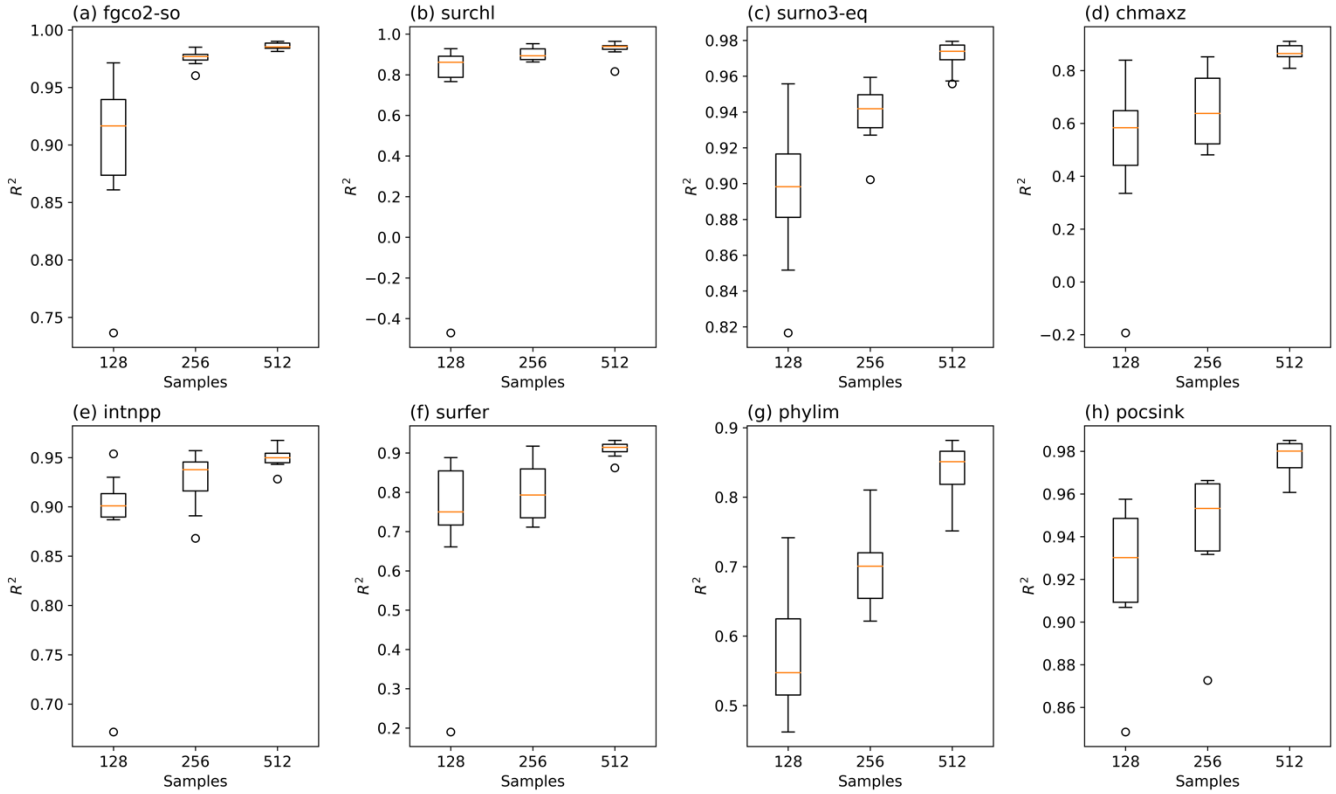


Figure S1. Sensitivity of the Gaussian Process Regression (GPR) model to true model output sample size during the sensitivity analysis. Performance of the GPR is measured as the R^2 when comparing predicted versus definite root mean square errors (RMSE). Variables are (a) downward flux of CO_2 in the Southern Ocean south of 30ºS, (b), surface chlorophyll concentrations, (c) surface nitrate concentrations in the equatorial zone between 20ºS and 20ºN, (d) the depth of the chlorophyll maximum, (e) depth-integrated net primary production, (f) surface concentrations of dissolved iron, (g) the primary limiting nutrient (iron or nitrogen) for phytoplankton growth, and (h) the sinking flux of particulate organic carbon.

