Supplement Information
Multidecadal trends in CO_2 evasion and aquatic metabolism in a large temperate river

An Truong Nguyen[1*], Gwenaël Abril[2,3], Jacob S. Diamond[1,4], Raphaël Lamouroux[5], Cécile Martinet[6], Florentina Moatar[1]

[1] INRAE, UR RiverLy, 5 Rue de la Doua, 69100 Villeurbanne, France

[2] Laboratoire de Biologie des Organismes et Ecosystèmes Aquatiques (BOREA), UMR 8067, Muséum National d'Histoire Naturelle, CNRS, IRD, SU, UCN, UA, Paris, France.

[3] Programa de Geoquímica, Universidade Federal Fluminense, Niterói, Rio de Janeiro, Brazil

[4] Dipartimento di Scienze Ambientali, Informatica e Statistica, University of Venice Ca' Foscari, Via Torino 155, 30172 Venezia Mestre, Italy

[5] EDF – Recherche et Développement, Laboratoire National d'Hydraulique et Environnement, Chatou, France

[6] EDF – Division Technique Générale, Electricité de France, Grenoble, France

[*] Corresponding author: truongan9393@yahoo.com, truong-an.nguyen@inrae.fr

1. Supplementary method

S1. Dataset

The combination of datasets from the continuous monitoring and grab sampling program is used to obtain long-term high-frequency datasets, including hourly temperature, conductivity, alkalinity, and dissolved oxygen in 1990-2021.

Continuous monitoring program

The EDF measurement system is a floating platform with a temperature sensor and sensors for pH (range 0–14 pH unit), DO (range 0–20 mg L^{-1}), and conductivity (range 0–1000 µS cm^{-1}) (Campbell l ®). The surface water at 20 cm depth is pumped (ca. 0.5 L s^{-1}) through the system and measurements are recorded every 5 seconds, with average values saved every hour. It should be noted that data was collected both upstream and downstream at each power plant, with the upstream station located at the entrance of the dam and the downstream station located approximately 2-5km downstream of the dam. The data used for data analysis in this study was upstream station because of its data completeness. Prior to 2008, estimated uncertainties from membrane sensors were ±0.3°C, ±0.3 pH units, ±8% mg O_2 L^{-1}, ±5% µS

cm^{-1} membrane sensors (Moatar et al., 2001). After 2008, new optical sensors have uncertainties of ± 0.1 pH units, 3% mg O_2 L^{-1}.

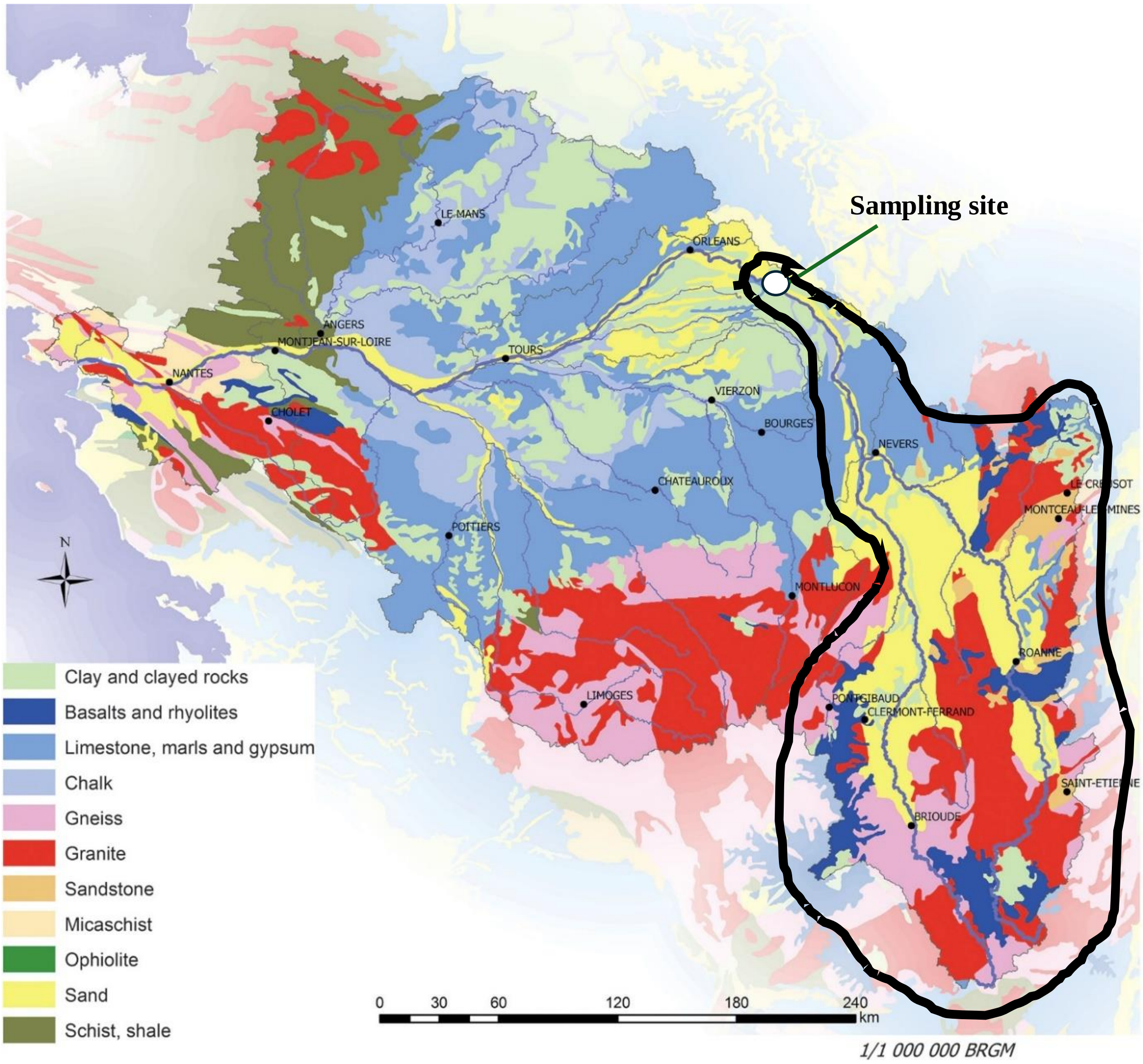

Figure S1. The sampling site and the lithology in the Loire basin (Moatar et al., 2022).

