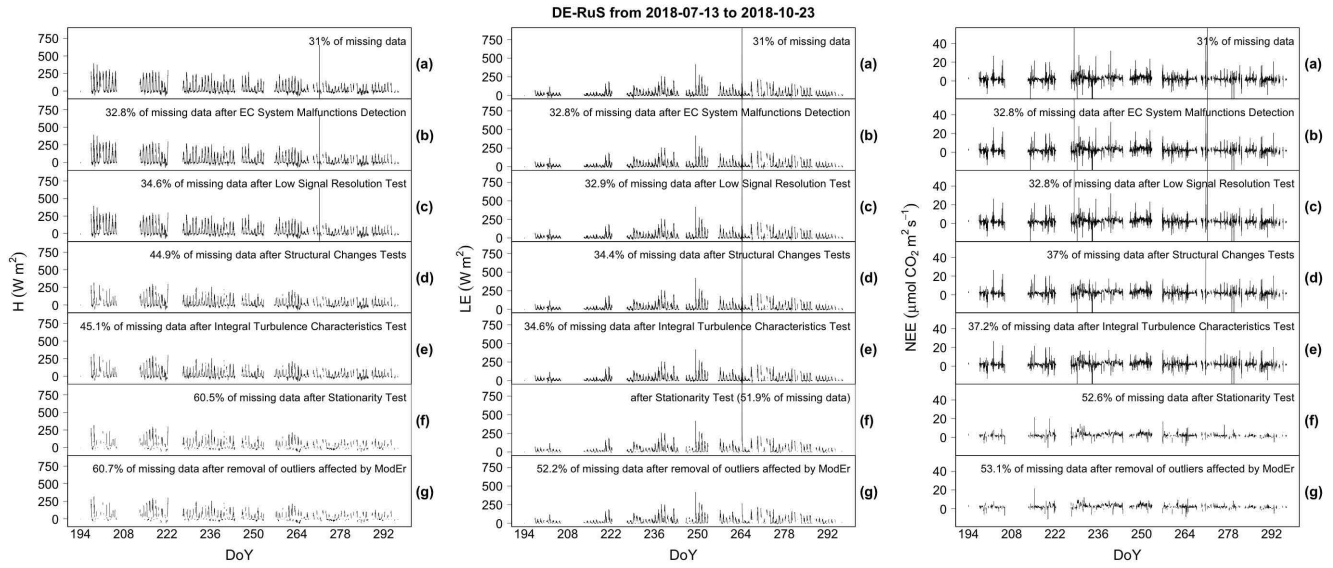
Figure D10. Sequential data cleaning procedure applied to H, LE and NEE fluxes at DE-RuS site.
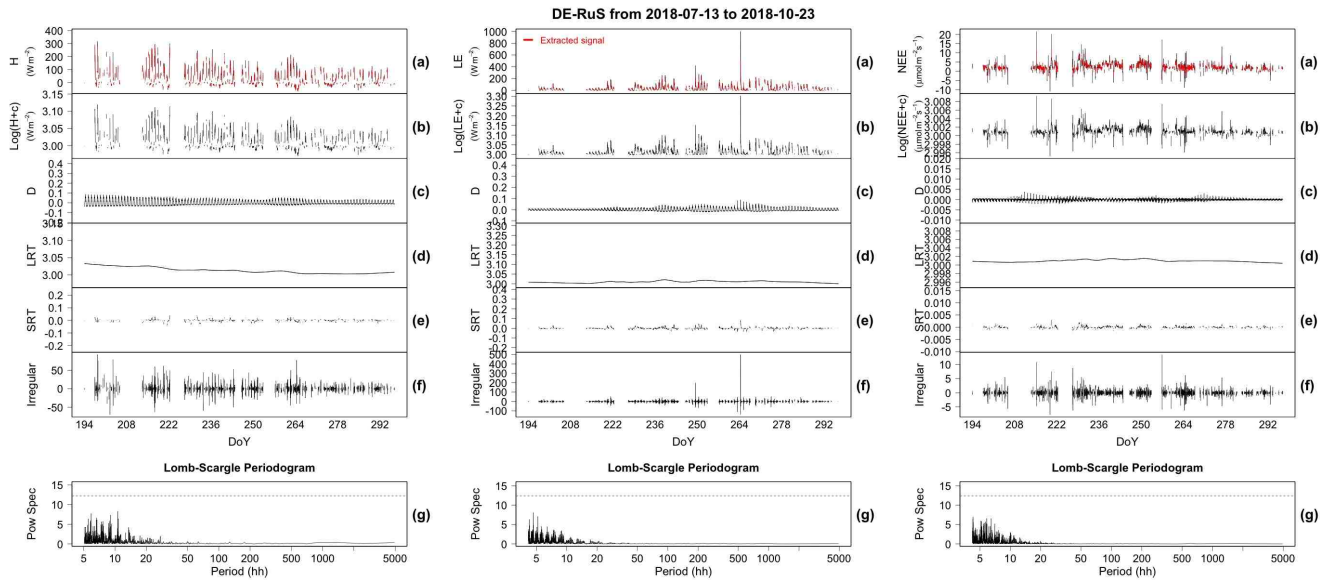
Figure D11. STL decomposition (panels a-f) of H, LE and NEE fluxes at DE-RuS site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.

