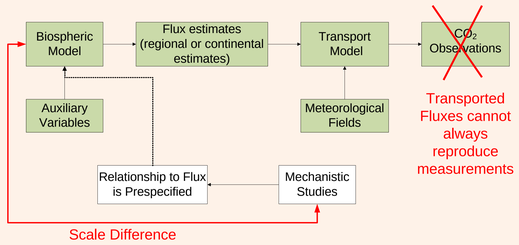
A Geostatistical Synthesis Study of Factors Affecting Net Ecosystem Exchange in Different Ecosystems of North America

Vineet Yadav [1](vineety@umich.edu), Anna M. Michalak[1,2] and Kim Mueller[1]
[1]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109, USA.
[2]Department of Atmospheric, Oceanic, and Space Science, University of Michigan, Ann Arbor, MI 48109, USA

Introduction

► Process-based models and satellite-derived products are frequently used to estimate fluxes of CO_2 for specific areas. Such estimates, however, cannot always reproduce measured CO_2 concentrations when coupled with an atmospheric transport model.



► Although there may be many reasons for these discrepancies, they point to a need to study the relationship between critical ecological processes and CO_2 flux at the spatio-temporal resolution of the biospheric models and satellite data.

► The presented analysis uses flux data collected at eight AmeriFlux sites from 2001 to 2005 shown in Figure 2, together with MODIS satellite data, to explore the relationship between Net Ecosystem Exchange (NEE) and parameters associated with atmospheric carbon cycling at different temporal scales.

A geostatistical regression analysis combined with Bayesian model selection is performed at various temporal scales (monthly, 8day, and daily), which quantifies the relationship between NEE and auxiliary data, while accounting for any temporal correlation in the flux residuals.

Statistical Methodology

► Geostatistical regression can be used to estimate linear drift coefficients (β's) associated with auxiliary variables within the model of the trend. The observation, therefore, can be expressed as the sum of the individual contribution of each variable ($X_i\beta_i$) and the flux residual **Q(x)** that is modeled as a stationary isotropic Gaussian process with exponential covariance function.

$$\mathbf{Z} = \beta_0 + \beta_1\mathbf{X}_1 + \beta_2\mathbf{X}_2 ... + \beta_{p-1}\mathbf{X}_{p-1} + \mathbf{Q(x)}$$

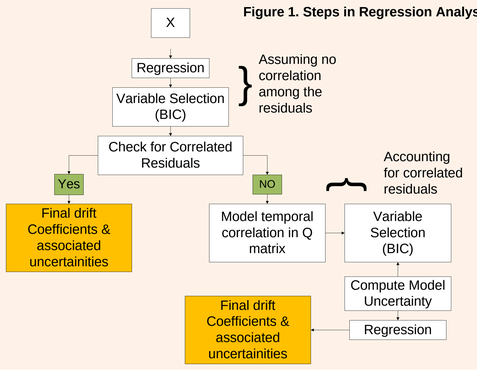
► In this analysis, the observation vector (**Z**) are the NEE measurements.

► The **X** vectors are known variables (e.g. temperature, precipitation, etc.) selected using the Bayes Information Criteria (BIC) approach. This approach is based on maximum likelihood estimation, and quantifies the degree to which the data support each candidate set of auxiliary variables, taking into account both the goodness of fit, i.e. the sum of squared residuals and the number of parameters in the model.

► Deriving inferences on the basis of a single model assuming it to be a true model ignores model uncertainty. This can lead to underestimation of uncertainty of regressors of interest and the model itself. A Bayesian solution to this problem was proposed by Leamer [1978] and implemented in a BIC framework by Raftery [1995].In this study we implement a modification of this method to determine posterior probabilities of the regressors to be in the true model and the probability of the model selected from BIC.

► In this work, the covariance matrix **Q** defines the temporal correlation between the portion of NEE not explained by **Xβ** (i.e. **Q(x)** or NEE residuals). If there is no temporal correlation within the residuals, Q becomes a diagonal matrix with the variance of the residuals as its diagonal entries.

