Validation metrics summary

Definitions

Assume that we have a model m and observations o of the same size N.

Root mean square deviation (RMSD)

$$ \text{RMSD}(m, o) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (m_i - o_i)^2} $$

Bias

$$ \text{Bias}(m, o) = \frac{1}{N} \sum_{i=1}^N (m_i - o_i) $$

Pearson correlation coefficient

$$ \text{Corr}(m, o) = \frac{\sum_{i=1}^N (m_i - \bar{m})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^N (m_i - \bar{m})^2 \sum_{i=1}^N (o_i - \bar{o})^2}} $$

where $\bar{m}$ and $\bar{o}$ are the mean values of m and o respectively.

Software used

Models were matched up with observational data using the Python package nctoolkit. Plots were created using the Python package nctoolkit, and the R package ggplot2. Data analysis was performed using the Python packages pandas, nctoolkit, and statsmodels.

Full domain summary statistics of model performance

Taylor diagrams for the sea surface

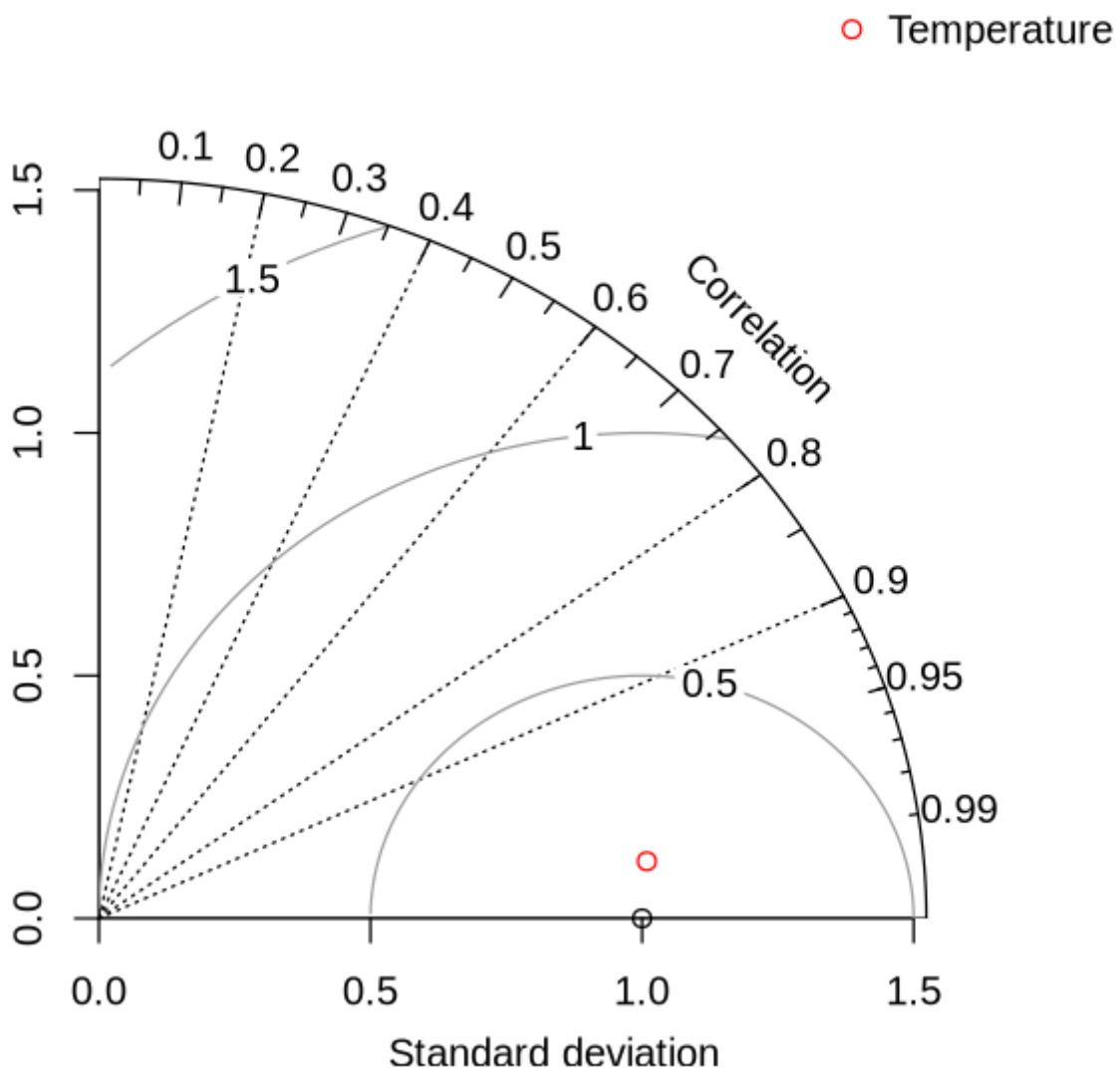


Figure 1: Taylor diagram for **sea surface** annual mean of Temperature. This diagram compares climatological annual averages of the model and observations across the model's spatial domain. Standard devaiation is normalized by the standard deviation of the observations, and a standard deviation below 1 indicates that the model is less variable than the observations. Note: This figure summarizes the overall ability of the model to reproduce climatological **spatial patterns**, and it does not represent temporal performance.