Grab sampling monitoring program

Grab sampling data was collected by EDF and Loire-Brittany Water Agency (AELB), including pH, conductivity, and alkalinity from 1990-2021, with frequency ranging from daily to monthly. Grab sampling data exists only in the upstream of the nuclear power plant. While AELB provided data for the period of 1990 to 2003 for these parameters, EDF supplied data from 2007 to 2021, so missing grab sampling data in the period 2004-2006. However, the primary objective of utilizing grab sampling data is to determine the correlation between total alkalinity and conductivity. This allows for the estimation of daily alkalinity based on the mean daily conductivity (calculated from the hourly dataset) for the period of 1990-2021 (Figure S2).

To reconstruct daily alkalinity from conductivity, we employed the *IterativeImputer* function with *BayesianRidge* estimator (i.e., regularized linear regression) by using *scikit-learn*[1], a Python package (Pedregosa et al., 2011). The *BayesianRidge* estimator filled the missing daily alkalinity by iteratively modeling the linear relationship between daily conductivity and available alkalinity data, while regularization accounts for potential changes in their relationship over 32 years. This process begins by estimating data for the period with fewest missing data, then continues iteratively until the imputed values converge, meaning subsequent iterations produce minimal changes in the estimates. This iterative process allows the imputer to adapt to underlying trends and shifts in the data. Besides, to verify the stability of the relationship between alkalinity and conductivity, we performed linear regressions on 4-year period of 32 years dataset which revealed quite similar slopes across all periods except 1990-1993 (Figure S2).

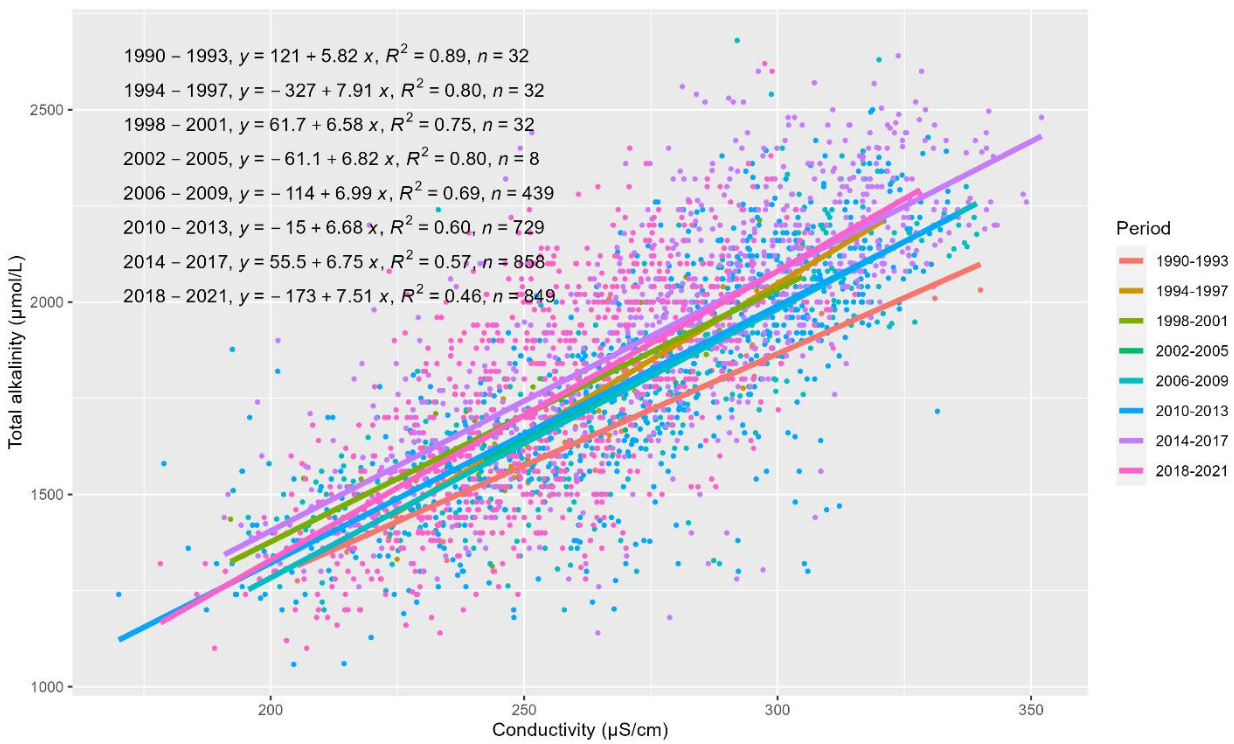

Figure S2. Relationship between measured total alkalinity and continuous measured conductivity in Loire River.

We estimated daily depth (m) using a local rating curve ($Depth\ (m) = 0.0716 \times \sqrt{Q(m^3s^{-1})} + 0.6171$), which was established based on modeled depth from a 1D hydraulic model for the Loire River (Camenen et al., 2016).

[1] https://scikit-learn.org/stable/modules/generated/sklearn.impute.IterativeImputer.html

S2. Data cleaning procedure

Although the sensors for pH and DO measurements were periodically calibrated by EDF, the dataset exhibited a notable number of anomalous values prior to 2008, which prompted the implementation of comprehensive data control procedures. These procedures addressed sensor drift and outlier removal which were proposed by (Jones et al., 2022; Moatar et al., 2001). Data cleaning was conducted for the hourly pH and dissolved oxygen data in this study, while the daily conductivity was carried out by EDF which based on visual inspection. Hourly temperature and alkalinity data from grab sampling were only checked through a range check to eliminate unrealistic values, but there was minimal significant removal of this data.

