A ML-based climatology of dissolved organic carbon

Thelma Panaïotis[1], Jamie Wilson[2] & BB Cael[1]

[1] National Oceanography Centre, Southampton, UK

[2] University of Liverpool, Liverpool, UK
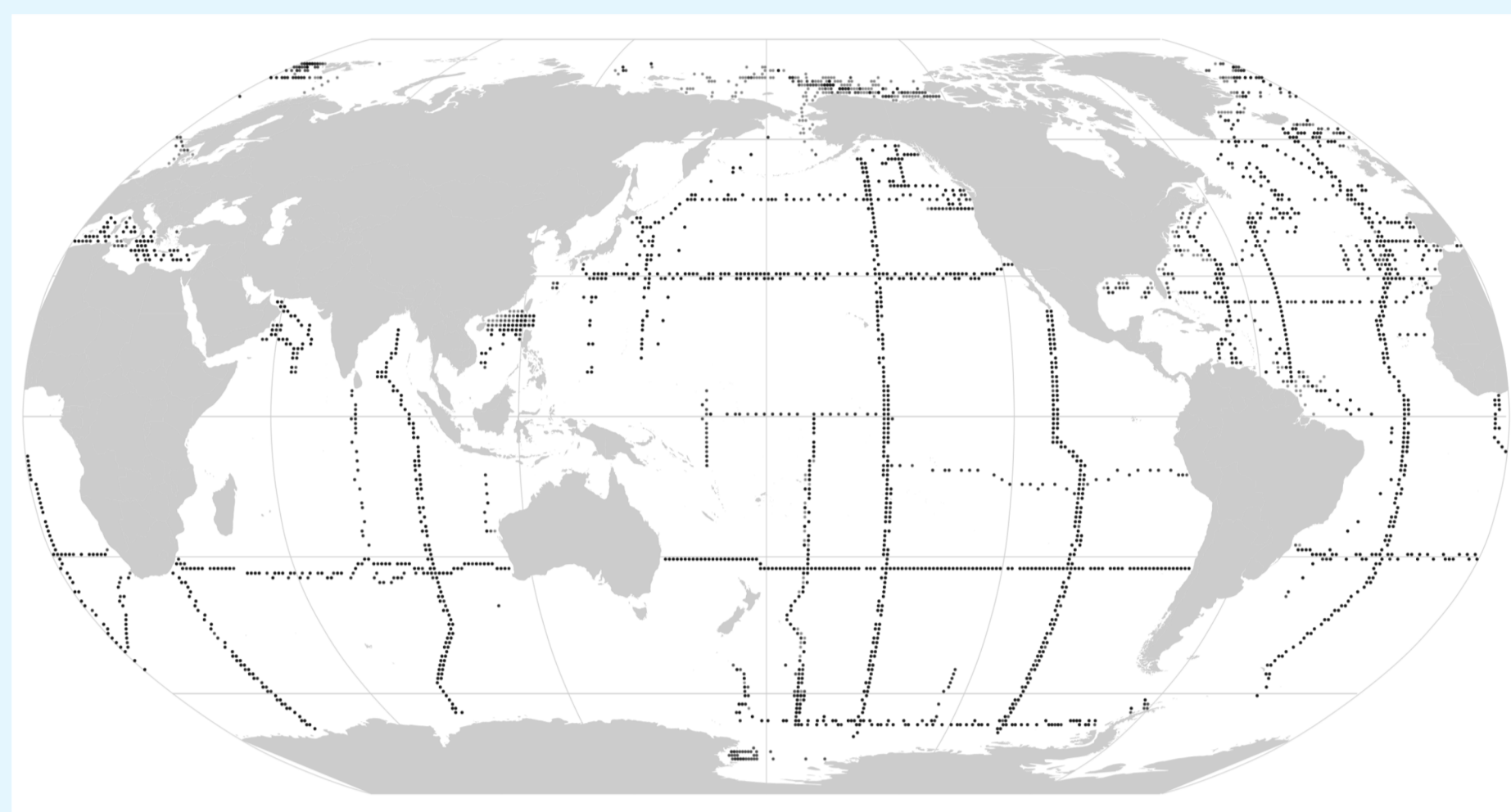
thelma.panaiotis@noc.ac.uk

Context

Dissolved organic carbon = large and reservoir of carbon, but poorly constrained in the biogeochemical carbon cycle.

