Daily Model of Xco_2 Based on Orbital Observations for the State of São Paulo – Brazil

Luis Miguel Da Costa (luism.costa00@gmail.com)
Sao Paulo State University Julio de Mesquita Filho https://orcid.org/0000-0002-0698-4616
Gustavo André de Araújo Santos
Sao Paulo State University Julio de Mesquita Filho
Alan Rodrigo Panosso
Sao Paulo State University Julio de Mesquita Filho
Glauco de Souza Rolim
Sao Paulo State University Julio de Mesquita Filho
Newton La Scala Jr.
Sao Paulo State University Julio de Mesquita Filho

Research Article

Keywords: Carbon cycle, Remote Sensing, OCO-2, Stepwise Regression Analysis, Climate Change, Meteorology.

Posted Date: January 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1244068/v1

Abstract

Background: Understanding the behavior of carbon dioxide (CO_2) it is crucial to create strategies to deal with climate change. Several space missions were designed to monitor this and other greenhouse gases and studies at global and large regional scales were developed using these remote sensing tools. However, there is still a time gap in orbital data, since the revisit time of a satellite at the same location for this type of observation is not daily. Daily measurements of CO_2 can be made using Eddy Covariance technique, although this type of study is at a very local scale. In this study, we aimed to build a daily model to estimate the natural CO_2 in the atmosphere.

Results: The data was retrieved from the Orbiting Carbon Observatory-2 (OCO-2), NASA-POWER and, Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) and consist in a time-series from 2015 to 2019. To summarize the most relevant factor we used the Variance Inflation Factor (VIF) after we performed the Pearson's Correlation and apply descriptive statistics in the significant variables ($p < 0.05$). The model was construed using the stepwise regression method and the selected model was defined by the lowest RMSE in training (~0.6 ppm). The most related variables with Xco_2 were Global Radiation (Qg), Sun Induced Chlorophyll Fluorescence (SIF) and, Relative Humidity (RH), all these factors were negatively correlated with the CO_2 concentration. The model has the best perform with the Qg and RH (RMSE = 0.47 ppm, $R^2 = 0.44$, $p < 0.01$).

Conclusion: In summary, the cycle of atmospheric CO_2 in the state of São Paulo has higher average values during April to October, and the lowest averages of Xco_2 were usually observed between December to March and the inverse behavior was observed for SIF 757, Global Radiation (Qg) and Relative Humidity (RH). Concerning the daily model, despite the differences between the spatial observations, the model derived here was capable to estimate the cycle of atmospheric CO_2 using only meteorological data.

Background

Understanding the variability of carbon dioxide (CO_2) in time and space is a crucial task so that we can adopt mitigation strategies. In this sense, several studies analyze the average concentration of this greenhouse gas not only on a global scale [1, 2] but also to estimate anthropogenic emissions in urban centers [3, 4]. Other studies focus on understanding the column-averaged of carbon in the atmosphere (Xco_2) above tropical forests [5], or above agriculture crops in different seasons of the year [6, 7]

In a recent regional study, da Costa et al. (2021) [7] analyze the spatio-temporal variability of Xco_2 in a sugarcane-producing area in the southeast region of Brazil. They observed an important inverse relationship between the average carbon concentration in the atmosphere with climatic and vegetative variables. Concluding that the dependence of the natural carbon cycle is related to the predominant agriculture crop in the region and how global radiation (Qg), relative humidity (RH), and the Sun Induced Chlorophyll Fluorescence (SIF) was related to this behavior. Similarly, Morais Filho et al. (2021) [6]

conducted a study that analyzed three different crops and the temporal variability of Xco_2 and SIF in these environments, they also found a significant negative correlation between these variables.

However, there is still a temporal gap in the Xco_2 data collected by remote sensing, since the measurements are not daily [8, 9]. This type of measurement is important to several factors, such as, estimate the potential capability of atmospheric CO_2 assimilation by vegetation, establishing public strategies at local levels for climate adaptation and mitigation, and even in economy incorporating daily trends in the carbon market and ecosystems services payments[10, 11, 12, 13, 14, 15, 16].

Daily CO_2 measurements can be made using the Eddy Covariance technique [17, 18, 19], although this has the disadvantage of being a point (local) study. In this sense using orbital data modeling to estimate the daily variation of several aspects (e.g, climate, meteorological, land-use changes, ecosystems services) for an entire region has become common [16, 20, 21, 22, 23].

Several studies confirm that photosynthesis is the main regulator of atmospheric carbon sinks [10, 24, 25, 26]. However, photosynthesis is a process sensitive to climatic variations such as relative humidity [27], precipitation [28], evapotranspiration [29], and incident solar irradiance [30].

Therefore, the natural cycle of CO_2 is dependent on several aspects, such as vegetation and climate, being necessary data from several different bases for understanding this dynamic [7], and with this the problem of which variables we should use to model it, turning pre-processing techniques and analysis of autocorrelations necessary. In this sense, we aim to model the atmospheric CO_2 cycle above the state of São Paulo to estimate the time changes on a daily scale, based on vegetative and climatic variables retrieved from different orbital platforms, improving in this way the regional understanding of CO_2.

Results

Variance Inflation Factor (VIF) analysis (Table 1) shows it was possible to reduce the number of variables related to Xco_2 (according to the adopted criterion, VIF < 10) as shown comparing Figure 1a with 1b, before and after the selection, respectively. Of the selected variables the only one that relates positively with Xco_2 is wind speed, however, it was not significant ($p > 0.05$). Variables most related to X_{CO2} were the Global Radiation (Qg), Sun-Induced chlorophyll Fluorescence at 757 nm (SIF 757), and Relative Humidity (RH).

Table 1
Variance Inflation Factor (VIF) of studied variables

Variable	VIF
Qg	**9.351368**
RH	**5.1076**
SIF 757	**1.811858**
Prec	10.12886
Temp	21.42606
Ws	**4.067037**
LST	19.54074
NDVI	22.1573
LAI	87.21371
Fpar	65.40795
ET	33.47052

Where: Qg=global radiation; RH=relative humidity; SIF 757 = Solar-Induced Chlorophyl Fluorescence at 757nm; Prec= Precipitation; Temp = Temperature at 2m; Ws = Wind Speed; LST = Land Surface Temperature (MODIS); NDVI = Normalized Difference Vegetation Index; LAI = Leaf Area Index; ET=Evapotranspiration.

