CALCULATING SEAWATER CARBON CONTENT USING SATELLITE SENSOR DATA

Dr C K Gomathy, Assistant Professor, SCSVMV Deemed to be University, India
Ms. Lekha Rajeswari, Ms. Sudha Mounika, Ms. Renu Sree, Ms. Geetha
UG Scholars- SCSVMV Deemed to be University, India.

ABSTRACT

There are numerous effects of chemical reactions in ocean surface waters. The importance of carbon dioxide and its constituents, as well as the potential consequences they may have on climate change, has recently increased. Key sea-surface data can be measured in close to real-time by remote sensing, allowing for the estimation of carbon levels. Hybrid satellite sensor data is used to illustrate. We train our algorithms using data gathered from cruise ships as the "ground truth" in order to validate our findings. The proposed approach can be utilised to estimate Carbon levels in any ocean because our predictor's error rate is shown to be low. The use of sea water salinity as a stand-in for carbon calculations is one way that this work advances prior research. The Carbon content of surface water is predicted using binary combinations of typically unary predictor variables, and the relationship is quantified using an intrinsically non-linear model.

Keywords: Climate Change, Carbon emissions, Ocean acidification

I. INTRODUCTION

Ocean acidification and climate change are both serious issues. The carbon cycle affects the climate, environment, and overall quality of life for humans. Furthermore, the main anthropogenic source of carbon causing climate change is CO2. Prior to civilization, atmospheric CO2 levels fluctuated between 200 and 280 ppm for 400,000 years. However, due to primarily anthropogenic sources, current levels are now approaching 300 ppm. The extent of fluctuation attributable to natural sources is greater than the amount of anthropogenic carbon sequestered by the seas during the past millennium, further indicating oceanic saturation. The longest record of in-situ carbon measurements, which dates back to 1960, demonstrates that the rate of rise now is up to 30 times greater than it was before industrialization. Approximately 48% of the world's carbon emissions are attributed to the manufacturing of cement and fossil fuels. Additionally, over the past 200 years, carbon emissions have increased at an unparalleled rate due to deforestation, industrialization, and changes in land use.
Surface-water the partial-pressure of carbon dioxide, or pCO2, at the surface is an accurate representation of carbon dioxide and reflects both its thermodynamic and biological influences. Physical mixing processes brought on by

- Sea Surface Temperature - SST,
- Sea Surface Salinity - SSS,
- Chlorophyll-a - Chl-a,
- Mixed Layer Depth - MLD,
- Coloured Dissolved Organic Matter - CDOM,
- Net Primary Productivity - NPP,
- Photo-synthetically Active Radiation - PAR, wind speed and other factors.

Observations at the surface are useful proxies for oceanic carbon content even though sea-surface measurements may not fully account for biological processes. This is because changes in carbonate chemistry brought on by atmospheric CO2 occur first at the surface. Thus, the use of orbiting platforms to collect data via remote sensing has considerable promise for tracking changes in ocean chemistry.

II. DESCRIPTION OF DATASETS

**The HYCOM dataset**

A multi-institutional project supported by the National Ocean Partnership Program (NOPP) as part of the US Global Ocean Data Assimilation Experiment is HYCOM, a data-assimilative hybrid isopycnal-sigma-pressure model (GODAE). HYCOM data are available on a standardised grid in non-extreme latitudes.

Data on the space-time variability of surface-wind stress, temperature, and specific humidity are gathered from the remote sensing platforms GFO, ENVISAT, and Jason-1. Maps of subsurface variability are improved by vertical profiles from expendable Bathythermographs, Argo floats, and Conductivity-Temperature-Depth sensors. These profiles, however, are often too brief to be used on their own.

**NOAA labelled data**

The National Oceanic and Atmospheric Administration’s (NOAA) Ocean Chemistry and Ecosystems Division (OCED) focuses on comprehending the function of the ocean in relation to the overall environment. The NOAA installed automated pCO2 measurement systems on cruise ships owned by Royal Caribbean
International Cruises and its affiliates Wanninkhof et al. From 2002 to the present, this system provides measurements of numerous ocean water parameters.

Equilibrators, a condenser, a water flow metre, drying tubes, and extra apparatus for analysing the equilibrator's output make up the instruments. Data from the Allure of the Seas for the Caribbean region for 2019–2020 was used for this study.