Figure 1. Steps in Regression Analysis

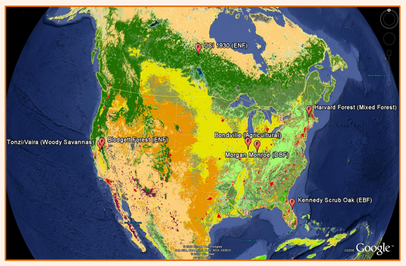


Site Locations

Auxiliary Variables:

1. LAI
2. FPAR
3. EVI
4. NDVI
5. Global Radiation (GRAD)
6. Vapor Pressure Deficit(VPD)
7. Air Temperature
8. Soil Temperature
9. Precipitation (Precip)
10. Soil Water Content (SWC)
11. Sensible Heat Flux (SHF)
12. Latent Heat Flux (LHF)

Figure 2: Spatial location of the AmeriFlux study sites (Base Map is Land Cover map of North America)



Relationship of NEE to auxiliary variables is only analyzed for years 2001-2005 due to data constraints of the study.

Table 1: Eco-climatic characteristics of the chosen AmeriFlux study sites

Site Name	IGBP Class	Köppen climate class	Canopy Height	Stand Age
UCI1930	ENF	Dfc	16	
Blodgett Forest	ENF	Cs	8	11
Bondville	Crop	Dfa	3	NA
Florida (Scrub oak)	CSH	Cfa	23	
Harvard Forest	DBF	Dfb	1	81
Morgan Monroe	DBF	Cfa	27	70
Tonzi Ranch	WSA	Csa	23	
Vaira Ranch	GRA	Csa	0.55 ± 0.12	NA

Results

Figure 3: Variance explained by the models selected from the BIC at three different temporal scales (Note: values on the bars show the probability of the models selected from BIC)

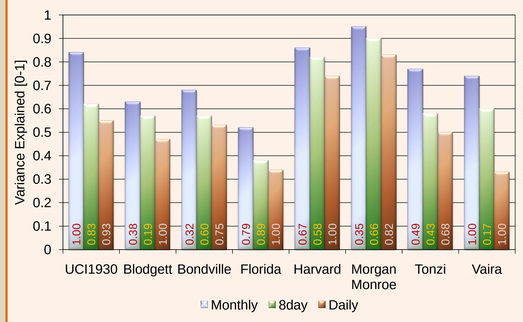


Figure 5: Contribution of individual regressors (for the model with maximum probability) at the daily scale for the study sites (Note: values above bars represent total source or sink at daily scale)

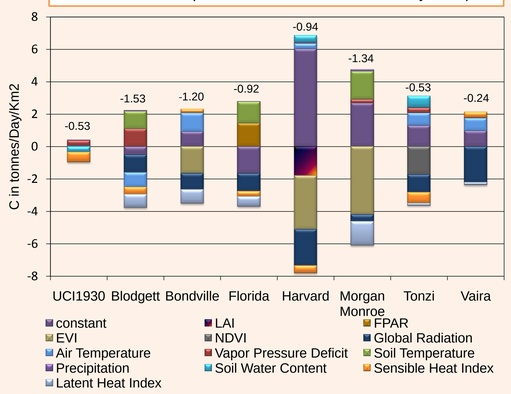
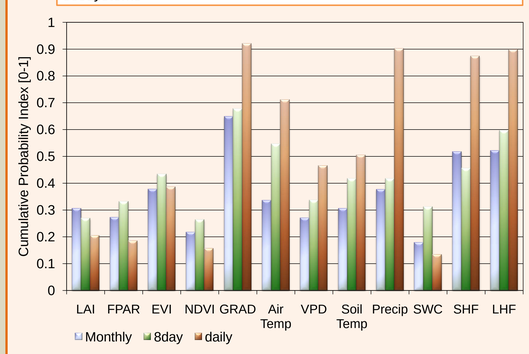


Table 2: Percent probabilities of auxiliary variables to be in the true model over full model space (i.e. 2^k-1; where k are total number of auxiliary variables): an example from Vaira flux site

