



## Ecosystem Service-Based Sensitivity Analyses of i-Tree Eco

By Jian Lin, Charles N. Kroll, and David J. Nowak

**Abstract.** Trees are known to provide various ecosystem services and disservices to urban communities, which can be quantified using models based on field and environmental data. It is often uncertain how tree structure and environmental variables impact model output. Here we perform a sensitivity analysis (SA) of i-Tree Eco, a common urban forest model, to analyze the relative impact of different model inputs on three module outputs: biogenic volatile organic compound (BVOC)(isoprene and monoterpenes) emissions, carbon storage and sequestration, and dry deposition of nitrogen dioxide, sulfur dioxide, and ozone. The SA methods included novel applications of the Morris one-at-a-time method and a variance-based decomposition method, which integrates Monte Carlo simulation with Latin hypercube sampling and Iman Conover analysis. A case study was performed in New York City, New York, USA, with field plot data collected in 2013. Genus has the largest influence on BVOC emissions by determining base emission rates and its high interactions with other input factors, and BVOC emissions are sensitive to leaf biomass in a concave manner and temperature in a convex manner, while isoprene emissions show a strong linear relationship with photosynthetically active radiation (PAR). Diameter at breast height plays the most important role for both carbon storage and sequestration estimators; crown light exposure and tree condition are also important for carbon sequestration. Dry deposition velocity is sensitive to leaf area index and relative humidity in a nearly linear way, while sensitive to temperature and PAR in a concave manner. The results provide guidance to facilitate future field plot campaigns and model development. The knowledge revealed by the SA is also beneficial for model uncertainty reduction, which in turn facilitates more effective urban forest management and decision-making.

**Keywords.** Air Pollutant; Carbon Storage and Sequestration; Monte Carlo; Urban Forests; Volatile Organic Compounds.

### INTRODUCTION

As a demographic trend and land transformation process, urbanization creates many environmental problems (e.g., increased runoff and nutrient export, increased temperatures, and increased material consumption and energy use)(Duh et al. 2008; Poumanyong and Kaneko 2010). One way to alleviate these impacts is through urban greening, and many cities have launched large urban tree planting initiatives (McPherson et al. 2011). Models of urban tree impacts on the environment can help develop more efficient and effective planting schemes (Morani et al. 2011), identify areas where existing forests should be maintained (Lin 2020), improve the overall management of urban forests, and better quantify the benefits of these forest resources (Nowak et al. 2008a). One popular model is i-Tree Eco (hereafter referred to as “Eco”)(<https://www.itreetools.org>), which can

quantify the structure, function, and ecosystem benefits trees provide (Nowak and Crane 2000), and has been used by thousands of researchers, urban planners, and others around the world to advocate for the benefits of urban trees (Lin et al. 2019).

While this tool has been extremely beneficial to the planning and management of urban trees, this model makes assumptions that simplify the relationships between the structure and function of urban forests and the representation of urban landscapes (Nowak and Crane 2000). While such assumptions are often necessary to model these complex systems, they can increase the uncertainty of model output and hinder the efficient and effective management of urban forests. Several studies point out that the uncertainty regarding various aspects of urban forest models (e.g., mortality rates of trees, transpiration rates, and meteorological conditions) should be better

addressed (Yang et al. 2005; Morani et al. 2011; Selmi et al. 2016). Current uncertainty estimation within Eco is limited to the standard error of the total number of trees within the study area (Nowak et al. 2008b) and its impact on selected model outputs (e.g., carbon storage)(Nowak et al. 2013). In addition, those standard errors usually come from sampling error instead of error of estimation (e.g., allometric equations) and therefore tend to underestimate the overall uncertainty (Nowak et al. 2013). This study focuses on developing an advanced sensitivity analysis (SA) of Eco, including assessing how changes in model outputs can be apportioned to different model inputs, differentiating the relative importance of different model inputs, and identifying specific input-output relationships.

SA can be used in different settings (e.g., variance cutting, factors prioritization and fixing)(Saltelli et al. 2004), and serves a variety of purposes (e.g., model development, scenario analyses, and comparative studies)(Saltelli et al. 2004; Hirabayashi et al. 2011; Steffens et al. 2012). A wide range of SA techniques have been developed, ranging from classical frequentist analyses to complex Bayesian inference, and from local (e.g., Morris one-at-a-time) to global (e.g., variance-based) methods (Marino et al. 2008; Saltelli et al. 2008). Different SA methods can also be integrated together to achieve a more complete understanding of the relationship between input and output variables (Van Griensven et al. 2006).

This study employs a Morris one-at-a-time method (MOAT) and a variance-based decomposition method (VD) which integrates Monte Carlo simulation with Latin hypercube sampling (LHS-MC) and the Iman Conover method to address the correlation structure of the input variables (Iman and Conover 1982). The novel applications of SA are illustrated with a case study in New York City (NYC) for 2013. The specific goals of this analysis are to: (1) determine the ranking and relative importance of input variables on Eco outputs; (2) improve the understanding of the input-output variable relationships in Eco (e.g., linear and additive effects, nonlinear and interaction effects, and threshold effects); and (3) explore the implications of the results for future research and urban forest management (e.g., areas where additional data collection and analyses may be needed).

## METHODS

### Model Description

Eco consists of modules which can quantify urban forest structure, function, and value using field plots, air pollution, and meteorological data as input variables (Nowak and Crane 2000; Nowak et al. 2008a) (Figure 1). The input module, Eco-A, characterizes urban forest “Anatomy,” or the structure and composition based on data from field plots. Other modules examined here include one ecosystem disservice and two ecosystem services of urban trees. The Eco-B module estimates biogenic volatile organic compounds (BVOCs) from trees based on tree species, leaf biomass, air temperature, and other environmental factors. Trees emit hundreds of species of BVOCs, but the two major BVOCs are isoprene and monoterpenes (Bonan 2015). The Eco-C module estimates total carbon storage and annual carbon sequestration based on allometric equations, tree growth, mortality, and decomposition rates. The Eco-D module estimates the hourly air pollution removal by urban forests based on dry deposition processes (Nowak et al. 2006). The interactions between input variables and these modules are outlined in Figure 1, and the outputs from Eco that will be examined in this study are listed in Table 1.

Eco-B estimates BVOC emissions ( $E$ ) as:

$$E = B_E \times B \times \gamma \quad (1)$$

where  $B_E$  is the base genus emission rate, which is defined as the emission level standardized to  $T = 30^\circ\text{C}$  and  $\text{PAR} = 1000 \mu\text{mol}/\text{m}^2\text{s}^{-1}$  (Nowak et al. 2008a),  $B$  is species leaf dry weight biomass, and  $\gamma$  is the temperature and light correction factor. The i-Tree species database contains data for more than 6,500 tree and shrub species and their corresponding attributes; among these attributes,  $B_E$  is compiled from the literature (Nowak et al. 2002). Users can also upload local site-specific  $B_E$  values to the i-Tree database (this needs to be validated by the i-Tree team). For isoprene,  $\gamma$  is estimated as:

$$\gamma = \frac{\exp\left(\frac{c_{T1} \times (T - T_S)}{R \times T_S \times T}\right)}{0.961 + \exp\left(\frac{c_{T2} \times (T - T_M)}{R \times T_S \times T}\right)} \times \frac{\alpha \times c_{L1} \times \text{PAR}}{\sqrt{1 + \alpha^2 \times \text{PAR}^2}} \quad (2)$$

while for monoterpenes it is:

$$\gamma = \exp(\beta \times (T - T_S)) \quad (3)$$

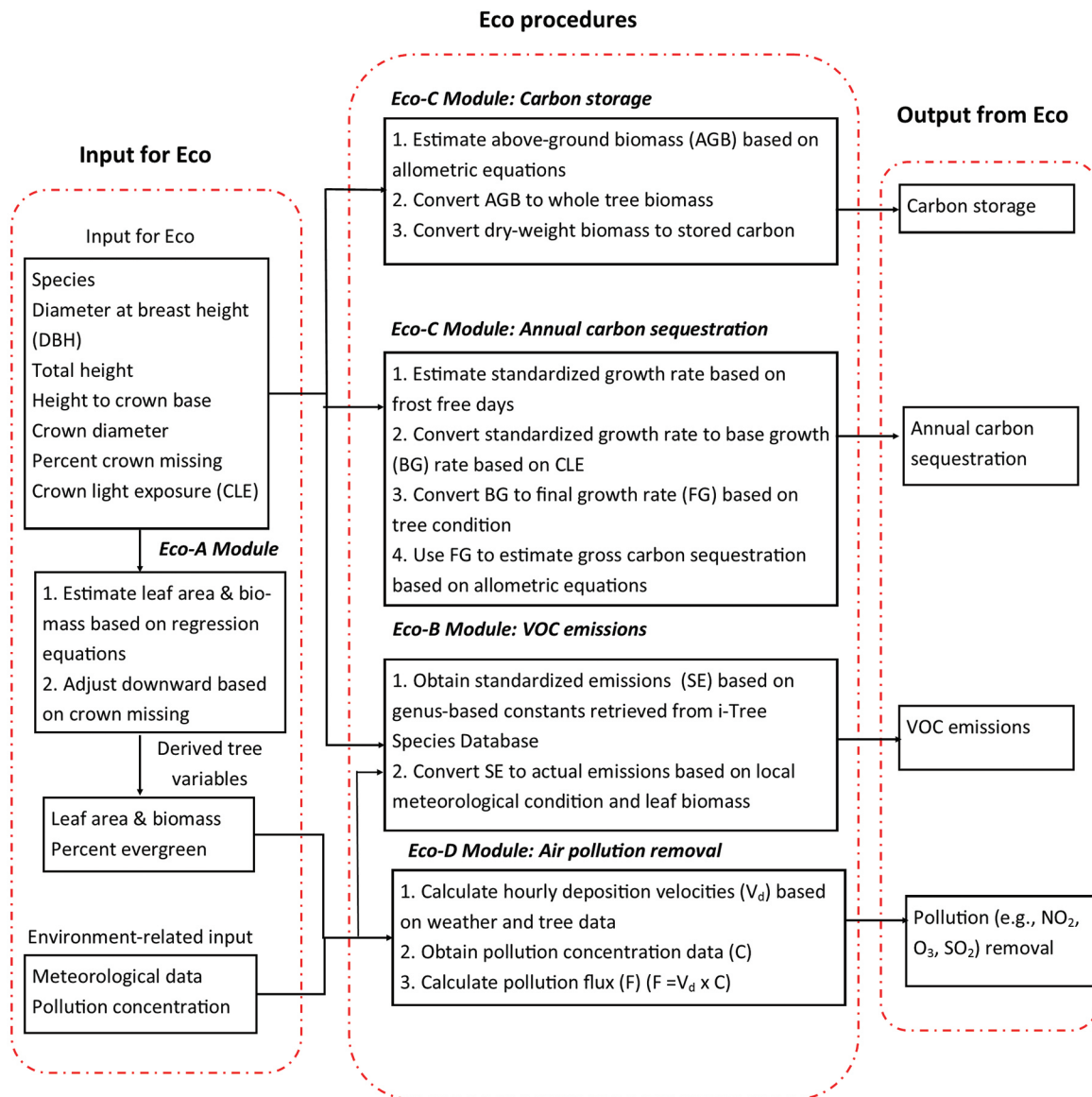


Figure 1. Input variables, procedures in each module, and output variables from Eco.

Table 1. A list of output variables used in this experiment.