Numerous DOC observations but **lack of available climatologies**. Can we use ML to build one?

Methods

DOC observations [1]
(several thousands)

Boosted regression trees

Existing env. climatologies [2, 3, 4]
(~ 10-20)
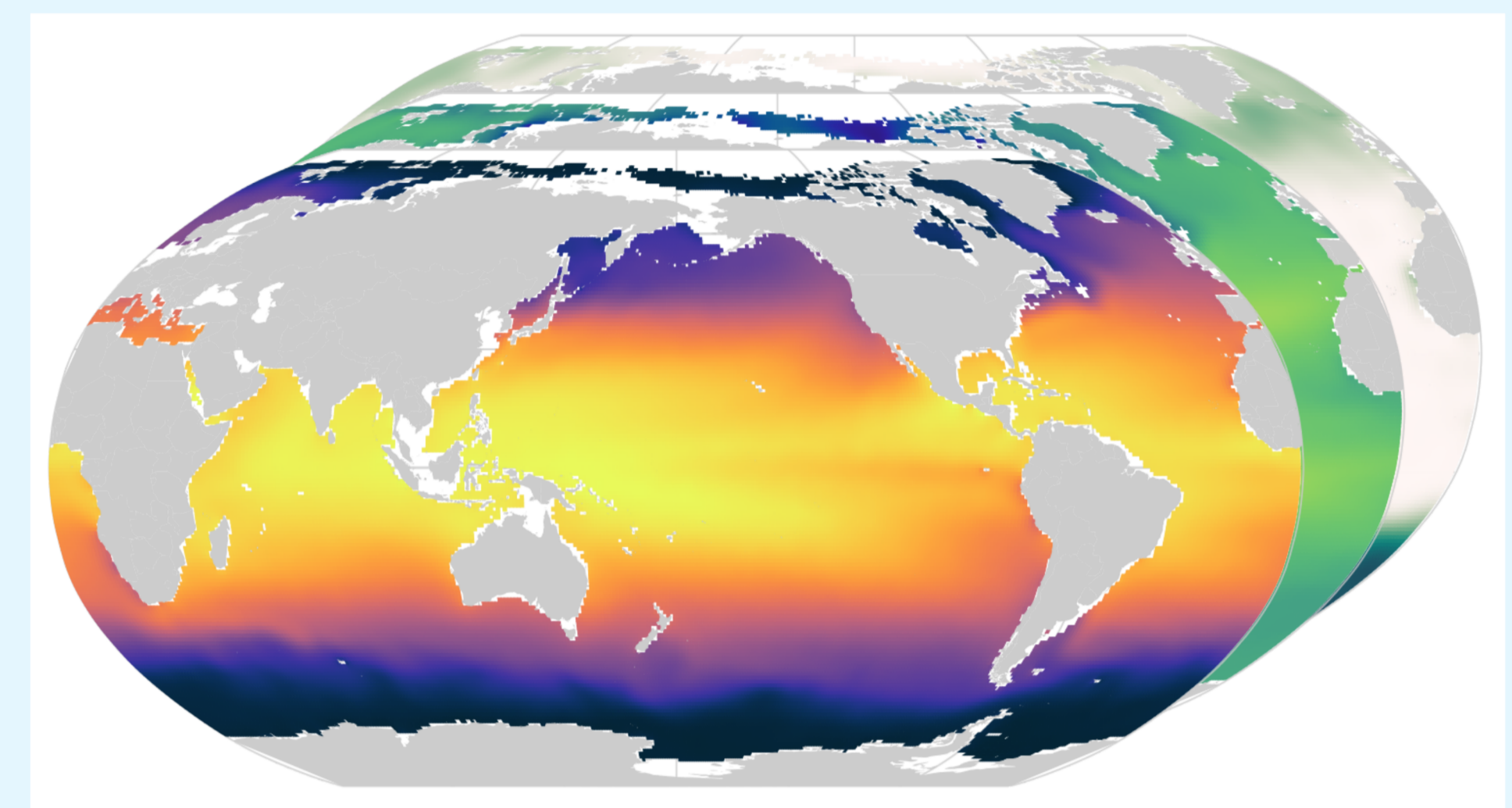

