Geostatistical inverse modeling for super-resolution mapping of continuous spatial processes

Jun Wang a,⁎, Daniel G. Brown a, Dorit Hammerling b

a School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI, USA
b Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI, USA

A R T I C L E   I N F O
Article history:
Received 18 October 2012
Received in revised form 4 August 2013
Accepted 9 August 2013
Available online 4 September 2013

Keywords:
Multi-resolution data fusion
Super-resolution mapping
Change-of-support
Spatial nonstationarity
Geostatistical inverse modeling
Spatial prediction
Uncertainty

A B S T R A C T
The increasing availability of satellite images derived from multiple sensors creates opportunities for broader spatial and temporal coverage but also methodological challenges. We present a geostatistical inverse modeling (GIM) approach for merging coarse-resolution images with variable resolutions and for super-resolution (i.e., predictions at the sub-pixel level) mapping of continuous spatial processes. GIM can explicitly account for the differences in spatial supports of multiple datasets. The restricted maximum likelihood method was used for parameter estimations associated with the change-of-support problem. We used GIM to produce both spatial predictions of a target image and prediction uncertainties, while preserving the values of original measurements. GIM is totally data driven, and covariance parameters for a target resolution can be directly derived from measurements. We also developed a moving-window GIM approach to accommodate spatial nonstationarity and reduce computational burden associated with large image data. First, we demonstrated GIM and moving-window GIM on synthetic images. Aggregated synthetic images with variable resolutions were merged to produce a single resolution image. The results show that the two approaches can produce accurate spatial predictions and generate prediction uncertainties. Second, we applied moving-window GIM for merging aerosol optical depth (AOD) data with variable resolutions, which were derived from two satellite sensors. The modeling results show that moving-window GIM can be applied for merging complementary AOD data from two sensors and for super-resolution mapping of global AOD distributions. Therefore, we can conclude that GIM is a practical solution for merging complementary coarse-resolution images and for super-resolution mapping of continuous spatial processes.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Synthesizing complementary information derived from multiple sensors prompts the need to study rigorous data fusion algorithms. Data fusion is a process that integrates information derived from different sensors or different spectral bands of the same sensor and produces a single image that contains complementary information from multiple sources, while minimizing loss or distortion of the original data (Hall, 2004; Pohl & Van Cenderen, 1998). In this work, we focus on statistical algorithms for merging measurements derived from multiple coarse-resolution sensors. Statistical data fusion combines statistically heterogeneous samples from marginal distributions to make statistical inference about the unobserved joint distributions or functions of them (Braverman, 2008). Statistical data fusion, including those based on geostatistics, can produce spatial predictions of pixel values (Atkinson, Pardo-Iguzquiza, & Chico-Olmo, 2008). Recently, several geostatistical algorithms, including fixed ranking kriging (Cressie & Johannesson, 2008; Shi & Cressie, 2007), fixed ranking filtering (Cressie, Shi, & Kang, 2010; Kang, Cressie, & Shi, 2010), spatial statistical data fusion (Nguyen, Cressie, & Braverman, 2012), space-time data fusion (Braverman, Nguyen, & Cressie, 2011), and moving-window kriging (Hammerling, Michalak, & Kawa, 2012), have been developed for mapping global distributions of environmental variables, such as aerosol optical depth (AOD) and carbon dioxide (CO2), with sparsely distributed remotely sensed data. The numerous algorithms for merging measurements from different spectral bands (e.g., pan-sharpening) are beyond the scope of this paper.

Measurements from remote sensing sensors are constantly influenced by factors like atmospheric conditions, electronic noise of sensors, and changes in illumination. In order to build geostatistical models of sensor measurements contaminated with measurement errors, we take a stochastic view of remote sensing images. We regard the true spatial process of interest (i.e., spectral radiance) as a random field, i.e., a spatial random process with a set of random variables that have certain probability distributions. Then, a remote sensing image covering an area can be conceived as a realization of the random field. In this work, we adopt the Gaussian random field model, which involves

http://dx.doi.org/10.1016/j.rse.2013.08.007
a set of Gaussian probability density functions for random variables. For remote sensing measurements, the values of continuous spatial processes of interest are regularized to discrete pixels by a weighted average process, with the spatial weights determined by point spread functions (PSFs) of sensors (Jupp, Strahler, & Woodcock, 1988). The effective instantaneous field of view (EIFOV) of the sensor is an area over which instantaneous measurements are averaged. EIFOV defines the spatial support of sensor measurements. Spatial support is a geostatistical concept that means the shape, size, and orientation of measurements (Gotway & Young, 2002, 2005). The value assigned to a pixel represents the average radiance arriving at the sensor from the EIFOV (Jupp, Strahler & Woodcock, 1988). Sensors with different sizes of pixels have different EIFOVs. As a result, measurements from these sensors have different spatial supports.

Merging remote sensing images with variable resolutions usually involves the change-of-support problem (Curran & Atkinson, 1999; Gotway & Young, 2002, 2005). Several geostatistical algorithms have been developed to solve the change-of-support problem. Area-to-point kriging was developed for downscaling areal data to point support (Kyriakidis, 2004; Kyriakidis & Yoo, 2005). Goovaerts (2008) developed a practical semivariogram deconvolution algorithm to derive point support variogram parameters from areal data. This algorithm solved one of the key problems in the practical application of area-to-point kriging. Moreover, a parallel computing algorithm has been developed for speeding-up the computations involved in the practical application of area-to-point kriging (Guan, Kyriakidis, & Goodchild, 2011). Area-to-point kriging has the potential for downscaling remote sensing data. Nguyen et al. (2012) developed a statistical data fusion approach, which integrates a change-of-support model into fixed rank kriging for multi-resolution data fusion. It was applied for merging variable-resolution Aerosol Optical Depth (AOD) images derived from the Multi-angle Imaging Spectrometer (MISR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. Sales, Souza, and Kyriakidis (2013) applied a Kriging with external drift (KED) method for increasing 500-m resolution (bands 3–7) MODIS images to 250-m resolution.