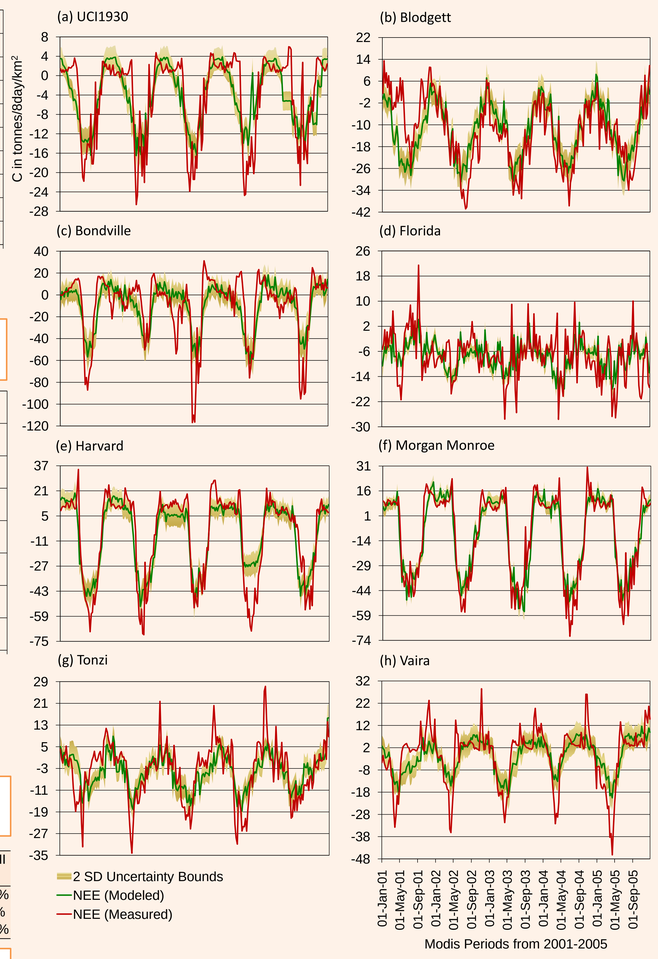
Flux Sites	LAI	FPAR	EVI	NDVI	GRAD	Air Temp	VPD	Soil Temp	Precip	SWC	SHI	LHI
Vaira (M)	14%	15%	20%	16%	98%	24%	16%	32%	24%	28%	98%	20%
Vaira (E)	22%	46%	10%	68%	100%	83%	23%	23%	14%	69%	94%	9%
Vaira (D)	19%	28%	4%	7%	100%	100%	4%	4%	100%	5%	94%	93%

Figure 6: Combined cumulative index of probabilities for auxiliary variables (as shown in Table 2) across 8 flux sites to be in the true model indicating their overall importance in governing flux across ecosystems



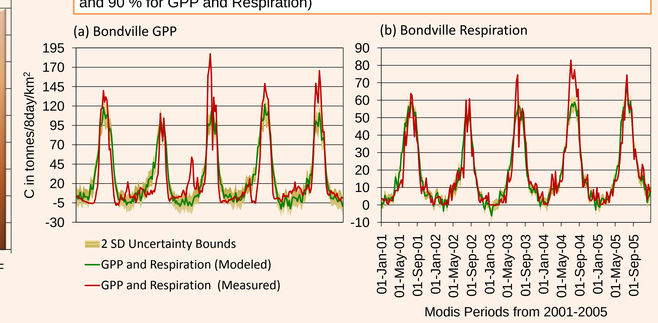
In figures and tables (+) values indicate source and (-) values indicate sink of carbon

Figure 4: 8day measured and modeled NEE with uncertainty bounds for the study sites



Division of NEE into flux components i.e. GPP and Respiration leads to better inference and higher explained variance .

Figure 7: 8day measured GPP and Respiration with confidence bounds for Bondville flux site (Variance explained rises from ~ 60 % in case of NEE to ~ 78 % and 90 % for GPP and Respiration)



Key Findings

► Across flux sites the variance explained declined from monthly to daily temporal scales (Figure 3). Furthermore, comparatively, fewer variables were required to explain flux at coarser temporal resolutions.

