Supporting Information for "Predictable variations of the carbon sinks and atmospheric CO_2 growth in a multi-model framework"

T. Ilyina[1], H. Li[1], A. Spring[1,2], W. A. Müller[1], L. Bopp[3], M. O. Chikamoto[4], G. Danabasoglu[5], M. Dobrynin[6], J. Dunne[7], F. Fransner[8], P. Friedlingstein[9], W. Lee[10], N. S. Lovenduski[11], W.J. Merryfield[10], J. Mignot[12], J.Y. Park[13], R. Séférian[14], R. Sospedra-Alfonso[10], M. Watanabe[15], S. Yeager[5]

[1]1Max Planck Institute for Meteorology, Bundesstraße 53, 20146 Hamburg, Germany

[2]International Max-Planck Research School of Earth System Modelling, Bundesstraße 53, 20146, Hamburg, Germany

[3]LMD-IPSL, CNRS, Ecole Normale Supérieure / PSL Res. Univ, Ecole Polytechnique, Sorbonne Université, Paris, France

[4]Institute for Geophysics, Jackson School of Geosciences, University of Texas at Austin, Austin, Texas USA

[5]National Center for Atmospheric Research, Boulder, Colorado, USA

[6]Deutscher Wetterdienst (DWD), Hamburg, Germany

[7]NOAA/OAR Geophysical Fluid Dynamics Laboratory, Princeton, NJ 08540 USA

[8]Geophysical Institute, University of Bergen, and Bjerknes Centre for Climate Research, Bergen, Norway

[9]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK

[10]Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, British Columbia, Canada

[11]Department of Atmospheric and Oceanic Sciences and Institute of Arctic and Alpine Research, University of Colorado, Boulder, Colorado, USA

[12]LOCEAN, Sorbonne Universites/IRD/CNRS/MNHN, Paris, France

[13]Department of Earth and Environmental Sciences, Jeonbuk National University, Jeollabuk-do 54896 Republic of Korea

December 15, 2020, 11:23pm

[14]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

[15]Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 3173-25, Showa-machi, Kanazawa-ku, Yokohama, Kanagawa, 236-0001, Japan

Contents of this file

1. Description of prediction systems

2. Anomaly correlation coefficient

3. Figures S1 to S5

4. Tables S1 to S2

1. Description of prediction systems

Details of the predictions systems used in this study are given below. Additionally, the Earth system models and corresponding initialization techniques are summarized in Table S1. Note that the models use different initialization and data assimilation designs. Prediction systems followed CMIP5 (historical extended by RCP4.5) or CMIP6 (historical extended by SSP2-4.5) forcing.

1.1. CanESM5

The Canadian Earth System Model version 5 (CanESM5, Swart et al. (2019)) developed at the Canadian Centre for Climate Modelling and Analysis couples version 5 of the Canadian Atmospheric Model (CanAM5) and the CanNEMO ocean component adapted from Nucleus for European Modelling of the Ocean (NEMO) version 3.4.1. CanAM5 incorporates version 3.6.2 of the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem model (CTEM), whereas CanNEMO represents ocean biogeochemistry (BGC) with the Canadian Model of Ocean Carbon (CMOC). Sea ice is simulated within the NEMO framework with the LIM2 model. CanAM5 is a spectral model with a T63 triangular truncation leading to a horizontal resolution of approximately 2.8°, and 49 hybrid vertical coordinate levels extending from the surface to 1hPa. CanNEMO is configured on the ORCA1 C-grid with 45 vertical levels ranging from about 6 meters thickness near the surface to about 250 meters in the abyssal ocean. The horizontal resolution is based on a 1 degree isotropic Mercator grid which is refined meridionally to 1/3 of a degree near the Equator, and includes a tripolar configuration to avoid the coordinate singularity in the Northern Hemisphere.

The CanESM5 ensemble of decadal hindcasts (Sospedra-Alfonso & Boer, 2020) is initialized each January 1st during 1961 to 2017 and run for 10 years. The ensemble members are initialized from separate data constrained coupled assimilation runs that span 1958 to 2016 and are started from consecutive years following a 80-year spinup run that assimilates repeating 1958-1967 data. For the ocean, the assimilation runs are nudged to 3D potential temperature and salinity from ECMWF's ORAS5 reanalysis, whereas sea surface temperature is relaxed to values interpolated from NOAA's OISSTv2 during November 1981 to 2016, and NOAA's ERSSTv3 prior 1981. Sea ice concentration is relaxed to values interpolated from HadISST.2 and the Canadian Meteorological Centre (CMC) analysis, whereas sea ice thickness uses assimilation of the SMv3 statistical model of Dirkson, Merryfield, and Monahan (2017). Atmospheric full-field temperature, horizontal wind components and specific humidity are nudged toward values from ERA-Interim during 1979 to 2016, and to ERA40 anomalies added to ERA-Interim climatology prior 1979. Land physical and BGC variables are initialized through response of CLASS-CTEM to the data-constrained atmosphere, whereas oceanic BGC variables are initialized through response of CMOC to data-constrained physical ocean variables and surface atmospheric forcing. CanESM5 uninitialized predictions are historical simulations extended after 2014 with SSP2-4.5 forcing scenario. Although CanESM5 incorporates CTEM and CMOC to simulate land and ocean carbon exchange with the atmosphere, initialized and uninitialized predictions have prescribed atmospheric CO_2 concentrations and thus ocean and land CO_2 are purely diagnostic without feedback onto the simulated physical climate.

1.2. CESM-DPLE

The CESM (Community Earth System Model) Decadal Prediction Large Ensemble (DPLE) is a collection of 40-member decadal hindcasts/forecasts using CESM version 1.1, run with prognostic ocean biogeochemistry (Yeager et al., 2018). The component models include: CAM5 atmosphere (nominal 1° with 30 vertical levels); POP2 ocean (nominal 1° with 60 vertical levels); CICE4 sea ice (nominal 1°, same horizontal grid as ocean); and CLM4 land (nominal 1°, same horizontal grid as atmosphere). The ocean biogeochemistry model used in CESM-DPLE has been described in detail elsewhere (Lindsay et al., 2014; Yeager et al., 2018; Lovenduski et al., 2019). The corresponding uninitialized historical simulations comprise the CESM Large Ensemble (CESM-LE; Kay et al. (2015)), 34 members of which include the biogeochemical fields of interest here.