Figure D12. Percentages of H, LE and NEE fluxes at DE-RuS site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
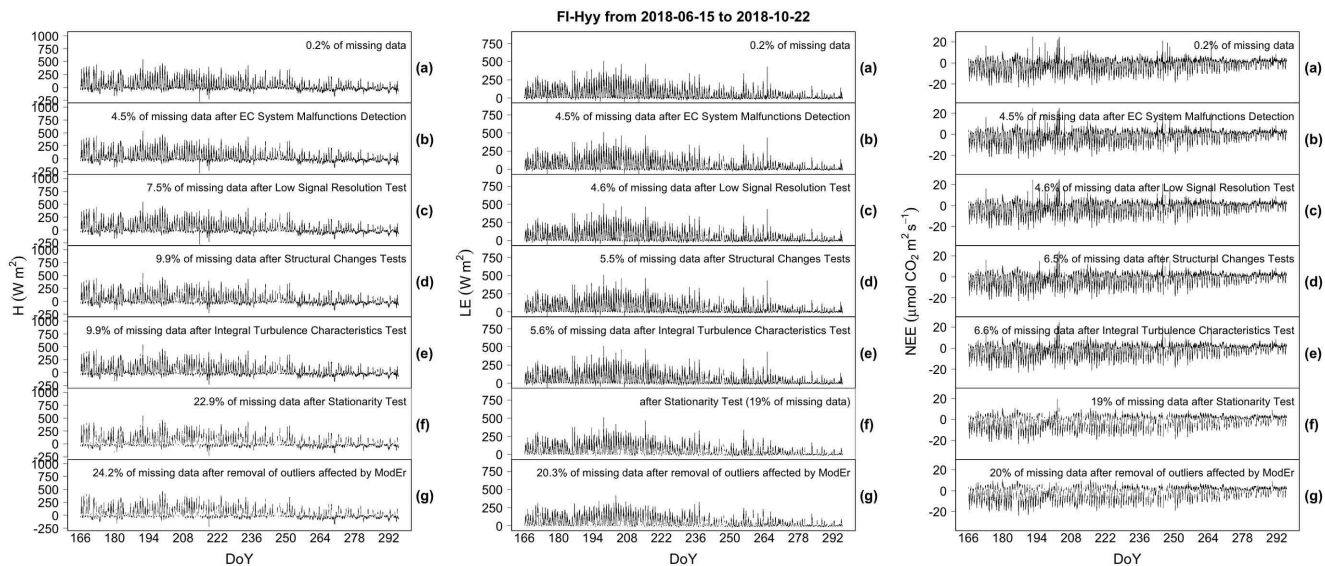
Figure D13. Sequential data cleaning procedure applied to H, LE and NEE fluxes at FI-Hyy site.
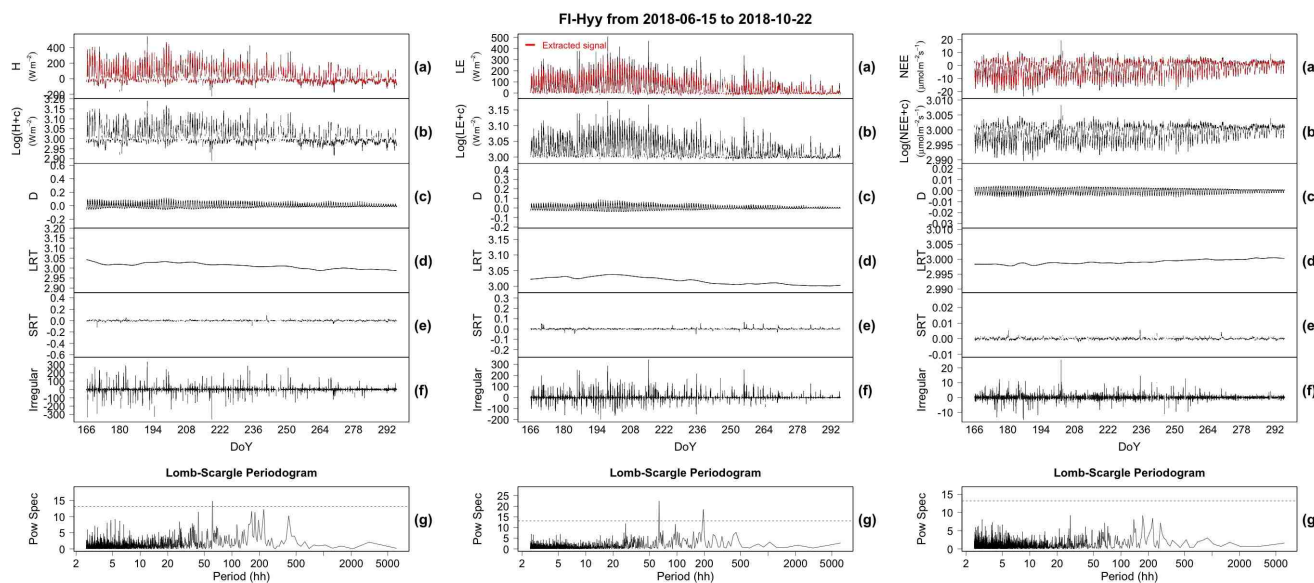
Figure D14. STL decomposition (panels a-f) of H, LE and NEE fluxes at FI-Hyy site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
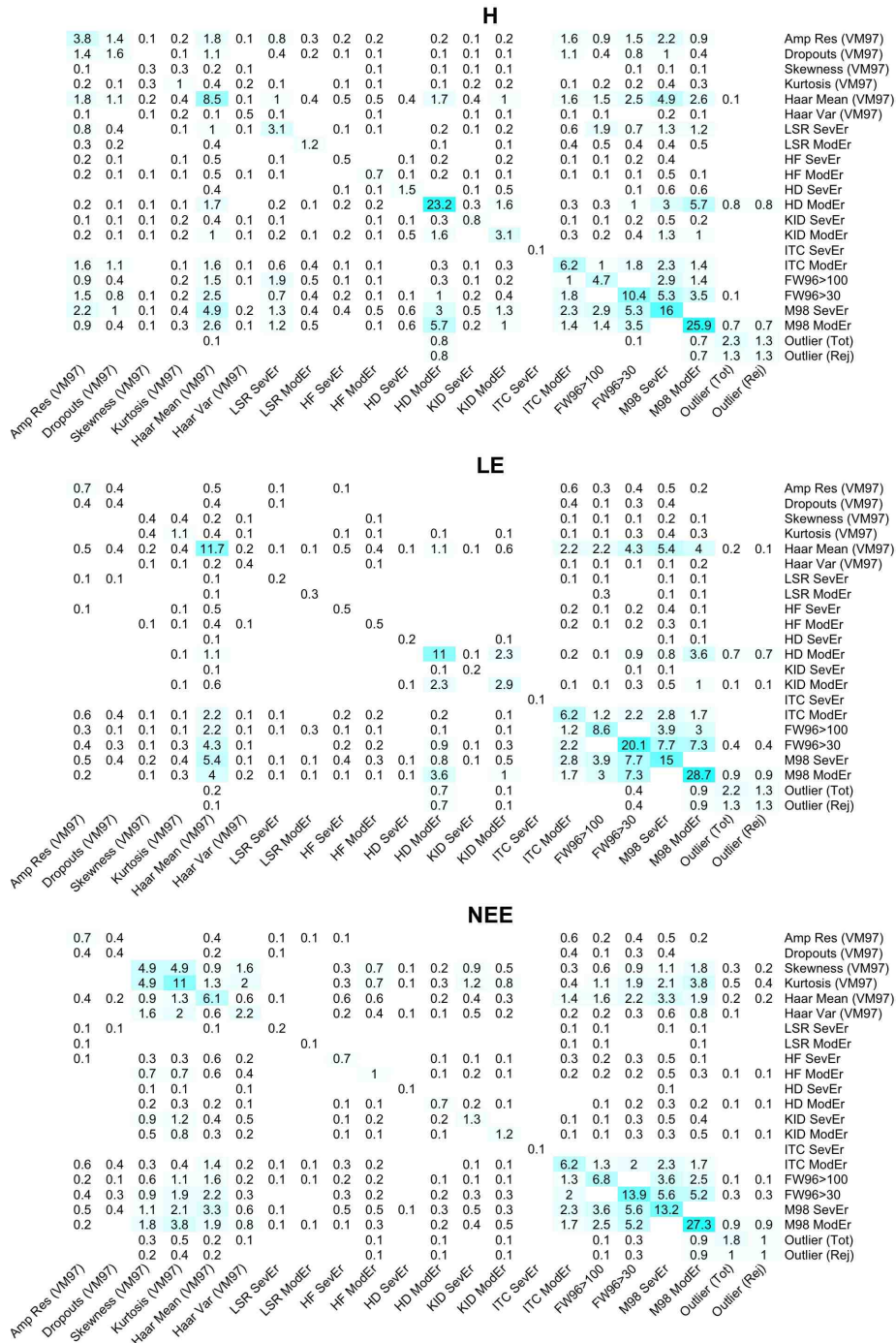