10-fold nested CV

- stratified & spatial splits
- model fine-tuning
- robust performance estimate
- prediction uncertainty

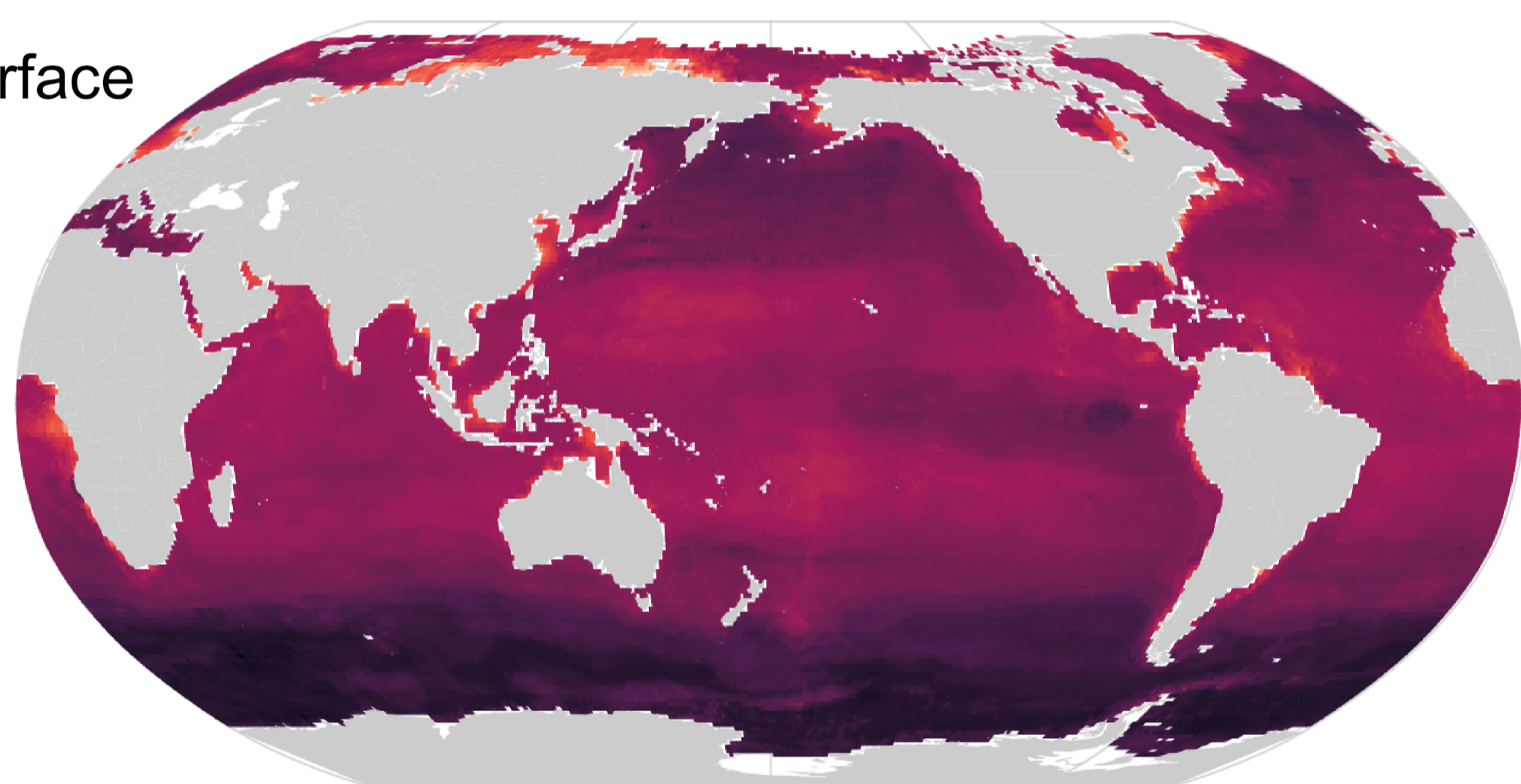Cut in 4 layers: surface, epi., meso. and bathy.
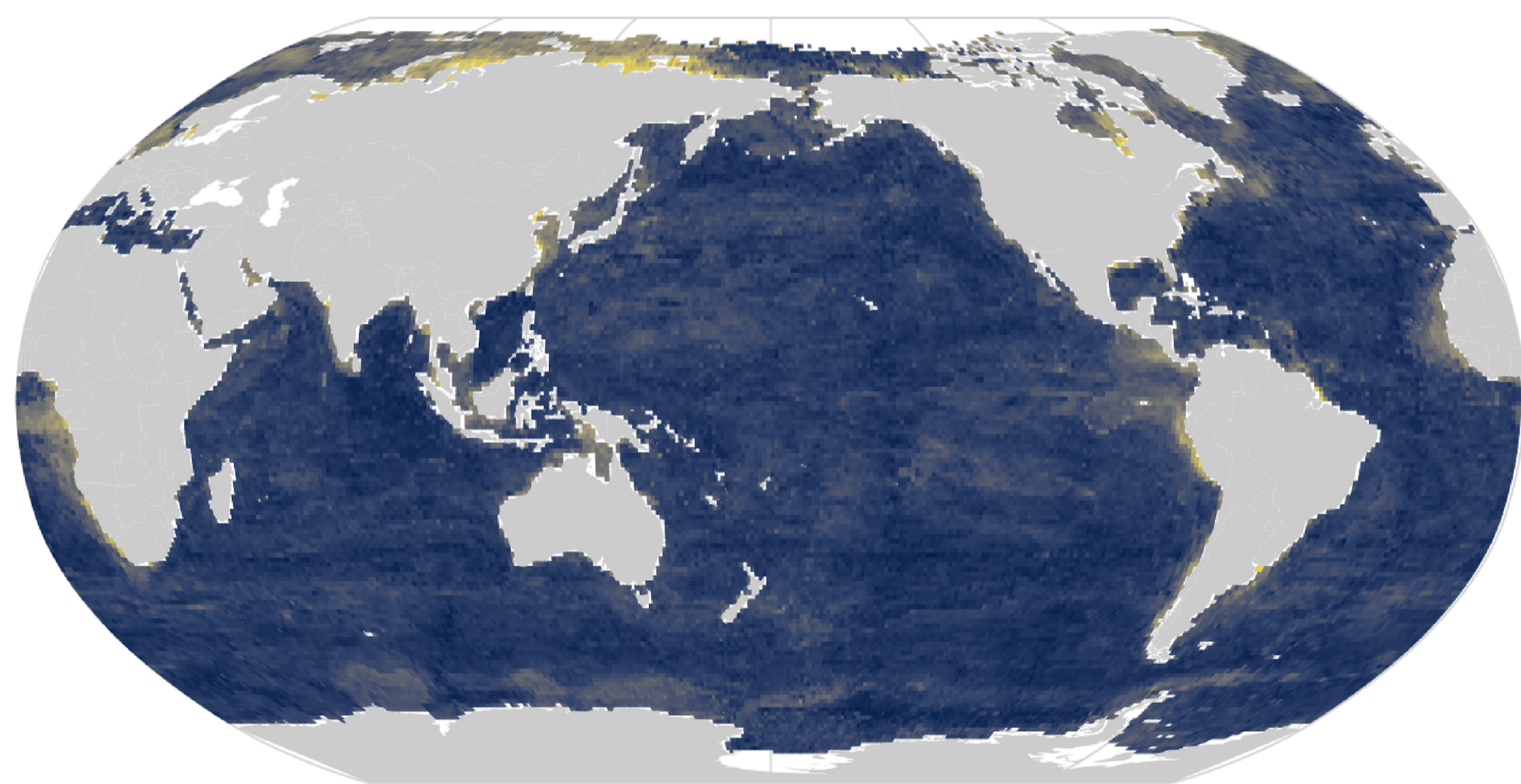Cut in 4 seasons for surface layer