The CESM-DPLE hindcasts are initialized on each November 1 between 1954-2015 and integrated for 122 months. The initial conditions for ocean and sea ice fields (including ocean biogeochemical fields) come from a coupled ocean-sea-ice historical reconstruction simulation forced with atmospheric reanalysis data combined with satellite-based flux data (Yeager et al., 2018). There is no assimilation of ocean or sea-ice observations in this reconstruction. Initial conditions for the atmosphere and land come from a randomly selected, single member of the CESM-LE. Full-field initialization is used, necessitating a drift correction step prior to analysis.

1.3. GFDL-ESM2

GFDL-ESM2 developed at the Geophysical Fluid Dynamics Laboratory builds on the fully coupled GFDL-CM2.1 atmosphere-land-sea ice-ocean model. The resolution of the

atmosphere and land models is 2.5° longitude x 2° latitude with 24 hybrid sigma/pressure vertical layers in the atmosphere. The resolution of ocean and ocean biogeochemical model is 1.0°, with telescoping to 1/3° near the equator, with 50 vertical levels with varying thickness ranging from 10 m near the surface to 400 m in the deep ocean. The physical ocean model incorporates the ocean biogeochemistry component, the GFDL's Carbon, Ocean Biogeochemistry and Lower Trophics (COBALT) marine biogeochemical model that simulates 33 tracers to resolve global-scale biogeochemical cycles. Note that this model version is different from ESM2M in key ways (different land model, different ocean code base and settings). For the purpose of this study we took the name used in the previous ocean biogeochemistry predictability studies by Park et al. (2018, 2019) which used the same model version.

The ESM prediction system using GFDL-ESM2 comprises 3 sets of simulations: i) An ensemble of 12-member uninitialized historical simulations, ii) assimilation run constrained by 3-D ocean in situ data and atmospheric reanalysis product, and iii) an ensemble of 10 year-long with 12 ensemble retrospective prediction runs initialized from the assimilation during the period 1961-2017. The initial conditions of ensemble prediction are taken from the GFDL ensemble coupled data assimilation (ECDA) system coupled with COBALT (Park et al., 2018, 2019). The ECDA system employs an ensemble Kalman filter (EKF) assimilation scheme. The ocean in the ECDA is constrained by satellite-retrieved surface temperature from NOAA optimum interpolation sea surface temperature v2 and in situ ocean temperature/salinity from World Ocean Database (WOD) and Argo profiles since 2000. The atmosphere in the ECDA is constrained by the 6 hourly temperature

and winds from National Centers for Environmental Prediction, Department of Energy (NCEP-DOE) Reanalysis 2 product. The ocean and atmosphere data constraints in the assimilation run are optimally calibrated to reduce spurious equatorial upwelling and the subsequent biogeochemical biases caused by assimilation-driven momentum imbalances (Park et al., 2018).

1.4. IPSL-CM6A-LR

The IPSL (Institut Pierre Simon Laplace) decadal prediction system used here is based on the IPSL-CM6A-LR version of the climate model described extensively in (Boucher et al., 2020). The component models include: LMDZ6 atmosphere (average 157km, 79 levels), NEMOv3.6STABLE ocean on the ORCA1 grid (nominal 1° with 75 vertical levels), LIM3 sea ice (on the same grid as the ocean) and ORCHIDEE (Cheruy et al., 2019) land (same grid as the atmosphere). The ocean biogeochemistry model used in IPSL-CM6A-LR is based on PISCESv2 (Aumont et al., 2015).

The uninitialized historical simulations comprises 32 members. The hindcasts are initialized from a global century long simulation in which anomalies of global EN4 sea surface temperature and Atlantic sea surface salinity presented by Reverdin et al. (2019) have been nudged into the climate model. The nudging procedure is described by Estella-Perez, Mignot, Guilyardi, Swingedouw, and Reverdin (2020). There is no assimilation of subsurface ocean, sea ice or atmospheric observations. Hindcasts start each December 1 during 1961-2014 and integrated for 10 years; 10 members are launched for each start date.

1.5. MIROC-ES2L

MIROC-ES2L (Hajima et al., 2020) is an ESM developed for CMIP Phase 6. The physical core of MIROC-ES2L is MIROC5.2 (Tatebe et al., 2018). The ocean biogeochemical component is OECO2, and the land biogeochemical component is VISIT-e, which has the same horizontal grid as atmosphere. Details of OECO2 and VIST-e are described in Hajima et al. (2020). The horizontal resolution of the atmospheric component has T42 spectral truncation (i.e., approximately 300 km) with 40 vertical levels up to 3 hPa. The oceanic component has a horizontal tripolar coordinate system. In the spherical coordinate portion south of 63°N, the longitudinal grid spacing is 1°, while the meridional grid spacing varies from approximately 0.5° near the equator to 1° in mid-latitude regions. There are 62 vertical levels in a hybrid σ–z coordinate system, the lowermost of which is located at the depth of 6300 m.

Using MIROC-ES2L, we conducted three sets of experiments, namely, uninitialized historical runs in 1850–2014, data assimilation runs in 1960–2016, and retrospective predictions starting from 1 January, every year from 1980 to 2017 with 10-yr-duration. All the experiments have ten ensemble members. In the assimilation run, the monthly objective analysis of ocean temperature and salinity (Ishii & Kimoto, 2009) at the depths between the sea surface and 3000 m are assimilated. The assimilation procedure is the same as that used in Tatebe et al. (2012) and Watanabe et al. (2020), but a full-field assimilation is adopted. In addition to that, absolute values of monthly sea-ice concentration based on satellite observations of Armstrong, Knowles, Brodzik, and Hardman (1994) are assimilated with the same procedure for ocean temperature and salinity. The

atmosphere is constrained by full-field assimilation with the 6 hourly temperature and winds from the JRA55 reanalysis (Kobayashi et al., 2015). With an initial condition at 1 January 1960 taken from a certain member of the historical runs, the spinup run for the assimilation run is conducted over a few hundred years with the fixed CMIP6 external forcings at the year 1960 until the air–sea and air–land carbon fluxes reach a quasi-steady state. Initial conditions for ten member ensemble of the assimilation runs are taken from arbitrary years of the spinup run.

1.6. MPI-ESM-LR

MPI-ESM-LR is a low-resolution configuration of the Max Planck Institute Earth System Model (MPI-ESM1.1; Giorgetta et al. (2013)), on which the coupled model intercomparison project phase 5 (CMIP5) simulations are based. The resolution of the ocean model MPIOM is about 150km with 40 vertical levels. The resolution of the atmosphere model ECHAM is T63 (200km) with 47 vertical levels. The ocean biogeochemistry component of MPI-ESM is represented by HAMOCC (Ilyina et al., 2013), and the land and vegetation component is represented by JSBACH.

