Supplements for model descriptions

Estimation of inflows

The load flowing to the box is expressed by:

$$I_j = \left(L_{\mathrm{river},j} + L_{\mathrm{piston},j} + L_{\mathrm{exch},j}\right) \tag{S1}$$

where L_{river} is the load from the inflowing rivers, L_{piston} is the load produced by extrusion flows from the upper boxes and L_{exch} is the load produced by exchange flows from the adjacent boxes (DSi amount in g hour^{-1}; diatom abundance in cm^3 hour^{-1}). The loads from groundwater input and precipitation could be negligible (Arai *et al.*, 2012). The L_{river} for DSi was calculated by the L-Q equation taken by the Kasumigaura River Office (KRO), while L_{river} for diatoms was regarded as zero.

DSi load from the inflowing rivers

The hourly DSi loads from Sakura and Ono Rivers (L_{Sakura} and L_{Ono}, respectively) were estimated by the L-Q equation based on the monitoring data for six rainfall events at the stations on the rivers (site information available at: http://www1.river.go.jp/) taken by the KRO in 2007 as follows ($N = 50$, $r^2 = 0.94$–0.98):

$$L_{\mathrm{Sakura}} = 24Q_{\mathrm{Sakura}}^{\ 0.88} \tag{S2}$$

$$L_{\mathrm{Ono}} = 42Q_{\mathrm{Ono}}^{\ 0.80} \tag{S3}$$

where Q_{Sakura} and Q_{Ono} are the river discharge on Sakura and Ono Rivers, respectively (m^3 hour^{-1}). The L-Q equation was not obtained on the other rivers. In the present study, the DSi load inflowing to the box j was estimated using the discharge ratio as the following equation:

$$L_{\text{river},\,j} = \frac{Q_{\text{river},\,j}\left(L_{\text{Sakura}} + L_{\text{Ono}}\right)}{\left(Q_{\text{Sakura}} + Q_{\text{Ono}}\right)} \tag{S4}$$

Discharge from the inflowing rivers to the box j, $Q_{\text{river},\,j}$, was estimated by:

$$Q_{\text{river},\,j} = Q_{\text{river, t}}\, q_j \tag{S5}$$

where $Q_{\text{river, t}}$ is the total river discharge from all influent rivers and the parameter q_j is the mean distribution ratio of discharge, which we determined using the water budgets in the lake as reported by Fukushima (1984) (q_1, q_2, q_3, and q_4 are 0.28, 0.32, 0.17, and 0.22, respectively). $Q_{\text{river, t}}$ was calculated based on the hourly river discharge observed on Sakura, Ono, and Koise Rivers whose catchment area of 740 km^2 accounts 52% of the entire catchment of the lake. $Q_{\text{river, t}}$ was calculated by dividing the sum of the discharges (Q_{Sakura} + Q_{Ono} + Q_{Koise}) by the catchment area ratio (0.52).

Load produced by extrusion flows

The load from the extrusion flow L_{piston} was calculated by multiplying the concentration C of the upper box by the flow rate determined on the assumption of the steady state of the water level of each box.

$$\left.\begin{aligned}
&\text{BOX 1}: L_{\text{piston, 1}} = 0\\
&\text{BOX 2}: L_{\text{piston, 2}} = 0\\
&\text{BOX 3}: L_{\text{piston, 3}} = C_1\, Q_{\text{river, 1}} + C_2\, Q_{\text{river, 2}}\\
&\text{BOX 4}: L_{\text{piston, 4}} = C_1\, Q_{\text{river, 1}} + C_2\, Q_{\text{river, 2}} + C_3\, Q_{\text{river, 3}}
\end{aligned}\right\} \text{(S6)}$$

Load produced by exchange flows

The exchange flow discharge Q_{exch} was determined by Fukushima (1984), based on both the mass balance of electric conductivity as a tracer and the flow simulation. Mean Q_{exch} between BOX 1 and BOX 3 ($Q_{\text{exch, 13}}$), BOX 2 and BOX 3 ($Q_{\text{exch, 23}}$), and BOX 3 and BOX 4 ($Q_{\mathrm{exch},\,34}$), were determined to be 3.0, 21.4, and 13.9 m^3 s^{-1}, respectively. We assumed that Q_{exch} is constant because the seasonal variation was not so large in the flow simulation by Fukushima (1984) (the coefficients of variance [CV] was less than 30%). L_{exch} was calculated by the following formulas.

$$\left.\begin{aligned}
&\text{BOX 1}: L_{\mathrm{exch},\,1} = (C_3 - C_1)\, Q_{\mathrm{exch},\,13}\\
&\text{BOX 2}: L_{\mathrm{exch},\,2} = (C_3 - C_2)\, Q_{\mathrm{exch},\,23}\\
&\text{BOX 3}: L_{\mathrm{exch},\,3} = (C_1 - C_3)\, Q_{\mathrm{exch},\,13} + (C_2 - C_3)\, Q_{\mathrm{exch},\,23} + (C_4 - C_3)\, Q_{\mathrm{exch},\,34}\\
&\text{BOX 4}: L_{\mathrm{exch},\,4} = (C_3 - C_4)\, Q_{\mathrm{exch},\,34}
\end{aligned}\right\} \quad \text{(S7)}$$

