Supplement of Biogeosciences, 23, 477–495, 2026
https://doi.org/10.5194/bg-23-477-2026-supplement

Biogeosciences Open Access EGU

Supplement of

Spatiotemporal variability and environmental controls on aquatic methane emissions in an Arctic permafrost catchment

Michael W. Thayne et al.

Correspondence to: Michael W. Thayne (m_thayne@me.com)

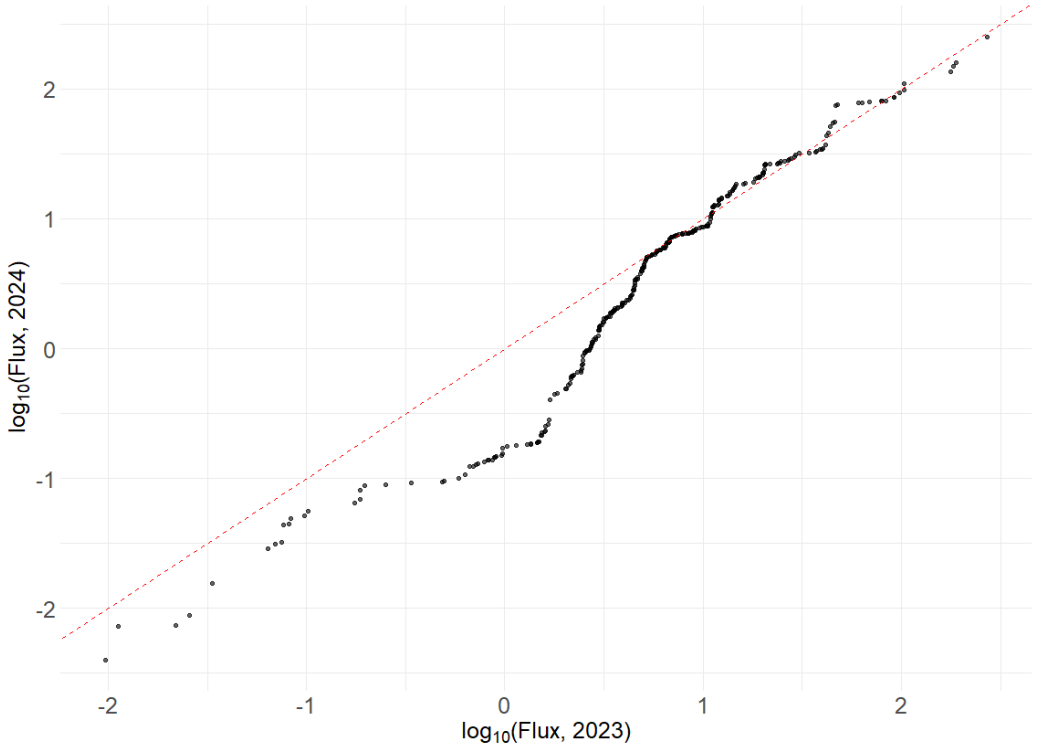

Figure S1 Quantile–quantile (Q–Q) plot comparing log_{10}-transformed diffusive CH_4 fluxes measured in 2023 (x-axis) and 2024 (y-axis). Each point represents a quantile-matched flux from the two years. The red dashed line indicates the 1:1 relationship. Fluxes align closely across most of the distribution, particularly in the mid-to-upper ranges, with minor divergence at the lower quantiles. These results support the conclusion that chamber construction differences between years did not substantially bias flux estimates.

Ebullition Detection

The algorithm has several parameters which are used to set the “ebullition” detection settings. We use quotations for ebullition as not all non-linear concentration increases are a result of gas bubbles entering into the chamber space, which is generally represented by abrupt step change in gas concentrations. More broadly speaking, the algorithm uses these parameters to classify non-linear concentration increases using the results of BRT. Before running the algorithm, the user will need to set a vector of ‘learning_rates’, ‘tree_complexity’, and ‘bag_fraction’ for BRT training, where default values for ‘tree_complexity’ and ‘bag_fraction’ are 5 and 1, respectively. Bag fraction is set to 1, so the model is deterministic. Learning rates for the work presented here were 0.001, 0.002, 0.003, 0.004, and 0.005. To avoid false detection and to compensate for issues related to problems that are often encountered in the field such as leakage, or the chamber fan not functioning, the user can set the parameter ‘window_size_concentration_smoother’ which is used to set a moving average window which smooths out the concentration time series before fitting the BRT. BRT training and parameter optimization were conducted using 10-fold cross-validation (Elith et al. 2008). Each BRT was fit with time of chamber measurement, chamber volume and area, air temperature and air pressure, and the gas temperature and pressure as measured by the gas analyzer. The user can define any vector of ‘predictors’ they would like, including other environmental variables that may help in classifying non-linear concentration increases, as this is the sole purpose of this algorithmic function. Final model selection was based on the combination of model hyper parameters; number of trees (i.e. ≥ 1000), tree complexity, and learning rate which results in the lowest mean deviance standard error. Residuals from the selected model were calculated by subtracting the model predictions from the observed concentrations, thus highlighting discrepancies between observed and predicted values.

