Supporting Information for "Improving GCM-based decadal ocean carbon flux predictions using observationally-constrained statistical models"

Gooya, P. [1], Swart, N. C. [1], Landschützer, P. [2]

[1]Canadian Centre for Climate Modeling and Analysis (CCCma)

[2]Flanders Marine Institute (VLIZ)

Contents of this file

Introduction The Supporting Information provides details of the statistical models used for this study, as well as the data reprocessing procedure for Earth System Model predictors. Supporting figures S1 to S7 provide complementary analysis and/or deeper insight to the main analysis of the manuscript as referenced and described in the main text.

S1. Statistical models

S1.1. Linear model

The linear model used in this study is a least square multi linear regression model. For this model, training is done on monthly mean time resolution at each grid cell on a normal one-by-one grid. The predictands are deseasonalized monthly mean ocean carbon flux time series at each ocean grid cell. For the linear model, the predictors are: SST, SSS, log(CHL), sfcWind squared, linear xCO_2 trend, and detrended xCO_2. Each of the predictors are monthly mean time series that are deseasonalized using a repeating seasonal cycle over 1990-2019 period. This combination of predictors was chosen to represent variability across different time scales. For instance, the linear atmospheric trend is the dominant driver of long term changes in ocean carbon flux, deviations of atmospheric forcing from the trend are the main drivers of the (multi)decadal variability of the sink, and other predictors are believed to drive variabilities on inter-annual to sub decadal scales. After trial and error with different combinations of our five predictors, this combination yielded best skills of reconstruction. Moreover, a repeating seasonal cycle over the period of study is removed to acquire the deseasonalized time series to reduce the variability of the variables. This showed however, to only marginally increase the skills. Finally, the training was done once with CHL and once without CHL and the results were combined with priority given to the model with CHL. This step was taken to account for possible missing CHL data point as satellite imaging of surface chlorophyll concentrations is not possible in time and space grids where clouds block the surface ocean.

S1.2. Neural Network model

NN models establish non-linear relationships between the target variable and the predictors through the use of non-linear activation functions and interconnected networks of neurons. Here, the predictant is the annual mean ocean carbon flux anomaly relative to the 1990-2019 period coming from each of the six SeaFlux data products (Fay et al., 2021). The predictors are annual mean anomalies of SST, SSS, log(CHL), sfcWind square, xCO_2 over the same period of time. These predictors are sufficient to reproduce the variability on different time scales on each data product with very high skill (Fig. S2). The NN model used in this study is a modified and simplified version of the SOM-FFN model from (Landschützer et al., 2016). The network was designed using Python Tensorflow as a dense fully connected Keras model with one hidden layer with sigmoid activation and an output layer with linear activation function. The criteria for the number of hidden layer neurons is not only minimizing the root mean square error in a randomly generated evaluation sample from training data, but more importantly, not overfitting over the forecast period, i.e., consistency of the forecast with the expected near term future behaviour of the global flux based on the evolution of the atmospheric forcing. More concisely, we already have observational references over the historical period. What we want are models that are consistent with these observation based estimates over the historical period, yet, are not overfitting to the same period of training and are extendable to future time period for actual forecasts. This is the ultimate goal of decadal prediction systems. The number of neurons was set to 15 after trial and error with a variety of neuron numbers. Comparison with the linear model where a different combination for external forcing is utilized, serve as a validation tool for the products, and against what theory suggests.

Unlike the linear model, the training resolution of the NN model is not grid scale. Here, data points are grouped into ocean biomes as used in the version 2021 of MPI-SOM-FFN product (Landschützer et al., 2020) and training is done at each biome. These biomes are acquired by a self organizing map that divides the ocean into 16 regions based on statistical similarities in the seasonal cycles of SST, SSS, mixed layer depth and surface partial pressure of CO_2. The details of the SOM-FFN method can be found in (Landschützer et al., 2016). This choice was made because grid scale resolution does not provide enough data point for the complex NN model and would end up in large overfitting. On biome scale resolution, training with monthly timeseries was very costly in terms of computational resources. Hence, annual means were used. The output of the NN model is comparable with the simple linear model both over the 1990-2019 period and for forecasts (refer to the manuscript). Finally, the method is not limited to the choice of biomes. For instance, we used (Fay & McKinley, 2014) biomes and trained the network using MPI-SOM-FFN as the target (not shown here). The results showed similar skill of reconstruction on the global scale, while differences were more detectable on regional scales. Lastly, to avoid sharp changes over the edges of the biomes, a 3-by-3 lat-lon moving window spatial smoothing was applied to the NN outputs after biomes were combined (Landschützer et al., 2016).