The following steps carried out the data cleaning and correction for hourly pH. Firstly, performing the rules-based anomaly detection and correction as a first pass at quality control, including range check (pH ranges from 6 to 10 in Loire River), data persistence check (pH relatively constant in few days), significant change check (jump or drop within few hours), calibration and drift event detection check. This step was performed automatically with the support of *pyhydroqc,* a python package for automating aquatic sensor data processing (Jones et al., 2022). Secondly, error detection was manually inspected by comparing values between upstream and downstream, together with daily discharge. This step used the interactive plot with the support of *plotly* package to check the errors which were identified in previous steps. The use of daily discharge was to eliminate false detection of abnormal data, especially in the case of high discharge where there are often sudden changes in pH and conductivity. There was 10.6% data (about 3 years of data) was assessed as anomalous and was discarded. Finally, missing data will be completed based on several cases. Linear interpolation was applied for missing data within 6 hours (2.3% data). Linear interpolation between upstream and downstream stations was applied for missing either upstream or downstream (7.5 % data). Linear interpolation between adjacent stations was applied when missing both data in upstream and downstream in Dampierre station but existed in adjacent stations (0.5% data). The remaining missing data was then filled based on the seasonal Kalman smoother, which estimates the missing values while considering the seasonal patterns and annual trend (*tsmoothie* package) (0.3% data).

The data cleaning and correction for hourly DO were carried out by following steps which were from Diamond et al. (in revision.). We first removed physically impossible values and then applied a lowpass filter to remove instrument noise in the DO signal. We then removed values that exceeded plausible hourly changes in DO (e.g., a leap from 10 to 15 mg L^{-1}) using 95% confidence intervals for hourly changes on a monthly basis as our cutoff. We finally used visual inspection to flag data of questionable validity and corrected for linear drift and anomalous drops or jumps in DO data. We then filled all remaining missing with a seasonal Kalman filter.

S3. Handling estimated GPP, ER and K600

GPP, ER

Although the *streamMetabolizer* model uses inputs such as light, water temperature, and river discharge to reduce the equifinality of GPP, ER, and K600, this model can produce unrealistic values, like negative GPP. This issue typically arises when diel variations in dissolved oxygen (DO) are weak—meaning the DO levels are similar between day and night—making it difficult for the model to accurately separate the contributions of GPP and ER (Appling et al., 2018). When the diel DO signal is minimal, the GPP is likely close to zero, which can lead to the model estimating a negative median GPP value. Consequently, it is common practice to set these negative GPP estimates to zero (Blaszczak et al., 2019). In our study, we used a different approach by replacing negative GPP estimates with the 75th percentile of GPP values estimated by the *streamMetabolizer* model rather than force to zero. However, this adjustment did not substantially alter the annual GPP calculations. Replacing negative GPP with the 75th percentile increased annual GPP by an average of 1.3% (ranging from 0.1% to 5.3%), while setting negative GPP to zero resulted in a smaller increase, ranging from 0.04% to 3.4% (Figure S3). Similarly, the annual ER calculations across different treatments for unrealistic ER values show no significant differences, with an average flux variation of around 1%, except in 1995, where the difference reaches 15% (Figure S4).

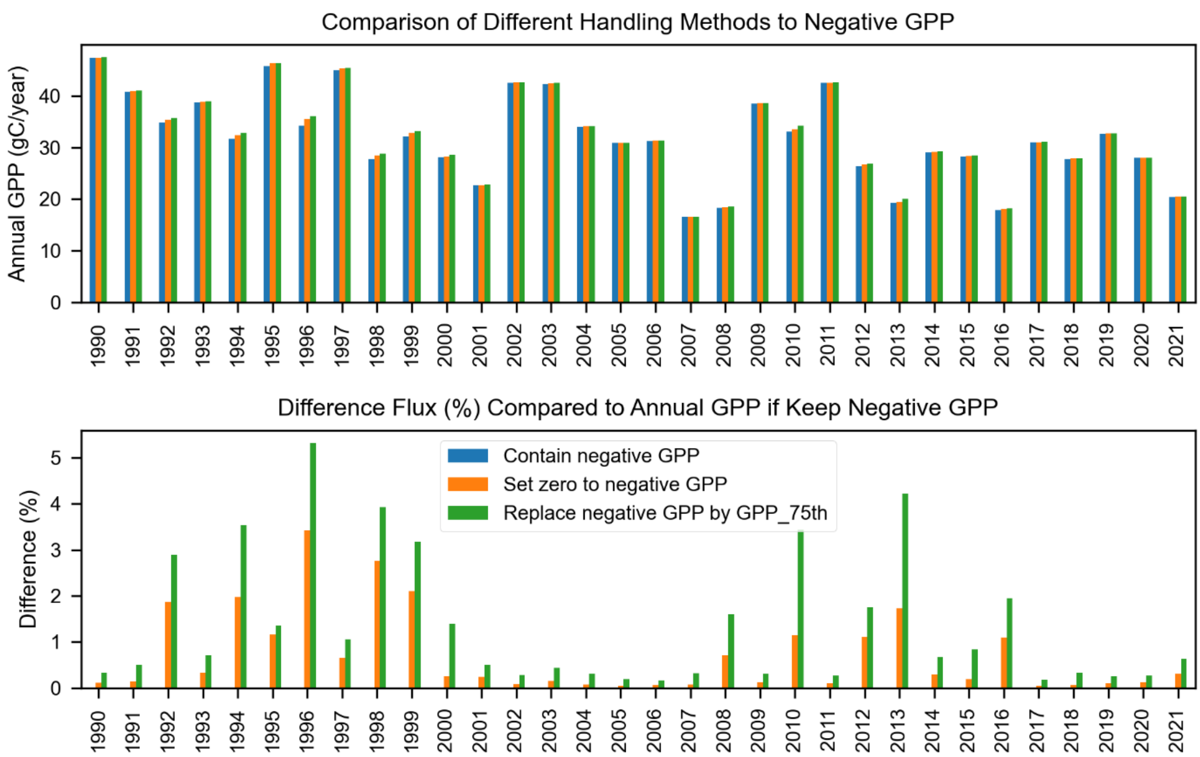

Figure S3. Comparison of annual GPP estimates based on different approach for handling negative GPP values: retaining negative GPP, setting negative GPP to zero, and replacing negative GPP with the 75th percentile of estimated GPP from the *streamMetabolizer* model.