To facilitate detection of data points which significantly deviate from one another, residuals from the selected model are smoothed using a moving average window with a default width of 5 data points. The moving average window is user-defined and can be changed by setting 'window_size_residual_smoother'. The standard deviation of the smoothed residuals is calculated and then used to create a dynamic threshold by multiplying it with a 'dynamic_multiplier', which has a default value of 2. The dynamic threshold is used to identify significant deviations in the smoothed residuals. The dynamic multiplier can be thought of as a handle on a water faucet; if the user increases the value, the algorithm becomes more restrictive on what data points will be classified as significant deviations. The algorithm then computes the absolute differences between consecutive concentration values. Using the stored absolute differences, the algorithm checks to see if the concentration differences exceed a default quantile value of 0.95 (Hoffman et al. 2017). Consequently, the algorithm by default considers the top 5% of the greatest concentration differences as significant deviations from the linearly increasing concentrations. However, the user is able to define the quantile by setting the parameter 'quantile_threshold'. Thus, a gas concentration data point is classified as significantly deviating from linearly increasing concentrations when either of these two conditions are met: (1) the absolute value of the smoothed residuals exceeds the dynamic threshold and/or (2) the concentration difference between two consecutive points exceeds the quantile threshold. Lastly, the algorithm generates a range diagnostic plots (Figures S2-S6), for example one showing which points during the chamber measurement met one or both of the conditions (Figure S2 and S5). The plot is saved for review in the user-defined 'save_directory'. The algorithm then initiates code to handle the classification of data sequences, or non-linear events, which are comprised of the points classified during the process described above, and may potentially indicate ebullition events.

The data sequence classification code first initializes variables to track the start and end of possible non-linear concentration increases and creates a buffer index to store up to 5 non-event points, or points which briefly interrupt an ongoing non-linear event. The algorithm loops through each data point of the chamber measurement, marking the start and end of events while resetting the buffer index each time a new event point is detected. When non-event points are encountered during an ongoing event sequence, they are added to the buffer index, where if the buffer reaches capacity, the event sequence length is checked against a user-defined 'min_ebullition_sequence' parameter, which has a default of value of 8 seconds. Event sequences are classified based on this length check, and ongoing events at the end of the dataset are handled similarly to account for edge effects. This process ensures ongoing events are not broken up by data points, and only sequences meeting the minimum length can be classified as events. The algorithm then stores each sequence in two data frames, one for diffusive (i.e. linear concentration increases) and another for ebullition events (i.e. non-linear concentration increases). Each sequence is given a numerical i.d. based on when it was encountered during the chamber measurement. For example, if CH_4 concentration increases linearly for the 240 seconds of the chamber measurement and then broken up by an ebullition event for 40 seconds, the first sequence will be labelled diffusive-1 and the second ebullition-2 (Figure S2). Thus, numerical values are given based in temporal order. At this stage the user is prompted by the algorithm to "review the combined plot" saved in the 'save_directory' to ensure the algorithm has accurately classified the sequences (Figure S3).

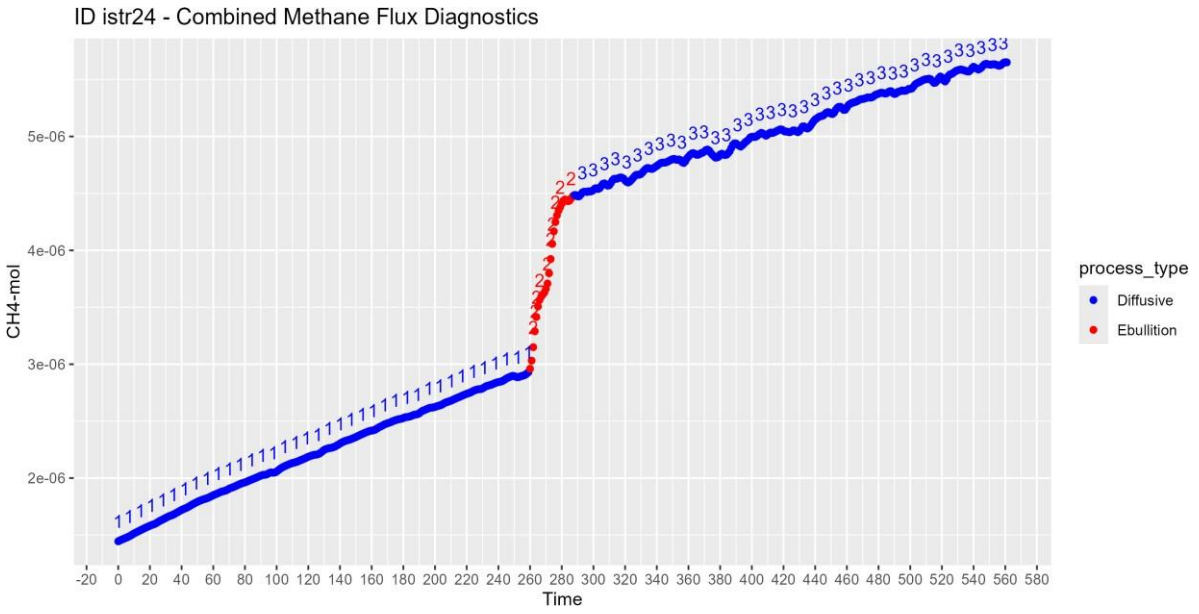

Figure S2. Diagnostic plot of for chamber ID istr24 over a period of 561 seconds. Time of the chamber measurement is given on the x-axis, while CH_4 concentration is given on the y-axis. The blue dots and corresponding numerical id represent those concentration sequences which the algorithm has deemed as linear diffusion, while the red dots represent concentration sequences that may represent ebullition. Overall, we see a linear increase in concentration which is disrupted by a step change in concentration increase. In the end the algorithm handles each concentration sequence independently from one another and calculates flux accordingly.