S2. Preprocessing of CanESM5 predictors

Except for the atmospheric CO_2 concentrations that is the same xCO_2 as seen by CanESM, when making historical, assimilation, hindcast, and forecast simulations using the statistical models, ensemble means of CanESM5 predictors from the corresponding model runs where selected. These predictors were regridded into normal one-by-one degree

resolution for compatibility. The CHL obsearvational data used for training (table S1), only extends back to 1998. To acquire estimates prior to this date (1982-1998), the time series are extended using the mean seasonal cycle of the observed period (Landschützer et al., 2016). To maintain consistency between the data that is used for training the statistical models and predictions using CanESM5 predictors, the same procedure is applied to CanESM5 CHL predictors.

Studies with ESMs have shown that initialized hindcasts simulations have biases and systematic errors when compared to the observations as a function of lead time (Kharin et al., 2012). Consequently, post processing bias correction is common practice for seasonal to decadal predictions. For each of the physical predictors and as a function of the lead time (number of years between the initialization year and prediction year), we perform a grid wise mean and trend adjustment to the corresponding observational data. The mean adjustment corrects for the mismatch between the mean over the period of the prediction at each grid cell with the mean of observations. Additionally, ESM hindcasts drift towards the preferred state of the model as represented in the historical simulation (Kharin et al., 2012). To counter this, trend adjustment based on the lead time is done to adjust for the systematic drifts of the predictors as a function of lead time. Please refer to (Kharin et al., 2012) for further details on the bias correction scheme. For CHL, only mean adjustment to the observation is applied as CHL does not exhibit a clear trend.

References

Bethke, I., Wang, Y., Counillon, F., Kimmritz, M., Fransner, F., Samuelsen, A., ... Keenlyside, N. (2019). *NCC NorCPM1 model output prepared for CMIP6 DCPP dcppA-hindcast.* Earth System Grid Federation [Dataset]. doi: 10.22033/ESGF/CMIP6.10865

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., ... Swingedouw, D. (2020). *IPSL IPSL-CM6A-LR model output prepared for CMIP6 DCPP dcppA-hindcast.* Earth System Grid Federation [Dataset]. doi: 10.22033/ESGF/CMIP6.5137

Danabasoglu, G. (2019). *NCAR CESM1-1-CAM5-CMIP5 model output prepared for CMIP6 DCPP.* Earth System Grid Federation [Dataset]. doi: 10.22033/ESGF/CMIP6.11542

Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M., ... Zeng, J. (2021). Seaflux: harmonization of air–sea co_2 fluxes from surface pco_2 data products using a standardized approach. *Earth System Science Data*, *13*(10), 4693–4710. Retrieved from `https://essd.copernicus.org/articles/13/4693/2021/` doi: 10.5194/essd-13-4693-2021

Fay, A. R., & McKinley, G. A. (2014). Global open-ocean biomes: mean and temporal variability. *Earth System Science Data*, *6*(2), 273–284. Retrieved from `https://essd.copernicus.org/articles/6/273/2014/` doi: 10.5194/essd-6-273-2014

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). En4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, *118*, 6704–6716.

Kharin, V. V., Boer, G. J., Merryfield, W. J., Scinocca, J. F., & Lee, W.-S. (2012). Statistical adjustment of decadal predictions in a changing climate. *Geophysical Research Letters*, *39*(19). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052647 doi: https://doi.org/10.1029/2012GL052647

Landschützer, P., Gruber, N., & Bakker, D. (2020). *An observation-based global monthly gridded sea surface pco2 and air-sea co2 flux product from 1982 onward and its monthly climatology.* NOAA National Centers for Environmental Information [Dataset]. (NCEI Accession 0160558, Version 6.6)

Landschützer, P., Gruber, N., & Bakker, D. C. E. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, *30*(10), 1396–1417. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GB005359 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015GB005359) doi: https://doi.org/10.1002/2015GB005359

Pohlmann, H., Müller, W., Modali, K., Pankatz, K., Bittner, M., Früh, B., ... Roeckner, E. (2019). *MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 DCPP dcppA-hindcast.* Earth System Grid Federation [Dataset]. doi: 10.22033/ESGF/CMIP6.6490

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite sst analysis for climate. *Journal of Climate*, *15*, 1609–1625.