The decadal prediction system comprises 3 set of simulations, i.e., i) an ensemble of 10-member uninitialized historical simulations extended to the RCP4.5 scenario; ii) assimilation run by nudging the ocean 3-D temperature and salinity anomalies from the ECMWF ocean reanalysis system 4 (ORAS4) (Balmaseda et al., 2013) and the atmospheric 3-D full-field temperature, vorticity, divergence, and surface pressure ECMWF Re-Analysis ERA40 (Uppala et al., 2005) during the period 1960-1989 and ERA-Interim (Dee et al., 2011) during the period 1990-2014; iii) An ensemble of 10-member retrospective predic-

tion simulations initialized from the assimilation which is close to the observations, the initialized prediction simulations run for 10 years starting annually from 1st January for the period 1961-2014. There is no assimilation of ocean biogeochemical data due to the limit of available data.

1.7. MPI-ESM1.2-HR

MPI-ESM1.2-HR is based on a latest MPI-ESM model version 1.2 (Müller et al., 2018; Mauritsen et al., 2019), which is used for CMIP6 simulations. The major model development in the physical climate components relative to the CMIP5 model versions is the new radiation and aerosol parameterizations (Mauritsen et al., 2019). The representation of the land vegetation component is extended by including wild fires, multi-layer soil hydrology scheme, and nitrogen cycle. A major development to the ocean biogeochemistry is the implementation of cyanobacteria as additional phytoplankton specie for prognostic representation of nitrogen fixation, improved detritus settling and a number of other refinements. MPI-ESM1.2-HR is configured with grid spacings of 40 km in the ocean and T127 (100 km) in the atmosphere, with 40 ocean vertical levels and 95 atmospheric vertical levels, respectively. The assimilation in the MPI-ESM1.2-HR decadal prediction system is in general the same as in the MPI-ESM-LR prediction system for the atmosphere and the ocean, the difference for nudging is the inclusion of assimilation of sea-ice concentration from the National Snow and Ice Data Center (NSIDC) satellite observations (as described in Bunzel, Notz, Baehr, Müller, and Fröhlich (2016)). In addition, we run a pre-assimilation to spinup the ocean biogeochemistry for about 50 years before the assimilation so that the ocean biogeochemical processes slowly adjust to the new assim-

ilated physical climate states (Li et al., 2019). The ensemble member for the initialized predictions and uninitialized historical of MPI-ESM1.2-HR simulations is 10. Note that the initialized 10-year long predictions of MPI-ESM1.2-HR system start annually from November 1 for the period 1960-2018.

1.8. NorCPM1

The latest version of the Norwegian Climate Prediction Model (NorCPM1) builds on NorCPM (Counillon et al., 2014, 2016; Wang et al., 2016, 2017), with the inclusion of an ocean biogeochemical component, atmospheric aerosol and cloud chemistry, an update to CMIP6 forcing, a retuning and some minor bug-fixes. The ESM in NorCPM1 is based on the CMIP5 version of the medium resolution Norwegian Earth System Model NorESM1-ME (Bentsen et al., 2013; Tjiputra et al., 2013), where the atmospheric, ocean physical, ocean biogeochemical, sea ice and land components are represented by CAM4-OSLO, NorESM-O, HAMOCC, CICE4 and CLM4, respectively. CAM4-OSLO has a 1.9x2.5 ° latitude-longitude resolution and 26 vertical layers. The ocean component NorESM-O has a 1° horizontal resolution and consists of 51 isopycnic layers, where the two uppermost layers represent the mixed layer.

NorCPM1 applies an Ensemble Kalman Filter to assimilate monthly anomalies of sea surface temperature (SST) and temperature and salinity depth profiles. For 1950-2010 and 2011-2018, SST from the HadISST2 (HadISST2.1.0.0) and OISSTV2 (Reynolds et al., 2002) datasets, respectively, are used. The temperature and salinity depth profiles come from the EN4.2.1 dataset (Good et al., 2013). Based on Fransner et al. (2020), who showed that the biogeochemical initial conditions have a minor impact on the predictability of

ocean biogeochemistry on interannual to decadal timescales, no assimilation of ocean biogeochemistry is done within NorCPM1.

For the CMIP6 DCPP two sets of decadal hindcasts have been produced with NorCPM1. Both sets consist of 10 members each and have been initialized on October 15th every year from 1959 to 2017. They are thereafter run for ten years plus three months. The two sets have been initialized from two different reanalysis products that have been integrated between 1950 and 2019. The first one uses 1980-2010 as reference climatology and applies weakly coupled assimilation, meaning that only the ocean state is updated during assimilation. The second one uses 1950-2010 as a reference climatology and uses strongly coupled data assimilation, implying that also the sea ice is updated during the assimilation. Note that a discontinuity in the atmospheric CO_2 in the years of 2015 and 2016 was discovered after all the simulations had been performed, which had arisen when merging the historical and the future atmospheric forcing. The effect of this on the prediction skill is avoided if benchmarking the skill of the predictions against the historical (uninitialized) runs. However, this does not affect the current study which stretches until the year of 2013.

2. Anomaly correlation coefficient

The quality of a prediction as an average over an ensemble can be assessed using the anomaly correlation coefficient (ACC) defined as:

$$ACC(t) = \frac{\sum_{i=1}^{n}[a_i(t)o_i]}{\sqrt{\sum_{i=1}^{n}[a_i(t)]^2 \sum_{i=1}^{n}[o_i]^2}}$$

where t is the hindcast lead years, and n is the number of starting years, a_i represents the ensemble mean anomaly of the quantity in the hindcasts starting from year i, and o_i

is the anomaly of reference data from the corresponding year (Pohlmann et al., 2013). In this study, t is from lead time of 1 to 10 years and n is the length of time series of 32 years.

If the ACC value equals one, both data sets are positively correlated, which means a perfect prediction quality. If the ACC value equals minus one, the data sets have an inverse relationship. If the ACC value is zero, there is no correlation between the data.

References

Armstrong, R. L., Knowles, K. W., Brodzik, M. J., & Hardman, M. A. (1994). *DMSP SSM/I-SSMIS Pathfinder daily EASE-grid brightness temperatures, version 2, Jan 1987–Dec 2016.* National Snow and Ice Data Center, Boulder, CO, USA. doi: 10.5067/3EX2U1DV3434

Aumont, O., Éthé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development Discussions*, *8*(2).