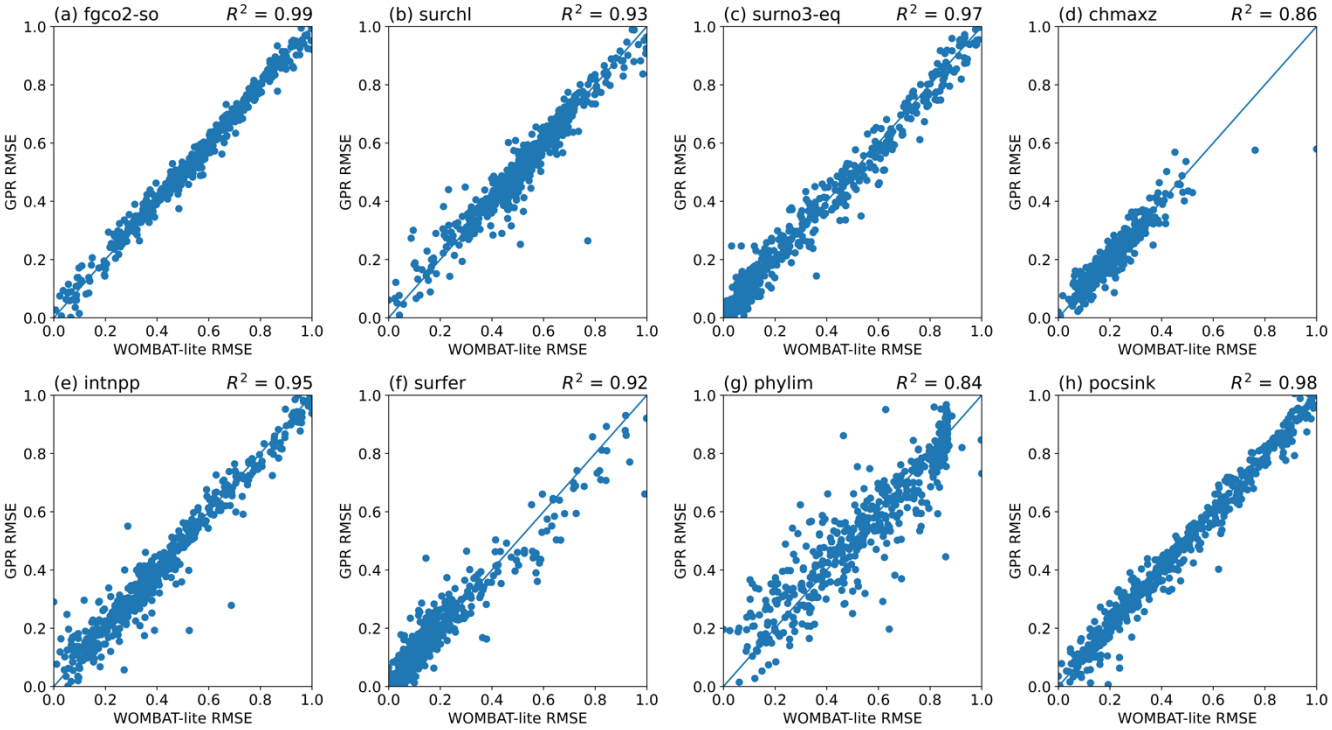


Figure S2. Accuracy of the Gaussian Process Regression (GPR) model to true model output (512 samples) during the sensivity analysis. Performance of the GPR is measured as the R^2 when comparing predicted versus definite root mean square errors (RMSE) and is shown above each panel for each target variable. Variables are (a) downward flux of CO_2 in the Southern Ocean south of 30ºS, (b), surface chlorophyll concentrations, (c) surface nitrate concentrations in the equatorial zone between 20ºS and 20ºN, (d) the depth of the chlorophyll maximum, (e) depth-integrated net primary production, (f) surface concentrations of dissolved iron, (g) the primary limiting nutrient (iron or nitrogen) for phytoplankton growth, and (h) the sinking flux of particulate organic carbon.