Figure D15. Percentages of H, LE and NEE fluxes at FI-Hyy site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
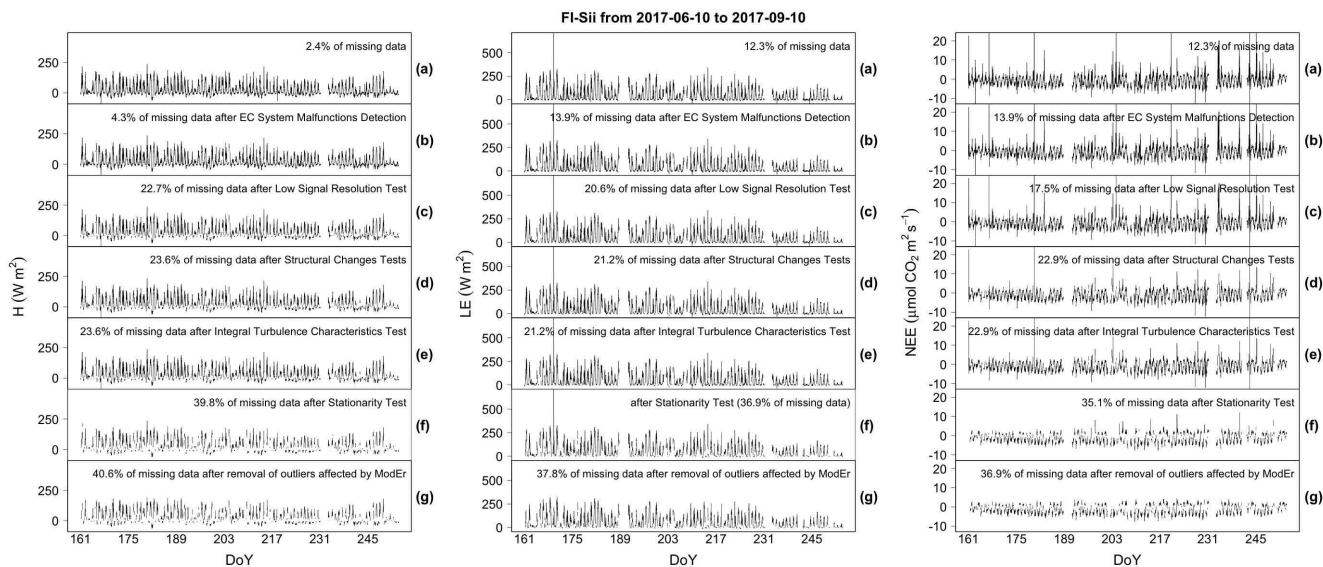
Figure D16. Sequential data cleaning procedure applied to H, LE and NEE fluxes at FI-Sii site.
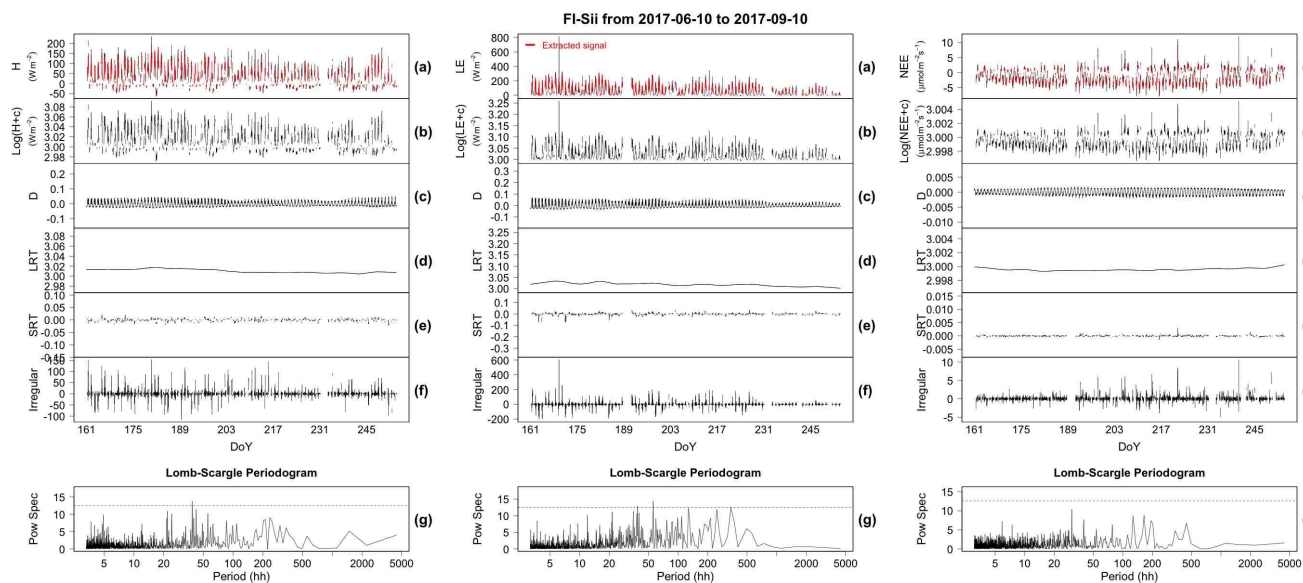
Figure D17. STL decomposition (panels a-f) of H, LE and NEE fluxes at FI-Sii site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
