

Principles and methods of scaling geospatial Earth science data

Yong Ge^{a,b*}, Yan Jin^{c,d}, Alfred Stein^e, Yuehong Chen^f, Jianghao Wang^a, Jinfeng Wang^a, Qiuming Cheng^g, Hexiang Bai^h, Mengxiao Liu^{a,b}, Peter M. Atkinsonⁱ

^a State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences & Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; gey@lreis.ac.cn; wangjh@lreis.ac.cn

^b University of Chinese Academy of Sciences, Beijing, 100049, China; liumx@lreis.ac.cn

^c School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; jinyan@njupt.edu.cn

^d Smart Health Big Data Analysis and Location Services Engineering Lab of Jiangsu Province, Nanjing 210023, China

^e Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, 7500 AE Enschede, The Netherlands; a.stein@utwente.nl

^f School of Earth Sciences and Engineering, Hohai University, Nanjing 210098, China

^g State Key Lab of Geological Processes and Mineral Resources, China University of Geosciences, Beijing 100083, China

^h School of Computer and Information Technology, Shanxi University, Taiyuan, Shanxi 030006, China

ⁱ Lancaster Environment Center, Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK; pma@lancaster.ac.uk

*Corresponding author: E-mail address: gey@lreis.ac.cn (Y. Ge); Tel: +86 10 64888053; Fax: +86 10 64889630

Abstract

The properties of geographical phenomena vary with changes in the scale of measurement. The information observed at one scale often cannot be directly used as information at another scale. Scaling addresses these changes in properties in relation to the scale of measurement, and plays an important role in Earth sciences by providing information at the scale of interest, which may be required for a range of applications, and may be useful for inferring geographical patterns and processes. This paper presents a review of geospatial scaling methods for Earth science data. Based on spatial properties, we propose a methodological framework for scaling addressing upscaling, downscaling and side-scaling. This framework combines scale-independent and scale-dependent properties of geographical variables. It allows treatment of the varying spatial heterogeneity of geographical phenomena, combines spatial autocorrelation and heterogeneity, addresses scale-independent and scale-dependent factors, explores changes in information, incorporates geospatial Earth surface processes and uncertainties, and identifies the optimal scale(s) of models. This study shows that the classification of scaling methods according to various heterogeneities has great potential utility as an underpinning conceptual basis for advances in many Earth science research domains.

Keywords: scaling, change-of-support, autocorrelation, heterogeneity

1. Introduction

1.1 Scale

Scale is a fundamental concept in Earth sciences and other disciplines concerned with spatial patterns and processes (e.g., ecology, soil science, hydrology, and climatology). The patterns and processes of geographical phenomena have an implicit scale of variation (Fu, 2014; Wu, 2007). Against geospace and time, scale can be regarded as an alternative dimension to interpret the patterns and processes of geographical phenomena and to understand the external features and internal mechanisms of the world. Scale has multiple definitions (Bierkens et al., 2000; Goodchild, 2011), but is most commonly used to refer to the size of a geographical process or phenomenon (Atkinson and Tate, 2000). Such a geographical process produces a pattern at the Earth's surface or sub-surface, which constrains changes due to the process within a certain space range and vice versa (Fu et al., 2011). However, scale can also refer to the spatial or temporal extent of an investigation, and to the typical spatial resolution or event frequency, corresponding to the size or extent of a phenomenon or process (Atkinson and Tate, 2000; Bierkens et al., 2000; Zhang et al., 2014a). In disciplines such as agrohydrology, environmental science, and ecology, the scale of investigations can be considered as the measurement scale (Lloyd, 2014) or observational scale (Bierkens et al., 2000), and the scale of phenomena or processes as the intrinsic or process scale (Schulze, 2000; Wu et al., 2006). Scales have also been classified into different types, and other kinds of scales have been identified, including the model scale, optional scale, and policy scale (Bierkens et al., 2000; Schulze, 2000). Subsequently, Li and Cai (2005) classified these scales into two categories in geography: intrinsic scales hidden in natural entities, processes and patterns and non-intrinsic scales which are artificially imposed and do not exist in nature.

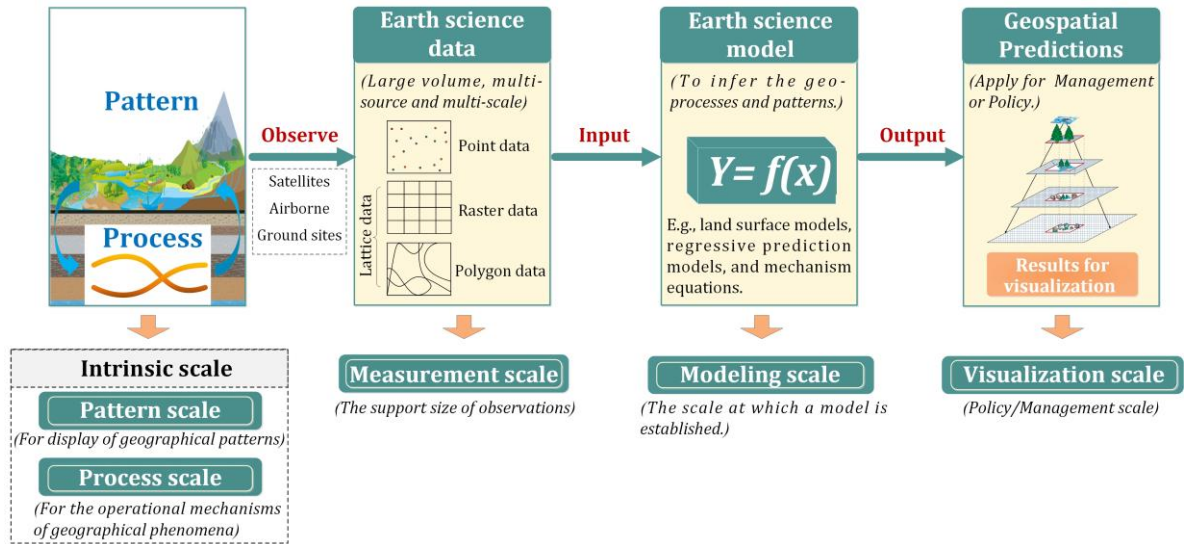


Fig. 1. A conceptual model illustrating the major categories of scales in Earth sciences and how they interrelate.

Taking the above studies as our point of departure, we classify scales into two major categories. One includes pattern scale and process scale, which refers to the scales of intrinsic patterns and processes of geographical phenomena. The other includes measurement scale, modeling scale and visualization scale, which refers to the observations, modeling, analysis and visualization to explore the patterns and processes. In other words, we need appropriate models based on the measurements to approximate the pattern and process at a certain scale. A scale analysis explores these patterns and processes, for example employing geospatial predictions that result from inputting the observations, those could be Earth science data, for example, including point data (quasi point observed data or area data represented as a point) and lattice data (i.e., regular raster data and irregular polygon data) that are used in an Earth science model (Fig. 1). The measurement scale defines the window through which we observe the world, realizing that the closer we look at the world, the more detail we can see. The most fundamental measurement scale (i.e., the support size), refers to the area or volume over which a measurement is taken to study spatial phenomena. Examples are the instantaneous point-scale observations of a wireless sensor network (Ge et al., 2015b; Yang et al., 2013) and the daily remotely

sensed products at 40-km spatial resolution ([Kerr et al., 2001](#)) for measuring soil moisture. The pattern scale is the scale for the display of geographical patterns, whereas the process scale refers to the operational mechanisms of geographical phenomena. To obtain the essentials of patterns and processes, models that reflect the processes and results into patterns can be used. Such process-based models are constructed using a mechanistic equation. The model scale refers to the scale at which a model is established. As an example, we consider a land surface model that provides daily evapotranspiration at 25-km spatial resolution ([Sun et al., 2017](#)), but subject to Richards' equation in modelling soil moisture dynamics at site scale or finer ([Western et al., 2003](#)). The visualization scale is the scale that is most useful for decision making, such as for crop yields on a country level per year ([Hulme et al., 1999](#)) and disease mortality at the city level ([Wang et al., 2017](#)).

1.2 Scaling

Geographical elements, which could be soil, vegetation and river, couple geographical patterns to geographical processes. Geographical variables are then quantitative or qualitative descriptions of these geographic elements, such as heavy metal content of surface soil, landslide hazard index and land cover as applied in land survey. On the basis of Earth observations and model simulations, predictions of geographical variables at different scales can be achieved by scaling, whereas visualizations can be constructed at various scale levels ([Fig. 2](#)). The measurement of geographical variables must occur at a given scale level, whereas modeling may be undertaken at another scale level. [Figure 2](#) shows the measurements for processes and patterns in Earth sciences involving the various scale categories.

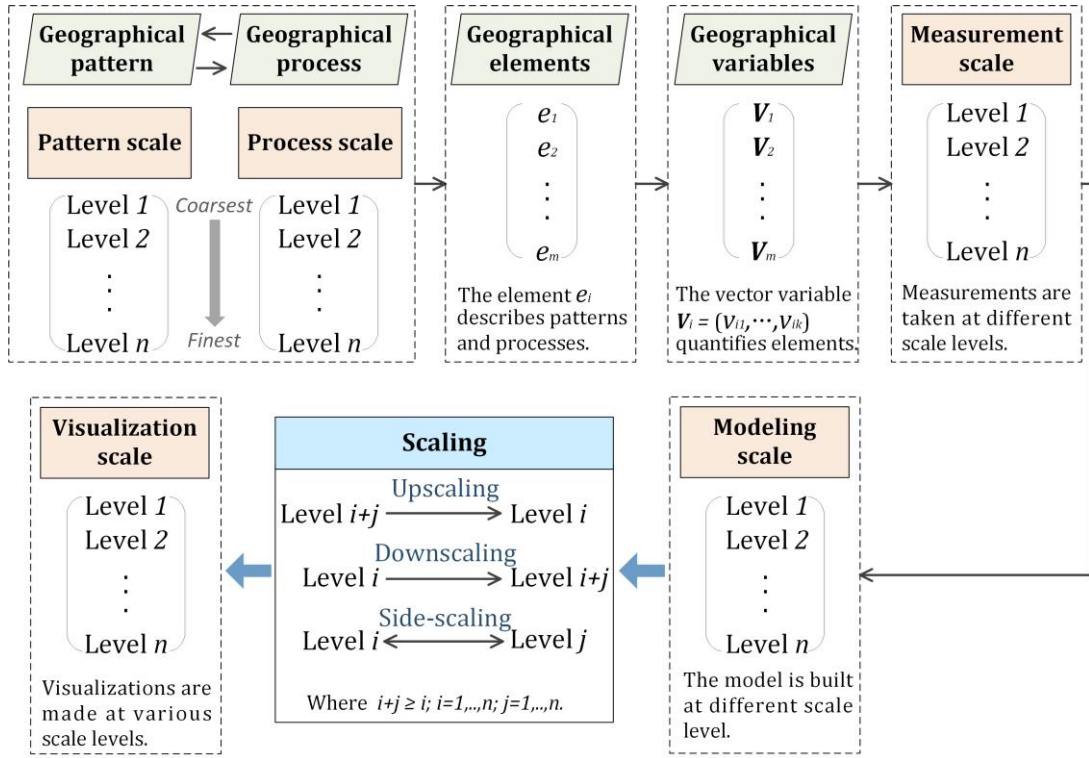


Fig. 2. Measurements of geographical variables and inference of geographical processes and patterns in Earth

sciences. From coarse to fine; the scale levels are labeled 1, 2, ..., n . (The n in various scales might be different.)

The observation of geographical variables, to create Earth science data, plays a crucial role in developing our understanding of geographical processes and in making predictions of geographical variables and their patterns. The spatial characteristics of a geographical variable can be manifested through scale-dependence and scale-independence (Cheng, 1999; Li et al., 2009). From a data viewpoint, scale-dependence makes the values of a geographical variable change with measurement scale whereas scale-independence indicates that the geographical pattern has scale-invariant properties, such as invariant spatial structure. Hence, Earth science studies often suffer from a scale discrepancy between data sources and models. As early as 1950, Robinson (1950) pointed out that ecological information cannot be used as a substitute for individual information, effectively highlighting the issue of scale mismatch in Earth sciences. A scale discrepancy causes a change in the spatial distribution of variables and their spatial relationships. In addition, the scales of multi-source Earth observation data

are often different (Li and Zhou, 2012), and these scales rarely satisfy the required scale of interest for modeling, understanding, or visualization. Scaling is, thus, required to obtain information at the scale of interest (Jarvis, 1995).

Scaling is the transfer of information between scales (Bierkens et al., 2000; Goodchild, 2011). According to the sizes of the scales before and after scale transfer, scaling refers to upscaling, side-scaling, or downscaling (Gotway and Young, 2007; Zhang et al., 2014b). Upscaling or aggregation predicts data or information at a coarse (spatial) or long (temporal) scale from data available at a fine (spatial) or shorter (temporal) scale (Atkinson and Tate, 2000; Ge et al., 2015a). Downscaling or disaggregation predicts data or information at a fine (spatial) or short (temporal) scale from data available at a coarser (spatial) or longer (temporal) scale (Atkinson, 2013). Side-scaling maintains the scale levels of the available and required data and information, and is used, for example, when obtaining area-to-area predictions (Gotway and Young, 2007; Kyriakidis, 2004; Young and Gotway, 2007). In the same study area, if the blocks in side-scaling are overlapping, no block is characterized as coarse or fine. However, when the blocks are nested, side-scaling can involve downscaling or upscaling based on block size, which is helpful to conceptualize side-scaling as relating to both coarse block-to-fine block scaling and fine block-to-coarse block scaling.

Scaling research has permeated into various disciplines, and scaling methods have appeared in hydrology (Lanza et al., 1997), environmental science (Stein et al., 2001), meteorology (Berrocal et al., 2012), ecology (Peterson, 2000), agriculture (Folberth et al., 2019), geography (Malenovsky et al., 2007) and epidemiology (Galukande et al., 2013). Statistical scaling models (Wood et al., 2004) and dynamic scaling models (Dai et al., 2017) have been applied to climate parameters, and the scale effects of physical models and algorithms have already been discussed at the pixel scale in remote sensing applications (Wu and Li, 2009). Models for scaling emphasize the use of information provided

by correlated variables ([Anderson et al., 2003](#); [Keil et al., 2018](#)). Serious challenges facing scaling in hydrology ([Wigmosta and Prasad, 2006](#)) are caused by the heterogeneity of geographical variables, scale dependence of parameters and geospatial geo-Earth surface processes, and incomplete knowledge of the Earth surface system. Although some reviews of scaling have addressed the spatial downscaling of satellite remotely sensed irrigation management ([Ha et al., 2013](#)) and soil moisture ([Peng et al., 2017](#)), there has been little research on the choice of an appropriate scaling method in a given context. [Bierkens et al. \(2000\)](#) and [Finke and Bierkens \(2002\)](#) have addressed several strategies in environmental research. The existence of scale-independent and, in particular, scale-dependent variation makes the scaling processes problematic ([Cheng, 1999](#)). The characteristics of some variation (e.g., spatial autocorrelation) also become irrelevant or undetectable at a certain scale ([Atkinson and Tate, 2000](#)). Moreover, spatial scaling might reduce to a simple averaging process for homogeneity, which would become more complex under heterogeneous conditions. [Schulze-Makuch and Cherkauer \(1998\)](#) demonstrated that the scale dependence of hydraulic conductivity would be affected by the heterogeneities within the aquifer. Heterogeneities have a significant influence on the distribution of geographical variables and geospatial geo-Earth surface processes ([Wu et al., 2006](#)). Therefore, in this paper, the spatial properties of Earth science data are discussed in detail.