Regarding the temporal variability of Xco_2, the maximum mean for the analyzed period was 393.09 ± 0.17 ppm and occurred in October 2019, while the minimum average was in November 2018, being 390.11 ± 0.15 ppm (Figure 2a). Meanwhile, the Qg (Figure 2b) ranged between 24.3 ± 0.09 and 13.07 ± 0.04 (MJ m^{-2} day^{-1}), with the maximum average occurring in December 2018 and the minimum in June of the same year.

SIF 757 (Figure 2d) had the highest average recorded in the period in November 2015 (1.1 ± 0.05 (Wm^{-2} sr^{-1} μm^{-1})) and the lowest in September 2017 (0.3 ± 0.06 (Wm^{-2} sr^{-1} μm^{-1})), while the Relative Humidity (Figure 2c) ranged from 84.86 ± 0.07 to 70.44 ± 0.19%, where the highest mean was observed in March 2016 and the lowest in October 2019.

Regarding SIF 757 the minimum averages occurred in June of 2015 and 2016, September 2017, November 2018, and July 2019, ranging from 0.3 to 0.46 Wm^{-2} sr^{-1} μm^{-1} (Figure 2d). The minimum Qg averages vary between May and June for the entire series approximately between 13.07 to 14.71 MJ m^{-2} day^{-1} (Figure 2b). Maximum Qg averages accumulate between December and January of each year, reaching 24 MJ m^{-2} day^{-1} in those months (Figure 2b). SIF 757 varies its maximum averages between November and February of each year, ranging from 0.8 to 1.1 Wm^{-2} sr^{-1} μm^{-1} (Figure 2d).

With multiple model building using the stepwise forward selection method, with training control, the best result was found with two variables, with RMSE of ~0.60 ppm for training (Figure 3), the selected variables being Qg and RH respectively (Eq 1).

$$Xco_{2\,(daily)} = 391.484(\pm 0.89) - 0.283\ (\pm 0.089) \times Qg - 0.263(\pm 0.09) \times RH \quad \text{(Eq. 1)}$$

After the selection of the best variables in the training, using as metric the best RMSE performance, the model built in the training was applied in the test sample of the variables cited (Qg and RH), and from the cross-validation of the estimated data with the observed data, we obtained an R^2 of 0.44, the values of the metrics MSE, RMSE, and MAE were respectively 0.22, 0.47, and 0.37 (ppm), and for MAPE we found a value of 1.54% (p-value < 0.01) (Figure 4a), with this we were able to reduce the time scale of the OCO-2 satellite from every 15 days to a daily scale (Figure 4b).

Discussion

The natural CO_2 cycle is affected by factors related to climate and vegetation aspects [6, 31, 32]. Due to the VIF analysis, we were able to summarize three main factors for São Paulo state: Global Radiation (Qg), Relative Humidity (RH) and Sun-Induced chlorophyll Fluoresce at 757 nm (SIF 757). Several studies have already been conducted using this method to identify which variables select for ecological studies [33], computational studies [34], and remote sensing studies [35].

Except for wind speed, all variables studied correlated negatively with Xco_2 (Figure 1), hence, related to the sink of atmospheric CO_2. In general, the highest concentrations of Xco_2 are observed in the months corresponding to the Brazilian autumn and winter (April to August) and lowest in the summer, from December to February. Studies such as by Siabi et al. (2019) [36] and Falahatkar et al. (2017) [37] reported how the different seasons affect the average CO_2 concentration in the atmosphere

Recently, researches were conducted at regional scales in Brazil such as by Morais Filho et al. (2021) [6] and da Costa et al. (2021) [7], indicating negative correlations between Xco_2 and SIF over agricultural areas, approximately -0.5 and -0.8 respectively. SIF is a variable directly related to the photosynthesis of plants, laboratory-scale experiments have demonstrated this relation [38], and remote sensing studies at the canopy and global level reported positive relations between SIF and Gross Primary Production, also a negative correlation between SIF and the Xco_2 [5, 39, 40, 41].

As a result of photosynthesis, it is expected that SIF increases during summer [7, 36], as in this season, higher precipitation events and higher temperatures are observed [42]. Our results show higher SIF average values in the months where summer occurs in São Paulo state, and an inverse relationship between SIF and Xco_2. The lowest average values of Xco_2 usually occur during the summer period in the study region. This is due to plant CO_2 assimilation [43], printing a quasi-periodical Xco_2 and SIF time changes as well as observed in other studies [5, 6, 36, 44].

Most of São Paulo's state has a wet summer and dry winter [42] resulting in a positive correlation between precipitation and SIF (Pearson's correlation = 0.61 and $p < 0.05$), while negative with Xco_2 ($r = -0.49$, $p < 0.05$) (Figure 1a). Precipitation is a photosynthetic control factor, so the greater availability of water that exists in the summer at São Paulo's state induces plants to perform more photosynthesis through primary productivity, which leads to a reduction of atmospheric CO_2. The opposite is observed in the dry winter because water availability is lower resulting in less photosynthesis, or less CO_2 assimilation by plants, either in natural or agricultural areas [7, 28, 45].

Another effect observed during summer in the region is the increase of relative humidity (RH), which reduces the water transfer between soil or plant to the atmosphere [46], inducing plants to keep their stomata open, where CO_2 assimilation occurs [47]. Studies have already shown the relationship of stomata opening in periods with good water availability is related to plant growth [48, 49]. Thus, establishing the negative relationship between RH and Xco_2, also previously observed by Golkar et al. (2020) [27].

In the same way, another requirement for photosynthesis occurs is sunlight which is the source of energy to carry out the biochemical processes of this phenomenon. Therefore, as the amount of radiation (Qg) is absorbed by the plant, photosynthesis tends to increase, and consequently higher assimilation of CO_2 and decreasing the concentration of this greenhouse gas on the atmosphere [7, 30]. We can observe these relationships in our results (Figure 3b), Qg correlates positively with SIF, and those variables relate negatively with Xco_2.

