Inland waters such as lakes are important CO$_2$ emitter, however, accurately estimate the emissions is still a big challenge, leading to large uncertainty to the terrestrial carbon sinks estimation. Compared to the abundant lakes with diverse types, the limited sampling data with strong randomness at global, national, regional-scale contributed to large uncertainty of lake CO$_2$ emissions estimation. Satellites have accumulated a large amount of long-term data at large spatial coverage with high temporal resolution, which not only can depict the current situation but also can reconstruct the past, and then play crucial role in lake CO$_2$ emissions estimation via mapping. In this perspective, we highlight the key opportunities and challenges of lake CO$_2$ estimation with satellite. We believe that satellite remote sensing holds the potential to address the existing knowledge gap in lake CO$_2$ emissions.

Lakes link the water, soil, air, and other spheres on the surface earth and are sites of intense biogeochemical processes of terrestrial carbon. The total area of lakes, including freshwater and salt lakes, is approximately 3.2x 10$^6$ km$^2$, accounting for only 2.2% of the Earth’s land area. However, total greenhouse gas emissions from lakes is almost 20% of CO$_2$ fossil fuel emissions at global scale. The global CO$_2$ emissions from lakes spans one order of magnitude (0.07-0.81 Pg C yr$^{-1}$) with large uncertainty, which brings uncertainty to the terrestrial carbon sinks estimation. In China, without integrating the lake CO$_2$ exchange with atmosphere, the terrestrial carbon sink is overestimated by 14.1% to 22.6%, while it is underestimated by one-third on the Qinghai Tibet Plateau. The role of lakes should be integrated to advance the assessments of land carbon budget. It is well known that lake CO$_2$ emissions exhibited high spatiotemporal heterogeneity, however, the existing monitoring system is incomplete and cannot meet the frequency and large-scale observation requirements. Therefore, the current biggest issue of lake CO$_2$ research at the global and national scale is the inability to calculate the amount, that is the questions of “how much”. The great uncertainty of lake CO$_2$ emissions estimation can be attributed to the lack of accurate description of highly dynamic lake ecosystem and associated CO$_2$ changes. First, CO$_2$ emissions from lakes varied greatly across sites within lakes, among lakes between regions, and span sampling date. Limited data with poor representativeness and strong randomness is the core issue leading to the uncertainty. For example, the earliest estimates of global lake CO$_2$ emissions was 0.11-0.14 Pg C yr$^{-1}$, but the value was reported as 0.81 Pg C yr$^{-1}$ with increasing/updating data. Meanwhile, it is summarized that temporally representative data are needed to accurately estimate the CO$_2$ emission. Second, spatially explicit understanding of the sizes and distribution of global lakes is a key issue in accurately estimate the CO$_2$ emissions. A recent study shows that the CO$_2$ emissions from global lakes may be overestimated by 60% due to the overestimation of lake area. The uneven measurements across regions and lake type in the global lake datasets also accounted for the uncertainty of CO$_2$ emissions estimation. First, lakes in tropical regions are less considered, although they occupied 13% of the global lake surface area. It is estimated that tropical lakes accounted for one third of the global lake CO$_2$ emission, but they were represented by only 1.5% of the dataset used in upscaling global CO$_2$ emissions from lakes. Meanwhile, field data used for lake CO$_2$ emissions estimation of tropical regions had been obtained mainly in South America. Second, eutrophic lakes are less covered. It is reported that about 63% of lakes are eutrophic with high productivity in the world. Eutrophic lakes played significant role in CO$_2$ budget estimation, however, less than 1% of lakes in the existing dataset represent eutrophic lakes. Third, large lakes are less included. Existing global datasets mainly focused on small lakes and ponds with size less than 10 km$^2$. For example, previous collected data on lakes worldwide aimed to estimate the global CO$_2$ emissions, with more than 90% of them are small lakes. In fact, mediate and large lakes with size larger than 10 km$^2$ occupy 71% of the global lake area, they may dominate the CO$_2$ emissions due to greater overall coverage. Obviously, the lack of a broad representation of lake types and regional distribution is bound to pose challenges for accurate estimate of global lake CO$_2$ budget.

Understanding the effects of climate change and human activity on lake CO$_2$ variability is hindered by the scarcity of long-term datasets spanning over multiple decades. Longer datasets more that multi-decades are required to filter out inter-annual anomalies and disentangle adequately the role of climate and human pressures on aquatic CO$_2$ changes. Unfortunately, few long-term continuous data series (more than 10 years) have been reported until now, posing challenges to accurately quantify effects of climate and human pressures and to understanding how their contributions shift over time. It should be noted that satellite remote sensing can provide a powerful tool in obtaining long-term time series of lake CO$_2$ datasets since 1980s, which are projected to shed light on this issue.

Satellites can effectively address the gap in previous assessments of lake carbon sources and sinks. At present, there are two main approaches for estimating CO$_2$ emissions/uptake in waters at large scale using satellite remote sensing. One is simply upscaling with the satellite-derived lakes area/distribution and snapshot field-measured areal CO$_2$ flux from lakes; however, large uncertainty of the estimation occurred in upsampling as a result of considerable heterogeneity across time and space, and snapshot sampling from limited lakes are hardly capture the CO$_2$ dynamic variability; Another is to relate the CO$_2$ variability to physical and biological parameters, such as surface water temperature, salinity, and chlorophyll-a concentration, which can be retrieved from satellite imagery. This method effectively solves the problem of temporal and spatial heterogeneity of CO$_2$, which fully agrees the requirements of high spatial-temporal coverage of lake CO$_2$ monitoring. However, the existing research on inland lakes is relatively rare, compared to the extensive focus on oceanic and coastal waters. In limited cases, it was worth noting that the estimation model established in Lake Taihu, a typical shallow and eutrophic lake, was relatively mature, which has been expanded to the middle-eutrophic lakes in the middle and lower reaches of the Yangtze-Huaihe River of China. However, it is unclear whether this model can be applied to the others regional worldwide as a result of processes and factors influencing the CO$_2$ varying among lake types.