The user is asked if the classified sequences are correct, where the user can respond accordingly; yes, no, reclassify, or skip. If the user responds yes, each classified sequence will be

independently handled, flux calculations made, and results stored following the methods previously described (Figure S3 or S4 when linear).

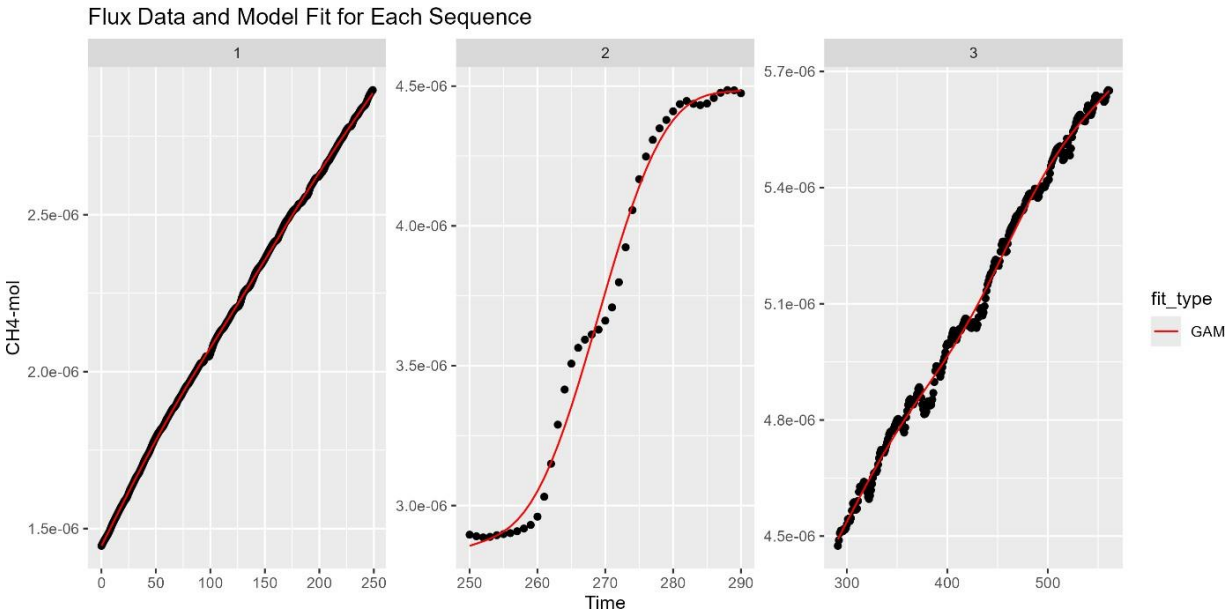

Figure S3. Diagnostic plot showing the flux data included for each sequence (black dots) and the predicted concentration given by the GAM (red line) for ID istr24 shown in figure 1. Time is given on the x-axis for each sequence, while CH_4 concentration is given on the y-axis. Overall, the figure shows that the GAM has robustly predicted the concentration increases from each sequence.

If the user responds no, they will be asked if they would like to manually enter sequences. If they respond yes they will enter a time range, whether the sequence is diffusive (D) or ebullition (E), and a numerical label is given to the sequence (e.g. 0-250,D,1; 251-290,E,2; 291-600,D,3) (Figure S2). The user is then prompted to ensure the sequences have been classified and labelled correctly and if so, fluxes are calculated for each sequence and results are stored in the results directory. When the user responds reclassify, they will be given the opportunity to reclassify a sequence from diffusive to ebullition, or vice versa. If a sequence is reclassified the user is prompted to review the updated plot and confirm sequences are correct before moving forward with flux calculations (Figure S6). All flux calculations are stored alongside the measurement i.d., whether it is diffusive or ebullitive flux, and the R^2 value and the root mean squared error for the GAM fits.