Since we are dealing with the natural cycle of CO_2 the main factor of the higher concentrations of this gas in the atmosphere is due to the lowest photosynthetic absorption by plants. The autumn and winter have low available water and sunlight for plants, leading to a decrease in photosynthesis, also another important factor is that the annual calendar for agriculture in the state of São Paulo have harvest periods between these seasons [50], and as consequence decreasing the cover area by vegetation. Shekhar et al. (2020) [51] show how the crop's grown in summer decrease the values of Xco_2 over Nile Delta and when the harvest starts the values of Xco_2 are higher, also, they found that SIF values are higher in the grown season.

Our model was based on Qg and RH, which are two variables related to the CO_2 assimilation process, or CO_2 sink. The model has lower RMSE values than have been reported in previous studies, such as by Guo et al. (2012) [52] where the values of this metric ranged from 0.7 to 1.1 ppm. In a more recent study by Taylor et al (2020) [53] when evaluating initial OCO-3 data results from the globe and model-related errors, they found an RMSE between 1 and 2 ppm. Another important measure is the MAPE, which shows in percentage how much we are getting wrong, studies with remote sensing have already demonstrated errors below 10% as being considered extremely low for predicting various plant and climate aspects [54, 55]. With this, we can evaluate that the performance of the model proposed in this work presents a very low error.

The coefficient of determination (R^2) was 0.44, an increment of almost 20% from the simple linear fit with Qg alone with a higher importance in the model. Although the R^2 is moderate, studies using other orbital sensors such as MODIS to model the average CO_2 concentration in the atmosphere have reported similar results (Guo et al., 2015) [23]. In addition, we should consider that although OCO-2 and NASA-POWER are two high quality and validated databases [8, 9, 56], the difference between grids and spatial resolution (see Table 2 in methods) cannot be disregarded, as it is an aspect that can influence these results, leading us to consider the coefficient of determination observed in this study as being high.

These differences between the databases can be suppressed by the greater temporal coverage of NASA-POWER, being able to estimate the daily temporal variability of the natural CO_2 cycle in the atmosphere for the state of São Paulo, besides reducing in the future the spatial scale of Xco_2 obtained from OCO-2 and gaining greater spatial resolution cover. Other vegetation index-based models aimed at reducing the spatial sampling of OCO-2 data, but focused on SIF, as is the case of Zhang et al. (2018) [57] and Yu et al. (2019) [58].

Despite the errors associated with the model and the uncertainty measures due to the difference in satellite resolution, an advantage of using models similar to the one proposed here is being able to have a daily measure of the variability of atmospheric CO_2 and how the climate parameters affect this dynamic, also serving as an indirect indicator of how is the daily assimilation capacity of this gas in a region.

Conclusions

In summary, the cycle of atmospheric CO_2 in the state of São Paulo has higher average values during April to October, periods of lower intensity of rainfall and considered as the winter in the state, in the other hand the lowest averages of Xco_2 were usually observed between December to March, this period corresponds to the summer, and the inverse behavior was observed for SIF 757, Global Radiation (Qg) and Relative Humidity (RH). This pattern is due to the relationship between photosynthesis and carbon assimilation, given that photosynthesis is a process sensitive to climate variation and a process that depends on water and light, in summer this process tends to be greater, leading to a decrease in CO_2.

Concerning the daily Xco_2 model presented, the model performed well when we looked at the set of metrics presented. Given this, we were able to estimate what the daily behavior of natural CO_2 is like in general for the state of São Paulo, a semi-periodical wave with a maximum peak between March and July, and a minimum peak between December to February. Therefore, we successfully estimate the behavior of natural carbon dioxide in a synthetic way using daily meteorological open access data, establishing a low coast way to estimate this greenhouse gas. There are still challenges in this aspect, but this study will serve as a basis for further implementations.

We suggest that for future work, that the relationship between soil respiration and factor controlling organic matter decay in soil with the Xco_2 would be needed to better understand CO_2 dynamics, as well

the addition of variables related to activities, such as in transports in big cities to improve predictions.

Methods

Study Region

The state of São Paulo (SP) (Figure 5), southern Brazil, has approximately 249×10^3 km^2 and 645 municipalities, with a demographic density of 179.84 habitants per km^2 [59] being one of the main agricultural hubs of Brazil, regarding the production of sugarcane and citrus [60]. According to Rolim et al. (2016) [42] the climate of the state, in general, has its areas characterized with humid subtropical climate with dry winter, followed by humid tropical dry winter and sub-humid tropical dry winter, according to the climate classification proposed by Camargo (1991) [61].

Products of remote sensing: acquisition and processing

Greenhouse gas, climate, and vegetation data were collected from different satellites (Table 2) for a time series from 2015 to 2019. The primary product of the Orbiting Carbon Observatory-2 (OCO-2) consists of georeferenced estimates of the mean atmospheric CO_2 concentration (Xco_2), in addition, the Sun Induced Chlorophyll Fluorescence (SIF), retrieved due to the overlap that occurs in the SIF wavelengths with the O_2 absorption wavelength (680~850nm) [8, 9, 38]. Data from this satellite have already been validated by Crisp et al. (2012) [8] and, according to O'Dell et al. (2012) [9], this satellite provides about 65000 quality observations per day worldwide.

MODIS sensor data were extracted from the “Application for Extracting and Exploring Analysis Ready Samples” (AppEEARS). This application allows users to obtain subsets of large databases using spatial and temporal parameters. Two types of sample requests are available: point samples by entering geographic coordinates and area samples using vector polygons. Sample requests submitted to AppEEARS provide users with not only data values but also associated quality data values. Interactive visualizations with summary statistics are provided for each sample within the application, which allows users to view and interact with their samples before downloading the data [62].

Nasa Power data (https://power.larc.nasa.gov) consists of precipitation (mm), surface solar shortwave irradiance (MJ m^{-2} d^{-1}), average air temperature (ºC), and relative humidity at 2 m (%). This platform consists of a NASA project entitled: Worldwide Energy Resource Forecast (POWER) and was initiated to enhance the current renewable energy dataset and create new datasets from new satellite systems [63].