Balmaseda, M. A., Mogensen, K., & Weaver, A. T. (2013). Evaluation of the ecmwf ocean reanalysis system oras4. *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1132–1161.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., ... Kristjánsson, J. E. (2013). The Norwegian Earth System Model, NorESM1-M â Part 1: Description and basic evaluation of the physical climate. *Geoscientific Model Development*, *6*(3), 687–720. Retrieved from `https://www.geosci-model-dev.net/6/687/2013/` doi: 10.5194/gmd-6-687-2013

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., ... others (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7).

Bunzel, F., Notz, D., Baehr, J., Müller, W. A., & Fröhlich, K. (2016). Seasonal climate forecasts significantly affected by observational uncertainty of arctic sea ice concentration. *Geophysical Research Letters*, *43*(2), 852–859.

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, E., Gastineau, G., ... others (2019). Improved near surface continental climate in IPSL-CM6A-LR by combined evolutions of atmospheric and land surface physics. *Journal of Advances in Modeling Earth Systems*, e2019MS002005.

Counillon, F., Bethke, I., Keenlyside, N., Bentsen, M., Bertino, L., & Zheng, F. (2014). Seasonal-to-decadal predictions with the ensemble Kalman filter and the Norwegian Earth System Model: a twin experiment. *Tellus A: Dynamic Meteorology and Oceanography*, *66*(1), 21074. Retrieved from `https://doi.org/10.3402/tellusa.v66.21074` doi: 10.3402/tellusa.v66.21074

Counillon, F., Keenlyside, N., Bethke, I., Wang, Y., Billeau, S., Shen, M. L., & Bentsen, M. (2016). Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model. *Tellus A: Dynamic Meteorology and Oceanography*, *68*(1), 32437. Retrieved from `https://doi.org/10.3402/tellusa.v68.32437` doi: 10.3402/tellusa.v68.32437

Dee, D. P., Uppala, S. M., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., ... others (2011). The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553–597.

Dirkson, A., Merryfield, W. J., & Monahan, A. (2017). Impacts of sea ice thickness initialization on seasonal Arctic sea ice predictions. *Journal of Climate*, *30*(3), 1001–1017.

Estella-Perez, V., Mignot, J., Guilyardi, E., Swingedouw, D., & Reverdin, G. (2020).

Advances in reconstructing the AMOC using sea surface observations of salinity. *Climate Dynamics*, 1–18.

Fransner, F., Counillon, F., Bethke, I., Tjiputra, J., Samuelsen, A., Nummelin, A., & Olsen, A. (2020). Ocean biogeochemical predictions—initialization and limits of predictability. *Frontiers in Marine Science*, *7*, 386.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., ... others (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, *5*(3), 572–597.

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*(12), 6704-6716. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JC009067` doi: 10.1002/2013JC009067

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., ... Kawamiya, M. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, *13*, 2197–2244. doi: 10.5194/gmd-13-2197-2020

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., & Núñez-Riboni, I. (2013). Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations. *Journal of Advances in Modeling Earth Systems*, *5*(2), 287–

315.

Ishii, M., & Kimoto, M. (2009). Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *Journal of Oceanography*, *65*, 287–299. doi: 10.1007/s10872-009-0027-7

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., ... others (2015). The community earth system model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... Takahashi, K. (2015). The JRA-55 Reanalysis: General specifications and basic characteristics. *Journal of Meteorological Society of Japan*, *93*, 5–48. doi: 10.2151/jmsj.2015-001

Li, H., Ilyina, T., Müller, W. A., & Landschützer, P. (2019). Predicting the variable ocean carbon sink. *Science advances*, *5*(4), eaav6471.

Lindsay, K., Bonan, G. B., Doney, S. C., Hoffman, F. M., Lawrence, D. M., Long, M. C., ... Thornton, P. E. (2014). Preindustrial-Control and Twentieth-Century Carbon Cycle Experiments with the Earth System Model CESM1(BGC). *Journal of Climate*, *27*(24), 8981-9005.

Lovenduski, N. S., Yeager, S. G., Lindsay, K., & Long, M. C. (2019). Predicting near-term variability in ocean carbon uptake. *Earth System Dynamics (Online)*, *10*(1).

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., ... others (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing co_2. *Journal of Advances in Modeling Earth Systems*,

11(4), 998–1038.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., ... others (2018). A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1. 2-HR). *Journal of Advances in Modeling Earth Systems*, *10*(7), 1383–1413.

Park, J.-Y., Stock, C. A., Dunne, J. P., Yang, X., & Rosati, A. (2019). Seasonal to multiannual marine ecosystem prediction with a global Earth system model. *Science*, *365*(6450), 284–288.

Park, J.-Y., Stock, C. A., Yang, X., Dunne, J. P., Rosati, A., John, J., & Zhang, S. (2018). Modeling global ocean biogeochemistry with physical data assimilation: a pragmatic solution to the equatorial instability. *Journal of Advances in Modeling Earth Systems*, *10*(3), 891–906.

Pohlmann, H., Smith, D. M., Balmaseda, M. A., Keenlyside, N. S., Masina, S., Matei, D., ... Rogel, P. (2013). Predictability of the mid-latitude atlantic meridional overturning circulation in a multi-model system. *Climate dynamics*, *41*(3-4), 775–785.

Reverdin, G., Friedman, A. R., Chafik, L., Holliday, N. P., Szekely, T., Valdimarsson, H., & Yashayaev, I. (2019). North atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. part 1: dataset validation and RMS variability. *Ocean Dynamics*, *69*(3), 385–403.

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*(13), 1609-1625.

Sospedra-Alfonso, R., & Boer, G. J. (2020). Assessing the impact of initialization on decadal prediction skill. *Geophysical Research Letters*, *47*(4), e2019GL086361.

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., ... others (2019). The Canadian Earth System Model version 5 (CanESM5. 0.3). *Geoscientific Model Development*, *12*(11), 4823–4873.