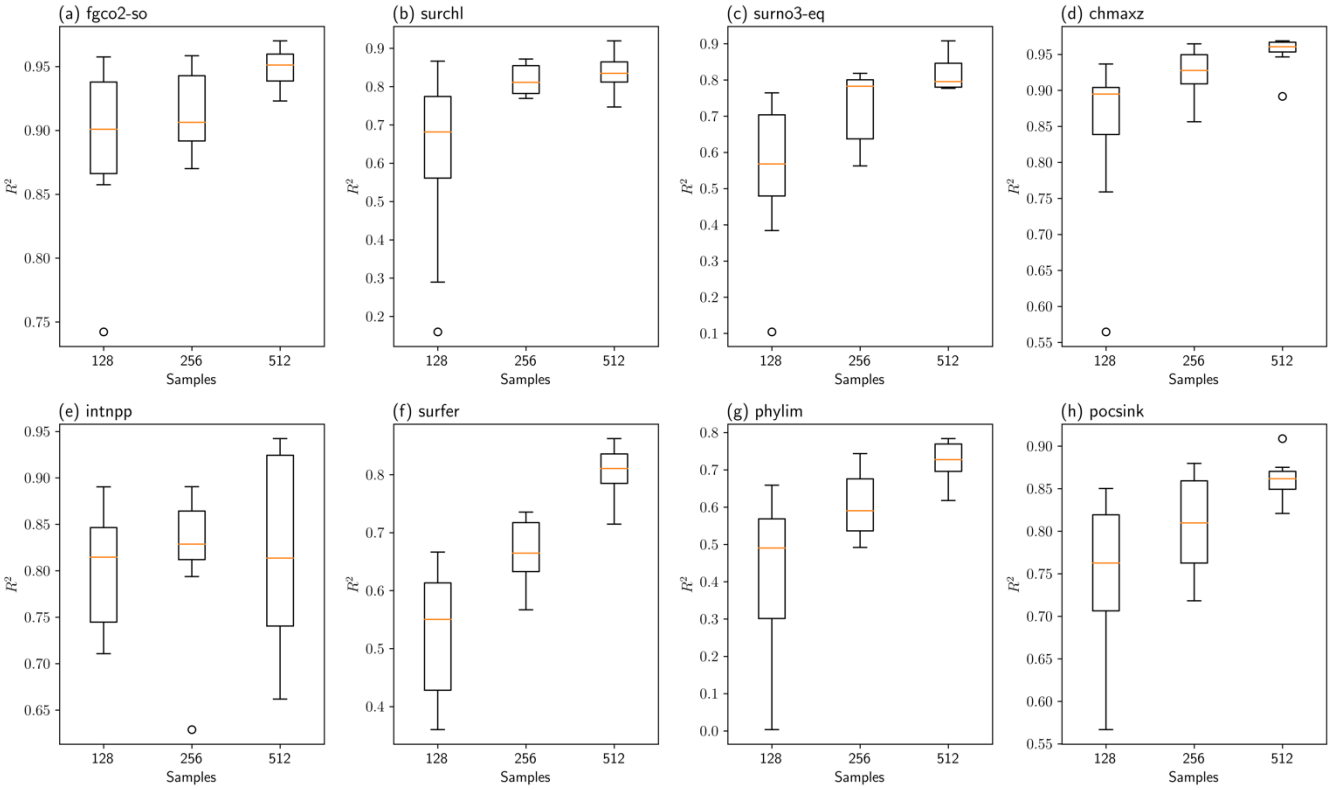


Figure S3. Sensitivity of the Gaussian Process Regression (GPR) model to true model output sample size during the optimization procedure. Performance of the GPR is measured as the R^2 when comparing the cost function of each variable. Variables are (a) downward flux of CO_2 in the Southern Ocean south of 30ºS, (b), surface chlorophyll concentrations, (c) surface nitrate concentrations in the equatorial zone between 20ºS and 20ºN, (d) the depth of the chlorophyll maximum, (e) depth-integrated net primary production, (f) surface concentrations of dissolved iron, (g) the primary limiting nutrient (iron or nitrogen) for phytoplankton growth, and (h) the sinking flux of particulate organic carbon.