Model biases based on gridded sea surface data

Variable	Model mean	Observed mean	Model bias	Percentage bias
Temperature	19.15	18.49	0.66	N/A

Table 1: Bias of model compared with **sea surface** observations. The bias is calculated as the modelled spatial mean minus the observational spatial mean. The percentage bias is calculated as the model bias divided by the observational spatial mean.

Spatial performance of the model at the sea surface

Variable	Spatial correlation between model and observations	Spatial RMSD between model and observations
Temperature	0.99	1.51

Table 2: Pearson correlation coefficient and RMSD between model and observations at the **sea surface** for annual mean of sea temperature. This table compares climatological annual averages of the model and observations across the model's spatial domain.

Temporal performance of the model at the sea surface

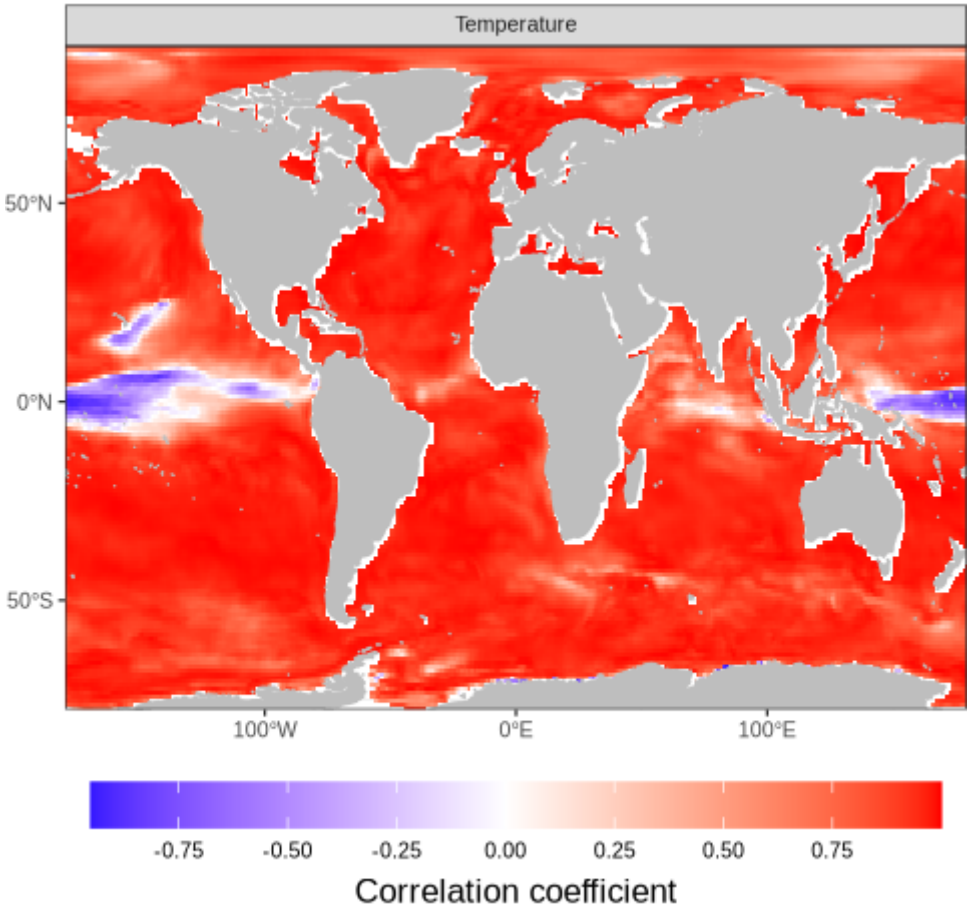


Figure 2: Spatial correlation (Pearson correlation coefficient) between model and observations for annual mean of sea temperature. This figure compares climatological monthly averages of the model and observations across the model's spatial domain.

