Improving Remote Sensed Data Products Using Bayesian Methodology for the Analysis of Computer Model Output Grant Number NNG06GI77G – Final Report

Robin D. Morris, USRA-RIACS
Athanasios Kottas, UCSC
Roberto Furfaro, Barry D. Ganapol, UofA

July 23, 2007

Abstract

We report on work performed under the start-up grant awarded under the 2005 ROSES NRA. This work was focused on addressing the concerns of the reviewers regarding the foundations and feasibility of the proposed methodology, in particular the sensitivity of the conclusions to particular modeling choices. We report on an extensive cross-validation study of the effects of choices of the mean and correlation function of the Gaussian Process (GP) used to emulate the Leaf-Canopy Model (LCM) Radiative Transfer Model, and some preliminary results on the use of the Dirichlet Process Mixture Model emulator.

We also completed the derivation of the main effects and sensitivity indices of the LCM using the GP model. These latter results have been submitted to IEEE Transactions on Geoscience and Remote Sensing, and the paper is attached to this report.

1 Verifying the Applicability of the Models Chosen, and the Insensitivity of the Results to Details of the Models

The main aim of the investigation was to verify the applicability of the nonparametric Bayesian methodology proposed, and, further, the impact of specific methodology choices on the results of the investigation – if the uncertainty characterization depends strongly on the details of the nonparametric model chosen, then clearly it would be difficult for the domain scientists to accept the results.

Figure 1 shows the reflectance spectrum generated by the LCM for wavelengths from 400-2500nm, for various LAI values[1]. A serious concern when modeling spectra of this type is the dependence of spectral features on the model input parameters. If, for example, the *position* of the sharp fall-off that is shown at 1400nm is a function of one of the LCM inputs, then the definition of a suitable mean function for the GP model would be extremely difficult. MODIS does not, however,

---

[1]The other inputs to the LCM are Sun angle - 0; LAD - planophile; Leaf thickness - 0.0134cm; soil reflectance - 0.3; water concentration - 72.3%; protein concentration - 6.46e-4 g/cm^2; lignin concentration - 1.21e-3 g/cm^2; chlorophyll concentration - 37.8 μg/cm^2
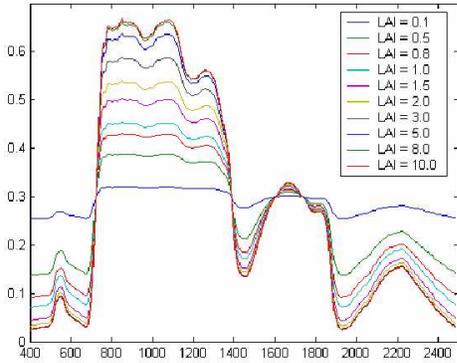
Figure 1: The spectral response of the LCM for various LAI values.

| | Power in correlation function | | | |
	2.0	1.9	1.7	1.0
linear	4.092	4.089	4.089	4.091
log	3.893	3.893	3.893	3.893
log-lin	4.176	4.179	4.179	4.178

Table 1: Mean log-posterior predictive probability. (Averaged over 90 design points and 8 spectral bands)

record a continuous spectrum. Rather, it makes measurements only in a number of discrete spectral bands. Further, only a subset of these bands are sensitive to vegetation, and are thus used for LAI estimation – the current production algorithm, MOD15, uses seven spectral bands as input. Thus, the interest is not in modeling continuous spectra, but, rather, modeling reflectance as a function of LAI, chlorophyll, etc., *at a small number of distinct wavelengths*.

Figure 2 shows, rather than a spectrum, the response of the LCM in MODIS band 4 (centered at 550nm), which corresponds to the small peak to the left of figure 1. The response is plotted as function of LAI and chlorophyll (the other model inputs being held constant) and can be seen to be a smooth function. This was also the case for the other MODIS bands. Finding a suitable mean function, $E(f(LAI, CHL))$, for the GP model is therefore the search for a basis that adequately represents this smooth function. From study of figure 1 (and the corresponding plots for the other spectral bands), the choice of basis function was narrowed to

- *linear* (A + B×LAI + C×CHL)

- *logarithmic* (A + B×log(LAI) + C×log(CHL))

- *log-lin* (A + B×LAI + C×CHL + D×log(LAI) + E×log(CHL)).

The GP model also has as a parameter the power in the exponential covariance function. This was chosen to be from the set {1, 1.7, 1.9, 2}, varying the function space from nowhere differentiable to infinitely differentiable.

To compare the space of models (3 bases × 4 powers), an extensive cross-validation study was performed. A 90 point d-optimal design was constructed over the LAI×CHL range, and leave-one-out cross validation was performed, consisting of 90 repetitions of training the GP model over the 89 included points, and predicting the 90-th point. The log-posterior predictive probability of the 90-th point was computed, and averaged over the 8 spectral bands considered. Figure 2 shows the result of one of the bands. The circles are the values computed from the LCM, and the lines are the 5%-95% predictive probability intervals. Table 1 shows that the modeling is insensitive to the power in the correlation function, and that the log-lin basis succeeds in modeling very well the shape
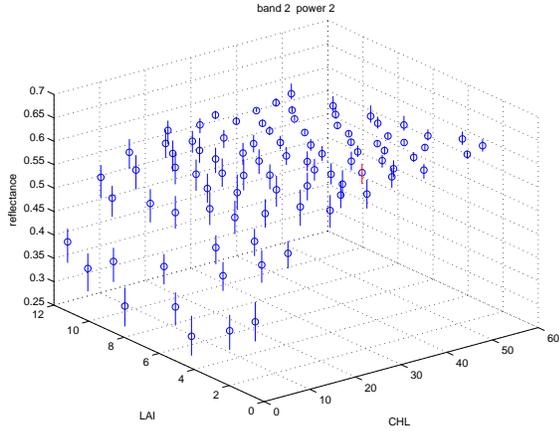
2

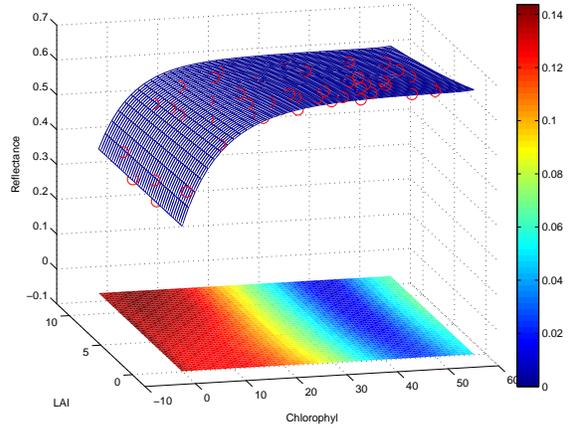Figure 2: Response of the LCM at MODIS band 4 (circles) and GP prediction (lines). See text for details.

Figure 3: Fit of the GP Model to the 90 points from the LCM and 5%-95% interval (colored plot at reflectance=0)

of the response surface. Figure 3 shows the predicted mean surface, with the 5%-95% probability interval plotted as the coloured surface. The size of these intervals is quite small compared with the function values.

To further ensure that the conclusions of our investigation are not strongly dependent on the choice of nonparametric model, we also obtained some initial results from the DP mixture model. For comparison with table 1, the mean log-posterior predictive probability over the 90 point leave-one-out cross validation was 3.794. This compares well, and is a good indication that we can model the output of the LCM in a manner that is insensitive to the details of the Bayesian nonparametric model chosen.

Some of these results were presented at the 2006 AGU Fall Meeting.

2   Main Effects and Sensitivity Indices

We also completed an initial study of the Main Effects and Sensitivity Indices of the LCM model, using only the GP approximation. This study used maximum-likelihood estimation of the parameters of the GP model, and as such underestimates the uncertainties associated with the Main Effects and Sensitivity Indices. However, this is expected to be a small effect, as the GP model is relatively insensitive to the exact parameter values in the region of parameter space that the maximum likelihood estimate produced. This initial study is also incapable of estimating the uncertainty of the Sensitivity Indices. Computation of these uncertainties will follow from the development of the fully inferential approach to sensitivity analysis that we will develop in future research.

A paper describing the Main Effects and Sensitivity Indices for the LCM model, and their implications for researchers in radiative transfer modeling and remote sensing has been submitted to IEEE Transactions on Geoscience and Remote Sensing. It is attached as an appendix to this report.

3   Presentations and Publications

- "A Gaussian Process Approach to Quantifying the Uncertainty of Biospheric Parameters from Remote Sensing Observations". Presented at the *AGU Fall Meeting*, San Francisco, 2006

- "An Analysis of the Uncertainties in Radiative Transfer Models Used in Remote Sensed Data Product Generation". NASA Science Technology Conference, Adelphi, MD, 2007.