In this perspective, we highlight the key opportunities and challenges of lake CO$_2$ estimation with satellite. We believe that satellite remote sensing holds the potential to address the existing knowledge gap in lake CO$_2$ emissions. First, lakes in tropical regions are less considered, although they occupied 13% of the global lake surface area. It is estimated that tropical lakes accounted for one third of the global lake CO$_2$ emission, but they were represented by only 1.5% of the dataset used in upscaling global CO$_2$ emissions from lakes. Meanwhile, field data used for lake CO$_2$ emissions estimation of tropical regions had been obtained mainly in South America. Second, eutrophic lakes are less covered. It is reported that about 63% of lakes are eutrophic with high productivity in the world. Eutrophic lakes played significant role in CO$_2$ budget estimation, however, less than 1% of lakes in the existing dataset represent eutrophic lakes. Third, large lakes are less included. Existing global datasets mainly focused on small lakes and ponds with size less than 10 km$^2$. For example, previous collected data on lakes worldwide aimed to estimate the global CO$_2$ emissions, with more than 90% of them are small lakes. In fact, mediate and large lakes with size larger than 10 km$^2$ occupy 71% of the global lake area, they may dominate the CO$_2$ emissions due to greater overall coverage. Obviously, the lack of a broad representation of lake types and regional distribution is bound to pose challenges for accurate estimate of global lake CO$_2$ budget.

Understanding the effects of climate change and human activity on lake CO$_2$ variability is hindered by the scarcity of long-term datasets spanning over multiple decades. Longer datasets more that multi-decades are required to filter out inter-annual anomalies and disentangle adequately the role of climate and human pressures on aquatic CO$_2$ changes. Unfortunately, few long-term continuous data series (more than 10 years) have been reported until now, posing challenges to accurately quantify effects of climate and human pressures and to understanding how their contributions shift over time. It should be noted that satellite remote sensing can provide a powerful tool in obtaining long-term time series of lake CO$_2$ datasets since 1980s, which are projected to shed light on this issue.

Satellites can effectively address the gap in previous assessments of lake carbon sources and sinks. At present, there are two main approaches for estimating CO$_2$ emissions/uptake in waters at large scale using satellite remote sensing. One is simply upscaling with the satellite-derived lakes area/distribution and snapshot field-measured areal CO$_2$ flux from lakes; however, large uncertainty of the estimation occurred in upsampling as a result of considerable heterogeneity across time and space, and snapshot sampling from limited lakes are hardly capture the CO$_2$ dynamic variability; Another is to relate the CO$_2$ variability to physical and biological parameters, such as surface water temperature, salinity, and chlorophyll-a concentration, which can be retrieved from satellite imagery. This method effectively solves the problem of temporal and spatial heterogeneity of CO$_2$, which fully agrees the requirements of high spatial-temporal coverage of lake CO$_2$ monitoring. However, the existing research on inland lakes is relatively rare, compared to the extensive focus on oceanic and coastal waters. In limited cases, it was worth noting that the estimation model established in Lake Taihu, a typical shallow and eutrophic lake, was relatively mature, which has been expanded to the middle-eutrophic lakes in the middle and lower reaches of the Yangtze-Huaihe River of China. However, it is unclear whether this model can be applied to the others regional worldwide as a result of processes and factors influencing the CO$_2$ varying among lake types.

The challenges and opportunities of measuring lake carbon dioxide from space are coexisting and link to satellite synchronous observation, principles and mechanisms of remote sensing, and water color products for model development (Figure 1). Firstly, the development of universal satellite remote sensing models heavily relies on satellite-to-field matchup observation datasets from long-term accumulation and diverse types of lakes. However, CO$_2$ is not a conventional observation parameter in lakes and therefore has been neglected for a long time. More field measurements of lake CO$_2$ are needed for satellite model development. Encouragingly, new technologies such as buoy and autonomous surface vehicles can help improve the spatial-temporal resolution of in-situ measurements, contributing to more sufficient satellite-to-field matchup observations.

Secondly, the mechanisms controlling the CO$_2$ variability of lakes are different to that in ocean which generally depends on primary production. Therefore, more indirect environmental variables, beside the direct signals on satellites such as optically active components of lakes, may be required to constraint the remote sensing model. Specifically, the organic and inorganic carbon input from catchment under different geographic background, can directly or indirectly affect the various biochemical processes of CO$_2$ production and consume. Therefore, how to distinguish the impact of internal metabolic activities and external loadings
Model development
Analytical model
Regression model
Machine learning

DOC: Dissolved organic carbon
LULC: Land use and land cover

Tw: Water temperature
ZSD: Transparency
PAR CDOM: Photosynthetically active radiation
UVR: Ultraviolet radiation
CDOM: Colored dissolved organic matter

Figure 1. The framework of satellite remote sensing of lake CO₂ considering main carbon cycling processes in different types of lakes.

REFERENCES

ACKNOWLEDGMENTS
This work was supported by the National Natural Science Foundation of China (41971309, 42271114, 42301422, and U2243205).

DECLARATION OF INTERESTS
The authors declare no competing interests.