The overall ability of the model reproduce the seasonality of each variable was estimated by calculating the spatial mean of the Pearson correlation coefficient between the model and the observations. The spatial mean was calculated by averaging the correlation coefficient of each grid cell.

Variable	***r***
Temperature	0.84

Table 3: Spatial average of the temporal correlation (Pearson correlation coefficient) between model and observations for annual mean of sea temperature. The correlation is calculated for each grid cell individually using monthly climatological averages. The spatial average is then calculated for each variable.

Sea surface sea temperature validation using gridded observations from COBE2

Matchup procedure: The model and observations were matched up as follows. First, the model dataset was cropped by a small amount to make sure cells close to the boundary were removed. The model was then regridded to the observational grid using bilinear remmaping. Only grid cells with model and observational data were maintained. The following model output was used to compare with the observational values: **tos**.

Baseline climatologies of sea surface sea temperature

Climatologies of model and observational sea surface sea temperature are shown in the figures below. The model climatology is calculated using the year **2014**.

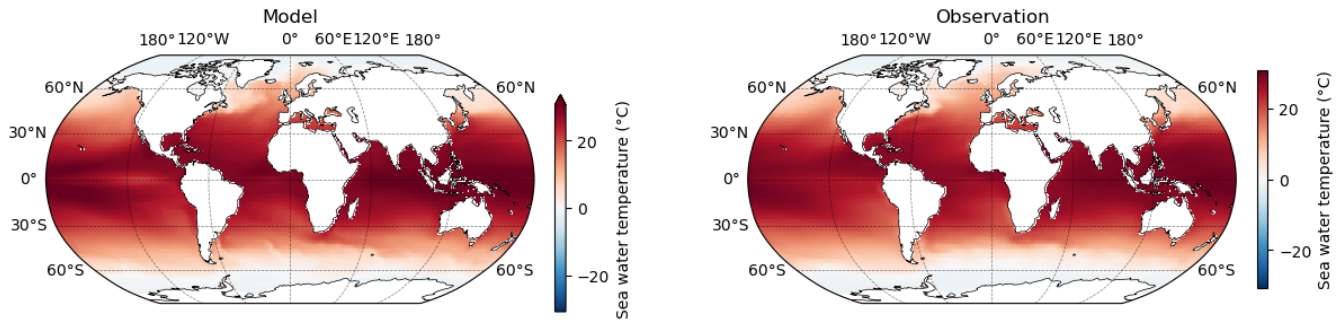


Figure 1: Annual average surface sea temperature from the model (2014) and observations. Data is limited to the 2^{nd} and 98^{th} percentile of the combined model and observational data. Arrows indicate that values can exceed the colorbar limits.

Assessing model bias for surface sea temperature

Figure 2 shows the average bias of surface sea temperature simulated by the model. A positive bias indicates that the model overestimates the observation, while a negative bias indicates that the model overpredicts the observation.

The spatial average bias of surface sea temperature is 0.66 degC. Overall, the model overestimates the observations in 74.5% of the model domain.