The spatially varying dependence structure of random variables (i.e., spatial nonstationarity) is another problem to be dealt with when working with remote sensing images covering large geographic areas. In geostatistics, a fundamental assumption of most models is that random fields are second-order stationary, i.e., in a relatively small region the mean values of random variables are constant and the covariance between the values of random variables only depends on the distance between them (Chilès & Delfiner, 1999). However, for spatial data covering large geographic areas, this assumption may not be true because remote sensing images covering large geographic areas usually include spatial nonstationarity. In this work, we differentiated two types of spatial nonstationarity: one is spatial nonstationarity in the mean values of regionalized variables, and the other is spatial nonstationarity in the covariance structure. The need to address spatial nonstationarity has been discussed in the field of geostatistics over the past decades. Universal kriging is one way to address spatial nonstationarity in the mean values of regionalized variables. Hass (1990) applied a moving-window kriging approach to model acid deposition. In the moving-window kriging model, measurements in local-windows are used for both parameter estimations and spatial predictions. This approach is simple to be implemented, and it alleviates the problem of spatial nonstationarity. Because of local fitting and computing, moving-window kriging is also computationally efficient. One caveat of the local-window approach is that there is no consistent covariance function over the whole study domain. Higdon, Swall, and Kern (1999) convolved spatially varying kernels to give a nonstationary version of the squared exponential stationary covariance function. This approach has been applied in modeling remote sensing images (D’Hondt, LópezMartínez, Ferro-Famil, & Pottier, 2007). Although this method can produce a consistent covariance function over the whole prediction domain, the Gaussian kernel applied in this method is too smooth for real spatial processes.

Besides spatial nonstationarity, the problem of computational burden also needs to be solved when applying geostatistical models for dealing with large spatial data. The local-window approach (e.g., moving-window kriging) is one way to solve the problem of computational burden. Data dimension reduction is another way to reduce computational burden associated with large spatial data (Wikle, 2010). Cressie and Johannesson (2008) developed a spatial mixed effects model (i.e., fixed rank kriging) with a flexible family of nonstationary covariance functions. For this approach, kriging can be done exactly, and the computational complexity is linear to the size of the data. Moreover, Cressie et al. (2010) developed a spatial–temporal random effect model (i.e., fixed rank filtering), which integrates fixed rank kriging and Kalman filter for dealing with large spatial–temporal data. Fixed rank kriging and fixed rank filtering are both approaches of data dimension reduction. These methods eliminate or reduce some components of spatial variability to improve computational efficiency.

In this work, we present a geostatistical inverse modeling approach for merging coarse-resolution remote sensing images with variable spatial supports. The geostatistical inverse model was designed to be statistically principled, and it can produce the best predictions (i.e., minimizing the squared errors between predictions and measurements) and prediction uncertainties, while honoring the original data (i.e., preserving the values of original measurements) (Kitanidis, 1995; Michalak, Bruhwiler, & Tans, 2004). In the geostatistical inverse modeling framework, the restricted maximum likelihood method was used for estimating covariance parameters related to the change-of-support problem. Moreover, we contributed a moving-window geostatistical inverse modeling approach to accommodate spatial nonstationarity and reduce computational burden associated with large spatial data. Following the introduction, we introduce the geostatistical inverse modeling methodology in Section 2. In Section 3, we illustrate the computer experiments using synthetic and real images. The modeling results are presented in Section 4. Finally, we discuss possible model improvements and summarize the major findings of this work.

**2. Methodology**

**2.1. Geostatistical inverse modeling**

Geostatistical inverse modeling (GIM) follows a Bayesian approach, and it is based on the principle of combining the prior information (i.e., spatial and/or temporal autocorrelation) with the information from available measurements (Michalak, Bruhwiler & Tans, 2004). Spatial and/or temporal autocorrelation can provide information about the structure of the data that can be used to reduce prediction uncertainty. GIM has been applied in ground water systems (Kitanidis, 1995), contaminant sources identification (Snodgrass & Kitanidis, 1997), estimating surface fluxes of atmospheric trace gases (Gourdji, Mueller, Schaefer, & Michalak, 2008; Michalak, Bruhwiler & Tans, 2004), characterizing attribute distributions in water sediments (Zhou & Michalak, 2009), and merging remote sensing images with variable resolutions (Erickson & Michalak, 2006). There remain unrealized opportunities in applying GIM for image scaling (i.e., downsampling and up-scaling) and multi-resolution data fusion. In comparison with area-to-point kriging, which also deals with predicting point values from areal data, the covariance parameters in the GIM framework can be inferred directly from measurements by the restricted maximum likelihood algorithm. Moreover, measurements with variable spatial supports can also be merged to produce a single resolution image using GIM.

We only present the key equations of GIM here. Readers are referred to Michalak et al. (2004) for an in-depth discussion about GIM. The spatial prediction problem of GIM can be expressed as

\[ z = h(s, \theta) + v \]  \hspace{1cm} (1)
where \( z \) is an \( n \times 1 \) vector of measurements, \( s \) is an \( m \times 1 \) vector of predictions, and the vector \( r \) contains other parameters needed by the transformation function \( h(s, r) \). The vector \( v \) describes the model-data mismatch (i.e., error). The error includes both measurement errors associated with collecting the data and any random numerical or conceptual inaccuracies associated with the evaluation of the function \( h(s, r) \). When the function is linear with respect to the unknown surface \( s \), it can be written as

\[
h(s, r) = Hs
\]

(2)

where \( H \) is a known \( n \times m \) matrix, the Jacobian representing the sensitivity of the measurements \( z \) to the predictions \( s \) (i.e., \( H_{ij} = \partial z_i / \partial s_j \), where \( i \) is the index of measurements, and \( j \) is the index of prediction locations). For remote sensing applications, the transformation function can be defined as the point spread functions (PSFs) of sensors. PSFs for many electro-optical sensors follow a two-dimensional Gaussian distribution (Huang, Townshend, Liang, Kalluri, & DeFries, 2002). The sensitivity matrix is calculated in the same way where one or multiple datasets are considered. In the case of multiple datasets with variable spatial supports, the sensitivity matrix is

\[
H = \begin{bmatrix} H_1 \\ \vdots \\ H_N \end{bmatrix}
\]

(3)

where \( N \) is the number of datasets. We assume \( v \) has zero mean and known covariance matrix \( R \). The covariance of the measurement errors is most commonly modeled as

\[
R = \sigma_v^2 I
\]

(4)

where \( \sigma_v^2 \) is the variance of the measurement error, and \( I \) is an \( n \times n \) identity matrix.