The aim of this paper was to provide a comprehensive review of scaling methods from an Earth science perspective based upon the scale-independent and scale-dependent properties of spatial variables. This review is intended to assist the development of multi-scale Earth science data, which may benefit regional applications in hydrology, meteorology, climatology, and water resource management. The remainder of this paper is organized as follows. Section 2 discusses the spatial properties that reflect certain characteristics with changing scales. Section 3 describes the scaling issue, including the specific expression of scaling. Section 4 presents several upscaling methods, before

Sections 5 and 6 introduce downscaling and side-scaling methods. The greatest challenges and most pressing problems in scaling are discussed in Section 7. Section 8 summarizes the conclusions to this Review.

2. Spatial properties

Geographical variables may exhibit scale-dependent and scale-independent variations in spatial, temporal, and spatiotemporal dimensions. The spatial properties consist of scale-invariance and scale-variation. Scale-invariance refers to self-similarity in scaling processes. [Halley et al. \(2004\)](#) indicated that a series of scale-invariant metrics for aspects of ecology that are fractal in nature have been developed by landscape ecologists. However, no scale-independent method of analysis is widely agreed upon for spatial statistics. It is rare for studies of scale-independent properties to consider the issue of scaling. Scale-variation means that the properties of geographical variables exhibit differences, such as different spatial distribution functions at different scales. Almost all environmental processes are scale-dependent ([Atkinson and Tate, 2000](#)). The question of scale-dependence is important, for example, in hydraulic conductivity studies, that typically do not scale in an additive manner ([Renard and Marsily, 1997](#); [Mahmud et al., 2015](#)). These two spatial properties are the foundations for achieving accurate scaling.

2.1 Scale-independent variations

Because of the scale-independent property of geographical variables, there is a common phenomenon in nature named self-similarity ([Cheng, 1999](#)) which refers to similar processes or patterns of geographical phenomena occurring at different scales ([Wu et al., 2006](#); [Zhang and Li, 2012](#)). Current Earth science data make it possible to explore the self-similarity of geographical variables, which is the key idea in fractal theory ([Mandelbrot, 1967](#)).

Statistical self-similarity may exist in local and global dimensions in terms of form, time, and space. For example, the statistical parameters of climatic series at one scale can be applied to another scale through a certain scaling relationship according to the principle of self-similarity (Chen et al., 2006). The degree of complexity can be described by the fractal dimension. This fractal dimension is used to describe scale-invariant features, which can be positive, negative, fuzzy, and multifractal. Also Zhou et al. (2008) have utilized an area weighted fractal dimension index to analyze the spatial pattern of land cover change. Conventionally, one parameter is used to model the self-similarity through a simple fractal, although sometimes it is better to employ a multifractal, which could be seen as an generalization of fractals (Gires et al., 2018; Xu et al., 2015). Multifractal spectra, structured by different fractal dimensions, are often applied to describe multifractals. There are two classes of multifractals according to the characteristics of their fractal dimension spectra. Continuous multifractals are intertwined spatially with a continuous spectrum of dimensions (Stanley and Meakin, 1988), whereas discrete multifractals have a discrete fractal spectrum (Cheng, 1997). Fractals and multifractals have been applied widely for information extraction (Ge et al., 2006; Hu et al., 2013), and several methods that combine fractal or multifractal dimensions and variograms have been proposed for the interpolation of mineral distributions and the analysis of remote sensing imagery (Cheng, 2005). Fractal geometry applications have been investigated in ecology to describe and analyze patch patterns and patch dynamics (Li, 2000). In many geophysical fields, fractal phenomena have been shown to exist in cloud formations, river networks, and turbulence. Fractal thinking can be employed to deal effectively with scale issues in geography and geospatial analysis (Jiang and Brandt, 2016).

Although self-similarity appears at all levels of scale for fractal geometry, the self-similarity of a specific object is approximate or statistical. That is, similar fractal features are displayed within a certain range (Fu et al., 2011). The fractal features obtained at certain scales cannot be generalized to

others. Moreover, all fractals can be reflected in a complex structure (Cheng, 1999) formed by a series of scales ranging from the finest to the coarsest (Jiang, 2018; Jiang and Ma, 2018). The scale-independent parts of the prediction of geographical variables can be obtained at the target scale using self-similarity theory, but some scale-dependent variation will still exist. Thus, scale-dependent variation should be considered.

2.2 Scale-dependent variations

Scale-dependent variation exhibits different properties at different scales (Lloyd, 2014). Spatial autocorrelation and heterogeneity are two general properties of Earth science data (Goodchild, 2011), describing the concentration and decentralization of geographical phenomena. These properties can help to infer the interactions between underlying processes and their responses at different scales. The details of spatial autocorrelation and heterogeneity are explained below.

2.2.1 Spatial autocorrelation

The first law of geography states that spatial events are influenced by neighboring events and are more closely related to nearby events; this is known as Tobler's First Law of geography (Tobler, 1970). It refers to the spatial autocorrelation of Earth science data, which means that similarity decays as a function of distance. Moran's I , Geary's C , and the semivariogram are common indices and functions for quantifying the degree of spatial autocorrelation.

Moran's I and Geary's C indices can be calculated by comparing the values at one location with the values at neighboring locations (Geary, 1954; Moran, 1950). In Geary's C , the difference between locations is the discrepancy of each observation; the equivalent value in Moran's I is the difference from the mean. However, both Moran's I and Geary's C give one value in each fixed zone. This provides limited scope to capture variation in spatial autocorrelation with distance. The semivariogram, or variogram, is a function of lag (i.e., distance and direction vector, but commonly reduced to distance

for omnidirectional cases) that is commonly used to represent spatial dependence by measuring the changes in semivariance with spatial distance. The semivariogram is the basis of geostatistics, and is known to be a useful tool for analyzing spatial data in ecology and environmental science (Rossi et al., 1992). Furthermore, as a statistical tool for the analysis of dependence and variability in space, the semivariogram is also applicable to spatiotemporal variability.

Because semivariograms consider only two-point correlations, multipoint statistics have been proposed for simultaneously computing correlations for more than two points (Guardiano and Srivastava, 1993; Journel, 1993). This idea borrows the spatial arrangements of values from training images and uses them to retrieve conditional probabilities for the simulation (Ge and Bai, 2011; Mariethoz et al., 2010). Many algorithms have been developed for multipoint simulations, such as SNESIM, FILTERSIM, and IMPALA (Strebel, 2002; Tahmasebi et al., 2012; Zhang et al., 2006), each of which effectively captures complex structural information over different scales. The multipoint simulations are sensitive to implementation-specific parameters and these parameters might be different depending on the method used (Bai et al., 2016). The number of multiple grids is also critical in achieving acceptable simulation results (Liu, 2006; Strebel, 2002).

2.2.2 Spatial heterogeneity

The second law of geography is the principle of spatial heterogeneity, or non-stationarity in terms of fitted models (Goodchild, 2004). Formally, stationarity indicates that a parameter of a model does not vary with space or time, and hence non-stationarity is the absence of this property (Chiles and Delfiner, 1999). Heterogeneity refers to the inhomogeneity or complexity of spatial distributions for a geographical variable. Thus, both non-stationarity and heterogeneity are related to spatial and temporal variability. In this paper, the generic term “heterogeneity” is employed to characterize the variation in a geographical variable in the spatial and temporal dimensions, whereas the term “non-stationarity”

refers to the spatially or temporally varying parameters of a model fitted to that variation. Spatial heterogeneity changes as the measurement scale varies, which makes the spatial distribution features of geographical variables different. Statistically, heterogeneity means that both the first- and second-order (and indeed higher-order) moments of geographical variables models may need to be non-stationary in space or time, which can cause unbounded variance.

There are different types of heterogeneities (see [Table 1](#)). According to their statistical features, these heterogeneities can be classified as either a tendency, anisotropy, or stratification. Tendency in heterogeneity refers to variation in the mean value for a geographical variable, assuming there is a spatial trend, anisotropy indicates a diversity of variation in a geographical variable in different directions, whereas stratified heterogeneity refers to patch-based variation in the variability in that geographical variable ([Wang et al., 2016](#)). From a different point of view, heterogeneities can be divided into sample heterogeneity, local heterogeneity, and global heterogeneity in terms of their spatial extent. Sample heterogeneity refers to heterogeneity among samples, for example, samples of unequal precision ([Wang et al., 2014a](#)) and heterogeneity within samples (e.g., footprint observations ([Hu et al., 2015b](#)) and mixed pixels ([Ge et al., 2016a](#))). Different heterogeneities are handled differently, and several methods have been proposed, such as using hotspot detection to identify statistically significant spatial clustering of high values (hot spots) or low values (cold spots) ([Ghahramani et al., 2018](#)), employing hierarchical Bayesian methods by deriving the parameter estimates in multiple levels to capture the unexplained variance in the average estimates of parameters ([Zhu et al., 2006](#)), and applying geographically weighted regression to estimate spatially varying regression coefficients ([Jin et al., 2018a](#)). Sometimes, various heterogeneities coexist because of the complex nature of the geographical process. Spatially varying coefficient models in spatial statistics take various forms that allow results to vary spatially ([Gelfand et al., 2003](#)), for example, the spatially

varying coefficients geostatistical model (Hamm et al., 2015). Moreover, in landscape ecology, scale variance analysis has been used to explore landscape heterogeneity at multiple scales, where the spatial data would be reconstructed hierarchically by resampling to detect multiple-scale patterns (Wu et al., 2000).

Table 1
Different types of various heterogeneities

	Types	Description
Statistical feature of heterogeneity	Tendency	Variation of mean value
	Anisotropy	Variations in different directions
	Stratification	Variation of variance
Spatial extent of heterogeneity	Sample heterogeneity	Heterogeneity among or within samples
	Local heterogeneity	Heterogeneity at local scale
	Global heterogeneity	Heterogeneity at global scale

Considering variability as a system property (Gaucherel, 2007), heterogeneity is absolute and ubiquitous, whereas homogeneity is relative (Fu et al., 2011; Wu et al., 2006). That is, if a geographical variable is heterogeneous (which leads to the need for a non-stationary statistical characterization) at a certain scale, then it might be homogeneous at a finer scale (i.e., stationary statistical characterization at a finer scale). The scale-dependent and scale-independent variation provides a bridge among multiple scales, thus confirming the possibility and rationality of scaling.

3. Description of scaling

Scaling (i.e., a change of measurement scale) is a complex process for characterizing phenomena (more precisely, variables) at multiple scales. The fact that the dominant patterns and their underlying processes in Earth sciences vary with scales of time and space make this operation particularly

challenging. It is difficult to recognize the form of the scaling function, although the flows of matter, energy and information across different scales provide objective bases for scaling. For geospatial Earth surface processes, there is commonly some nonlinearity at one scale or among multiple scales, further increasing the complexity of scaling (Jarvis, 1995). Considering that the scale-invariant and scale-varying characteristics discussed in Section 2 allow us to describe geographical phenomena, the corresponding parts are considered to be involved in scaling and the scaling function.

3.1 An expression of scaling (or abstraction of scaling)

Assume that the vectors $\{Z_O = Z_O(\mathbf{s}^a, \mathbf{t}^a); \mathbf{s}^a \in S, \mathbf{t}^a \in T\}$ and $\{Z_D = Z_D(\mathbf{s}^b, \mathbf{t}^b); \mathbf{s}^b \in S, \mathbf{t}^b \in T\}$ are the random variables of a geographical variable Z at the scales before and after the scale transfer. They vary within a spatial domain S and time interval T . The generic formulation for predicting variable Z at the target scale D can be expressed as follows:

$$Z_D = \Phi[f(Z_O), g(Z_O)] + \varepsilon, \quad (1)$$

where ε is a random function with a mean of zero (i.e., random error), the function $f(\cdot)$ is the scale-independent part, and $g(\cdot)$ is the scale-dependent part in the scaling process. The scaling function Φ is the composition of f and g . Let Φ_{\uparrow} , Φ_{\downarrow} , and Φ_{\updownarrow} be the upscaling, downscaling, and side-scaling functions, respectively. In the absence of temporal or spatial information, Φ is a purely spatial scaling or a purely temporal scaling function. The simplest form of Φ is a linear function with equal weights (e.g., the weights of f and g are 1). If we assume that the scale-invariant part is zero, Eq. (1) simplifies as:

$$Z_D = f(Z_O) + g(Z_O) + \varepsilon = g(Z_O) + \varepsilon. \quad (2)$$

Geospatial data can be considered to consist of a trend and residual components. Let $m(Z_O)$ and $r(Z_O)$ be the trend component and the residual component, respectively, of the geographical variable at the original scale, and $m(Z_D)$ and $r(Z_D)$ be the corresponding components at the target scale. The

estimation in Eq. (2) can then be rewritten as:

$$Z_D = m(Z_D) + r(Z_D) = g(m(Z_O) + r(Z_O)) + \varepsilon. \quad (3)$$

It is common to model the trend $m(\cdot)$ using regression analyses by incorporating auxiliary variables, whereas the residual $r(Z_D)$ is often estimated by employing weighted methods. In some scaling studies (i.e., [Jin et al., 2018a](#); [Jin et al., 2018b](#)), the coefficients of $m(Z_D)$ and $m(Z_O)$ were assumed to be the same, and $r(Z_D)$ was calculated using the linear weights of $r(Z_O)$. However, the coefficients and dominant auxiliary variables would be different at various scales, as discussed in Section 7.5.

3.2 Scaling based on scale-independent variation

Fractals represent scale-independent variation and are, therefore, useful for scaling if the data behave this way. The premise for applying fractals for scaling is the presence of fractal features of the geographic variable. Assuming that the geographic variable is characterized by some scale-independent properties, the goal of scaling can be achieved by using these properties to predict the values at other scales. Fractals and multifractals provide a suitable way of describing the scale-independence of Earth science variables (e.g., topography ([Pradhan et al., 2010](#)) and climate variables ([Deidda, 2000](#))). In particular, many fractal and multifractal methods have been used to scale leaf area index ([Jiang et al., 2015](#); [Wu et al., 2015](#)), and downscale satellite-based remote sensing rainfall data ([Espinosa et al., 2017](#); [Tao and Barros, 2010](#)). Unlike the above fractal methods, several scale-independent properties have been used in scaling. [Molnar and Burlando \(2005\)](#) achieved stochastic rainfall disaggregation by preserving the distribution, intermittency and extremes properties of coarse resolution rainfall data in cascade models. Fractal characteristics have also been employed in multiple-point geostatistical simulation to interpolate values at fine scale ([Mariethoz et al., 2011](#)). These fractal-based methods require relatively few parameters, are highly efficient and produce

realistic simulation results. However, the scale-independent property cannot always be applied at any scale, and is often invalid beyond a certain scale.

3.3 Scaling based on scale-dependent variation

Based on the scale-dependent property, many methods have been developed to transfer the scales of Earth science data. Most of them benefit from Tobler's First Law of geography. Numerous scaling methods were developed using the autocorrelation function and spatial weight function, for example, the geostatistical scaling methods (Goovaerts, 2008), spatial statistics and spatial interpolation (Yue et al., 2007). Scaling methods based on scale-dependent variation are described in the following sections. In addition, some methods do not use the spatial properties of the geographical variables (i.e., data-driven approaches). These include general-statistical, machine learning, and probability-based methods.

Bierkens et al. (2000) divided upscaling and downscaling methods in environmental research into four and three major classes, respectively, by using decision trees to determine whether a linear model or a process model should be employed. As scaling methods based on scale-independent variations are rare, we focus mainly on scaling methods in which such variation is not included. In most of the available Earth science research on techniques for scaling, the specific scaling methods fall mainly into one of six categories: general-statistical methods, spatial statistical analysis, machine learning methods, resampling methods, process-based methods, and hybrid methods. The following sections discuss these scaling categories as applied to upscaling, downscaling and side-scaling. The characteristics of different scaling methods, also including fractals, are shown in Table 2.

Table 2

Characteristics of different scaling methods.

