

Dear editor and reviewers,

Thank you for your time and for the careful evaluation of our revised manuscript titled "*Breathing Storms: Enhanced Ecosystem Respiration During Storms in a Heterotrophic Headwater Stream*". We have followed all recommendations from both the reviewer and the editor and have revised the manuscript accordingly. All changes are indicated in the track-changed version, and below, we respond to each comment. The original text of the decision letter is in *black and italics*, while our responses are in dark blue.

General comments: The revised manuscript represents a substantial improvement over the initial submission.

The authors now report the frequency of high-flow days (less than 5% of storm days). The new Table S3 in the supplement is excellent. It explicitly links QC failures to discharge ranges and details the specific reasons for failure (e.g., poor model fit, implausible values). This directly addresses my request and improves the robustness of the paper.

The discussion appropriately states that the stopper hypothesis could not be empirically tested with reliable data from their study, and they provide plausible reasons for model failure at high flows (e.g., sensor burial). The updated discussion clearly states that no clear saturation pattern in ER was identified, contrasting with typical nutrient uptake patterns.

My last concern is: the re-analysis is described but not shown, making it unverifiable. The finding from this re-analysis show that 8 plausible ER estimates could be generated and that they did not show a stopper effect.

From your response letter: "For these days, high Q ranged from 100 to 1370 L s⁻¹ and ER values were similarly elevated (range: 20 to 41g O₂ m⁻² d⁻¹) as on days with large ΔQ increases, providing no evidence for a metabolic stopper effect within our data".

--> Please provide the data from this re-analysis. A new supplementary table is required that lists the results for the 8 successful high-flow runs mentioned in your response letter.

We thank the reviewer for the positive feedback and constructive suggestions. In response, we have added a new supplementary table (Table S5) presenting the metabolic activity for all storm days with discharge greater than 100 L s⁻¹, including their quality check statistics (R², RMSE). The table highlights both the unsuccessful runs and the 8 successful high-flow runs from the re-analysis with a constrained K_{600} .

Table S5. Results of the reanalysis of storm days with daily discharge above 100 L s⁻¹ using *StreamMetabolizer* with constrained K_{600} to 100 d⁻¹. For each date, the table shows daily discharge (Q), model fit metrics (R², RMSE), and daily estimates of gross primary production (GPP, g O₂ m⁻² d⁻¹), ecosystem respiration (ER, g O₂ m⁻² d⁻¹), and reareation rates (K_{600} , d⁻¹). Values in bold indicate days that passed the quality check. Missing values (NA) indicate cases where reliable metabolic estimates could not be obtained..

date	Q	R ²	RMSE	GPP	ER	K600
dd/mm/yyyy	L s ⁻¹			g O ₂ m ⁻² d ⁻¹	g O ₂ m ⁻² d ⁻¹	d ⁻¹
15/10/2018	157.2	98.40	6.74	-0.03	-20.34	99.99
01/11/2018	158.6	6.67	57.93	0.62	-20.44	100.00
15/11/2018	104	-54.44	69.32	0.19	-20.63	100.00
16/11/2018	356.4	-186.36	98.23	-1.30	-37.59	100.00
17/11/2018	218.8	-16.32	61.90	0.53	-33.13	100.00
18/11/2018	744.1	-16009.87	714.99	-3.95	-67.99	99.86
19/11/2018	1037	-2748.18	295.03	4.86	-106.70	99.86
20/11/2018	283.2	80.42	24.05	-1.85	-54.05	100.00
21/11/2018	154.1	-113.10	80.63	0.84	-42.65	99.99
14/04/2020	193.2	-223.85	105.56	1.25	-39.60	99.98
15/04/2020	135	-50.69	71.85	1.45	-38.95	100.00
16/04/2020	108.8	-222.15	104.70	1.30	-35.49	99.99
19/04/2020	183.9	-2341.77	284.57	0.50	-34.16	99.97
20/04/2020	295.5	-15.73	62.04	0.48	-40.08	100.00
21/04/2020	2319.7	-20530.77	828.47	0.77	-51.19	99.81
22/04/2020	3883.7	-3218.55	332.92	3.42	-69.40	99.81
23/04/2020	364.8	-1.19	57.75	0.33	-48.12	100.00
24/04/2020	227.7	76.31	27.80	0.29	-41.12	99.99
25/04/2020	169.1	96.06	11.23	0.05	-38.53	99.98
26/04/2020	136.3	97.90	8.15	0.13	-37.14	100.00
27/04/2020	115.3	98.25	7.39	0.06	-36.86	100.00
21/03/2022	108.2	NA	NA	NA	NA	NA
28/04/2020	100.2	93.27	14.41	0.19	-35.51	99.98
21/04/2022	259.5	64.28	35.16	0.49	-36.35	99.96
22/04/2022	150.3	95.04	13.34	0.81	-22.13	99.99
23/04/2022	104	74.67	29.90	0.61	-19.69	99.99

Minor comments:

Line 25: ...was found for ER ($R^2 > 0.37$)...recovery times were positively related to the size of the event only for ER ($R^2 > 0.46$). The wording is slightly imprecise. Why not acknowledge the precise range/values?