Following the Bayes’ theorem, the posterior probability density function of a state vector \( s \) given an observation vector \( z \) is proportional to the likelihood of the state given the data (or, conversely, the probability density function of the state given the state) times the prior probability density function of the state (Eq. 5). The drift parameters \( \beta \) are estimated along with \( s \).

\[
p(s|\beta|z) \propto |R|^{-1/2} \sum p(s|\beta) \exp \left[ -\frac{1}{2} (z-Hs)^T R^{-1} (z-Hs) \right]
\]

(5)

The prior probability density function represents the assumed spatial structure of the prediction field. The likelihood of the data represents the degree to which a prediction of the unknown function \( s \) represents the measurements \( z \). The prior probability density function is modeled as

\[
p(s, \beta) = p(s|\beta) p(\beta) \propto |Q|^{-1/2} \exp \left[ -\frac{1}{2} (s-X\beta)^T Q^{-1} (s-X\beta) \right]
\]

(6)

where \( X \) is an \( m \times t \) matrix of auxiliary variables related to the distribution of predictions, \( t \) is the number of auxiliary variables, \( \beta \) is a \( t \times 1 \) vector of drift coefficients on \( X \), \( X\beta \) is the model of the trend, and \( Q \) is an \( m \times m \) matrix representing the spatial autocorrelation of residuals that are not explained by the model of the trend. The prior probability density function of \( \beta \) is assumed to be uniform over all values (\( p(\beta) \propto 1 \)). The likelihood of measurements is modeled as

\[
p(z|s) \propto |R|^{-1/2} \exp \left[ -\frac{1}{2} (z-Hs)^T R^{-1} (z-Hs) \right]
\]

(7)

The posterior probability density of the unknown surface distributions \( s \) becomes

\[
p(s, \beta|z) \propto |R|^{-1/2} |Q|^{-1/2} \exp \left[ -\frac{1}{2} (z-Hs)^T R^{-1} (z-Hs) - \frac{1}{2} (s-X\beta)^T Q^{-1} (s-X\beta) \right]
\]

(8)

By taking the negative logarithm of the above formula, we can get the objective function of GIM.

\[
L_{\lambda\beta} = (z-Hs)^T R^{-1} (z-Hs) + (s-X\beta)^T Q^{-1} (s-X\beta)
\]

(9)

The best predictions are obtained by minimizing the objective function with respect to \( s \) and \( \beta \). By taking the first order derivative of the above objective function with respect to \( s \) and \( \beta \) and expressing the result in the matrix form, we can get the solution of the GIM.

\[
\begin{bmatrix} HQH^T + R & HX \\ (HX)^T & M \end{bmatrix} \begin{bmatrix} \lambda z \\ \beta \end{bmatrix} = \begin{bmatrix} HQ \\ X^T \end{bmatrix}
\]

(10)

where \( \lambda \) is an \( m \times n \) matrix of estimated weights assigned to related observations, and \( M \) is a \( p \times m \) matrix of Langmuir multipliers. By solving \( \lambda \) and \( M \), we can get the best predictions of \( s \) and \( \beta \)

\[
\hat{s} = \lambda z
\]

(11)

\[
\hat{\beta} = (X^T Q^{-1} X)^{-1} X^T Q^{-1} \hat{s}
\]

(12)

where \( \hat{s} \) and \( \hat{\beta} \) are the best predictions of \( s \) and \( \lambda \), respectively. The posterior covariance of \( \hat{s} \) is calculated by the inverse of the Hessian of the objective function (Kitanidis, 1995; Michalak, Bruhwiler & Tans, 2004).

\[
\begin{bmatrix} V_{\hat{s}} & V_{\hat{s}\beta} \\ V_{\hat{s}\beta} & V_{\hat{\beta}} \end{bmatrix} = \begin{bmatrix} Q^{-1} + H^T R^{-1} H & Q^{-1} X^{-1} \\ X^T Q^{-1} & X^T Q^{-1} X \end{bmatrix}^{-1}
\]

(13)

The portion of the matrix defining the posterior covariance of \( \hat{s} \) is

\[
V_{\hat{s}} = -XM + Q^{-1} Q^T \lambda
\]

(14)

The diagonal elements of \( V_{\hat{s}} \) represent the predicted error variance of individual elements in \( \hat{s} \). An overview of GIM for synthesizing information from multiple image datasets is shown in Fig. 1.