- "A Statistical Framework for the Sensitivity Analysis of Radiative Transfer Models Used in Remote Sensed Data Product Generation", Submitted to *IEEE Transactions on Geoscience and Remote Sensing*, June 2007

A   Paper submitted to IEEE Transactions on Geoscience and Remote Sensing

A Statistical Framework for the Sensitivity Analysis of Radiative Transfer Models Used in Remote Sensed Data Product Generation

Robin D. Morris, Athanasios Kottas, Matt Taddy, Roberto Furfaro and Barry D. Ganapol

Abstract

The process of predicting a satellite observation of a vegetated region (e.g. a MODIS scene) involves running a Radiative Transfer Model (RTM). The RTM takes as input various biospheric and illumination parameters and computes the upwelling radiation at the top of the canopy that is ultimately observed by the satellite mounted sensor. The question we address is: which of the inputs to the RTM has the greatest impact on the computed observation?

We study the Leaf Canopy Model (LCM) RTM. The LCM was designed to study the feasibility of observing leaf chemistry remotely. It takes as input leaf chemistry variables (chlorophyll, water, lignin, cellulose) and canopy structural parameters (leaf area index, leaf angle distribution, soil reflectance, sun angle). We present a statistical approach to sensitivity analysis of RTMs, to answer the question posed above. The focus is on global sensitivity analysis, which studies how the RTM output changes as the inputs vary continuously according to a probability distribution over the input space. The influence of each input variable is captured through the determination of the "main effects" and "sensitivity indices". Direct computation requires extensive runs of the RTM, which is computationally expensive. We develop a Gaussian Process approximation to the RTM output to enable efficient computation. We illustrate how the approach can effectively determine the input variables that are vital for accurate prediction. The methods are applied to the LCM with 7 inputs and output obtained at 8 wavelengths associated with specific MODIS bands that are sensitive to vegetation.

Index Terms

radiative transfer model, MODIS, sensitivity analysis, main effects, sensitivity index, Gaussian Process

R. D. Morris is with USRA-RIACS, 444 Castro St, Suite 320, Mountain View, CA 94041. rdm@riacs.edu. Tel (650) 966 5035; fax (650) 966 5021. Corresponding author.

A. Kottas and M. Taddy are with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz.

R. Furfaro and B.D. Ganapol are with the Department of Aerospace and Mechanical Engineering, University of Arizona.

A Statistical Framework for the Sensitivity Analysis of Radiative Transfer Models Used in Remote Sensed Data Product Generation

I. INTRODUCTION

The accurate estimation of properties of the biosphere is critical for our understanding of the Earth's coupled system. The atmosphere, oceans and land comprise a complex, coupled dynamical system, and valid statistical prediction of the properties of this system, and its changes, require inputs that are both accurate and have their uncertainties accurately quantified.

Global models require global observations, and the only effective method for making routine global measurements is via sensors mounted on orbiting satellites. Typically, however, satellite mounted sensors do not measure directly the quantity of interest. Passive visible/near infra-red sensors measure upwelling radiation, and it is from these measurements that the biospherical parameters of interest must be inferred.

This inference process is complex. It is the inversion of the process of sunlight passing through the atmosphere, being reflected off vegetation on the ground, and then passing again through the atmosphere before being detected by the satellite mounted sensor. Clearly in this scenario the uncertainty introduced by traditional "noise" in the actual sensor will be swamped by the number of places at which uncertain process models enter the estimation. In the brief outline above we have two, one for the propagation of light through the atmosphere, and the second for the reflection of light by the vegetation on the ground. It is the uncertainty characteristics of the second of these process models that we will analyze in this paper.

Analyzing, quantifying and reporting the uncertainty in remote sensed data products is of great importance. It is the only way in which the uncertainty of further analyses using these data products as inputs can be quantified. Analyzing the source of the data product uncertainties can identify where the models must be improved, or where better input information must be obtained. Both of these aspects are known; the editorial for the Special Issue on Global Land Product Validation [1] wrote

> users need access to quantitative information on product uncertainties

and that

> [m]aking quantified accuracy information available to the user can ultimately provide developers the necessary feedback for improving the products.

In terms of actually implementing these ideas, there is still work to be done. For example, the current MODIS LAI/fPAR (Leaf Area Index/fraction of Photosynthetically Active Radiation) algorithm has been improved continuously since the satellite's launch. The main improvements have been in the use of a better biome map (reducing

the uncertainty in that input); improvements in atmospheric correction; and improved models of surface reflectance from different biomes [2].

These improvements have reduced the uncertainty in the resulting data product, but have not necessarily improved the quantification of the uncertainties, and have not specifically addressed the statistical identification of the sources of the uncertainties. Here we will address one aspect of this overall process. Models of surface reflectance are typically Radiative Transfer Models (RTMs). We analyze in detail the effects of the inputs to an RTM in terms of the sensitivity of the RTM's output to each of the inputs. Specifically we analyze the Leaf Canopy Model (LCM) RTM [3], used as a surrogate for the RTM used as the basis for the MODIS production algorithm [4]. See section II for a discussion of the LCM. In section III, we use and develop methods from the statistical literature on sensitivity analysis [5] to compute the main effects, which graphically show the relative importance of each input on the RTM output, and the sensitivity indices, which give a measure of the expected amount by which the uncertainty in the output would be reduced if the true value of the input was known. See section III.

A 1999 paper, [6], discussed the state of sensitivity analysis in the remote sensing and geoscience domains. A that time the analyses were typically very basic, looking only at one variable at a time, and based around a fixed operating point. A number of suggestions as to better methods were suggested, principally the Fourier Amplitude Sensitivity Test [5]. This suggestion does not seem to have been adopted – the number of papers that cite [6] is small, and the number that adopt the suggestions, smaller still. For example, [7] uses ideas from the design of experiments, but does not compute sensitivity indices. While discussing sensitivity indices, the analysis in [8] is based on local sensitivity computations. In [9] sensitivity indices are computed, but the methods used required large numbers of model runs. In this paper we give explicit, computationally efficient methods for computing the main effects and sensitivity indices, as part of a global sensitivity analysis.

Computing the main effects and sensitivity indices requires the evaluation of multidimensional integrals over the input space of the model. Evaluating RTMs is typically computationally expensive, and so standard numerical integration methods (e.g. multidimensional quadrature or Monte Carlo integration) would be computationally prohibitive in terms of the number of times the RTM would have to be run. Instead, we adopt the approach of approximating the RTM by a Gaussian Process (GP) model [10], [11], [12]. A GP provides a very flexible nonparametric function approximation that has found wide application as a replacement for neural networks [10]. Early work involving GP response-surface approximations for the analysis of computer experiments includes [13], [14], [15]. We refer to [16] for background and further references. The GP model approximation can be constructed using a comparatively small number of carefully chosen RTM evaluations. See section IV. Using the GP approximation instead of the actual RTM will introduce uncertainty into the evaluation of the main effects, and the sensitivity indices, but this can also be quantified [17]. See section V.

Finally, in section VI we present the main effects and sensitivity indices for the LCM RTM, and show how they enable the identification of the relative importance of each input to the model output. This also gives information as to how well these inputs can be predicted from observations of the model output at different wavelengths.

II. A Coupled Leaf-Canopy Radiative Transfer Model

Over the past decade, in collaboration with the Ecosystem science and technology branch at NASA Ames, the Vegetation Modeling Transport Group (University of Arizona) has developed a coupled Leaf-Canopy Model (LCM) in order to capture the essential biophysical processes associated with the interaction between light and vegetation [3]. LCM was developed to provide a tool to aid in remote sensing as applied to ecosystem dynamics in support of the TERRA platform and it is specifically used to investigate the feasibility of observing chemistry remotely. The model combines two different radiative transfer models, one at leaf level (LEAFMOD) and one at canopy level (CANMOD) to predict the radiative regime inside the vegetation canopy under consideration.

LEAFMOD [18] is the model that simulates the radiative regime inside the single leaf. From a morphological point of view, the leaf element is an extremely complex and rich object. Any model that attempts to describe each single interaction process for the light moving in such a medium will face this enormous complexity. The strength of the LEAFMOD algorithm is its simplicity through natural averaging. The model relies on the fact that while light is moving in a complicated medium, natural averaging occurs in such way that the simpler assumption of isotropic scattering and uniform absorption seems to capture the transport effects. Moreover, the model has the ability to include chemistry as a key element dominating the absorption process. Different concentrations of chlorophyll, water, lignin and cellulose can be specified to model the optical properties of the single leaf species. The model is calibrated over the LOPEX leaf species archive [19], where experimental leaf property data are stored. The calibration occurs in the sense that the optical properties required by the canopy model are retrieved through a procedure that uses the LOPEX archive as input data.

The CANMOD (CANopy Model) algorithm [20], [3] takes the information coming from LEAFMOD regarding the single leaf characteristic (transmittance and reflectance) and together with canopy structural parameters, (LAI and Leaf Angle Distribution), soil reflectance and sun angle inclination, computes, at any given wavelength, the radiative regime within and at the top of the canopy by solving a radiative transfer equation. The strengths of the model are simplicity and the ability to take into account leaf chemistry, which is important to properly describe the light absorption environment.

Figure 1 shows a flowchart that demonstrates the operation of the coupled algorithm. The algorithm can be explained as follows. The first module uses LEAFMOD in the forward and inverse mode to compute the leaf optical properties (i.e. leaf reflectance and transmittance). The second module uses the CANMOD forward mode to compute the spectral canopy hemispherical reflectance factor. The code requires the specification of the input parameters. In addition to the parameters listed in table I, the model also takes as input wavelength (between 400nm and 2100nm), canopy architecture (LAD - leaf angle distribution, which takes one of 5 discrete values) and the sun angle.