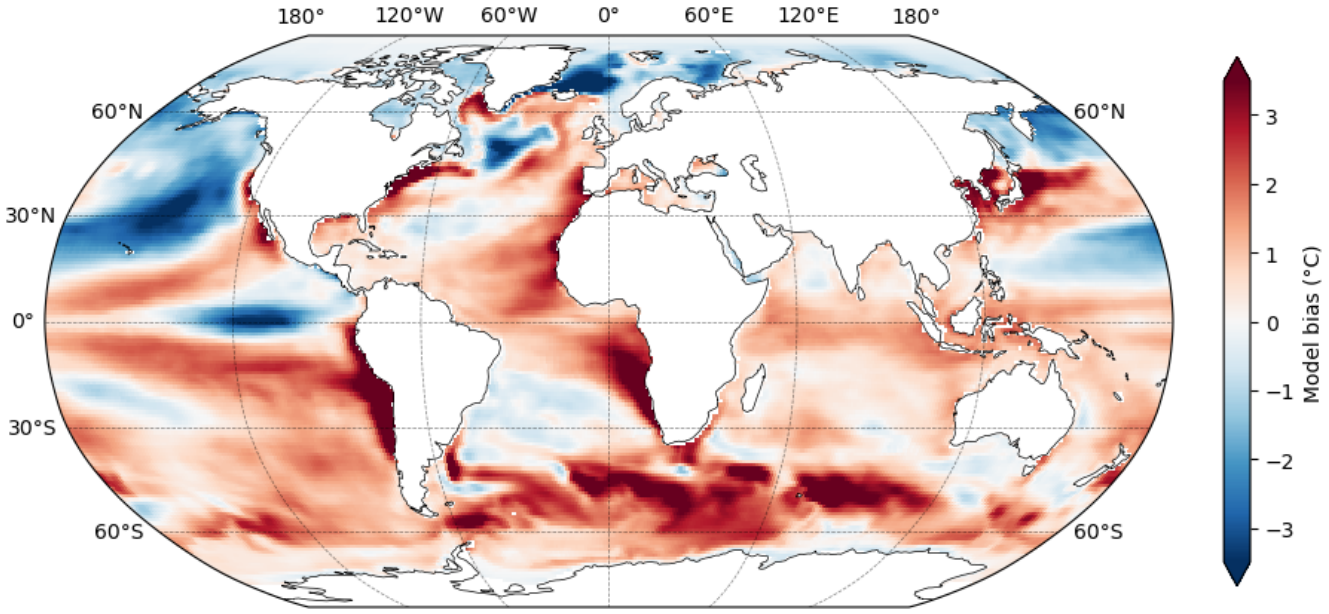


Figure 2: Bias of surface sea temperature from the model. A positive bias indicates that the model overestimates the observation. For clarity, the colorbar is limited to the 2nd and 98th percentile of the data.

Can the model reproduce seasonality of sea surface sea water temperature?

The ability of the model to reproduce seasonality of sea surface sea temperature was assessed by comparing the modelled and observed seasonal cycle of sea temperature. First, we derive a monthly climatology for the model data. Then, we calculate the Pearson correlation coefficient between the modelled and observed sea temperature at each grid cell.

Note: we are only assessing the ability of the model to reproduce the ability of the model to reproduce seasonal changes, not long-term trends.