Main category	Methods	References	Characteristics	Scaling
Fractal method	Fractal and multifractals	Deidda, 2000; Pradhan et al., 2010; Tao and Barros, 2010; Wu et al., 2015; Jiang et al., 2015; Espinosa et al., 2017	Based on the self-similarity, the fractal methods are high efficient which need only a few parameters. However, the scale-independent property cannot always be applied at any scale.	Upscaling and downscaling
Probability-based scaling	Area proportion, area-weighted average and footprint-weighted method	Beyrich et al., 2006; Serafimovich et al., 2018	It is area-based weighting, with no assumptions about the spatial variability of the aggregation variables.	Upscaling and side-scaling
Resample method	Bilinear interpolation, inverse distance weighted method, function-based interpolation, and High Accuracy Surface Modeling method	Bindlish and Barros, 2000; Muñoz et al., 2001; Steinacker, et al., 2006; Yue et al., 2007; Liu and Zuo, 2012	It is the simplest scaling method which doesn't require any auxiliary information, but ignoring the change of support in scaling.	Upscaling and downscaling
Geostatistical method	Block kriging, area-to-point Poisson kriging, area-to-point kriging, downscaling cokriging and area-to-point regression kriging	Yoo and Kyriakidis, 2006; Goovaerts, 2006; Pardo-Igúzquiza et al., 2006; Atkinson et al., 2008; Kerry et al., 2010; Ling et al., 2013; Wang et al., 2014b; Wang et al., 2015a; Wang et al., 2015b	It considers spatial autocorrelation and the change of support in scaling with or without auxiliary information. However, only two-point correlations are involved and the cross-semivariogram and point semivariogram would make scaling computationally intensive.	Upscaling and downscaling
	Area-to-area kriging, area-to-area Poisson kriging and area-to-area regression kriging	Kyriakidis, 2004; Skøien and Blöschl, 2007; Kerry et al., 2010; Ge et al., 2015a; Parajka et al., 2015		Side-scaling
	Multiple-point geostatistics	Strebelle, 2002; Jha et al., 2013; Tang et al., 2015	It characterizes the spatial structure by more than two points simultaneously. The multipoint simulations are sensitive to implementation-specific parameters.	Downscaling
General-statistical method	Regression methods, canonical correlation analysis and principal component analysis	Landman and Tennant, 2000; Bergant et al., 2002; Rolland, 2003; Friederichs and Hense, 2007; Bergin et al., 2012; Piles et al., 2011; Kang et al., 2018	It establishes relationship between the target variable and auxiliary variables without considering the spatial correlation.	Upscaling and downscaling
Machine learning method	Support vector regression, random forest, regression tree, artificial neural network, deep belief network and patch-pair	Dibike and Coulibaly, 2006; Yang et al., 2006; Nigussie et al., 2011; Landeras et al., 2012; Zhang et al., 2014c; Hutengs and Vohland, 2016; Ling et al., 2016; Song et al.,	Based on a large number of observations, it could generate reliable predictions, with an implicit expression which has less explanatory capability. It also provides prior knowledge of variables from training datasets at different	Upscaling and downscaling

	learning-database	2016	scales.	
Bayesian-based method	Bayesian regression, Bayesian hierarchical model, Bayesian maximum entropy	Lee et al., 2009; Jameson and Heymsfield, 2013; Li et al., 2013; Qin et al., 2013; Song et al., 2015; Zhang and Yan, 2015; Kou et al., 2016	It could integrate various prior information and account for the uncertainty of the estimated parameters. However, it is a challenge to explore the probability of the prior knowledge.	Upscaling and downscaling
Process-based method	Statistical-dynamical downscaling, dynamical model and data simulations	Bindlish and Barros, 2000; Kim et al., 2000; Margulis et al., 2002; Crow et al., 2012; Mechri et al., 2016; Ramzan et al., 2017	The process model has the obvious physical significance, however, it is always complicated and won't remain the same at coarse and fine scales.	Upscaling and downscaling
Super-resolution mapping (SRM)	Fusing subpixel-scale and pixel-scale spatial autocorrelation of categorical variables (e.g., land cover/land use)	Tatem et al., 2001; Atkinson, 2005; Kasetkasem et al., 2005; Mertens et al., 2006; Ge et al., 2009, 2014; Ling et al., 2013; Chen et al., 2014; Wang et al., 2015b; Zhong et al., 2015	The spatial autocorrelation represents the structure of features of categorical variables at different scales. These methods ignore the local details or the structure of coarse patch features.	Downscaling categories
	Modeling spatial distribution patterns of geographical features	Tatem et al., 2002; Atkinson, 2004; Thornton et al., 2007; Xu et al., 2014; Ai et al., 2014; Ge et al., 2016a	The features of categorical variables are divided into areal pattern, point pattern and linear pattern, which are described by building different models. The accurate division of the three patterns is key and prerequisite for SRM.	

Note: The mentioned scaling methods and the references are limited. There are also other and many hybrid methods for scaling. More details can be found in Sections 3, 4, 5 and 6. The downscaling and upscaling methods can also be applied for side-scaling. The super-resolution mapping methods are only used for downscaling categories.

4. Upscaling

This section reviews several potential methods for addressing the upscaling problem. Upscaling algorithms generally use the spatial and temporal variability of *in situ* measurements, as well as auxiliary information, to provide predictions for a given spatiotemporal support (i.e., the average values for a pixel over a given period).

Various upscaling strategies have been proposed (see the excellent review papers by [Gotway and Young \(2002\)](#) and [Gelfand et al. \(2001\)](#)) for disciplines including hydrology process upscaling ([Vereecken et al., 2007](#)), soil process modeling ([Heuvelink and Pebesma, 1999](#)), soil moisture ([Crow et al., 2012](#)) and evapotranspiration ([Liu et al., 2016](#)). The most direct approaches are based on minimizing the prediction error by either optimizing the sampling location of point-scale observations or improving the upscaling function Φ_{\uparrow} . The sampling design will influence upscaling accuracy. Thus, prior to upscaling samples for a given support, it is essential to develop efficient procedures for designing information-effective sampling. Spatial sampling methods for sampling design fall into one of the following categories: geometry-based, probability-based and model-based sampling ([de Gruijter et al., 2006](#)). Among these methods, the model-based ones are generally more efficient, especially when the target has a significant spatial structure. As spatial sampling optimization is beyond the scope of this review, we focus mainly on the upscaling function Φ_{\uparrow} . Theoretically, Φ_{\uparrow} can be divided into five categories: probability-based, regression model-based, geostatistical model-based, machine learning model-based, and process model-based methods.

4.1 Probability-based upscaling methods

Probability-based upscaling assumes that every observation has a known and equal chance of being selected as a sample into the upscaling model. It can only be used when at least one point value

is given in each block and when the point locations are derived from a probabilistic sample design (Heuvelink and Pebesma, 1999), as mentioned before. The simplest way to scale up sparse ground-based measurements is to define Φ_{\uparrow} as a simple linear average, that is, $\hat{Z}_D = \Phi_{\uparrow}(Z_O) = N^{-1} \cdot \sum_{i=1}^N Z_O(\mathbf{s}^i, \mathbf{t}^i)$, where N is the sample number. Here, all point measurements are given equal weight. This would be optimal if the observation locations had been selected using a probability-based sampling scheme. In practice, the efficiency of this method can be increased by weighting (e.g., the area-weighted average (Beyrich et al., 2006) and footprint-weighted method (Serafimovich et al., 2018)). An important property of probability-based approaches is that they make no assumptions about the spatial variability of the aggregation variables, which would be limited by the default probabilities of the samples involved in the scaling.

4.2 Regression model-based methods

When the target variable Z exhibits strong heterogeneity in the mean (i.e., trend heterogeneity), additional relative information should always be incorporated to capture the spatiotemporal variability, especially when covariates are available with continuous spatial coverage. In practice, the spatial regression model (linear or nonlinear) is commonly fitted between sparse ground observations and multisource remote sensing information related to the target variable. Various regression model-based upscaling methods have been proposed, such as ordinary least squares, Bayesian linear regression and ridge regression (Kang et al., 2018). For example, Qin et al. (2013) developed an upscaling algorithm based on Bayesian linear regression for sparse ground observations, which did not need to measure the representativeness of ground-based sites. It has been applied in soil moisture upscaling (Zhao et al., 2014; Kang et al. 2017). However, the relationship between ground observations of target variable and pixel auxiliary variables, which might be fitted at mismatched scales, would lead to inaccurate upscaled samples. The validity of the regression coefficient and the selection of regression variables

also require further exploration, discussed in Section 7.5.

4.3 Geostatistical model-based methods

When the observations are spatially autocorrelated, block kriging (BK) (Goovaerts, 2001) can be used to derive the optimal weights w_i for each observation, thereby improving the estimates of $\hat{Z}(B)$ by $\hat{Z}_D = \Phi_{\uparrow}(Z_O) = \sum_{i=1}^N w_i \cdot Z_O(\mathbf{s}^i, \mathbf{t}^i)$. Theoretically, BK will provide the best linear unbiased estimate, as it takes the spatial autocorrelation into account. However, this unbiased estimate is based on the assumption that all observations are error-free or have the same random measurement errors. Several upscaling methods have been proposed to overcome the problem of observations bearing unequal precision instrument error (Wang et al., 2014a). For a spatiotemporal process, the upscaling strategy can benefit from the spatiotemporal characteristics of the point-support observations, as well as from related auxiliary information such as spatiotemporal regression BK (Wang et al., 2015a). The upscaling of area or footprint observations to areal blocks is a form of side-scaling, and is thus described in Section 6. These geostatistical methods have been applied widely in the scaling process, but the stationarity hypothesis in the above geostatistical model-based methods would be unsupported by the occurrence of heterogeneity.

4.4 Machine learning model-based methods

Upscaling methods based on machine learning include support vector machines (Yang et al., 2006), artificial neural networks (Landeras et al., 2012), regression trees (Jung et al., 2011), random forests (Shortridge et al., 2016), and deep belief networks (Song et al., 2016). These methods have been used in soil moisture and surface flux upscaling. Comparison of the results in these references (i.e., Yang et al., 2006; Landeras et al., 2012; Jung et al., 2011; Shortridge et al., 2016; Song et al., 2016) also shows that machine learning model-based methods can yield more accurate upscaling than traditional regression- and geostatistical model-based methods. However, these data-driven approaches

also have limitations. For example, machine learning methods can only generate reliable predictions for conditions with a large number of observations, and this situation can be difficult to achieve in practice.

4.5 Process model-based methods

In cases where the sample observations are not available, the upscaling function Φ_{\uparrow} can be established using land surface models that incorporate process mechanisms. The application of process model-based methods assumes that spatially distributed forcing data are fed into the model and that the subsequent model physics acting upon these data can capture accurately the processes generating spatial variability (Crow et al., 2012). If this assumption holds, then the process model can help to define an appropriate Φ_{\uparrow} function (e.g., soil moisture model (Crow et al., 2005) and Priestley–Taylor equation (Liu et al., 2016)). The process model has obvious physical significance, but it is always complicated and is unlikely to be consistent at coarse and fine scales, as discussed in Section 7.3.

5. Downscaling

Downscaling is more complicated than upscaling, which presents an ill-posed problem with a non-unique solution. It can be divided into downscaling continua or downscaling categories based on whether a continuous or categorical variable is predicted (Atkinson, 2013). Although the same methods may be applied to predict continuous or categorical variables, the various downscaling methods can be classified in different ways for these two goals.

5.1 Downscaling continua

Continuous Earth science variables, such as temperature, have an infinite number of possible values between two given points. If the distribution function of Earth science data can be obtained at the desired scale, it is easy to predict the required values. Because the distributional characteristics

vary with decreasing scale, they must be deduced from the information at the original scale. However, in many cases, the target geographic variable at the fine scale cannot be fitted well using only a single information source. In downscaling continua, the first consideration is whether or not to use auxiliary data; auxiliary information can be employed to explain some of the unknown temporal or spatial variation of the target geographic variable. Specifically, fine-scale auxiliary information can aid in exploring the properties at a fine scale. The next consideration is whether or not to involve a process model in downscaling. The process model is one of the main ways to estimate the geographical variable at finer scales. Although the process model is available at a coarse scale, it is still a huge challenge to explore the form and parameters of geographic processes at a fine scale. Therefore, the downscaling continua methods can be divided into downscaling without and with fine-scale auxiliary information, and downscaling without and with a process model. Most specific downscaling methods fall into the above classes. [Figure 3](#) shows the major classes of downscaling continua methods. This schema provides guidance for choosing from the many available downscaling methods in practical applications.

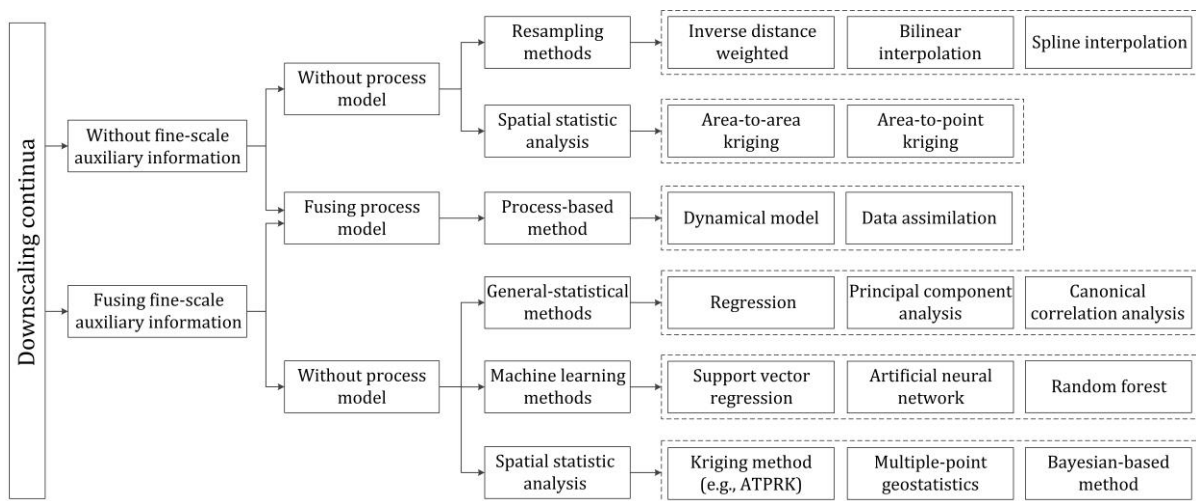


Fig. 3. Major classes of methods for downscaling continua.

5.1.1 Downscaling without fusing both fine-scale auxiliary information and process model

For downscaling without fusing both fine-scale auxiliary information and a process model, several methods have been developed to downscale continuous Earth science data: the resampling method and spatial statistical analysis.

(a) Downscaling using resampling methods with no other information

The simplest downscaling method is the resampling method, which refers mainly to interpolation here. Fine-scale Earth science data can be predicted by interpolating and extrapolating samples. Interpolation methods are used commonly to predict meteorological variables at a fine scale, such as bilinear interpolation ([Bindlish and Barros, 2000](#)), inverse distance-weighted methods ([Liu and Zuo, 2012](#)), and function-based interpolation ([Steinacker et al., 2006](#)). Other interpolation methods, such as finite differences ([Muñoz et al., 2001](#)) and the High Accuracy Surface Modeling method ([Yue et al., 2007](#)) can also predict fine-scale information. However, the above interpolation methods do not account explicitly for the change of support in downscaling. The downscaled predictions may not, therefore, have sufficiently high accuracy.

