Carbonate weathering as a driver of CO\textsubscript{2} supersaturation in lakes

Rafael Marcé, Biel Obrador, Josep-Anton Morguí, Joan Lluís Riera, Pilar López, Joan Armengol

Supplementary Methods

Data collection. We used data collected during a national-wide sampling program during 1987-1988 including 101 reservoirs\textsuperscript{15}. This database offers a unique opportunity due to the verifiable accuracy of in-situ chemical analyses, the homogeneity of methods applied across lakes, and the wide range of DIC content and trophic states it covers (Supplementary Table 1 and Supplementary Data). Every system was visited during winter and summer for a total of 202 samples. We measured surface DO concentration by performing Winkler titrations in the field, and DIC and dissolved CO\textsubscript{2} concentrations were estimated from pH (Ross combination electrode Orion 81-04), water temperature, and alkalinity (in situ Gran titration with HCl) along with the dissociation constants for carbonic acid in freshwater systems\textsuperscript{31}. CO\textsubscript{2} concentration calculations are very sensitive to errors in pH measurements\textsuperscript{2}. In our study this was minimized by the fact that all pH determinations were performed with the same probe and by the same person (J. A. M.) after calibration with buffer solutions in the field. The accuracy of the alkalinity determinations was assessed elsewhere by comparing alkalinity with the ionic balance of the samples\textsuperscript{32}. In calculating values for DO, DIC, and CO\textsubscript{2} disequilibrium relative to the atmosphere, elevation effects on atmospheric pressure and a baseline of 350 ppmv CO\textsubscript{2} for 1987-1988 were considered. CO\textsubscript{2} concentrations plotted in Fig. 2 were solved at 25ºC after standardization of pH values at the same temperature to discount the effect of varying water temperatures.
Regression analyses between DO and DIC disequilibrium. To investigate the relationships between DO and DIC disequilibrium relative to the atmosphere in our dataset we performed linear regressions grouping the data considering the 33% and 66% alkalinity percentiles: low alkalinity (n = 67, median = 0.4, range = 0.06-0.83 meq L\(^{-1}\)), mid alkalinity (n = 68, median = 1.6, range = 0.88-2.22 meq L\(^{-1}\)), and high alkalinity (n = 67, median = 3.0, range = 2.27-4.69 meq L\(^{-1}\)). The distribution of some variables inside these three groups showed non-normality. Considering the overly conservative Shapiro-Wilk W test only two out of 6 distributions at play (3 groups for DO and DIC) showed conspicuous deviations from normality (DIC at low and medium alkalinity). We repeated all regression analyses showed in Figure 1b normalizing all variables using Box-Cox transformation. After transformation, all distributions were normal (Shapiro-Wilk W test) except DIC at low alkalinity, that still showed a moderate deviation from normality (skewness=-1.08); and alkalinity, that showed a flat distribution (skewness=0.07) that did not compromise the regression analysis. The analyses showed in Figure 1b but using the transformed variables resulted in an identical result when compared to the analyses using the original variables.

We also performed a factorial regression analysis to test the effect of alkalinity on the DO vs. DIC relationships. Results using original and Box-Cox transformed variables also rendered identical results.

DO disequilibrium as a proxy of NEP. The keystone of our modeling approach is the use of dissolved DO disequilibrium relative to the atmosphere as a surrogate for NEP. This is based on the fact that the time scale for gas equilibration with the atmosphere is large compared to the velocity of metabolic reactions\(^{33}\). Therefore, our DO disequilibrium estimates are representative of conditions at the short to medium term, because the supposition of a system closed to the atmosphere would not hold for longer periods (e.g., year). Despite the fact that some recent
works support the use of discrete sampling data to estimate lake NEP\textsuperscript{34} there are obvious concerns regarding the magnitude of the diel variations in DO, CO\textsubscript{2}, and pH. Consequently, we do not calculate NEP values in mass per volume per time units. However, it is important to note that we do not aim to characterize a water mass in terms of a particular time scale, but rather assess the relationship between carbon and oxygen deviations from equilibrium for a particular moment. In addition, the need for considering spatial heterogeneity in carbon processing in lakes has been increasingly stressed, either in the vertical\textsuperscript{35} or horizontal\textsuperscript{36,37} dimensions. Spatial heterogeneity appears to be particularly relevant for CO\textsubscript{2} (refs. 38-41) and CH\textsubscript{4} (refs. 42-45) emission estimates from lakes and reservoirs. However, we must note again that our aim is not to characterize the carbon budget of a particular water body, but to assess the relationships between CO\textsubscript{2} and a proxy of NEP in surface water parcels showing contrasted DIC contents. This means that our data should not be considered as representative of the whole system.

