MOCCA

Multi-sourced Oceanic Carbonate Chemistry Assimilation

Friedrich Burger

May 22, 2024

Contents

1 Introduction

neural networks...Goodfellow et al. (2016)

2 Model architecture

2.1 Overview

2.2 Surrogate models for mocsy fCO2 and pH

In a first step, surrogate multi-layer perceptron models were trained to replace numerical solution of the oceanic carbonate system. To do so, samples for total alkalinity (A_T), dissolved inorganic carbon (C_T), temperature (T), salinity (S), total silicate (Si_T), and total phosphate (P_T) were randomly generated from uniform distributions (Table 1).
The training data size was set to 5 000 000 samples and the validation data size was set to 1 000 000 samples. Mocsy 2.0 (Orr et al., 2015a) was then used to calculate fCO_2 and pH for these samples. After normalizing per feature with the means and standard deviations given in Table 1, multilayer perceptron models with three identical hidden layers were trained with mocsy fCO_2 and pH as labels.

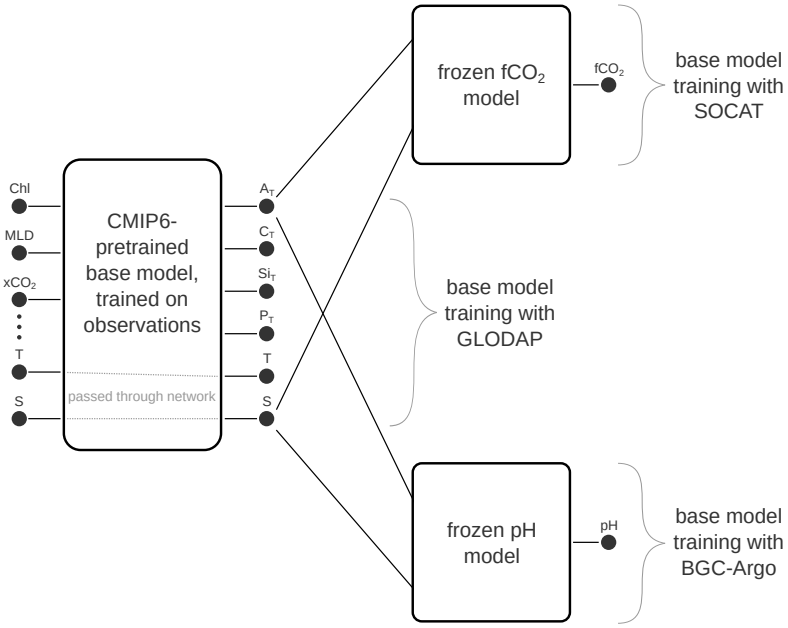

Figure 1: Scheme displaying the overall model architecture of MOCCA.

For fCO_2, model complexity was iteratively increased until a desired maximum deviation of less than $1\,\mu$atm was reached (the measurement uncertainty for pCO_2 as reported by Orr et al, 2015b)[1]. During training, learning rate was decreased from 10^{-3} to 10^{-5} following an exponential learning rate schedule over 10 000 epochs. The decay in learning rate was chosen to shift from an initial identification of an optimal region in the parameter space to finding an optimal set of parameters for which the mean squared error over the training and validation sets converges to a similar and low value.
Hidden layer size was increased from an initial 64 hidden layer units (8833 trainable parameters), 80 units (13601 parameters), 96 units (19393 parameters), 128 units (34049 parameters), to 160 units (52801 parameters). The largest model[2] hit a maximum deviation of $1.08\,\mu$atm over the validation set (Figure 2). The same model architecture was then also used to train the pH model, resulting in a maximum deviation over the validation set of 0.0044.
With root mean squared errors (RMSE) of $0.026\,\mu$atm (fCO_2 model) and 0.00008 (pH model), these surrogate models provide a precision that is comparable to numerical carbonate chemistry packages: Orr et al., 2015b report a desired numerical uncertainty of $0.1\,\mu$atm and 0.0003, respectively.
The accurracy of the fCO_2 and pH neural network models is similar for the six million random samples for A_T, C_T, T, S, Si_T, and P_T from the CMIP6 models (section 2.3). Specifically, RMSE is $0.029\,\mu$atm and 0.00006, respectively, and the maximum deviations are $0.24\,\mu$atm and 0.0006, respectively.