(b) Downscaling using spatial statistical analysis with no other information

Considering the spatial correlation among Earth science data, spatial statistical analysis has been applied in downscaling of continuous variables, involving spatial and spatiotemporal interpolation. Simple kriging and ordinary kriging, as with the spatial interpolation methods mentioned above, cannot solve the modifiable areal unit problem in downscaling. To solve this challenge, area-to-point kriging (ATPK) and area-to-area kriging (ATAK) have been developed ([Gotway and Young, 2002](#); [Kyriakidis, 2004](#)). Both ATPK and ATAK provide predictions on a finer support (point or area) than those of the original data, ensuring the coherence of predictions; for example, the sum of downscaled predictions within any area is equal to the original aggregated count. The ATPK method has been

further improved by taking into account the inequality constraints in spatial interpolation (Yoo and Kyriakidis, 2006; Yoo et al., 2010) and developing area-to-point Poisson kriging by considering sample heterogeneity (Goovaerts, 2006; Kerry et al., 2010).

Temporal downscaling approaches, such as the wavelet transform and Fourier transform (Martinez and Gilabert, 2009; Wegener et al., 2012; Zhang et al., 2016), which transform a spatial domain image into the frequency domain or wavelet domain, have been applied on spatial downscaling through a decomposition and reconstruction process (Mascaro et al., 2013). These downscaling methods preserve the waveform structure characteristics, but are highly susceptible to the accuracy of pixel matching.

5.1.2 Downscaling fusing fine-scale auxiliary information

Downscaling fusing fine-scale auxiliary information, but without a process model, usually involves a general-statistical method, machine learning method, or spatial statistical analysis. The idea behind these methods is to establish one or more relationships (ignoring or considering structure information) between the coarse-scale target variable and the fine-scale auxiliary variables.

(a) Downscaling using general-statistical methods with auxiliary information

Based on the scale-invariance of the model between the target variable and auxiliary variables at different scales, the general-statistical method and machine learning method can be used for spatial or temporal downscaling. The general-statistical methods, including principal component analysis, canonical correlation analysis, and regression methods, have seen a huge growth in application for the downscaling of continuous variables (Bergant et al., 2002; Landman and Tennant, 2000; Rolland, 2003). Many regression models (both linear and nonlinear) have been applied for downscaling, such as the multivariate linear model (Murphy, 2000), polynomial regression (Piles et al., 2011), ridge regression (Cannon, 2009), quantile regression (Friederichs and Hense, 2007) and generalized linear

regression ([Bergin et al., 2012](#)).

(b) Downscaling using machine learning methods with auxiliary information

When the explicit expression in the general-statistical method does not produce good fitting results, an alternative is the machine learning method. The machine learning method may be considered as a branch of nonlinear regression, and has strong potential for complex, nonlinear, and time-varying input–output mapping. The three most common methods in downscaling are support vector regression ([Dibike and Coulibaly, 2006](#)), artificial neural networks ([Hoai et al., 2011](#)), and random forests ([Hutengs and Vohland, 2016](#)). The lack of explanatory capabilities is a significant drawback for machine learning methods because of their implicit expression.

(c) Downscaling using spatial statistical analysis with auxiliary information

To consider the spatial or temporal correlation of variables or among variables, spatial statistical analysis techniques, including the kriging method, multi-point geostatistics and methods using Bayesian-based inference framework, have been developed. Here, the geospatial data can be considered to consist of trends and stochastic residuals, with both components being scale-dependent. Methods such as downscaling cokriging ([Atkinson et al., 2008](#); [Pardo-Igúzquiza et al., 2006](#)), area-to-point residual kriging ([Liu et al., 2008](#)), ATPK with external drift ([Yoo and Kyriakidis, 2009](#)), and area-to-point regression kriging ([Kerry et al., 2012](#); [Wang et al., 2015b](#)) have been proposed for downscaling. In these regression-based downscaling methods, the different regression models generally reproduce the mean of the predictions, with the probabilistic approaches recapturing the variability of the process by building the whole distribution ([Fasbender and Ouarda, 2010](#)). However, the required calculation of the cross-semivariogram and point semivariogram make downscaling computationally intensive. To integrate several sources of information while accounting for the uncertainty of the estimated parameters, methods that use Bayesian inference such as Bayesian

regression, Bayesian hierarchical modeling, and Bayesian maximum entropy exhibit great potential for downscaling applications (Jameson and Heymsfield, 2013; Kou et al., 2016; Lee et al., 2009; Li et al., 2013; Song et al., 2015; Zhang and Yan, 2015), in which the definition of the prior knowledge is still a challenge. Furthermore, beyond the traditional two-point geostatistics (such as kriging methods), multiple-point geostatistics has been proposed for downscaling (Jha et al., 2013; Strebel, 2002), characterizing the spatial structure by involving three or more points simultaneously.

Moreover, many hybrid downscaling methods have been developed by integrating the advantages of various individual techniques (Jin et al., 2018b; Kaheil et al., 2008; Schoof et al., 2007; Tang et al., 2015). However, in regression models, while independent variables can increase explanatory power, they might also introduce errors that would decrease prediction accuracy.

5.1.3 Downscaling with a process model

For downscaling with a process model and with or without fusing fine-scale auxiliary information, process-based downscaling methods such as dynamical downscaling and data assimilation have been developed.

(a) Dynamical downscaling with a process model

Process-based methods allow for better physical consistency among the downscaled variables in dynamical downscaling than in statistical downscaling (Freybuss et al., 1995). There has been considerable research on dynamical downscaling (Kim et al., 2000; Ramzan et al., 2017) and statistical-dynamical downscaling (Najac et al., 2011; Schulze, 2000). Hydrological models can be coupled in climate simulations (Bindlish and Barros, 2000), and hydrological simulations can be realized through dynamical downscaling (Fujihara et al., 2008; Hay and Clark, 2003). Dynamical downscaling of climate variables is computationally expensive and time-dependently driven by global circulation models, whereas statistical-dynamical downscaling requires less computer resources

without being restricted to short periods, but ignoring the temporal variability ([Praskievicz, 2017](#)).

(b) Downscaling using data simulations with a process model

Data assimilation, as an iterative approach to adjust the modeled states based on multisource observations, has been a growing concern in the downscaling of geophysical variables ([Margulis et al., 2002](#); [Peng et al., 2017](#)). Combined with fine spatial resolution modeling, data assimilation can extract information from multiple-source observations to downscale satellite sensor observations ([Mechri et al., 2016](#); [Reichle et al., 2001](#); [Zupanski et al., 2011](#)). In data assimilation procedures, it is important to account for both measurement and model errors, and to consider bias correction and rigorous quality control for increased understanding of process behaviors.

5.2 Downscaling categories

Categorical Earth science data are enumerated Earth science data that may be divided into groups or levels, such as ecological zones, soil types, socio-economic statistics data and land cover/land use (LCLU). When downscaling categorical variables, the input data could be continuous variables ([Atkinson, 2013](#)). Hence, the mentioned scaling methods for downscaling continua can be applied on categorical variables (e.g., [Wang et al., 2014b](#)). In addition, there are dedicated methods for downscaling categories, e.g., super-resolution mapping (SRM, also termed subpixel mapping) methods, which are used widely for the fine resolution mapping of Earth science variables such as tree crowns, waterlines and inundation ([Foody et al., 2005](#); [Ardila et al., 2011](#); [Li et al., 2015](#)). Among various categorical Earth science variables, LCLU is considered as one of the most important for downscaling, as it can provide critical auxiliary information on land surface types for using in a wide range of Earth science applications ([Atkinson, 2013](#)). Thus, efforts at downscaling to predict categories have been directed at downscaling LCLU more than other categorical Earth science variables. Therefore, in the process of downscaling categorical Earth science data, the categorical data usually represent land

cover/land use (LCLU) classes (but not limited to this variable), as extracted from remote sensing images (Atkinson, 2013). Here, the review of downscaling categories mainly focuses on LCLU. The earliest SRM methods downscale the coarse fraction images obtained from soft classification (i.e., spectral unmixing) of raw images to generate fine hard-classified LCLU maps (i.e., categorical data). Note that they generate categorical data (i.e., LCLU maps), differently to image super-resolution in image processing which outputs continuous data (i.e., images) based on the superposition of input images (Chen et al., 2018a). In the past two decades, many SRM methods have been developed for downscaling categorical data. These can be classified into four main types: (1) spatial autocorrelation-based SRM methods, (2) spatial distribution pattern-based SRM methods, (3) learning-based SRM methods, and (4) auxiliary data-based SRM methods, as shown in Fig. 4.

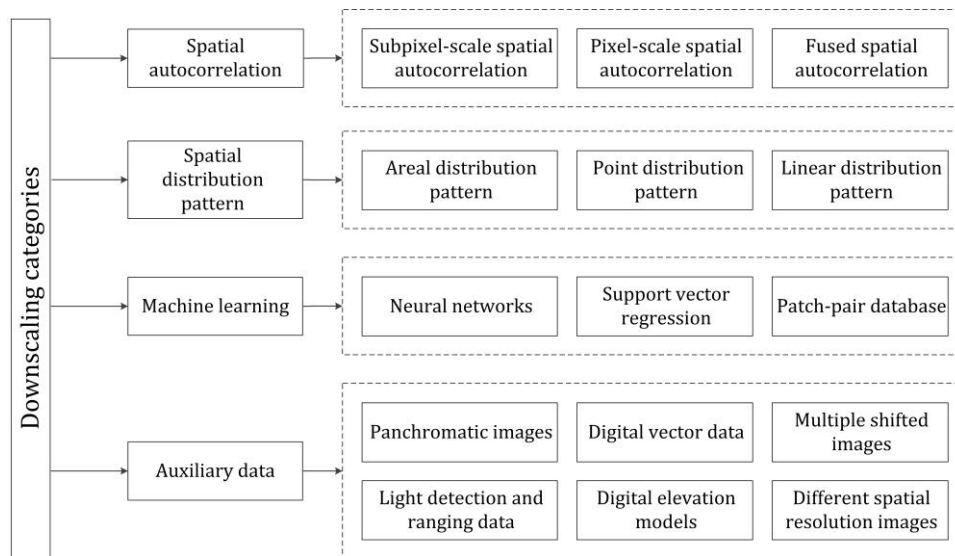


Fig. 4. Major classes of downscaling methods.

5.2.1 Downscaling categories based on spatial autocorrelation

Most SRM methods are based on the principle of spatial autocorrelation, which is considered as the key basis for SRM. Three types of LCLU spatial autocorrelation are used widely in SRM: (1)

subpixel-scale spatial autocorrelation, (2) pixel-scale spatial autocorrelation, and (3) a fusion of these two spatial autocorrelation types (Chen et al., 2014; Ling et al., 2013; Wang et al., 2014b).

(a) Downscaling with subpixel-scale spatial autocorrelation

Subpixel-scale spatial autocorrelation of LCLU classes is characterized by the relationship between the subpixel under consideration (i.e., the fine support of the downscaling prediction) and its neighboring subpixels. The exponential function (Atkinson, 2005), spatial clustering function (Tatem et al., 2001), mean filter (Mertens et al., 2003), and prior energy function (Ardila et al., 2011; Kasetkasem et al., 2005) have been used to characterize subpixel-scale spatial autocorrelation. These methods are able to account for the size of the support in the prediction process.

(b) Downscaling with pixel-scale spatial autocorrelation

Pixel-scale spatial autocorrelation of LCLU classes is extracted from the relationship between the subpixel under consideration and its neighboring pixels. The characterization of pixel-scale spatial autocorrelation can be completed by interpolation algorithms (e.g., kriging) (Chen et al., 2015; Ling et al., 2013; Verhoeye and Wulf, 2002; Wang et al., 2015b), spatial attraction models (Mertens et al., 2006), and geometric models (Ge, 2013; Ge et al., 2014; Ge et al., 2009). Pixel-scale spatial autocorrelation approaches are often linear and fast. However, neither subpixel-scale spatial autocorrelation nor pixel-scale spatial autocorrelation can effectively represent the local details or structure of LCLU patches.

(c) Downscaling with subpixel-scale and pixel-scale spatial autocorrelations

Combining the advantages of the above two types of spatial autocorrelation, there are two ways to increase the prediction accuracy of SRM results. The first is to use the pixel-scale spatial autocorrelation to generate an initial SRM result, and then fine-tune these results based on subpixel-scale spatial autocorrelation (Shen et al., 2009). The other is to combine both types of spatial

autocorrelation to produce the SRM results (Chen et al., 2014; Zhong and Zhang, 2012), whereby a spatial regularization model is generally used to consider not only the spatial autocorrelation of LCLU classes, but also the spectral constraints (Hu et al., 2015a; Zhong et al., 2015).

Although spatial autocorrelation-based SRM methods have achieved acceptable accuracy, they are more appropriate for fine-resolution LCLU features in which the size of the object of interest is coarser than that of the pixels (Chen et al., 2018b). For coarse-resolution LCLU features in which the object sizes are finer than the pixels, they usually fail to yield satisfactory results (Chen et al., 2018a). These spatial autocorrelation-based SRM methods cannot capture the spatial pattern of LCLU features.

5.2.2 Downscaling categories based on spatial distribution pattern

There are three main spatial distribution types of LCLU features: areal patterns (referring to H-resolution LCLU features), point patterns (referring to L-resolution LCLU features), and linear patterns (Ge et al., 2016a). Because of the different spatial distribution characteristics, it is difficult to handle all types of LCLU features at the same time. Many efforts have been made to develop specific SRM methods for different spatial distribution patterns of LCLU features, for example, predicting spatial point patterns using a Hopfield neural network (Tatem et al., 2002), two-point histograms (Atkinson, 2004), and Moran's I index (Ge et al., 2016a). Moreover, neither spatial autocorrelation-based nor spatial pattern-based SRM methods effectively preserve the structure of linear patterns. Anisotropic models (Thornton et al., 2007), linear subpixel mapping agents (Xu et al., 2014), central line control (Ai et al., 2014), and linear templates (Ge et al., 2016a) have been developed to enhance the mapping of linear features. However, as for spatial autocorrelation-based SRM, the spatial distribution pattern-based SRM methods are prior knowledge approaches that require specific models to characterize the spatial distribution pattern or the spatial information of LCLU features.

5.2.3 Downscaling categories based on machine learning

Existing LCLU maps contain a variety of spatial distribution characteristics and provide abundant prior knowledge of LCLU features. Therefore, learning prior knowledge from existing LCLU maps is an attractive approach for downscaling categorical data. Compared with spatial autocorrelation-based and spatial distribution pattern-based SRM methods, the learning-based SRM methods are indirect approaches which focus on learning the relationships between fine LCLU maps and coarse fraction images of LCLU classes. Instead of requiring specific models, learning-based SRM methods need only a learning model to downscale the coarse images to obtain the fine categorical LCLU maps, such as a back-propagation neural network (Nigussie et al., 2011; Wang et al., 2006; Zhang et al., 2008), support vector regression (Zhang et al., 2014c), or the patch-pair learning-database (Ling et al., 2016).

5.2.4 Downscaling categories based on auxiliary data

Although various SRM methods have been developed, SRM is essentially an ill-posed problem, and there may be multiple plausible solutions when only a single remote sensing image is applied to the process of downscaling LCLU categorical data (Chen et al., 2015; Zhong et al., 2014). Therefore, to reduce the uncertainty of SRM, several methods have been developed that use auxiliary data, such as panchromatic images (Li et al., 2014; Nguyen et al., 2011), digital vector data (Aplin and Atkinson, 2001; Su et al., 2012), light detection and ranging data (Nguyen et al., 2005), multiple shifted images (Chen et al., 2015; Wang and Wang, 2013; Xu et al., 2011; Zhong et al., 2014), digital elevation models (Ling et al., 2008), and different spatial resolution images (Li et al., 2017; Ling et al., 2014).

6. Side-scaling

In general, side-scaling methods are similar to upscaling and downscaling methods, and include spatial statistical analysis (Laaha et al., 2013) and hybrid methods (Jin et al., 2018a). In recent decades,

side-scaling methods have been developed and applied in various fields, gradually expanding to the humanities and social sciences (e.g., cancer mortality risk (Goovaerts, 2006) and crime data (Kerry et al., 2010)) as well as the natural sciences (e.g., evapotranspiration (Ge et al., 2015a) and cranberry yield (Kerry et al., 2017)).