⟶ **4 layer-wise annual climatologies**
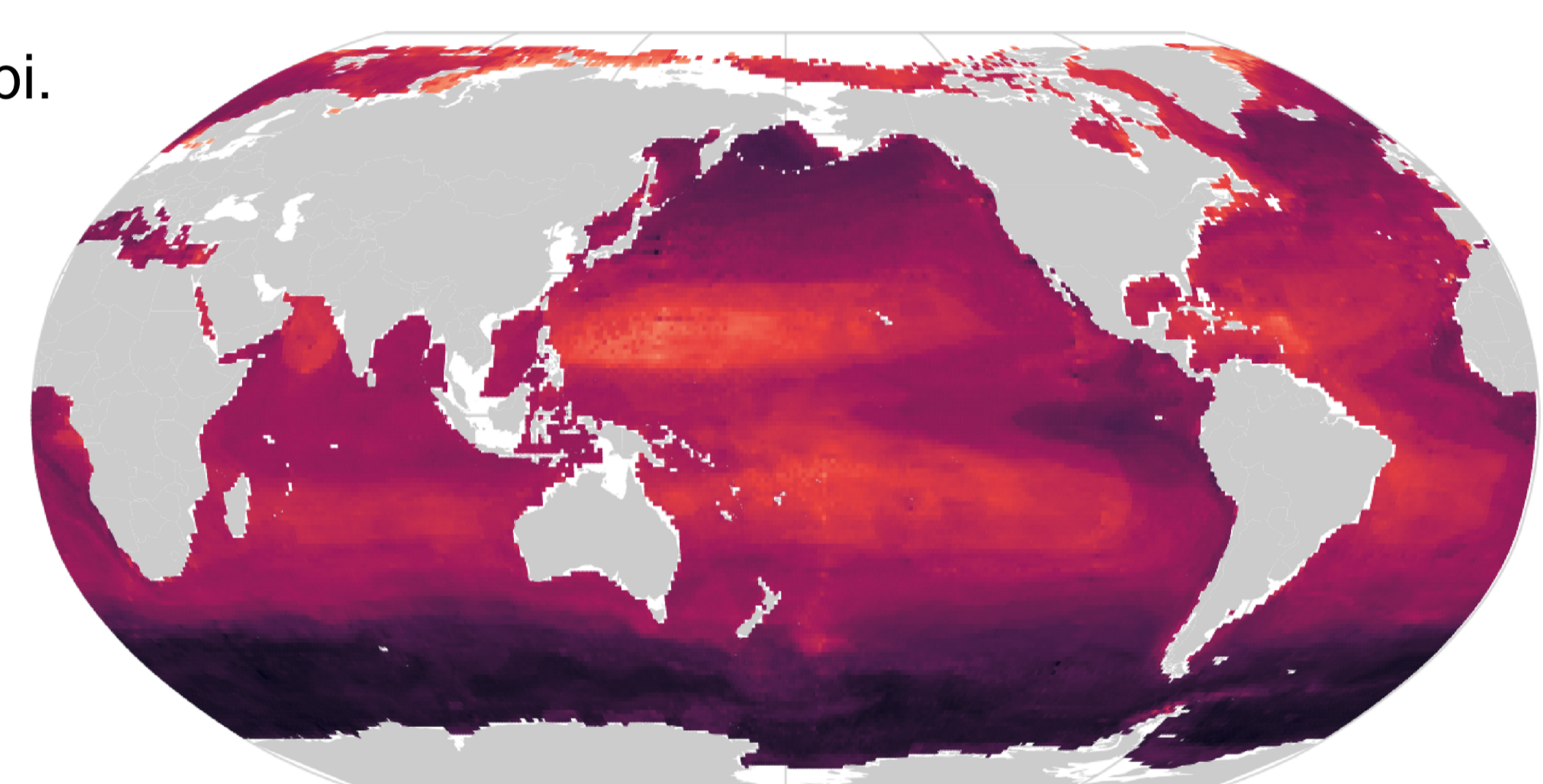⟶ **4 surface seasonal climatologies**

