# Spatio-temporal statistical analysis of the carbon footprint of the terrestrial vegetation

Patrick Gneuss[1,*], Wolfgang Schmid[2] and Reimund Schwarze[3]

[1] *Department of Statistics, European University Viadrina, PO Box 1786, 15207 Frankfurt (Oder); vetter@europa-uni.de*
[2] *Department of Statistics, European University Viadrina, PO Box 1786, 15207 Frankfurt (Oder); schmid@europa-uni.de*
[3] *Department of Environmental Economics, European University Viadrina, PO Box 1786, 15207 Frankfurt (Oder); schwarze@europa-uni.de*
[*] *Corresponding author*

---

***Abstract.*** *The study of sources and sinks of carbon dioxide is of interest in many research disciplines and in political negotiations on climate change mitigations. The most important source/sink for the global carbon dioxide balance is the global vegetation, which acts as a sink during the photosynthesis and at the same time as a source of $CO_2$ as plants use the produced chemical energy for building up biomass and for cell respiration. In natural science, the Gross Primary Productivity (GPP) of the terrestrial vegetation, in effect the produced chemical energy from photosynthesis has been analyzed frequently. However the net effect of vegetation on $CO_2$ emissions (Net Primary Productivity (NPP)) on a global spatial scale and on an intra-annual time basis has not yet been well discovered. This study addresses this problem from a spatio-temporal statistical point of view. We make use of remotely sensed observations of the vertical profile of $CO_2$ concentrations obtained from the Greenhouse Gases Observing Satellite (GOSAT) and observations of the GPP derived from data of the MODIS satellite mission on the primary production of vegetation. A space-time linear mixed effects model was fitted to the data, which is able to capture spatial and temporal auto-correlation of ground $CO_2$ concentrations and as well the spatial and temporal cross-correlation between $CO_2$ and the GPP through latent spatial and temporal random processes. In that way we were able to obtain spatio-temporal predictions of the influence of vegetation on surface $CO_2$ concentrations and discover the source/sink activity of vegetation on a global spatial scale with $1^\circ \times 1^\circ$ resolution and in a nearly weekly time resolution.*

***Keywords.*** *Spatio-temporal statistics; Carbon dioxide concentrations; Net primary productivity; Remote sensing; Environmental processes*

---

# 1 Motivation

As a main driver of climate change, the study of sources and sinks of the greenhouse gas carbon dioxide is of interest in many research disciplines and in political negotiations on climate change mitigation. The most important source/sink for the global carbon dioxide balance is the global vegetation. It acts as a sink during the photosynthesis, through the intake of $CO_2$ and its conversion to biomass, and at the same time as a source of $CO_2$, as plants use the produced chemical energy for building up biomass and for cell respiration. In natural science, the Gross Primary Productivity (GPP) of the terrestrial vegetation, in effect the produced chemical energy from photosynthesis has been analyzed frequently. However the net effect of vegetation on $CO_2$ concentrations (Net Primary Productivity (NPP)) on a global spatial scale and on an intra-annual time basis has not yet been well discovered. This study addresses this problem from a spatio-temporal statistical point of view. We make use of remotely sensed observations of the vertical profile of CO2 concentrations obtained from the Greenhouse Gases Observing Satellite (GOSAT) and observations of the GPP derived from data of the MODIS satellite mission on the primary production of vegetation in a $1^\circ \times 1^\circ$ global spatial grid and in nearly weekly time intervals reaching from May 2009 to May 2013. In addition the global landcover dataset, which can be found in [1], was used in order to account for spatial and temporal non-stationarities in the covariance structure. In Figure 1 the GOSAT surface observations of the $CO_2$ concentration and in Figure 2 the GPP observations derived from the MODIS product (MOD17A2), which were collected during May 2010, are shown.