Module	Output variable	Unit	Description
Eco-B	Isoprene	µgC/hr	Isoprene emission
Eco-B	Monoterpenes	µgC/hr	Monoterpenes emission
Eco-C	Carbon storage	kg	Carbon storage
Eco-C	Carbon gross sequestration	kg/yr	Annual gross carbon sequestration
Eco-D	$V_d$ of NO <sub>2</sub>	cm/s	Dry deposition velocity of nitrogen dioxide (NO <sub>2</sub> )
Eco-D	$V_d$ of SO <sub>2</sub>	cm/s	Dry deposition velocity of sulfur dioxide (SO <sub>2</sub> )
Eco-D	$V_d$ of O <sub>3</sub>	cm/s	Dry deposition velocity of ozone (O <sub>3</sub> )

where  $c_{T1} = 95000 \text{ J} \times \text{mol}^{-1}$ ,  $c_{T2} = 230000 \text{ J} \times \text{mol}^{-1}$ ,  $T_M = 314 \text{ K}$ ,  $\alpha = 0.0027 \text{ } \mu\text{mol}^{-1}/\text{m}^2\text{s}$ ,  $c_{L1} = 1.066$  (dimensionless),  $\beta = 0.09 \text{ K}^{-1}$  (empirical coefficient),  $R$  is the ideal gas constant ( $8.314 \text{ J} \times \text{K}^{-1} \times \text{mol}^{-1}$ ),  $T_S$  is a standard temperature ( $303 \text{ }^\circ\text{K}$ ), and  $T$  is leaf temperature ( $^\circ\text{K}$ ), which is assumed to be equal to air temperature (Guenther et al. 1995; Guenther 1997).

Eco-C estimates forest biomass (Bio) using allometric equations from the literature (Nowak et al. 2013). The two most commonly used equations have the forms:

$$\text{Bio} = \exp(A + B \times \text{LN}(X)) \text{ and} \quad (4)$$

$$\text{Bio} = A \times (X^B)$$

where  $A$  and  $B$  are coefficients whose values vary based on species information. Species also determine the selection of equation forms.  $X$  is the predictor variable. Two forms of  $X$  employed in Eco-C are:

$$X = \text{DBH} \text{ and } X = (\text{DBH}^2) \times \text{HEIGHT} \quad (5)$$

For Eco-D, detailed equations used to calculate aerodynamic resistance, boundary layer resistance, canopy resistance, and dry deposition velocity can be found in Hirabayashi et al. (2011).

## Study Area and Data Employed

A sensitivity analysis case study of three Eco modules was performed for New York City (NYC). Tree

species, diameter at breast height (DBH), tree height, crown height and width, tree condition, crown light exposure (CLE), percent crown missing, and land use type associated with 1075 trees were obtained from 296 randomly sampled field plots measured in 2013, which is more than the 200 plots used by many i-Tree studies (Nowak et al. 2008b). Although some rare species are likely omitted by random sampling, random sampling produces statistically accurate estimators of urban forest structure (e.g., number of trees and tree sizes) with known bounds of error. Tree condition is estimated as the percent of the crown that is composed of dead branches (to nearest 5%), which has a total of 7 categories ranging from dead to excellent; CLE is the number of sides (4 cardinal directions and 1 top side) of the tree receiving sunlight from above (ranging from 0 to 5), which is employed to estimate light environment and consequently growth rates (Nowak et al. 2008a). Eight land use types were identified across the field plots; land use affects carbon storage and gross carbon sequestration, which is addressed by assigning land use biomass adjustment factors (Nowak et al. 2008a). Leaf area index (LAI) statistics for this study area were taken from Breuer et al. (2003)(Table 2), which summarize LAI for 26 temperate regions. Model meteorological inputs (e.g., temperature, relative humidity [RH], wind speed, air pressure, and photosynthetically

**Table 2. Summary statistics of input variables (each integer represents one category for categorical variables).**

Module	Parameter	Type	Mean	Standard deviation	Minimum	Maximum	Distribution	P-value for Kolmogorov-Smirnov Test
<b>Eco-B</b>	Genus	Categorical					Uniform	
	Leaf biomass (kg)	Continuous	8.18	18.5	0	267	Lognormal	0.19
	PAR (W/m <sup>2</sup> )	Continuous	84.3	122	0	483	Gamma	0.08
	Temperature (°C)	Continuous	13.3	10.1	-10	37.2	Gumbel	0.09
<b>Eco-C</b>	Species	Categorical					Uniform	
	DBH (cm)	Continuous	16.3	18.3	2.5	122	Lognormal	0.12
	Height (m)	Continuous	7	4.98	1.2	30.5	Lognormal	0.1
	Land use	Categorical			1	8	Uniform	
	Tree condition	Categorical			0	6	Uniform	
	CLE*	Categorical			0	5	Uniform	
<b>Eco-D</b>	LAI	Continuous	5.8	1.7	2	10	Uniform	
	Pressure (Mbar)	Continuous	1010	7.79	978	1030	Lognormal	0.05
	PAR (W/m <sup>2</sup> )	Continuous	84.3	122	0	483	Gamma	0.08
	Relative humidity (%)	Continuous	61	19	8	100	Beta	0.04
	Temperature (°C)	Continuous	13.3	10.1	-10	37.2	Gumbel	0.09
	Wind speed (m/s)	Continuous	5.27	2.83	0	20.6	Weibull	0.09

\* CLE affects the calculation of leaf biomass and therefore is also employed as an input parameter for Eco-B.



active radiation [PAR]) were obtained or derived from data from the nearest airport (NCDC 2018). PAR, which is the visible part (400 to 700 nm) of the solar spectrum, is calculated as 46% of total solar irradiance (Hirabayashi et al. 2011). The summary statistics for each input parameter are presented in Table 2.

## Sensitivity Analyses

### Morris One-at-a-Time Analysis

Two SA methods were employed in this study: MOAT and a variance-based decomposition (VD) method. MOAT and VD, along with regression analyses to support the SA, were conducted using the R statistical software package. MOAT is a local sensitivity method which is computationally inexpensive and can differentiate input variables as negligible, linear/nonlinear, or having interaction effects (Saltelli et al. 2004; 2008). The method starts by random sampling  $k$  parameters,  $X_1, \dots, X_i, \dots, X_k$ , from predefined levels of all input variables and calculating the subsequent model output  $Y(X_1, \dots, X_i, \dots, X_k)$ . For different Eco modules, the number of model parameters for ( $k$ ) varies ( $k = 3$  for isoprene, 2 for monoterpenes, 3 for carbon storage, 5 for carbon sequestration, and 6 for dry deposition). In the second run, only one parameter,  $X_i$ , is increased by step size  $\Delta$  with all other parameters remaining at their starting values, and model output is again calculated  $Y(X_1, \dots, X_i + \Delta, \dots, X_k)$ . From this, the elementary effect (EE) (Van Griensven et al. 2006; Saltelli et al. 2008) of the  $i^{\text{th}}$  input variable is estimated as:

$$EE_i = \frac{\frac{Y(X_1, \dots, X_i + \Delta, \dots, X_k) - Y(X_1, \dots, X_i, \dots, X_k)}{Y(X_1, \dots, X_i, \dots, X_k)}}{\frac{\Delta}{X_i}} \quad (6)$$

$i$  is an elasticity, i.e., the percent change of the output divided by the percent change of the input. For each random sample of parameters, every parameter is subsequently increased with all other parameters back at their starting values, and the EE for that parameter is calculated. The levels of the parameters are usually identified by the midpoint of 4, 6 or 8 divisions of the parameter range, with each division of equal probability (Saltelli et al. 2004). In this study, 8 divisions were used (so  $\Delta = 1/8 \times$  parameter range), and the above procedure was repeated 20 times, which leads to a total of  $20(1 + k)$  runs. The mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the EE for each

parameter across all runs is then calculated and plotted.  $\mu$  measures the overall influence individual parameters have on the output, while  $\sigma$  indicates whether the interaction between input and output are linear across the parameter space (low  $\sigma$ ) or if nonlinear or interaction effects are present (high  $\sigma$ ) (Saltelli et al. 2008). When positive and negative values of  $EE_i$  cancel each other out, a low  $\mu$  may be obtained for parameters which have a large impact on the output. Campolongo et al. (2007) proposed to instead calculate the mean of the absolute value of the  $EE_i$  ( $\mu^*$ ) to avoid this problem; this approach is implemented in this study.

When applying MOAT to Eco-B and C, three things should be noted. First, we dropped the cases when the original model output  $Y(X_1, \dots, X_i, \dots, X_k) = 0$ , because it will lead to an error (division by zero) when calculating  $EE_i$ . For example, when a tree is dead, the annual carbon sequestration is equal to zero; this also occurs when the BVOC emissions are estimated as zero for some species. Second, for the categorical input variables “land use,” “tree condition,” and “CLE,” which have limited categories and can be effectively represented using levels (each category is one level),  $\frac{\Delta}{X_i}$  was always set as 1 when changing individual input parameters from one level to the next in the  $EE_i$  calculation. Third, the unordered categorical input variables “species” was not examined because its levels cannot be effectively determined and chosen from hundreds of species in the study site; species were randomly chosen and fixed for each simulation.

### Variance-Based Decomposition (VD) Analysis

A flowchart to perform VD by integrating Monte Carlo simulation with Latin hypercube sampling (LHS) and Iman Conover is presented in Figure 2. Stage 1 is to determine appropriate probability distribution functions (PDFs) to describe input variables, and then to perform LHS on the PDFs. Stage 2 is a Monte Carlo simulation where model outputs are estimated using input variables obtained from the LHS in Stage 1. Stage 3 is a variance decomposition analysis, which decomposes the variance of the output into different fractions that can be attributed to different inputs, as well as a quantile regression analysis to explore the general input-output relationships. Each of these stages is briefly described below.

Stage 1 has two components. First, probability distributions were fit to input variables. Specific probability distributions were identified based on either empirical data sets from the study site or literature

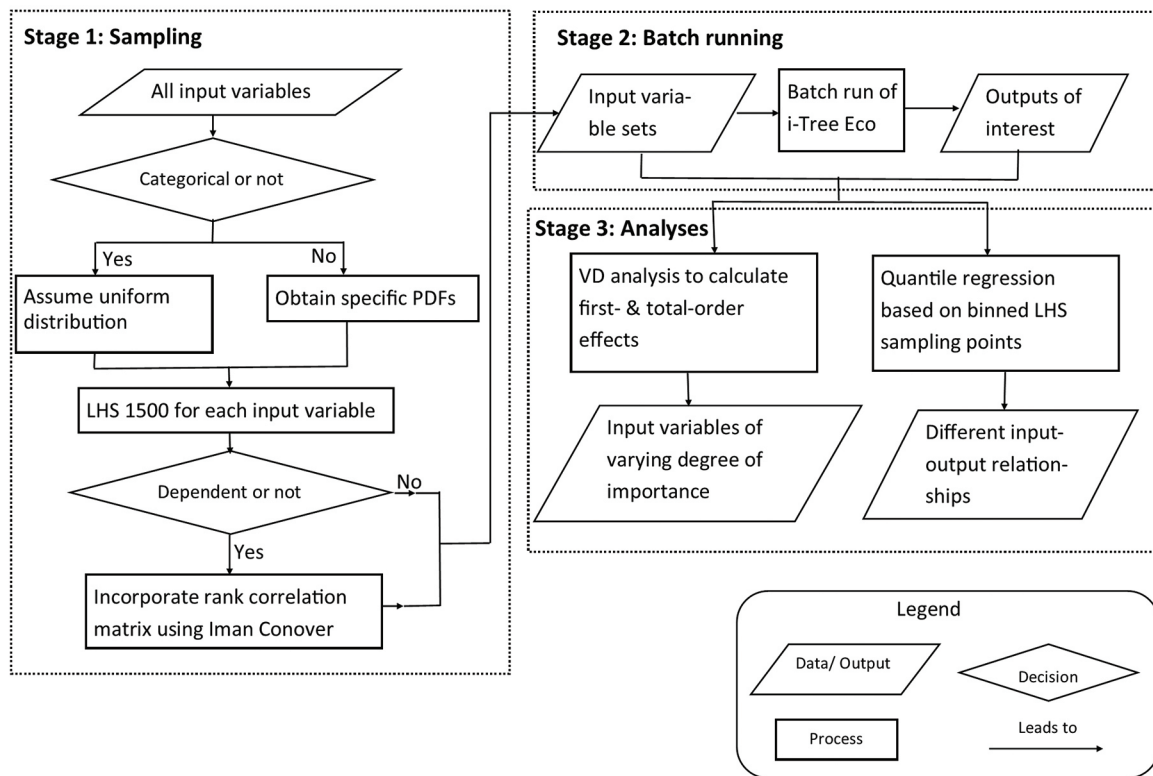


Figure 2. Flowchart to perform sensitivity analyses.

recommendations (Table 2). The use of the Weibull distribution for wind speed and the beta distribution for relative humidity were suggested from the literature (Yao 1974; Celik et al. 2010), and the PDFs of other identified input variables were empirically identified and estimated based on observational data sets from the study site. The assessment of PDF fit was performed by employing a Kolmogorov-Smirnov goodness-of-fit test and quantile-quantile plots (Wasserman 2013). For LAI, a uniform distribution was assumed with the minimum and maximum values taken from Breuer et al. (2003). Ordered categorical input variables were assumed to follow a uniform distribution with an equal probability of being within each category. Table 2 contains the distribution used for each input variable and the  $P$ -values for the Kolmogorov-Smirnov tests, which were all less than 0.2 and generally less than 0.1.