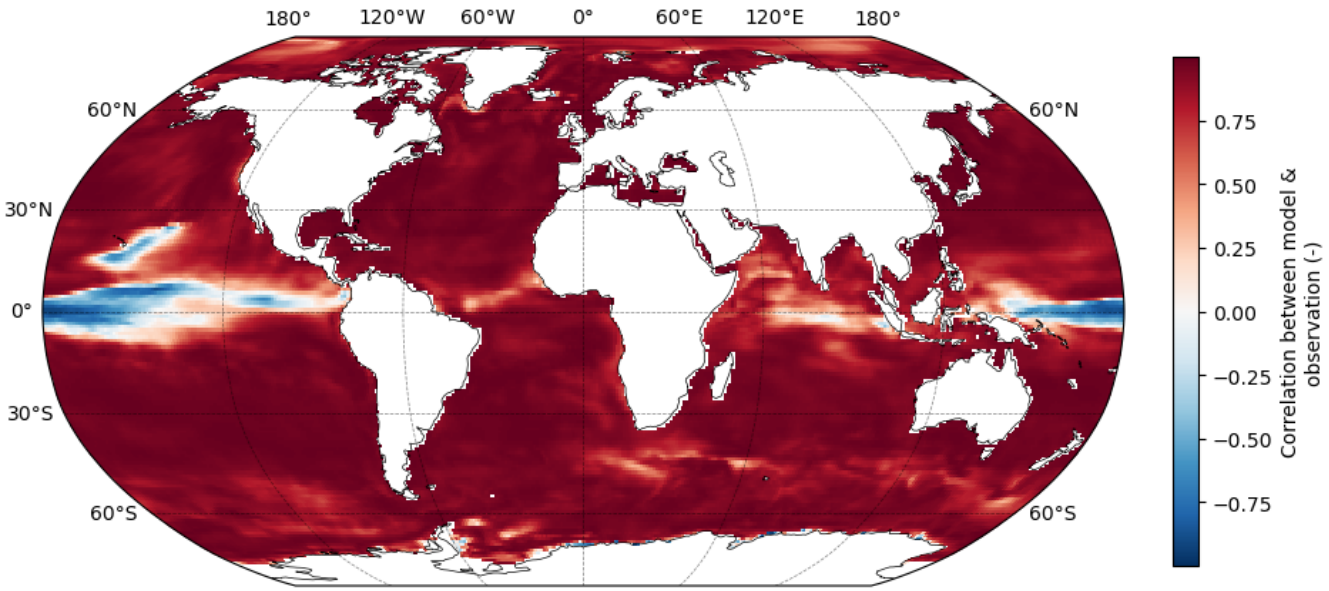
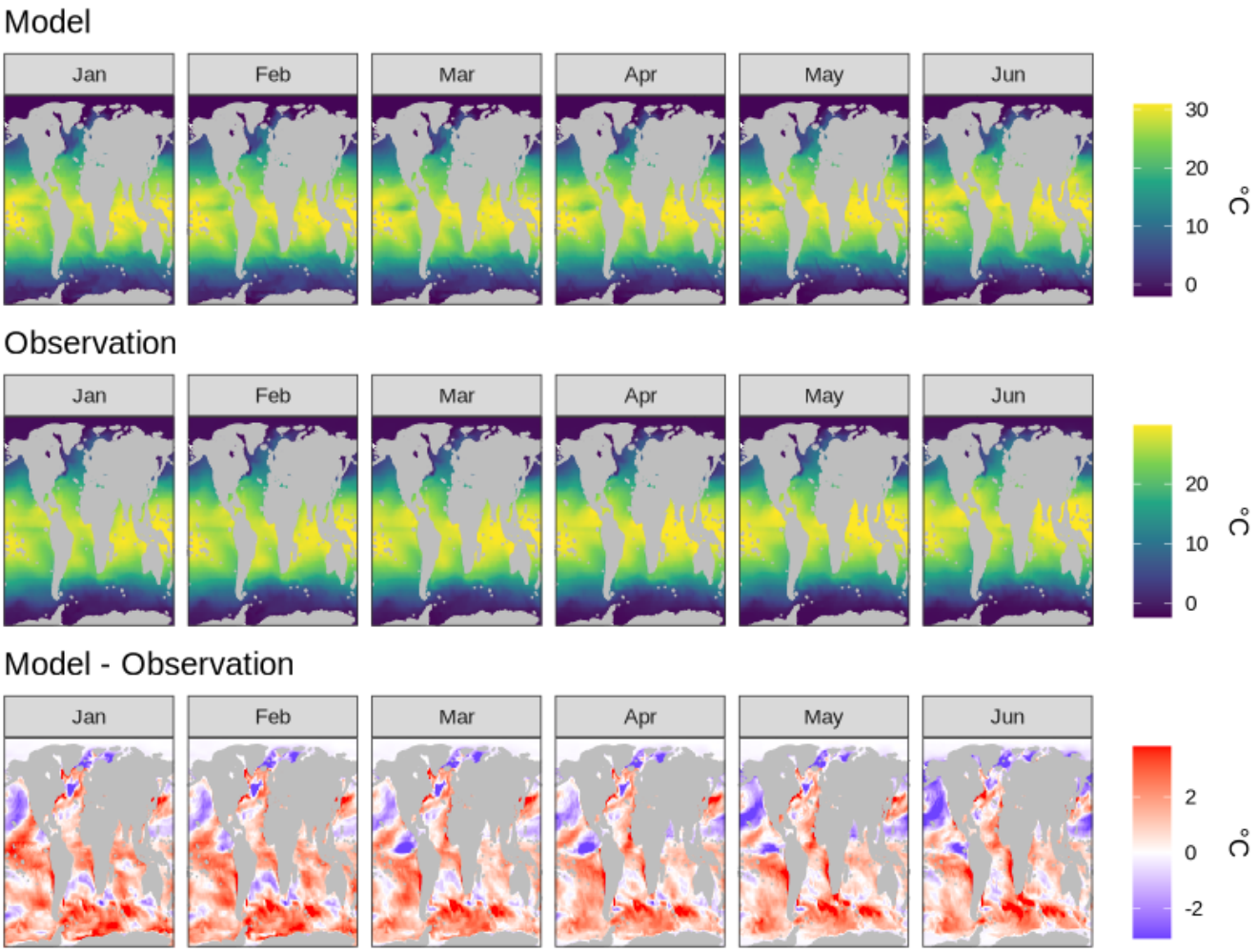


Figure 3: Seasonal temporal correlation between model and observations for surface sea temperature. This is the Pearson correlation coefficient between climatological monthly mean values in the model and observations.