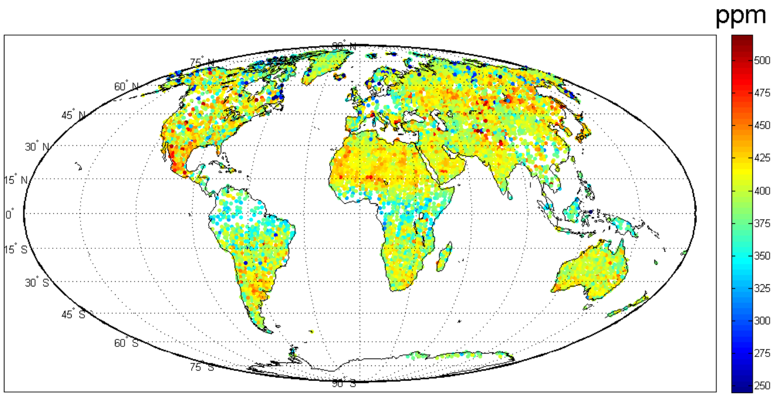


Figure 1: Surface level observations of the GOSAT $CO_2$ vertical profile in May 2010

In comparing these figures, it can already be seen, how the primary production of the terrestrial vegetation affects ground $CO_2$ concentrations. In the tropical regions around the equator, the high gross primary productivity leads to lower levels of the $CO_2$ concentration. However in the mid-latitude regions, characterized by the seasonal cycle of local vegetation, the influence of the vegetation on the surface $CO_2$ concentrations is not clear. Despite the high photosynthesis activity in May 2010 (Figure 2) the $CO_2$ concentration levels are moderate or even high (Figure 1). This is due to the counteracting effects of photosynthesis and plant respiration (GPP vs. NPP). Although photosynthesis activity is high in May, plants also use high amounts of chemical energy in order to grow and for the maintenance of their current biomass. Consequently the net effect on the $CO_2$ concentration remains unclear. In order to discover this net effect, a space-time linear mixed effects model as in [3] was fitted to the data, which is able to capture spatial and temporal auto-correlation of ground $CO_2$ concentrations and as well the spatial and temporal cross-correlation between $CO_2$ and the GPP through latent spatial and temporal random processes. In

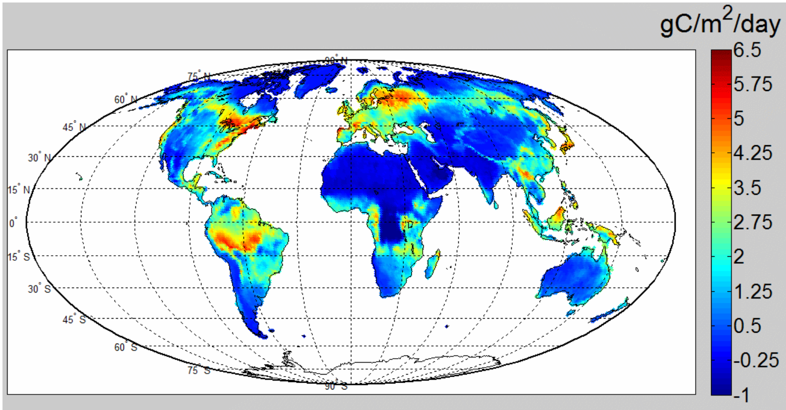


Figure 2: Gross primary productivity derived from the MOD17A2 product in May 2010

that way spatio-temporal predictions of the influence of vegetation on surface $CO_2$ concentrations are derived and the source/sink activity of vegetation on a global spatial scale with $1^\circ \times 1^\circ$ resolution and in a nearly weekly time resolution is discovered.

## 2 Spatio-temporal linear mixed effects model

In order to adequately capture the spatial and temporal dependence between the surface $CO_2$ concentrations and the gross primary productivity, a spatio-temporal linear mixed effects model was considered, as in Equation 1. Let $\mathrm{CO2}(\mathbf{s},t)$ denote the surface $CO_2$ concentration at location $\mathbf{s}$ and time point $t$, then

$$\begin{aligned} \mathrm{CO2}(\mathbf{s},t) &= \sum_{k=1}^{K} \beta_k \cdot LC_k \cdot \mathrm{GPP}(\mathbf{s},t) + \beta_{gas}\frac{P}{T} + \sum_{k=1}^{K} \alpha_k \cdot LC_k \cdot \mathrm{GPP}(\mathbf{s},t) \cdot w_{GPP,k}(\mathbf{s},t) \qquad (1) \\ &\quad + \sum_{k=1}^{K} LC_k \cdot \mathrm{GPP}(\mathbf{s},t) \cdot z_{GPP,k}(t) + z(t) + \gamma w(\mathbf{s},t) + \varepsilon(\mathbf{s},t) \quad , \end{aligned}$$