Table 2. Studied Variables, Data Base, Temporal and Spatial resolution

	Variable	Data base	Temporal Resolution	Spatial Resolution
GHG	X_{CO2} (ppm)	OCO-2 “OCO-2 Data product user’s guide, 2016” V9	16 days	1,29 km ×2,25 km
CLIMATE	Surface solar shortwave irradiance (Global radiation, Qg) (MJ m^{-2} d^{-1})	FLASH Flux Version 3 (A, B,C) NASA/POWER	Daily	111,3 km × 111,3 km
	Average Air Temperature at 2m (Temp) (ºC)	GEO-5 FP-IT (NASA/POWER)	Daily	111,3 km × 111,3 km
	Land Surface Temperature (LST) (ºC)	MOD11A1.006 V6 MODIS-TERRA	Daily	1200 km × 1200 km
	Wind Speed at 10 meters (WS) (m s^{-1}) (ºC)	GEO-5 FP-IT (NASA/POWER)	Daily	111,3 km × 111,3 km
	Relative Humidity (RH) (%)	GEO-5 FP-IT (NASA/POWER)	Daily	111,3 km × 111,3 km
	Precipitation (Prec) (mm day^{-1})	GEO-5 FP-IT (NASA/POWER)	Daily	111,3 km × 111,3 km
VEGETATION	SIF 757	OCO-2 “OCO-2 Data product user’s guide, 2016” V9	16 days	1,29 km × 2,25 km
	LAI (m^2 m^{-2})	MCD15A2H.006 V6 MODIS- CFPAR	8 days	500 m × 500 m
	Fraction of Photosynthetically Active Radiation (Fpar) (%)			
	Evapotranspiration (ET) (kg m^{-2} day^{-1})	MOD16A2.006 V6 MODIS-TERRA	8 days	500 m × 500 m
	NDVI	MOD13A1.006 V6 MODIS-TERRA	16 days	500 m × 500 m

To minimize the differences between the spatial and temporal resolutions of the different orbital sensors used in this study, the process described in Figure 5 was employed, which establishes a standard in the acquisition of data from the coordinates obtained in the OCO-2 platform. We emphasize that several studies have been conducted using different time and spatial scales [6, 7, 27].

Pre-process of the data

Using the regression method proposed by Gujarati and Potter (2011) [64], we removed the trend from X_{CO2} data, in order to understand the regional variability of Xco_2 and its relationships with other factors [6, 7]. The other variables were standardized using the function *scale* from the R language [65].

Variance Inflation Factor (VIF)

Variance Inflation Factor (VIF) analysis was performed. This analysis is a method of detecting multicollinearity within a database since the relation between the predictors for a multi-regression model can affect the estimative and the standard errors associated with the regression model [66]. The VIF is based on the R^2 value (Eq. 2), and should not be greater than 10, however, this can vary according to the study [66, 67].

$$VIF = \frac{1}{1-R^2} \qquad \text{(Eq. 2)}$$

where R^2 is the coefficient of determination.

Temporal variability, Pearson's correlation, and dependency analysis

The data was processed using month averages for the analysis period, except precipitation, which consists of monthly sums for the entire state of SP (ST.1). The means were subjected to analysis of variance (F-test) to obtain the mean standard errors. Simultaneously, the basic assumptions of analysis of variance and, normality of errors, and homogeneity of variances were tested for the selected variables by VIF analysis. To understand the variation of Xco_2 with the other variables, Pearson correlation analyses were performed.

Stepwise: forward selection

The stepwise method used in this study was the forward selection method being performed in R language [65], as can be seen in the flow chart (Figure 6), the variables selected in the VIF analysis were separated into a training and test sample. The training sample was submitted to the *train ()* function of the *caret* package, with training control using the Cross-Validation (cv) method. The number of variables that make up the model is based on the lowest RMSE. From variables selected in training, the generated model is applied to the test sample and Xco_2 is estimated with these independent data. Finally, cross-validation between the estimated data and observed data in the test sample was performed and from this, we derive the metrics Mean squared error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2, and Mean absolute percentage error (MAPE).

Abbreviations

AppEEARS = Application for Extracting and Exploring Analysis Ready Samples

CO_2 = carbon dioxide

ET=Evapotranspiration

LAI = Leaf Area Index

LST = Land Surface Temperature (MODIS)

MAE = Mean Absolute Error

MAPE = Mean absolute percentage error

MSE = Mean squared error

NASA-POWER = NASA project entitled: Worldwide Energy Resource Forecast

NDVI = Normalized Difference Vegetation Index

OCO-2 = Orbiting Carbon Observatory-2

OCO-3 = Orbiting Carbon Observatory-3

Prec= Precipitation

Qg=global radiation;

RH=relative humidity

RMSE = Root Mean Squared Erro

SIF 757 = Solar-Induced Chlorophyl Fluorescence at 757nm

Temp = Temperature at 2m

VIF = Variance Inflation Factor

Ws = Wind Speed at 2m

Xco_2 = Column average of carbon dioxide in the atmosphere

Declarations

Ethical Approval: Not applicable

Consent to Participate: Not applicable

Consent to Publish: Not applicable

Data and Software availability: The processed data can be found in supplementary tables attached to this paper. The Xco_2 and SIF was retrieved from: https://co2.jpl.nasa.gov/build?mission=oco-2&dataset=OCO2L2Stdv10; The MODIS products from: https://lpdaacsvc.cr.usgs.gov/appeears/; the NASA-POWER from: https://power.larc.nasa.gov/; The VIF and regression was made in R language and the code is available at: https://github.com/lm-costa/stepwise_fapesp

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Funding: This work was supported by the São Paulo Research Foundation (FAPESP) [Grant number: 2019/25812-4] and CNPq - National Council for Scientific and Technological Development [Grants: 304075/2018-3, 311981/2020-8].

Authors Contributions

L.M.C: Writing-Original draft, Conceptualization, Methodology, Investigation, Writing-Review, and Editing; **G.A.A.S:** Methodology, Investigation, Writing-Review, and Editing; **A.R.P.**: Conceptualization, Methodology Writing-Review, Editing, and Supervision; **G.S.R.:** Conceptualization, Methodology Writing-Review, Editing, and Supervision; **N.L.S.Jr.:** Conceptualization, Methodology Writing-Review, Editing, and Supervision

Acknowledgments: All authors are grateful to the São Paulo State University (UNESP), Jaboticabal Campus, for the academic support and the open data provide by the NASA projects: OCO-2, NASA-POWER, and AppEEARS.