The core of side-scaling is area-based weighting, for which the simplest approach is to use the area proportion. This allows coarse-resolution data to be derived from fine-resolution data according to the areal extent of different land cover types within the coarse-resolution pixels. These methods are simple in principle and can select from a wide range of models according to the characteristics of the specific problem. However, parameters such as the mean and variance usually only characterize the global picture of a certain feature, and cannot reflect local features.

The geostatistical method is another common side-scaling approach based on the scale-dependent property. ATAK-based methods are a representative technique, for example, conventional ATAK (Kyriakidis, 2004) and area-to-area Poisson kriging (Goovaerts, 2006; Kerry et al., 2010). In addition, top-kriging (also known as topological kriging) takes both the area and the nested nature of catchments into account in streamflow-related predictions (Skøien et al., 2005; Skøien and Blöschl, 2007; Parajka et al., 2015). To fuse the available auxiliary variables, area-to-area regression kriging and its extensions can be applied (Ge et al., 2015a). Moreover, weighted area-to-area regression kriging (Hu et al., 2015b) and geographical weighted area-to-area regression kriging (Jin et al., 2018a) have been proposed to increase the prediction accuracy by considering the sample heterogeneity and local spatial heterogeneity, respectively.

7. Discussion

Scaling methods rely largely on statistical-driven or physics-based methods, sometimes with a

certain assumption of scale-invariance and commonly based on autocorrelation. The more sophisticated scaling methods are scale-dependent when the actual heterogeneity can be captured, even to a limited extent (Hewitt et al., 2007). We suggest that the scaling methods should be tuned according to the specific type of heterogeneity (see Section 2). We now discuss different aspects of scaling, including the main problems still faced and possible developing trends.

7.1 Scaling under spatial autocorrelation and heterogeneity

As scale-dependent properties, spatial autocorrelation and spatial heterogeneity of Earth science data may change from one scale to another. Taking spatial autocorrelation as an example, the semivariograms of soil moisture observations at 0–4 cm depth for three different scales were estimated at coarse, medium and fine measurement scales (i.e., spatial extents) (Fig. 5b-d). The ground stations were installed in the Yingke-Daman irrigation district, in the middle reaches of the Heihe River Basin, China. Detailed information on the three scales and stations (Fig. 5a) was reported by Ge et al. (2015b) and Wang et al. (2015a). It is obvious that the semivariogram changes as the scale or spatial extent varies in Fig. 5b-d.

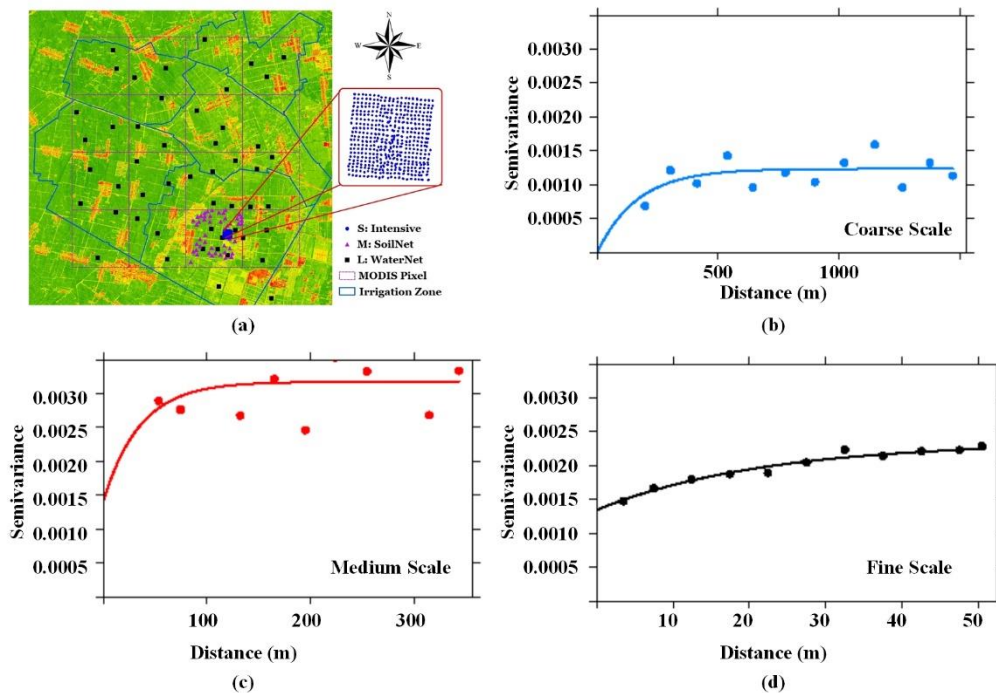


Fig. 5. (a) Study area in the Yingke-Daman irrigation district; and various semivariograms at different scales: (b)

coarse scale; (c) medium scale; (d) fine scale.

However, most of the abovementioned scaling methods are based on spatial autocorrelation and ignore its variation. Although some scaling methods have been developed by considering spatial heterogeneity ([Pettorelli et al., 2005](#)), it is rare to involve both autocorrelation and heterogeneity in the scaling process. Spatial autocorrelation or spatial heterogeneity alone can only approximately describe the Earth's surface ([Goodchild, 2004](#)) and, thus, it is best to consider the two laws simultaneously when quantitatively representing geographic variables. Moreover, sometimes the spatial variation in a geographic variable may be characterized as autocorrelation at one scale, and heterogeneity at another scale. The spatial properties might change at a certain scale, which would bring challenges for scaling. Therefore, the development of scaling methods for Earth science data should include use of spatial autocorrelation and heterogeneity, and requires some consideration of the interchanging relationships between different scales.

7.2 Scaling with scale-independent and scale-dependent factors

The proposed scaling function in Eq. (1) consists of scale-independent and scale-dependent parts, and the corresponding scaling methods have been described in previous sections. However, scale-independent features are easily neglected in scaling, and few scaling methods take both scale-independent and scale-dependent properties into consideration ([Atkinson, 2013](#)). In the field of geospatial Earth sciences, some attributes of a variable will vary with a change of scale, but others will remain unchanged. Thus, there may exist scale-dependent factors and scale-independent factors in characterizing a geographical variable through multiple scales. In the scaling of Earth science data, using either scale-independent or scale-dependent factors for scaling is unlikely to satisfy the high-accuracy requirement for prediction. To explore these factors, one could design an experiment in a selected study area. The multi-source and multi-scale data of the target variable could be collected

from multiple remote sensing products and airborne observations, and sometimes from auxiliary variables. After exploring the scale characteristics, one could deduce the trend in variation with scale change using various indices (e.g., statistical moments and regression parameters). Then, the identified scale-dependent and scale-independent factors could be entered into Eq. (1) where the simplest form refers to $Z_D = f(Z_0) + g(Z_0) + \varepsilon$. A conceptual experiment design for this scaling is shown in Fig.6.

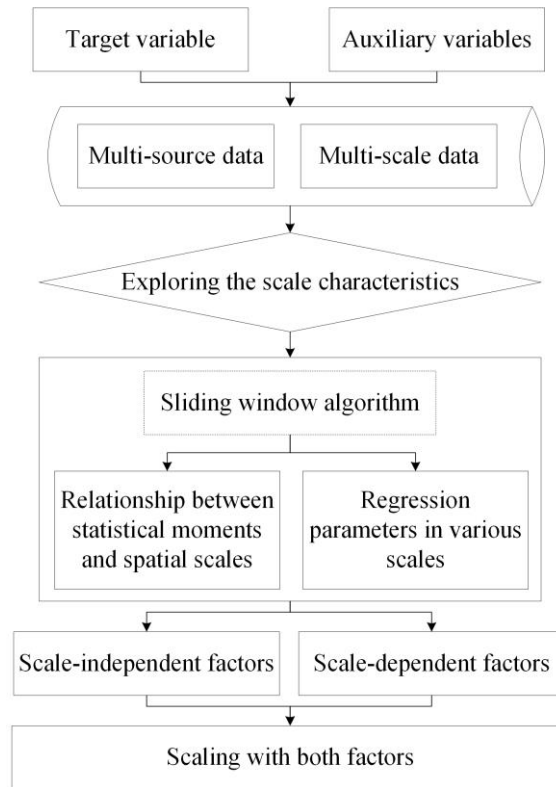


Fig. 6. A conceptual experiment design for scaling with scale-independent and scale-dependent factors

7.3 Scaling incorporating geospatial Earth surface processes

As it is crucial to explore physical dependencies and structural correlations in the observed data in spatiotemporal analysis, we should adopt models that can incorporate a dynamical mechanism, such as fusing geographical processes in Bayesian maximum entropy, hierarchical Bayesian methods, and data assimilation (Kyriakidis and Journel, 1999). However, the methodology is still insufficient for estimating geographical variables at the target scale, and applications are limited by the complexity

and operability of exploring the geographical processes.

In our opinion, there are three possible aspects for determining the spatiotemporal scaling method incorporating geospatial Earth surface processes. The first is to develop a theoretical system based on the main existing geostatistical spatiotemporal models, such as integrating stochastic partial differential equations into kriging, or fusing the laws of nature, high-order physical quantities, and various uncertain information into a Bayesian maximum entropy model. The second way is to explore the geographical pattern and process from a systematic point of view. A regional geographical system might be established by analyzing multiple geographical units (i.e., landscape-type units), and then the dynamical geographical processes could be identified to reveal the scaling law for obtaining information at different scales. The third approach is to improve the data assimilation method by involving multi-source and multi-scale data. Moreover, capturing the real geographical system is a serious challenge because of its complexity and size. Another potentially useful direction is to realize scaling through data assimilation and stochastic partial differential equations (Cressie and Wikle, 2015).

7.4 Scaling considering information changes in the scaling process

Because the causes and formation mechanisms of geographical processes and patterns are generally different at different measurement scales (Wu et al., 2000), there is likely to be a transformation in energy and information. These flows across different scales could provide the basis for scaling. In ecology, the Markov model of the transfer of energy between different species, and the fluctuation–dissipation theorem and scaling rules of species numbers, can be derived based on the local Damuth symmetry (Li and Charnov, 2001), which might be referential for capturing the energy and information changes in scaling. Let $k(\cdot)$ be the gains or losses of energy and information, and write the scaling function as $Z_D = \Phi[f(Z_0), g(Z_0), k(Z_0)] + \varepsilon$, in which the scaling function Φ is

the composition of f , g , and k . Ideally, predictions are derived at the target scale involving the energy and information at the available scale and the exchanges between them. During scaling, the energy system might be in the equilibrium or non-equilibrium state, and may be transferred from dominant lower modes into higher modes. The information might be measured from several aspects (such as geometric features and attribute features) using various indices. Through the change in regularity of these indices, the gains or losses of information in the scaling process could be obtained, contributing to the process of scaling.

Moreover, for spatiotemporal scaling, limitations in the input data coverage and validation of scaling predictions also need to be considered. The geographical patterns and processes are often spatially and temporally continuous with strong dependence in both space and time. This leads to the modifiable areal unit problem ([Openshaw, 1984](#)) and the modifiable temporal unit problem ([Cheng and Adepeju, 2014](#)). Separating space and time and neglecting the characteristics of spatiotemporal variability can result in the loss of a great deal of valuable information. Although the nonseparable covariance functions could support space-time interaction, the nonseparable covariance structures may not be feasible at any space-time lags ([Gneiting, 2002](#)). The spatiotemporal analysis is vital in studying the scaling of geographic patterns and processes. Moreover, due to observing conditions, it is not always possible to obtain the full spatial coverage of geographic variables, for example, visible/infrared remotely sensed products are often negatively affected by clouds. Thus, a scaling method based on these data may not produce continuous predictions. If the scaling occurs with a lack of information at the target scale, there will not be sufficient reference data for validation.

7.5 Suitability scale for models

Scaling models vary widely in terms of their complexity and data requirements, and cannot be applied directly to any place and at any scale without calibration or adjustments. There are different

properties and models at each scale for characterizing the geographical variables. The dominant variables for the geographical pattern and process will change with the change of scale. To establish the relationship among different scales, it is possible to presuppose scale-invariance during scaling. Models based on observations could be considered as constant, hence ignoring the change, and the scale could be allowed to vary within a certain range to describe the characteristics at the scale of interest. Beyond this range, the assumption of scale-invariance does not work, as it results in the discredited information captured by established models. Similarly, the dynamic variation of spatiotemporal distribution characteristics also faces this problem. The parameters or forms of scaling model need to be updated. However, to the best of our knowledge, only a few researchers have studied the upper and lower limits of scaling in subpixel mapping (e.g., [Ge et al., 2016b](#)). Thus, it is important to identify the suitability of the scaling model at the target scale. The exploration of sudden changes in variable characteristics might provide a way to solve this issue. Stepwise scaling can be applied to obtain multi-scale predictions, such as downscaling a remote sensing soil moisture product from 25-km spatial resolution to 20, 15, 10, 5, 3 and 1 km. According to the autocorrelation (e.g., by using autocorrelation coefficients and semivariograms) of the predictions at different scales, the most appropriate scale could be determined, particularly if the target is known, for example for a test area. If it were possible to determine the most suitable target scales for the scaling model, it would be possible to select various scaling methods for a specific application.

7.6 Uncertainties in the scaling process

Uncertainties in the input data and models used will increase the error in predictions from scaling. The uncertainties of the scaling model, mainly statistical models, are embodied in the model formula and model parameters. The errors of all input data may propagate to the following phases in the scaling process, and can significantly influence the accuracy of the predictions. For combined data-driven and

physics-based scaling methods, in particular, the predicted results are questionable under the complex interaction among the errors in the variables and the process model. In the validation of scaling methods, direct observations are always the reference data. It is common for the scale of these observations to be different from the target scale. Hence, the differences between the predictions and the validation data are caused not only by scaling model errors, but also by the representativeness errors of different supports and observation errors in the validation data. By investigating the sources of uncertainties in the scaling process by using methods such as Monte Carlo simulation (Kuczera and Parent, 1998) and generalised likelihood uncertainty estimation (Beven and Freer, 2001), it is feasible to reduce and control the errors in predictions, and thus increase the scaling accuracy. Although the repeated measurements would remove random errors and increase the reliability of the observations, it is a computationally and challenging task to observe all input variables repeatedly. By employing a priori distributions and likelihood functions of input variables, the corresponding posteriori distributions can be obtained for exploring the error propagation. Then, the different influences of different input variables on scaling prediction error can be derived. A further approach for error reduction is to increase the accuracy of the most significant input variable with the multiple mean method.

8. Summary

Scaling is often necessary to provide the information required at the scale of interest. This paper summarized the main existing methods for upscaling, downscaling and side-scaling, which should help researchers select the most appropriate technique for their specific research objectives and data conditions. Considering various spatial properties, a general form of scaling function was proposed, involving scale-invariant and scale-variant characteristics. The review of current scaling methods

revealed a focus on one component of the scaling function. A specific formulation of the scaling function incorporating two components should be explored in future research. Moreover, various heterogeneities were classified and should be employed for future scaling processes. The challenges faced by current scaling approaches were discussed. The information, energy, matter and characteristics of geospatial Earth science phenomena can be preserved and propagated effectively into the data at the transformed scale. This Review addressed scaling primarily in the geospatial Earth science domain and the conclusions need further investigation for other domains. Moreover, studies on scaling, as a persistent issue in the geospatial analysis of Earth science data, should be developed further.

Acknowledgments

This work was supported by the National Natural Science Foundation for Distinguished Young Scholars of China under Grant 41725006, and two Key Programs of the National Natural Science Foundation of China under Grant 41531174 and Grant 41531179.