where $\mathrm{GPP}(\mathbf{s},t)$ is the gross primary productivity at location $\mathbf{s}$ and time $t$, $LC_k$ is a dummy variable indicating whether the observation is located in the land cover class $k$, $w_{GPP,k}(\mathbf{s},t)$ is a spatial random effect related to GPP in the land cover class $k$, $z_{GPP,k}(t)$ is a temporal random effect related to GPP in the land cover class $k$, $z(t)$ is a temporal random effect covering the residual temporal auto-correlation, $w(\mathbf{s},t)$ is a spatial random effect covering the residual spatial auto-correlation, $\alpha_k$,$\beta_k$ and $\gamma$ are unknown scale coefficients and $\varepsilon(\mathbf{s},t)$ denotes the residual process at location $s$ and time $t$. Since we are modeling a gas concentration, we have also controlled for the local surface temperature and surface air pressure. $w_{GPP,k}(\mathbf{s},t)$ and $w(\mathbf{s},t)$ are modeled as a gaussian spatial random field with an exponential spatial correlation function with range parameters $\theta_k$ and $\theta$. $z_{GPP,k}(t)$ and $z(t)$ are modeled as auto-regressive processes of order one. With $\mathbf{z}_t$ being the vector of temporal random effects at time $t$, the following state-space representation is assumed

$$\mathbf{z}_t = G \cdot \mathbf{z}_{t-1} + \eta_t, \qquad (2)$$

where $G$ is the corresponding transition matrix and $\eta_t$ denotes the matrix of temporal innovations, which are assumed to be diagonal matrices. For estimating the parameters, a Maximum-Likelihood approach is applied using an EM-algorithm, as in [2].

## 3 Results

Using kriging predictions based on the estimated parameters and the smoothed random effects, an estimate of the the influence of the terrestrial vegetation on the surface $CO_2$ can be obtained by considering only the predicted/smoothed random and fixed effects related to GPP. The resulting prediction for the global $1^\circ \times 1^\circ$ prediction grid for May 2010 is shown in Figure 3.[1] As can be seen, the predictions

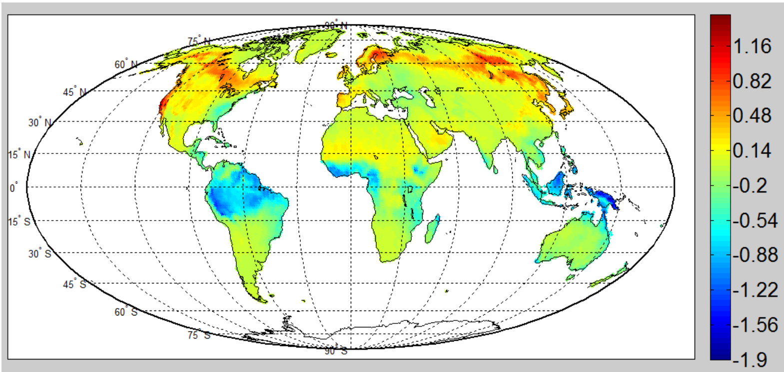


Figure 3: Predicted influence of the terrestrial vegetation on the surface $CO_2$ concentrations for May 2010

show the expected pattern of the net effect of the vegetation on the surface $CO_2$ concentrations. In the equatorial regions the high photosynthesis activity leads to a drop in local $CO_2$ concentration levels and the contrary holds true for the mid-latitude region in the northern hemisphere. Although local photosynthesis activity is very high in May, the net influence of vegetation is positive here, leading to higher $CO_2$ concentrations. Through fixed and random effects we were able to describe the spatio-temporal cross-correlation between the GPP and the $CO_2$ concentrations. In that way, we arrived at an estimate of the net effect of terrestrial vegetation on surface $CO_2$ concentrations, which to our knowledge has not been achieved at a global spatial and on a nearly weekly time scale up to now.

### References

[1] Global Land Cover 2000 database. (2003).*European Commission, Joint Research Centre*.

[2] Fassò, Alessandro and Finazzi, Francesco (2011). Maximum likelihood estimation of the dynamic coregionalization model with heterotopic data. *Environmetrics* **22**(6), 735–748.

[3] Fassò, Alessandro and Finazzi, Francesco (2013). A varying coefficients space-time model for ground and satellite air quality data over Europe. *Statistica & Applicazioni - Special Issue 2013*, 45–56.

[1]Throughout the analysis standardized and logarithmized data were used and accordingly the values for the predictions in Figure 3 are unit-free.