2.3 CMIP6-pretrained DIC and Alk model

2.4 SOCAT-, GLODAP-, and BGC-Argo-based model tuning

3 Model performance

References

I. J. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, Cambridge, MA, USA, 2016. `http://www.deeplearningbook.org`.

[1] For reference, fCO_2 across the generated samples varies between $0.001\,\mu$atm and $5108\,\mu$atm.

[2] To give some context about the model complexity: The number of paramters of this model is comparable to that of a Taylor expansion of a function with six arguments to 15th order. Assuming an efficient use of the MLP parameters, a good fit to the numerical solution from mocsy is thus expected.

	minimum	maximum
A_T	$1000\,\mu\text{mol}\,\text{kg}^{-1}$	$3000\,\mu\text{mol}\,\text{kg}^{-1}$
C_T	$1000\,\mu\text{mol}\,\text{kg}^{-1}$	A_T
T	-2 °C	35 °C
S	10 PSU	50 PSU
Si_T	$0\,\mu\text{mol}\,\text{kg}^{-1}$	$134\,\mu\text{mol}\,\text{kg}^{-1}$
P_T	$0\,\mu\text{mol}\,\text{kg}^{-1}$	$4\,\mu\text{mol}\,\text{kg}^{-1}$

Table 1: Minima and maxima of the uniform distributions used to generate samples for the fCO_2 and pH surrogate models. The range for A_T was chosen such that it easily encompasses open ocean variations in A_T. That for C_T is limited to values lower A_T since larger C_T do not occur in the ocean. The ranges for T and S are chosen according to Lueker et al., 2000, whose parameterizations for K_1 and K_2 are used in mocsy. Finally, the maximum values for Si_T and P_T were chosen to be the global maxima found in the monthly climatologies for Si_T and P_T from World Ocean Atlas 2023. These uniform distributions have means $(max + min)/2$ and standard deviations $(max - min)/\sqrt{12}$, expcept for C_T where mean and standard deviation are given by $min + (max - min)/4$ and $(max - min) \cdot \sqrt{7/144}$, respectively. These means and standard deviations are used for feature normalization.

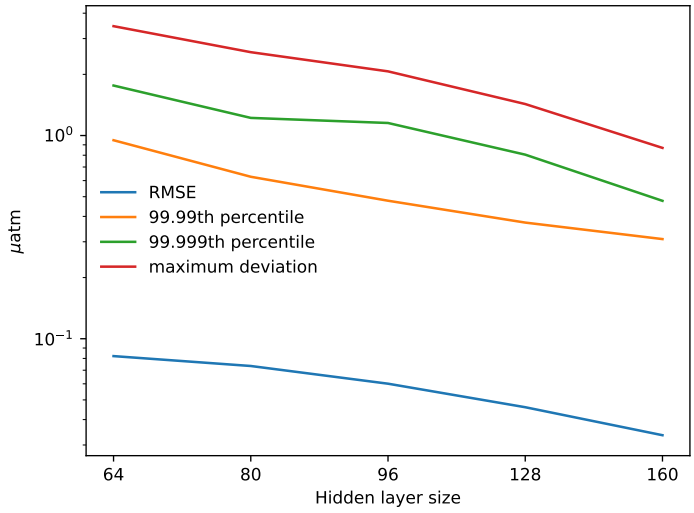

Figure 2: The evolution of root mean squared error (RMSE), the 99.99th and 99.999th percentiles of deviation and the maximum deviation between the surrogate model and mocsy fCO_2 over the validation data set with 1000 000 randomly generated samples.