In order to accommodate spatial nonstationarity and reduce computational burden associated with large spatial data, we developed a moving-window GIM approach, which is an integration of GIM and the moving-window approach. The general idea of moving-window GIM is that instead of using all available measurements to predict the value at a prediction location, only measurements within local-windows are used. Local-windows are centered at prediction locations. The rationale behind this approach is that measurements distant from a prediction location may contribute very little to the prediction, except long-range dependence in sparse data situations. In most remote sensing applications, this is particularly true because images usually cover everywhere in the study domain, except the missing pixels caused by measurement errors or clouds. Moreover, the point spread functions (PSFs) of optical sensors also have a blurring effect because pixel values are calculated using data from neighboring pixels (Huang, Townshend, Liang, Kalluri & DeFries, 2002). In this work, the size of local-windows was chosen based on pre-calculations of covariance parameters over the prediction domain in order to guarantee the size of local-windows meets both computational requirements and the accuracy of approximation. We used measurements in local-windows for both parameter estimates and spatial predictions.
2.2. Covariance parameter estimations

Parameter estimations refer to the statistical inference of covariance parameters for a stochastic model of measurements. There are several approaches for estimating covariance parameters of geostatistical models, such as ordinary-least-square (OLS; Bogaert & Russo, 1999), weighted-least-square (WLS; Cressie, 1985), and restricted maximum likelihood (REML; Kitanidis, 1995; Michalak, Bruhwiler & Tans, 2004).

In this work, we applied REML for optimizing covariance parameters by maximizing the likelihood of measurements. REML is a variant of the maximum likelihood (ML) algorithm. REML leads to less biased estimates of covariance parameters when the sample size is small (Diggle & Ribeiro, 2007). Image scaling and multi-resolution data fusion involve two types of parameter estimations. If there are available measurements at the prediction resolution, we can directly estimate covariance parameters using the measurements. Otherwise, we have to infer covariance parameters at the prediction resolution using data measured at other spatial resolutions. REML can be applied for the two types of parameter estimations.

First, when measurements are available at the prediction resolution, REML for parameter estimations is referred to as “REML-Kriging (Gouradj, Hirsch, Mueller, Yadav, Andrews & Michalak, 2010; Nagle, Sweeney, & Kyriakidis, 2011).” The objective function of “REML-Kriging” is defined as

\[
L = -\frac{1}{2} \ln |Q| + \frac{1}{2} \ln |X^T Q^{-1} X| + \frac{1}{2} \text{trace} \left( Q^{-1} - Q^{-1} X (X^T Q^{-1} X)^{-1} X^T Q^{-1} \right)^2
\]  

In this case, the gradient-based searching routine can be used for parameter optimizations because the number of measurements is usually quite large for remote sensing applications. The uncertainties of the covariance parameters (i.e., sill, range, and nugget) in Q can be calculated by the Hessian computations.

Second, when no measurements are available at the prediction resolution, REML for parameter estimations is referred to as “REML-Inverse Modeling (Kitanidis, 1995; Michalak, Bruhwiler & Tans, 2004).” The objective function of “REML-Inverse Modeling” is defined as

\[
L = \frac{1}{2} \ln |\Psi| + \frac{1}{2} |X^T H^T \Psi^{-1} H X| + \frac{1}{2} \text{trace} \left( R^{-1} - \Psi^{-1} X (X^T H^T \Psi^{-1} H X)^{-1} X^T H^T \Psi^{-1} \right)^2
\]

If the number of measurements is small, we can use the unconstrained nonlinear optimization routine for parameter optimizations. Otherwise, the gradient-based searching routine can be applied for parameter optimizations. The two optimization algorithms are both provided in the software package MATLAB (Mathworks Inc., Natick, Massachusetts, USA). In the moving-window GIM model, we used “REML-Inverse Modeling” for estimating covariance parameters in local-windows because “REML-Inverse Modeling” leads to less biased estimates of covariance parameters when the sample size is small (Diggle & Ribeiro, 2007). The unconstrained nonlinear optimization routine was applied for parameter optimizations in local-windows.

In this work, we applied the exponential covariance function for all geostatistical models. The exponential covariance function and the corresponding semivariogram model are defined as

\[
\Psi = HQH^T + R
\]

\[
\Xi = \Psi^{-1} - \Psi^{-1} X H \left( X^T H^T \Psi^{-1} H X \right)^{-1} X^T H^T \Psi^{-1}
\]

where \(\sigma^2\) is the variance, \(l\) is the integral scale, and \(h\) is the distance between two measurements. The straight-line distance (i.e., Euclidian distance) and the great-circle distance were used in this work for synthetic and real data modeling experiments, respectively. The great-circle distance between \(x_i\) and \(x_j\) is defined as (Diggle & Ribeiro, 2007)

\[
h(x_i, x_j) = \theta \left( \sin \varphi_1 \sin \varphi_2 + \cos \varphi_1 \cos \varphi_2 \cos (\lambda_1 - \lambda_2) \right)
\]

where \(\theta\) is the radius of the Earth, \(\varphi_1\) and \(\varphi_2\) are the latitude and longitude of \(x_i\), respectively, and \(\lambda_1\) and \(\lambda_2\) are the latitude and longitude of \(x_j\), respectively.

2.3. Unconditional and conditional simulations

Unconditional and conditional simulations are spatially consistent Monte Carlo simulations. They are geostatistical approaches for describing spatial variability of random fields. Conditional simulations generate realizations of a random field that possess the same structural characteristics as measurements. It can also reproduce measurements. Conditional simulations are equally likely realizations that have the same spatial structure defined by the covariance function. Unconditional simulations also follow the same spatial dependence structure defined by measurements, but they cannot reproduce measurements (Chilès & Delfiner, 1999; Diggle & Ribeiro, 2007). Unconditional and conditional simulations can generate multiple outcomes, and each of which is an equally likely realization. When modeling a stationary Gaussian random field over an area that is much larger than the range, a single simulation can give a view of a variety of possible local situations (Chilès & Delfiner, 1999).

In order to demonstrate the application of GIM for image downsampling and multi-resolution data fusion on synthetic images with spatially stationary characteristics, we used unconditional simulations to generate such synthetic data. In this case, the Gaussian random field model was used in modeling the spatial random process. The Gaussian models are widely used because they are simple, and they can capture a wide range of spatial behaviors according to the specification of their correlation structures (Diggle & Ribeiro, 2007). The random field is spatially stationary (i.e., second-order stationary) if the mean values of regionalized variables are constant, and the covariance is only determined by the distance between two measurements (Chilès & Delfiner, 1999).