All in all, since no temporal and spatial scale can be unequivocally adopted in our calculations, our NEP proxy stays as DO disequilibrium expressed in mass per liter units. Note that this limitation defines our surrogate NEP values as semi-quantitative estimates, because two similar values can indeed be the result of rather different net metabolic rates expressed in the customary mass per volume per time units.

**Rationale of the models and the calcification hypothesis.** We applied three different models combining different assumptions in a heuristic framework. The aim was to show how the different assumptions explained different patterns in the observed data, in order to guide the reader to the main conclusions of our paper. We understand that this was better accomplished by serially combining our assumptions (i.e., heuristically) in three models, which included the relevant steps of our reasoning.
Although the effects of calcification on the partial pressure of carbon dioxide is a classical topic in oceanography and freshwater research, the molar ratios between carbonate precipitation and NEP have been extensively studied in the ocean but they are almost absent in the limnological literature. The molar ratios between carbonate precipitation and NEP used in Model 3 ($\alpha$ in Eq. 2) are lower than those found in marine benthic ecosystems (1.3 for coral reefs), but comparable to ratios measured for marine planktonic assemblages (between 0.2 and 1, ref. 19). Our data come mainly from systems poor in benthic primary producers, hence it is possible that macrophyte-dominated lakes show higher rates of carbonate precipitation to NEP ratios, as has been found in incubation experiments with freshwater macrophytes. In any case, although the assumption of carbonate precipitation to NEP ratios dependent on fixed alkalinity thresholds is clearly not realistic for particular lakes, our results show that it is a reasonable approach when modelling processes across a large population of lakes.

Although our hypothesis that calcite reactions impact the DIC vs. DO disequilibrium is consistent with our results and with the thermodynamic state of calcium carbonate in our samples (Supplementary Figure 1), we acknowledge that we cannot unequivocally assign the observed effect to calcite reactions. Anaerobic metabolic processes like sulfate reduction and denitrification may also impart a similar change in the stoichiometry between DIC and DO, although this is unlikely to be responsible for a widespread effect on DIC levels in the surface layer of deep systems like the ones in our dataset. Another process that may be responsible for the observed effect on the DIC vs. DO relationship would be an exchange between water and the atmosphere faster for DO than for $\text{CO}_2$. Actually, the time needed to equilibrate a volume of water with the atmosphere is larger for $\text{CO}_2$ than for DO, due to the effects of the DIC pool on partial pressure of $\text{CO}_2$. Although we cannot discard this effect in our high alkalinity samples, it would be difficult to explain why this process is not affecting the DIC vs. DO relationship in low alkalinity systems as well, because even at low alkalinitiess the effect of DIC on $\text{CO}_2$ equilibration is not negligible. In any case, although our results are not at odds with the
hypothesis of carbonate precipitation included in Model 3, further research providing direct evidences on the effect of carbonate precipitation and dissolution reactions on lake CO$_2$ dynamics is required to reach a solid conclusion.