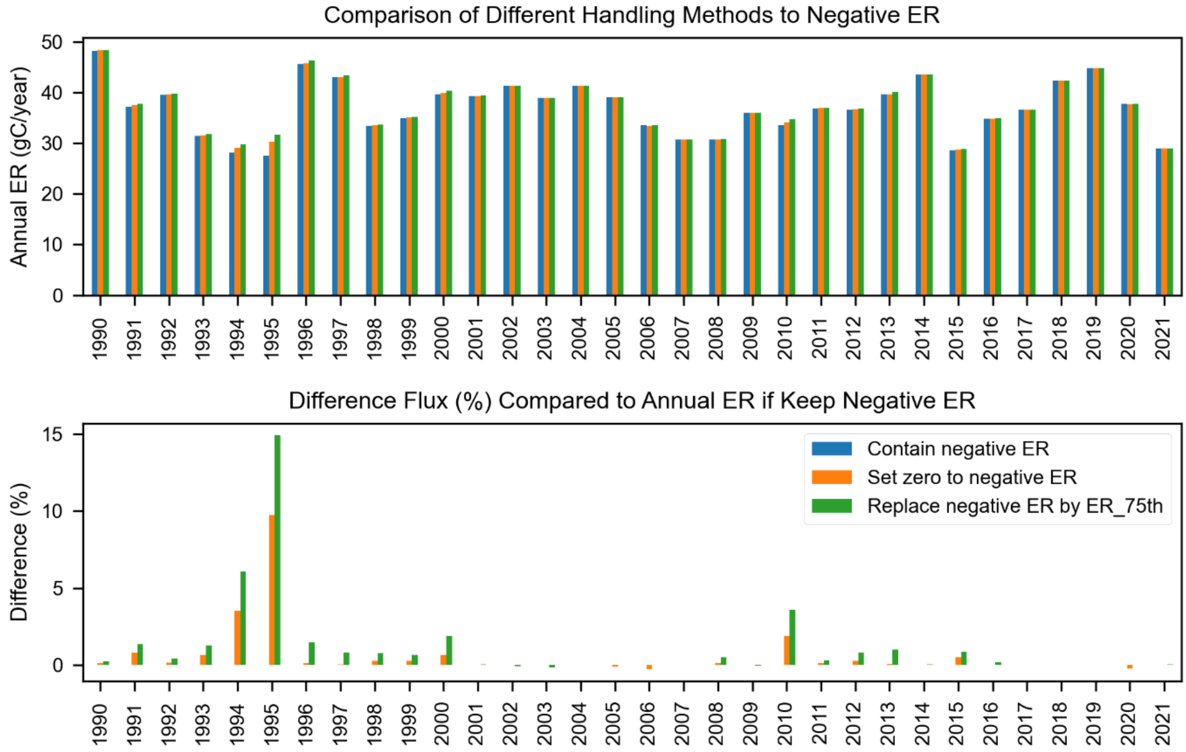

Figure S4. Comparison of annual ER estimates based on different approach for handling negative ER values: retaining negative ER, setting negative ER to zero, and replacing negative ER with the 75th percentile of estimated ER from the *streamMetabolizer* model.

K600

The k600 values estimated by the StreamMetabolizer model were compared with the mean k600 (m d^{-1}) calculated from seven fitted equations proposed by Raymond et al. (2012) for streams and small rivers (Table S1). Both k600 estimates exhibited similar seasonal fluctuations, with the lowest values occurring in summer and the highest in winter. The comparison revealed that the mean absolute percentage error (MAPE) between the StreamMetabolizer estimates and the mean k600 from the seven fitted equations ranged from 36% to 62%. Specifically, the Raymond et al. (2012) k600 estimates tended to be higher in summer and lower in winter compared to those estimated by the StreamMetabolizer model. However, the k600 values derived from StreamMetabolizer fall within the same order of magnitude as those from the seven fitted equations (Figure S5). The k600 estimates from the StreamMetabolizer model were selected for FCO2 calculations to ensure consistency with the NEP calculations.

Table S1. Seven fitted equations for predicting the k600 (m d^{-1}) for stream/rivers based on velocity (V, in m s^{-1}), slope (S; unitless), depth (D, in meters), discharge (Q, in $m^3 s^{-1}$), and the Froude number (Fr; unitless) (Raymond et al., 2012).

Model equation
1. $k_{600} = (VS)^{0.89 \pm 0.0.020} \times D^{0.54 \pm 0.030} \times 5037 \pm 604$
2. $k_{600} = 5937 \pm 606 \times (1 - 2.54 \pm 0.223 \times Fr^2) \times (VS)^{0.89 \pm 0.017} \times D^{0.58 \pm 0.027}$
3. $k_{600} = 1162 \pm 192 \times S^{0.77 \pm 0.028} V^{0.85 \pm 0.045}$
4. $k_{600} = (VS)^{0.76 \pm 0.027} \times 951.5 \pm 144$
5. $k_{600} = VS \times 2841 \pm 107 + 2.02 \pm 0.209$
6. $k_{600} = 929 \pm 141 \times (VS)^{0.75 \pm 0.027} \times Q^{0.011 \pm 0.016}$
7. $k_{600} = 4725 \pm 445 \times (VS)^{0.86 \pm 0.016} \times Q^{-0.14 \pm 0.012} \times D^{0.66 \pm 0.029}$