Second, LHS, which combines the advantages of random sampling and stratified sampling (Helton and Davis 2003), was employed to sample probability distributions from Stage 1. In LHS, the parameter

space of each input variable is divided into  $N$  ( $N = 1500$  in this study) intervals of equal marginal probability,  $1/N$ , and one sample of each variable is made randomly within each interval. Thus,  $N$  non-overlapping values for each input parameter are generated (Hirabayashi et al. 2011). The LHS method typically assumes that the sampling is performed independently for each parameter (Marino et al. 2008), though many of the input variables are correlated. To avoid this assumption, the Iman Conover method is used to enforce a dependence structure on input variables (Iman and Conover 1982). The Iman Conover method is based on rearranging the values of the parameters so that a desired correlation structure, which is derived from Spearman's rank correlation coefficients of empirical data sets (Table 3), is imposed on the  $k$  parameters. This technique has advantages of being distribution free and can be used with any type of sampling scheme (Helton and Davis 2003). At Stage 2, Eco-B, C, and D were batch run using input variables from Stage 1 to generate model outputs.

**Table 3. Spearman's rank correlation coefficient structure among input variables.**

	Height	Pressure	PAR	Relative humidity	Temperature
DBH	0.66				
Pressure		1.00			
PAR		0.14	1.00		
Relative humidity		-0.24	-0.28	1.00	
Temperature		0.04	0.58	-0.10	1.00

Stage 3 has two components. First, VD was employed to characterize and quantify the relative importance of input variables, and then for those input variables which were identified as important, quantile regression models were fit to binned data to capture specific input-output relationships.

VD is a global SA which can decompose the variance of model output into fractions attributed to each model input and to input interactions (Saltelli et al. 2004; 2008). The effects of VD are commonly measured by two sensitivity indices: a first order index and a total effect index. Model output,  $Y$ , is a function of a vector of  $k$  model inputs,  $X_1, \dots, X_k$ . The variance decomposition of  $Y$  can be expressed as:

$$\text{Var}(Y) = \sum_{i=1}^k V_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^k V_{ij} + \dots + V_{1,\dots,k} \quad (7)$$

where  $V_i = \text{Var}[\sum(Y|X_i)]$  is the contribution of  $X_i$  to the variance of  $Y$ , and  $V_{ij} = \text{Var}[\sum(Y|X_i, X_j)] - V_i - V_j$  is a part of the total variance caused by the interaction between  $X_i$  and all other  $X_j$ , namely the joint impact of  $X_i$  and all  $X_j$  on the variance of the output minus their first-order effects (Saltelli et al. 2008). The first-order index ( $S_i$ ) of  $X_i$  on  $Y$  is:

$$S_i = \frac{V_i}{\text{var}(Y)} \quad (8)$$

and represents the percentage of the total variance accounted by the first-order effect of  $X_i$ . The total effect index ( $S_{Ti}$ ) is:

$$S_{Ti} = 1 - \frac{V_{\sim i}}{\text{var}(Y)} \quad (9)$$

where  $V_{\sim i}$  is the total contribution of all parameters except for  $X_i$ .  $S_{Ti}$  accounts for the total contribution of  $X_i$  to the variance of model output (e.g., its first-order effect plus all higher-order effects due to interactions) (Saltelli et al. 2008). By definition,  $S_i \leq S_{Ti}$ ,  $\sum S_i \leq 1$ , and  $\sum S_{Ti} \geq 1$ .

Following Saltelli et al. (2010), two base matrices ( $A$  and  $B$ ) with dimension  $(N, k)$  were generated by LHS, where  $N$  is the number of simulation ( $N = 1500$ )

and  $k$  is the number of input variables.  $S_i$  and  $S_{Ti}$  are calculated as:

$$f_o = \frac{1}{N} \sum_{j=1}^N f(A)_j \quad (10)$$

$$S_i = \frac{1}{N} \sum_{j=1}^N \frac{f(B)_j (f(A_B^{(i)})_j - f(A)_j)}{\frac{1}{N} \sum_{j=1}^N (f(A)_j)^2 - f_o^2} \quad (11)$$

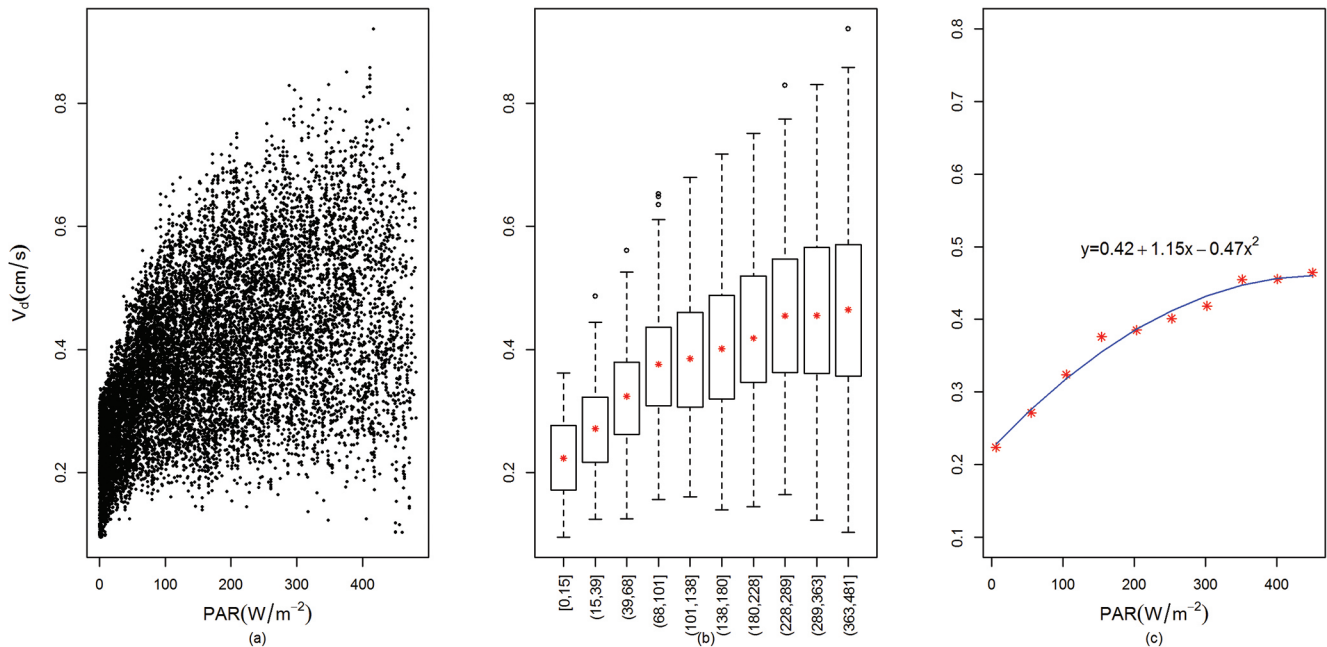
$$S_{Ti} = \frac{1}{2N} \sum_{j=1}^N \frac{(f(A)_j - f(A_B^{(i)})_j)^2}{\frac{1}{N} \sum_{j=1}^N (f(A)_j)^2 - f_o^2} \quad (12)$$

where  $A_B^{(i)}$  represents all columns from  $A$ , except the  $i^{\text{th}}$  column which is from  $B$ ,  $f()$  is a function to estimate a corresponding ecosystem service, and  $f(A)_j$  (or  $f(B)_j$ ) is the estimate an ecosystem service using the  $j^{\text{th}}$  row of base matrix  $A$  (or  $B$ ) as input variables.  $f_o$  represents the expected value of the ecosystem service across all parameter simulations represented in  $A$ , while the denominator of Equations 11 and 12 represents the total variance of the ecosystem service across all parameters in  $A$ . The main disadvantage of VD is the computational cost (here  $N(k + 2)$  simulations).

General input-output variable relationships were explored and identified using a binned quantile regression approach (Jucker et al. 2017). We divided the scatterplot points into ten bins with an equal number of points by calculating corresponding quantile points (10%, 20%, ..., 90%) and using these quantile points to separate scatterplot points; we then estimated the median values of output and input for each bin, and performed a regression exploring both linear and nonlinear relationships (Figure 3). Three possible forms of quantile regression for the medians (e.g., linear [ $c = 0$ ], convex [ $c > 0$ ], concave [ $c < 0$ ]) were employed to differentiate different input-output relationships using:

$$Y = a + bX + cX^2 \quad (13)$$

Note that the binned quantile regression approach provides us with a “median” response which is less impacted by outliers.



**Figure 3.** Bin and quantile regression process: (a) scatterplot of PAR vs  $V_d$  of  $\text{NO}_2$ ; (b) scatterplot divided into 10 bins, and each bin represented as a boxplot with median values shown as a red star; and (c) regression curve fit to the median values to express a specific input-output relationship.

The best form for each input-output relationship was selected based on the Akaike information criterion (Wasserman 2013), with the constraint that all the parameters should also be statistically significant. To compare the degree of concavity or convexity among different input-output relationships, regression models were built by standardizing the response and explanatory variables by subtracting the mean from each value and dividing by the standard deviation. A larger  $|c|$  value indicates greater concavity or convexity. This approach does more than simply explore the input-output relationships from the equations within Eco, but allows us to discover general relationships between correlated variables and model output. An example of the binned regression for the deposition velocity of  $\text{NO}_2$  as a function of PAR is shown in Figure 3.