Equations of DSi release rate

The DSi release rate from SS is expressed using the following equation:

$$R_{\mathrm{SS},\,j} = \frac{B_{0,\,j}}{1+\gamma\cdot SS_{\mathrm{sed},\,j}}\left[k_{\mathrm{f}}\beta_{\mathrm{f},\,j}\exp\left(-k_{\mathrm{f}}\tau_j\right)+k_{\mathrm{o}}\beta_{\mathrm{o},\,j}\right]\exp\left[-\left(\frac{a_1}{WT_j+273.15}-\frac{a_1}{298.15}\right)\right] \quad \text{(S8)}$$

where B_0 is the BSi content of resuspended sediments ($B_{0,\,1}$, $B_{0,\,2}$, $B_{0,\,3}$, and $B_{0,\,4}$ are 0.023, 0.029, 0.041, and 0.010 g g^{-1}, respectively), γ is a constant of 1.2×10^{-4} m^3 g^{-1}, k_{f} and k_{o} are the dissolution rate constants of fresh diatom frustules (5.4×10^{-2} hour^{-1}) and old diatom frustules (2.1×10^{-4} hour^{-1}), respectively, β_{f} and β_{o} are the ratio of the BSi amount consisting of fresh frustules to the total BSi amount (0.024) and the ratio of the BSi amount consisting old frustules to the total BSi amount (0.976), respectively, τ is the elapsed time of sediment resuspension caused by strong wind (hour), a_1 is a constant of 4.2×10^3 K, and WT is the water temperature (°C).

The DSi release rate from bottom sediments was also determined by:

$$R_{\mathrm{bottom},j} = k_{\mathrm{bottom}}\left[a_2\exp\left(a_3 WT_j\right) - C_{\mathrm{DSi},j}\right] \quad \text{(S9)}$$

where k_{bottom} is the rate constant of 2.0×10^{-4} m hour^{-1}, a_2 is 4.9×10^{-4} g m^{-3}, and a_3 is 3.6×10^{-2} K^{-1}.

Estimation of input variables

SS concentration derived from the sediment resuspension

The hourly SS_{sed} was estimated by the method developed by Seki *et al.* (2006) since 1998, which uses the automatically monitored hourly turbidity and the chlorophyll *a* concentrations at the four sites taken by the KRO. To improve the accuracy, we used the linear regression model between those values and the monthly concentrations of SS and chlorophyll *a* taken by the manual sampling by the National Institute for Environmental Studies (NIES) for each year (the CV of the root mean squared error [RMSE] for SS: 29%–38%; chlorophyll *a*: 33%–42%). We subtracted the minimum value of the SS concentrations in 2004 simulated by Seki *et al.* (2006), 10 g m^{-3}, from the estimated SS_{sed} by assuming that the SS remaining in the water column (consisting mainly of clay minerals) might not attribute to the DSi release.

Water temperature

The hourly WT was estimated by a linear interpolation of the monthly water temperature at the depth of 0.5 m at the four sites obtained by the NIES.

Solar irradiance

The hourly IRR_0 was estimated by the daily irradiance values at the Tsukuba Weather

Station taken by the Japan Meteorological Agency (JMA; http://www.jma.go.jp/jma/). For estimating the hourly values, we used the hourly observed irradiance values at Kasumigaura Water Research Station recorded by the NIES to calculate the ratio of hourly to daily irradiance values during the years 1998–2010. Hourly *IRR* was estimated by multiplying the JMA irradiance by this ratio (the CV of the RMSE is 42%).

Light attenuation coefficient

The monthly K_{d} was determined by applying Lambert-Beer's law to the profiles of light intensity in the water column taken at the four sites by the NIES. The hourly K_{d} was estimated from the hourly SS_{sed} values using the relationships between SS_{sed} and the monthly observed K_{d} at three sites (except site D; r^2 = 0.36–0.50), because the correlation coefficient was low at site D (r^2 = 0.14). We therefore estimated the K_{d} at site D from that at site C (r^2 = 0.36).

Model calibration

The object function *OFUNC* was calculated using the following equations:

$$OFUNC = \overline{E}_{\mathrm{DSi}}\overline{E}_{\mathrm{diatoms}} \tag{S10}$$

$$\overline{E} = \frac{\sum_j E_j V_j}{\sum_j V_j} \tag{S11}$$

where E is the RMSE (DSi in g m^{-3}, diatoms in cm^3 m^{-3}). We calibrated the DSi concentrations in the four boxes using the observation data at the four sites, but since the diatom abundances were monitored at only two sites (A and C), we calibrated the parameters in BOX 1 and BOX 3. E_{DSi} was determined as the RMSE of the monthly DSi concentration, and we calculated E_{diatoms} by three different methods: (1) the RMSE of the monthly diatom abundance (N = 12 months × 6 years), (2) the RMSE of the annual maximum diatom abundance (N = 6), and (3) the average of (2) and the RMSE of the annual minimum diatom abundance (N = 6). We refer to the three calibration methods as CM1, CM2 and CM3, respectively.

As results, different shapes of diatom bloom patterns were obtained using the three calibration methods. Some model predictions by CM1 and 2 represented relatively flat peaks of diatom blooms which were different from the field observations, especially during the 2000s. In contrast, the predictions calibrated by CM3 showed sharp peaks of blooms which were similar to the field observations. We therefore used the results calibrated by CM3, as **Table 1** in the manuscript.