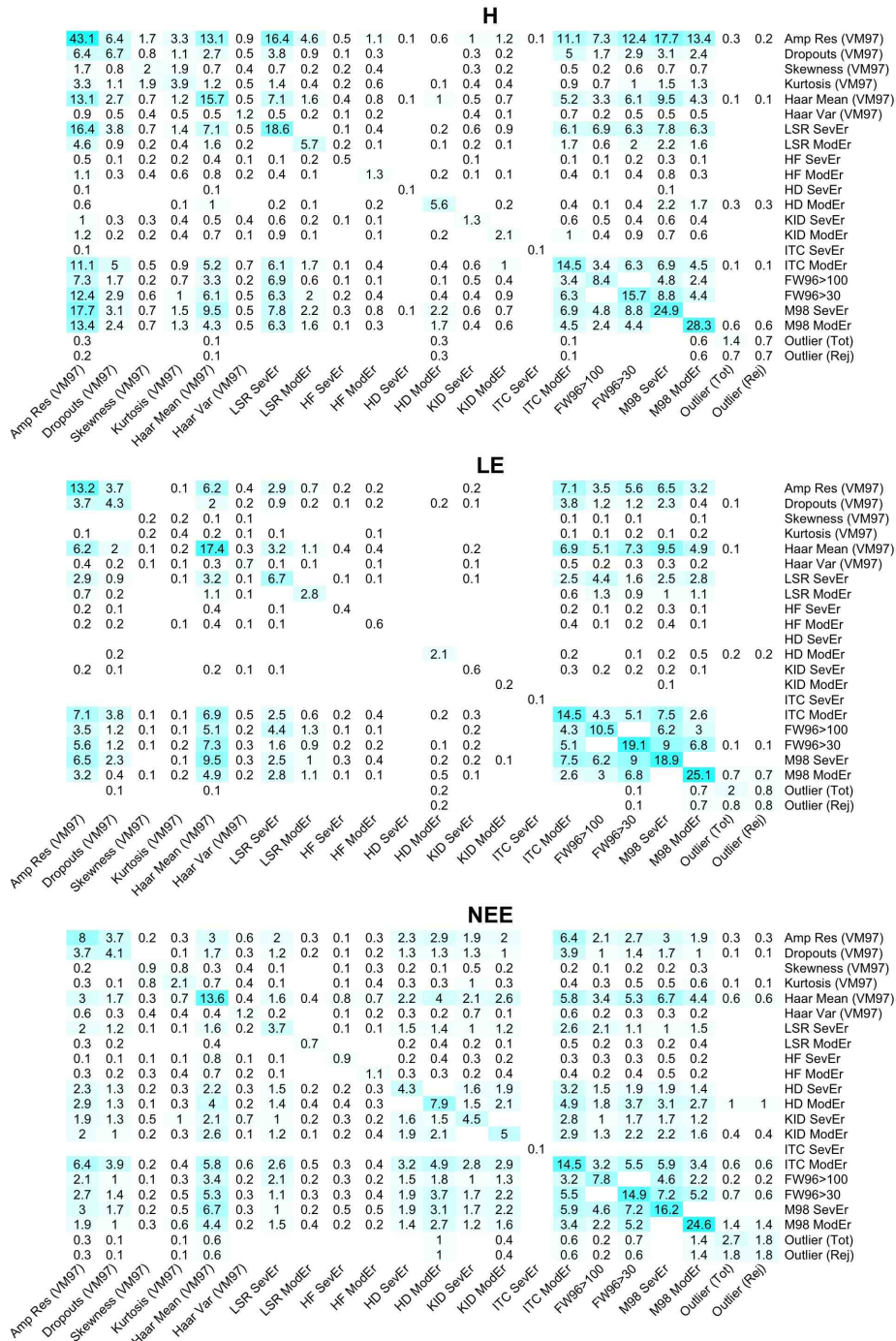
Figure D18. Percentages of H, LE and NEE fluxes at FI-Sii site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
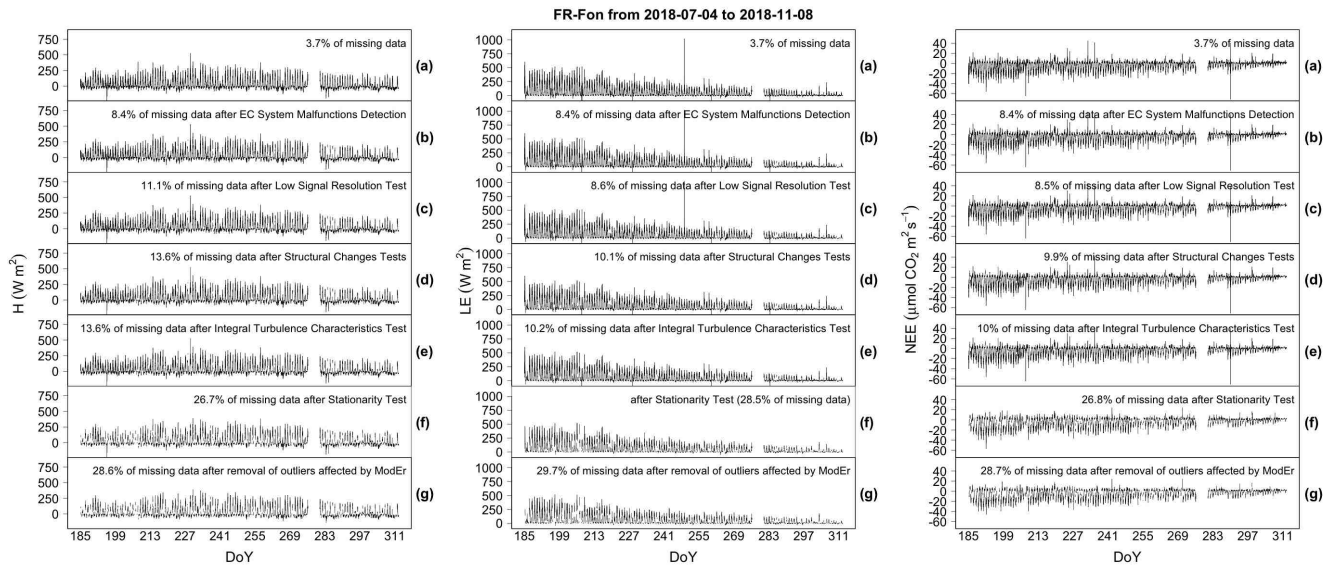
Figure D19. Sequential data cleaning procedure applied to H, LE and NEE fluxes at FR-Fon site.
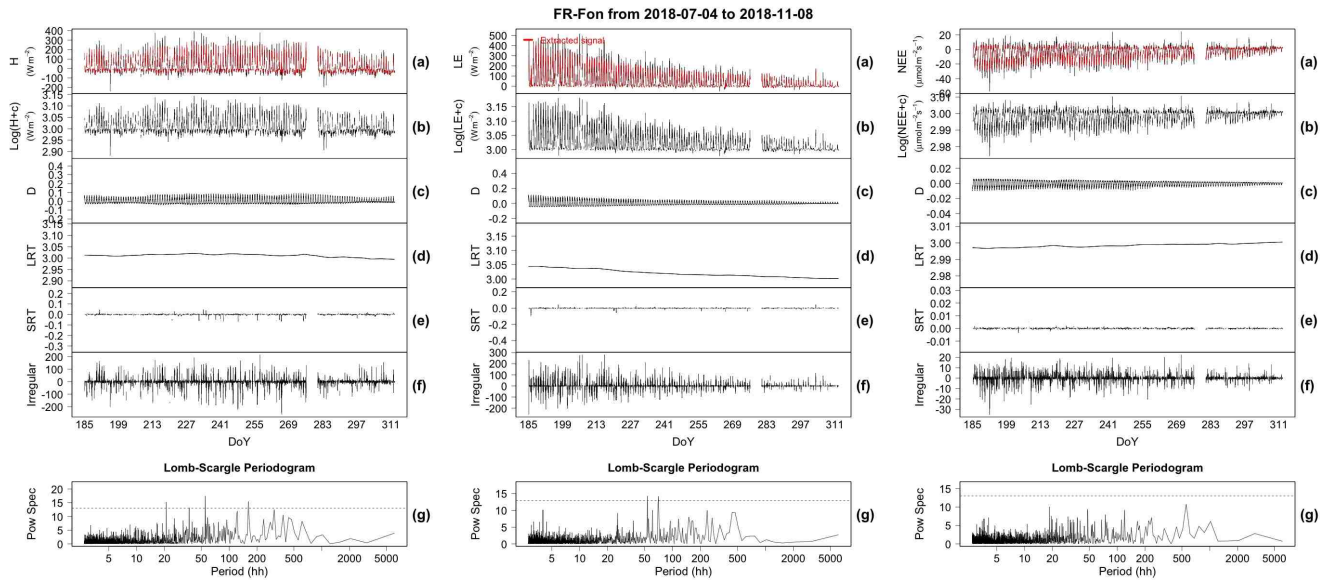
Figure D20. STL decomposition (panels a-f) of H, LE and NEE fluxes at FR-Fon site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.