## RESULTS

### MOAT Analysis

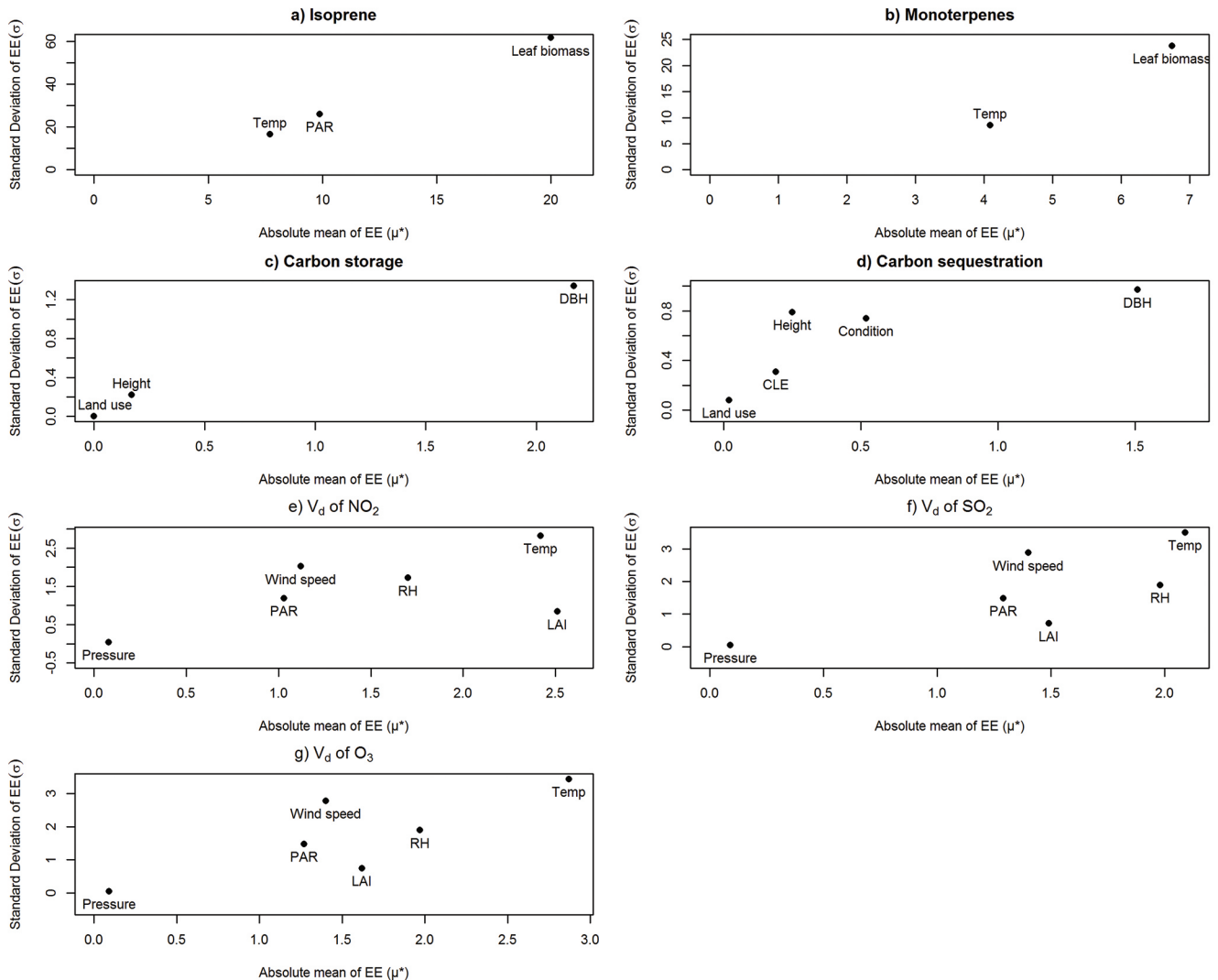
The MOAT analysis produced a mean absolute value ( $\mu^*$ ) and standard deviation ( $\sigma$ ) of the output elementary effect for each input variable across the 20 simulations. For isoprene (Figure 4a), leaf biomass had the

greatest  $\mu^*$ , indicating its large impact on model output; temperature and PAR had around the same magnitude of  $\mu^*$ , indicating their moderate effects on model output. For monoterpenes (Figure 4b), leaf biomass ( $\mu^* = 6.74$ ) was slightly more important than temperature ( $\mu^* = 4.09$ ), and both had high values of  $\sigma$ .

Among the three input variables examined for carbon storage (Figure 4c), the effect of DBH on the outputs (high  $\mu^*$  and  $\sigma$ ) was much larger than height and land use. The high  $\sigma$  for DBH indicated nonlinear effects of this input variable on carbon storage. For carbon sequestration (Figure 4d), two additional input variables related to tree health (condition) and site condition (CLE) were also examined. The five input variables can be grouped into three categories according to  $\mu^*$  and  $\sigma$ : most important (DBH), moderate importance (condition, height, CLE), and negligible importance (land use).

Eco-D calculates dry deposition of air pollution to trees based on the dry deposition velocity and the air pollutant concentration and assumes that the pollutants do not damage plant functions. For dry deposition velocity ( $V_d$ ) of  $\text{NO}_2$  (Figure 4e), temperature and LAI were the most important, as indicated by a high  $\mu^*$ .





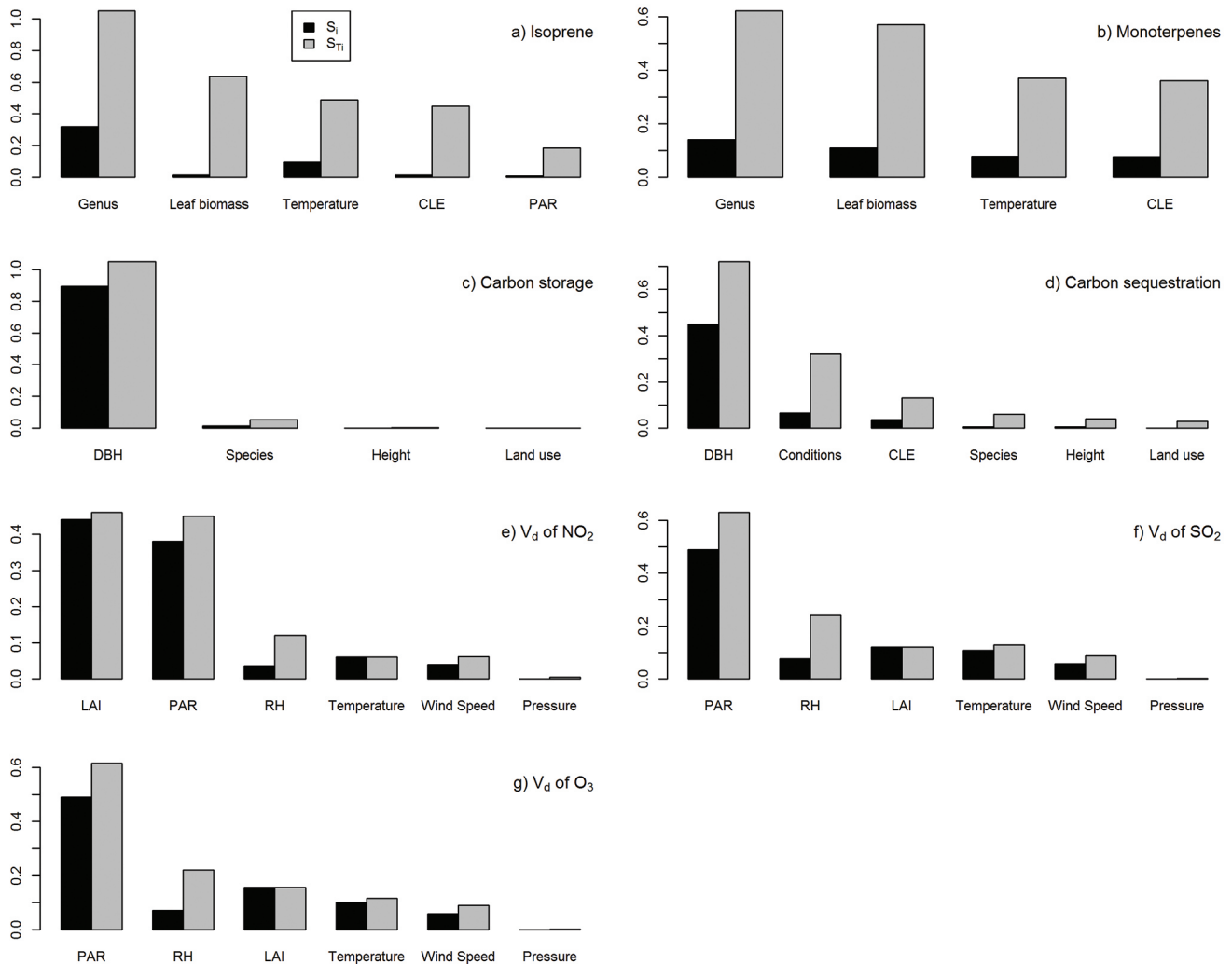
**Figure 4.** Output from MOAT analysis showing standard deviation of elementary effect ( $\sigma$ ) versus the mean of the absolute value of the elementary effect ( $\mu^*$ ) for Eco-B (a) isoprene and (b) monoterpenes, Eco-C (c) carbon storage and (d) sequestration, and Eco-D dry deposition velocity ( $V_d$ ) for (e)  $\text{NO}_2$ , (f)  $\text{SO}_2$ , and (g)  $\text{O}_3$ .

While they had similar  $\mu^*$ , temperature had a higher  $\sigma$  than LAI, indicating that its relationship to  $V_d$  of  $\text{NO}_2$  was much less linear than for LAI. RH, wind speed, and PAR all had smaller  $\mu^*$  than temperature and LAI, indicating their more moderate and similar effects on  $V_d$  of  $\text{NO}_2$ ; these variables all had a higher  $\sigma$  than LAI, indicating a less linear relationship to  $V_d$  of  $\text{NO}_2$ . Pressure had a negligible effect, as indicated by low  $\mu^*$  and  $\sigma$ . For  $V_d$  of  $\text{SO}_2$  (Figure 4f) and  $\text{O}_3$  (Figure 4g), a similar pattern of ranking based on  $\mu^*$  was displayed: temperature > RH > LAI > wind speed > PAR > pressure. Among these, temperature had the greatest  $\mu^*$  and  $\sigma$ , indicating its large impact on model output and

a highly nonlinear relationship with output; RH, LAI, wind speed, and PAR clustered in the middle of the ( $\mu^*$ ,  $\sigma$ ) plane, while pressure had negligible effect.

### Variance-Based Decomposition (VD) Analysis

While MOAT is computationally cheap, it cannot fully explore the input space, has difficulty with categorical input variables, and assumes input variables are uncorrelated. To overcome these disadvantages, VD was employed in the study. In addition to four direct input variables (genus, leaf biomass, temperature, and PAR), the sensitivity of one indirect input



**Figure 5. Sensitivity measures ( $S_i$  and  $S_{Ti}$ ) of (a) isoprene and (b) monoterpenes emissions, (c) carbon storage and (d) sequestration, and dry deposition velocity ( $V_d$ ) for (e)  $NO_2$ , (f)  $SO_2$ , and (g)  $O_3$ .**

variable (CLE) was also examined, because previous studies have demonstrated that altering CLE values could greatly affect BVOC emissions (Pace et al. 2018). For isoprene (Figure 5a), the order of importance measured by  $S_{Ti}$  was genus > leaf biomass > temperature > CLE > PAR. Genus played a dominant role, since base emission rates were assumed to be based on genera. The differences between  $S_i$  and  $S_{Ti}$  for all five input variables were large, indicating significant interaction effects between input variables. For monoterpenes (Figure 5b), the order of importance measured by  $S_{Ti}$  was genus > leaf biomass > temperature > CLE; large interaction effects also existed for all input variables as indicated by the large differences between  $S_i$  and  $S_{Ti}$ .

Among the four input variables that determine carbon storage (Figure 5c), DBH played the most dominant role, species had a smaller role, while height and land use had negligible effects. For carbon storage, the model interaction effects were small, as indicated by the small difference between  $S_i$  and  $S_{Ti}$ . For carbon sequestration (Figure 5d), DBH again had the largest effect, while condition and CLE had moderate effects; the other three input variables (species, height, and land use) all had negligible effects. The differences between  $S_i$  and  $S_{Ti}$  for DBH, condition, and CLE indicated the existence of important interaction effects between these variables and carbon sequestration.