Note that the soil reflectance depends on the wavelength. Indeed, usually the spectral soil reflectance is specified depending on the type of soil of interest but once the wavelength is set, the algorithm will work with the value of soil reflectance relative to the specified wavelength. We assumed a typical visible/near infra-red spectrum for a dry
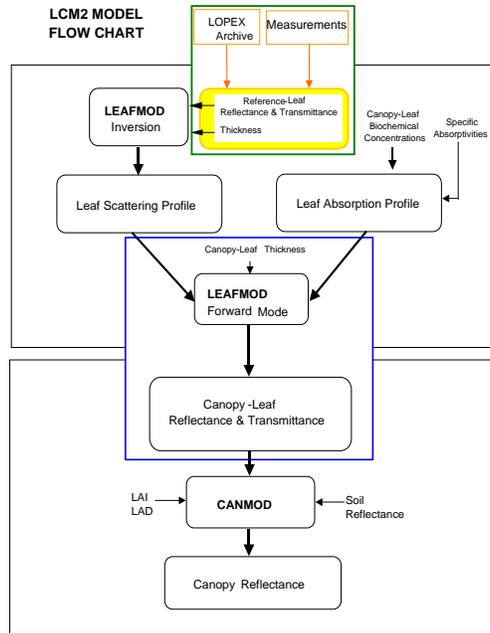
Fig. 1. LCM2 Flow Chart

input	min	max
LAI	0	8
chlorophyll $(\mu g/cm^2)$	0	100
water fraction	0.1	0.8
protein (g/cm^2)	0.0001	0.001
lignin/cellulose (g/cm^2)	0.0001	0.006
thickness (cm)	0.001	0.01
soil	0.3	1.3

TABLE I

RANGES OF VALUES OF THE INPUTS TO THE LCM. LAI AND WATER FRACTION ARE DIMENSIONLESS. THE SOIL PARAMETER IS

MULTIPLIED BY A STANDARD SOIL SPECTRUM. SEE TEXT.

soil, and we considered a multiplicative brightness parameter varying between 0.3-1.3 (see table I) to account for the possible variations of the background (soil) reflectance level [21]. This multiplicative parameter is assumed to be wavelength independent.

Once the leaf type is specified, the LOPEX database contains the measured leaf optical properties for the leaf of interest. Nevertheless, we can tune the canopy by considering leaves that are of the same type but with different biochemistry and thickness. This gives the code great flexibility in modeling the effect of biochemistry on the overall canopy reflectance. The algorithm begins by analyzing the leaf under consideration. Assume, for example

that the canopy of interest is a maple canopy. The LOPEX database is accessed to retrieve the measured spectral reflectance and transmittance for a nominal maple leaf. Note that, as before, since the wavelength is set, reflectance and transmittance for the nominal leaf are selected for the specific wavelength of interest. The LEAFMOD inverse mode accepts the reflectance and transmittance and retrieves scattering and absorption coefficients. It is assumed that, to first order, the scattering depends on the anatomical structure of the leaf, while the absorption depends only on the biochemical components [18]. Thus, the scattering coefficient for maple leaves is assumed to be the same and it is retained. A new maple leaf having the biochemical components and thickness specified by the inputs is constructed retaining the same scattering coefficient and constructing the new absorption coefficient for the wavelength of interest. Both absorption and scattering coefficient are fed to the LEAFMOD forward mode to compute the reflectance and transmittance of the desired leaf, i.e. the leaf with thickness, water, chlorophyll, lignin and protein specified by the inputs. Reflectance and transmittance are fed to the second module together with LAI, LAD, soil reflectance and sun angle to compute the hemispherical reflectance.

III. SENSITIVITY ANALYSIS

Sensitivity analysis aims to determine how the variation in the output of a model can be apportioned amongst the inputs [16, ch 7]. That is, it attempts to determine how much of the variation seen in the output is due to variation in each of the inputs. The type of sensitivity analysis we are interested in here is *global* sensitivity analysis, looking at how the output changes as all the inputs vary continuously, rather than the more common *local* sensitivity analyzes, which look at how the output changes as the inputs are each varied about a fixed point. Clearly this latter type of analysis will give limited information about how the output varies for substantial changes in the inputs.

How the inputs vary is determined by a probability distribution that defines the expected distributions of the inputs. Using v to denote the vector of model inputs, this distribution is $H(v)$. The actual form of this distribution is problem dependent, and dependent on the amount of knowledge available about each input variable. It may be that for some inputs all that can be given is a physically plausible range (e.g. water fraction is limited to the range 0-1), whereas for others a more precise distribution may be known (e.g. the distribution of leaf thickness for a particular tree type may be known from field measurements). The distribution $H(v)$ also encodes correlations between variables that are known to vary together. The authors in [21] give truncated Gaussian distributions for the variables in table I. In this work we use the simpler formulation of independent uniform distributions over the ranges given in table I for each input variable.

A. Main Effects

Denote the response of the model to input v as $y = f(v)$. The function $f(v)$ can be decomposed as

$$y = f(v) \quad = \quad \mathrm{E}(Y) + \sum_{i=1}^{d} z_i(v_i) + \sum_{i<j} z_{i,j}(v_i, v_j) + \ldots$$
$$+ z_{1,2,\ldots,d}(v_1, v_2, \ldots, v_d) \tag{1}$$

where $\boldsymbol{v} = (v_1, ..., v_d)$ is $d-$dimensional (with $d = 7$ in our sensitivity analysis of the LCM). The first term is the expected value of $f(\boldsymbol{v})$, i.e.,

$$\mathrm{E}(Y) = \int_{v_j, j=1...d} f(\boldsymbol{v}) \mathrm{d}H(\boldsymbol{v})$$

and the next d terms are the *main effects*, given by

$$\begin{aligned} z_i(v_i) &= \mathrm{E}(Y|v_i) - \mathrm{E}(Y) \\ &= \int_{\boldsymbol{v}_{-i}} f(\boldsymbol{v}) \mathrm{d}H(\boldsymbol{v}_{-i}|v_i) - \mathrm{E}(Y) \end{aligned} \tag{2}$$

where $\boldsymbol{v}_{-i}$ denotes all the elements of $\boldsymbol{v}$ except v_i. The later terms of the decomposition are the interactions. They give information about the combined influence of two or more inputs taken together. We will not consider them further here.

Plotting the main effects, $z_i(v_i)$ for each i gives a visual impression of the relative importance of each input to the variation in the output. This visual impression is heightened if the inputs are normalized (to the range 0-1, for example, for uniformly distributed inputs), allowing all the main effects to be plotted together on the same plot. See section VI where we present main effects plots for the LCM RTM.

To compute the main effects requires the evaluation of a $(d - 1)-$dimensional integral. For even moderately complex functions $f(\boldsymbol{v})$ it will be impossible to evaluate this integral analytically; indeed, for most cases of interest an analytic form for $f(\boldsymbol{v})$ does not exist, rather, $f(\boldsymbol{v})$ only exists as a computer program. In these cases the $z_i(v_i)$ must be computed numerically. If evaluating $f(\boldsymbol{v})$ for a given $\boldsymbol{v}$ requires appreciable computation then the standard methods of numerical integration, multidimensional quadrature and Monte Carlo integration, will be too computationally intensive to be practical. In these cases we can approximate $f(\boldsymbol{v})$ and compute the main effects of the approximation, and also compute the uncertainty introduced by the approximation to $f(\boldsymbol{v})$. This is given in sections IV and V. Details of the Gaussian Process approximation we use for $f(\boldsymbol{v})$ is given in section IV, and its application to computing the main effects and sensitivity indices in section V.

B. Sensitivity Indices

The sensitivity indices are based on the variances of the terms in the decomposition of $f(\boldsymbol{v})$ given in equation 1. Consider

$$V_i = \mathrm{Var}\{\mathrm{E}(Y|v_i)\}.$$

This is the expected amount by which the uncertainty in y will be reduced if we learn the true value of v_i [17]. It thus gives a measure of how much of the variance of y is due to input v_i. The V_i's can be normalized to

$$S_i = V_i/\mathrm{Var}(Y)$$

and the sum of all the S_i's and higher-order terms ($S_{i,j}$, $S_{i,j,k}$, etc.) is unity. Thus the value of S_i gives the relative importance of input x_i. The S_i's can also be used to direct improvements – reducing the uncertainty on the input with the largest S_i will have the greatest effect in reducing the uncertainty of the model output. This can be used to direct data collection work.

Computing the S_i's and V_i's can be complex, even under the Gaussian Process approximation to $f(\boldsymbol{v})$. See section V for details.

IV. APPROXIMATING THE LCM USING A GAUSSIAN PROCESS

Gaussian Processes (GPs) are probability distributions over *functions*. Rather than placing a distribution over a (small) set of parameters, a GP places a distribution directly over the function of interest. Under a GP probability model for function $f(\cdot)$, the joint distribution of $(f(\boldsymbol{v}_1), ..., f(\boldsymbol{v}_k))$ is multivariate Gaussian, for any finite set of input points $\boldsymbol{v}_1, ..., \boldsymbol{v}_k$. It is this property that allows for tractable computation – whilst the GP is defined over an infinite dimensional quantity (the continuous function, $f(\boldsymbol{v})$), any computation is necessarily done over only a finite set of locations.