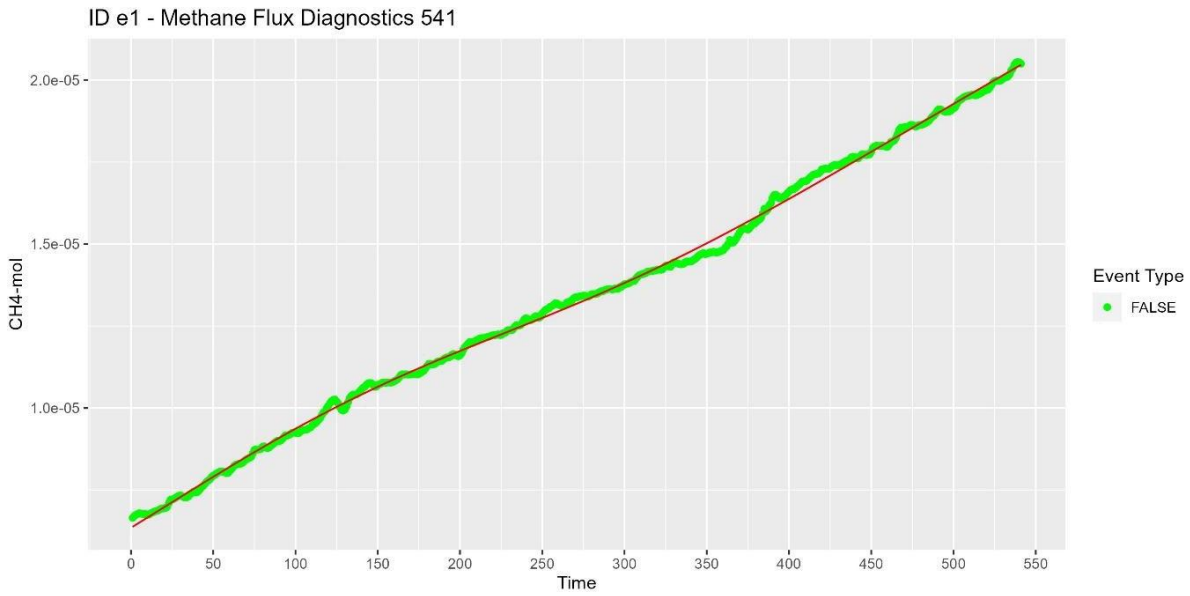

Figure S4. Diagnostic plot showing the outcome of a fitted GAM where the chamber measurement met the condition to be considered a linear concentration increase and sufficient for further flux calculations. Time steps of the chamber measurement are given on the x-axis and CH_4 concentration is given in moles on the y-axis. The red line shows the model prediction, while the green dots labeled false in the legend, are those that do not meet the requirements to be considered significant deviations from the linear concentration increase. The number 541 in the main figure title represents the maximum time interval recorded during the chamber measurement, in this case 541 s.

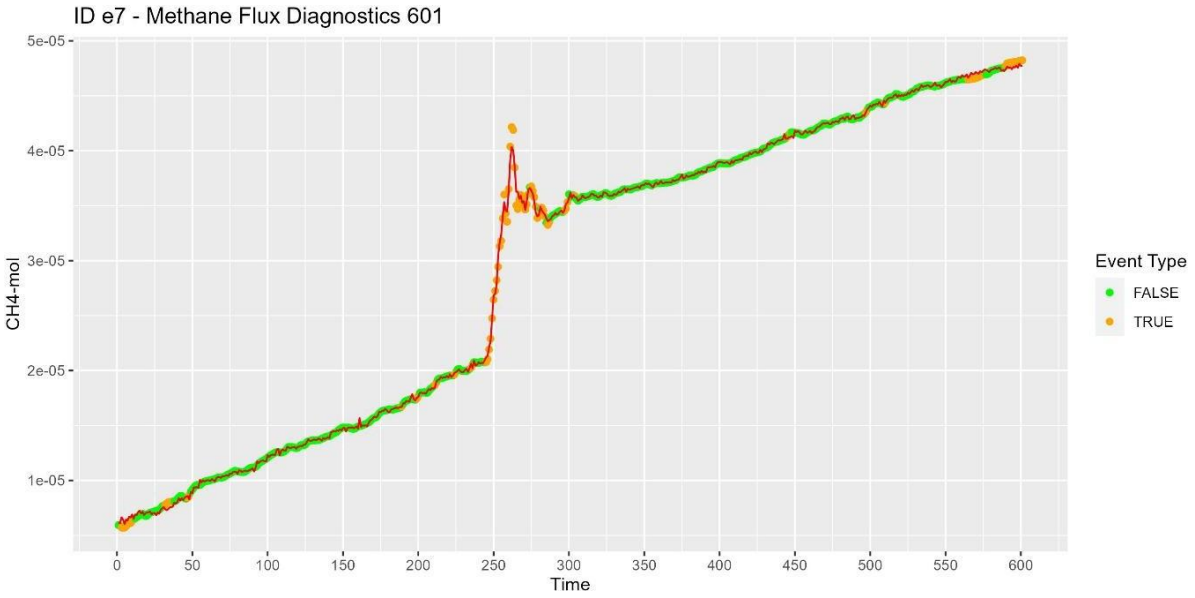

Figure S5. Diagnostic plot showing the outcome of a trained BRT used to detect data points which meet the conditions for a data point which significantly deviates from what is expected. Time steps of the chamber measurement are given on the x-axis and CH_4 concentration is given in moles on the y-axis. The red line shows the model prediction, while the orange dots represent those points that have been considered by the algorithm to have truly met one or both of the two conditional criteria, while green represent data points which have not met the conditional criteria (see legend). The number 601 in the main figure title represents the maximum time interval recorded during the chamber measurement, in this case 601s.