References

1. Eldering A, O'Dell CW, Wennberg PO, Crisp D, Gunson MR, Viatte C, Avis C, Braverman A, Castano R, Chang A, Chapsky L, Cheng C, Connor B, Dang L, Doran G, Fisher B, Frankenberg C, Fu D, Granat R, Hobbs J, Lee RAM, Mandrake L, McDuffie J, Miller CE, Myers V, Natraj V, O'Brien D, Osterman GB, Oyafuso F, Payne VH, Pollock HR, Polonsky I, Roehl CM, Rosenberg R, Schwandner F, Smyth M, Tang V, Taylor TE, To C, Wunch D, Yoshimizu J. The Orbiting Carbon Observatory-2: first 18 months of science data products. Atmos Meas Tech. 2017;10:549–63. https://doi.org/10.5194/amt-10-549-2017.
2. Boesch H, Baker D, Connor B, Crisp D, Miller C. Global Characterization of CO2 Column Retrievals from Shortwave-Infrared Satellite Observations of the Orbiting Carbon Observatory-2 Mission. Remote Sensing 2011. 2011;3(2):270–304. https://doi.org/10.3390/RS3020270. Pages 270-304, 3) .
3. Ye X, Lauvaux T, Kort EA, Oda T, Feng S, Lin JC, Yang EG, Wu D. (2020). Constraining Fossil Fuel CO2 Emissions From Urban Area Using OCO-2 Observations of Total Column CO2. Journal of Geophysical Research: Atmospheres, 125(8), e2019JD030528. https://doi.org/10.1029/2019JD030528.

4. Kort EA, Frankenberg C, Miller CE, Oda T. (2012). Space-based observations of megacity carbon dioxide. Geophys Res Lett, 39(17). https://doi.org/10.1029/2012GL052738.
5. Parazoo NC, Bowman K, Frankenberg C, Lee JE, Fisher JB, Worden J, Jones DBA, Berry J, Collatz GJ, Baker IT, Jung M, Liu J, Osterman G, O'Dell C, Sparks A, Butz A, Guerlet S, Yoshida Y, Chen H, Gerbig C. Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of XCO2 and chlorophyll fluorescence from GOSAT. Geophys Res Lett. 2013;40(11):2829–33. https://doi.org/10.1002/grl.50452.
6. Morais Filho LFF, de Meneses KC, de Santos GA, Bicalho AE da, Rolim GS de., & La Scala N. (2021). xCO2 temporal variability above Brazilian agroecosystems: A remote sensing approach. Journal of Environmental Management, 288, 112433. https://doi.org/10.1016/j.jenvman.2021.112433.
7. da Costa LM, de Araújo Santos GA, de Mendonça GC, et al. Spatiotemporal variability of atmospheric CO2 concentration and controlling factors over sugarcane cultivation areas in southern Brazil. Environ Dev Sustain. 2021. https://doi.org/10.1007/s10668-021-01677-6.
8. Crisp D, Fisher BM, O'dell C, Frankenberg C, Basilio R, Bösch H, Brown LR, Castano R, Connor B, Deutscher NM, Eldering A, Griffith D, Gunson M, Kuze A, Mandrake L, Mcduffie J, Messerschmidt J, Miller CE, Morino I, Natraj V, Notholt J, O'Brien DM, Oyafuso F, Polonsky I, Robinson J, Salawitch R, Sherlock V, Smyth M, Suto H, Taylor TE, Thompson DR, Wennberg PO, Wunch D, Yung YL. The ACOS CO2 retrieval algorithm – Part II: Global XCO2 data characterization. Atmospheric Measurement Techniques. 2012;5:687–707. https://doi.org/10.5194/amt-5-687-2012.
9. O'Dell CW, Connor B, Bösch H, O'Brien D, Frankenberg C, Castano R, Christi M, Eldering D, Fisher B, Gunson M, McDuffie J, Miller CE, Natraj V, Oyafuso F, Polonsky I, Smyth M, Taylor T, Toon GC, Wennberg PO, Wunch D. Description and validation against synthetic observations. The ACOS CO2 retrieval algorithm – Part 1. Atmospheric Measurement Techniques. 2012;5:99–121. https://doi.org/10.5194/amt-5-99-2012.
10. Li X, Xiao J, Fisher JB, Baldocchi DD. (2021). ECOSTRESS estimates gross primary production with fine spatial resolution for different times of day from the International Space Station. Remote Sens Environ, 258. https://doi.org/10.1016/j.rse.2021.112360.
11. Wassie YT, Adaramola MS. Socio-economic and environmental impacts of rural electrification with Solar Photovoltaic systems: Evidence from southern Ethiopia. Energy Sustain Dev. 2021;60:52–66. https://doi.org/10.1016/j.esd.2020.12.002.
12. Rupp M, Rieke C, Handschuh N, Kuperjans I. (2020). Economic and ecological optimization of electric bus charging considering variable electricity prices and CO2eq intensities. Transportation Research Part D: Transport Environment, 81. https://doi.org/10.1016/j.trd.2020.102293.
13. Smith P, Calvin K, Nkem J, et al. Which practices co-deliver food security, climate change mitigation and adaptation, and combat land degradation and desertification? Glob Change Biol. 2020;26:1532–75. https://doi.org/10.1111/gcb.14878.
14. Li Q, Wu S, Lei Y, Li S. (2020). Dynamic features and driving forces of indirect CO2 emissions from Chinese household: A comparative and mitigation strategies analysis. Sci Total Environ, 704.

https://doi.org/10.1016/j.scitotenv.2019.135367.