For  $V_d$  of  $NO_2$  (Figure 5e), LAI and PAR had the most dominant effect, while RH, temperature, and

wind speed had smaller effects, and pressure had a negligible effect. The small differences between  $S_i$  and  $S_{Ti}$  for these variables indicated minimal interaction effects between these input variables and the output. For  $V_d$  of  $SO_2$  (Figure 5f) and  $O_3$  (Figure 5g), the pattern of rankings of importance were similar: PAR had the largest effect, followed by RH, LAI, temperature, and wind speed with moderate effects, and pressure with a negligible effect. For PAR and RH, interaction effects were detected by the differences in  $S_i$  and  $S_{Ti}$ .

### Bin Regression Analysis

While VD can quantify the effects of individual input variables and their interaction impacts on output variability, it does not identify general input-output relationships. For those input variables identified as important in VD (unordered categorical input variables were excluded), a bin quantile regression analysis was employed, and the general input-output relationships at the aggregated level were described using regression form, regression coefficients, and the adjusted  $R^2$  (Table 4). Note that by binning the data we do inflate the adjusted  $R^2$  of the models; our goal here is to generally describe the median model response.

Leaf biomass showed a concave relationship with both emissions of isoprene and monoterpenes (Eco-B outputs), as indicated by a significant negative parameter on  $X^2$  (a significant positive parameter would be convex). With an increase in leaf biomass, the emissions increased at a faster rate initially, then at a slower rate, and finally became relatively insensitive to the change in leaf biomass. The concavity of the monoterpenes vs. leaf biomass relationship was greater than that of the isoprene vs. leaf biomass relationship. Although leaf biomass of an individual genus at a specific temperature affected BVOC emissions in a linear manner (Equation 1), with different base rates from different genera and the nonlinear relationship between temperature and the temperature correction factors (Equations 2 and 3), the combined effect across all genera and temperature in the simulation produced a concave relationship. CLE showed a linear relationship with both isoprene and monoterpenes emissions, and their adjusted  $R^2$  values were lower than that of leaf biomass, indicating a weaker relationship with BVOC emissions when compared with the leaf biomass and BVOC relationship. By contrast, temperature showed a convex relationship with both BVOC emissions, indicating that the

**Table 4. Bin quantile regression analysis of i-Tree Eco-B, C, and D outputs for most important input variables.**

Module	Output variable $Y$	Input variable $X$	Regression equation	Regression form	Adjusted $R^2$
Eco-B	Isoprene	Leaf biomass	$Y = 0.16 + 1.32X - 0.18X^2$	Concave	0.99
Eco-B	Isoprene	Temperature	$Y = -0.45 + 0.82X + 0.49X^2$	Convex	0.96
Eco-B	Isoprene	PAR	$Y = 0.94X$	Linear	0.87
Eco-B	Isoprene	CLE	$Y = 0.95X$	Linear	0.87
Eco-B	Monoterpenes	Leaf biomass	$Y = 0.34 + 1.66X - 0.38X^2$	Concave	1.00
Eco-B	Monoterpenes	Temperature	$Y = -0.35 + 0.89X + 0.39X^2$	Convex	0.99
Eco-B	Monoterpenes	CLE	$Y = 0.89X$	Linear	0.74
Eco-C	Carbon storage	DBH	$Y = -0.18 + 0.65X + 0.2X^2$	Convex	0.99
Eco-C	Carbon gross sequestration	DBH	$Y = 0.99X$	Linear	0.98
Eco-C	Carbon gross sequestration	Condition	$Y = 0.91X$	Linear	0.78
Eco-C	Carbon gross sequestration	CLE	$Y = 0.82X$	Linear	0.59
Eco-D	$V_d, NO^2$	PAR	$Y = 0.42 + 1.148X - 0.467X^2$	Concave	0.94
Eco-D	$V_d, NO^2$	Relative humidity	$Y = 0.77X$	Linear	0.55
Eco-D	$V_d, NO^2$	LAI	$Y = 0.99X$	Linear	0.98
Eco-D	$V_d, NO^2$	Temperature	$Y = 0.2 + 0.97X - 0.22X^2$	Concave	0.96
Eco-D	$V_d, SO^2$	PAR	$Y = 0.35 + 1.15X - 0.39X^2$	Concave	0.99
Eco-D	$V_d, SO^2$	Relative humidity	$Y = 0.86X$	Linear	0.71
Eco-D	$V_d, SO^2$	Temperature	$Y = 0.15 + 0.98X - 0.16X^2$	Concave	0.98
Eco-D	$V_d, SO^2$	LAI	$Y = 0.975X$	Linear	0.95
Eco-D	$V_d, O^3$	PAR	$Y = 0.37 + 1.16X - 0.41X^2$	Concave	0.98
Eco-D	$V_d, O^3$	Relative humidity	$Y = 0.85X$	Linear	0.69
Eco-D	$V_d, O^3$	LAI	$Y = 0.9775X$	Linear	0.95
Eco-D	$V_d, O^3$	Temperature	$Y = 0.15 + 0.98X - 0.17X^2$	Concave	0.98

output responded relatively insensitively at low and moderate values of temperature, but then increased strongly as the temperature increased; the degrees of convexity, measured by  $|c|$  were 0.49 and 0.39 for isoprene and monoterpenes, respectively. This result makes sense, since emission increased nearly logarithmically at higher temperatures. Unlike leaf biomass and temperature, PAR showed a strong linear relationship with isoprene emissions. The relationships between binned median inputs and median Eco-B output variables were generally strong, with adjusted  $R^2$  values ranging from 0.87 to 1.00.

DBH showed a convex relationship with carbon storage, while it had a linear relationship with gross carbon sequestration; these relationships had a high adjusted  $R^2$ . By contrast, the linear form used to describe the relationships between tree condition and CLE to gross carbon sequestration had a relatively low adjusted  $R^2$ , indicating either a weaker input-output relationship, or the model we proposed to represent the binned quantile regression (Equation 8) was inadequate to capture this relationship.

PAR showed concave relationships with  $V_d$  of  $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{O}_3$ , with the degree of concavity of  $\text{NO}_2 > \text{O}_3 > \text{SO}_2$ ;  $V_d$  of all three gases also showed concave relationships with temperature with similar degrees of concavity. For all other input variables, linear relationships were detected. LAI showed the strongest linear relationship with  $V_d$  (high adjusted  $R^2$ ), while RH showed the weakest linear relationship (low adjusted  $R^2$ ). This may indicate  $V_d$  responds to RH and interacts with other input variables in more complex ways. The dry deposition processes of all three gases showed similar regression patterns with input variables.

## DISCUSSION

### The Emissions of Isoprene and Monoterpenes

The amount of BVOCs emitted was affected by tree physiology (e.g., genus, leaf biomass) and environmental factors (e.g., temperature). The VD analysis indicated that genus had the largest influence on the emissions of isoprene and monoterpenes, and there were strong interaction effects among input variables, as indicated by the large differences between  $S_i$  and  $S_{Ti}$  and the high  $\sigma$  in the MOAT analysis. In Eco-B, each genus has a specific  $B_E$  value, varying from 0

(e.g., *Pyrus*) to 70 (e.g., *Liquidambar*)  $\mu\text{g}/\text{g}^{-1}/\text{hr}^{-1}$  for isoprene emissions, and from 0 (e.g., *Pyrus*) to 7.9 (e.g., *Pistacia*)  $\mu\text{g}/\text{g}^{-1}/\text{hr}^{-1}$  for monoterpenes emissions. Misidentifications of tree species may or may not result in large changes in BVOC emissions; however, misidentifications of tree genera could produce large deviations in emissions, as demonstrated in one study in Munich, Germany, that showed isoprene emissions decreased and monoterpene emissions increased primarily due to misclassifications of tree species (Pace et al. 2018). Genus-level BVOC emission rates are used. If species-specific emissions are demonstrated to provide more accurate estimates, the model can be updated to use species-specific BVOC emission rates. Note that in the MOAT analysis,  $\mu^*$  for both Eco-B outputs were much larger in magnitude than for Eco-C and Eco-D outputs. This difference may be because a strong interaction existed among these input variables (high  $\sigma$ ). High VOC-emitting genera can increase the magnitude of EE.

The roles of other factors (e.g., leaf biomass, temperature, PAR, and CLE) were highly affected by the selection of a specific genus. As expected, the effects of CLE were smaller than leaf biomass, because CLE's role is mainly through its effect on leaf biomass. Eco-A calculates leaf biomass using different approaches when CLE changes between classes (e.g., 0~1, 2~3, and 4~5). The same environmental factors may have a smaller influence if low-emitting genera are selected, and a larger influence when high-emitting genera are selected. Leaf biomass affected BVOC emissions in a concave manner (Table 4), which contradicts the linear relationship expressed in Equation 1. This is probably because in the SA, we resampled values of genus, leaf biomass, PAR, and temperature across the entire parameter space, and therefore these other parameters, in addition to leaf biomass, also varied across different simulations. The concave relationships may be due to the simultaneous changes of all the variables and the interactions between the variables across the sample space. The effect of tree physiology (e.g., genus and leaf biomass) is greater than that of environmental factors (e.g., temperature, sunlight), which indicates planting of low-emitting genus is a good strategy to help prevent BVOC emissions (Nowak and Crane 2000). Temperature showed a convex relationship with both BVOC emissions, and extreme temperature can strongly increase BVOC emissions (Sharkey et al. 1991; Calfapietra et al.



2013). This may be due to the physiological functions of plants to protect against heat stress (Bonan 2016), and cooler environments can also reduce BVOC emissions. While genus and temperature both impact BVOC emissions, an urban planner generally has limited control over temperature; thus, maximizing the use of low VOC-emitting trees is an efficient strategy to prevent and reduce BVOC emissions.

### Carbon Storage and Sequestration

DBH played a dominant role for both carbon storage and sequestration (high  $\mu^*$  and  $S_{Ti}$ ) while land use and tree height had negligible effects (low  $\mu^*$  and  $S_{Ti}$ ). These findings indicate that in terms of data collection, increasing the accuracy of DBH measurements is an effective way to reduce model output uncertainty. DBH is important because allometric equations typically have a nonlinear relationship (e.g., exponential or log-log) between DBH and carbon storage, which means that even small errors in DBH measurements can lead to inaccurate carbon estimators (Jucker et al. 2017). These results suggest large trees play a more important role in estimators of carbon than small trees; large trees (DBH > 77 cm) store approximately 1,000 times more carbon than small trees (DBH < 7 cm) (Nowak 1994). Sustaining the health of large trees, and accurately measuring the DBH of large trees, is critical for carbon management and estimation.

Given the high correlation between DBH and height, it is counterintuitive that height had a small influence on carbon estimation (low  $\mu^*$  and  $S_{Ti}$ ). This difference is likely because some allometric equations do not use height as an explanatory variable (Equation 5). To compete for sunlight, trees often initially allocate resources to maximize tree height growth and approach their maximum height rapidly and then invest resources in diameter growth (Jucker et al. 2017). This makes carbon estimation based on height alone problematic, because trees of similar height can have very different woody biomass. In addition, urban trees are more open-grown than trees in forested stands. However, McPherson et al. (2016) state that allometric equations using both DBH and height as predictor variables typically have higher accuracy than corresponding allometric equations employing only DBH. Although not as large as previously thought, species also affected the estimations of carbon storage ( $S_{Ti} = 5.4\%$ ) and sequestration ( $S_{Ti} =$

6.5%) mainly through the selection of allometric equations of different forms and coefficients, which reflects differences among wood density and the water content of species (McPherson et al. 2013). This finding indicates that employing species-specific allometric equations could improve the accuracy of carbon estimators. Some studies also indicate that developing urban-specific allometric relationships is necessary (McHale et al. 2009). However, urban-specific allometric equations are scarce and location-specific, and the very few that exist are usually developed based on street trees (McPherson et al. 2016). Employing allometric equations developed for street trees for other urban trees may be problematic, as growing conditions vary widely across the urban environment. The current approach adopted in Eco-C, namely to use forest-based allometric equations and apply a biomass correction factor if it is an open-grown environment (Nowak et al. 2013), is popular in urban forestry, although some studies criticize that this simple correction results in conservative estimates of biomass (Aguaron and McPherson 2012). Future studies to develop urban-specific allometric equations and to perform uncertainty analyses of the biomass correction factor are needed.