The generation of synthetic images proceeded as follows. The user-
specified covariance function is first decomposed by the Cholesky decomposition

\[ Q = CC^T \]

where \( Q \) is the covariance function. The unconditional simulations are generated by

\[ s_{ui} = x_i \beta + u_i \]  

(23)

where \( x_i \) is the model of spatial trend and can be set as zero, and \( u_i \) is a vector of normally distributed random numbers with zero-mean and unit-variance. In the case of generating synthetic images, the model of spatial trend was set as zero.

The conditional simulations of the geostatistical inverse model are defined as (Kitanidis, 1995; Michalak, Bruhwiler & Tans, 2004)

\[ s_{ui} = s_{ui} + \lambda (z + v - Hs_{ui}) \]  

(24)

where \( v \) is a vector of normally distributed random numbers, which is sampled from a multivariate normal distribution with mean zero and the model-data mismatch error covariance \( K = \sigma_u^2L \). \( s_{ui} \) is the unconditional simulation of GIM (Eq. 23).

2.4. Fixed Rank Kriging

In order to demonstrate the application of the moving-window GIM approach for image downscaling and multi-resolution data fusion on synthetic data with spatially nonstationary characteristics, we used fixed ranking kriging (FRK; Cressie & Johannesson, 2008) to generate such synthetic data. FRK is a low-rank representation of spatial continuous random processes, and it eliminates or reduces some components of spatial variability to improve computational efficiency. This data dimension reduction approach was developed for dealing with the spatial prediction problem associated with large spatial data. Remote sensing images covering large geographic areas usually have spatially nonstationary characteristics. FRK can accommodate spatial nonstationarity by using a set of multi-scale basis functions in the model. Readers are referred to Cressie et al. (2010), Cressie and Johannesson (2008), and Kang et al. (2010) for detailed discussions about FRK. Only brief descriptions are provided here because we only used FRK to generate synthetic images.

The goal is to make statistical inferences of a spatial process \( \{Y(s) : s \in D \subseteq \mathbb{R}^d\} \) based on measurements with errors. The measurements are given by the data model

\[ Z(s) = Y(s) + \varepsilon(s) \]  

(25)

where \( \varepsilon(s) \in D \) is a white-noise Gaussian process with zero-mean and variance \( \sigma_\varepsilon^2 > 0 \). We assume that \( Y(s) \) has the following structure,

\[ Y(s) = \mu(s) + \upsilon(s) \]  

(26)

where \( \mu(\cdot) \) is a deterministic trend function representing the large-scale spatial variability (e.g., \( \mu(\cdot) = X(\cdot) \)). The small-scale spatial variability is modeled as a Gaussian process. We assume \( \upsilon(\cdot) \) is with zero-mean, and it follows a spatial random effect model. The unknown random variables to be predicted are fixed at a number, which is equal to the number of spatial basis functions. It can be expressed as

\[ \upsilon(s) = S(s) \eta + \xi(s) \]  

(27)

where \( S(\cdot) \equiv (S_1(\cdot), ..., S_l(\cdot))^T \) is a set of basis-functions, which can capture different scales of spatial dependence. The basis-functions are not necessarily orthogonal, but they should represent information at multiple resolutions. The multi-resolution wavelet basis function

\[ C(h) = 0.01 \times \left( 1 - \frac{h}{10} \right) \]  

(30)

\[ where \ v_{jl}^{(n)} \ is one of the center points of the \ l_n \ resolution (l = 1, 2, 3, N), \ and n = 1.5 \ (i.e., \ the \ shortest \ arc \ distance \ between \ center \ points \ of \ the \ l_n \ resolution). The \ generated \ synthetic \ images \ can \ have \ more \ spatial \ variability \ by increasing \ the \ number \ of \ resolutions \ in \ the \ model. However, \ this \ will \ increase \ computational \ requirements \ for \ generating \ such \ synthetic \ data. The \ spatially \ nonstationary \ covariance \ function \ is modeled \ as

\[ C(u, v) = S(u) KS(v), \quad u, v \in \mathbb{R}^d \]  

(29)

where \( K \) is a positive definite \( r \times r \) unknown matrix, \( \eta = (\eta_1, ..., \eta_r)^T \) is a zero-mean Gaussian random vector with \( \text{cov}(\eta) = K \), and \( \xi(\cdot) \) captures the fine-scale spatial variability. The fraction of the fine-scale spatial variability to the total spatial variability and the signal-to-noise ratio of the generated synthetic image can also be defined in the FRK model (Cressie, Shi & Kang, 2010).

3. Computer experiments

In this work, the overall geostatistical modeling experiments proceeded as follows. First, in order to isolate various problems associated with real remote sensing images (e.g., measurement errors), we demonstrated GIM and moving-window GIM for image downscaling and multi-resolution data fusion on synthetic images. Second, we applied moving-window GIM for merging AOD measurements derived from two sensors and for super-resolution mapping of global AOD distributions. For all experiments on synthetic and real data, we used REML for estimating covariance parameters. In order to assess prediction accuracies and the correctness of prediction uncertainties, we calculate the root mean square error (RMSE) between spatial predictions and measurements, and the percentage of pixels in original images with their values falling within two standard deviations of the mean. We call the first measure the RMSE index and the second measure the percentage index. In addition, image registration is an important step prior to pixel-level data fusion. We assumed that the images used for multi-resolution data fusion have been registered correctly. Analyzing the influence of image mis-registration on the accuracy of spatial predictions is beyond the scope of this work. All of the geostatistical models discussed in the paper (Section 2) and the following experiments were coded in MATLAB by the authors.