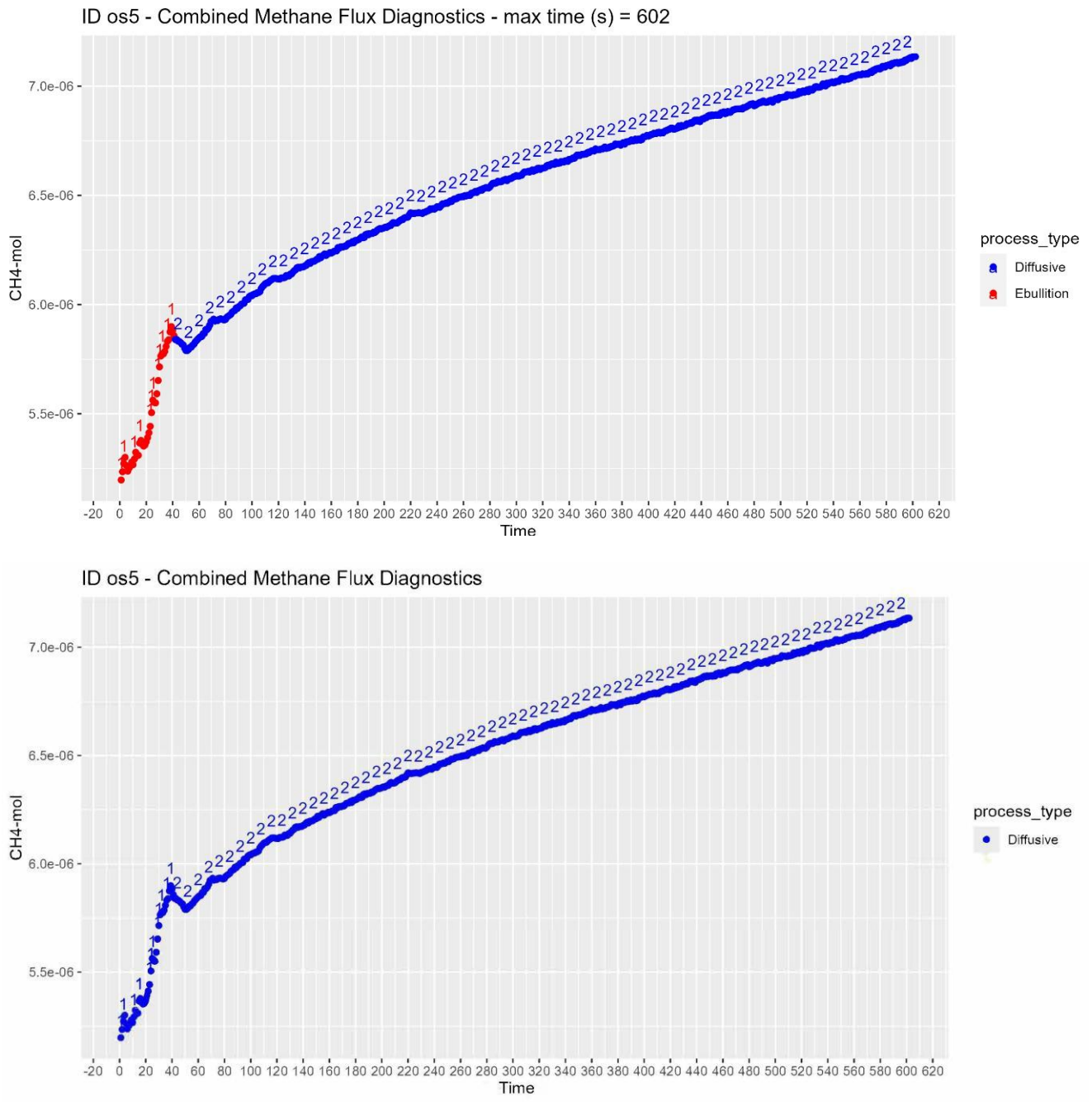

Figure S6. Diagnostic plots showing how the reclassify function works if one chooses to reclassify, for example, ebullition to diffusive as seen in the figures. In the top figure we see that the initial sequence was considered ebullition (labeled 1), as it is a non-linear rise in CH_4 concentrations. However, the concentration rise is not reflective of an ebullition event, but rather a rapid rise in concentrations to a steady linear rise in concentrations, reflective of a diffusive flux. Thus, in this case the user can reclassify the initial rise from ebullition to diffusive.