For carbon sequestration, two additional input factors (CLE and condition) were examined. CLE represent site characteristics that impact sunlight to the tree crown, while condition is a measure of tree health. These two factors affect carbon estimators by adjusting tree growth rates. Both MOAT and VD showed that CLE and condition played moderate roles in model output (Figures 4 and 5), and the bin quantile regression analysis showed that carbon sequestration tended to increase linearly with the increase of CLE and condition (Table 4). These linear relationships may inadequately capture input-output relationships, as indicated by the low values of adjusted  $R^2$  (0.59 and 0.78). Consistent with the results from the bin quantile regression, VD also showed that the effects of CLE and condition may depend on the selection of other input factors, as indicated by the large difference between  $S_i$  and  $S_{Ti}$ . This difference may be due to the responses of different species, which may have different tolerances to abiotic and biotic stressors. This finding suggests that adaptive management approaches are needed to enhance and sustain forest health to maximize carbon sequestration and storage.

## Air Pollution Removal by Dry Deposition Processes

Both MOAT and VD analyses revealed that LAI was the most important factor for  $V_d$  of  $\text{NO}_2$ , and an important factor for  $V_d$  of  $\text{SO}_2$  and  $\text{O}_3$  (Figures 4 and 5). Bin quantile regression analysis showed that dry deposition of all three gases tended to respond linearly to LAI, which is consistent with the analysis performed by Hirabayashi et al. (2011). However, this linear relationship is unlikely to be maintained with an increase of LAI. As LAI increases, there is more chance of overlap among the leaf distribution, as well as resource (e.g., nutrients) and microenvironment (e.g., light) limitations (Chapin et al. 2011). Although trees adapt to these limitations, the leaves tend to behave at a lower efficiency as LAI increases, and thus are unlikely to maintain a linear relationship with  $V_d$  at higher LAI values (Van der Zande et al. 2009).

RH was an important factor for  $V_d$  of  $\text{NO}_2$ ,  $\text{SO}_2$ , and  $\text{O}_3$  (Figures 4 and 5). Eco calculates stomatal conductance by employing the Ball-Berry model, which is based on empirical relationships between stomatal conduction and photosynthesis from numerous leaf gas exchange measurements (Medlyn et al. 2011). During the process, Eco assumes RH is equivalent to the relative humidity at the leaf surface (Hirabayashi et al. 2011). Therefore, RH affects the stomatal conduction directly as an input to the Ball-Berry model, as well as indirectly through the influence of net leaf photosynthesis. Dry deposition of all three gases showed linear behavior with the change of RH (Table 4), which is consistent with the result obtained by Hirabayashi et al. (2011). However, a linear model may be inadequate to fully capture the response of  $V_d$  to RH, as indicated by relatively low adjusted  $R^2$  values (Table 4). Bonan (2015) shows that stomatal conductance responds to vapor pressure deficit linearly but at different rates with changing temperature, which indicates the linear model is not completely justified to represent the response of  $V_d$  to RH. The large differences between  $S_i$  and  $S_{Ti}$  also indicate that the response of  $V_d$  to RH may depend on other factors. Our conclusion regarding the relationship between RH and  $V_d$  is based on applying SA to the Ball-Berry model. However, the Ball-Berry model has been criticized for employing leaf surface RH as a direct input, because stomata sense and respond directly to transpiration water fluxes (Medlyn et al. 2011). Therefore, it is the leaf-to-air vapor pressure deficit rather than RH that

actually drives this process. Future model development should directly incorporate vapor pressure deficit into the calculation of the stomatal conductance.

Wind speed and pressure were among the least important factors for dry deposition (Figures 4 and 5). Wind speed is important for determining aerodynamic and quasi-laminar boundary layer resistances, because high winds reduce the thickness of the boundary layer of still air around each leaf and produce steeper gradients for the exchange of elements (e.g., water vapor, carbon dioxide) between the leaf surface and the atmosphere (Chapin et al. 2011). However, canopy resistance, which is the main driver of dry deposition, does not appear to be affected by wind speed (Hirabayashi et al. 2011); as a result, wind speed plays a minor role in determining dry deposition. Pressure also appeared to have a limited relationship with  $V_d$  (Figures 4 and 5). Pressure affects stomatal conduction indirectly through the computation of the direct and diffuse PAR (Hirabayashi et al. 2011).

Temperature and PAR affect the opening and closing of stomata during photosynthesis and transpiration (Bonan 2015). For the photosynthetic process, PAR mainly affects light-harvesting reactions, while temperature mainly influences carbon-fixing reactions through the control of enzymes (Bernacchi et al. 2013). For transpiration processes, water is taken up by the tree's roots, transported vertically by the xylem, and evaporates from the leaf surface to the ambient atmosphere through a water potential gradient. The entire process is driven by vapor pressure deficit and radiative forcing, which is closely correlated to temperature and light intensity (Will et al. 2013). The MOAT and VD analyses gave different conclusions about the  $V_d$  response to temperature and PAR. MOAT indicated that temperature was the most important factor, while VD indicated that PAR had the largest effect. The differences may be due to the different approaches of the two methods, as well as the nonlinear response curves of  $V_d$  to temperature and PAR. MOAT is based on changing one factor at a time across its entire range while fixing all other factors at certain values, and therefore may not fully explore the parameter space (a 6-dimensional space in this case). VD is a global method based on the simultaneous change of all input factors among their PDFs. The typical response curve of stomatal conduction to PAR for many species is that conduction increases with increasing PAR up to about 300 to 400  $\text{W/m}^2$ , and is

relatively insensitive to further increases in PAR (Jones 2013). When the randomly selected levels used to calculate EE lie outside the range of 300 to 400 W/m<sup>2</sup>, a conclusion that conduction is insensitive to PAR may be obtained. The response curve of  $V_d$  to PAR was concave in the bin quantile regression analysis (Table 4).  $V_d$  also responded to temperature in a concave way but with less concavity than with PAR (Table 4). Bonan (2015) shows that there exists an optimal temperature for the response of stomatal conductance, with the values being 15 to 25 °C for most C<sub>3</sub> plants (plants that don't have photosynthetic adaptations to reduce photorespiration). The different response mechanisms of  $V_d$  to PAR and temperature indicate potential model inadequacies to differentiate these response curves. More systematic modeling practices, which can capture different  $V_d$  response mechanisms (e.g., threshold and optimal points), are needed in the future.

PAR, LAI, RH, and temperature all play important roles in dry deposition processes. The VD analysis

showed that PAR and RH were the most sensitive variables for  $V_d$  of SO<sub>2</sub> and O<sub>3</sub>, while they were less sensitive for  $V_d$  of NO<sub>2</sub> (Figure 5). This may be because there were greater changes in  $V_d$  for SO<sub>2</sub> and O<sub>3</sub> than for NO<sub>2</sub> when input variables were altered similarly. For example, when increasing PAR from 50 to 600 W/m<sup>2</sup>,  $V_d$  of SO<sub>2</sub> and O<sub>3</sub> increased by 0.31 and 0.33 cm/s<sup>-1</sup>, respectively, which is larger than that of NO<sub>2</sub> (0.18 cm/s<sup>-1</sup>). The variations of  $V_d$  with respect to PAR for these three gases are detailed in Figure 6. Hirabayashi et al. (2011) also observed similar patterns in their sensitivity analysis of Eco-D in Baltimore. These different degrees of sensitivity may be due to different parametrization schemes for the calculation of the quasi-laminar boundary layer and canopy resistances, where NO<sub>2</sub> shows higher values for mesophyll and cuticular resistances compared with SO<sub>2</sub> and O<sub>3</sub> (Hirabayashi et al. 2015). For PAR and LAI, Eco employs one of the best available processes in the literature to scale up from the leaf to canopy,

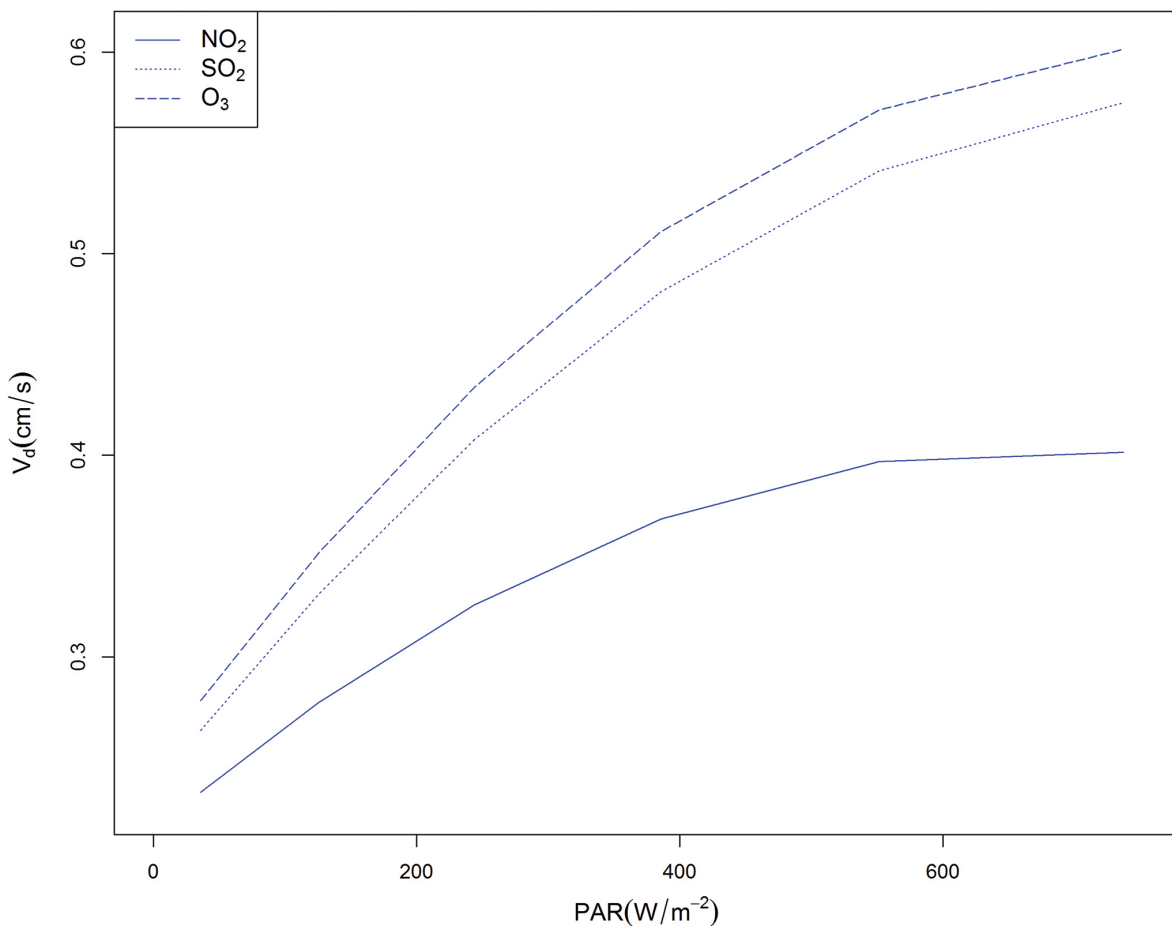


Figure 6. Relationships between  $V_d$  and PAR.

and the interaction between PAR and LAI is fully captured by different components of PAR (e.g., direct, diffuse) and LAI (e.g., sunlit, shaded) (Hirabayashi et al. 2011). For RH and temperature, single values, instead of vertical profiles around canopy height, are employed in Eco, which may constrain the model performance. In addition, the model assumes that leaf temperature is equal to air temperature, while leaf traits (e.g., hair) and properties (e.g., latent heat exchange) may make leaf temperature differ from air temperature (Yu et al. 2015). Future model development may focus on the improved representation of RH and leaf temperature to reduce the uncertainties of model outputs.