References

- Ai, B., Liu, X.P., Hu, G.H., Li, X., 2014. Improved sub-pixel mapping method coupling spatial dependence with directivity and connectivity. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7(12), 4887-4896.
- Anderson, M.C., Kustas, W.P., Norman, J.M., 2003. Upscaling and downscaling - A regional view of the soil-plant-atmosphere continuum. *Agron. J.* 95(6), 1408-1423.
- Aplin, P., Atkinson, P.M., 2001. Sub-pixel land cover mapping for per-field classification. *Int. J. Remote Sens.* 22(14), 2853-2858.
- Ardila, J.P., Tolpekin, V.A., Bijker, W., Stein, A., 2011. Markov-random-field-based super-resolution mapping for identification of urban trees in VHR images. *ISPRS J. Photogramm. Remote Sens.* 66(6), 762-775.
- Atkinson, P.M., 2004. Super-resolution land cover classification using the two-point histogram, In: X. SanchezVila, J. Carrera and J.J. GomezHernandez (Eds), *Geoenv Iv - Geostatistics for Environmental Applications: Proceedings. Quantitative Geology and Geostatistics*. Springer, Dordrecht, pp. 15-28.
- Atkinson, P.M., 2005. Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogramm. Eng. Remote Sens.* 71(7), 839-846.
- Atkinson, P.M., 2013. Downscaling in remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* 22, 106-114.
- Atkinson, P.M., Pardo-Iguzquiza, E., Chica-Olmo, M., 2008. Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. *IEEE Trans. Geosci. Remote Sens.* 46(2), 573-580.
- Atkinson, P.M., Tate, N.J., 2000. Spatial scale problems and geostatistical solutions: A review. *Prof. Geogr.* 52(4), 607-623.
- Bai, H.X., Ge, Y., Mariethoz, G., 2016. Utilizing spatial association analysis to determine the number of multiple grids for multiple-point statistics. *Spat. Stat.* 17, 83-104.
- Bergant, K., Kajfez-Bogataj, L., Crepinsek, Z., 2002. Statistical downscaling of general-circulation-model-simulated average monthly air temperature to the beginning of flowering of the dandelion (*Taraxacum officinale*) in Slovenia. *Int. J. Biometeorol.* 46(1), 22-32.
- Bergin, E., Buytaert, W., Onof, C., Wheeler, H., 2012. Downscaling of rainfall in Peru using generalised linear models, *World congress on water, climate and energy*.
- Berrocal, V.J., Craigmiller, P.F., Guttorp, P., 2012. Regional climate model assessment using statistical upscaling and downscaling techniques. *Environmetrics* 23(5), 482-492.
- Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. *J. Hydrol.* 249(1-4), 11-29.
- Beyrich, F., Leps, J.P., Mauder, M., Bange, J., Foken, T., Huneke, S., Lohse, H., Ludi, A., Meijninger, W.M.L., Mironov, D., Weissensee, U., Zittel, P., 2006. Area-averaged surface fluxes over the litfass region based on eddy-covariance measurements. *Bound.-Layer Meteor.* 121(1), 33-65.
- Bierkens, M., Finke, P.A., De Willigen, P., 2000. Upscaling and downscaling methods for environmental research (developments in plant and soil sciences) Kluwer Academic, Dordrecht, Boston, London.
- Bindlish, R., Barros, A.P., 2000. Disaggregation of rainfall for one-way coupling of atmospheric and hydrological models in regions of complex terrain. *Glob. Planet. Change* 25(1-2), 111-132.
- Cannon, A.J., 2009. Negative ridge regression parameters for improving the covariance structure of multivariate linear downscaling models. *Int. J. Climatol.* 29(5), 761-769.
- Chen, Y.D., Chen, X., Xu, C.Y., Shao, Q., 2006. Downscaling of daily precipitation with a stochastic weather generator for the subtropical region in South China. *Hydrol. Earth Syst. Sci.* 3(3), 1145-1183.
- Chen, Y.H., Ge, Y., Chen, Y., Jin, Y., An, R., 2018a. Subpixel land cover mapping using multiscale spatial dependence. *IEEE Trans. Geosci. Remote Sens.* 56(9), 5097-5106.
- Chen, Y.H., Ge, Y., Heuvelink, G.B.M., An, R., Chen, Y., 2018b. Object-based superresolution land-cover mapping from remotely sensed imagery. *IEEE Trans. Geosci. Remote Sens.* 56(1), 328-340.
- Chen, Y.H., Ge, Y., Heuvelink, G.B.M., Hu, J.L., Jiang, Y., 2015. Hybrid constraints of pure and mixed pixels for

- soft-then-hard super-resolution mapping with multiple shifted images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8(5), 2040-2052.
- Chen, Y.H., Ge, Y., Wang, Q.M., Jiang, Y., 2014. A subpixel mapping algorithm combining pixel-level and subpixel-level spatial dependences with binary integer programming. *Remote Sens. Lett.* 5(10), 902-911.
- Cheng, Q., 2005. A new model for incorporating spatial association and singularity in interpolation of exploratory data, *Geostatistics Banff 2004. Quantitative Geology and Geostatistics*, Dordrecht.
- Cheng, Q.M., 1999. Multifractality and spatial statistics. *Comput. Geosci.* 25(9), 949-961.
- Cheng, Q.M., 1997. Discrete multifractals. *Math. Geol.* 29(2), 245-266.
- Cheng, T., Adepeju, M., 2014. Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *Plos One* 9(6), e100465.
- Chiles, J., Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New York.
- Cressie, N.A., Wikle, C.K., 2015. *Statistics for spatio-temporal data*. John Wiley & Sons, United States.
- Crow, W.T., Berg, A.A., Cosh, M.H., Loew, A., Mohanty, B.P., Panciera, R., de Rosnay, P., Ryu, D., Walker, J.P., 2012. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Rev. Geophys.* 50, 20.
- Crow, W.T., Ryu, D., Famiglietti, J.S., 2005. Upscaling of field-scale soil moisture measurements using distributed land surface modeling. *Adv. Water Resour.* 28(1), 1-14.
- Dai, A., Rasmussen, R.M., Ikeda, K., Liu, C.H., 2017. A new approach to construct representative future forcing data for dynamic downscaling. *Clim. Dyn.*, 1-9.
- de Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Kotters, M., 2006. *Sampling for natural resource monitoring*. Springer-Verlag, New York.
- Deidda, R., 2000. Rainfall downscaling in a space-time multifractal framework. *Water Resour. Res.* 36(7), 1779-1794.
- Dibike, Y.B., Coulibaly, P., 2006. Temporal neural networks for downscaling climate variability and extremes. *Neural Netw.* 19(2), 135-144.
- Espinosa, L.D., Posadas, A., Carbajal, M., Quiroz, R., 2017. Multifractal downscaling of rainfall using normalized difference vegetation index (NDVI) in the Andes Plateau. *PloS one* 12(1), e0168982.
- Fasbender, D., Ouara, T.B.M.J., 2010. Spatial Bayesian model for statistical downscaling of AOGCM to minimum and maximum daily temperatures. *J. Clim.* 23(19), 5222-5242.
- Finke, P.A., Bierkens, M.F.P., 2002. Choosing appropriate upscaling and downscaling methods for environmental research, In: J. Steenvoorden, F. Claessen and J. Willems (Eds), *Agricultural Effects on Ground and Surface Waters: Research at the Edge of Science and Society*. IAHS Publication, pp. 405-409.
- Folberth, C., Baklanov, A., Balkovic, J., Skalsky, R., Khabarov, N., Obersteiner, M., 2019. Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. *Agric. For. Meteorol.* 264, 1-15.
- Foody, G.M., Muslim, A.M., Atkinson, P.M., 2005. Super - resolution mapping of the waterline from remotely sensed data. *Int. J. Remote Sens.* 26(24), 5381-5392.
- Freybuness, F., Heimann, D., Sausen, R., 1995. A statistical-dynamical downscaling procedure for global climate simulations. *Theor. Appl. Climatol.* 50(3-4), 117-131.
- Friederichs, P., Hense, A., 2007. Statistical downscaling of extreme precipitation events using censored quantile regression. *Mon. Weather Rev.* 135(6), 2365-2378.
- Fu, B., 2014. The integrated studies of geography: Coupling of patterns and processes. *Acta Geographica Sinica* 69(8), 1052-1059.
- Fu, B.J., Chen, L.D., Ma, K.M., Wang, Y.L., 2011. *Landscape ecology: principles and applications* Science Press, Beijing.
- Fujihara, Y., Tanaka, K., Watanabe, T., Nagano, T., Kojiri, T., 2008. Assessing the impacts of climate change on the water resources of the Seyhan River Basin in Turkey: Use of dynamically downscaled data for hydrologic simulations. *J. Hydrol.* 353(1-2), 33-48.
- Galukande, M., Kaggwa, S., Sekimpi, P., Kakaire, O., Katamba, A., Munabi, I., Runumi, F.M., Mills, E., Hagopian, A., Blair, G., 2013. Use of surgical task shifting to scale up essential surgical services: a feasibility analysis at facility

- level in Uganda. *BMC Health Serv. Res.* 13(1), 292.
- Gaucherel, C., 2007. Multiscale heterogeneity map and associated scaling profile for landscape analysis. *Landscape Urban Plan.* 82(3), 95-102.
- Ge, Y., 2013. Sub-pixel land-cover mapping with improved fraction images upon multiple-point simulation. *Int. J. Appl. Earth Obs. Geoinf.* 22, 115-126.
- Ge, Y., Bai, H., 2011. Multiple-point simulation-based method for extraction of objects with spatial structure from remotely sensed imagery. *Int. J. Remote Sens.* 32(8), 2311-2335.
- Ge, Y., Chen, Y., Li, S., Jiang, Y., 2014. Vectorial boundary-based sub-pixel mapping method for remote-sensing imagery. *Int. J. Remote Sens.* 35(5), 1756-1768.
- Ge, Y., Chen, Y., Stein, A., Li, S., Hu, J., 2016a. Enhanced subpixel mapping with spatial distribution patterns of geographical objects. *IEEE Trans. Geosci. Remote Sens.* 54(4), 2356-2370.
- Ge, Y., Du, Y., Cheng, Q., Li, C., 2006. Multifractal filtering method for extraction of ocean eddies from remotely sensed imagery. *Acta Oceanologica Sinica* 25(5), 27-38.
- Ge, Y., Jiang, Y., Chen, Y.H., Stein, A., Jiang, D., Jia, Y.X., 2016b. Designing an experiment to investigate subpixel mapping as an alternative method to obtain land use/land cover maps. *Remote Sens.* 8(5), 360.
- Ge, Y., Li, S., Lakhan, V.C., 2009. Development and testing of a subpixel mapping algorithm. *IEEE Trans. Geosci. Remote Sens.* 47(7), 2155-2164.
- Ge, Y., Liang, Y., Wang, J., Zhao, Q., Liu, S., 2015a. Upscaling sensible heat fluxes with area-to-area regression kriging. *IEEE Geoscience and Remote Sens. Lett.* 12(3), 656-660.
- Ge, Y., Wang, J.H., Heuvelink, G.B.M., Jin, R., Li, X., Wang, J.F., 2015b. Sampling design optimization of a wireless sensor network for monitoring ecohydrological processes in the Babao River basin, China. *Int. J. Geogr. Inf. Sci.* 29(1), 92-110.
- Geary, R.C., 1954. The contiguity ratio and statistical mapping. *The Incorporated Statistician* 5, 115-127.
- Gelfand, A.E., Kim, H. J., Sirmans, C. F., Banerjee, S., 2003. Spatial modeling with spatially varying coefficient processes. *J. Am. Stat. Assoc.* 98(462), 387-396.
- Gelfand, A.E., Zhu, L., Carlin, B.P., 2001. On the change of support problem for spatio-temporal data. *Biostatistics* 2(1), 31-45.
- Ghahramani, M., Zhou, M. C., Hon, C. T., 2018. Mobile phone data analysis: a spatial exploration toward hotspot detection. *IEEE Trans. Autom. Sci. Eng.* 16(1), 351-362.
- Gires, A., Abbes, J.B., Paz, I.D.R., Tchiguirinskaia, I., Schertzer, D., 2018. Multifractal characterisation of a simulated surface flow: A case study with Multi-Hydro in Jouy-en-Josas, France. *J. Hydrol.* 558, 482-495.
- Gneiting, T., 2002. Nonseparable, stationary covariance functions for space-time data. *J. Am. Stat. Assoc.* 97(458), 590-600.
- Goodchild, M.F., 2004. The validity and usefulness of laws in geographic information science and geography. *Ann. Assoc. Am. Geogr.* 94(2), 300-303.
- Goodchild, M.F., 2011. Scale in GIS: An overview. *Geomorphology* 130(1-2), 5-9.
- Goovaerts, P., 2001. Geostatistical modelling of uncertainty in soil science. *Geoderma* 103(1-2), 3-26.
- Goovaerts, P., 2006. Geostatistical analysis of disease data: accounting for spatial support and population density in the isopleth mapping of cancer mortality risk using area-to-point Poisson kriging. *Int. J. Health Geogr.* 5, 52.
- Goovaerts, P., 2008. Kriging and semivariogram deconvolution in the presence of irregular geographical units. *Math. Geosci.* 40(1), 101-128.
- Gotway, C.A., Young, L.J., 2002. Combining incompatible spatial data. *J. Am. Stat. Assoc.* 97(458), 632-648.
- Gotway, C.A., Young, L.J., 2007. A geostatistical approach to linking geographically aggregated data from different sources. *J. Comput. Graph. Stat.* 16(1), 115-135.
- Guardiano, F.B., Srivastava, R.M., 1993. Multivariate geostatistics - beyond bivariate moments, *Geostatistics Troia* 92, Vols 1 and 2, 5 Kluwer Academic Publ, Dordrecht.
- Ha, W., Gowda, P.H., Howell, T.A., 2013. A review of downscaling methods for remote sensing-based irrigation