Figure D21. Percentages of H, LE and NEE fluxes at FR-Fon site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
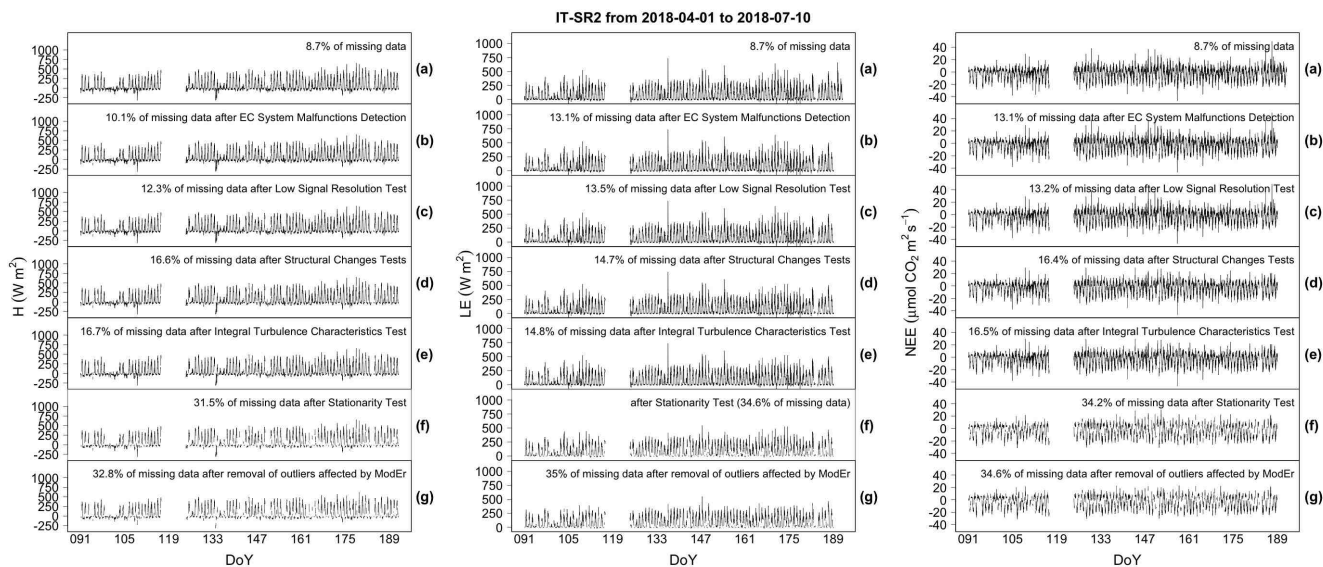
Figure D22. Sequential data cleaning procedure applied to H, LE and NEE fluxes at IT-SR2 site,
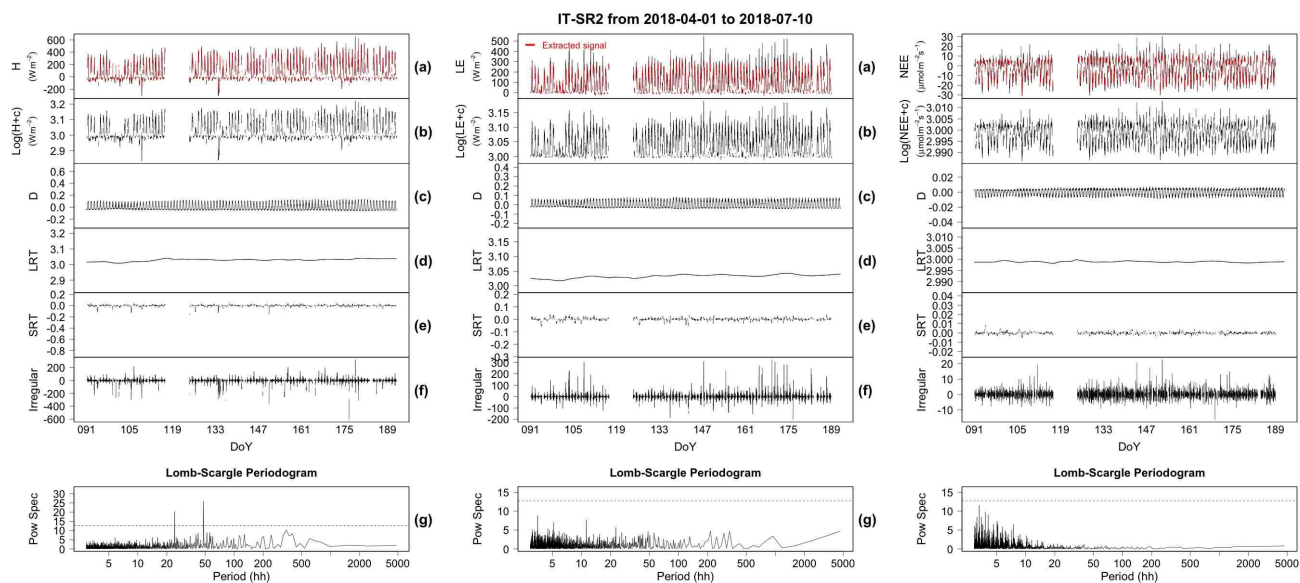
Figure D23. STL decomposition (panels a-f) of H, LE and NEE fluxes at IT-SR2 site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
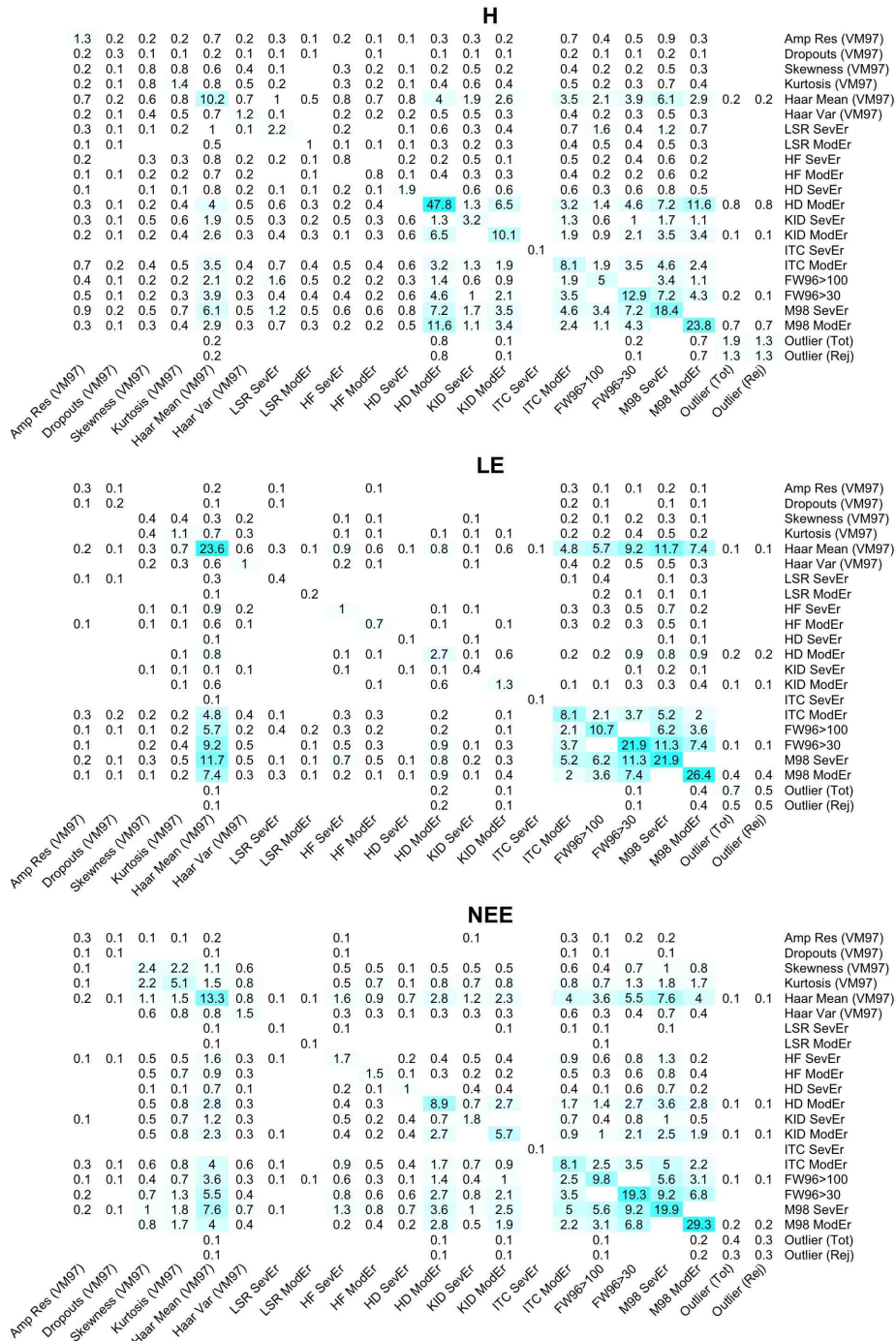
Figure D24. Percentages of H, LE and NEE fluxes at IT-SR2 site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
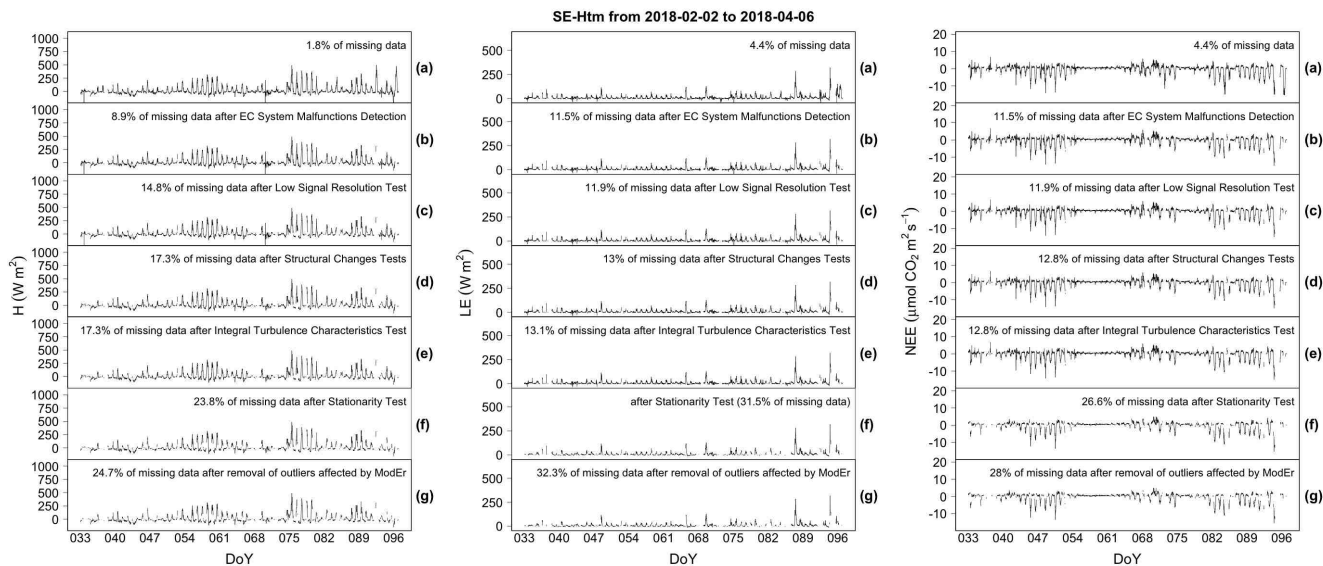
Figure D25. Sequential data cleaning procedure applied to H, LE and NEE fluxes at SE-Htm site.