A GP is specified by its mean function, $\mathrm{E}(f(\boldsymbol{v}))$, and its covariance function $\mathrm{Cov}(f(\boldsymbol{v}), f(\boldsymbol{v}'))$. The flexibility of choosing and adapting the mean and covariance functions allows a GP model to be successfully used to approximate a wide spectrum of functions $f(\boldsymbol{v})$, based on a set of training examples, $\boldsymbol{d} = \{\boldsymbol{y}, \boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$, where $\boldsymbol{y} = (y_1, ..., y_n)$ and y_i is the response $f(\boldsymbol{x}_i)$ at observed input point $\boldsymbol{x}_i$, $i = 1, ..., n$. The set of training examples is chosen carefully to optimally sample the input space. Here we used a Latin Hypercube design [22] to choose the set of inputs to the LCM. The other choices made were to use a constant mean function, $\mathrm{E}(f(\boldsymbol{v})) = \mu$, a constant variance $\mathrm{Var}(f(\boldsymbol{v})) = \sigma^2$, and the product Gaussian correlation function

$$\mathrm{Corr}(f(\boldsymbol{v}), f(\boldsymbol{v}'); \boldsymbol{\theta}) = \exp\left(-\sum_{\ell=1}^{d} \frac{(v_\ell - v'_\ell)^2}{\gamma_\ell}\right)$$

where $\boldsymbol{\theta} = (\gamma_1, ..., \gamma_d)$, and d is the number of dimensions in the inputs, $\boldsymbol{v}$. The γ parameters give a measure of the scale over which the function $f(\boldsymbol{v})$ varies in each input dimension, and σ^2, the variance of the GP, determines the overall scale of $f(\boldsymbol{v})$. Using this mean and correlation function, the GP defines the joint distribution

$$p(\boldsymbol{y}|\boldsymbol{\theta}, \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}|C|^{1/2}} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{y} - \mu\boldsymbol{1}_n)^T C^{-1}(\boldsymbol{y} - \mu\boldsymbol{1}_n)\right)$$

where C is the correlation matrix with (i, j)-th element $\mathrm{Corr}(f(\boldsymbol{x}_i), f(\boldsymbol{x}_j))$, and $\boldsymbol{1}_n$ denotes an n-dimensional vector with all elements equal to 1.

We use the set of training examples, $\boldsymbol{d}$, to estimate the parameters $\{\boldsymbol{\theta}, \mu, \sigma^2\}$ of the GP model using maximum likelihood estimation. From equation 3 the log-likelihood is

$$\mathcal{L} = -\frac{1}{2\sigma^2}(\boldsymbol{y} - \mu\boldsymbol{1}_n)^T C^{-1}(\boldsymbol{y} - \mu\boldsymbol{1}_n) - \frac{1}{2}\log|C| - \frac{n}{2}\log(2\pi\sigma^2) \tag{3}$$

where C depends on the parameters $\boldsymbol{\theta}$. The derivatives of $\mathcal{L}$ with respect to each of the parameters can be straightforwardly derived [11]. Maximizing $\mathcal{L}$ results in a point estimate for the parameters, denoted by $\{\hat{\boldsymbol{\theta}}, \hat{\mu}, \hat{\sigma^2}\}$, that we use when evaluating the main effects. Note that using point estimates for these parameters will cause the uncertainty of the main effects to be underestimated. In future work we will consider a fully inferential Bayesian approach where expectations are also taken with respect to these parameters.

Once the GP model parameters are estimated, the first quantities of interest are the predictive distributions for sets of new inputs, conditioned on the training examples. From the definition of the GP, these distributions will be Gaussian. For a single new input v the predictive distribution for $f(v)$ has mean

$$m \equiv m(v; \hat{\mu}, \hat{\theta}, d) = \hat{\mu} + r^T(v)C^{-1}(y - \hat{\mu}\mathbf{1}_n)$$

and variance

$$S \equiv S(v; \hat{\mu}, \hat{\sigma}^2, \hat{\theta}, d) = \hat{\sigma}^2 \left(1 - r^T(v)C^{-1}r(v)\right).$$

Here $r(v)$ is the $n \times 1$ vector with i-th element given by $\mathrm{Corr}(f(v), f(x_i)) = \exp(-\sum_{\ell=1}^{d}(v_\ell - x_{i\ell})^2/\hat{\gamma}_\ell)$, and C is the observed $n \times n$ correlation matrix with (i, j)-th element given by $\exp(-\sum_{\ell=1}^{d}(x_{i\ell} - x_{j\ell})^2/\hat{\gamma}_\ell)$. Recall that the x_i are the input values of the training examples.

The joint predictive distribution for $(f(v), f(v'))$ corresponding to generic inputs $v = (v_1, ..., v_d)$ and $v' = (v'_1, ..., v'_d)$ is bivariate normal with (2×1) mean vector

$$w = \hat{\mu}\mathbf{1}_2 + R^T(v, v')C^{-1}(y - \hat{\mu}\mathbf{1}_n) \tag{4}$$

and (2×2) covariance matrix

$$W = \hat{\sigma}^2 \left(B(v, v') - R^T(v, v')C^{-1}R(v, v')\right), \tag{5}$$

where $B(v, v')$ is the (2×2) correlation matrix for $(f(v), f(v'))$ with off-diagonal element given by $\exp(-\sum_{\ell=1}^{d}(v_\ell - v'_\ell)^2/\hat{\gamma}_\ell)$, and $R(v, v')$ is the $(n \times 2)$ matrix with first row elements $\exp(-\sum_{\ell=1}^{d}(v_\ell - x_{i\ell})^2/\hat{\gamma}_\ell)$, $i = 1, ..., n$, and analogously for the second row elements replacing v_ℓ with v'_ℓ.

V. APPROXIMATING THE MAIN EFFECTS AND SENSITIVITY INDICES USING THE GAUSSIAN PROCESS APPROXIMATION TO THE LCM

To compute the main effects requires evaluating $\mathrm{E}(Y \mid v_j)$ and $\mathrm{E}(Y)$, as indicated in equation 2. However, we recall that we are approximating the function $y = f(v)$ by a GP model, and we must account for this approximation by computing $\mathrm{E}^*\{\mathrm{E}(Y \mid v_j)\}$ and $\mathrm{E}^*\{\mathrm{E}(Y)\}$, where we use $\mathrm{E}^*\{\}$, $\mathrm{Var}^*\{\}$ and $\mathrm{Cov}^*\{\}$ to indicate expectation, variance and covariance, respectively, with respect to the GP predictive distributions. We give details of these quantities here.

For the global mean, we have

$$\mathrm{E}(Y) = \int_v f(v) \prod_{\ell=1}^{d} \mathrm{d}H_\ell(v_\ell)$$

where $H(v) = \prod_{\ell=1}^{d} H_\ell(v_\ell)$ is the input distribution, comprising independent components $H_\ell(v_\ell)$, which are

uniform distributions over ranges (a_ℓ, b_ℓ), $\ell = 1, ..., d$. Therefore,

$$
\begin{aligned}
\mathrm{E}^* \{\mathrm{E}(Y)\} &= \int \mathrm{E}(Y) \mathrm{d}N(f(\boldsymbol{v}); m, S) \\
&= \int_{\boldsymbol{v}} m(\boldsymbol{v}) \prod_{\ell=1}^{d} \mathrm{d}H_\ell(v_\ell) \\
&= \int_{\boldsymbol{v}} \{\hat{\mu} + \boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y} - \hat{\mu}\mathbf{1}_n)\} \prod_{\ell=1}^{d} \mathrm{d}H_\ell(v_\ell) \\
&= \hat{\mu} + \boldsymbol{T}^T C^{-1}(\boldsymbol{y} - \hat{\mu}\mathbf{1}_n),
\end{aligned}
\tag{6}
$$

where $\boldsymbol{T}$ is the $n \times 1$ vector with i-th element given by $\prod_{\ell=1}^{d} \left\{ \int_{a_\ell}^{b_\ell} \exp(-(v_\ell - x_{i\ell})^2/\hat{\gamma}_\ell)(b_\ell - a_\ell)^{-1}\mathrm{d}v_\ell \right\}$.