**Global map of alkalinity in runoff.** The basis for this calculation are the studies of global atmospheric CO$_2$ consumption by chemical weathering of minerals$^{30-31}$. Those studies collect information on alkalinity fluxes from watersheds of known lithology to build empirical relationships relating runoff and alkalinity fluxes for different lithologies. We assigned different empirical relationships (Supplementary Table 2) to the lithologic classes found in a global lithological map$^{52}$ with a resolution of 1 km$^2$. We used global composite runoff fields$^{53}$ to solve the equations for the local generation of alkalinity at every pixel (1 km$^2$) considering the corresponding lithology, and we converted the figures to alkalinity concentration in meq L$^{-1}$. All calculations and alignments were performed in ESRI ArcGIS Spatial Analyst. At this step, however, calculations still miss the fact that runoff accumulates along river networks, and this may dramatically change the values of alkalinity in large watersheds draining areas with very different lithologic classes and local runoff generation. To overcome this limitation, we solved the transport of alkalinity along river networks using the Dominant River Tracing (DRT), a global river network database designed to perform macroscale hydrologic calculations$^{54}$. This database merges the HydroSHEDS database$^{55}$ with HYDRO1k (USGS, https://lta.cr.usgs.gov/HYDRO1K) to cover high latitude regions not included in the former. We used the flow direction raster at 1/16 of a degree (aprox. 300 m) in ftp://ftp.ntsg.umt.edu/pub/data/DRT to generate an area accumulation raster in ESRI ArcGIS Spatial Analyst (the area accumulation raster at this resolution is incorrect or corrupted in the aforementioned ftp site). After alignment of the different rasters at 300 m resolution (flow direction, area accumulation, local runoff generation, and local alkalinity), the routing of runoff and alkalinity was solved in Matlab 7.2 after exporting all maps in ASCII format.
The area accumulation raster was used to hierarchically solve calculations, using the flow direction raster to define the cells contributing to each pixel. Alkalinity for accumulated runoff was calculated as the runoff-weighted average of the contributing cells (including the local pixel and considering accumulation effects as calculations progress). Raster maps, ASCII files, and Matlab scripts are available under request to R. M. See Supplementary Fig. 3b for a comparison between the map of alkalinity in locally generated runoff and the same map after considering water routing along river networks.

The final map of alkalinity in accumulated runoff was checked against alkalinity values stored in the GEMSTAT database maintained by UNEP/GEMS/Water Programme (www.gemstat.org). We initially collected 60848 measurement of alkalinity from surface water sampling sites (rivers, lakes, and reservoirs) around the globe, but we finally worked with 55055 measurements from 584 stations (Supplementary Fig. 3a) after discarding several measurements. The main reason to discard measurements was an ambiguous definition of the methodology used to measure alkalinity and of the final units of the results in the database, which coincided with apparently inconsistent results for some stations (extremely wide ranges or unrealistic upper or lower bounds). We calculated the mean alkalinity for each station (the average number of measurements at each station was 99) and compared this value with the alkalinity value of the corresponding pixel in our map of alkalinity in accumulated runoff, using the reported geographic coordinates of the site to extract the value from the map.

The comparison between values from the GEMSTAT database and those extracted from our map gave an average error of 0.26 meq L\(^{-1}\). The distribution of the alkalinity values obtained with the map was very similar to the measured values (Supplementary Fig. 3c), although the map slightly underestimates the prevalence of high alkalinity values. This was expected because our procedure does not account for concentration mechanisms promoted by evaporation, especially relevant in endorheic basins. For this reason the map is not suitable for predicting
alkalinity in saline environments. Another source of error potentially explaining why high alkalinity values are not adequately predicted in our map is the fact that evaporites are always poorly represented in analyses of alkalinity generation from runoff\textsuperscript{60-32}. In any case, the overall quality of the results is appropriate for a global analysis focused on global distributions like the one presented here.