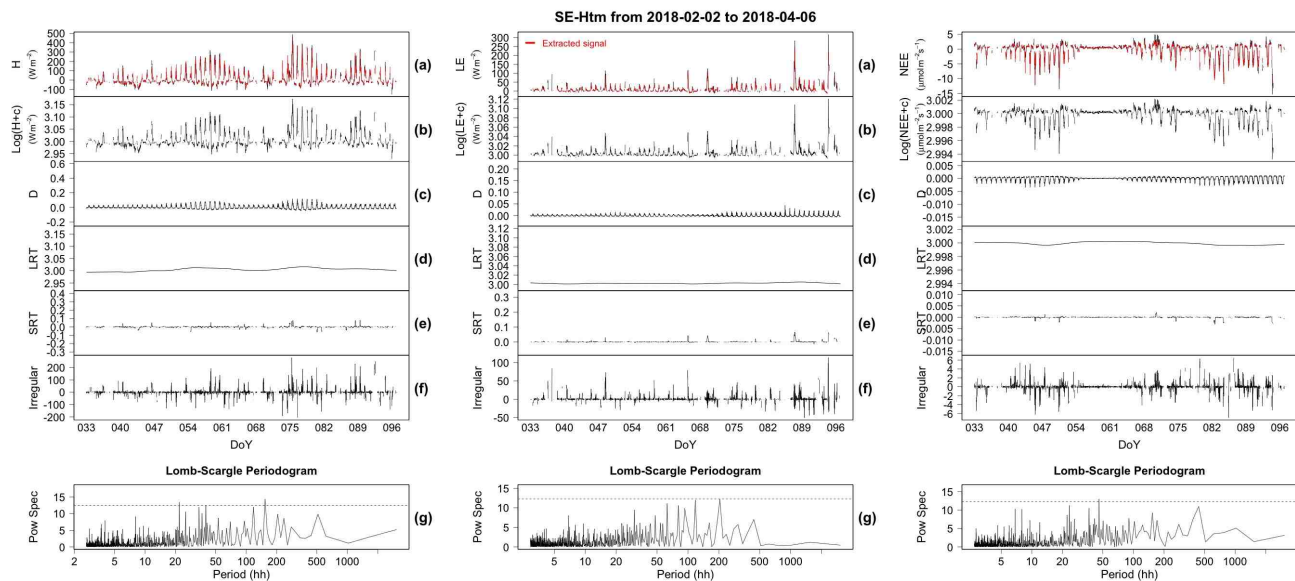
Figure D26. STL decomposition (panels a-f) of H, LE and NEE fluxes at SE-Htm site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
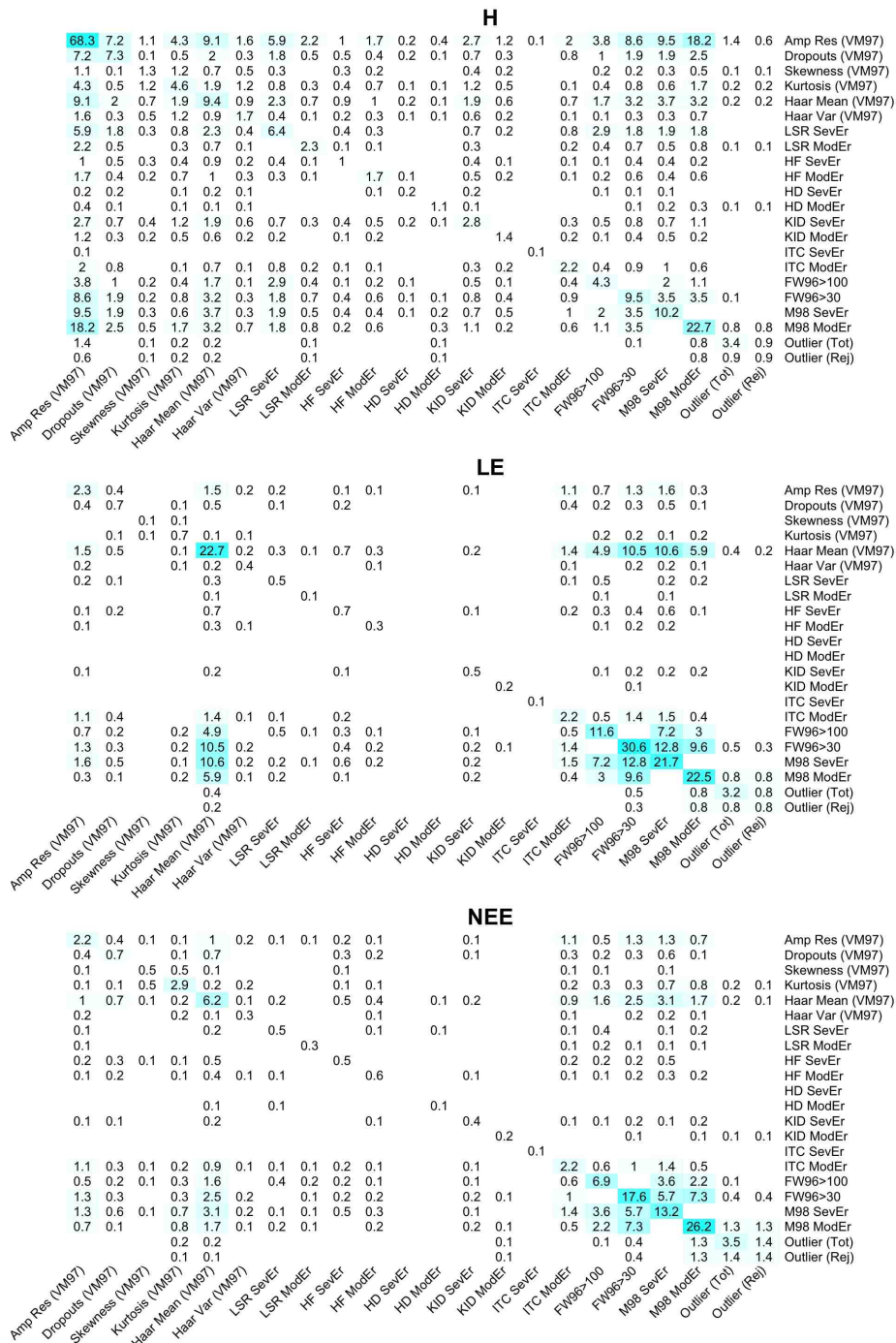
Figure D27. Percentages of H, LE and NEE fluxes at SE-Htm site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.
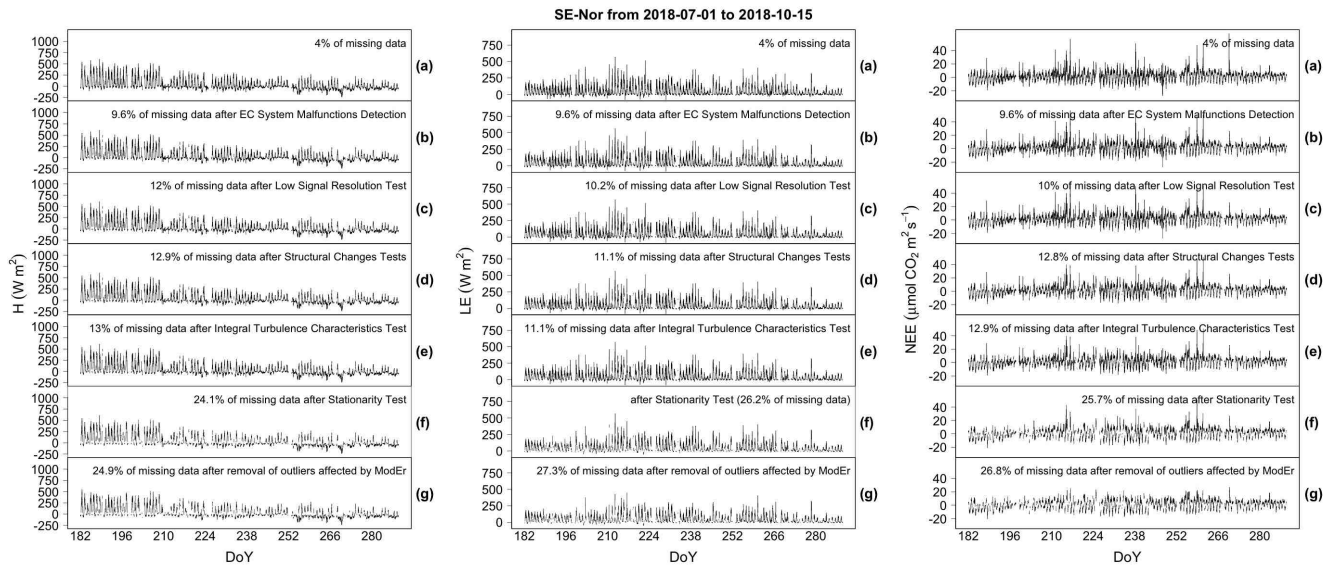
Figure D28. Sequential data cleaning procedure applied to H, LE and NEE fluxes at SE-Nor site.
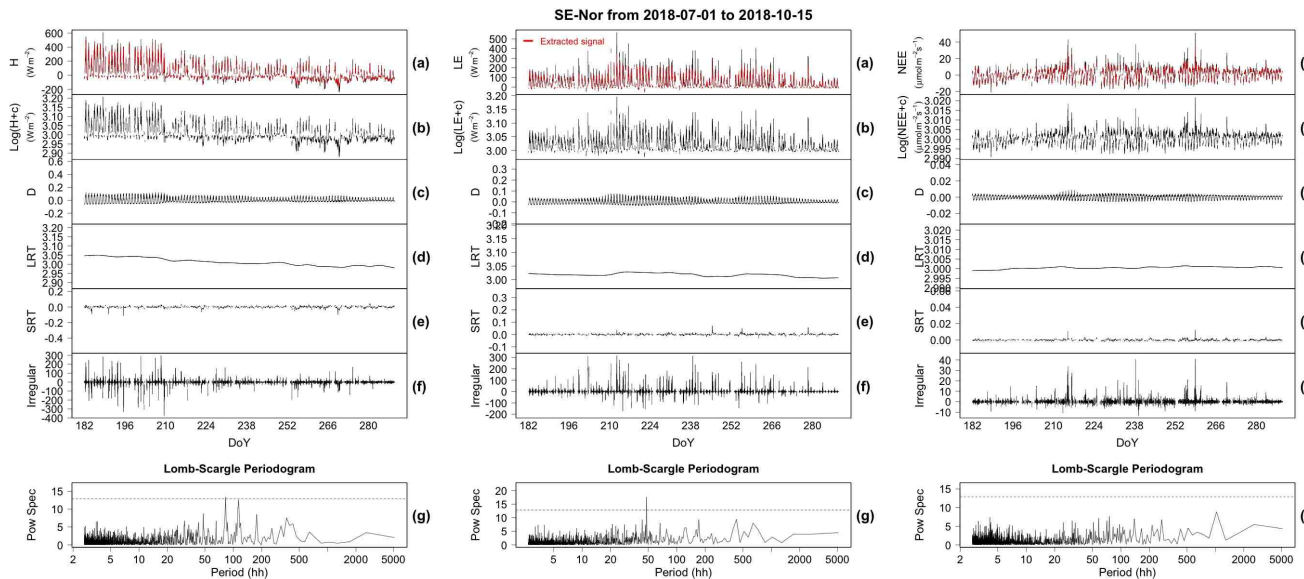
Figure D29. STL decomposition (panels a-f) of H, LE and NEE fluxes at SE-Nor site and power spectral density of the irregular component estimated by means