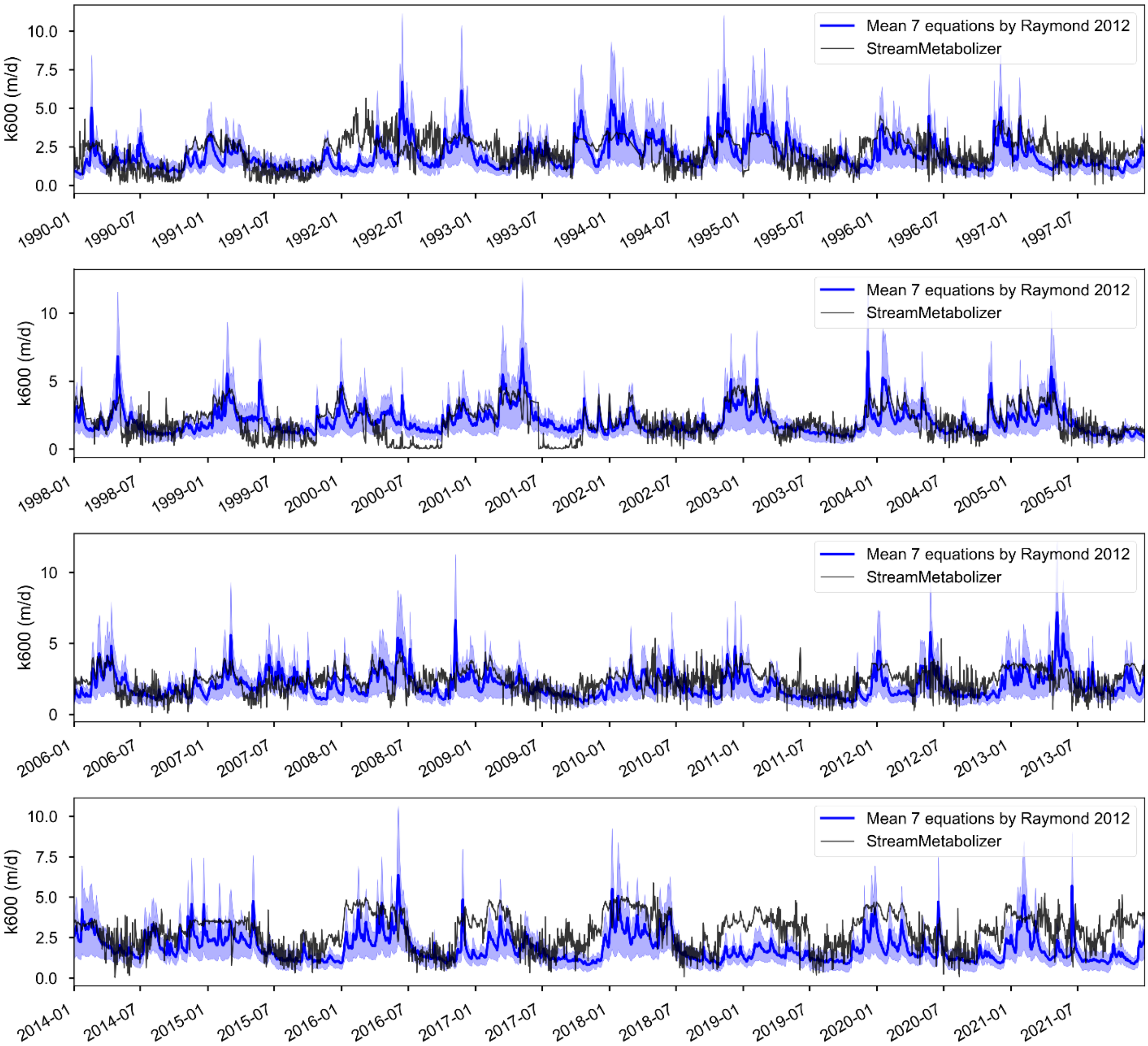

Figure S5. Comparison of estimated k600 from *StreamMetabolizer* model and mean of seven fitted equations from Raymond et al. (2012) for streams/rivers.

S4. Uncertainties in FCO2 and NEP estimation

Estimating FCO2 and NEP using models such as PyCO2SYS and streamMetabolizer often involves large uncertainties, particularly when considering the propagation of errors in all model input data and the summing/multiplying of these uncertainties in calculating fluxes (Battin et al., 2023; Kirk & Cohen, 2023). In this study, we assumed that after a careful data treatment process, the continuous datasets of DO, pH, conductivity, water temperature, discharge, and solar radiation were accurate. While both the PyCO2SYS and streamMetabolizer models provide a range of uncertainty, we used the average of these distributions as the best daily estimates, using the default input data accuracy. However, it is important to note that the daily total alkalinity (TA) data did not cover the entire 32-year period, unlike the other variables. Consequently, the error in TA reconstruction could introduce uncertainty in FCO2 estimation and potentially affect conclusions regarding the temporal distribution of CO2 sink/source states throughout the year, as well as comparisons with NEP. However, our analysis indicates that the uncertainty in the estimated TA (±190 μmol/L) only leads to ±11% uncertainty in pCO2 estimation by PyCO2SYS. As shown in Table S1, the statistical results comparing the annual distribution of trophic states remain consistent, with a maximum deviation of only 3%. Moreover, the dominance of the CO2 source–heterotrophic state throughout the year remains almost unchanged, with less than a 1% difference under any range of TA uncertainty, even though the magnitude of FCO2 could vary up to 20%.

Table S1. Comparison of the occurrence and fluxes of each trophlux state within the uncertainty range of estimated alkalinity.

		CO2 source - Heterotrophic			**CO2 source - Autotrophic**		
	Period	Min	Mean	Max	Min	Mean	Max
Occurrence (% of days)	**1990-2000**	47	47.3	47.7	15.6	16.7	17.6
	2001-2010	60.3	61.2	61.3	23.6	25.3	27
	2011-2021	65.4	65.7	65.7	24.7	26.2	27.6
FCO2 (gC/m2/y)	**1990-2000**	830.7	954.2	1100.2	87.2	102.6	118.9
	2001-2010	1266.5	1453.5	1668.7	75	87.7	102.4
	2011-2021	602.2	717.3	840.8	48.6	58.7	70.6

		CO2 sink - Heterotrophic			**CO2 sink - Autotrophic**		
	Period	Min	Mean	Max	Min	Mean	Max
Occurrence (% of days)	**1990-2000**	7.6	7.3	6.9	29.8	28.7	27.8
	2001-2010	2.3	1.7	1.5	14.1	15.5	13.4
	2011-2021	1.4	1.1	1.3	8.8	7.3	5.9
FCO2 (gC/m2/y)	**1990-2000**	-4.6	-4.4	-3.8	-22.4	-21	-19.6
	2001-2010	-1.3	-0.6	-1	-7.4	-7.8	-6.6
	2011-2021	-1.5	-0.9	-1.2	-3.4	-2.6	-2.1