Tatebe, H., Ishii, M., Mochizuki, T., Chikamoto, Y., Sakamoto, T. T., Komuro, Y., ... Kimoto, M. (2012). The initialization of the MIROC climate models with hydrographic data assimilation for decadal prediction. *Journal of Meteorological Society of Japan*, *90A*, 275–294. doi: 10.2151/jmsj.2012-A14

Tatebe, H., Tanaka, Y., Komuro, Y., & Hasumi, H. (2018). Impact of deep ocean mixing on the climatic mean state in the Southern Ocean. *Scientific Reports*, *8*, 14479. doi: 10.1038/s41598-018-32768-6

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., ... Heinze, C. (2013). Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). *Geoscientific Model Development*, *6*(2), 301–325. Retrieved from `https://www.geosci-model-dev.net/6/301/2013/` doi: 10.5194/gmd-6-301-2013

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V. D. C., Fiorino, M., ... others (2005). The era-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology*

and physical oceanography, *131*(612), 2961–3012.

Wang, Y., Counillon, F., & Bertino, L. (2016). Alleviating the bias induced by the linear analysis update with an isopycnal ocean model. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 1064-1074. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2709` doi: 10.1002/qj.2709

Wang, Y., Counillon, F., Bethke, I., Keenlyside, N., Bocquet, M., & lin Shen, M. (2017). Optimising assimilation of hydrographic profiles into isopycnal ocean models with ensemble data assimilation. *Ocean Modelling*, *114*, 33 - 44. Retrieved from `http://www.sciencedirect.com/science/article/pii/S146350031730063X` doi: https://doi.org/10.1016/j.ocemod.2017.04.007

Watanabe, M., Tatebe, H., Koyama, H., Hajima, T., Watanabe, M., & Kawamiya, M. (2020). Importance of el niño reproducibility for reconstructing historical co_2 flux variations in the equatorial pacific. *Ocean Science*, *16*(6), 1431–1442. Retrieved from `https://os.copernicus.org/articles/16/1431/2020/` doi: 10.5194/os-16-1431-2020

Yeager, S., Danabasoglu, G., Rosenbloom, N., Strand, W., Bates, S., Meehl, G., ... others (2018). Predicting near-term changes in the earth system: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bulletin of the American Meteorological Society*, *99*(9), 1867–1886.

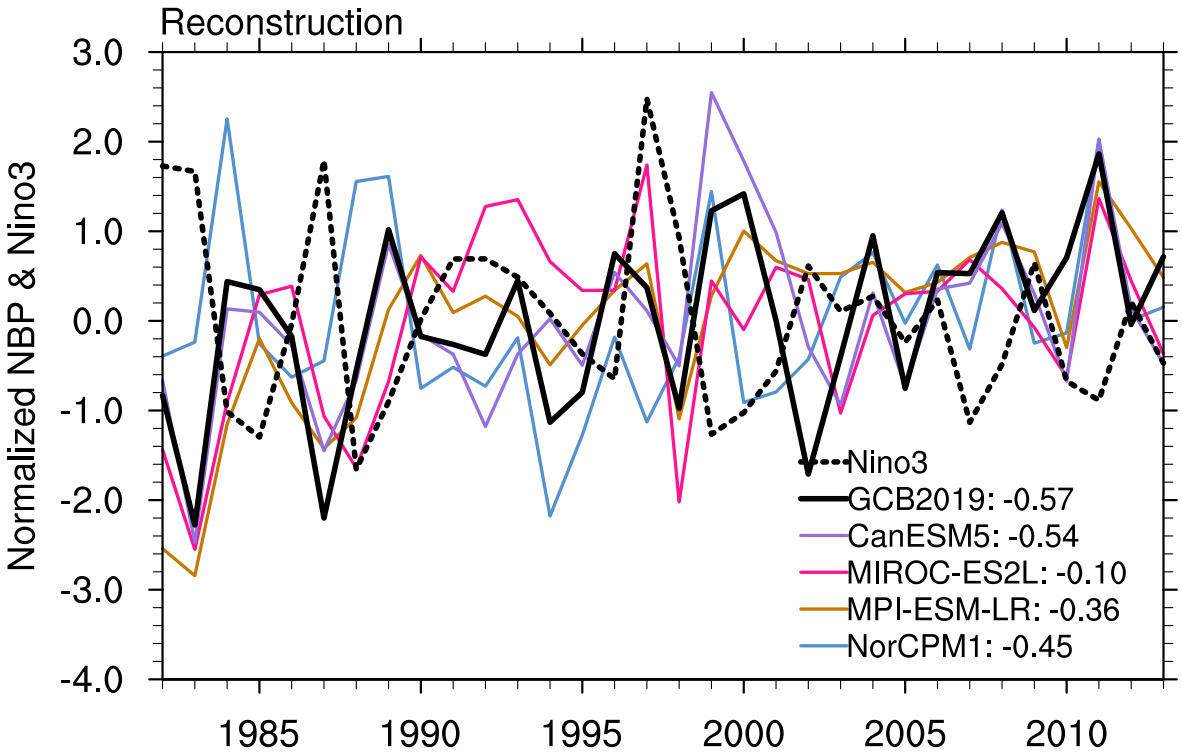

Figure S1. Time series of Nino3 index and the air-land CO_2 flux (NBP) from model reconstruction. The Nino3 index from HadISST is shown with black dash line, and the GCB2019 NBP is shown with black solid line. The numbers after the legend show the correlation with HadISST Nino3 index.