## CONCLUSIONS

In this study, sensitivity analyses (SAs) were performed to investigate how the characteristics of the Eco inputs impact the ecosystem services and disservices of urban trees predicted by this model. Here the focus is on the inputs to three Eco modules: BVOC emissions (Eco-B), carbon storage and sequestration (Eco-C), and dry deposition velocity of air pollutants (Eco-D). Two SAs with different theoretical foundations are employed. Morris one-at-a-time (MOAT) is based on changing one factor at a time across its entire range to see what effect it has on the output, while variance decomposition (VD) is based on decomposing the variance of the output into different fractions which can be attributed to different inputs or their interactions. The results provide useful information for future urban Forest Inventory and Analysis (FIA) data collections (<https://www.nrs.fs.fed.us/fia/urban>), model uncertainty analyses, and urban forest management.

Genus has the largest influence on BVOC emissions by determining base emission rates and the interaction between genus and other input factors. Temperature shows a convex relationship with both BVOC emissions, indicating that BVOCs increase at a greater rate with temperature as temperature increases (Sharkey et al. 1991). High temperatures can strongly increase BVOC emissions. Leaf biomass has a concave relationship with BVOCs, indicating that the emissions increase at a faster rate initially, then at a slower rate, and finally become relatively insensitive to the change in leaf biomass; understanding these inflection points from sensitive to insensitive is important to control BVOC emissions while maximizing other leaf related ecosystem

services. PAR has a linear relationship with isoprene emissions. These findings indicate that maximizing the use of low VOC-emitting trees is an efficient strategy to prevent and reduce BVOC emissions, and maintaining cooler environments (e.g., through tree transpiration) can also help to reduce BVOC emissions.

DBH has the greatest influence on carbon storage and sequestration provided by urban trees, and carbon storage tends to increase in a convex manner as DBH increases. Unlike relationships among input variables and BVOC emissions, which show strong interactions, the combined interactions between DBH and the other input variables and their influence on carbon storage and sequestration appears minimal. These results indicate that increasing the accuracy of DBH measurements, especially for larger trees, is critical for accurate carbon estimators. Employing species-specific allometric equations can also improve the accuracy of carbon estimators. Effort should be spent on improving current or developing new biomass equations in urban environments whenever possible. By contrast, tree height and land use appear to play minimal roles in carbon storage and sequestration. For carbon sequestration, tree condition and CLE are also important. Maintaining a good site environment and tree health are critical to maximizing carbon storage and sequestration provided by trees.

For the dry deposition processes, PAR, LAI, RH, and temperature all play important roles, with PAR and LAI generally having the largest influence. Dry deposition velocity is sensitive to LAI and RH in a nearly linear manner, while it is sensitive to temperature and PAR in a concave manner. Air pressure has almost no influence on dry deposition, while wind speed has a minimal influence. The interaction between the input variables and their influence on dry deposition velocity is also minimal. There exists an optimal temperature for maximum dry deposition velocity, while PAR affects dry deposition velocity up to a certain threshold value. Representation of RH and temperature around the entire canopy space as a single value may constrain model performance. Future model development should focus on the improved representative of RH and leaf temperature.

Trees provide important ecosystem services and disservices to urban communities. This analysis provided insight into the relationships between tree characteristics, environmental factors, and model parameters on the estimated ecosystem services and disservices



of urban trees. Results identified the most important input variables to i-Tree Eco, and future field plot campaigns should emphasize accurate measurements for these variables. In addition, several gaps regarding the representativeness of input variables and model parameters in Eco were identified, which may help to assist future model development. The knowledge gained from this analysis is beneficial to subsequent model uncertainty analyses, which in turn contributes to improved urban forest planning and management.

## LITERATURE CITED

- Aguaron E, McPherson EG. 2012. Comparison of methods for estimating carbon dioxide storage by Sacramento's urban forest. In: Lal R, Augustin B, editors. *Carbon sequestration in urban ecosystems*. Dordrecht (Netherlands): Springer.
- Bernacchi CJ, Bagley JE, Serbin SP, Ruiz-Vera UM, Rosenthal DM, Vanlooocke A. 2013. Modelling C 3 photosynthesis from the chloroplast to the ecosystem. *Plant, Cell and Environment*. 36(9):1641-1657.
- Bonan G. 2015. *Ecological climatology: concepts and applications*. New York (NY, USA): Cambridge University Press. 617 p.
- Bonan GB. 2016. Forests, climate, and public policy: a 500-year interdisciplinary odyssey. *Annual Review of Ecology, Evolution, and Systematics*. 47:97-121.
- Breuer L, Eckhardt K, Frede HG. 2003. Plant parameter values for models in temperate climates. *Ecological Modelling*. 169(2-3):237-293.
- Calfapietra C, Fares S, Manes F, Morani A, Sgrigna G, Loreto F. 2013. Role of biogenic volatile organic compounds (BVOC) emitted by urban trees on ozone concentration in cities: a review. *Environmental Pollution*. 183:71-80.
- Campolongo F, Cariboni J, Saltelli A. 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*. 22(10):1509-1518.
- Celik A, Makkawi A, Muneer T. 2010. Critical evaluation of wind speed frequency distribution functions. *Journal of Renewable and Sustainable Energy*. 2(1):013102.
- Chapin III F, Matson P, Vitousek P. 2011. *Principles of terrestrial ecosystem ecology*. New York (NY, USA): Springer-Verlag. p. 139.
- Duh JD, Shandas V, Chang H, George LA. 2008. Rates of urbanisation and the resiliency of air and water quality. *Science of the Total Environment*. 400(1):238-256.
- Guenther A. 1997. Seasonal and spatial variations in natural volatile organic compound emissions. *Ecological applications*. 7(1):34-45.
- Guenther A, Hewitt CN, Erickson D, Fall R, Geron C, Graedel T, Harley P, Klinger L, Lerdau M, McKay W. 1995. A global model of natural volatile organic compound emissions. *Journal of Geophysical Research: Atmospheres*. 100(D5):8873-8892.
- Helton JC, Davis FJ. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*. 81(1):23-69.
- Hirabayashi S, Kroll CN, Nowak DJ. 2011. Component-based development and sensitivity analyses of an air pollutant dry deposition model. *Environmental Modelling & Software*. 26(6):804-816.
- Hirabayashi S, Kroll CN, Nowak DJ. 2015. i-Tree Eco Dry Deposition Model Descriptions. [Accessed 2019 July]. [http://www.itreetools.org/eco/resources/iTree\\_Eco\\_Dry\\_Deposition\\_Model\\_Descriptions.pdf](http://www.itreetools.org/eco/resources/iTree_Eco_Dry_Deposition_Model_Descriptions.pdf)
- Iman RL, Conover WJ. 1982. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics-Simulation and Computation*. 11(3):311-334.
- Jones HG. 2013. *Plants and microclimate: a quantitative approach to environmental plant physiology*. Cambridge (UK): Cambridge University Press. 456 p.
- Jucker T, Caspersen J, Chave J, Antin C, Barbier N, Bongers F, Dalponte M, Ewijk KY, Forrester DI, Haeni M. 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology*. 23(1):177-190.
- Lin J. 2020. Developing a composite indicator to prioritize tree planting and protection locations. *Science of the Total Environment*. 717:137269.
- Lin J, Kroll CN, Nowak DJ, Greenfield EJ. 2019. A review of urban forest modeling: implications for management and future research. *Urban Forestry & Urban Greening*. 43:126366.
- Marino S, Hogue IB, Ray CJ, Kirschner DE. 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*. 254(1):178-196.
- McHale M, Burke I, Lefsky M, Peper P, McPherson E. 2009. Urban forest biomass estimates: is it important to use allometric relationships developed specifically for urban trees? *Urban Ecosystems*. 12(1):95-113.
- McPherson EG, Simpson JR, Xiao Q, Wu C. 2011. Million trees Los Angeles canopy cover and benefit assessment. *Landscape and Urban Planning*. 99(1):40-50.
- McPherson EG, van Doorn NS, Peper PJ. 2016. *Urban tree database and allometric equations*. Albany (CA, USA): US Department of Agriculture, Forest Service, Pacific Southwest Research Station. General Technical Report PSW-GTR-253. 86 p.
- McPherson EG, Xiao Q, Aguaron E. 2013. A new approach to quantify and map carbon stored, sequestered and emissions avoided by urban forests. *Landscape and Urban Planning*. 120:70-84.
- Medlyn BE, Duursma RA, Eamus D, Ellsworth DS, Prentice IC, Barton CV, Crous KY, De Angelis P, Freeman M, Wingate L. 2011. Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Global Change Biology*. 17(6):2134-2144.
- Morani A, Nowak DJ, Hirabayashi S, Calfapietra C. 2011. How to select the best tree planting locations to enhance air pollution removal in the MillionTreesNYC initiative. *Environmental Pollution*. 159(5):1040-1047.
- National Climate Data Center (NCDC). 2018. World's Largest Archive of Climate Data: National Climate Data Center. Silver Spring (MD, USA): National Oceanic and Atmospheric Administration. [Accessed 2018 December]. <http://www.ncdc.noaa.gov/oa/ncdc.html>
- Nowak DJ. 1994. Atmospheric carbon dioxide reduction by Chicago's urban forest. In: McPherson EG, Nowak DJ, Rowntree