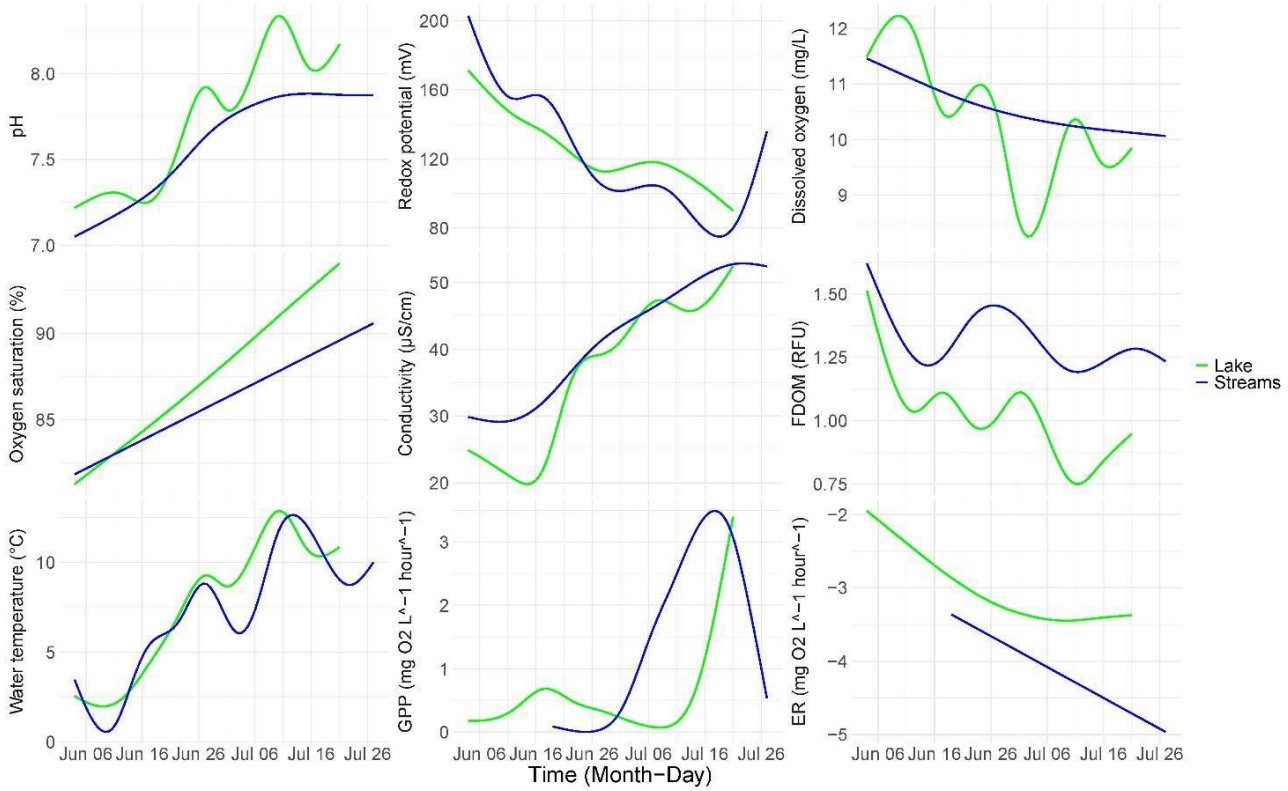

Figure S7. Illustrates the trend of water column conditions in the lake (green lines) and in the streams (blue lines).

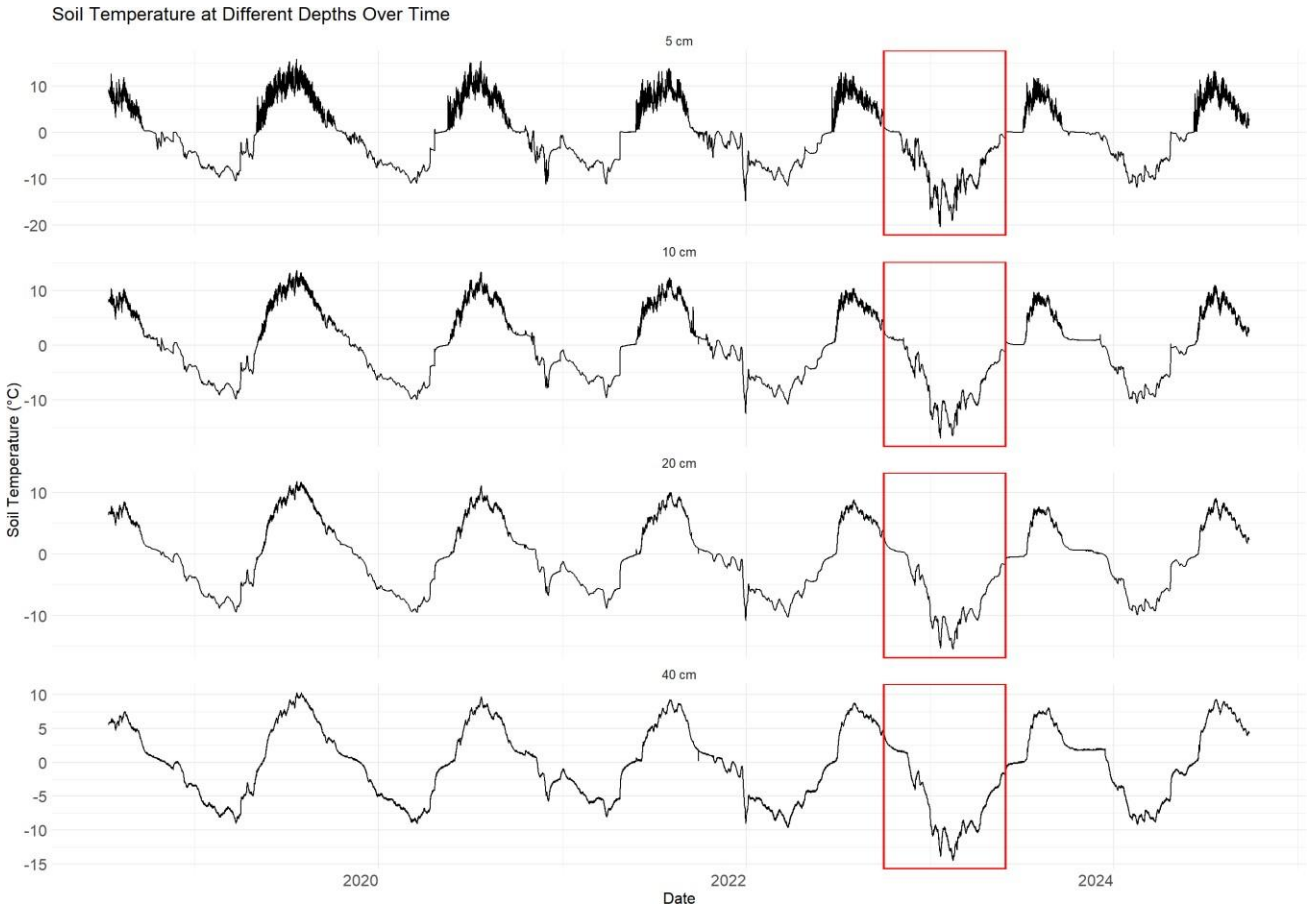

Figure S8. Illustrates 6-year time series (x axis) of soil temperature (y axis) measured at 5, 10, 20, and 40 cm. Soil Temperatures at each depth rarely reach below -10 °C, but in 2023, soil temperatures dropped to between 15-20 °C (red box). The relative importance of soil temperatures suggest they play a significant role in regulating CH_4 fluxes from the catchment.