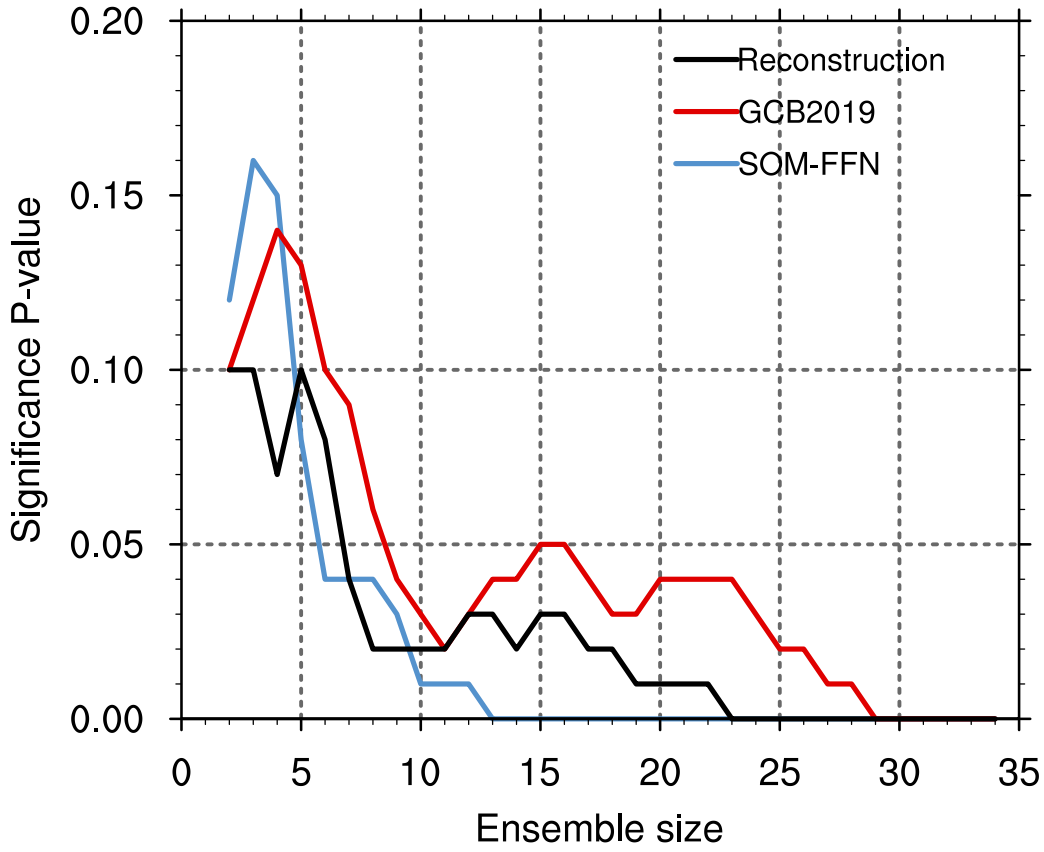

Figure S2. Significance P-values of air-sea CO_2 flux predictions evolving with ensemble size. The significance of skill of CESM-DPLE air-sea CO_2 flux predictions at lead time of 3 years relative to the corresponding reconstruction simulation is shown in black. The colors indicate results relative to different reference data, i.e. GCB (red) and SOM-FFN (blue).

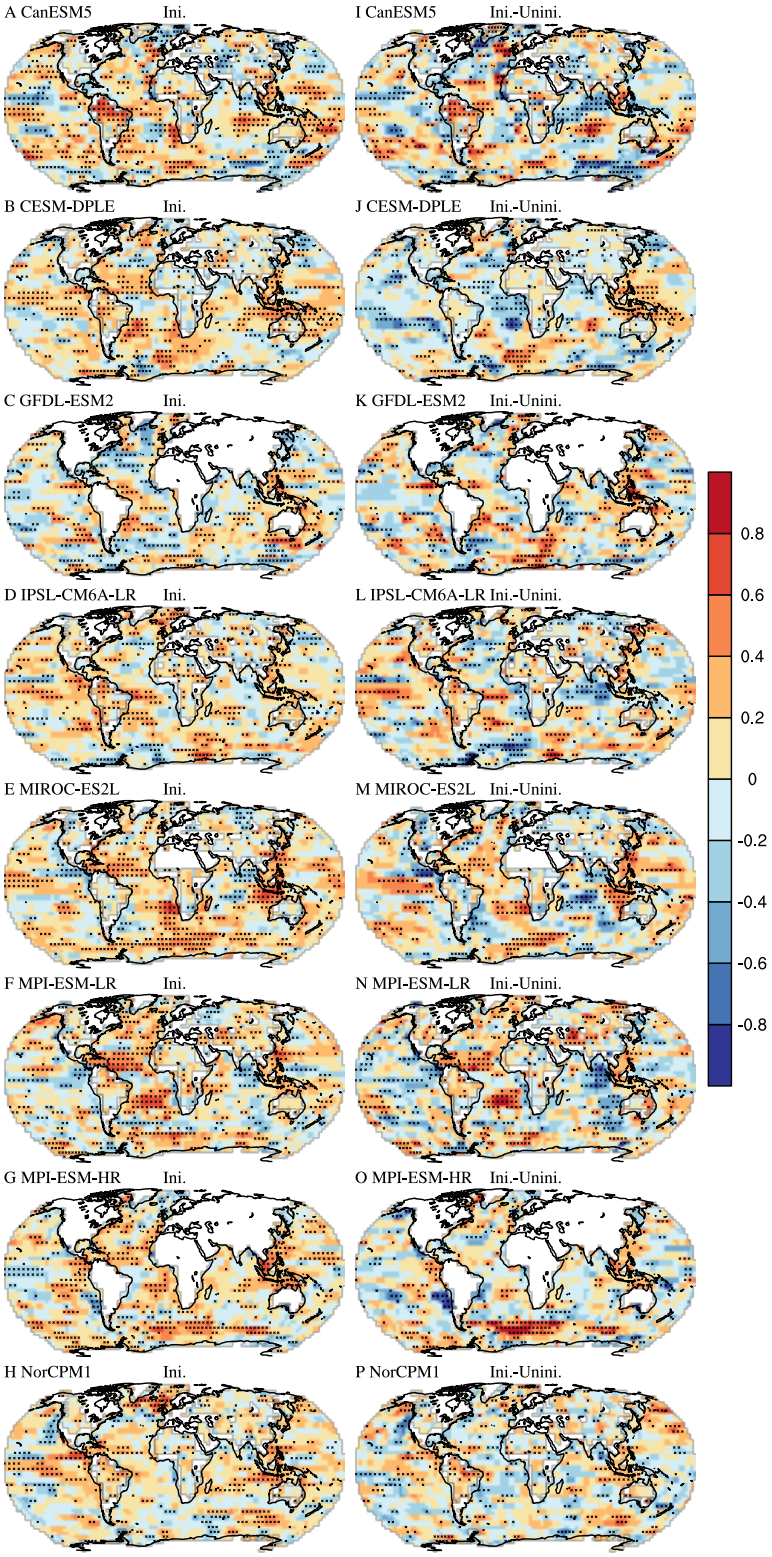

Figure S3. Correlation for air-sea and air-land CO_2 flux at lead time of 2 years relative to SOM-FFN for ocean and GCB2019 for land, respectively. The skill is quantified with anomaly correlation coefficient. Shown are the correlations of the initialized retrospective predictions (left) and the difference between initialized and uninitialized simulations(right). The crosses show significance at 95% level.