► The number of models and the probabilities of the BIC selected model increased from monthly to daily temporal scales except for the site UCI1930 (Figure 3).

► For all flux sites the variables chosen at monthly and 8day scale by the statistical algorithm were from the mix chosen at the daily scale, though probability based evidence-strength relationships (Table 2) were different for different temporal resolutions.

► At most sites Precip, SWC, Air Temp and Soil Temp were correlated with C efflux (*Davidson et al.,* 2006; *Lee et al.*, 2002) whereas GRAD, LHF and SHF were associated with C sink (Figure 5).

► Across sites GRAD followed by LHF had the highest cumulative probability (summing across sites) at all three temporal resolutions (Figure 6). At daily scale other than GRAD and LHF significant governors of NEE identified through BIC based probabilities were SHF and Precip. Similarly at 8day and monthly scale these variables were GRAD, LHF, SHF, Air Temp and SWC.

► The probabilities of the variables to be in the BIC selected model increased from monthly to daily scale (Table 2). However explained variance at the daily scale was substantially lower. This we believe can be either due to omission of certain variables explaining flux, more noise or enhanced non-linearity of NEE cycles at this scale.

► The algorithm proposed here was more effective in explaining variations in NEE at the study sites that have a definite seasonal growing period (Harvard, Morgan Monroe and UCI-1930) in contrast to site like Florida (Scrub Oak) where there is an annual GPP – Respiration cycle.

Conclusion

► In this research a novel methodological implementation of the Bayesian model selection in a geostatistical framework is proposed and utilized, to identify primary regressors and their contribution in explaining NEE in different ecosystems. Through the use of the proposed methodology an enhanced understanding of the environmental stochasticity of these regressors in governing NEE is achieved for the eight AmeriFlux sites representing different Ecosystems

► Results indicate that the primary governors and their stochasticity in explaining NEE varies at different temporal resolutions. Search for scale based controls of NEE in this research, is not a search for different physiological processes but for dominant controls & stimuli that can explain the observed pattern at a particular temporal interval. Once these patterns and their governors are understood etiological exploration on the basis of adjustments in physiology to environmental stimuli can begin and that can lead to lessening of uncertainty surrounding global C cycle.

Future Work:

Future work will focus on investigating the state space of non-linearity in NEE time series of different ecosystems and also examining the non linear relationships between NEE and auxiliary variables.

Findings of preliminary research:

Preliminary investigation using dynamical systems methods/tests (Brock, W. A., Hsieh D.A. and LeBaron B. 1991) on NEE time series show absence of nonlinearity at monthly and 8day temporal scales for most flux sites. Though there is strong evidence of nonlinearity at daily scales in NEE time series of all AmeriFlux study sites. This primarily happens due to increasing aperiodicity of NEE time series at smaller (daily to sub-daily) temporal scales

Acknowledgements and References

This research is funded by the NASA SMD ROSES North American Carbon Program, under contract number NNX06AE84G. The authors also want to gratefully acknowledge the principal investigators of eight AmeriFlux study sites for their data that made this research viable.

Brock, W. A., Hsieh D.A., LeBaron B. (1991). *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence,* MIT Press, Cambridge, Massachusetts

Davidson, E. A., Janssens, I. A., and Luo, Y. (2006). On the variability of respiration in terrestrial ecosystems: moving beyond, *Global. Change Biology.*, 12(2), 154-164.

Leamer, E. E. (1978). *Specification Searches: Ad Hoc Inference With Nonexperimental Data*, John Wiley, New York.

Lee, M.-s., Nakane, K., Nakatsubo, T., Mo, W.-h. and Koizumi, H., (2002). Effects of rainfall events on soil CO 2 flux in a cool temperate deciduous broad-leaved forest. *Ecological Research.*, 17(3), 401-409.

Raftery, A. E. (1995). Bayesian model selection in social research, *Sociological Methodology.,* 25, 111-163.