15. Cernusak LA. Gas exchange and water-use efficiency in plant canopies. Plant Biol. 2020;22(S1):52–67. https://doi.org/10.1111/plb.12939.
16. Parras R, de Mendonça GC, Araújo Costa RC, Pissarra TCT, Valera CA, Fernandes LFS, Leal Pacheco FA. The Configuration of Forest Cover in Ribeirão Preto: A Diagnosis of Brazil's Forest Code Implementation. Sustainability. 2020;12(14):5686. https://doi.org/10.3390/su12145686.
17. de Oliveira ML, dos Santos CAC, de Oliveira G, Perez-Marin AM, Santos CAG. Effects of human-induced land degradation on water and carbon fluxes in two different Brazilian dryland soil covers. Science of The Total Environment. 2021;792:148458. https://doi.org/10.1016/J.SCITOTENV.2021.148458.
18. Cabral OMR, Freitas HC, Cuadra SV, de Andrade CA, Ramos NP, Grutzmacher P, Galdos M, Packer APC, da Rocha HR, Rossi P. (2020). The sustainability of a sugarcane plantation in Brazil assessed by the eddy covariance fluxes of greenhouse gases. Agric For Meteorol, 282–3 107864. https://doi.org/10.1016/j.agrformet.2019.107864.
19. Oliveira RR, Pezzi LP, Souza RB, Santini MF, Cunha LC, Pacheco FS. First measurements of the ocean-atmosphere CO2 fluxes at the Cabo Frio upwelling system region, Southwestern Atlantic Ocean. Cont Shelf Res. 2019;181:135–42. https://doi.org/10.1016/J.CSR.2019.05.008.
20. Tedesco D, de Oliveira MF, dos Santos AF, Costa Silva EH, de Souza Rolim G, da Silva RP. Use of remote sensing to characterize the phenological development and to predict sweet potato yield in two growing seasons. Eur J Agron. 2021;129:126337. https://doi.org/10.1016/J.EJA.2021.126337.
21. Aparecido LE, de O, Lorençone, Lorençone JA, Rolim PA, Meneses GdeS, de Moraes KC, J. R. da S. C. de, & Torsoni GB. Can nonlinear agrometeorological models estimate coffee foliation? J Sci Food Agric. 2021. https://doi.org/10.1002/JSFA.11387.
22. Helman D, Lensky IM, Osem Y, Rohatyn S, Rotenberg E, Yakir D. A biophysical approach using water deficit factor for daily estimations of evapotranspiration and CO2 uptake in Mediterranean environments. Biogeosciences. 2017;14(17):3909–26. https://doi.org/10.5194/BG-14-3909-2017.
23. Guo M, Xu J, Wang X, He H, Li J, Wu L. (2015). Estimating CO2 concentration during the growing season from MODIS and GOSAT in East Asia. 36(17), 4363–4383. https://doi.org/10.1080/01431161.2015.1081305.
24. Li X, Hu X-M, Cai C, Jia Q, Zhang Y, Liu J, et al. Terrestrial CO2 fluxes, concentrations, sources and budget in Northeast China: Observational and modeling studies. Journal of Geophysical Research: Atmospheres. 2020;125:e2019JD031686. https://doi.org/10.1029/2019JD031686.
25. Chhabra A, Gohel A. Dynamics of atmospheric carbon dioxide over different land cover types in India. Environ Monit Assess. 2019;191:799. https://doi.org/10.1007/s10661-019-7681-z.
26. Walker AP, et al. Predicting long-term carbon sequestration in response to CO2 enrichment: How and why do current ecosystem models differ? Global Biogeochem Cycles. 2015;29:476–95. https://doi.org/10.1002/2014GB004995.

27. Golkar F, et al. Using OCO-2 satellite data for investigating the variability of atmospheric CO2 concentration in relationship with precipitation, relative humidity, and vegetation over Oman. Water. 2020;12(1):101. https://doi.org/10.3390/w12010101.
28. Barbosa HA, Kumar L, T. V., & Silva. LRM. Recent trends in vegetation dynamics in the South America and their relationship to rainfall. Nat Hazards. 2015;77:883–99. https://doi.org/10.1007/s11069-015-1635-8.
29. Wagle P, Gowda PH, Billesbach DP, Northup BK, Torn MS, Neel JPS, Biraud SC. Dynamics of CO2 and H2O fluxes in Johnson grass in the U.S. Southern Great Plains. Science of The Total Environment Volume. 2020;739(140077):ISSN 0048–9697. https://doi.org/10.1016/j.scitotenv.2020.140077.
30. Liu Q, Fu YH, Zeng Z, Huang M, Li X, Piao S. Temperature, precipitation, and insolation effects on autumn vegetation phenology in temperate China. Glob Change Biol. 2016;22:644–55. https://doi.org/10.1111/gcb.13081.
31. Sá MMF, Schaefer CEGR, Loureiro DC, Simas FNB, Alves BJR, de Sá Mendonça E, de Figueiredo EB, La Scala N, Panosso AR. Fluxes of CO2, CH4, and N2O in tundra-covered and Nothofagus forest soils in the Argentinian Patagonia. Science of The Total Environment. 2019;659:401–9. https://doi.org/10.1016/J.SCITOTENV.2018.12.328.
32. Vicentini ME, Pinotti CR, Hirai WY, de Moraes MLT, Montanari R, Filho MCMT, Milori DMBP, Júnior NLS, Panosso AR. CO2 emission and its relation to soil temperature, moisture, and O2 absorption in the reforested areas of Cerrado biome, Central Brazil. Plant Soil. 2019;2019 444(1):1. 193–211. https://doi.org/10.1007/S11104-019-04262-Z. 444).
33. Graham MH (2003), CONFRONTING MULTICOLLINEARITY IN ECOLOGICAL MULTIPLE REGRESSION. Ecology. 84: 2809-2815. https://doi.org/10.1890/02-3114.
34. Tamura R, Kobayashi K, Takano Y, ·, Miyashiro R, Nakata K, ·, & Matsui T. Mixed integer quadratic optimization formulations for eliminating multicollinearity based on variance inflation factor. J Global Optim. 2019;73:431–46. https://doi.org/10.1007/s10898-018-0713-3.
35. Rafiei Sardooi E, Azareh A, Choubin B, Barkhori S, Singh VP, Shamshirband S. Applying the remotely sensed data to identify homogeneous regions of watersheds using a pixel-based classification approach. Appl Geogr. 2019;111:102071. https://doi.org/10.1016/J.APGEOG.2019.102071.
36. Siabi Z, Falahatkar S, Alavi SJ. (2019). Spatial distribution of XCO2 using OCO-2 data in growing seasons. Journal of Environmental Management, Volume 244, 2019, Pages 110–118, ISSN 0301–4797, https://doi.org/10.1016/j.jenvman.2019.05.049.
37. Falahatkar S, Mousavi SM, Farajzadeh M. Spatial and temporal distribution of carbon dioxide gas using GOSAT data over IRAN. Environ Monit Assess. 2017;189:627. https://doi.org/10.1007/s10661-017-6285-8.
38. Mohammed GH, Colombo R, Middleton EM, Rascher U, van der Tol C, Nedbal L, Goulasf Y, Pérez-Priegog O, Dammh A, Meronij M, Joinerc J, Cogliatib S, Verhoefe W, Malenovskýk Z, Gastellu-Etchegorryl J-P, Millerm JR, Guantern L, Morenoo J, Moyaf I, Berryp JA, Frankenbergq C, Zarco-Tejada

PJ. Remote sensing of solar-induced chlorophyll fluorescence (SIF) in vegetation: 50 years of progress. Remote Sens Environ. 2019;231:111177. https://doi.org/10.1016/j.rse.2019.04.030.