Sospedra-Alfonso, R., Lee, W., Merryfield, W. J., Swart, N. C., Cole, J. N., Kharin, V. V., ... Sigmond, M. (2019). *CCCma CanESM5 model output prepared for CMIP6 DCPP*

dcppA-hindcast. Earth System Grid Federation [Dataset]. doi: 10.22033/ESGF/CMIP6.3557

Table S1. Observational products used for training

Variable	Source
Sea surface temperature	*NOAA OISST V*2.1 [a]
Sea surface salinity	*Hadley centre* EN4.2.1[b]
Surface Chlorophyll − a concentration	*GlobColour project*
Surface wind speed	*ERA*5
Atmospheric CO_2 *concentrations*	*NOAA ESRL*

[a] (Reynolds et al., 2002)

[b] (Good et al., 2013)

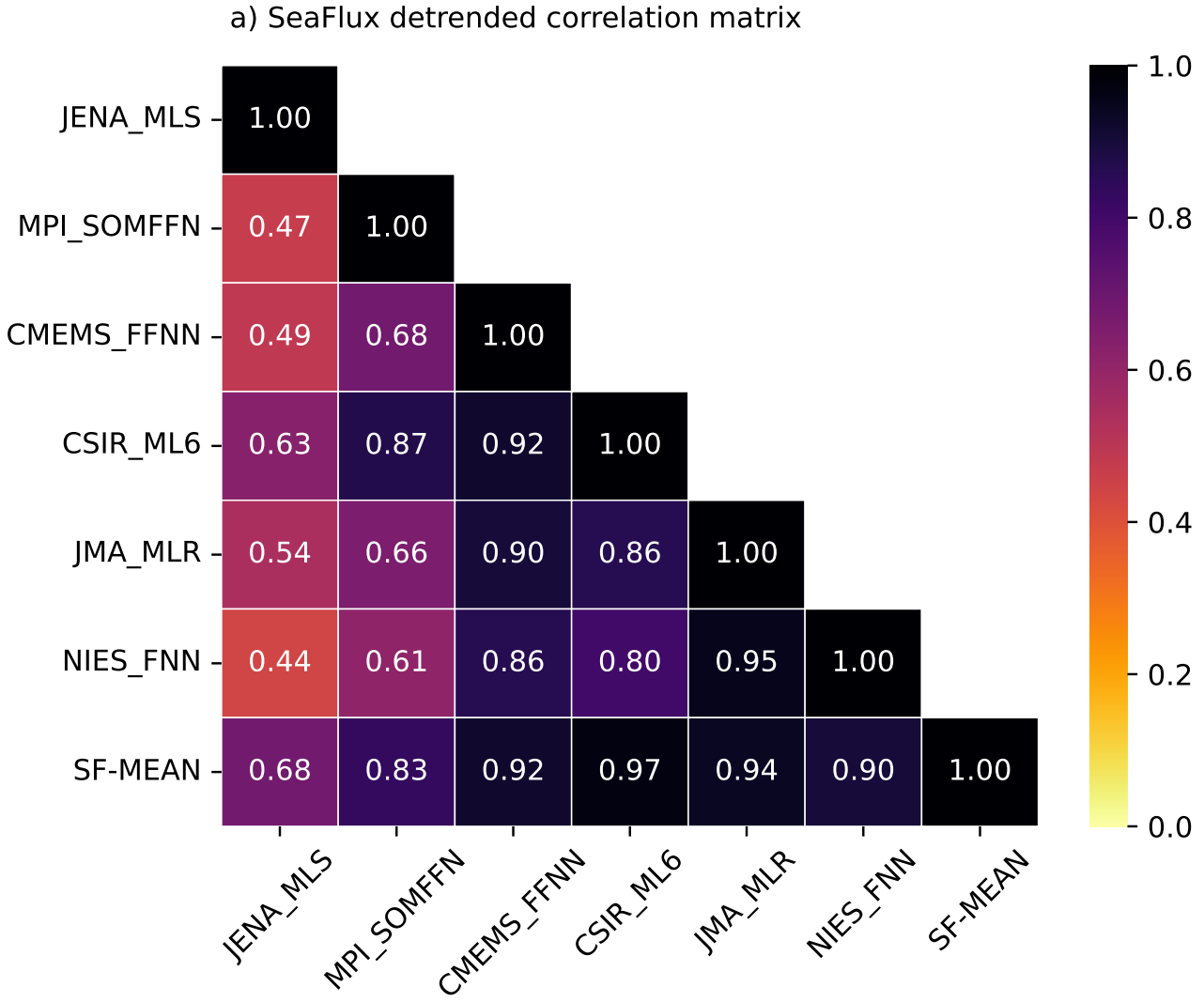

Figure S1. Cross-correlation matrix for detrended global SeaFlux observation-based ocean carbon flux products using ERA5 wind product.