Surface climatology

Surface

Predicted DOC (µmol kg^{-1}) R^2 = 77.5 ± 3.3%

Prediction uncertainty (µmol kg^{-1})

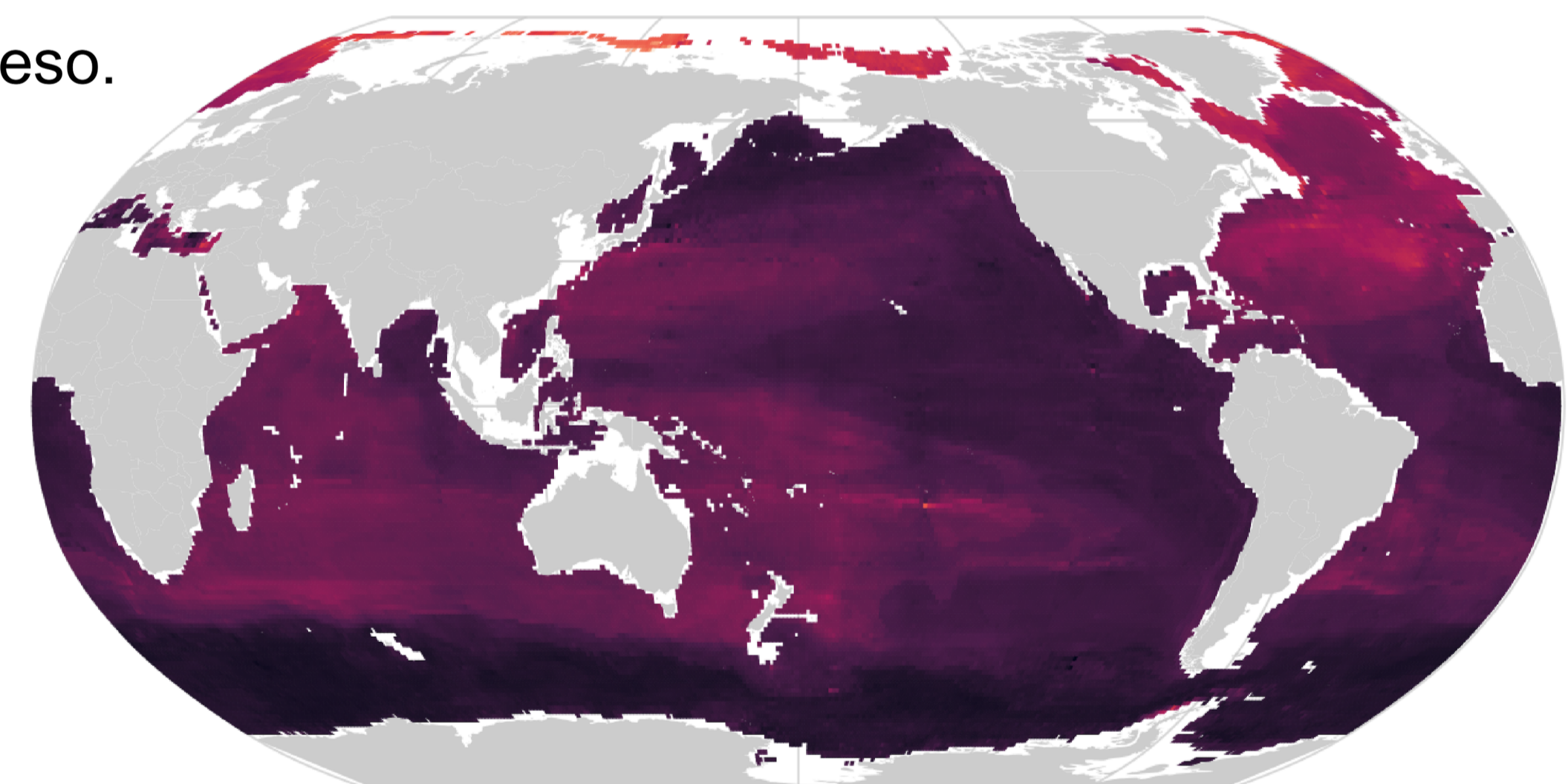## Deeper climatologies

Epi.