**Global distribution of alkalinity in lakes and reservoirs.** We used two different sources of information to account for lakes and reservoirs. For reservoirs we used the GRAND database\textsuperscript{24}, while for lakes we used the GLWD database\textsuperscript{23}. The polygons of the GRAND database were rasterized at 300 m resolution, and areas corresponding to very large reservoirs (>600 km\textsuperscript{2}) were removed because the alkalinity map is void inside the margins of very large water bodies. Alkalinity and area values for those very large reservoirs were treated separately in this analysis using data collected from the GEMSTAT database. For lakes we considered the water bodies labeled as lakes in the Levels 1 and 2 of the GLWD database, also rasterized at 300 m resolution. Unfortunately, this includes some reservoirs already present in the GRAND database. To avoid this overlapping, we removed all waterbodies coincident with the GRAND database from our lake map. Very large lakes (>600 km\textsuperscript{2}) were treated separately following the same procedure as for very large reservoirs. To calculate the areal distribution of alkalinity across lakes and reservoirs we first projected all the layers using the cylindrical equal area Behrmann projection. Then we estimated the areal distribution of alkalinity in lakes and reservoirs by overlapping our map of alkalinity in accumulated runoff with our lake and reservoirs rasters. All analyses were performed in ESRI ArcGIS Spatial Analyst. Results for very large lakes and reservoirs were added to this calculation afterwards. We performed calculations for the entire Earth in 10° latitudinal strips. It is worth mentioning that this analysis does not include the area of very small water bodies not covered by the GRAND and GLWD databases. The area of lakes and reservoirs larger than 0.001 km\textsuperscript{2} has been estimated at 3.8 million km\textsuperscript{2} (ref. 56), while our lake and reservoir rasters (including very large water bodies) account for 2.5
million km$^2$. Although the fact that our analysis misses small water bodies is an obvious limitation of our approach, we preferred to restrict our analysis to lakes for which the explicit location was available. In any case, this implies that our estimates of weathering-related CO$_2$ emissions are indeed conservative figures.

**Potential emissions from the supply of alkalinity to lakes and reservoirs.** Adopting the assumptions that rock weathering is the main source of alkalinity in surface waters and that this DIC loading reaches lakes and reservoirs without significant equilibration with the atmosphere, we can use DIC$^W$ as defined in this work to calculate the potential emissions of CO$_2$ in lakes related to chemical weathering in the watershed. DIC$^W$ values across lakes were calculated assuming DIC$^W$ = alkalinity, using our global map of alkalinity in accumulated runoff as the source of alkalinity values. We used values from the GEMSTAT database for very large water bodies (>600 km$^2$) for which no alkalinity was available from our map. Dissolved CO$_2$ concentration was calculated from DIC$^W$ using CO2SYS$^{31}$ and assuming that the system was not in equilibrium with the atmosphere (i.e., closed system). We used the mean annual air temperature over the continents for the corresponding latitude (NOAA GHCN_CAMS database) to solve calculations in CO2SYS. Although it can be argued that we have applied our findings to natural lakes while our model assumptions have been tested with reservoirs, CO$_2$ concentrations in lakes and reservoirs older than 10-15 years are comparable$^2$. In our dataset, only two reservoirs are less than 15 years old, and average age was 33 years (Supplementary Table 1 and Supplementary Data). We assumed $k_{600} = 4$ cm hr$^{-1}$ (ref. 3), a value almost identical to recent estimates of the global $k_{600}$ average in lacustrine systems$^2$. The actual emissions were calculated by adjusting $k$ and Henry’s constant to the mean annual air temperature for the corresponding latitude. Current CO$_2$ concentration in the atmosphere was assumed at 390 ppmv. We did not consider the effects of chemical enhancement$^{57}$ on $k$ values for three reasons. First, chemical enhancement has been neglected in virtually all global assessments of CO$_2$ emissions from freshwaters, and its use would render our calculations less comparable to other
studies. Second, including chemical enhancement will increase the calculated emissions, and we wanted to keep our calculations as a conservative estimate. Finally and more fundamental, chemical enhancement is more likely to hold in systems showing very high biological activity. However, our calculations are intended to represent the emissions from the supply of weathering related carbon to lakes, without consideration of metabolic activity. Without metabolic activity and considering our assumption that DIC loading reaches lakes and reservoirs without significant equilibration with the atmosphere, pH values cannot rise above 8.5 irrespective of temperature and alkalinity values considered in this paper (calculations performed with CO2SYS). At pH below 8.5, the effect of chemical enhancement on k values is expected to be very small except at unreasonably high temperatures.