For $\mathrm{E}(Y \mid v_j)$, for each value u_j of the $j-$th input, we have

$$
\mathrm{E}(Y \mid u_j) = \int_{\{v_\ell : \ell \neq j\}} f(v_1, ..., u_j, ..., v_d) \prod_{\{\ell : \ell \neq j\}} \mathrm{d}H_\ell(v_\ell)
$$

and thus

$$
\begin{aligned}
\mathrm{E}^* \{\mathrm{E}(Y \mid u_j)\} &= \int \mathrm{E}(Y \mid u_j) \mathrm{d}N(f(v_1, ..., u_j, ..., v_d); m, S) \\
&= \int_{\{v_\ell : \ell \neq j\}} m(v_1, ..., u_j, ..., v_d) \prod_{\{\ell : \ell \neq j\}} \mathrm{d}H_\ell(v_\ell) \\
&= \hat{\mu} + \boldsymbol{T}_j^T(u_j)C^{-1}(\boldsymbol{y} - \hat{\mu}\mathbf{1}_n),
\end{aligned}
$$

where $\boldsymbol{T}_j(u_j)$ is the $(n \times 1)$ vector with i-th element given by the following expression

$$
\exp\left(-\frac{(u_j - x_{ij})^2}{\hat{\gamma}_j}\right) \prod_{\{\ell : \ell \neq j\}} \left\{ \int_{a_\ell}^{b_\ell} \exp\left(-\frac{(v_\ell - x_{i\ell})^2}{\hat{\gamma}_\ell}\right) \frac{1}{b_\ell - a_\ell}\mathrm{d}v_\ell \right\}.
\tag{7}
$$

The previous expressions provide point estimates for all main effects associated with the d inputs. In particular, for each input $j = 1, ..., d$, $\mathrm{E}^* \{\mathrm{E}(Y \mid u_j)\}$ can be computed over a grid of u_j values to obtain point estimates for the functions $\mathrm{E}(Y \mid u_j)$ (or for $\mathrm{E}(Y \mid u_j) - \mathrm{E}(Y)$ using also $\mathrm{E}^* \{\mathrm{E}(Y)\}$). These estimates can be compared graphically (linear transformations can be applied so that all inputs are on the same scale).

For a measure of the uncertainty associated with these estimates, we use

$$
\mathrm{Var}^* \{\mathrm{E}(Y \mid u_j)\} = \mathrm{E}^* \left\{ (\mathrm{E}(Y \mid u_j))^2 \right\} - (\mathrm{E}^* \{\mathrm{E}(Y \mid u_j)\})^2.
$$

Because we already have the expression for $\mathrm{E}^* \{\mathrm{E}(Y \mid u_j)\}$ through (7) and (7), what is needed is an expression for $\mathrm{E}^* \left\{ (\mathrm{E}(Y \mid u_j))^2 \right\}$. This derivation is given in Appendix I, resulting in

$$
\mathrm{Var}^* \{\mathrm{E}(Y \mid u_j)\} = \hat{\sigma}^2 \left(e - \boldsymbol{T}_j^T(u_j)C^{-1}\boldsymbol{T}_j(u_j) \right)
$$

where e is given by 11 in Appendix I.

The sensitivity indices are defined by

$$
S_j = \frac{\mathrm{Var}(\mathrm{E}(Y \mid u_j))}{\mathrm{Var}(Y)}, \quad j = 1, ..., d.
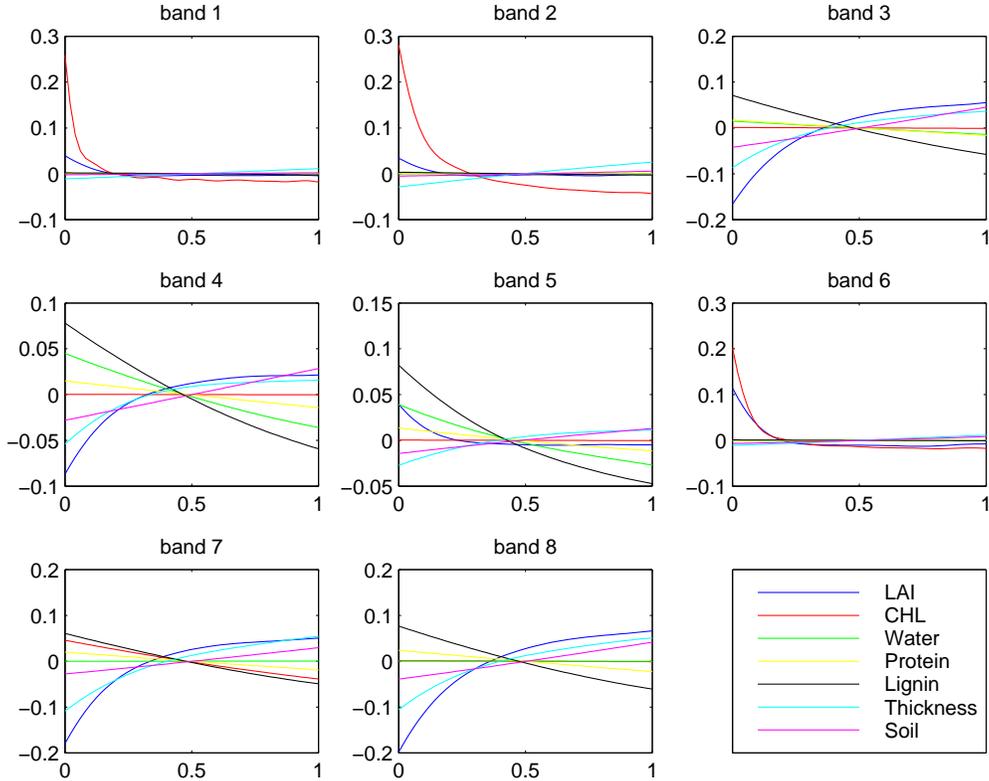$$

Fig. 2. The main effects for the LCM RTM

Computing $E^*\{S_j\}$ cannot be done analytically, even under the GP approximation, so we approximate it by computing the ratio of $E^*\{Var(E(Y \mid u_j))\}$ and $E^*\{Var(Y)\}$. (In future work we will use a Bayesian approach implemented via Markov chain Monte Carlo (MCMC) methods [23] to estimate the entire distribution of S_j under the GP approximation, allowing the uncertainty of the sensitivity indices to also be determined.) We have that

$$E^*\{Var(E(Y \mid u_j))\} = E^* \left\{ E \left[(E(Y \mid u_j))^2 \right] \right\} - E^*\{(E(Y))^2\}$$

and

$$E^*\{Var(Y)\} = E^*\{E(Y^2)\} - E^*\{(E(Y))^2\}.$$

The expressions for these terms are not difficult to derive, though care is needed. They are given in Appendix II.

VI. RESULTS

The proposed methodology has been applied to execute a global sensitivity analysis and to analyze both the sensitivity of the spectral hemispherical reflectance to the defined input parameters and the relative contribution of each parameters to the model output

To generate the training data for the GP model we generated a 250 point Latin Hypercube design over the 7-dimensional space of inputs given in table I. The Leaf Angle Distribution (LAD) variable was set to planophile
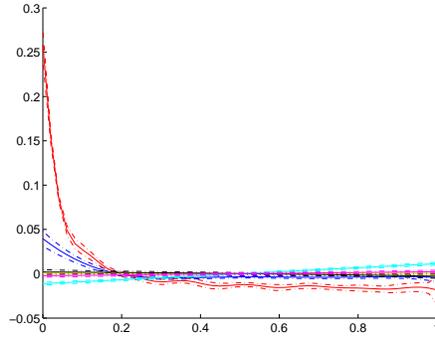
Fig. 3. The uncertainty in the main effects due to using the GP approximation to the LCM RTM. Band 1. Line colours as in figure 2.
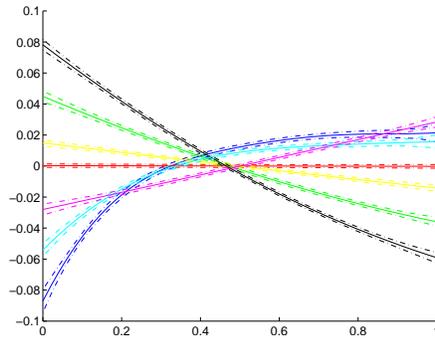

Fig. 4. The uncertainty in the main effects due to using the GP approximation to the LCM RTM. Band 4. Line colours as in figure 2.

(leaves mostly horizontal) and the sun angle was set to zenith. While the sun angle will vary, for any given satellite scene it will be known, and so we do not consider it as one of the inputs for this analysis. The LCM was run at 8 wavelengths, given in table II, corresponding to eight of the MODIS bands that are sensitive to vegetation. The corresponding MODIS band is also given in table II. Note that the bands are in MODIS band order, not in wavelength order.

Figure 2 shows plots of the main effects for the 7 input variables for each of the 8 bands. The larger the variation of the main effect plot, the greater the influence of that input on the LCM response. To display the main effects for all parameters on a single plot, the range of each input (given in table I has been normalized to 0-1. The slope of each main effect plot gives information as to whether the output is an increasing or decreasing function of that input. The relative scale of the main effects can be easily compared visually. The absolute scale depends on the absolute magnitude of the model output.

To correctly interpret the results and to put them in the right perspective, we divide the input parameters in two categories, i.e. spectrally and radiative transfer (RT) driven. Biochemical inputs, i.e. chlorophyll, water, lignin and protein are spectrally driven since their effect heavily depends on wavelength, and mainly affect the absorption characteristic of vegetation canopies [18]. Conversely, LAI, leaf thickness and soil brightness can be categorized

band number	wavelength (nm)	MODIS band
1	469	ref3
2	555	ref4
3	1240	ref5
4	1640	ref6
5	2130	ref7
6	667	ref13
7	748	ref15
8	870	ref16

TABLE II

THE WAVELENGTH FOR EACH BAND USED, AND THE CORRESPONDING MODIS BAND NUMBER.

as RT driven since they directly influence the transport of photons in the medium. The main effects and sensitivity indices are analyzed next.