- management: part I. *Irrig. Sci.* 31(4), 831-850.
- Halley, J.M., Hartley, S., Kallimanis, A.S., Kunin, W.E., Lennon, J.J., Sgardelis, S.P., 2004. Uses and abuses of fractal methodology in ecology. *Ecol. Lett.* 7(3), 254-271.
- Hamm, N. A. S., Finley, A. O., Schaap, M., Stein, A., 2015. A spatially varying coefficient model for mapping pm10 air quality at the European scale. *Atmos. Environ.* 102, 393-405.
- Hay, L.E., Clark, M.P., 2003. Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *J. Hydrol.* 282(1-4), 56-75.
- Heuvelink, G.B.M., Pebesma, E.J., 1999. Spatial aggregation and soil process modelling. *Geoderma* 89(1-2), 47-65.
- Hewitt, J.E., Thrush, S.F., Dayton, P.K., Bonsdorff, E., 2007. The effect of spatial and temporal heterogeneity on the design and analysis of empirical studies of scale-dependent systems. *Am. Nat.* 169(3), 398-408.
- Hoai, N.D., Udo, K., Mano, A., 2011. Downscaling Global Weather Forecast Outputs Using ANN for Flood Prediction. *J. Appl. Math.* 1, 223-236.
- Hu, J.L., Ge, Y., Chen, Y.H., Li, D.Y., 2015a. Super-resolution land cover mapping based on multiscale spatial regularization. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8(5), 2031-2039.
- Hu, M.G., Wang, J.H., Ge, Y., Liu, M.X., Liu, S.M., Xu, Z.W., Xu, T.R., 2015b. Scaling flux tower observations of sensible heat flux using weighted area-to-area regression Kriging. *Atmosphere* 6(8), 1032-1044.
- Hu, S.G., Cheng, Q.M., Wang, L., Xu, D.Y., 2013. Modeling land price distribution using multifractal IDW interpolation and fractal filtering method. *Landsc. Urban Plan.* 110, 25-35.
- Hulme, M., Harrison, P., Arnell, N., 1999. Climate variability and crop yields in Europe - Reply. *Nature* 400(6746), 724-724.
- Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sens. Environ.* 178, 127-141.
- Jameson, A.R., Heymsfield, A.J., 2013. A Bayesian approach to upscaling and downscaling of aircraft measurements of ice particle counts and size distributions. *J. Appl. Meteorol. Climatol.* 52(9), 2075-2088.
- Jarvis, P.G., 1995. Scaling processes and problems. *Plant Cell Environ.* 18(10), 1079-1089.
- Jha, S.K., Mariethoz, G., Evans, J.P., McCabe, M.F., 2013. Demonstration of a geostatistical approach to physically consistent downscaling of climate modeling simulations. *Water Resour. Res.* 49(1), 245-259.
- Jiang, B., 2018. Spatial heterogeneity, scale, data character and sustainable transport in the big data era. *ISPRS Int. Geo-Inf.* 7(5), 167.
- Jiang, B., Brandt, S.A., 2016. A fractal perspective on scale in geography. *ISPRS Int. Geo-Inf.* 5(6), 95.
- Jiang, B., Ma, D., 2018. How complex is a fractal? Head/tail breaks and fractional hierarchy. *J. geovis. spat. anal.* 2(1), 6.
- Jiang, J.L., Liu, X.N., Liu, C.H., Wu, L., Xia, X.P., Liu, M.L., Du, Z.H., 2015. Analyzing the spatial scaling bias of rice leaf area index from hyperspectral data using wavelet-fractal technique. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8(6), 3068-3080.
- Jin, Y., Ge, Y., Wang, J.H., Chen, Y.H., Heuvelink, G.B.M., Atkinson, P.M., 2018a. Downscaling AMSR-2 soil moisture data with geographically weighted area-to-area regression kriging. *IEEE Trans. Geosci. Remote Sens.* 56(4), 2362-2376.
- Jin, Y., Ge, Y., Wang, J.H., Heuvelink, G.B.M., 2018b. Deriving temporally continuous soil moisture estimations at fine resolution by downscaling remotely sensed product. *Int. J. Appl. Earth Obs. Geoinf.* 68, 8-19.
- Journel, A.G., 1993. *Geostatistics-roadblocks and challenges*, Geostatistics Troia 92, Vols 1 and 2, 5 Kluwer Academic Publ, Dordrecht.
- Jung, M., Reichstein, M., Margolis, H.A., Cescatti, A., Richardson, A.D., Arain, M.A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J.Q., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B.E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E.J., Papale, D., Sottocornola, M., Vaccari, F., Williams, C., 2011. Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res. Biogeosci.* 116, G00J07.
- Kaheil, Y.H., Rosero, E., Gill, M.K., McKee, M., Bastidas, L.A., 2008. Downscaling and forecasting of evapotranspiration

- using a synthetic model of wavelets and support vector machines. *IEEE Trans. Geosci. Remote Sens.* 46(9), 2692-2707.
- Kang, J., Jin, R., Li, X., Zhang, Y., Zhu, Z.L., 2018. Spatial upscaling of sparse soil moisture observations based on ridge regression. *Remote Sens.* 10(2), 192.
- Kang, J., Jin, R., Li, X., Ma, C.F., Qin, J., Zhang, Y., 2017. High spatio-temporal resolution mapping of soil moisture by integrating wireless sensor network observations and modis apparent thermal inertia in the babao river basin, china. *Remote Sens. Environ.* 191, 232-245.
- Kasetkasem, T., Arora, M.K., Varshney, P.K., 2005. Super-resolution land cover mapping using a Markov random field based approach. *Remote Sens. Environ.* 96(3-4), 302-314.
- Keil, P., Pereira, H.M., Cabral, J.S., Chase, J.M., May, F., Martins, I.S., Winter, M., 2018. Spatial scaling of extinction rates: Theory and data reveal nonlinearity and a major upscaling and downscaling challenge. *Glob. Ecol. Biogeogr.* 27(1), 2-13.
- Kerr, Y.H., Waldteufel, P., Wigneron, J.P., Martinuzzi, J.M., Font, J., Berger, M., 2001. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.* 39(8), 1729-1735.
- Kerry, R., Goovaerts, P., Gimenez, D., Oudemans, P.V., 2017. Investigating temporal and spatial patterns of cranberry yield in New Jersey fields. *Precis. Agric.* 18(4), 507-524.
- Kerry, R., Goovaerts, P., Haining, R.P., Ceccato, V., 2010. Applying geostatistical analysis to crime data: car-related thefts in the Baltic States. *Geogr. Anal.* 42(1), 53-77.
- Kerry, R., Goovaerts, P., Rawlins, B.G., Marchant, B.P., 2012. Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. *Geoderma* 170, 347-358.
- Kim, J., Miller, N.L., Farrara, J.D., Hong, S.Y., 2000. A seasonal precipitation and stream flow hindcast and prediction study in the western United States during the 1997/98 winter season using a dynamic downscaling system. *J. Hydrometeorol.* 1(4), 311-329.
- Kou, X.K., Jiang, L.M., Bo, Y.C., Yan, S., Chai, L.N., 2016. Estimation of land surface temperature through blending MODIS and AMSR-E data with the Bayesian maximum entropy method. *Remote Sens.* 8(2), 105.
- Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the metropolis algorithm. *J. Hydrol.* 211(1-4), 69-85.
- Kyriakidis, P.C., 2004. A geostatistical framework for area-to-point spatial interpolation. *Geogr. Anal.* 36(3), 259-289.
- Kyriakidis, P.C., Journel, A.G., 1999. Geostatistical space-time models: A review. *Math. Geol.* 31(6), 651-684.
- Laaha, G., Skoien, J.O., Nobilis, F., Blöschl, G., 2013. Spatial prediction of stream temperatures using top-kriging with an external drift. *Environ. Model. Assess.* 18(6), 671-683.
- Landeras, G., Lopez, J.J., Kisi, O., Shiri, J., 2012. Comparison of Gene expression programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain). *Energy Conv. Manag.* 62, 1-13.
- Landman, W.A., Tennant, W.J., 2000. Statistical downscaling of monthly forecasts. *Int. J. Climatol.* 20(13), 1521-1532.
- Lanza, L.G., Schultz, G.A., Barrett, E.C., 1997. Remote sensing in hydrology: some downscaling and uncertainty issues. 22(3-4), 215-219.
- Lee, S.J., Yeatts, K.B., Serre, M.L., 2009. A Bayesian Maximum Entropy approach to address the change of support problem in the spatial analysis of childhood asthma prevalence across North Carolina. *Spatial Spatio-temporal Epidemiol.* 1(1), 49-60.
- Li, A.H., Bo, Y.C., Zhu, Y.X., Guo, P., Bi, J., He, Y.Q., 2013. Blending multi-resolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method. *Remote Sens. Environ.* 135, 52-63.
- Li, B.L., 2000. Fractal geometry applications in description and analysis of patch patterns and patch dynamics. *Ecol. Model.* 132(1-2), 33-50.
- Li, B.L., Charnov, E.L., 2001. Diversity-stability relationships revisited: scaling rules for biological communities near equilibrium. *Ecol. Model.* 140(3), 247-254.
- Li, L.Y., Chen, Y., Xu, T.B., Liu, R., Shi, K.F., Huang, C., 2015. Super-resolution mapping of wetland inundation from

- remote sensing imagery based on integration of back-propagation neural network and genetic algorithm. *Remote Sens. Environ.* 164, 142-154.
- Li, S.C., Cai, Y.L., 2005. Some scaling issues of geography. *Geographical Research(China)* 24(1), 11-18.
- Li, X.D., Ling, F., Du, Y., Zhang, Y.H., 2014. Spatially adaptive superresolution land cover mapping with multispectral and panchromatic images. *IEEE Trans. Geosci. Remote Sens.* 52(5), 2810-2823.
- Li, X.D., Ling, F., Foody, G.M., Ge, Y., Zhang, Y.H., Du, Y., 2017. Generating a series of fine spatial and temporal resolution land cover maps by fusing coarse spatial resolution remotely sensed images and fine spatial resolution land cover maps. *Remote Sens. Environ.* 196, 293-311.
- Li, X.W., Cao, C.X., Zhang, H., 2009. Study on scale problem. *Journal of remote sensing (China)* 13(s1), 1993-2002.
- Li, Z.L., Zhou, Q., 2012. Integration of linear and areal hierarchies for continuous multi-scale representation of road networks. *Int. J. Geogr. Inf. Sci.* 26(5), 855-880.
- Ling, F., Li, X.D., Du, Y., Xiao, F., 2013. Sub-pixel mapping of remotely sensed imagery with hybrid intra- and inter-pixel dependence. *Int. J. Remote Sens.* 34(1), 341-357.
- Ling, F., Li, X.D., Du, Y., Xiao, F., 2014. Super-resolution land cover mapping with spatial-temporal dependence by integrating a former fine resolution map. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7(5), 1816-1825.
- Ling, F., Xiao, F., Du, Y., Xue, H.P., Ren, X.Y., 2008. Waterline mapping at the subpixel scale from remote sensing imagery with high-resolution digital elevation models. *Int. J. Remote Sens.* 29(6), 1809-1815.
- Ling, F., Zhang, Y.H., Foody, G.M., Li, X.D., Zhang, X.H., Fang, S.M., Li, W.B., Du, Y., 2016. Learning-based superresolution land cover mapping. *IEEE Trans. Geosci. Remote Sens.* 54(7), 3794-3810.
- Liu, D.L., Zuo, H.P., 2012. Statistical downscaling of daily climate variables for climate change impact assessment over New South Wales, Australia. *Clim. Change* 115(3-4), 629-666.
- Liu, S.M., Xu, Z.W., Song, L.S., Zhao, Q.Y., Ge, Y., Xu, T.R., Ma, Y.F., Zhu, Z.L., Jia, Z.Z., Zhang, F., 2016. Upscaling evapotranspiration measurements from multi-site to the satellite pixel scale over heterogeneous land surfaces. *Agric. For. Meteorol.* 230, 97-113.
- Liu, X.H., Kyriakidis, P.C., Goodchild, M.F., 2008. Population-density estimation using regression and area-to-point residual kriging. *Int. J. Geogr. Inf. Sci.* 22(4), 431-447.
- Liu, Y.H., 2006. Using the snesim program for multiple-point statistical simulation. *Comput. Geosci.* 32(10), 1544-1563.
- Lloyd, C.D., 2014. *Exploring spatial scale in geography* John Wiley & Sons.
- Mahmud, K., Mariethoz, G., Baker, A., Sharma, A., 2015. Integrating multiple scales of hydraulic conductivity measurements in training image-based stochastic models. *Water Resour. Res.* 51(1), 465-480.
- Malenovsky, Z., Bartholomeus, H.M., Acerbi-Junior, F.W., Schopfer, J.T., Painter, T.H., Epema, G.F., Bregt, A.K., 2007. Scaling dimensions in spectroscopy of soil and vegetation. *Int. J. Appl. Earth Obs. Geoinf.* 9(2), 137-164.
- Mandelbrot, B., 1967. How long is coast of britain? Statistical self-similarity and fractional dimension. *Science* 156(3775), 636.
- Margulis, S.A., McLaughlin, D., Entekhabi, D., Dunne, S., 2002. Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *Water Resour. Res.* 38(12), 18.
- Mariethoz, G., Renard, P., Straubhaar, J., 2011. Extrapolating the fractal characteristics of an image using scale-invariant multiple-point statistics. *Math. Geosci.* 43(7), 783-797.
- Mariethoz, G., Renard, P., Straubhaar, J., 2010. The direct sampling method to perform multiple-point geostatistical simulations. *Water Resour. Res.* 46, 14.
- Martinez, B., Gilabert, M.A., 2009. Vegetation dynamics from NDVI time series analysis using the wavelet transform. *Remote Sens. Environ.* 113(9), 1823-1842.
- Mascaro, G., Vivoni, E.R., Gochis, D.J., Watts, C.J., Rodriguez, J.C., 2013. Temporal downscaling and statistical analysis of rainfall across a topographic transect in Northwest Mexico. *J. Appl. Meteorol. Climatol.* 53(4), 910-927.
- Mechri, R., Ottele, C., Pannekoucke, O., Kallel, A., Maignan, F., Courault, D., Trigo, I.F., 2016. Downscaling Meteosat land surface temperature over a heterogeneous landscape using a data assimilation approach. *Remote Sens.* 8(7), 18.
- Mertens, K.C., De Baets, B., Verbeke, L.P.C., De Wulf, R.R., 2006. A sub-pixel mapping algorithm based on

- sub-pixel/pixel spatial attraction models. *Int. J. Remote Sens.* 27(15), 3293-3310.
- Mertens, K.C., Verbeke, L.P.C., Ducheyne, E.I., De Wulf, R.R., 2003. Using genetic algorithms in sub-pixel mapping. *Int. J. Remote Sens.* 24(21), 4241-4247.
- Molnar, P., Burlando, P., 2005. Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model. *Atmos. Res.* 77(1-4), 137-151.
- Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. *Biometrika* 37(1-2), 17-23.
- Muñoz, A., Blu, T., Unser, M., 2001. Least-squares image resizing using finite differences. *IEEE Trans. Image Process* 10(9), 1365-1378.
- Murphy, J., 2000. Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *Int. J. Climatol.* 20(5), 489-501.
- Najac, J., Lac, C., Terray, L., 2011. Impact of climate change on surface winds in France using a statistical - dynamical downscaling method with mesoscale modelling. *Int. J. Climatol.* 31(3), 415-430.
- Nguyen, M.Q., Atkinson, P.M., Lewis, H.G., 2005. Superresolution mapping using a hopfield neural network with LIDAR data. *IEEE Geosci. Remote Sens. Lett.* 2(3), 366-370.
- Nguyen, Q.M., Atkinson, P.M., Lewis, H.G., 2011. Super-resolution mapping using Hopfield Neural Network with panchromatic imagery. *Int. J. Remote Sens.* 32(21), 6149-6176.
- Nigussie, D., Zurita-Milla, R., Clevers, J.G.P.W., 2011. Possibilities and limitations of artificial neural networks for subpixel mapping of land cover. *Int. J. Remote Sens.* 32(22), 7203-7226.
- Openshaw, S., 1984. The modifiable areal unit problem.
- Parajka, J., Merz, R., Skøien, J.O., Viglione, A., 2015. The role of station density for predicting daily runoff by top-kriging interpolation in Austria. *J. Hydrol. Hydrol.* 63(3), 228-234.
- Pardo-Igúzquiza, E., Chica-Olmo, M., Atkinson, P.M., 2006. Downscaling cokriging for image sharpening. *Remote Sens. Environ.* 102(1), 86-98.
- Peng, J., Loew, A., Merlin, O., Verhoest, N.E.C., 2017. A review of spatial downscaling of satellite remotely sensed soil moisture. *Rev. Geophys.* 55,
- Peterson, G.D., 2000. Scaling ecological dynamics: self-organization, hierarchical structure, and ecological resilience. *Clim. Change.* 44(3), 291-309.
- Pettorelli, N., Mysterud, A., Yoccoz, N.G., Langvatn, R., Stenseth, N.C., 2005. Importance of climatological downscaling and plant phenology for red deer in heterogeneous landscapes. *Proc. Royal Soc. B* 272(1579), 2357-2364.
- Piles, M., Camps, A., Vall-Llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y.H., Walker, J., 2011. Downscaling SMOS-derived soil moisture using MODIS Visible/Infrared data. *IEEE Trans. Geosci. Remote Sens.* 49(9), 3156-3166.
- Pradhan, N.R., Tachikawa, Y., Takara, K., 2010. A downscaling method of topographic index distribution for matching the scales of model application and parameter identification. *Hydrol. Processes* 20(6), 1385-1405.
- Praskievicz, S., 2017. Downscaling climate-model output in mountainous terrain using local topographic lapse rates for hydrologic modeling of climate-change impacts. *Phys. Geogr.*, 1-19.
- Qin, J., Yang, K., Ning, L., Chen, Y., Long, Z., Han, M., 2013. Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote Sens. Environ.* 138(138), 1-9.
- Ramzan, M., Ham, S., Amjad, M., Chang, E.-C., Yoshimura, K., 2017. Sensitivity evaluation of spectral nudging schemes in historical dynamical downscaling for South Asia. *Adv. Meteorol.* 2017,
- Reichle, R.H., Entekhabi, D., McLaughlin, D.B., 2001. Downscaling of radio brightness measurements for soil moisture estimation: A four - dimensional variational data assimilation approach. *Water Resour. Res.* 37(9), 2353-2364.
- Renard, P., Marsily, G.D., 1997. Calculating equivalent permeability: a review. *Adv. Water Resour.* 20(5), 253-278.
- Robinson, W.S., 1950. Ecological correlations and the behavior of individuals. *Int. J. Epidemiol.* 15(3), 351-357.
- Rolland, C., 2003. Spatial and seasonal variations of air temperature lapse rates in Alpine regions. *J. Clim.* 16(7), 1032-1046.
- Rossi, R.E., Mulla, D.J., Journel, A.G., Franz, E.H., 1992. Geostatistical tools for modeling and interpreting ecological