In these two cases, changes in ER and recovery time of ER both exhibited a positive relationship with ΔQ , with linear and logarithmic models providing equally strong fits based on AIC values. To maintain the abstract's conciseness and focus, we have chosen to report these results in a more general form, while the detailed model fits and values are presented in the main text.

L 20 (track-change version) “35 of them”: Please rewrite it to indicate how you opted for 35 “selected” (?) storm events.

L 21 “considering all events”: Do you mean “all selected 35 events”?

We have changed the sentence to “Due to data and model constraints, we were able to estimate metabolic rates for 35 of the events” to improve clarity. (L21)

L 23 “unrelated”: Given the hydrologic theme of this study, this expression doesn’t make sense logically. Please clarify whether you meant “not significantly related”.

We have changed to “not statistically significant” as suggested. (L25)

L 51-55 & associated discussion (e.g., L 333-341; 360-365): Given the reviewer’s concern about the “stopper” effect, I would invite you to revise the terms and definitions in line with the literature and other descriptions in the manuscript: For instance, refer to Fig. 5 in Covino, 2017 (<http://dx.doi.org/10.1016/j.geomorph.2016.09.030>), where high flows (connectivity) function as a kind of modulator or suppressor of reactivity, not as “stoppers”. Unlike triggers or stoppers, I got the impression that the terms “stimulation” and “suppression” in Fig. 1 would not elicit any unnecessary misunderstanding regarding the often subtle, storm-induced metabolic shifts.

We have revised the terminology throughout the manuscript to align with terms used in the literature, including Covino’s (2017) conceptual framework. Specifically, we replaced the terms “metabolic trigger” and “metabolic stopper” with the more process-oriented terms “stimulation” and “suppression” that are already used in our hypotheses (Fig. 1). We believe these revisions clarify the underlying processes, avoid misunderstandings, and ensure consistency with the hydrologic connectivity–reactivity framework described in the literature.

L 118 “are”: Please use the past tense consistently when you describe methods and findings.

We have changed to “were” as suggested.

L 145 “DO”: Please define this at its first use (L 131?) Given the importance of DO measurements in this study, it would be helpful if more details are provided, particularly with regard to sensor principle (optical or old membrane-type?) and maintenance (e.g., cleaning measures, especially to handle any sensor biofouling issue)

First use of “DO” in L49 has been defined. The MiniDOT used in this study is an optical (fluorescent optode) logger designed for robust, long-term field deployments with minimal maintenance requirements. To ensure data quality, sensors were inspected every 15–20 days for debris or biofilm, and the sensing surface and housing were gently rinsed with stream water when needed. In any case, note that sensors exhibited reduced biofouling due to oligotrophic water and low light inputs, which restricts the growth of organisms that form biofilms. These details have been added to the Methods section.

L 251: Just to double check whether $r^2 = 0.06$ was significant at $p < 0.001$? By comparing the numbers described in L 251-253, I wondered about the significance of this marginal r^2 number.

We appreciate the editor’s attention to this point. We have double-checked the analysis and confirmed that the reported values are correct. Furthermore, we now specify in the manuscript that the relationship was weak, although statistically significant.

L 310: Please use “by” instead of “over”.

We have changed to “by” as suggested. (L311)

L 387-390: Please rephrase or provide more supporting evidence to respond to the reviewer’s concern.

We have revised the manuscript to better answer the reviewers' concern about the need to revisit the River Network Saturation concept and to differentiate the observed saturation in recovery time from the absence of saturation in the ER magnitude response. Updated text (L377-386).

“Moreover, the relationship between ΔER and ΔQ was equally well explained by both linear and logarithmic models, indicating no statistically supported evidence for a saturation pattern in ER. This finding contrasts with the saturation responses often

reported for nutrient uptake with increasing discharge (sensu River Network Saturation concept; Wollheim et al., 2018). Given that the largest analysed storms fell within a moderate ΔQ range compared to extreme hydrological events in similar systems, it is possible that the ecosystem's processing capacity was not fully exceeded, and thus the available range was insufficient to reveal a true saturation response. Nevertheless, we observed apparent constraints in ER during storms. For instance, ER rates were never below $-36.4 \text{ g O}_2 \text{ m}^{-2} \text{ d}^{-1}$, and the maximum ΔER was around 100% indicating that storm-driven increases in ER never exceeded twice the prior baseflow rates. These observed constraints on the heterotrophic response to storm disturbances should be interpreted as empirical bounds within our dataset rather than definitive evidence of metabolic saturation in Fuirosos."

L 424 "carbon cycling": Please articulate specific aspects of the aquatic carbon cycling related to your findings.

We appreciate the reviewer's suggestion to be more specific regarding the aspects of aquatic carbon cycling related to our findings. We have revised the conclusion (L418–L423) to explicitly link our results to carbon processing pathways. The text now specifies that storm-driven stimulation of ER represents rapid pulses of organic carbon mineralization and CO₂ emissions, and that the magnitude and frequency of these "breathing storms" influence both the total annual CO₂ flux and its temporal variability, with potential effects on downstream C transport and emissions. We also clarify that changes in hydrological regimes could alter the timing, intensity, and cumulative magnitude of these pulses, thereby modifying key pathways of aquatic carbon cycling.