3.1. Generating synthetic images

We applied unconditional simulations with user-specified covariance parameters to generate synthetic images with spatially stationary characteristics. The synthetic images with and without nugget effect were generated and used in the modeling experiments. The simulated reference pixel values were distributed on a 120 × 120 regular grid (of unit cell size). In this work, we differentiated fine-scale spatial variability and measurement errors, and we defined nugget effect as fine-scale variability of image data (Cressie, Shi & Kang, 2012). Moreover, we did not introduce simulated measurement errors into the synthetic images. The covariance functions used for generating synthetic images are

\[ C(h) = 0.01 \times \left( 1 - \frac{h}{10} \right) \]  

(30)
\[ C(h) = 0.005 + 0.02 \times \left(1 - \frac{h}{5}\right) \]  

where \( C(h) \) is the covariance function, and \( h \) is the Euclidian distance between two locations. Eq. (30) is the exponential covariance function, and the values of sill and range are 0.01 and 30, respectively. Eq. (31) is a combination of the exponential covariance function and nugget effect, and the values of sill and range are 0.025 and 15, respectively.

We used FRK to generate a synthetic image with spatially nonstationary characteristics. The values of sill and range for the exponential covariance function in FRK were set as 0.05 and 30, respectively. The specified covariance function was used to parameterize the K matrix in the spatially nonstationary covariance function (Eq. 29). In order to simulate spatial variability of real remote sensing images, we used a large number of basis-functions for characterizing multi-scale spatial variability. In this experiment, we set the resolution number \( (l)^2 \) as five. The number of bi-square basis-functions used for generating the synthetic image was 1364, which was the sum of basis-functions in each of the five lower resolutions: \( 4 \times \) (1 × 1), 16 (4 × 4), 64 (16 × 4), 256 (64 × 4), and 1024 (256 × 4). We set the fraction of fine-scale spatial variability to the total spatial variability as 0.05. We did not introduce measurement errors into the synthetic image. Therefore, the value of the signal-to-noise ratio was \( \infty \). The simulated pixels were distributed on a 320 × 320 regular grid (of unit size).

3.2. Merging synthetic images with variable resolutions

We designed four experiments for demonstrating the application of GIM for image downscaling and multi-resolution data fusion on synthetic images with spatially stationary characteristics: two were on the synthetic image without nugget effect, and the other two were on the synthetic image with nugget effect. The modeling experiments proceeded as follows. Given space limitations, we only used the experiment on the synthetic image without nugget effect as an example. First, we averaged the synthetic image to two coarser resolutions using a non-overlapping moving-window. The first aggregated image was with the spatial support of 64 (i.e., the size of the local-window was 8 × 8), and the second aggregated image was with the spatial support of four (i.e., the size of the local-window was 2 × 2). We only used the diagonal part of the second aggregated image for the experiment. We set two relationship matrices (i.e., the H matrix in GIM) for the two aggregated images because they had different spatial supports. We used the uniform distribution function to represent the relationships between coarse-resolution measurements and predictions, and we assumed sub-pixels within a coarse-resolution pixel had the same contribution to the value of the coarse-resolution pixel. The size of the prediction field was set as the same size of the original simulated image. We also generated conditional simulations to show spatial uncertainties of the predictions. In the second experiment, the aggregated image with the spatial support of 64 was merged with the diagonal part of the aggregated image with the spatial support of 16 (i.e., the size of the local-window was 4 × 4). We also ran two similar experiments on the synthetic image with nugget effect.

We demonstrated the application of moving-window GIM for multi-resolution data fusion on the synthetic image with spatially nonstationary characteristics. Similar to the above experiments, the synthetic image generated by FRK was averaged using non-overlapping moving windows to two coarser resolutions: one aggregated image was with the spatial support of 16 (i.e., the size of the local-window was 4 × 4), and the other aggregated image was with the spatial support of four (i.e., the size of the local-window was 2 × 2). Only the diagonal part of the second aggregated image was used in the experiment. We also set two relationship matrices for the two aggregated images with different spatial supports. The uniform distribution function was also used for representing the relationships between coarse-resolution measurements and fine-resolution predictions. In order to assess the modeling results, the size of the prediction field was also set as the same size of the original simulated image.

3.3. Merging AOD measurements with variable resolutions

We applied moving-window GIM for merging Level-3 AOD data derived from the MISR and the Terra-MODIS sensors and for super-resolution mapping of global AOD distributions in short time-intervals. The spatial resolutions of Level-3 MISR AOD and Terra-MODIS AOD are 0.5° × 0.5° and 1° × 1°, respectively. MISR and Terra-MODIS are both carried by the Terra satellite but with different sensor characteristics, such as the size of swath and the instantaneous field of view. Second, the algorithms for retrieving AOD from original MISR and Terra-MODIS data are different. MODIS land retrieval algorithm does not operate over desert (e.g., northern Africa and Middle East) or other bright land surfaces, and MODIS has difficulties in computing AOD in part of its swath over dark water due to sun glint. MISR can measure AOD in the above conditions (Kahn et al., 2009). Moreover, MODIS has much wider spatial coverage than MISR does due to its wider swath. However, MISR has measurements in some areas where MODIS does not have. Given the differences in instrumental designs and retrieval algorithms, AOD measurements from the MISR and MODIS sensors were used in conjunction with one another to exploit their complementary strengths, especially for the middle and low latitudes (Kahn et al., 2009). In addition, AOD is changing quickly in space and time, measurements from one sensor in a short time-interval may not be enough to capture the functional features of the continuous AOD process. The complementary coverage makes the two AOD datasets suitable for data fusion.