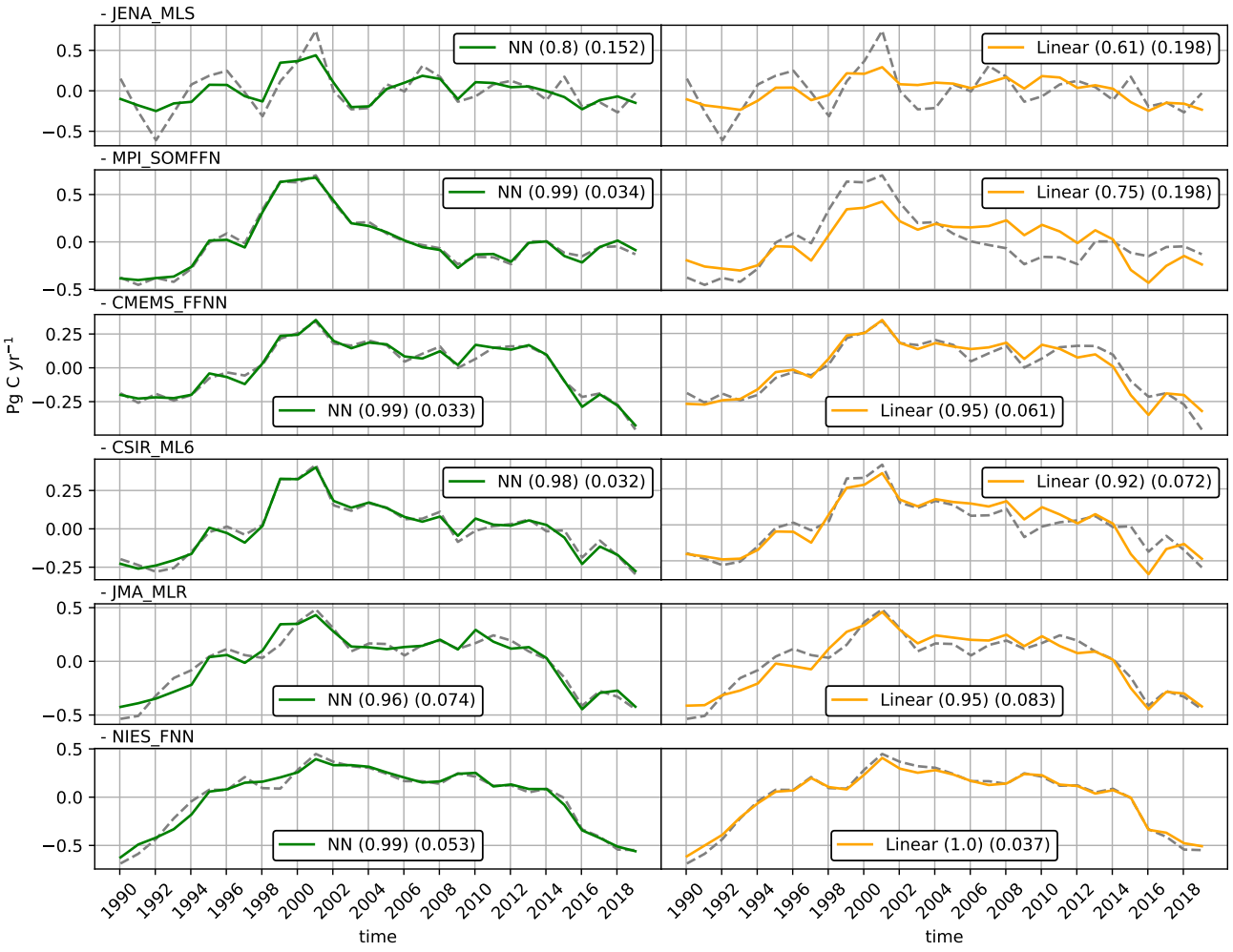

Figure S2. Time series of the detrended global ocean carbon flux reconstruction using observational predictors. Columns represent NN and linear models trained on individual products. Numbers in the legends are correlation (first number) skills versus the same product as used for training (dashed black lines), and root mean square error for the same time series (second number).

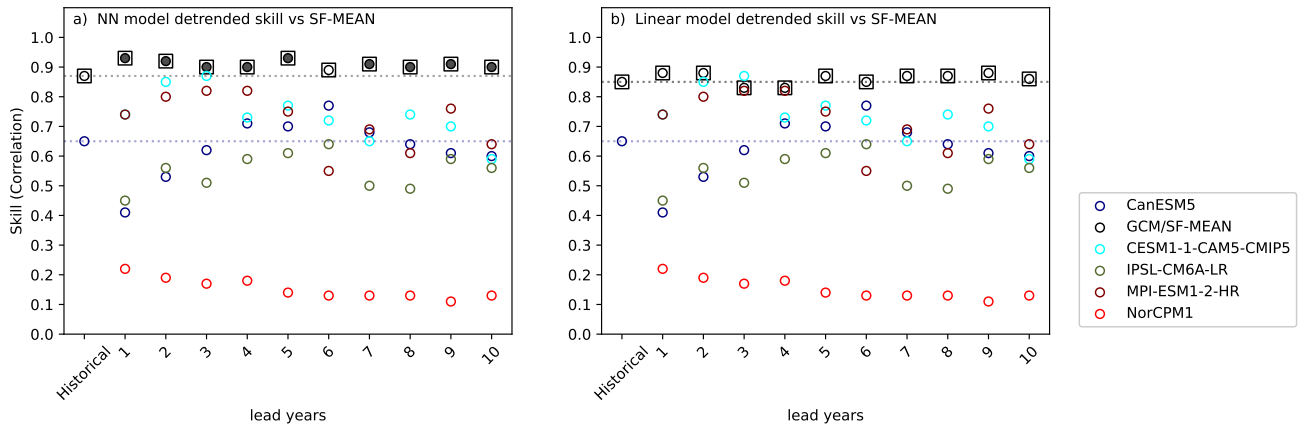

Figure S3. Detrended global ocean carbon flux skills based on bias corrected historical/hindcast predictors from CanESM5 (black dots) as well as raw CanESM5 scores (blue dots) for the hybrid model trained and evaluated using SF-MEAN. The scores that are statistically better than the raw CanESM5 score based on 1000 iteration bootstrap tests are shown with black boxes and the lead years where scores are significantly better than the historical score are filled. Colored dots are hindcast skills from ensemble means of all available CMIP6 models. The time period of this analysis is 1990-2017 as this is the common time period to all available CMIP6 models and our hybrid models.