The seasonal cycles of simulated and observed sea temperature are compared in Figure 4 below. This figure shows the model and observation average in each month of the year, and the differences between the two each month



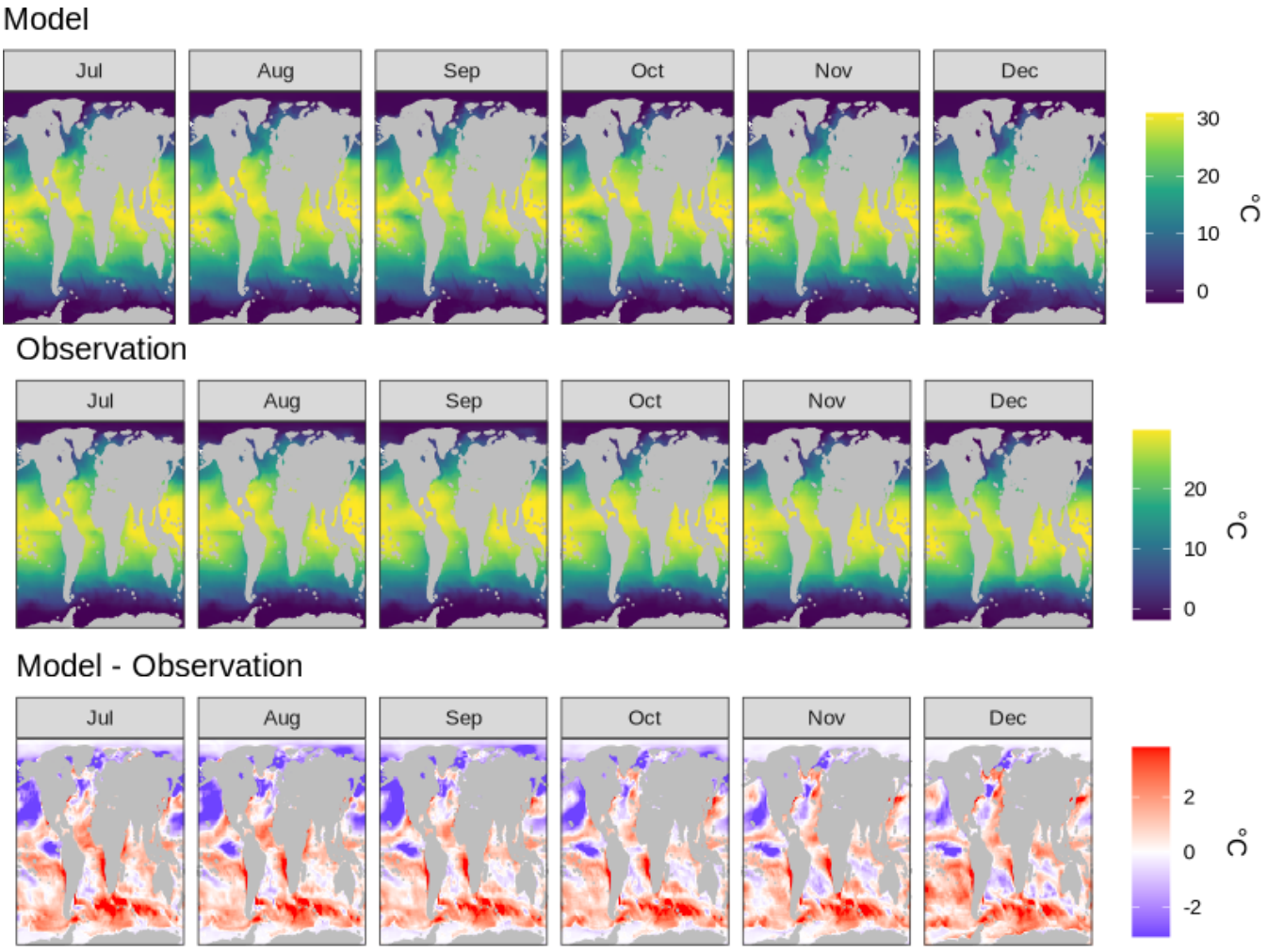


Figure 4: Monthly mean surface sea temperature for the model, observation and the difference between model and observations. For clarity, the maximum values are capped to the 98th percentiles.