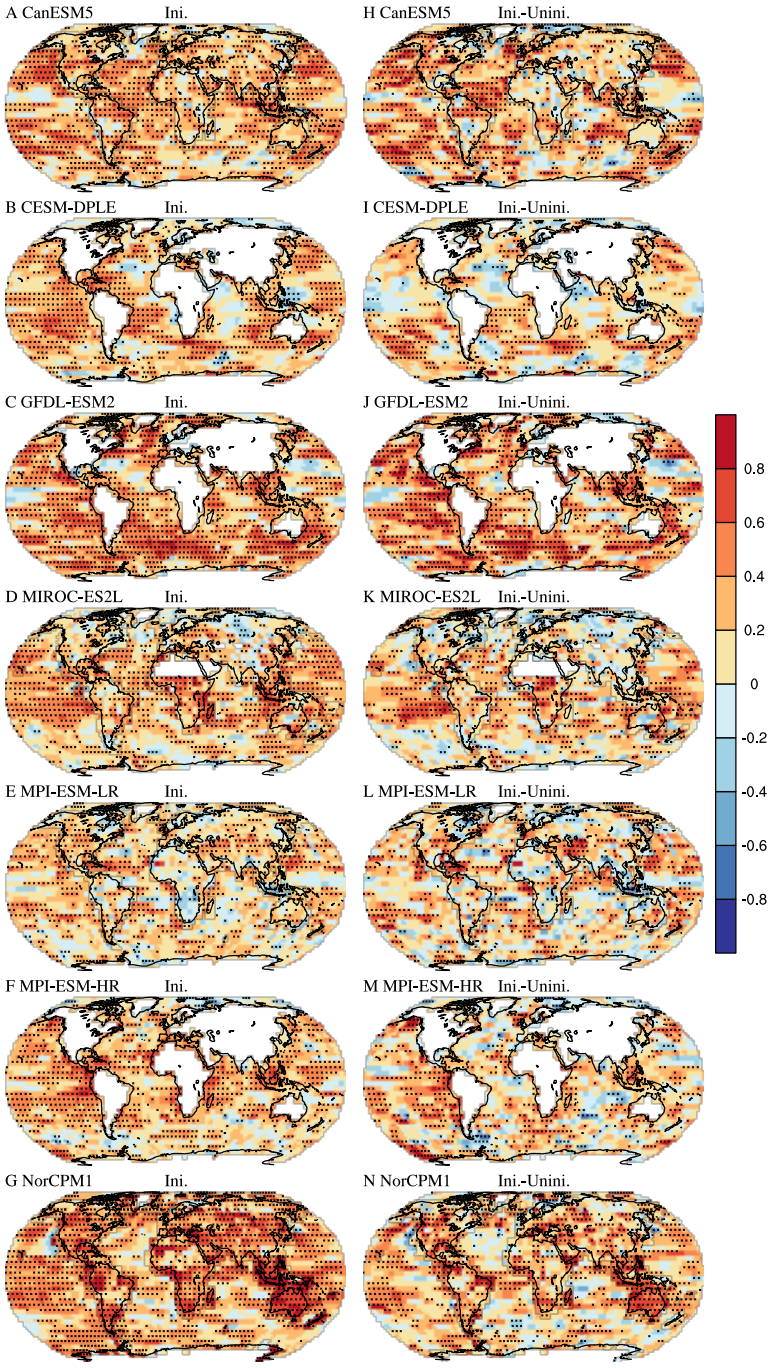

Figure S4. Same as Fig. S3, but based on correlation with the reconstruction simulation.

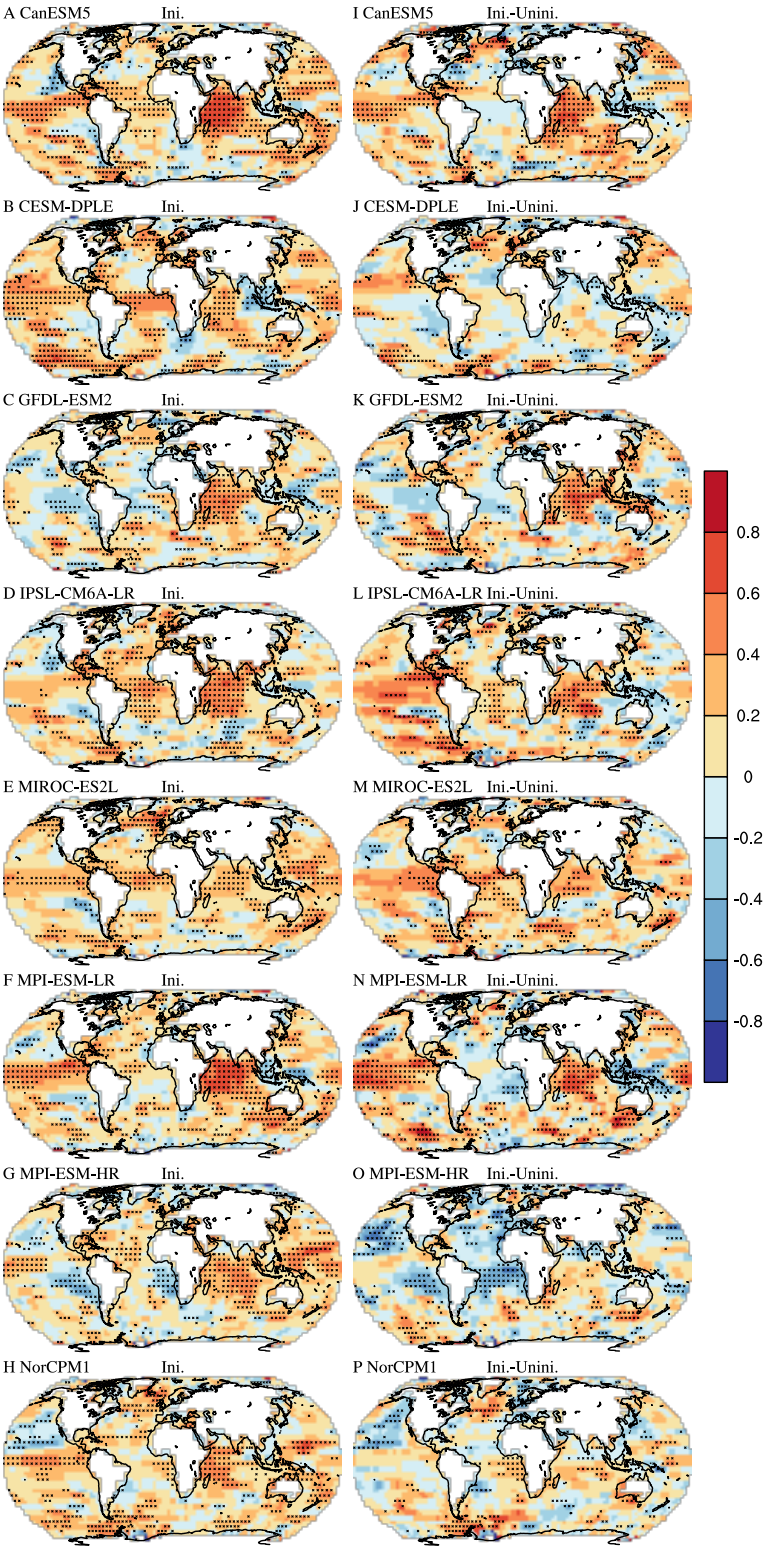

Figure S5. Predictive skill of sea surface temperature for the initialized simulation at lead time of 2 years and the difference between the the initialized and uninitialized simulations for the period from 1982-2013. Observations are HadISST.