- spatial dependence. *Ecol. Monogr.* 62(2), 277-314.
- Schoof, J.T., Pryor, S.C., Robeson, S.M., 2007. Downscaling daily maximum and minimum temperatures in the midwestern USA: A hybrid empirical approach. *Int. J. Climatol.* 27(4), 439-454.
- Schulze-Makuch, D., Cherkauer, D.S., 1998. Variations in hydraulic conductivity with scale of measurement during aquifer tests in heterogeneous, porous carbonate rocks. *Hydrogeol. J.* 6(2), 204-215.
- Schulze, R., 2000. Transcending scales of space and time in impact studies of climate and climate change on agrohydrological responses. *Agric. Ecosyst. Environ.* 82(1), 185-212.
- Serafimovich, A., Metzger, S., Hartmann, J., Kohnert, K., Zona, D., Sachs, T., 2018. Upscaling surface energy fluxes over the North Slope of Alaska using airborne eddy-covariance measurements and environmental response functions. *Atmos. Chem. Phys.* 18(13), 10007-10023.
- Shen, Z.Q., Qi, J.G., Ke, W., 2009. Modification of pixel-swapping algorithm with initialization from a sub-pixel/pixel spatial attraction model. *Photogramm. Eng. Remote Sens.* 75(5), 557-567.
- Shortridge, J.E., Guikema, S.D., Zaitchik, B.F., 2016. Machine learning methods for empirical streamflow simulation: A comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrol. Earth Syst. Sci.* 20(7), 2611-2628.
- Skøien, J.O., Blöschl, G., 2007. Spatiotemporal topological kriging of runoff time series. *Water Resour. Res.* 43(9), 2253-2286.
- Skøien, J.O., Merz, R., Blöschl, G., 2005. Top-kriging? geostatistics on stream networks. *Hydrol. Earth Syst. Sci. Discuss.* 2(6), 2253-2286.
- Song, X.D., Zhang, G.L., Feng, D.C., Zhao, Y.G., Yang, J.L., 2016. Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model. *J. Arid. Land.* 8(5), 734-748.
- Song, Y., Li, Y., Bates, B., Winkle, C.K., 2015. A Bayesian hierarchical downscaling model for south-west Western Australia rainfall. *J. Royal Stat. Soc.* 63(5), 715-736.
- Stanley, H. E., Meakin, P., 1988. Multifractal phenomena in physics and chemistry. *Nature* 335(6189), 405-409.
- Stein, A., Riley, J., Halberg, N., 2001. Issues of scale for environmental indicators. *Agric. Ecosyst. Environ.* 87(2), 215-232.
- Steinacker, R., Ratheiser, M., Bica, B., Chimani, B., Dorninger, M., Gepp, W., Lotteraner, C., Schneider, S., Tschannett, S., 2006. A mesoscale data analysis and downscaling method over complex terrain. *Mon. Weather Rev.* 134(10), 2758-2771.
- Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Math. Geol.* 34(1), 1-21.
- Su, Y.F., Foody, G.M., Muad, A.M., Cheng, K.S., 2012. Combining hopfield neural network and contouring methods to enhance super-resolution mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 5(5), 1403-1417.
- Sun, S.B., Chen, B.Z., Shao, Q.Q., Chen, J., Liu, J.Y., Zhang, X.J., Zhang, H.F., Lin, X.F., 2017. Modeling evapotranspiration over China's landmass from 1979 to 2012 using multiple land surface models: evaluations and analyses. *J. Hydrol.* 18(4), 1185-1203.
- Tahmasebi, P., Hezarkhani, A., Sahimi, M., 2012. Multiple-point geostatistical modeling based on the cross-correlation functions. *Comput. Geosci.* 16(3), 779-797.
- Tang, Y.W., Atkinson, P.M., Zhang, J.X., 2015. Downscaling remotely sensed imagery using area-to-point cokriging and multiple-point geostatistical simulation. *ISPRS J. Photogramm. Remote Sens.* 101, 174-185.
- Tao, K., Barros, A.P., 2010. Using fractal downscaling of satellite precipitation products for hydrometeorological applications. *J. Atmos. Oceanic Technol.* 27(3), 409-427.
- Tatem, A.J., Lewis, H.G., Atkinson, P.M., Nixon, M.S., 2001. Multiple-class land-cover mapping at the sub-pixel scale using a Hopfield neural network. *Int. J. Appl. Earth Obs. Geoinf.* 3(2), 184-190.
- Tatem, A.J., Lewis, H.G., Atkinson, P.M., Nixon, M.S., 2002. Super-resolution land cover pattern prediction using a Hopfield neural network. *Remote Sens. Environ.* 79(1), 1-14.
- Thornton, M.W., Atkinson, P.M., Holland, D.A., 2007. A linearised pixel-swapping method for mapping rural linear land cover features from fine spatial resolution remotely sensed imagery. *Comput. Geosci.* 33(10), 1261-1272.
- Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Econ. Geog.* 46(sup1), 234-240.

- Vereecken, H., Kasteel, R., Vanderborght, J., Harter, T., 2007. Upscaling hydraulic properties and soil water flow processes in heterogeneous soils: A review. *Vadose Zone J.* 6(1), 1-28.
- Verhoeve, J., Wulf, R.D., 2002. Land cover mapping at sub-pixel scales using linear optimization techniques. *Remote Sens. Environ.* 79(1), 96-104.
- Wang, J.F., Yin, Q., Tong, S.L., Ren, Z.P., Hu, M.G., Zhang, H.R., 2017. Prolonged continuous exposure to high fine particulate matter associated with cardiovascular and respiratory disease mortality in Beijing, China. *Atmospheric Environ.* 168, 1-7.
- Wang, J.F., Zhang, T.L., Fu, B.J., 2016. A measure of spatial stratified heterogeneity. *Ecol. Indic.* 67, 250-256.
- Wang, J.H., Ge, Y., Heuvelink, G.B.M., Zhou, C.H., 2015a. Upscaling in situ soil moisture observations to pixel averages with spatio-temporal geostatistics. *Remote Sens.* 7(9), 11372-11388.
- Wang, J.H., Ge, Y., Song, Y.Z., Li, X., 2014a. A geostatistical approach to upscale soil moisture with unequal precision observations. *IEEE Geosci. Remote Sens. Lett.* 11(12), 2125-2129.
- Wang, L., Wang, Q., 2013. Subpixel mapping using markov random field with multiple spectral constraints from subpixel shifted remote sensing images. *IEEE Geosci. Remote Sens. Lett.* 10(3), 598-602.
- Wang, L., Zhang, Y., Li, J., 2006. BP neural network based subpixel mapping method, In: H. DS., L. K. and I. G.W. (Eds), *Intelligent Computing in Signal Processing and Pattern Recognition*. Springer, Berlin, Heidelberg, pp. 755-760.
- Wang, Q.M., Shi, W.Z., Atkinson, P.M., 2014b. Sub-pixel mapping of remote sensing images based on radial basis function interpolation. *ISPRS J. Photogramm. Remote Sens.* 92(92), 1-15.
- Wang, Q.M., Shi, W.Z., Atkinson, P.M., Zhao, Y.L., 2015b. Downscaling MODIS images with area-to-point regression kriging. *Remote Sens. Environ.* 166, 191-204.
- Wegener, J., Lave, M., Luoma, J., Kleissl, J., 2012. Temporal downscaling of irradiance data via Hidden Markov Models on Wavelet coefficients: Application to California Solar Initiative data.
- Western, A.W., Grayson, R.B., Blöschl, G., 2003. Scaling of soil moisture: A hydrologic perspective. *Annu. Rev. Earth Planet. Sci.* 8(30), 149-180.
- Wigmosta, M., Prasad, R., 2006. Upscaling and downscaling—dynamic models. *Encyc. hydrol. sci.*
- Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change* 62(1-3), 189-216.
- Wu, H., Li, Z.L., 2009. Scale issues in remote sensing: a review on analysis, processing and modeling. *Sensors* 9(3), 1768-1793.
- Wu, J.G., 2007. *Landscape Ecology: Pattern, Process, Scale and Hierarchy* Higher Education Press, Beijing.
- Wu, J.G., Jelinski, D.E., Luck, M., Paul, T.T., 2000. Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics. *Geogr. Inf. Sci.* 6(1), 6-19.
- Wu, J.G., Jones, B.K., Li, H.b., Loucks, O.L., 2006. *Scaling and uncertainty analysis in ecology* Springer, Dordrecht, The Netherlands.
- Wu, L., Liu, X.N., Zheng, X.P., Qin, Q.M., Sun, Y.J., 2015. Spatial scaling transformation modeling based on fractal theory for the leaf area index retrieved from remote sensing imagery. *J. Appl. Remote Sens.* 9(1), 096015.
- Xu, G.H., Xu, X.L., Liu, M.X., Sun, A., Wang, K.L., 2015. Spatial downscaling of TRMM precipitation product using a combined multifractal and regression approach: demonstration for south China. *Water* 7(6), 3083-3102.
- Xu, X., Zhong, Y.F., Zhang, L.P., 2011. Sub-pixel mapping with multiple shifted remotely sensed images based on attraction model, Sino-foreign-interchange Conference on Intelligent Science and Intelligent Data Engineering. Springer, Berlin, Heidelberg
- Xu, X., Zhong, Y.F., Zhang, L.P., 2014. Adaptive subpixel mapping based on a multiagent system for remote-sensing imagery. *IEEE Trans. Geosci. Remote Sens.* 52(2), 787-804.
- Yang, F.H., White, M.A., Michaelis, A.R., Ichii, K., Hashimoto, H., Votava, P., Zhu, A.X., Nemani, R.R., 2006. Prediction of continental-scale evapotranspiration by combining MODIS and AmeriFlux data through support vector machine. *IEEE Trans. Geosci. Remote Sens.* 44(11), 3452-3461.
- Yang, K., Qin, J., Zhao, L., Chen, Y.Y., Tang, W.J., Han, M.L., Zhu, L., Chen, Z.Q., Lv, N., Ding, B.H., 2013. A multi-scale

- soil moisture and freeze-thaw monitoring network on the Third Pole. 94(12), 1907-1916.
- Yoo, E.H., Kyriakidis, P.C., 2006. Area-to-point kriging with inequality-type data. *J. Geog. Sys.* 8(4), 357-390.
- Yoo, E.H., Kyriakidis, P.C., 2009. Area-to-point kriging in spatial hedonic pricing models. *J. Geog. Sys.* 11(4), 381.
- Yoo, E.H., Kyriakidis, P.C., Tobler, W., 2010. Reconstructing population density surfaces from areal data: a comparison of Tobler's pycnophylactic interpolation method and area - to - point kriging. *Geog. Anal.* 42(1), 78-98.
- Young, L.J., Gotway, C.A., 2007. Linking spatial data from different sources: the effects of change of support. *Stochastic Environ. Res. Risk Assess.* 21(5), 589-600.
- Yue, T.X., Du, Z.P., Song, D.J., Gong, Y., 2007. A new method of surface modeling and its application to DEM construction. *Geomorphology* 91(1), 161-172.
- Zhang, H., Li, Z.L., 2012. Fractality and self-similarity in the structure of road networks. *Ann Assoc Am Geogr.* 102(2), 350-365.
- Zhang, J.X., Atkinson, P.M., Goodchild, M.F., 2014a. *Scale in spatial information and analysis* Taylor & Francis Group, LLC.
- Zhang, J.X., Zhu, T., Tang, Y.W., Zhang, W.L., 2014b. Geostatistical approaches to refinement of digital elevation data. *Geo Spat Inf Sci.* 17(4), 181-189.
- Zhang, L.P., Ke, W., Zhong, Y., Li, P.X., 2008. A new sub-pixel mapping algorithm based on a BP neural network with an observation model. *Neurocomputing* 71(10), 2046-2054.
- Zhang, T.F., Switzer, P., Journel, A., 2006. Filter-based classification of training image patterns for spatial simulation. *Math. Geol.* 38(1), 63-80.
- Zhang, T.P., Stackhouse, P.W., Gupta, S.K., Cox, S.J., Mikovitz, J.C., 2016. A generalized formulation for downscaling data based on fourier transform and inversion: mathematical rationale and application to the Max-Planck-Institute aerosol climatology data. *J. Quant. Spectrosc. Radiat. Transfer.* 188, S0022407316300462.
- Zhang, X.L., Yan, X.D., 2015. A new statistical precipitation downscaling method with Bayesian model averaging: a case study in China. *Clim. Dyn.* 45(9-10), 2541-2555.
- Zhang, Y.H., Du, Y., Ling, F., 2014c. Example-based super-resolution land cover mapping using support vector regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7(4), 1271-1283.
- Zhao, L., Yang, K., Qin, J., Chen, Y.Y., Tang, W.J., Lu, H., Yang, Z.L., 2014. The scale-dependence of smos soil moisture accuracy and its improvement through land data assimilation in the central Tibetan plateau. *Remote Sens. Environ.* 152, 345-355.
- Zhong, Y.F., Wu, Y.Y., Xu, X., Zhang, L.P., 2015. An adaptive subpixel mapping method based on MAP model and class determination strategy for hyperspectral remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* 53(3), 1411-1426.
- Zhong, Y.F., Wu, Y.Y., Zhang, L.P., Xu, X., 2014. Adaptive MAP sub-pixel mapping model based on regularization curve for multiple shifted hyperspectral imagery. *ISPRS J. Photogramm. Remote Sens.* 96(4), 134-148.
- Zhong, Y.F., Zhang, L.P., 2012. Remote sensing image subpixel mapping based on adaptive differential evolution. *IEEE Trans. Syst., Man, Cybern. B, Cybern.* 42(5), 1306.
- Zhou, Q.M., Li, B.L., Kurban, A., 2008. Spatial pattern analysis of land cover change trajectories in Tarim Basin, northwest China. *Int. J. Remote Sens.* 29(19), 5495-5509.
- Zhu, L., Gorman, D. M., Horel, S., 2006. Hierarchical bayesian spatial models for alcohol availability, drug 'hot spots' and violent crime. *Int. J. Health Geogr.* 5(1), 54.
- Zupanski, D., Zhang, S.Q., Zupanski, M., Hou, A.Y., Cheung, S.H., 2011. A prototype WRF-based ensemble data assimilation system for dynamically downscaling satellite precipitation observations. *J. Hydrol.* 12(1), 118-134.