<table>
<thead>
<tr>
<th>Lat</th>
<th>lon</th>
<th>To</th>
<th>So</th>
<th>T2</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.520</td>
<td>-68.327</td>
<td>26.590</td>
<td>35.784</td>
<td>26.476</td>
<td>35.788</td>
</tr>
<tr>
<td>25.148</td>
<td>-77.325</td>
<td>25.252</td>
<td>36.602</td>
<td>25.071</td>
<td>36.608</td>
</tr>
<tr>
<td>22.559</td>
<td>-71.743</td>
<td>26.549</td>
<td>36.584</td>
<td>26.405</td>
<td>36.585</td>
</tr>
</tbody>
</table>

III. METHODS

A method for supervised learning simulates the implicit link found in labelled data. Data using a series of equations. In light of a collection of labelled data D, supervisedThe purpose of a learning algorithm is to discover the connection between input qualities X and an order to forecast the output y given previously unobserved X, output attribute y. A loss function provides feedback to supervised learning algorithms, which quantitatively determines how well the model fits the patterns in the data. In our if the output property y is the fugacity of the carbon dioxide partial-pressure this is determined by a feature-selection process and depends on the input attributes. We employ a gradient-boosting regression tree that offers a continuous-valued regression model. Output as a nonlinear function of the characteristics of its input.

IV. DECISION-TREE REGRESSION

The supervised gradient-boosting regression tree (GBRT) gradually improves itself using an explicit loss function after iteratively defining a series of mappings from a labelled training dataset. Through a sequence of conditional, thresholding processes, the GBRT learns an input-output relation. In other words, an information criterion is used to narrow the output domain at each stage of the regression tree. The stochastic Gradient Boosted Regression Tree generates models that are very accurate, adaptive, and simple to understand. The Stochastic GBRT begins with a single decision-tree and incrementally adds new trees dependent on how well they perform relative to an objective function.
V. LOSS FUNCTION

A learning algorithm's performance is measured using a loss function. The Mean-Absolute Error and Root Mean Squared Error have gained popularity in the field of geosciences despite the fact that many loss functions have been developed for supervised learning. The MAE measure gives each error an equal weighting (known as l1 optimization). The loss value is directly inversely proportional to the size of divergence between the anticipated value $y$ and the ground truth label $y$, if we represent ground-truth values as $y$ and forecast values as $y$.

VI. RESULTS

Our model was validated using in-situ measurements that are grounded in reality. This target data was expressed as fCO$_2$ in microatm (-atm) units. The equilibrator's measurement of fCO$_2$ had a range of values with a mean of 400.73 and a standard deviation of 19.78. For model validation, about one-third of the training data were set aside. Cross-validation was used to train the model, and the Huber Loss function was used to assess it. The Huber Loss for fCO$_2$ was found to be 3.98, or roughly 1% of the mean value for the target characteristic, which is highly accurate using the chosen attributes and the best GBRT.

VII. DISCUSSION

As of this writing, there is no universal agreement on how SST, SSS, and fCO$_2$ are related. This data period captured a whole annual cycle, protecting against seasonal oceanic occurrences while accounting for annual meteorological events. Utilizing in-situ surface data, our model's low variance and error terms are confirmed. The validation mentioned in Section 5.3 above demonstrates accuracy higher than that attained by a comparable global model developed by Liu W. Timothy and Krishna et al. Our work enhances both of these methods by utilising SSS for all data. Additionally, by taking into account binary aspects of influence and applying an intrinsically non-linear model, this work improves upon earlier methodologies.

VIII. CONCLUSION

Using in-situ observations in the Caribbean, a non-linear model is constructed using satellite data and validated in this article. This model utilised the connection between to remotely anticipate the carbon content of surface water, consider temperature, salinity, and sea-surface carbon concentration. When compared to worldwide techniques, we discovered that estricting the scope of observations to non-extreme latitudes produced smaller losses.
Additionally, we discovered that using binary predictor features greatly decreased prediction errors as compared to earlier methods. Additionally, the Huber loss-function and use of an intrinsically non-linear model outperformed earlier methods that relied on linear penalised models. Lastly, our findings were verified utilising a plentiful source of measures taken in-situ and made available by the NOAA.

IX. REFERENCES