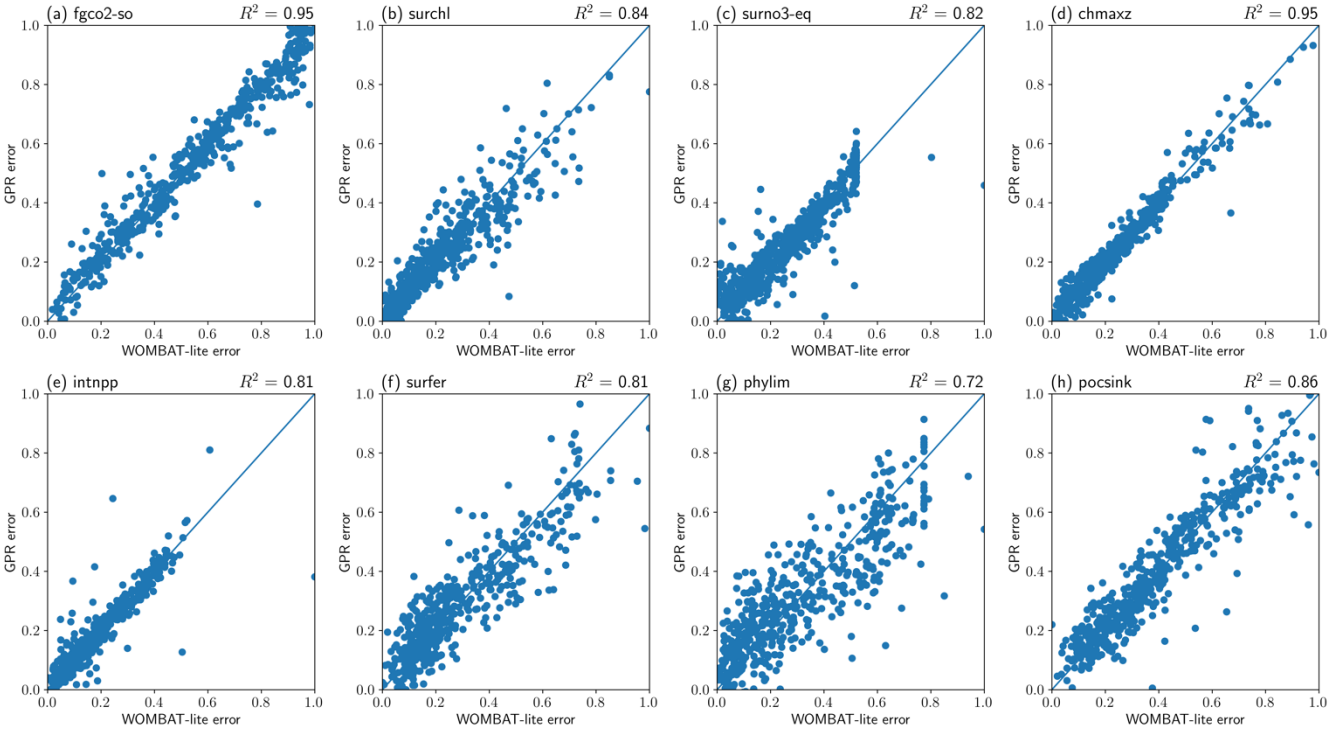
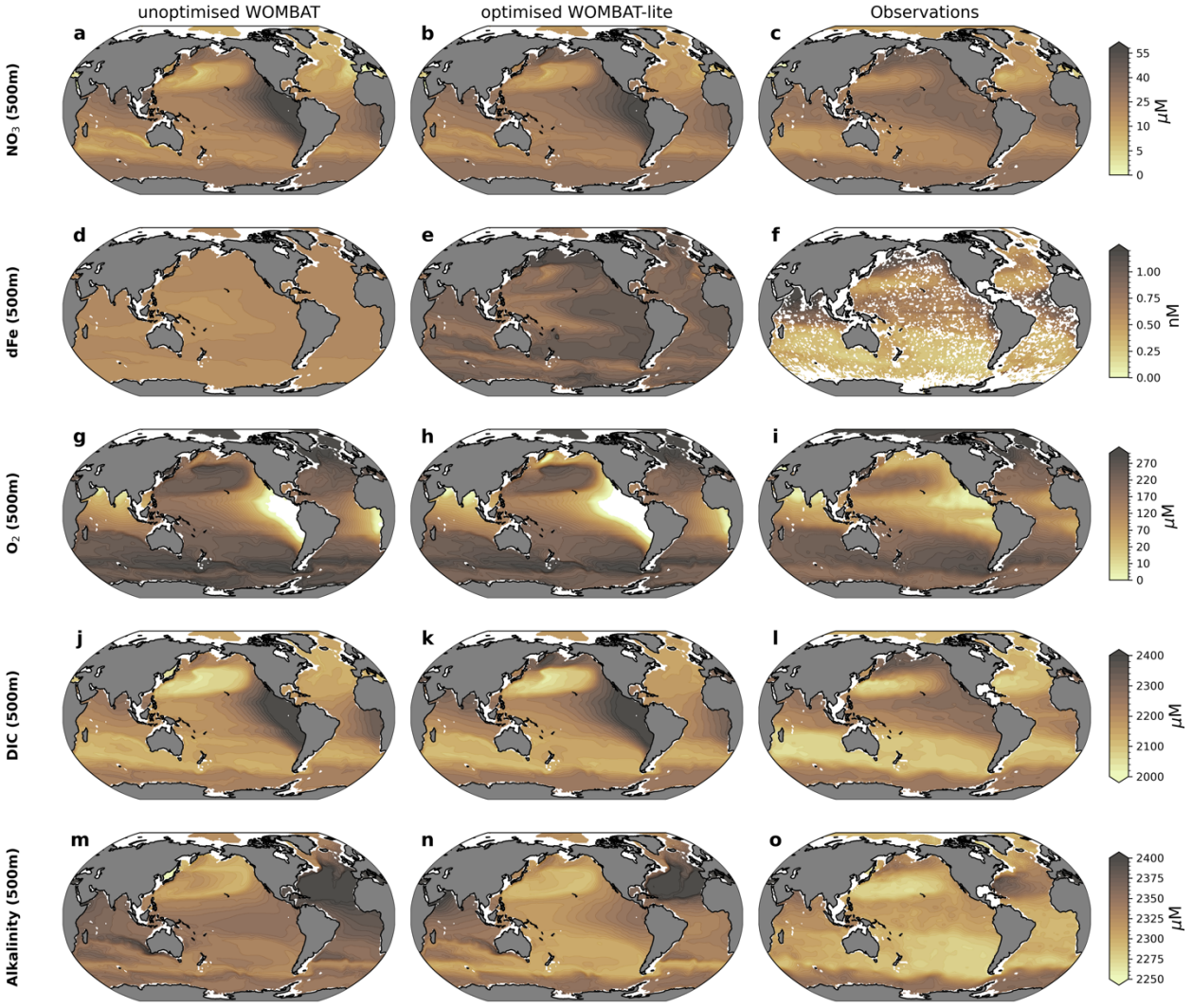