of the Lomb-Scargle periodogram (panel g). Horizontal dashed lines represent the critical value at $\alpha = 0.01$ significance level.
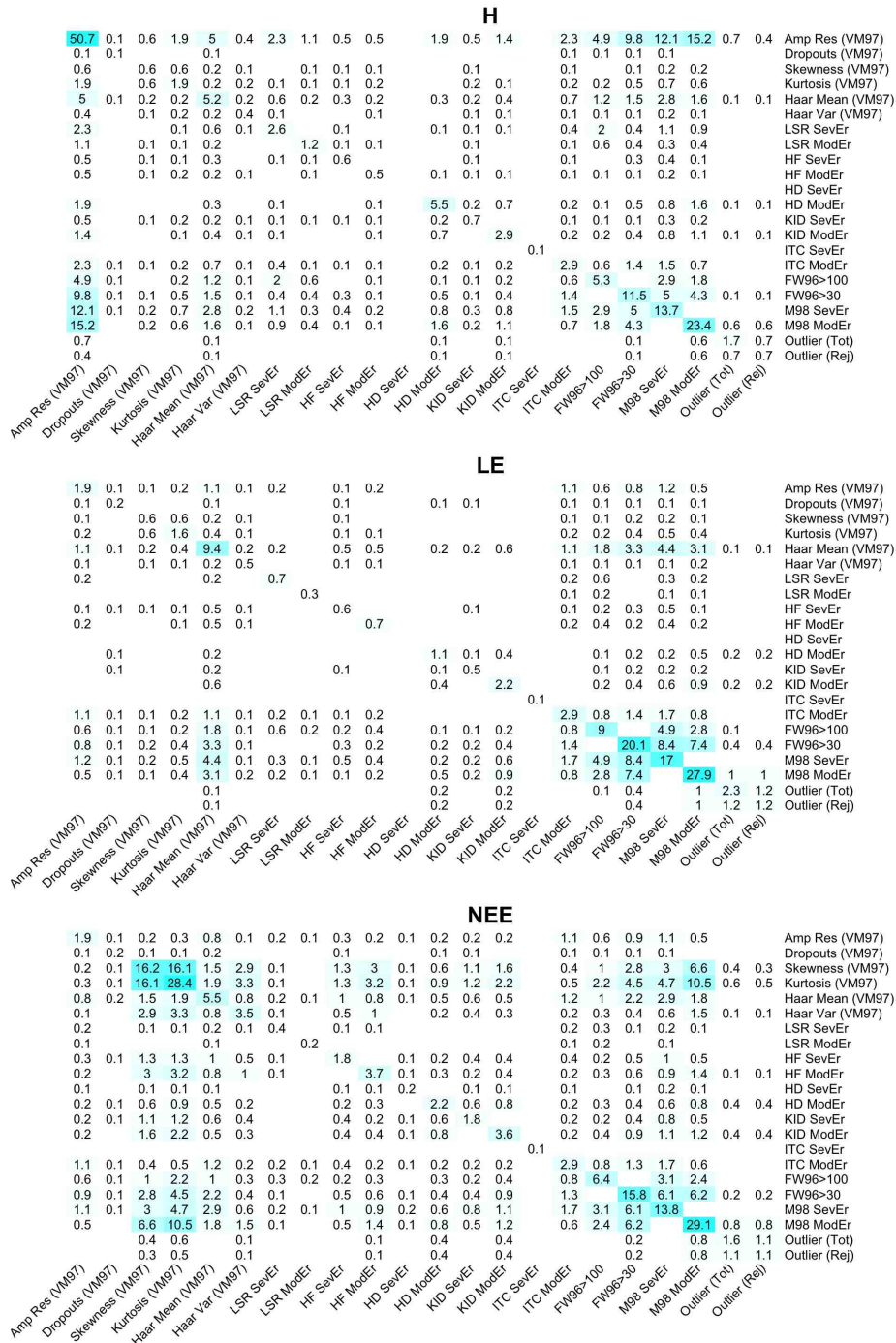
Figure D30. Percentages of H, LE and NEE fluxes at SE-Nor site affected by specific sources of systematic error according to several QC tests. Each value indicates the percentage of data receiving the status indicated in the ith row and the jth column (on diagonal, the percentage of data identified by each individual test). Cyan color highlights the highest values, while lower values are lighter. Empty cells indicate no data was identified (NoEr). The results of tests by VM97 refers to only *hard-flagged* data.