We also calculated emissions considering the lower and upper bounds of $k_{\text{w0}}$ reported in Ref. 2 (2.25 and 7.9 cm hr$^{-1}$) to assess the sensitivity of the calculated emissions to this parameter. Emissions were substantially sensitive to this parameter in the considered range, because emissions values are linearly dependent on k. Values varied 100% from the central estimate (0.09 P g C yr$^{-1}$), from 0.05 to 0.17 P g C yr$^{-1}$. This clearly indicates that reliable regional estimates of k are paramount for a correct assessment of carbon emissions from freshwaters. Emissions were aggregated in 10º latitudinal strips, and compared to previous latitudinal CO$_2$ emissions estimates. For this comparison, we calculated areal emissions to discount the effect of the different lake and reservoir area considered in Ref. 3 and our work. It should be noted that emissions from DIC$^W$ in this work must be understood as potential emissions. For instance, some high alkalinity systems are highly productive as well, showing CO$_2$ undersaturation.
Supplementary References


55. Lehner, B., Verdin, K. & Jarvis, J. New global hydrograph derived from spaceborne 

56. Mcdonald, C. P., Rover, J. A., Stets, E. G. & Striegl, R. G. The regional abundance and 
size distribution of lakes and reservoirs in the United States and implications for 

57. Wanninkhof, R. & Knox, M. Chemical enhancement of CO₂ exchange in natural waters. 

58. Portielje, R. & Lijklema, L. Carbon dioxide fluxes across the air-water interface and its 
Supplementary Figure 1. Relationship between calcium carbonate saturation in solution and the partial pressure of CO₂ in 101 reservoirs of the Iberian Peninsula. Data were grouped following two criteria: the concentration of alkalinity and the balance of NEP. Green symbols identify those groups of samples showing dissolved oxygen supersaturation (that is, NEP imbalanced towards net autotrophy), while black symbols identify those showing dissolved oxygen undersaturation (i.e., NEP imbalanced towards net heterotrophy). Symbols represent average values for samples grouped following the same alkalinity thresholds as in Fig. 1 (low, mid, and high alkalinity). Bars are standard deviations around the mean value inside each group. Dashed lines indicate $pCO_2=350 \mu atm$ and $\Omega=0$. Calcium carbonate saturation is an index that expresses whether calcium carbonate in solution is chemically saturated or undersaturated. If $\Omega>0$, calcium carbonate is saturated in solution, and the precipitation of calcite is thermodynamically facilitated. $\Omega$ is proportional to calcium concentration times carbonate concentration, and was calculated for each of the 202 samples using observed calcium concentration and carbonate concentration calculated with CO2SYS. Note that at high alkalinity almost all samples were supersaturated in calcium carbonate and CO₂ irrespective of NEP values. By contrast, at low and moderate alkalinity the prevalence of CO₂ supersaturation conspicuously decreased when NEP was imbalanced towards net autotrophy.
Supplementary Figure 2. Seasonal evolution of CO$_2$ supersaturation in reservoirs of different alkalinity. The box-whisker plots show the median, and the 10$^{th}$, 25$^{th}$, 75$^{th}$, and 90$^{th}$ percentiles. The points represent the 5$^{th}$ and 95$^{th}$ percentiles. The dashed lines indicate the CO$_2$ concentration at equilibrium with the atmosphere. The 202 samples were grouped in the same alkalinity groups as in Fig. 1. Note that although differences between seasons were significant in the three sets of reservoirs (Mann-Whitney Rank Sum Test, P<0.007), the seasonal shift towards CO$_2$ undersaturation promoted by primary production during summer is less evident as alkalinity increases. NEP values in Fig. 2a and chlorophyll content across systems (Supplementary Data) clearly indicate that primary production in high alkalinity waters was similar than at low and intermediate alkalinity. Non significant differences in neither total phosphorus, chlorophyll concentration, or DO disequilibrium were actually observed between alkalinity groups (Kruskall Wallis tests, P>0.05). Therefore, it is unlikely that the lack of CO$_2$ undersaturation at higher alkalinity is promoted by a differential primary production.
Supplementary Figure 3. Global map of alkalinity in runoff. a, 300 m resolution map for alkalinity in accumulated runoff elaborated in this study, and location of the 584 surface water stations (filled circles) from the GEMSTAT database (www.gemstat.org) used to validate the map. b, An example of the comparison between the map of alkalinity in locally generated runoff (left panel) and the final map considering the routing of alkalinity along river networks (right panel). This example is for the Upper Amazonas (the red area roughly corresponds to the Peruvian Andes). Note how the river routing transports alkalinity from the Andes to the downstream Amazonas until inputs of low alkalinity waters dilute the signal. Color legend as in panel a. c, Comparison between observed and map-generated distributions for mean alkalinity in the 584 GEMSTAT surface water sampling stations showed in panel a.
Supplementary Table 1. Chemical characteristics of the surface layer of the reservoirs considered in this study. Data are detailed for both winter 1987 and summer 1988 sampling campaigns, and include Age (years since first filling in 1987), trophic state indicators (chlorophyll-a (Chl-a) and total phosphorus (TP) concentrations) and measurements of DO, alkalinity, and partial pressure of dissolved CO₂ (pCO₂).