Table S1. Overview of prediction systems and initialization techniques.

Model	CanESM5	CESM-DPLE	GFDL-ESM2	IPSL-CM6A-LR	MIROC-ES2L	MPI-ESM-LR	MPI-ESM1.2-HR	NorCPM1
Resolution Atmosphere	T63, 47 levels	1.0°, 30 levels	2.5° lon 2.0° lat, 24 levels	2.5°x1.3°, 79 levels	T42, 40 levels	T63, 47 levels	T127, 95 levels	1.9x2.5°, 26 levels
Resolution Ocean	ORCA1, 45 levels	1°, 60 levels	1°, 50 levels	1°, 75 levels	Tripolar (∼1°), 62 levels	1.5°, 40 levels	0.4°, 40 levels	1°, 51 levels
Initialization ocean	ORAS5 3D T-S anomalies, SST relaxed to OISSTv2; sea-ice concentration relaxed to HadISST.2, CMC analysis; thickness assimilation	Ocean-sea-ice forced at the surface with atmospheric states and fluxes (modified COREv2)	GFDL's ECDA for WOD, argo, SST	EN4 SST and Atlantic SSS	Full-field 3D T-S	ORAS4 3D T-S anomalies	ORAS4 3D T-S anomalies, sea-ice concentration anomalies from NSIDC	EKF for HadISST2 + OISSTV2 SST, EN4 T,S profiles
Initialization atmosphere	ERA-40 and ERA-Interim: vorticity, divergence, log(p), T; full field	CESM Large Ensemble	GFDL's ECDA with NCEP-DOE reanalysis 2	N/A	JRA55 wind and T; full field	ERA-40 and ERA-Interim: vorticity, divergence, log(p), T; full field	ERA-40 and ERA-Interim: vorticity, divergence, log(p), T; full field	N/A
Ensemble size	20	40	12	10	10	10	10	20
Start years	1961-2017 yearly from 1 Jan. for 10 years	1954-2015 yearly from 1 Nov. for 10 years	1961-2017 yearly from 1 Jan. for 10 years	1961-2014 yearly from 1 Jan.	1980-2017 yearly from 1 Jan. for 10 years	1961-2014 yearly from 1 Jan. for 10 years	1960-2018 yearly from 1 Nov. for 10 years	1959-2017 yearly from 15 Oct. for 10 years
Forcing	cmip6	cmip5	cmip5	cmip6	cmip6	cmip5	cmip6	cmip6
References	Swart et al. (2019)	Yeager et al. (2018)	Park et al. (2018)	Boucher et al. (2020)	Watanabe et al. (2020)	Giorgetta et al. (2013)	Mauritsen et al. (2019)	(Counillon et al., 2016)

Table S2. Significance P-values for predictive skills of the detrended CO_2 flux into the ocean, CO_2 flux into land, and variations in the inferred atmospheric CO_2 growth presented in Figure 3. LY1 to LY10 refer to lead years 1 to 10, respectively.

Model	LY1	LY2	LY3	LY4	LY5	LY6	LY7	LY8	LY9	LY10
air-sea CO_2 flux										
CanESM5	0.97	0.94	0.57	0.11	0.12	0.01	0.73	0.73	0.64	0.90
CESM-DPLE	0.03	0.00	0.00	0.01	0.02	0.02	0.19	0.05	0.09	0.64
GFDL-ESM2	0.00	0.01	0.09	0.62	0.82	0.71	0.90	0.89	0.93	0.78
IPSL-CM6A-LR	0.34	0.29	0.32	0.21	0.10	0.10	0.16	0.23	0.11	0.14
MIROC-ES2L	0.01	0.04	0.08	0.19	0.26	0.29	0.30	0.31	0.89	0.93
MPI-ESM-LR	0.13	0.09	0.26	0.12	0.06	0.12	0.12	0.03	0.09	0.28
MPI-ESM1-2-HR	0.00	0.00	0.00	0.00	0.00	0.10	0.04	0.16	0.05	0.07
NorCPM1	0.00	0.00	0.01	0.00	0.05	0.05	0.12	0.38	0.34	0.34
air-land CO_2 flux										
CanESM5	0.00	0.00	0.47	0.86	0.82	0.91	0.61	0.97	0.82	0.70
CESM-DPLE	0.24	0.07	0.62	0.77	0.71	0.80	0.87	0.76	0.82	0.89
IPSL-CM6A-LR	0.00	0.00	0.13	0.23	0.40	0.19	0.27	0.21	0.31	0.14
MIROC-ES2L	0.01	0.56	0.35	0.39	0.63	0.73	0.50	0.36	0.56	0.80
MPI-ESM-LR	0.00	0.03	0.35	0.59	0.79	0.48	0.21	0.59	0.35	0.48
NorCPM1	0.00	0.01	0.74	0.46	0.25	0.46	0.48	0.27	0.26	0.21
atmospheric CO_2 growth										
CanESM5	0.00	0.00	0.55	0.86	0.77	0.92	0.66	0.97	0.83	0.73
CESM-DPLE	0.33	0.07	0.57	0.70	0.61	0.81	0.88	0.76	0.82	0.90
IPSL-CM6A-LR	0.00	0.01	0.12	0.24	0.46	0.22	0.31	0.28	0.39	0.19
MIROC-ES2L	0.01	0.65	0.42	0.41	0.65	0.69	0.40	0.32	0.52	0.71
MPI-ESM-LR	0.00	0.05	0.46	0.67	0.80	0.51	0.19	0.53	0.31	0.44
NorCPM1	0.00	0.00	0.70	0.40	0.23	0.38	0.48	0.28	0.31	0.25