Can the model reproduce spatial patterns of sea surface sea water temperature?

The ability of the model to reproduce spatial patterns of sea surface sea temperature was assessed by comparing the modelled and observed sea temperature at each grid cell. We calculated the Pearson correlation coefficient between the modelled and observed sea temperature at each grid cell.

This was carried out monthly and using the annual mean in each grid cell

Time period	***r***
Annual mean	0.99
Jan	0.99
Feb	0.99
Mar	0.99
Apr	0.99
May	0.99
Jun	0.99
Jul	0.98
Aug	0.98
Sep	0.98
Oct	0.99
Nov	0.99
Dec	0.99

Table 1: Pearson correlation coefficient between modelled and observed sea surface sea water temperature at each grid cell. The correlation was calculated monthly and using the annual mean in each grid cell.

Can the model reproduce regional spatial averages?

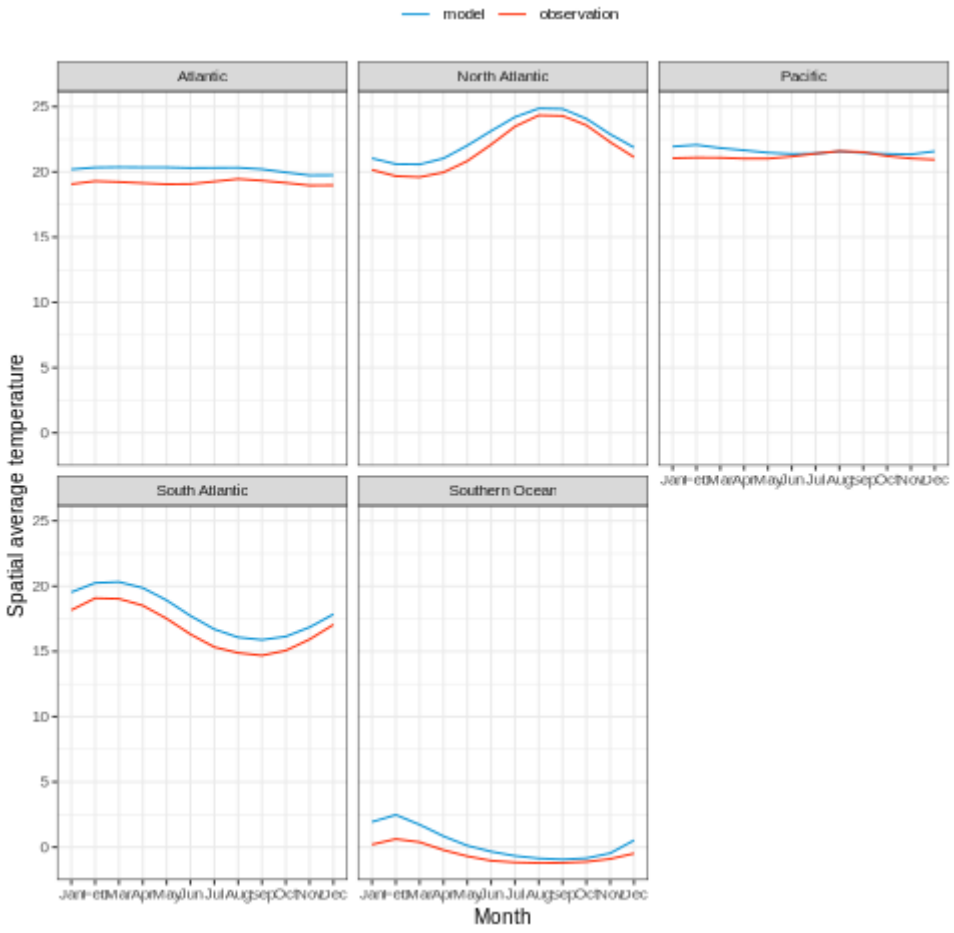


Figure 5: Climatological monthly spatial averages of sea surface sea water temperature for model and observations in each region.

Data Sources for validation of sea temperature

COBE-SST 2 and Sea Ice data provided by the NOAA PSL, Boulder, Colorado, USA, from their website at https://psl.noaa.gov/data/gridded/data.cobe2.html.