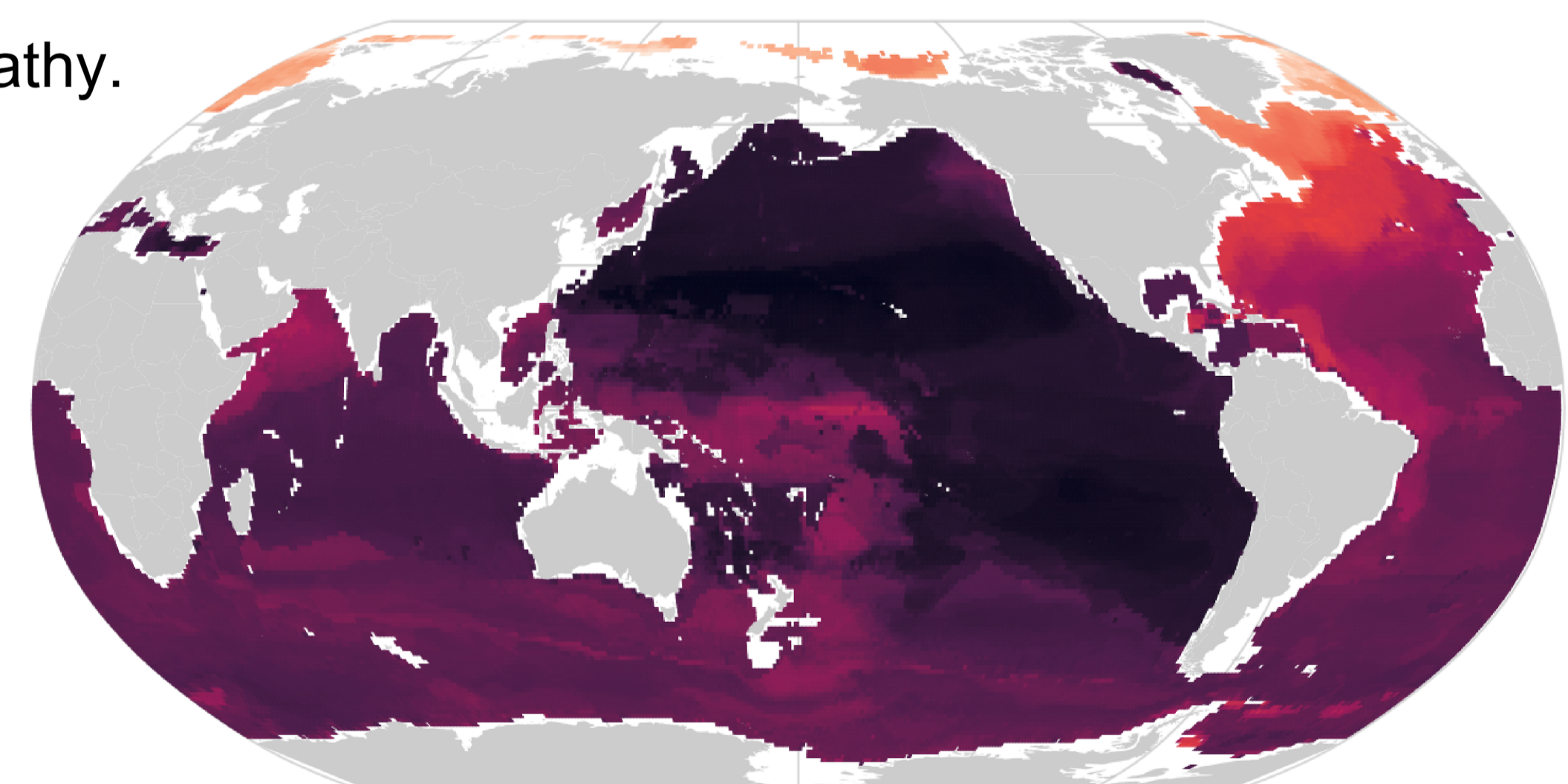Predicted DOC (µmol kg^{-1}) R^2 = 78.4 ± 2.2%

Meso.