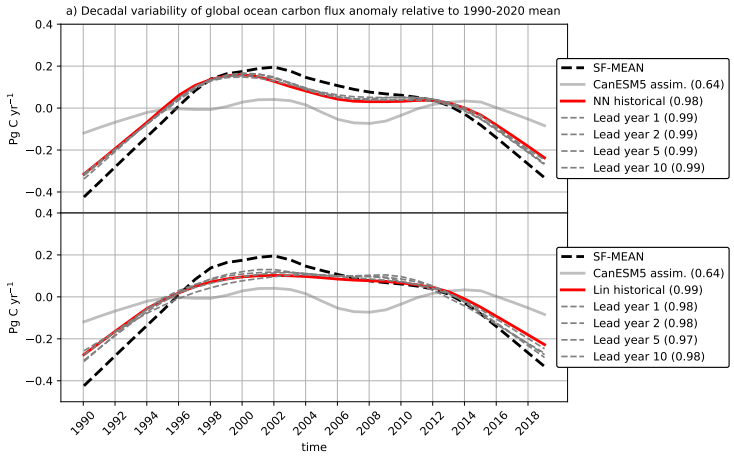

Figure S4. Time series of the smoothed detrended global ocean carbon flux simulations from NN models (top) trained on SF-MEAN for hindcast years 1,2,5, and 10 (dashed grey), historical (red), as well as CanESM5 assimilation (solid grey). Numbers are correlation coefficients with the smoothed detrended SF-MEAN (dashed black). Bottom) same as top but for linear models.

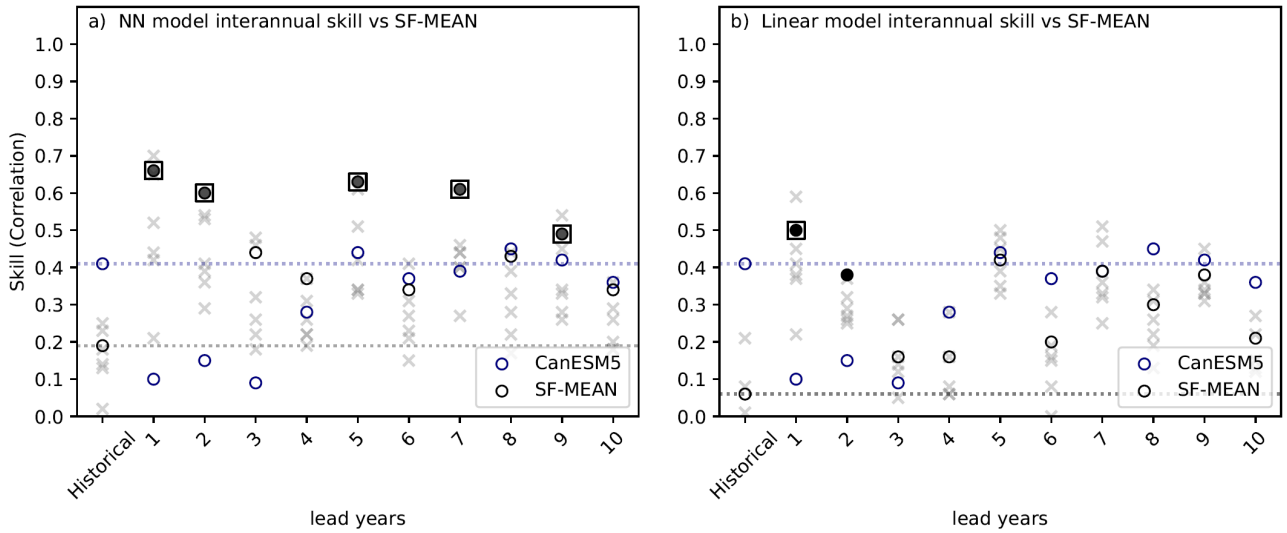

Figure S5. Interannual skills of global ocean carbon flux based on bias corrected historical/hindcast predictors from CanESM5 (black dots) as well as raw CanESM5 scores (blue dots) for the hybrid model trained and evaluated using SF-MEAN. The scores that are statistically better than the raw CanESM5 score based on 1000 iteration bootstrap tests are shown with black boxes and the lead years where scores are significantly better than the historical score are filled.