Figure S4. Accuracy of the Gaussian Process Regression (GPR) model to true model output (512 samples) during the optimization procedure. Performance of the GPR is measured as the cost function (Eq. 84) for each target variable. Variables are (a) downward flux of CO_2 in the Southern Ocean south of 30ºS, (b), surface chlorophyll concentrations, (c) surface nitrate concentrations in the equatorial zone between 20ºS and 20ºN, (d) the depth of the chlorophyll maximum, (e) depth-integrated net primary production, (f) surface concentrations of dissolved iron, (g) the primary limiting nutrient (iron or nitrogen) for phytoplankton growth, and (h) the sinking flux of particulate organic carbon.



30 **Figure S5. Comparison of the unoptimised WOMBAT (left) and the optimized WOMBAT-lite (middle) with observed distributions of major tracers at 500 metres depth (right).** (**a-c**) Nitrate concentration (μM). (**d-f**) Dissolved iron concentration (nM). (**g-i**) Dissolved oxygen concentration (μM). (**j-l**) Dissolved inorganic carbon (μM). (**m-o**) Alkalinity (μM Eq.). We show annual means of output after 100 years of forward simulation from initialization with repeat atmospheric forcing for calendar year 1990-1991 conditions using the JRA-55do (Tsujino et al., 2018). Atmospheric CO_2 set at 315 ppm in the experiments.

35