S5. Change-point analysis

We evaluated the long-term changes in FCO2 and metabolism using a statistical change point analysis, which identifies points in a time series where the statistical properties, such as the mean or variance, undergo significant shifts. We first applied seasonal decomposition on daily time series to extract trend, seasonal, and residual components using the *statsmodels* Python package (Seabold & Perktold, 2010). Subsequently,

The long-term trend component was analyzed using a piecewise linear regression method (model ="linear" in *ruptures*, a Python package), while shift point detection by standard deviation (changes in variance by model="normal" in *ruptures*) was employed for the seasonal components. This process was also applied on related parameters including daily discharge, temperature, GPP, ER

2. Supplementary results

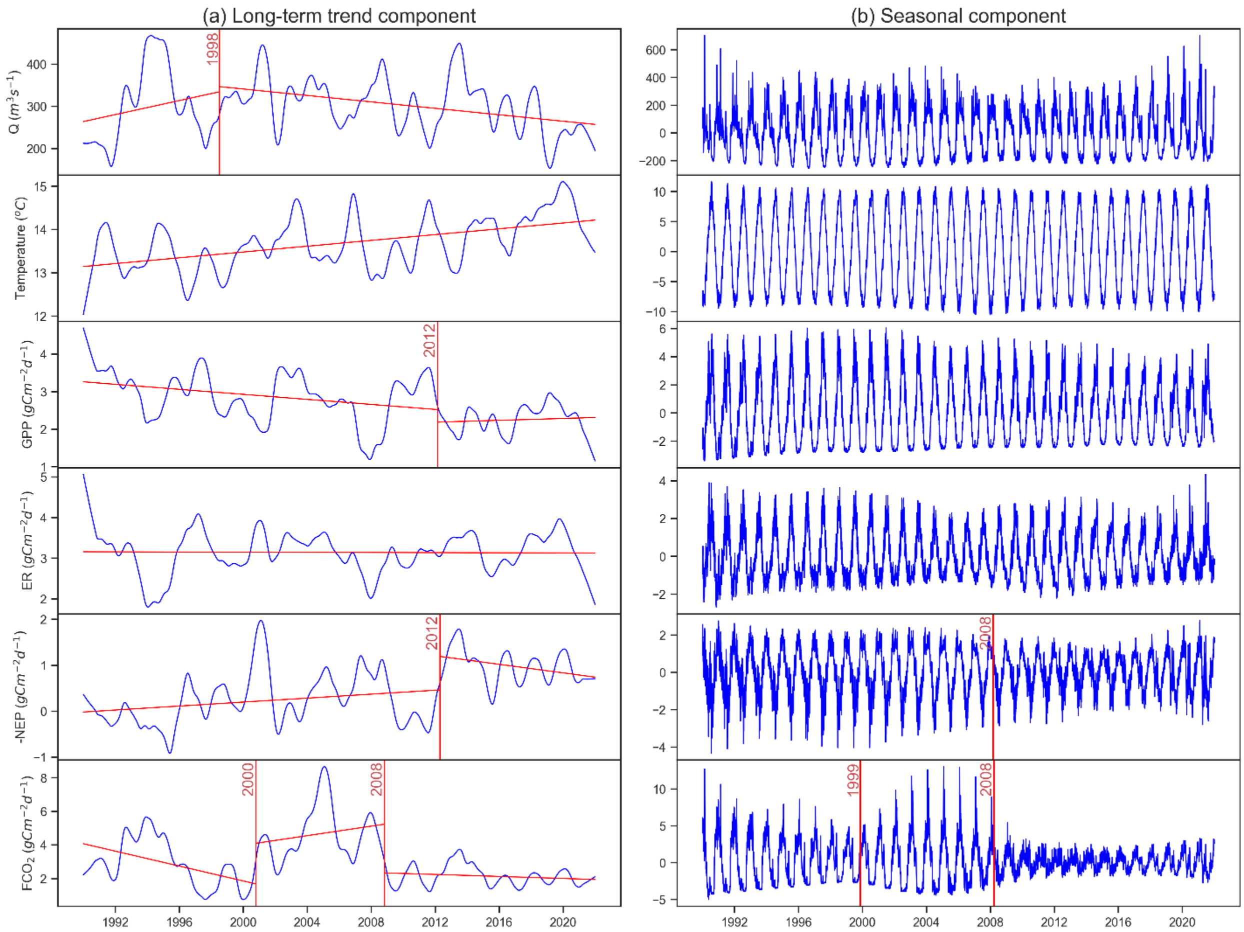

Figure S6. Change-point analysis on the (a) long-term trend components and (b) seasonal components of daily discharge, temperature, and fluxes of GPP, ER, -NEP, and FCO_2. The red vertical lines indicate the change periods.

Table S2. The correlations between annual FCO2, -NEP, and hydroclimatic conditions (discharge, temperature) in each trophlux state

CO2_NEP_state	Parameters	Days	Discharge	Temp	FCO2	-NEP	-NEP/CO2
Autotrophic Sink	Days		0.08	0.01	0.07	0.14	0.00
	Discharge	0.08		0.07	0.19	0.03	0.19

	Temp	0.01	0.07		0.04	0.00	0.00
	FCO2	0.07	0.19	0.04		0.14	0.32
	-NEP	0.14	0.03	0.00	0.14		0.13
	-NEP/CO2	0.00	0.19	0.00	0.32	0.13	
Autotrophic Source	**Days**		0.00	0.06	0.06	0.16	0.00
	Discharge	0.00		0.34	0.41	0.00	0.36
	Temp	0.06	0.34		0.14	0.00	0.21
	FCO2	0.06	0.41	0.14		0.13	0.32
	-NEP	0.16	0.00	0.00	0.13		0.01
	-NEP/CO2	0.00	0.36	0.21	0.32	0.01	
Heterotrophic Sink	**Days**		0.01	0.21	0.00	0.03	0.05
	Discharge	0.01		0.40	0.67	0.00	0.16
	Temp	0.21	0.40		0.40	0.10	0.90
	FCO2	0.00	0.67	0.40		0.01	0.18
	-NEP	0.03	0.00	0.10	0.01		0.12
	-NEP/CO2	0.05	0.16	0.90	0.18	0.12	
Heterotrophic Source	**Days**		0.00	0.69	0.01	0.30	0.07
	Discharge	0.00		0.04	0.36	0.11	0.11
	Temp	0.69	0.04		0.09	0.07	0.06
	FCO2	0.01	0.36	0.09		0.06	0.59
	-NEP	0.30	0.11	0.07	0.06		0.04
	-NEP/CO2	0.07	0.11	0.06	0.59	0.04	