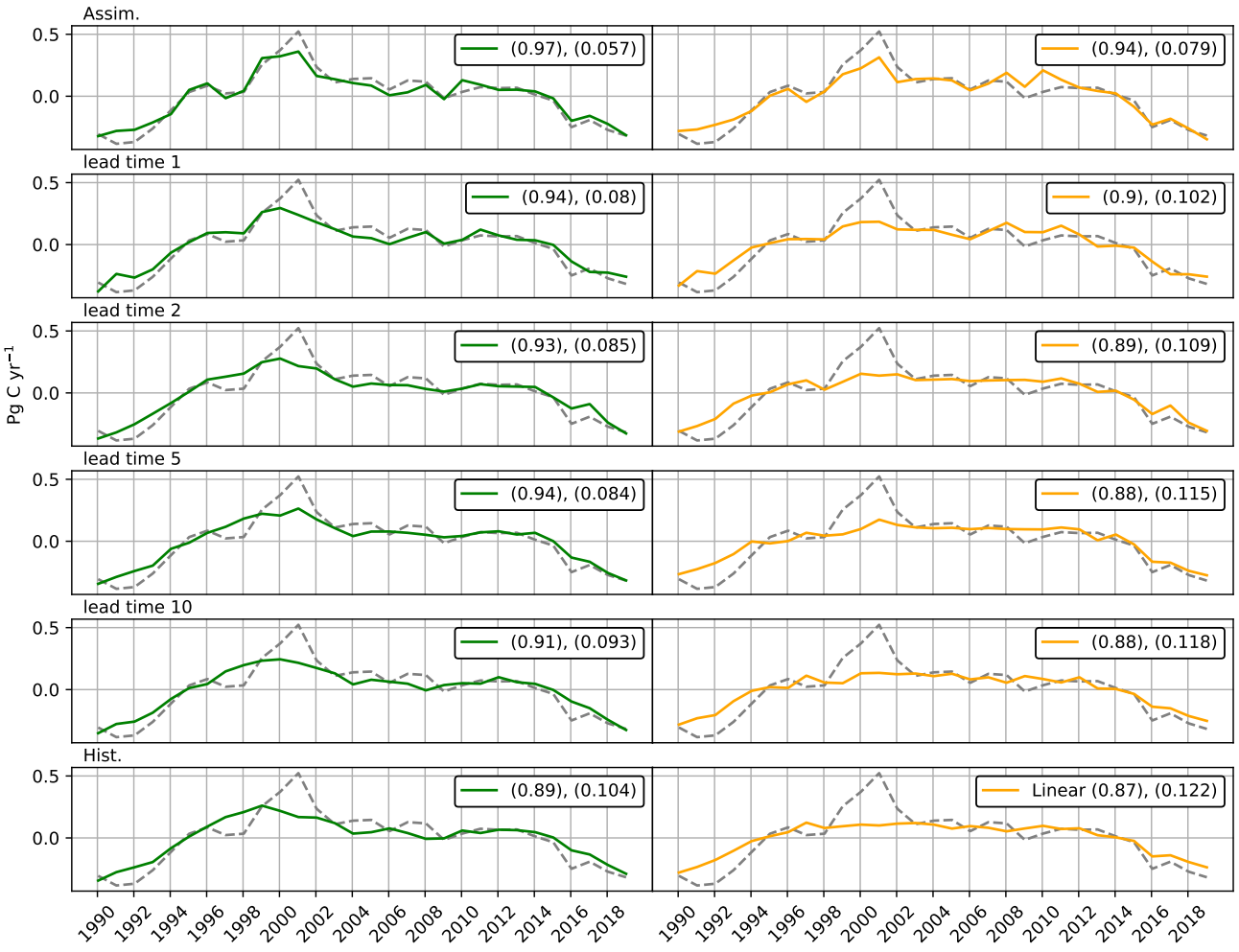

Figure S6. Detrended global ocean carbon flux time series for assimilation, hindcast years 1, 2, 5, 10, and historical simulations from NN (left column) and Linear (right column) models trained on SF-MEAN. The dashed line in the background is the detrended SF-MEAN and numbers in legends are correlation coefficients (first number) and root mean square error of the time series (second number). The plot shows how on longer lead times, the time series grow smoother and more similar to the historical time-series. They indicate less year to year variability, and are closer to the smooth (multi)decadal scale signal.

