Large ensemble testbed
Evaluating pCO$_2$ interpolation methods

Luke Gloege
Air-Sea CO$_2$ exchange

$pCO_2^{atm.}$  \[ \text{arrow} \]  $pCO_2^{ocean}$

$pCO_2^{atm.}$  \[ \text{arrow} \]  $pCO_2^{ocean}$
Models and data products mismatch

- Data products are more variable than models
- Different interpretation for evolution of ocean carbon sink
- Which is correct: models or data products?

Decadal variability south of 35°S

Gruber, Landschützer, and Lovenduski (2019).
Two-step neural network (SOM-FFN)

SOM : Self organizing map
FFN : feed-forward neural network

1. Climatological pCO$_2$, SSS, SST, and MLD
2. Data grouped using self organizing map

- Learns a non-linear relationship between global features and pCO$_2$
- Inputs are proxies for processes affecting ocean pCO$_2$

Comparing SOM-FFN and SOCAT

SOCATv2019 data

SOM-FFN product

Bias (global avg. 2.15 μatm)

Bakker et al. 2016 EESD; Sabine et al. 2013

How large ensembles can help

• Evaluate performance across different climate states

• 100 randomly selected ensemble members
  • CanESM2 (25)
  • CESM (25)
  • GFDL (25)
  • MPI (25)

• Each climate state is equally likely
• Only considering 1982-2015
Large ensemble pCO$_2$ testbed

1. Sample model member as SOCAT monthly pCO$_2$ product

2. Train, evaluate, test reconstruction method at sampled locations

3. Estimate monthly varying pCO$_2$ on global scale using trained model

4. Statistically compare reconstructed pCO$_2$ to model truth. Each spatial point is temporally decomposed

Neural network reconstruction method:
- SST
- SSS
- MLD
- Chl
- xCO$_2$

pCO$_2$
Large ensemble pCO$_2$ testbed

1. Sample model member as SOCAT monthly pCO$_2$ product

2. Train, evaluate, test reconstruction method at sampled locations

3. Estimate monthly varying pCO$_2$ on global scale using trained model

4. Statistically compare reconstructed pCO$_2$ to model truth. Each spatial point is temporally decomposed

Neural network reconstruction method

- SST
- SSS
- MLD
- Chl
- $x$CO$_2$
- pCO$_2$
Large ensemble pCO$_2$ testbed

1. Sample model member as SOCAT monthly pCO$_2$ product

2. Train, evaluate, test reconstruction method at sampled locations

3. Estimate monthly varying pCO$_2$ on global scale using trained model

4. Statistically compare reconstructed pCO$_2$ to model truth. Each spatial point is temporally decomposed

Neural network reconstruction method
Large ensemble pCO$_2$ testbed

1. Sample model member as SOCAT monthly pCO$_2$ product

2. Train, evaluate, test reconstruction method at sampled locations

3. Estimate monthly varying pCO$_2$ on global scale using trained model

4. Statistically compare reconstructed pCO$_2$ to model truth. Each spatial point is temporally decomposed

Neural network reconstruction method

- SST
- SSS
- MLD
- Chl
- xCO$_2$
Large ensemble pCO\textsubscript{2} testbed

1. Sample model member as SOCAT monthly pCO\textsubscript{2} product

2. Train, evaluate, test reconstruction method at sampled locations

3. Estimate monthly varying pCO\textsubscript{2} on global scale using trained model

4. Statistically compare reconstructed pCO\textsubscript{2} to model truth. Each spatial point is temporally decomposed
Temporal decomposition

Each location in each member and reconstruction decomposed into seasonal, decadal, and sub-decadal variability

Cleveland et al. (1990) *Journal of official statistics.*
Statistical metrics

Phasing

Is reconstruction in phase with model?

Long-term mean

is there a systemic offset?

Variability

does reconstruction capture observed variability?

Correlation

\[ r = \frac{\text{cov}(m,r)}{\sigma_r \sigma_m} \]

Bias

\[ \text{Bias} = \bar{m} - \bar{r} \]

Normalized STD

\[ \sigma^* = \left( \frac{\sigma_r}{\sigma_m} - 1 \right) \times 100\% \]
Bias and number of observations

Is the reconstruction offset from the model?

Global avg = -0.44 μatm

- Bias spatially varying but global average is low
- Less spread in bias in regions with more data
- Higher decadal correlation in regions with more data
Normalized standard deviation
Does the reconstruction capture the model’s variability?

- Overestimates decadal variability in Southern Ocean by about 12%
- Reconciles gap between models and SOM-FFN
Conclusions

performance dependent on data density

Global avg = -0.44 μatm

Bias is spatially heterogenous, but negligible on a global scale

Overestimation of decadal variability reconciles the discrepancy with models

This testbed serves a platform to develop and test new methodologies
Acknowledgements

Galen McKinley
Columbia University / LDEO

Peter Landschützer
Max Planck Institute for Meteorology

Nikki Lovenduski
University of Colorado

Keith Rodgers
IBS center for climate physics

Tatiana Ilyina
Max Planck Institute for Meteorology

John Fyfe
Environment and climate change Canada

Amanda Fay
Columbia University / LDEO

Steve Jones
University of Bergen

Christian Rödenbeck
Max Planck Institute for biogeochemistry

Thomas Frölicher
Physics Institute, University of Bern

Yohei Takano
Max Planck Institute for Meteorology

Sarah Schlunegger
Princeton University
Take a picture to access the testbed

Testbed URL
figshare.com/s/4337ae68dccbcf34e14c

Email
gloege@ldeo.columbia.edu
Extra slides
Correlation
Is the reconstruction in phase with the model?

- Higher correlations in sub-tropics
- Correlations are consistent across ensemble members
Calculate air-sea CO₂ exchange

\[
F_{CO_2} = k_w S_{CO_2} (1 - f_{ice}) (pCO_2^{ocn} - pCO_2^{atm})
\]

- ERA-interim 6-hourly global atmospheric reanalysis used to estimate monthly varying wind-speed covariance
- Calculated air-sea CO₂ exchange following Landschützer et al. 2014