39. Duveiller G, Filipponi F, Walther S, Köhler P, Frankenberg C, Guanter L, Cescatti A. A spatially downscaled sun-induced fluorescence global product for enhanced monitoring of vegetation productivity. Earth System Science Data. 2020;12(2):1101–16. https://doi.org/10.5194/essd-12-1101-2020.
40. Campbell PKE, Huemmrich KF, Middleton EM, Ward LA, Julitta T, Daughtry CST, Burkart A, Russ AL, Kustas WP. Diurnal and Seasonal Variations in Chlorophyll Fluorescence Associated with Photosynthesis at Leaf and Canopy Scales. Remote Sens. 2019;11:488. https://doi.org/10.3390/rs11050488.
41. Sun Y, Frankenberg C, Jung M, Joiner J, Guanter L, Köhler P, Magney T. Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP. Remote Sens Environ. 2018;209:808–23. https://doi.org/10.1016/j.rse.2018.02.016.
42. Rolim GS, Aparecido LEO. Camargo, Köppen and Thornthwaite climate classification systems in defining climatical regions of the state of São Paulo, Brazil. The International Journal of Climatology. 2016;36:636–43. https://doi.org/10.1002/joc.4372.
43. Taiz L, Zeiger E. Fisiologia Vegetal. 4th ed.: Editora Artmed; 2009.
44. Frankenberg C, O'Dell C, Guanter L, McDuffie J. Remote sensing of near-infrared chlorophyll fluorescence from space in scattering atmospheres: Implications for its retrieval and interferences with atmospheric CO 2 retrievals. Atmospheric Measurement Techniques. 2012;5(8):2081–94. https://doi.org/10.5194/amt-5-2081-2012.
45. Ishizawa M, Mabuchi K, Shirai T, Inoue M, Morino I, Uchino O, Yoshida Y, Belikov D, Maksyutov S. Inter-annual variability of summertime CO2 exchange in Northern Eurasia inferred from GOSAT XCO2. Environmental Research Letters. 2016;11(10):105001. https://doi.org/10.1088/1748-9326/11/10/105001.
46. Hansen R, Mander Ü, Soosaar K, et al. Greenhouse gas fluxes in an open air humidity manipulation experiment. Landscape Ecol. 2013;28:637–49. https://doi.org/10.1007/s10980-012-9775-7.
47. Snyder RL, Spano D. Phenology and Evapotranspiration. In: Schwartz M, editor. Phenology: An Integrative Environmental Science. Dordrecht: Springer; 2013. https://doi.org/10.1007/978-94-007-6925-0_28.
48. Aparecido LEDO, Ferreira RB, Rolim GDS, Souza BS, De, Souza PS, De. (2017). Nonlinear agrometeorological models for estimating lychee fruit growth. Revista Brasileira de Fruticultura, 39(2). https://doi.org/10.1590/0100-29452017169.
49. Kovalskyy V, Henebry GM, Roy DP, Adusei B, Hansen M, Senay G, Mocko DM. Evaluation of a coupled event-driven phenology and evapotranspiration model for croplands in the United States northern Great Plains. Journal of Geophysical Research Atmospheres. 2013;118(11):5065–81. https://doi.org/10.1002/JGRD.50387.

50. CONAB, Companhia Nacional de Abastecimento. (2019). Calendario de plantio e colheita de grãos no Brasil 2019. https://www.conab.gov.br.
51. Shekhar A, Chen J, Paetzold JC, Dietrich F, Zhao X, Bhattacharjee S, Ruisinger V, Wofsy SC. Anthropogenic CO2 emissions assessment of Nile Delta using XCO2 and SIF data from OCO-2 satellite. Environmental Research Letters. 2020;15(9):095010. https://doi.org/10.1088/1748-9326/AB9CFE.
52. Guo M, Wang X, Li J, Yi K, Zhong G, Tani H. Assessment of Global Carbon Dioxide Concentration Using MODIS and GOSAT Data. Sensors. 2012;12(12):16368–89. https://doi.org/10.3390/s121216368.
53. Taylor, T. E., Eldering, A., Merrelli, A., Kiel, M., Somkuti, P., Cheng, C., Rosenberg,R., Fisher, B., Crisp, D., Basilio, R., Bennett, M., Cervantes, D., Chang, A., Dang,L., Frankenberg, C., Haemmerle, V. R., Keller, G. R., Kurosu, T., Laughner, J. L.,… Yu, S. (2020). OCO-3 early mission operations and initial (vEarly) XCO2 and SIF retrievals. Remote Sensing of Environment, 251, 112032. https://doi.org/10.1016/J.RSE.2020.112032.
54. Rolim GS, de Oliveira Aparecido LE, de Souza PS, et al. Climate and natural quality of Coffea arabica L. drink. Theor Appl Climatol. 2020;141:87–98. https://doi.org/10.1007/s00704-020-03117-3.
55. de O Aparecido, de Souza Rolim LE, da Silva Cabral G, de Moraes JR. Validation of ECMWF climatic data, 1979–2017, and implications for modelling water balance for tropical climates. Int J Climatol. 2020;40:6646–65. https://doi.org/10.1002/joc.6604.
56. White JW, Hoogenboom G, Stackhouse PW, Hoell JM. Evaluation of NASA satellite- and assimilation model-derived long-term daily temperature data over the continental US. Agric For Meteorol. 2008;148(10):1574–84. https://doi.org/10.1016/J.AGRFORMET.2008.05.017.
57. Zhang Y, Joiner J, Hamed Alemohammad S, Zhou S, Gentine P. A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks. Biogeosciences. 2018;15(19):5779–800. https://doi.org/10.5194/bg-15-5779-2018.
58. Yu L, Wen J, Chang CY, Frankenberg C, Sun Y. High-Resolution Global Contiguous SIF of OCO-2. Geophys Res Lett. 2019;46(3):1449–58. https://doi.org/10.1029/2018GL081109.
59. SEADE – Fundação estadual de Análise de dados – Perfil dos municípios paulista. 2020. https://perfil.seade.gov.br/.
60. Camargo FP, et al (2020). Previsões e Estimativas das Safras Agrícolas do Estado de São Paulo, Ano Agrícola 2019/20. Análises e Indicadores do Agronegócio, São Paulo, v. 15, n. 9, set. http://www.iea.sp.gov.br/out/TerTexto.php?codTexto=14839.
61. Camargo AP. Classificação climática para zoneamento de aptidão agroclimática. In: Heldwein AB, Schneider FM, Buriol GA, Petter Medeiros SL, Estefanel V, editors. Congresso Brasileiro de Agrometeorologia. Vol. 7. Brazil: Viçosa Sociedade Brasileira de Agrometeorologia/Universidade Federal de Viçosa; 1991. pp. 126–31.
62. AppEEARS Team. (2020). Application for Extracting and Exploring Analysis Ready Samples (AppEEARS). Ver. 6. NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC),

USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, USA. Accessed 11, 11, 2020. https://lpdaacsvc.cr.usgs.gov/appeears/.

63. Stackhouse PW Jr, Westberg D, Chandler WS, Zhang T, Hoell JM. Prediction of worldwide energy resource (POWER) agroclimatology methodology, version 1.1.0, May 30. NASA Langley Research Center; 2015.
64. Gujarati DN, Porter DC. Econometria básica-5. Amgh Editora; 2011.
65. R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.
66. Miles J. (2014). Tolerance and Variance Inflation Factor. Wiley StatsRef: Statistics Reference Online. https://doi.org/10.1002/9781118445112.STAT06593.
67. Kalnins A. Multicollinearity: How common factors cause Type 1 errors in multivariate regression. Strateg Manag J. 2018;39(8):2362–85. https://doi.org/10.1002/SMJ.2783.

Figures

Figure 1

Heatmap of the Pearson's correlation matrix, Where: a) before the Variance Inflation Factor (VIF) selection and b) after the selection by Variance Inflation Factor (VIF).

Figure 2

Temporal variability between 2015 and 2019 where: a) Xco_2 (ppm), b Qg (MJ m^{-2} day^{-1}), c) RH (%) and d) SIF 757 (Wm^{-2} sr^{-1} μm^{-1}). Where Xco_2 = column average of carbon dioxide in the atmosphere; Qg=global radiation; RH=relative humidity; SIF 757 = Solar-Induced Chlorophyl Fluorescence at 757nm.

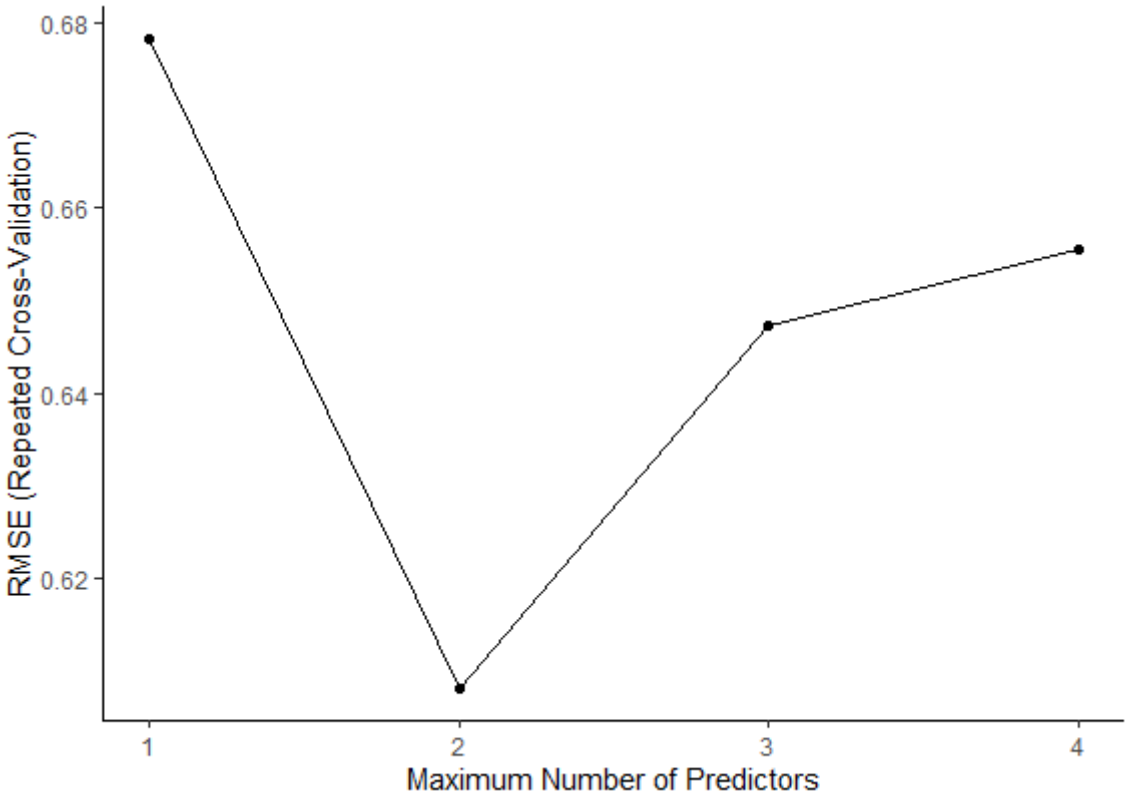

Figure 3

RMSE score for the training sample.

Figure 4

a) Cross-validation between Xco_2 estimated by stepwise and Xco_2 observed by OCO-2 and b) Daily downscale of natural Xco_2.

Figure 5

Flowchart of data acquisition, processing, and analysis.

Figure 6

Flowchart of the stepwise construction.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ST.xlsx