The LCM is most sensitive to LAI in the near-infrared (NIR) region of the spectrum (bands 7 and 8). It is shown in the figure that the LAI effect is highly non-linear and the behavior is such that in bands 7, 8 and 3 an increase in LAI produces an increase in reflectance. Conversely, the effect is opposite in the visible, i.e. increases in LAI produces a decrease in the hemispherical reflectance. This trend is known. [24], [25]. The sensitivity indices (table III) show that LAI is the major contributor in the bands which are most sensitive.

Chlorophyll, on the other end, is expected to be extremely influential in the visible, i.e. it is the prevailing factor that dominates the reflectance. Its effect is strong in the visible (bands 1, 2 and 6) while it dramatically decrease at the red-edge (band 7) to eventually disappear in the rest of the spectrum. Indeed, chlorophyll does not absorb light after 760nm. As shown in figure 4, band 4 show basically no sensitivity for chlorophyll with small quantified uncertainty in the result.

Water contribution occurs mainly in the short-wave infrared where it exhibits resonant phenomenon that increase the probability of light absorption around the 1445nm wavelength. Indeed, water is ranked as second and third major contributor to the reflectance in band 5 and 4 respectively. Conversely the reflectance is weakly sensitive in the visible.

Protein is shown to be insensitive to most of the spectrum. Its effect as well as contribution are extremely small and can be only detected with difficulties in the NIR (e.g. band 8).

Lignin is one of the major surprising results. It is extremely sensitive in the short-wave infrared (bands 4 and 5) where it is also the major contributor to the hemispherical reflectance. This is mainly due to the strong absorption features in this part of the spectrum where lignin absorption coefficient features a peak around 2110nm.

Leaf thickness demonstrates a true RT effect and its response shows interesting features. It is mainly sensitive and has the major contribution in bands 3 and 8. We believe that what we are seeing is that changing the leaf thickness has more influence on scattering than on absorption. Specifically, as we change the leaf thickness the model assumes

	band; wavelength (nm)							
	1	2	3	4	5	6	7	8
input	469	555	1240	1640	2130	667	748	870
LAI	0.05	0.01	0.43	0.16	0.04	0.28	0.41	0.48
CHL	0.80	0.83	0.00	0.00	0.00	0.56	0.08	0.00
Water	0.00	0.00	0.01	0.12	0.14	0.00	0.00	0.00
Protein	0.00	0.00	0.01	0.02	0.02	0.00	0.02	0.02
Lignin	0.00	0.00	0.19	0.36	0.53	0.00	0.13	0.16
Thickness	0.02	0.05	0.14	0.07	0.05	0.02	0.24	0.18
Soil	0.00	0.00	0.08	0.06	0.03	0.01	0.03	0.06
Total	0.88	0.90	0.86	0.80	0.81	0.87	0.90	0.90

TABLE III

THE SENSITIVITY INDICES FOR EACH INPUT FOR EACH SPECTRAL BAND.

that the leaf mass is unchanged, meaning that the absorption has little effect as can be seen specifically in the NIR part of the spectrum.

The soil brightness has generally little effect. The spectrum for a typical soil was spectrally defined and the brightness parameter is responsible for increasing the soil hemispherical reflectance therefore simulating the dry-wet effect.

That the sensitivity indices do not sum to one indicates that interaction effects between two or more inputs are important in some bands, particularly band 4. In future work we will compute the second order sensitivity indices that quantify which interactions are important.

Our results are consistent with, and extend, previous statistically-based work. For example, [7] presented a methodology for sensitivity analysis based on design of numerical experiments aimed at providing a comparison between four canopy RT models coupled with a leaf-based RT model (PROSPECT, [26]). Their results are consistent with ours regarding LAI, chlorophyll and soil brightness sensitivity behavior. That the response in bands 1, 2 and 6 is dominated by LAI and chlorophyll. is consistent with the results of a much more restricted sensitivity analysis in [27].

These results show that analyzing the uncertainty characteristics of RTMs used in remote sensed data product generation is practical and important. It gives information on the level of accuracy needed in the model's inputs, can guide data collection efforts to most effectively reduce the uncertainties, and can guide further development effort for the RTMs themselves. It also gives information as to which of the model's inputs affect the output, and hence which inputs it may be possible to determine from remotely-sensed observations.

APPENDIX I

VARIANCE OF THE MAIN EFFECTS

We give here the derivation of $E^* \left\{ (E(Y \mid u_j))^2 \right\}$ required in the expression for $\text{Var}^* \left\{ E(Y \mid u_j) \right\}$, the uncertainty associated with estimates of the main effects

Note that,

$$
(E(Y \mid u_j))^2 = \left(\int\limits_{\{v_\ell : \ell \neq j\}} f(v_1, ..., u_j, ..., v_d) \prod_{\{\ell : \ell \neq j\}} dH_\ell(v_\ell) \right)^2
$$

$$
= \iint\limits_{\substack{\{v_\ell : \ell \neq j\} \\ \{v'_\ell : \ell \neq j\}}} f(v_1, ..., u_j, ..., v_d) f(v'_1, ..., u_j, ..., v'_d) \prod_{\{\ell : \ell \neq j\}} dH_\ell(v_\ell) \prod_{\{\ell : \ell \neq j\}} dH_\ell(v'_\ell)
$$

and thus we need to take $E^* \{ \cdot \}$ with respect to the bivariate predictive distribution for $(f(v_1, ..., u_j, ..., v_d), f(v'_1, ..., u_j, ..., v'_d))$. Specifically,

$$
E^* \left\{ (E(Y \mid u_j))^2 \right\} =
$$

$$
\iint\limits_{\substack{\{v_\ell : \ell \neq j\} \\ \{v'_\ell : \ell \neq j\}}} E^* \left\{ f(v_1, ..., u_j, ..., v_d) f(v'_1, ..., u_j, ..., v'_d) \right\} \times \prod_{\{\ell : \ell \neq j\}} dH_\ell(v_\ell) \prod_{\{\ell : \ell \neq j\}} dH_\ell(v'_\ell), \quad (8)
$$

where

$$
E^* \left\{ f(v_1, ..., u_j, ..., v_d) f(v'_1, ..., u_j, ..., v'_d) \right\} =
$$

$$
\text{Cov}^* \left\{ f(v_1, ..., u_j, ..., v_d), f(v'_1, ..., u_j, ..., v'_d) \right\} + \left(E^* \left\{ f(v_1, ..., u_j, ..., v_d) \right\} E^* \left\{ f(v'_1, ..., u_j, ..., v'_d) \right\} \right). \quad (9)
$$

Denote by $\boldsymbol{R}_1^T \equiv \boldsymbol{R}_1^T(v_1, ..., u_j, ..., v_d)$ and $\boldsymbol{R}_2^T \equiv \boldsymbol{R}_2^T(v'_1, ..., u_j, ..., v'_d)$ the first and second rows, respectively, of the $(n \times 2)$ matrix $R(\boldsymbol{v}, \boldsymbol{v}')$ defined in section IV. Note that here the input vectors we are working with, $(v_1, ..., u_j, ..., v_d)$ and $(v'_1, ..., u_j, ..., v'_d)$, have common element u_j. Therefore, $\boldsymbol{R}_1$ is the $n \times 1$ vector with elements

$$
\exp \left(-\frac{(u_j - x_{ij})^2}{\hat{\gamma}_j} - \sum_{\{\ell : \ell \neq j\}} \frac{(v_\ell - x_{i\ell})^2}{\hat{\gamma}_\ell} \right), \quad i = 1, ..., n,
$$

and analogously for $\boldsymbol{R}_2$, replacing v_ℓ with v'_ℓ. Then, using (4) and (5), we obtain

$$
E^* \left\{ f(v_1, ..., u_j, ..., v_d) \right\} = \hat{\mu} + \boldsymbol{R}_1^T (C^{-1}(\boldsymbol{y} - \hat{\mu} \boldsymbol{1}_n)
$$

$$
E^* \left\{ f(v'_1, ..., u_j, ..., v'_d) \right\} = \hat{\mu} + \boldsymbol{R}_2^T C^{-1}(\boldsymbol{y} - \hat{\mu} \boldsymbol{1}_n)
$$

$$
\text{Cov}^* \left\{ f(v_1, ..., u_j, ..., v_d), f(v'_1, ..., u_j, ..., v'_d) \right\} = \hat{\sigma}^2 \left\{ \exp \left(- \sum_{\{\ell : \ell \neq j\}} \frac{(v_\ell - v'_\ell)^2}{\hat{\gamma}_\ell} \right) - \boldsymbol{R}_1^T C^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{R}_2 \right\}. \quad (10)
$$