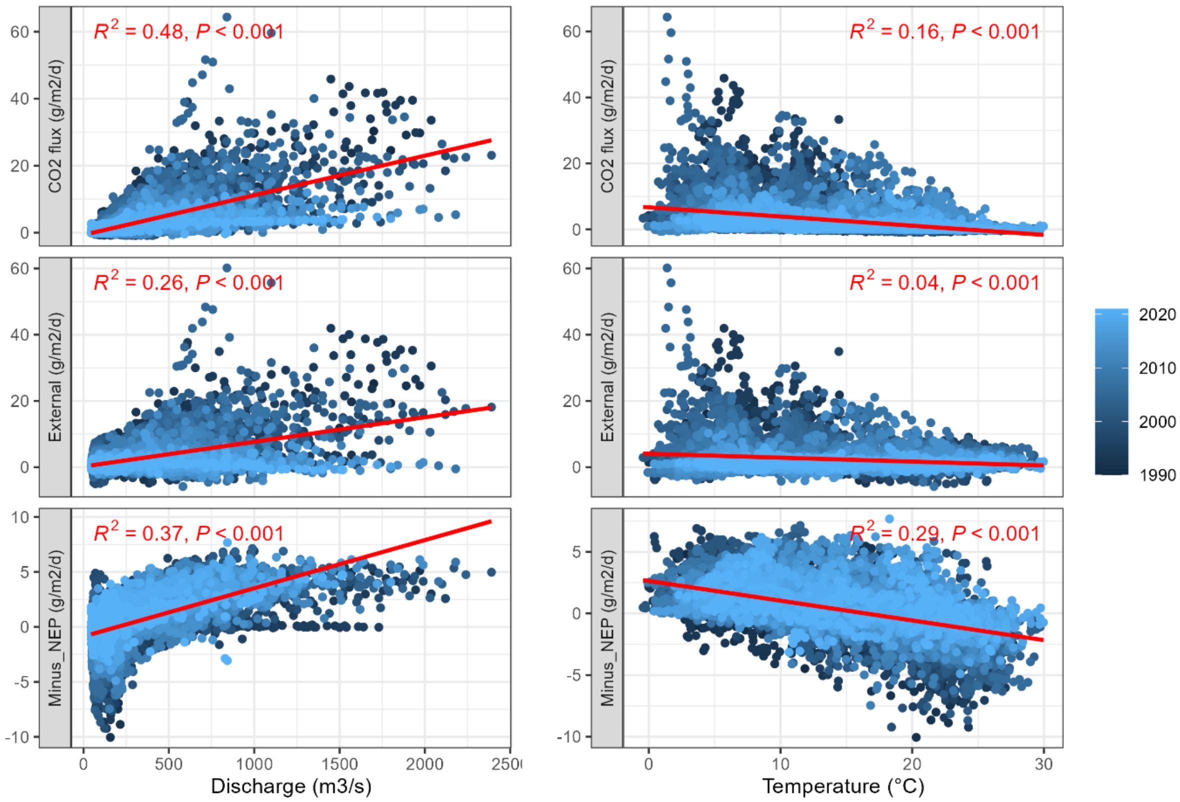

Figure S7. Relationship of daily fluxes and annual discharge or annual water temperature

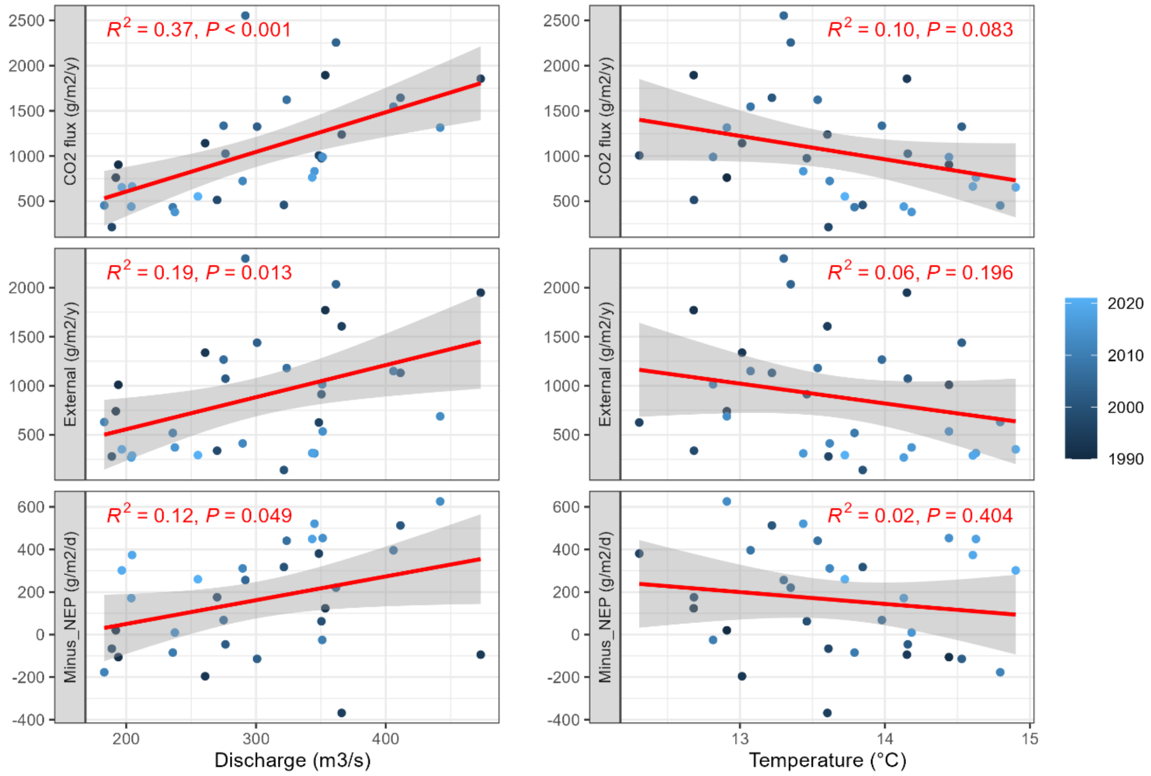

Figure S8. Relationship of annual fluxes and annual discharge or annual water temperature