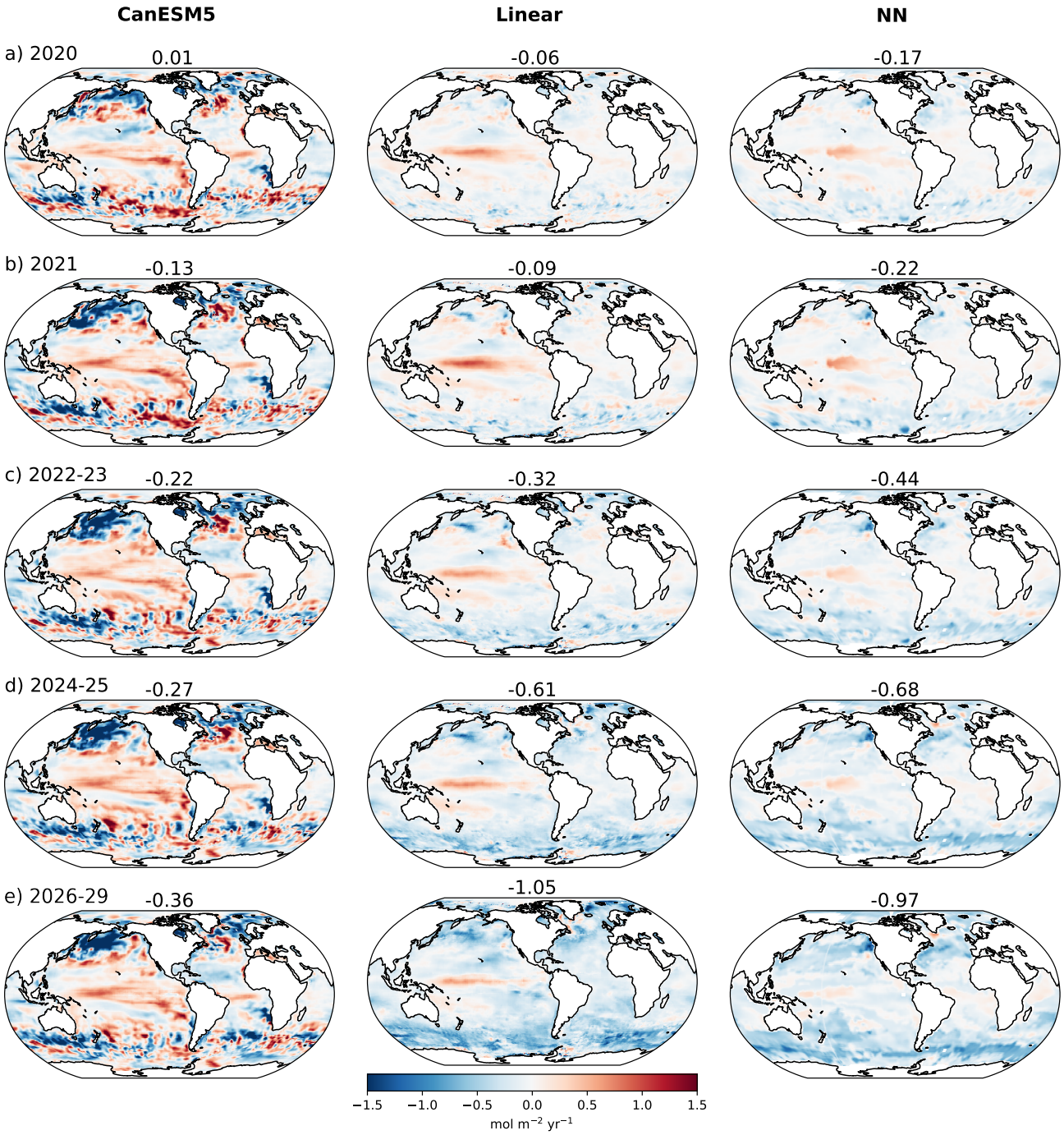

Figure S7. Regional patterns of forecasted changes in the ocean carbon flux for bias corrected CanESM5 (left column), hybrid NN model trained on SF-MEAN (middle column), and hybrid linear model trained on SF-MEAN (right column), relative to each product's 2019 projection. Numbers above each panel are global ocean carbon flux anomaly relative each product's 2019 in Pg C yr^{-1} over the same time periods of the maps.