Finally, substituting (9) and (10) in (8), we obtain for each $j = 1, ..., d$,

$$
E^* \left\{ (E(Y \mid u_j))^2 \right\} = \hat{\sigma}^2 \left(e - \boldsymbol{T}_j^T(u_j) C^{-1} \boldsymbol{T}_j(u_j) \right) + \left(\hat{\mu} + \boldsymbol{T}_j^T(u_j) C^{-1}(\boldsymbol{y} - \hat{\mu} \boldsymbol{1}_n) \right)^2,
$$

where $\boldsymbol{T}_j(u_j)$ is the $n \times 1$ vector with elements given in (7), and

$$e = \prod_{\{\ell:\ell\neq j\}} \left\{ \int_{a_\ell}^{b_\ell} \int_{a_\ell}^{b_\ell} \exp\left(-\frac{(v_\ell - v'_\ell)^2}{\hat{\gamma}_\ell}\right) \frac{\mathrm{d}v_\ell \mathrm{d}v'_\ell}{(b_\ell - a_\ell)^2} \right\}. \tag{11}$$

Note that the second term in equation 11 is $(\mathrm{E}^*\{E(Y \mid u_j)\})^2$, and so the required variance has the simpler expression

$$\mathrm{Var}^*\{\mathrm{E}(Y \mid u_j)\} = \hat{\sigma}^2 \left(e - \boldsymbol{T}_j^T(u_j)C^{-1}\boldsymbol{T}_j(u_j)\right)$$

APPENDIX II

THE SENSITIVITY INDICES

To compute the sensitivity indices, we must compute $\mathrm{E}^*\{\mathrm{Var}(\mathrm{E}(Y \mid u_j))\}$ and $\mathrm{E}^*\{\mathrm{Var}(Y)\}$. Regarding the unconditional variance

$$\mathrm{E}^*\{\mathrm{Var}(Y)\} = \mathrm{E}^*\{\mathrm{E}(Y^2)\} - \mathrm{E}^*\{(\mathrm{E}(Y))^2\}.$$

For the first term, we have

$$\begin{aligned}
\mathrm{E}^*\{\mathrm{E}(Y^2)\} &= \mathrm{E}^*\left\{\int_{\boldsymbol{v}} f^2(\boldsymbol{v}) \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\right\} \\
&= \int_{\boldsymbol{v}} \mathrm{E}^*\{f^2(\boldsymbol{v})\} \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell) \\
&= \int_{\boldsymbol{v}} (S + m^2) \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell) \\
&= \hat{\sigma}^2 \int_{\boldsymbol{v}} \left(1 - \boldsymbol{r}^T(\boldsymbol{v})C^{-1}\boldsymbol{r}(\boldsymbol{v})\right) \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell) \\
&\quad + \int_{\boldsymbol{v}} \left\{\hat{\mu}^2 + 2\hat{\mu}\boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n) + (\boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n))^2\right\} \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell) \\
&= \hat{\sigma}^2 - \hat{\sigma}^2 \left(\int_{\boldsymbol{v}} \boldsymbol{r}^T(\boldsymbol{v})C^{-1}\boldsymbol{r}(\boldsymbol{v}) \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\right) + \hat{\mu}^2 + 2\hat{\mu}\boldsymbol{T}^T C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n) \\
&\quad + \left(\int_{\boldsymbol{v}} (\boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n))^2 \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\right)
\end{aligned}$$

where $\boldsymbol{T}$ is the $n \times 1$ vector defined earlier in the expression for $\mathrm{E}^*\{\mathrm{E}(Y)\}$ after equation 6. Regarding the two integrals above, if we expand the quadratic form $\boldsymbol{r}^T(\boldsymbol{v})C^{-1}\boldsymbol{r}(\boldsymbol{v})$ and apply the integral, we obtain

$$\int_{\boldsymbol{v}} \boldsymbol{r}^T(\boldsymbol{v})C^{-1}\boldsymbol{r}(\boldsymbol{v}) \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} q_{ij}$$

where c_{ij} is the (i, j)-th element of matrix C^{-1}, and

$$q_{ij} = \prod_{\ell=1}^d \left\{ \int_{a_\ell}^{b_\ell} \exp\left(-\frac{(v_\ell - x_{i\ell})^2 + (v_\ell - x_{j\ell})^2}{\hat{\gamma}_\ell}\right) \frac{1}{b_\ell - a_\ell} \mathrm{d}v_\ell \right\}, \quad i, j = 1, ..., n.$$

(The q_{ij} are symmetric in (i, j).) Analogously, expanding the square $(\boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n))^2$ and taking the integral, we get

$$\int_{\boldsymbol{v}} (\boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n))^2 \prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell) = \sum_{i=1}^n z_i^2 q_{ii} + 2\sum_{i=1}^n \sum_{j=i+1}^n z_i z_j q_{ij}$$

where z_i denotes the i-th element of the $n \times 1$ vector $C^{-1}(\boldsymbol{y} - \hat{\mu}\boldsymbol{1}_n)$.

For the second term, we can write

$$
\begin{aligned}
\mathrm{E}^*\{(\mathrm{E}(Y))^2\} &= \mathrm{E}^*\left\{\left(\int_{\boldsymbol{v}} f(\boldsymbol{v})\prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\right)^2\right\} \\
&= \mathrm{E}^*\left\{\int_{\boldsymbol{v}}\int_{\boldsymbol{v}'} f(\boldsymbol{v})f(\boldsymbol{v}')\prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\prod_{\ell=1}^d \mathrm{d}H_\ell(v'_\ell)\right\} \\
&= \int_{\boldsymbol{v}}\int_{\boldsymbol{v}'}\mathrm{E}^*\{f(\boldsymbol{v})f(\boldsymbol{v}')\}\prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\prod_{\ell=1}^d \mathrm{d}H_\ell(v'_\ell)
\end{aligned}
$$

and thus $\mathrm{E}^*\{(\mathrm{E}(Y))^2\}$ can be expressed as

$$
\int_{\boldsymbol{v}}\int_{\boldsymbol{v}'}\mathrm{Cov}^*\{f(\boldsymbol{v}),f(\boldsymbol{v}')\}\prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\prod_{\ell=1}^d \mathrm{d}H_\ell(v'_\ell) + \int_{\boldsymbol{v}}\int_{\boldsymbol{v}'}\{\mathrm{E}^*(f(\boldsymbol{v}))\mathrm{E}^*(f(\boldsymbol{v}'))\}\prod_{\ell=1}^d \mathrm{d}H_\ell(v_\ell)\prod_{\ell=1}^d \mathrm{d}H_\ell(v'_\ell). \quad (12)
$$

Let $\boldsymbol{r}'(\boldsymbol{v}')$ denote the $n\times 1$ vector with i-th element given by $\exp(-\sum_{\ell=1}^d (v'_\ell - x_{i\ell})^2/\hat{\gamma}_\ell)$. Then, analogously to the expressions in (10), we have

$$
\begin{aligned}
\mathrm{E}^*\{f(\boldsymbol{v})\} &= \hat{\mu} + \boldsymbol{r}^T(\boldsymbol{v})C^{-1}(\boldsymbol{y}-\hat{\mu}\boldsymbol{1}_n) \\
\mathrm{E}^*\{f(\boldsymbol{v}')\} &= \hat{\mu} + \boldsymbol{r}'^T(\boldsymbol{v}')C^{-1}(\boldsymbol{y}-\hat{\mu}\boldsymbol{1}_n) \\
\mathrm{Cov}^*\{f(\boldsymbol{v}),f(\boldsymbol{v}')\} &= \hat{\sigma}^2\left\{\exp\left(-\sum_{\ell=1}^d \frac{(v_\ell - v'_\ell)^2}{\hat{\gamma}_\ell}\right) - \boldsymbol{r}^T(\boldsymbol{v})C^{-1}\boldsymbol{r}'(\boldsymbol{v}')\right\}.
\end{aligned}
\quad (13)
$$

Therefore, substituting (13) in (12), and applying the integrations, we finally obtain

$$
\mathrm{E}^*\{(\mathrm{E}(Y))^2\} = \hat{\sigma}^2\left(e^* - \boldsymbol{T}^T C^{-1}\boldsymbol{T}\right) + \left(\hat{\mu} + \boldsymbol{T}^T C^{-1}(\boldsymbol{y}-\hat{\mu}\boldsymbol{1}_n)\right)^2,
$$

where

$$
e^* = \prod_{\ell=1}^d\left\{\int_{a_\ell}^{b_\ell}\int_{a_\ell}^{b_\ell}\exp\left(-\frac{(v_\ell - v'_\ell)^2}{\hat{\gamma}_\ell}\right)\frac{1}{(b_\ell - a_\ell)^2}\mathrm{d}v_\ell \mathrm{d}v'_\ell\right\}.
$$