References

Brockwell, P. J. and Davis, R. A.: Introduction to Time Series and Forecasting, Springer New York, https://doi.org/10.1007/b97391, 2002.

Cryer, J. D. and Chan, K.-S.: Time Series Analysis, Springer New York, https://doi.org/10.1007/978-0-387-75959-3, 2008.

Delpierre, N., Berveiller, D., Granda, E., and Dufrêne, E.: Wood phenology, not carbon input, controls the interannual variability of wood growth in a temperate oak forest, New Phytol., 210, 459–470, https://doi.org/10.1111/nph.13771, 2015.

Haapanala, S., Rinne, J., Pystynen, K.-H., Hellén, H., and Hakola, H.: Measurements of hydrocarbon emissions from a boreal fen using the REA technique, Biogeosciences, 3, 103–112, https://doi.org/10.5194/bg-3-103-2006, 2006.

Hamilton, J. D.: Time Series Analysis, Princeton University Press, Princeton, NJ, 1994.

Lagergren, F., Eklundh, L., Grelle, A., Lundblad, M., Molder, M., Harry, L., and Lindroth, A.: Net primary production and light use efficiency in a mixed coniferous forest in Sweden, Plant Cell Environ., 28, 412–423, https://doi.org/10.1111/j.1365-3040.2004.01280.x, 2005.

LI-COR Biosciences: EddyPro 6.2.1: Help and User's Guide, LI-COR Biosciences, Lincoln, Nebraska USA, www.licor.com/EddyPro, 2019.

Lütkepohl, H.: New Introduction to Multiple Time Series Analysis, Springer Berlin Heidelberg, https://doi.org/10.1007/978-3-540-27752-1, 2005.

Matteucci, M., Gruening, C., Ballarin, I. G., Seufert, G., and Cescatti, A.: Components, drivers and temporal dynamics of ecosystem respiration in a Mediterranean pine forest, Soil Biol. Biochem., 88, 224–235, https://doi.org/10.1016/j.soilbio.2015.05.017, 2015.

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Rebmann, C., Schmid, H. P., Schmidt, M., and Steinbrecher, R.: A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements, Agr. Forest Meteorol., 169, 122–135, https://doi.org/10.1016/j.agrformet.2012.09.006, 2013.

Moureaux, C., Debacq, A., Bodson, B., Heinesch, B., and Aubinet, M.: Annual net ecosystem carbon exchange by a sugar beet crop, Agr. Forest Meteorol., 139, 25–39, https://doi.org/10.1016/j.agrformet.2006.05.009, 2006.

Plosser, C. I. and Schwert, G.: Estimation of a non-invertible moving average process, J Econometrics, 6, 199–224, https://doi.org/10.1016/0304-4076(77)90015-x, 1977.

Shumway, R. H. and Stoffer, D. S.: Time Series Analysis and Its Applications, Springer International Publishing, https://doi.org/10.1007/978-3-319-52452-8, 2017.

Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Maso, M., Kulmala, M., and Vesala, T.: Long-term measurements of surface fluxes above a Scots pine forest in Hyytiälä, southern Finland, 1996-2001, Boreal Environ. Res., 8, 287–302, 2003.

van Meeningen, Y., Wang, M., Karlsson, T., Seifert, A., Schurgers, G., Rinnan, R., and Holst, T.: Isoprenoid emission variation of Norway spruce across a European latitudinal transect, Atmos. Environ., 170, 45–57, https://doi.org/10.1016/j.atmosenv.2017.09.045, 2017.

Wollschläger, U., Attinger, S., Borchardt, D., Brauns, M., Cuntz, M., Dietrich, P., Fleckenstein, J. H., Friese, K., Friesen, J., Harpke, A., et al.: The Bode hydrological observatory: a platform for integrated, interdisciplinary hydro-ecological research within the TERENO Harz/Central German Lowland Observatory, Environmental Earth Sciences, 76, 29, https://doi.org/10.1007/s12665-016-6327-5, 2017.

Zielis, S., Etzold, S., Zweifel, R., Eugster, W., Haeni, M., and Buchmann, N.: NEP of a Swiss subalpine forest is significantly driven not only by current but also by previous year's weather, Biogeosciences, 11, 1627–1635, https://doi.org/10.5194/bg-11-1627-2014, 2014.