<table>
<thead>
<tr>
<th>Moments and quantiles</th>
<th>Summer</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age (years)</td>
<td>Chl-a (µg L⁻¹)</td>
</tr>
<tr>
<td>Average</td>
<td>33</td>
<td>12.0</td>
</tr>
<tr>
<td>Minimum</td>
<td>13</td>
<td>0.1</td>
</tr>
<tr>
<td>25% quartile</td>
<td>20</td>
<td>1.6</td>
</tr>
<tr>
<td>Median</td>
<td>27</td>
<td>3.9</td>
</tr>
<tr>
<td>75% quartile</td>
<td>38</td>
<td>12.3</td>
</tr>
<tr>
<td>Maximum</td>
<td>104</td>
<td>178.1</td>
</tr>
</tbody>
</table>
### SUPPLEMENTARY INFORMATION

**Carbonate weathering as a driver of carbon dioxide supersaturation in lakes**

*Rafael Marcé, Biel Obrador, Josep-Anton Marguí, Joan Lluís Riera, Pilar López, Joan Armengol*

#### Supplementary Table 2. Model formulation for alkalinity fluxes considering the different global lithologic classes.

<table>
<thead>
<tr>
<th>Lithologic classes (ref. 51)</th>
<th>Alkalinity flux* (refs 49,51)</th>
<th>Definition in lithologic map (ref. 52)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconsolidated sediments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dune sands</td>
<td>0.0158 × Runoff 0.647</td>
<td>Semiconsolidated or unconsolidated sedimentary rocks</td>
</tr>
<tr>
<td>Loess</td>
<td></td>
<td>Dunes and shifting sand</td>
</tr>
<tr>
<td>Alluvial deposits</td>
<td></td>
<td>Loess</td>
</tr>
<tr>
<td>Mixed sedimentary rocks</td>
<td>0.0158 × Runoff 1.134</td>
<td>Mixed sedimentary consolidated rocks</td>
</tr>
<tr>
<td>Carbonate sedimentary rocks</td>
<td>0.0158 × Runoff 1.746</td>
<td>Carbonate rocks</td>
</tr>
<tr>
<td>Metamorphic rocks</td>
<td>0.0158 × Runoff 1.930</td>
<td>Metamorphic rocks, Precambrian basement, Complex lithology</td>
</tr>
<tr>
<td>Basic plutonic rocks</td>
<td>0.0158 × Runoff 0.724</td>
<td>Basic-ultrabasic plutonic rocks</td>
</tr>
<tr>
<td>Intermediate plutonic rocks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acid plutonic rocks</td>
<td>0.0158 × Runoff 1.559</td>
<td>Acid plutonic rocks</td>
</tr>
<tr>
<td>Basic volcanic rocks</td>
<td>0.0158 × Runoff 0.683</td>
<td>Basic and intermediate volcanic rocks</td>
</tr>
<tr>
<td>Pyroclastics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acid volcanic rocks</td>
<td>0.0158 × Runoff 0.520</td>
<td>Acid volcanic rocks</td>
</tr>
<tr>
<td>Intermediate volcanic rocks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaporites</td>
<td>0.0006 × Runoff 1.830</td>
<td>Evaporites</td>
</tr>
</tbody>
</table>

*Runoff stands for local runoff generation in mm yr⁻¹ (ref. 53)*