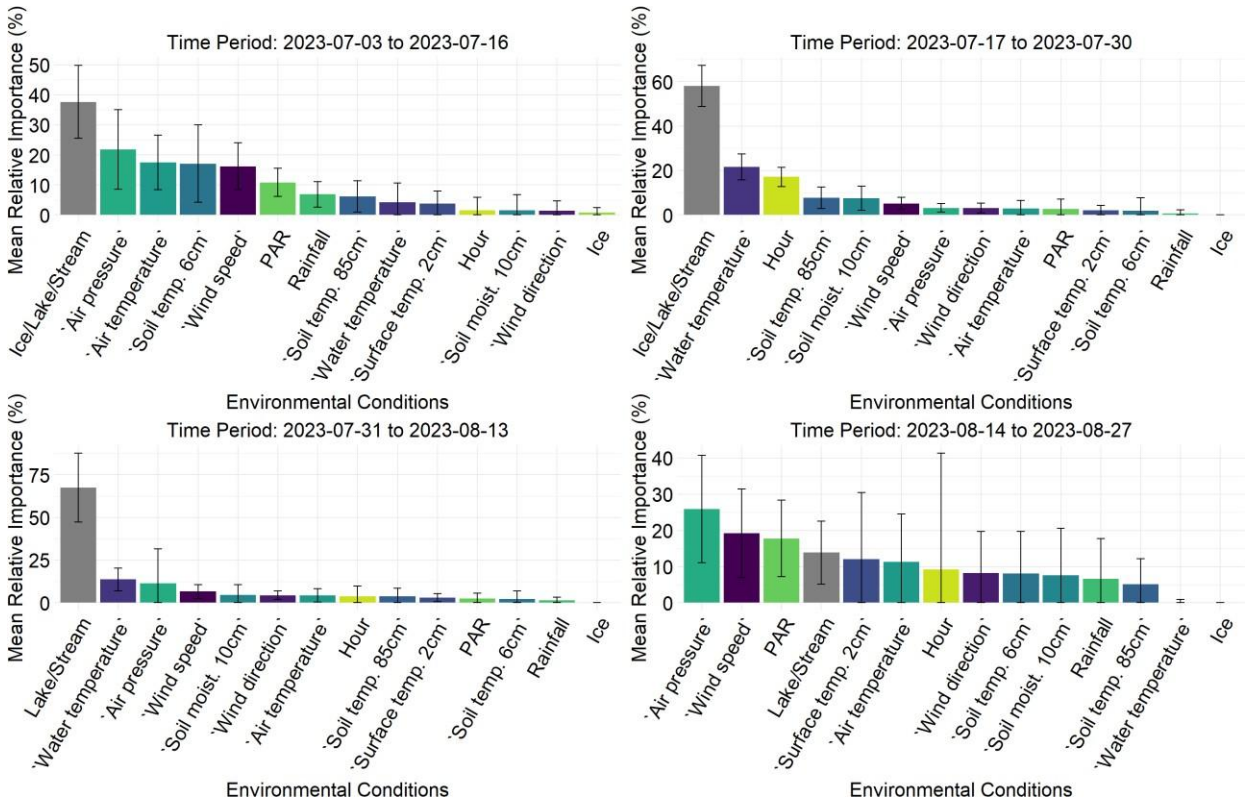

Figure S9. The figure illustrates the relative importance of environmental conditions predicting diffusive CH_4 fluxes using bar-plots and standard error bars. Each predictor variable is on the x-axis, while it's percent importance for its inclusion in a fitted BRT is given on the y-axis. Each bar color represents a specific environmental condition. Overall, we see that without water column characteristics that fluxes were strongly dependent on whether we were measuring in the lake versus the streams. While we expected water temperature to play a primary role in regulating fluxes from the catchment, we see in the figure that local climate and soil conditions influencing the catchment generally played a greater role in regulating fluxes from the water bodies.