Turning to the estimate for $\mathrm{Var}(\mathrm{E}(Y\mid u_j))$, we have

$$
\mathrm{E}^*\{\mathrm{Var}(\mathrm{E}(Y\mid u_j))\} = \mathrm{E}^*\left\{\mathrm{E}\left[(\mathrm{E}(Y\mid u_j))^2\right]\right\} - \mathrm{E}^*\{(\mathrm{E}(Y))^2\}
$$

and therefore we only need the expression for $\mathrm{E}^*\left\{\mathrm{E}\left[(\mathrm{E}(Y\mid u_j))^2\right]\right\}$. In particular,

$$
\begin{aligned}
\mathrm{E}^*\left\{\mathrm{E}\left[(\mathrm{E}(Y\mid u_j))^2\right]\right\} &= \mathrm{E}^*\left\{\int (\mathrm{E}(Y\mid u_j))^2 \mathrm{d}H_j(u_j)\right\} \\
&= \int \mathrm{E}^*\left\{(\mathrm{E}(Y\mid u_j))^2\right\}\mathrm{d}H_j(u_j) \\
&= \int\left\{\hat{\sigma}^2\left(e - \boldsymbol{T}_j^T(u_j)C^{-1}\boldsymbol{T}_j(u_j)\right) + \left(\hat{\mu} + \boldsymbol{T}_j^T(u_j)C^{-1}(\boldsymbol{y}-\hat{\mu}\boldsymbol{1}_n)\right)^2\right\}\mathrm{d}H_j(u_j) \\
&= \hat{\sigma}^2 e - \hat{\sigma}^2\left\{\int \boldsymbol{T}_j^T(u_j)C^{-1}\boldsymbol{T}_j(u_j)\mathrm{d}H_j(u_j)\right\} + \hat{\mu}^2 + 2\hat{\mu}\boldsymbol{T}^T C^{-1}(\boldsymbol{y}-\hat{\mu}\boldsymbol{1}_n) \\
&\quad + \left\{\int (\boldsymbol{T}_j^T(u_j)C^{-1}(\boldsymbol{y}-\hat{\mu}\boldsymbol{1}_n))^2 \mathrm{d}H_j(u_j)\right\}.
\end{aligned}
$$

The two integrals above can be computed as follows. First,

$$
\int \boldsymbol{T}_j^T(u_j)C^{-1}\boldsymbol{T}_j(u_j)\mathrm{d}H_j(u_j) = \sum_{m=1}^n\sum_{k=1}^n A_m A_k c_{mk}\left\{\int_{a_j}^{b_j}\exp\left(-\frac{(u_j - x_{mj})^2 + (u_j - x_{kj})^2}{\hat{\gamma}_j}\right)\frac{1}{b_j - a_j}\mathrm{d}u_j\right\}
$$

where, again, c_{mk} is the (m,k)-th element of matrix C^{-1}, and

$$
A_m = \prod_{\{\ell:\ell\neq j\}}\left\{\int_{a_\ell}^{b_\ell}\exp\left(-\frac{(v_\ell - x_{m\ell})^2}{\hat{\gamma}_\ell}\right)\frac{1}{b_\ell - a_\ell}\mathrm{d}v_\ell\right\}, \quad m = 1,...,n.
$$

Moreover,

$$
\begin{aligned}
\int (\boldsymbol{T}_j^T(u_j) C^{-1}(\boldsymbol{y} - \hat{\mu}\mathbf{1}_n))^2 \mathrm{d}H_j(u_j) \;=\; & \sum_{m=1}^{n} z_m^2 \left\{ A_m^2 \int_{a_j}^{b_j} \exp\left(-\frac{2(u_j - x_{mj})^2}{\hat{\gamma}_j} \right) \frac{1}{b_j - a_j} \mathrm{d}u_j \right\} \\
& + 2 \sum_{m=1}^{n} \sum_{k=m+1}^{n} z_m z_k \left\{ A_m A_k \int_{a_j}^{b_j} \exp\left(-\frac{(u_j - x_{mj})^2 + (u_j - x_{kj})^2}{\hat{\gamma}_j} \right) \frac{1}{b_j - a_j} \mathrm{d}u_j \right\}
\end{aligned}
$$

where, again, z_m denotes the m-th element of the vector $C^{-1}(\boldsymbol{y} - \hat{\mu}\mathbf{1}_n)$.

REFERENCES

[1] J. Morisette, F. Baret, and S. Liang, "Special issue on global land product validation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1695–1697, 2006.

[2] W. Yang, B. Tan, D. Huang, M. Rautiainen, N. Shabanov, Y. Wang, J. Privette, K. Huemmrich, R. Fenshold, I. Sandholt, M. Weiss, D. Ahl, S. Glower, R. Nemani, Y. Knyazikhin, and R. Myneni, "MODIS leaf area index products: From validation to algorithm improvement," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1885–1989, 2006.

[3] B. Ganapol, L. Johnson, C. Hlavka, D. Peterson, and B. Bond, "LCM2: A coupled leaf/canopy radiative transfer model," *Remote Sensing of the Environment*, vol. 70, pp. 153–166, 1999.

[4] Y. Knyazikhin, J. Martonchik, R. Myneni, D. Diner, and S. Running, "Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data," *Journal of Geophysical Research*, vol. 103, pp. 32 257–32 276, 1998.

[5] A. Saltelli, K. Chan, and E. Scott, *Sensitivity Analysis*. Chichester: John Wiley and Sons, 2000.

[6] A. Saltelli, "Sensitivity analysis: Could better methods be used?" *Journal of Geophysical Research*, vol. 104, no. D3, pp. 3789–3793, February 1999.

[7] C. Bacour, S. Jaquemoud, Y. Tourbier, M. Dechambre, and J.-P. Frangi, "Design and analysis of numerical experiments to compare four reflectance models," *Remote Sensing of the Environment*, vol. 79, pp. 72–83, 2002.

[8] R. Brun, P. Reichert, and H. Kuensch, "Practical identifiability analysis of large environmental simulation models," *Water Resources Research*, vol. 37, no. 4, pp. 1015–1030, April 2001.

[9] P. Bowyer, F. Danson, and N. Trodd, "Methods of sensitivity analysis in remote sensing: Implications for canopy reflectance model inversion," in *Geoscience and Remote Sensing Symposium*, vol. 6, July 2003, pp. 3839–3841.

[10] D. Mackay, "Introduction to Gaussian process," *NATO ASI Series F Computer and Systems Sciences*, vol. 168, pp. 133–166, 1998.

[11] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[12] R. Neal, "Regression and classification using Gaussian process priors," in *Bayesian Statistics 6*, J. Bernardo, J. Berger, A. Dawid, and A. Smith, Eds. Oxford University Press, 1998, pp. 475–501.

[13] J. Sacks, W. Welch, W. Mitchell, and H. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, pp. 409–435, 1989.

[14] C. Currin, T. Mitchell, M. Morris, and D. Ylvisaker, "Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments," *Journal of the American Statistical Association*, vol. 86, pp. 953–963, 1991.

[15] W. Welch, R. Buck, J. Sacks, H. Wynn, T. Mitchell, and M. Morris, "Screening, predicting and computer experiments," *Technometrics*, vol. 34, pp. 15–25, 1992.

[16] T. Santner, B. Williams, and W. Notz, *The Design and Analysis of Computer Experiments*, ser. Springer Series in Statistics. Springer, 2003.

[17] J. Oakley and A. O'Hagan, "Probabilistic sensitivity analysis of complex models: a Bayesian approach," *Journal of the Royal Statistical Society, Series B*, vol. 66, pp. 751–769, 2004.

[18] B. Ganapol, L. Johnson, C. Hlavka, and D. Peterson, "LEAFMOD: A new within-leaf radiative transfer model," *Remote Sensing of the Environment*, vol. 63, pp. 182–193, 1998.

[19] B. Hosgood, S. Jacquemound, G. Andreoli, J. Verdebout, G. Peredrini, and G. Schmuck, "Leaf OPtical Properties EXperiment (LOPEX93: Report EUR16095EN," Joint Research Center-European Commission, Institute for Remote Sensing Applications, Tech. Rep., 1995.

[20] B. Ganapol and R. Myneni, "The FN method for the one-angle radiative transfer equation applied to plant canopies," *Remote Sensing of the Environment*, vol. 39, pp. 213–231, 1992.

[21] C. Bacour, F. Baret, D. Béal, M. Weiss, and K. Pavageau, "Neural network estimation of LAI, $fAPAR$, $fCover$ and $LAI \times C_{ab}$ from top of canopy MERIS reflectance data: Principles and validation," *Remote Sensing of the Environment*, vol. 105, pp. 313–325, 2006.

[22] M. McKay, R. Beckman, and W. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, pp. 239–245, 1979.

[23] W. Gilks, S. Richardson, and D. Spiegelhalter, Eds., *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC, 1995.

[24] N. Goel, "Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data," *Remote Sensing Review*, vol. 4, pp. 1–222, 1988.

[25] J. Clevers and W. Verhoef, "Modeling and synergistic use of optical and microwave remote sensing. LAI estimation from canopy reflectance and WDVI: a sensitivity analysis with the SAIL model," Netherlands Remote Sensing Board (BCRS), Tech. Rep. 90-39, 1991.

[26] S. Jacquemoud and F. Baret, "PROSPECT: a model of leaf optical properties spectra," *Remote Sensing of the Environment*, vol. 34, no. 75-91, 1990.

[27] J. Dungan and B. Ganapol, "Sources of uncertainty in the prediction of LAI/fPAR from MODIS," in *AGU Fall Meeting*, San Francisco, 2002.