However, there were several difficulties in merging the AOD measurements from the MISR and MODIS sensors. We have discussed spatial nonstationarity and computational burden associated with large spatial data. In the real data modeling experiments, we had both of the problems. In this section, we ran two experiments. One experiment was merging the one-day AOD measurements (i.e., August 1, 2008) derived from the two sensors, and the other experiment was merging the eight-day AOD measurements (i.e., from July 28, 2008 to August 4, 2008) derived from the two sensors. These data were selected randomly with no specific reasons. For the one-day AOD measurements, the numbers of pixels for the MISR AOD and MODIS AOD measurements were 15,612 and 23,253, respectively. For the eight-day AOD measurements, the numbers of pixels for the MISR AOD and MODIS AOD measurements were 129,664 and 40,766, respectively. Most of the AOD measurements from the two sensors were made between 70° N and 50° S, and between 180° W and 180° E. For the two real data modeling experiments, the MISR AOD and MODIS AOD measurements were merged and downscaled to 0.25° × 0.25° for super-resolution mapping of global AOD distributions. Similar to the experiments on synthetic data, we also used two relationship matrices for the two AOD measurements with different spatial supports. We did not know exactly the relationships between measurements and predictions for the aggregated AOD data. We assumed that sub-pixels within a coarse-resolution pixel had the same contribution to the value of the coarse-resolution pixel. We used the uniform distribution function to calculate the relationship matrix (i.e., the H matrix in GIM). We applied moving-window GIM to accommodate spatial nonstationarity and reduce computational burden associated with large spatial data. By pre-calculations and analyses, we set a local-window with the size of 2000-km for both parameter estimations and spatial predictions. Even for the areas with the AOD measurements from both the MODIS and MISR sensors, merging them through the moving-window GIM approach can provide more accurate and robust predictions of the true AOD process. We also calculated prediction uncertainties for the two real data experiments.
4. Results

4.1. Image downscaling and multi-resolution data fusion

For the synthetic images generated by unconditional simulations, the image without nugget effect and with a longer correlation length and a lower sill (Fig. 2A) shows less spatial variability than the other image (Fig. 2B). The synthetic image generated by FRK is shown in Fig. 2C.

The results of the five synthetic data modeling experiments are shown in Table 1. The values of RMSE index and the percentage index show that the spatial predictions of the five experiments were all accurate and the prediction uncertainties were all correct. The values of RMSE were all lower than 10% of the original measurements. Theoretically, for realizations of Gaussian random fields, 95% of the pixel values of the original images should fall within the two standard deviation bounds of the best spatial predictions. The prediction accuracies of the modeling experiments using the synthetic images without nugget effect (Experiments 1 and 2) were better than the results of the experiments using the synthetic images with nugget effect (Experiments 3 and 4). The prediction accuracies of the modeling experiments using fine-resolution aggregated images (Experiments 1 and 3) were better than the results of the experiments using coarse-resolution aggregated images (Experiments 2 and 4). The prediction accuracy of Experiment 5 was better than the results of the first four experiments, and the percentage index of this experiment was also higher than the results of the first four experiments. This may be caused by the fact that the synthetic image generated by FRK has less spatial variability than the synthetic images generated by unconditional simulations.

Given space limitations, we only show the figures of the modeling results of Experiments 1, 3, and 5 (Fig. 3). The spatial predictions show more spatial details in the diagonal parts with fine-resolution aggregated data added in (Fig. 3B, F, and J). Moreover, the diagonal parts had lower prediction uncertainties than the other parts (Fig. 3C, G, and K). These results were reasonable because all spatial predictions benefit from the presence of nearby measurements. Spatial predictions made in the densely measured regions should have lower errors than those made in the sparsely measured regions. The conditional simulations (Fig. 3D, H, and L) look more realistic as the original simulated data (Fig. 2A, B, and C), and they show less smoothing effect than the spatial predictions (Fig. 3B, F, and J). From Fig. 3K we can see that the values of prediction uncertainties (i.e., standard errors) varied across the domain. This demonstrates that moving-window GIM could accommodate the spatial nonstationarity of the synthetic image generated by FRK. Therefore, using spatially stationary covariance functions for the realizations of spatial nonstationary random fields may introduce more prediction errors and generate incorrect prediction uncertainties.

4.2. Super-resolution mapping of global AOD distributions

From the one-day AOD measurements derived from the MISR and MODIS sensors (Fig. 4A and C), we can see that MODIS had more complete spatial coverage than MISR did. MODIS AOD was missing in some of the low and middle latitudes (Fig. 4A), MISR had measurements in those regions, but the field of view of MISR was narrow (Fig. 4C). The eight-day MODIS AOD measurements had more complete coverage than the one-day MODIS AOD measurements. However, there were still gaps in the low and middle latitudes (i.e., North Africa and Middle East) (Fig. 4B). The eight-day MODIS AOD measurements had better coverage than the one-day MISR AOD measurements, but they still did not cover the whole globe (Fig. 4D). Visually, the data gaps of the eight-day MODIS AOD measurements can be mostly covered by the eight-day MISR AOD measurements. Merging the two datasets can produce a more complete picture of global AOD distributions.

By comparing the spatial predictions using the one-day and eight-day AOD measurements (Fig. 4E and F), we can find that although the general patterns of the two images were similar, there were still differences between the two figures, especially in North Eurasia and North America. This indicates that aerosol concentrations were shifting quickly over space and time. Therefore, mapping AOD distributions over short time-intervals may be important. Previous studies found that the high concentrations of AOD were from different sources: high AOD values in North Africa and Middle East, Central Africa, and East Asia (e.g., China) and North America were mainly from dust, smoke, and pollution, respectively (Ichoku, Kaufman, Remer, & Levy, 2004). The standard errors of spatial predictions show that the errors of the experiment using the eight-day AOD measurements (Fig. 4G) were lower than the errors of the experiment using the one-day AOD measurements (Fig. 4H). In addition, the prediction uncertainties in some of the low and middle latitudes with fewer measurements were higher than the prediction uncertainties in the regions with dense measurements. These were reasonable because there were few measurements to constrain spatial predictions in the low and middle latitudes.