Predicted DOC (µmol kg^{-1}) R^2 = 61.1 ± 7.1%

Bathy.

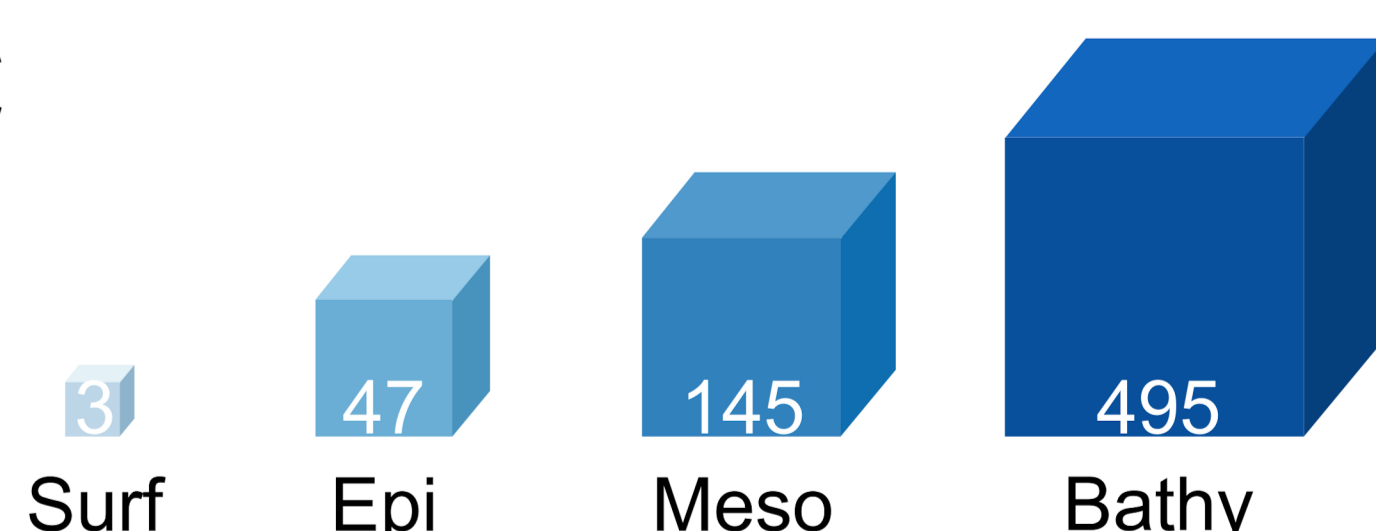Predicted DOC (µmol kg^{-1}) R^2 = 76.8 ± 4.4%

Total DOC estimate

690 PgC

Surf	Epi	Meso	Bathy
3	47	145	495

[1] Hansell, D. A. et al. Compilation of dissolved organic matter (DOM) data obtained from global ocean observations from 1994 to 2021. Version 2 (NCEI Accession 0227166) (2021).

[2] Garcia, H. et al. World Ocean Atlas 2018, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Dissolved Oxygen Saturation. NOAA Atlas NESDIS 83 (2019).

[3] Key, R. M. et al. Global Ocean Data Analysis Project, Version 2 (GLODAPv2). ORNL/CDIAC-162, NDP-093 http://cdiac.ornl.gov/oceans/GLODAPv2/NDP_093.pdf (2015).

[4] Lauvset, S. K. et al. A new global interior ocean mapped climatology: the 1° × 1° GLODAP version 2. Earth System Science Data 8, 325–340 (2016).

CALIPSO.

National Oceanography Centre