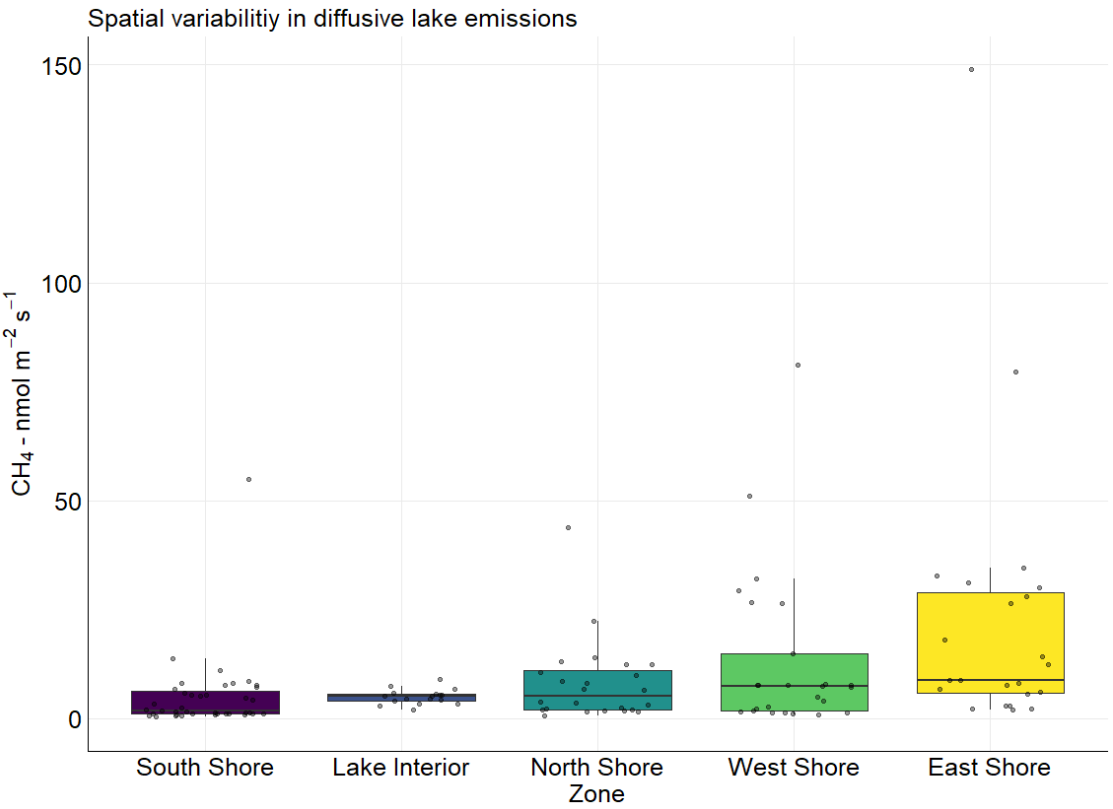

Figure S10. Boxplot of diffusive CH_4 fluxes (y-axis) across different lake zones at Sanningasup Tasia. Zones are classified as East Shore, West Shore, North Shore, South Shore, and Lake Interior (x-axis). Each point represents an individual chamber measurement. The figure highlights consistent spatial variability in emissions, with elevated fluxes generally observed near shorelines.

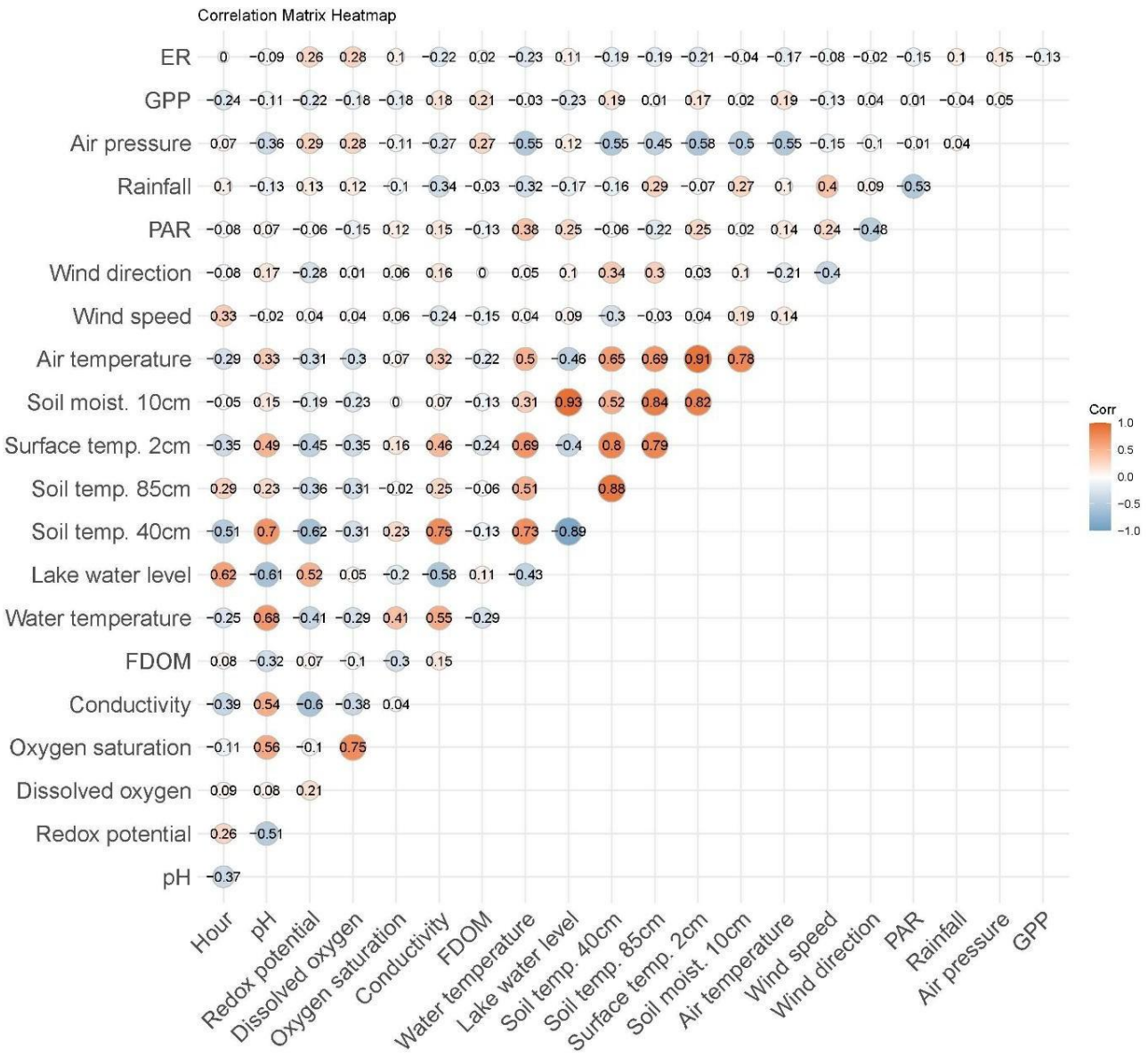

Figure S11. Illustrates a Pearson's correlation matrix showing the relationships among local meteorology, soil characteristics, and water parameters, with warmer orange-to-red cells indicating positive correlations and cooler blue cells indicating negative correlations.