4.3. Moving-window GIM versus spatial statistical data fusion

In order to gain insights into the relative performance and computational efficiency of moving-window GIM, we compared this approach with spatial statistical data fusion (SSDF; Nguyen, Cressie & Braverman, 2012). SSDF is a recently developed geostatistical approach for merging large spatial data with variable resolutions. SSDF was developed based on FRK, and it uses a data dimension reduction approach to improve computational efficiency. For the model comparison, we selected the subsets of the eight-day AOD measurements (i.e., from July 28, 2008 to August 4, 2008) from the MISR and MODIS sensors within the latitudes 20° N–60° N and the longitudes

Fig. 2. Synthetic images: (A) and (B) are the images without and with nugget effect, respectively. These two images were generated by unconditional simulations. (C) is the image generated by FRK.
The numbers of pixels for the MISR AOD and MODIS AOD measurements within the subsets were 4412 and 1477, respectively. In order to evaluate both short-range and long-range predictions, we used two sampling schemes here. First, we randomly selected 10% of the AOD measurements within the subsets and treated them as testing data. The remaining 90% of the measurements were used for spatial predictions. This was for evaluating short-range predictions. Second, we randomly selected a test region within the latitudes 30° N–40° N and longitudes 90° E–120° E (i.e., 10% of the total selected subset area). The numbers of pixels for the MISR AOD and MODIS AOD measurements within the test region were 397 and 159, respectively. This was for evaluating long-range predictions (e.g., big gaps in remote sensing images).

In the two experiments, spatial predictions were made at the resolution of 0.25° × 0.25°. For SSDF, the number of bi-square basis functions used was 340 (i.e., the resolution number was four). For moving-window GIM, we used the uniform distribution function to calculate the relationship matrix (i.e., the $H$ matrix in GIM). The spatial predictions were averaged back to the spatial resolutions of the AOD measurements for cross-validations. In the first experiment, the values of RMSE for moving-window GIM and SSDF were 0.3519 and 0.8106, respectively. In the second experiment, the values of RMSE for moving-window GIM and SSDF were 0.4973 and 0.8112, respectively. The value of RMSE for moving-window GIM increased significantly in the second experiment. This was reasonable because spatial predictions made by moving-window GIM were based on nearby measurements in local-windows, and SSDF used all of the available data for spatial predictions. The values of RMSE for the two models may vary across different test regions because the availabilities of the MODIS AOD and MISR AOD measurements vary across space (Fig. 4A–D). SSDF was more computationally efficient than moving-window GIM. For each of the above two experiments, SSDF took about 117 s on a 2.8 GHz machine with an Intel Core Processor, and moving-window GIM took about 212 s.

### Table 1

Results of merging multi-resolution synthetic images.

<table>
<thead>
<tr>
<th>Experiments with aggregated images</th>
<th>RMSE index</th>
<th>Prediction uncertainty</th>
<th>Percentage index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
<tr>
<td>Images with spatially stationary characteristics</td>
<td>Images without the nugget effect</td>
<td>Images with the nugget effect</td>
<td>Images with spatially nonstationary characteristics</td>
</tr>
</tbody>
</table>

Fig. 3. Results of merging synthetic data with variable resolutions: (A), (E), and (I) are overlays of synthetic images with different resolutions for Experiments 1, 3, and 5, respectively; (B), (F), (J) are spatial predictions with data from (A), (E), and (I), respectively; (C), (G), and (K) are prediction uncertainties for (B), (F), and (J), respectively; and (D), (H), and (L) are conditional simulations for (B), (F), and (J), respectively.
6 min on a 16-processor workstation (i.e., the computing process was parallelized). These were reasonable because SSDF eliminates some components of spatial variability to improve computational efficiency.

5. Discussion and conclusions

We have demonstrated GIM and moving-window GIM for image downscaling and multi-resolution data fusion on synthetic and real images. There were several limitations in the above experiments. First, we did not deal with measurement errors of remote sensing images. Measurement errors vary with spatial resolutions of remote sensing images, and they also vary for data derived from different sensors. For the experiments on synthetic and real images, we treated nugget effect as fine-scale variability, and we built geostatistical models to model it. However, real remote sensing images are always contaminated with measurement errors. Future work needs to examine the influence of measurement errors on spatial predictions. Fortunately, GIM can account for measurement errors by the model-data mismatch matrix (i.e., the $R$ matrix in GIM). Second, we applied the Gaussian random field model for both synthetic and real images. In some cases, measurements may have non-Gaussian characteristics. In those occasions we need to extend the geostatistical inverse model to accommodate non-Gaussian spatial data by using a set of models known as generalized linear mixed models. The actual probability distributions of measurements need to be explored in order to define the link function of generalized linear mixed models. Third, we used exponential covariance functions for both synthetic and real data modeling experiments. However,
and spatial statistical data fusion.

Other studies found that including auxiliary variables in GIM can improve prediction accuracy (Gourji, Mueller, Schaefer & Michalak, 2008). However, including inappropriate auxiliary information may bias the model. Usually, a statistical test of the contributions of auxiliary variables is needed (Gourji, Mueller, Schaefer & Michalak, 2008).

We used REML for estimating covariance parameters related to the change-of-support problem. The estimated covariance parameters always have errors, and these errors can consequently propagate to spatial predictions. Therefore, analyzing error propagations in spatial predictions is very important for understanding the accuracy of spatial predictions. By running modeling experiments, we found that the errors associated with the estimated covariance parameters increased with the increase in the resolution gap between measurements and spatial predictions. More in-depth analyses about error propagations in spatial predictions is very important for understanding the accuracy of spatial predictions. Therefore, analyzing error propagations in spatial predictions is very important for understanding the accuracy of spatial predictions.

Acknowledgements

The work was conducted with the financial support from NASA Land-Cover and Land-Use Change Program (NNX09AK87G). The authors would like to thank the science teams for producing and sharing the products of MODIS AOD and MISR AOD. The authors also would like to thank Dr. Hai Nguyen from the Jet Propulsion Laboratory, California Institute of Technology, and Dr. Emily Kang from University of Cincinnati for providing the sample codes related to fixed rank kriging and spatial statistical data fusion.

References