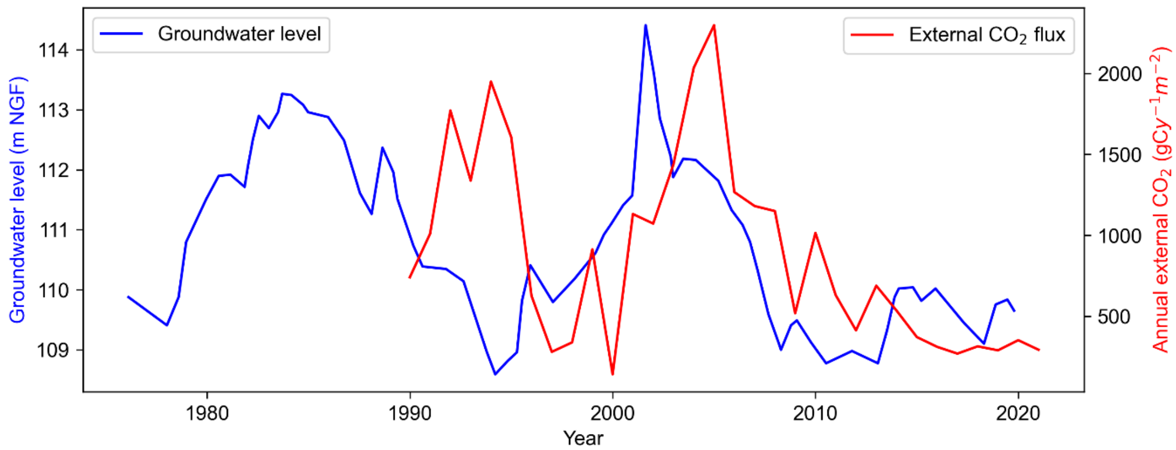

Figure S9. Multi-annual patterns of annual external CO2 source in Loire River (this study) and mean annual groundwater level in France (data extracted from Baulon et al., (2022))

References

Appling, A. P., Hall, R. O., Yackulic, C. B., & Arroita, M. (2018). Overcoming Equifinality: Leveraging Long Time Series for Stream Metabolism Estimation. Journal of Geophysical Research: Biogeosciences, 123(2), 624–645. https://doi.org/10.1002/2017JG004140

Battin, T. J., Lauerwald, R., Bernhardt, E. S., Bertuzzo, E., Gener, L. G., Hall, R. O., Hotchkiss, E. R., Maavara, T., Pavelsky, T. M., Ran, L., Raymond, P., Rosentreter, J. A., & Regnier, P. (2023). River ecosystem metabolism and carbon biogeochemistry in a changing world. Nature, 613(7944), 449–459. https://doi.org/10.1038/s41586-022-05500-8

Baulon, L., Allier, D., Massei, N., Bessiere, H., Fournier, M., & Bault, V. (2022). Influence of low-frequency variability on groundwater level trends. Journal of Hydrology, 606, 127436. https://doi.org/10.1016/j.jhydrol.2022.127436

Blaszczak, J. R., Delesantro, J. M., Urban, D. L., Doyle, M. W., & Bernhardt, E. S. (2019). Scoured or suffocated: Urban stream ecosystems oscillate between hydrologic and dissolved oxygen extremes. Limnology and Oceanography, 64(3), 877–894.

Camenen, B., Grabowski, R. C., Latapie, A., Paquier, A., Solari, L., & Rodrigues, S. (2016). On the estimation of the bed-material transport and budget along a river segment: Application to the Middle Loire River, France. Aquatic Sciences, 78(1), 71–81. https://doi.org/10.1007/s00027-015-0442-3

Jones, A. S., Jones, T. L., & Horsburgh, J. S. (2022). Toward automating post processing of aquatic sensor data. Environmental Modelling and Software, 151, 105364. https://doi.org/10.1016/j.envsoft.2022.105364

Kirk, L., & Cohen, M. J. (2023). River corridor sources dominate CO2 emissions from a lowland river network. Journal of Geophysical Research: Biogeosciences, 128(1), e2022JG006954. https://doi.org/10.1029/2022JG006954

Moatar, F., Descy, J.-P., Rodrigues, S., Souchon, Y., Floury, M., Grosbois, C., Minaudo, C., Leitao, M., Wantzen, K. M., & Bertrand, F. (2022). Chapter 7—The loire river basin. In K. Tockner, C. Zarfl, & C. T. Robinson (Eds.), Rivers of Europe (Second Edition) (pp. 245–271). Elsevier. https://doi.org/10.1016/B978-0-08-102612-0.00007-9

Moatar, F., Miquel, J., & Poirel, A. (2001). A quality-control method for physical and chemical monitoring data. Application to dissolved oxygen levels in the river loire (france). Journal of Hydrology, 252(1–4), 25–36. https://doi.org/10.1016/S0022-1694(01)00439-5

Raymond, P. A., Zappa, C. J., Butman, D., Bott, T. L., Potter, J., Mulholland, P., Laursen, A. E., McDowell, W. H., & Newbold, D. (2012). Scaling the gas transfer velocity and hydraulic geometry in streams and small rivers. Limnology and Oceanography: Fluids and Environments, 2(1), 41–53. https://doi.org/10.1215/21573689-1597669

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. 92–96. https://doi.org/10.25080/Majora-92bf1922-011