- RA, editors. Chicago's urban forest ecosystem: results of the Chicago urban forest climate project. Newtown Square (PA, USA): US Department of Agriculture, Forest Service, Northeastern Forest Experiment Station. General Technical Report NE-186. p. 83-94.
- Nowak DJ, Crane DE. 2000. The Urban Forest Effects (UFORE) Model: quantifying urban forest structure and functions. In: Hansen M, Burk T, editors. *Proceedings: integrated tools for natural resources inventories in the 21st century*. IUFRO Conference; 1998 August 16-20; Boise (ID, USA). St. Paul (MN, USA): US Department of Agriculture, Forest Service, North Central Research Station. General Technical Report NC-212. p. 714-720.
- Nowak DJ, Crane DE, Stevens JC. 2006. Air pollution removal by urban trees and shrubs in the United States. *Urban Forestry & Urban Greening*. 4(3):115-123.
- Nowak DJ, Crane DE, Stevens JC, Ibarra M. 2002. Brooklyn's urban forest. Newtown Square (PA, USA): Department of Agriculture, Forest Service, Northeastern Research Station. General Technical Report NE-290.
- Nowak DJ, Crane DE, Stevens JC, Hoehn RE, Walton JT, Bond J. 2008a. A ground-based method of assessing urban forest structure and ecosystem services. *Arboriculture & Urban Forestry*. 34(6):347-358.
- Nowak DJ, Greenfield EJ, Hoehn RE, Lapoint E. 2013. Carbon storage and sequestration by trees in urban and community areas of the United States. *Environmental Pollution*. 178:229-236.
- Nowak DJ, Walton JT, Stevens JC, Crane DE, Hoehn RE. 2008b. Effect of plot and sample size on timing and precision of urban forest assessments. *Arboriculture & Urban Forestry*. 34(6):386-390.
- Pace R, Biber P, Pretsch H, Grote R. 2018. Modeling ecosystem services for park trees: sensitivity of i-Tree Eco simulations to light exposure and tree species classification. *Forests*. 9(2):89.
- Poumanyong P, Kaneko S. 2010. Does urbanization lead to less energy use and lower CO<sub>2</sub> emissions? A cross-country analysis. *Ecological Economics*. 70(2):434-444.
- Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*. 181(2):259-270.
- Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. 2008. *Global sensitivity analysis: the primer*. Chichester, West Sussex (UK): John Wiley & Sons Ltd. p. 110.
- Saltelli A, Tarantola S, Campolongo F, Ratto M. 2004. *Sensitivity analysis in practice: a guide to assessing scientific models*. Chichester, West Sussex (UK): John Wiley & Sons Ltd. p. 49.
- Selmi W, Weber C, Rivière E, Blond N, Mehdi L, Nowak D. 2016. Air pollution removal by trees in public green spaces in Strasbourg city, France. *Urban Forestry & Urban Greening*. 17:192-201.
- Sharkey TD, Holland EA, Mooney HA. 1991. *Trace gas emissions by plants*. New York (NY, USA): Academic Press.
- Steffens JT, Wang YJ, Zhang KM. 2012. Exploration of effects of a vegetation barrier on particle size distributions in a near-road environment. *Atmospheric Environment*. 50:120-128.
- Van der Zande D, Mereu S, Nadezhkina N, Cermak J, Muys B, Coppin P, Manes F. 2009. 3D upscaling of transpiration from leaf to tree using ground-based LiDAR: application on a Mediterranean Holm oak (*Quercus ilex* L.) tree. *Agricultural and Forest Meteorology*. 149(10):1573-1583.
- Van Griensven A, Meixner T, Grunwald S, Bishop T, Diluzio M, Srinivasan R. 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*. 324(1-4):10-23.
- Wasserman L. 2013. *All of statistics: a concise course in statistical inference*. Heidelberg (Germany): Springer Science & Business Media. p. 293.
- Will RE, Wilson SM, Zou CB, Hennessey TC. 2013. Increased vapor pressure deficit due to higher temperature leads to greater transpiration and faster mortality during drought for tree seedlings common to the forest-grassland ecotone. *New Phytologist*. 200(2):366-374.
- Yang J, McBride J, Zhou J, Sun Z. 2005. The urban forest in Beijing and its role in air pollution reduction. *Urban Forestry & Urban Greening*. 3(2):65-78.
- Yao AY. 1974. A statistical model for the surface relative humidity. *Journal of Applied Meteorology*. 13(1):17-21.
- Yu MH, Ding GD, Gao GL, Sun BP, Zhao YY, Wan L, Wang DY, Gui ZY. 2015. How the plant temperature links to the air temperature in the desert plant *Artemisia ordosica*. *PLoS One*. 10(8).

## ACKNOWLEDGMENTS

This research was supported by the United States Department of Agriculture (USDA) Forest Service with grant 16-DG-11132544-034 recommended by the National Urban and Community Forestry Advisory Council, grant 11-JV-11242308-112 from the USDA Forest Service Northern Research Station, and additional support from the SUNY ESF Department of Environmental Resources Engineering. Special thanks to Robert Hoehn for helping with i-Tree Eco runs.

Jian Lin (corresponding author)

Graduate Program in Environmental Science  
State University of New York College of Environmental Science and Forestry (SUNY-ESF)  
321 Baker Lab  
1 Forestry Dr  
Syracuse, NY, USA  
jlin27@syr.edu  
315-708-9886

Charles N. Kroll

Department of Environmental Resources Engineering  
State University of New York College of Environmental Science and Forestry (SUNY-ESF)  
424 Baker Lab  
1 Forestry Dr  
Syracuse, NY, USA  
cnkroll@esf.edu  
315-470-6699

David J. Nowak  
 Northern Research Station, USDA Forest Service  
 5 Moon Library  
 1 Forestry Dr  
 Syracuse, NY, USA  
 dnowak@fs.fed.us  
 315-448-3212

**Conflicts of Interest:** David J. Nowak reports originally developing and overseeing the current development of i-Tree Eco, a public domain model that is distributed for free. The remaining authors made no disclosures.

**Résumé.** Il est connu que les arbres fournissent divers services écosystémiques et causent aussi certains préjudices aux communautés urbaines pouvant être quantifiés par le recours à des modèles fondés sur des données de terrain et environnementales. L'impact des variables environnementales et de la structure des arbres sur les résultats du modèle est souvent incertain. Nous réalisons une analyse de sensibilité (AS) de i-Tree Eco, un modèle courant pour la forêt urbaine, afin d'analyser l'impact relatif de différentes entrées de modèle sur trois sorties de module: les émissions de composé organique volatil biogénique (COVB) (isoprène et monoterpènes); le stockage et la séquestration du carbone; et le dépôt sec de dioxyde d'azote, de dioxyde de soufre et d'ozone. Les méthodes AS incluaient des applications originales de la méthode Morris et une méthode de décomposition fondée sur la variance qui intègre la simulation Monte Carlo avec un échantillonnage par hypercube latin et une analyse Iman Conover. Une étude de cas fut réalisée dans la ville de New-York, États-Unis avec des données de terrain recueillies en 2013. Le genre de l'arbre a eu la plus grande influence sur les émissions de COVB en établissant les taux d'émissions de base et ses interactions élevées avec les autres facteurs d'entrée, alors que les émissions de COVB sont sensibles à la biomasse des feuilles d'une manière concave et à la température d'une manière convexe, tandis que les émissions d'isoprène montrent une forte relation linéaire avec le rayonnement photosynthétiquement actif (RPA). Le diamètre à hauteur de poitrine joue le rôle le plus important quant à l'estimation du stockage et de la séquestration du carbone; l'exposition à la lumière du houppier et la condition de l'arbre sont également importants pour la séquestration du carbone. La vitesse de dépôt sec est sensible à l'indice de surface foliaire et à l'humidité relative d'une manière quasi linéaire ainsi que sensible à la température et au RPA d'une manière concave. Les résultats fournissent des orientations afin de faciliter les futures campagnes avec parcelles expérimentales sur le terrain et le développement de modèles. Les informations révélées par les AS sont également bénéfiques pour réduire les incertitudes des modèles ce qui, à son tour, facilite une gestion et une prise de décision plus efficaces en matière de forêts urbaines.

**Zusammenfassung.** Bäume sind dafür bekannt, dass sie verschiedene Ökosystemleistungen und Fehlleistungen für städtische Gemeinschaften erbringen, die mit Hilfe von Modellen auf der Grundlage von Feld- und Umweltdaten quantifiziert werden können. Es ist oft ungewiss, wie sich die Baumstruktur und Umweltvariablen auf das Ergebnis der Modelle auswirken. Hier

führen wir eine Sensitivitätsanalyse von i-Tree Eco, einem gängigen städtischen Waldmodell, durch, um die relative Auswirkung verschiedener Modelleingaben auf drei Modulausgaben zu analysieren: Emissionen biogener flüchtiger organischer Verbindungen (BVOC) (Isopren und Monoterpene), Kohlenstoffspeicherung und -bindung und trockene Deposition von Stickstoffdioxid, Schwefeldioxid und Ozon. Die Methoden der Sensitivitätsanalyse umfassten neuartige Anwendungen der Morris-Einzelmethode und eine varianzbasierte Zersetzungsmethode, die die Monte-Carlo-Simulation mit der lateinischen Hypercube-Probenahme und der Iman-Conover-Analyse integriert. Eine Fallstudie wurde in New York City, New York, USA, durchgeführt, bei der 2013 Feldflächen-daten gesammelt wurden. Die Gattung hat den größten Einfluss auf die BVOC-Emissionen, indem sie die Basis-Emissionsraten bestimmt und ihre hohen Wechselwirkungen mit anderen Eingangsfaktoren. Die BVOC-Emissionen reagieren auf die Blatt-Biomasse in konkaver Weise und die Temperatur in konvexer Weise. Die Isopren-Emissionen zeigen eine starke lineare Beziehung mit der photosynthetisch aktiven Strahlung (PAR). Der Durchmesser in Brusthöhe spielt die wichtigste Rolle sowohl für die Kohlenstoffspeicherung als auch den [sequestration estimators]: Die Belichtung der Krone und der Zustand des Baumes sind ebenfalls wichtig für die Kohlenstoffbindung. Die Trocken-depositionsgeschwindigkeit ist, in nahe zu linearer Weise, sensibel für den Blattflächenindex und die relative Feuchte, während sie in konkaver Weise für die Temperatur und die PAR empfindlich ist. Die Ergebnisse bieten eine Anleitung zur Erleichterung künftiger Feldparzellenkampagnen und Modellentwicklungen. Das von der Sensitivitätsanalyse aufgedeckte Wissen kann Unsicherheiten im Umgang mit dem Modell reduzieren, was wiederum Entscheidungsfindungen erleichtert und eine effektivere städtische Waldbewirtschaftung ermöglicht.

**Resumen.** Se sabe que los árboles proporcionan diversos servicios ecosistémicos a las comunidades urbanas, que se pueden cuantificar utilizando modelos basados en datos de campo y ambientales. A menudo es incierto cómo la estructura del árbol y las variables ambientales afectan la salida del modelo. Aquí realizamos un análisis de sensibilidad (SA) de i-Tree Eco, un modelo común de bosque urbano, para analizar el impacto relativo de diferentes entradas de modelo en tres salidas de módulo: emisiones de compuesto orgánico volátil biogénico (BVOC) (isopreno y monoterpénos), almacenamiento y secuestro de carbono, deposición en seco de dióxido de nitrógeno, dióxido de azufre y ozono. Los métodos SA incluyeron nuevas aplicaciones del método uno a la vez de Morris y un método de descomposición basado en varianza, que integra la simulación de Monte Carlo con el muestreo de hipercubos latinos y el análisis de Iman Conover. Se realizó un estudio de caso en la ciudad de Nueva York, Nueva York, EE. UU., con datos de la gráfica de campo recopilados en 2013. El género tiene la mayor influencia en las emisiones de BVOC al determinar las tasas de emisión base y sus altas interacciones con otros factores de entrada. Las emisiones de BVOC son sensibles a la biomasa de las hojas de una manera cóncava y la temperatura de una manera convexe, mientras que las emisiones de isopreno muestran una fuerte relación lineal con la radiación fotosintéticamente activa (PAR). El diámetro a la altura del pecho juega el papel más importante tanto para el almacenamiento como para los estimadores de secuestro de carbono; la exposición

a la luz de la copa y el estado de los árboles también son importantes para el secuestro de carbono. La velocidad de deposición en seco es sensible al índice de área foliar y a la humedad relativa de una manera casi lineal, mientras que sensible a la temperatura y al PAR de una manera cóncava. Los resultados proporcionan

orientación para facilitar futuras campañas de trazado de campo y desarrollo de modelos. El conocimiento revelado por la SA también es beneficioso para la reducción de la incertidumbre del modelo, que a su vez facilita una gestión y toma de decisiones más eficaces de los bosques urbanos.

### *Arboriculture & Urban Forestry Quiz Questions*

To complete this quiz, go to the ISA website, log into your MyISA account, and make your way to the page for *Arboriculture & Urban Forestry* CEU Quizzes ([www.isa-arbor.com/store/ceuquizzes/113](http://www.isa-arbor.com/store/ceuquizzes/113)).

Add the quiz to your cart, proceed through checkout, and look for the content to appear on your personal dashboard under the header, "My Quizzes." If you need a username and password, send us an e-mail ([isa@isa-arbor.com](mailto:isa@isa-arbor.com)).

A passing score for this quiz requires sixteen correct answers. Quiz results will display immediately upon quiz completion. CEU(s) are processed immediately. You may take the quiz as often as is necessary to pass.

*CEU quiz by Eric North, University of Nebraska-Lincoln, Lincoln, Nebraska, USA*

