Spatial Sampling, Data Models, Spatial Scale and Ontologies: Interpreting Spatial Statistics and Machine Learning Applied to Satellite Optical Remote Sensing

Peter M. Atkinson[a,b,c*], A. Stein[d] and C. Jeganathan[e]

[a] Lancaster Environment Centre, Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK; pma@lancaster.ac.uk
[b] Geography and Environmental Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[c] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A Datun Road, Beijing 100101, China
[d] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, 7500 AE Enschede, The Netherlands; a.stein@utwente.nl.
[e] Department of Remote Sensing, Birla Institute of Technology (BIT), Mesra, Ranchi 835215, India.

[*]Corresponding author: E-mail address: pma@lancaster.ac.uk.

This paper summarizes the development and application of spatial statistical models in satellite optical remote sensing. The paper focuses on the development of a conceptual model that includes the measurement and sampling processes inherent in remote sensing. We organized this paper into five main sections: introducing the basis of remote sensing, including measurement and sampling; spatial variation, including variation through the object-based data model; advances in spatial statistical modelling; machine learning and explainable AI; a hierarchical ontological model of the nature of remotely sensed scenes. The paper finishes with a summary. We conclude that optical remote sensing provides an important source of data and information for the development of spatial statistical techniques that, in turn, serve as powerful tools to obtain important information from the images.

Keywords: remote sensing; spatial statistical modeling; sampling; scale; ontology

1. Introduction

Remote sensing is a measurement-based discipline and as such it leads to the creation of data that have specific characteristics (Curran, 1985). The specific nature of these data has led to (or benefitted from) the development and application of many explicitly spatial statistical techniques. This paper serves to provide a historical look, over the last decade, at some key developments and applications of spatial statistical models in satellite optical remote sensing. Appendix 1 provides a classification of the references in this paper. However, more fundamentally, this paper focuses on the development of conceptual models of the measurement and sampling processes inherent in remote sensing, the nature of spatial information in remotely sensed images, and the nature of the real scenes that remotely sensed data are created from (Quattrochi and Goodchild, 1997; Stein et al., 1999; Atkinson and Tate, 2000; Goodchild, 2004; Miller, 2004; Ge et al., 2019). We also consider briefly some of the concepts underlying the spatial statistical techniques themselves. Thus, this paper seeks to question the nature of remotely sensed data and information, arguing that future research in spatial statistics for remote sensing should be guided by the concepts that emerge. This paper is not a review of methods that have emerged over the last decade. Such a task would be challenging given the volume of activity and production.

We organized this paper into seven sections as follows. Following this introduction, we consider measurement and sampling processes in remote sensing after briefly introducing the basis of remote sensing as a tool. We then consider the spatial variation and potential information in these data, using the spatial covariance function as a very crude lens with which to analyze continuous variation. We also consider variation through the object-based data model. Advances in spatial statistical modelling are considered in section 4, with examples of key developments being multiple-point geostatistics, mixed (spatial) regression models using the Bayesian inference paradigm, and fuzzy objects. In section 5, we examine machine learning, deep learning and explainable AI, drawing out some key concepts from these methods that we use to assess the appropriateness of these approaches for certain tasks, and develop further our conceptual models. Section 6 extends the learning from earlier sections to develop a hierarchical ontological model of the nature of remotely sensed scenes of interest, which then allows us to reflect further on the appropriateness of techniques and gaps that may demand new spatial statistical modelling approaches. Section 7 provides a summary.

2. Remote Sensing as a Source of Data

In this section, we develop a conceptual model of remote sensing as a source of spatial (and space-time) environmental *data*. In so doing, we set the basis for subsequent chapters which aim to analyze the data produced. The basic tenet is that principled methods for handling remotely sensed data should consider the ways that the data were produced.

2.1 *The basic concept underlying remote sensing*

We start by reminding readers of *Spatial Statistics* of the basic concepts underlying satellite optical remote sensing (Curran, 1985). The material is rudimentary, but it serves to build the proposition that this paper majors on in later sections.

In satellite optical remote sensing, light from the Sun traverses the atmosphere (where it is marginally scattered, refracted and so on), and eventually reaches the Earth's surface. Depending on the surface material and its properties, the light is (i) absorbed, (ii) transmitted through to a subsequent layer and (iii) reflected in three proportions summing to one. Light exists across a continuum of wavelengths referred to as the electromagnetic spectrum (EMS), with optical light representing the visible and infrared wavelengths. Conditional upon the wavelength, the light may be absorbed, transmitted and reflected in different proportions, thus, producing spectra. The *reflected* light traverses through the atmosphere again, where it is scattered, refracted and so on, before exiting the Earth's atmosphere. Satellite optical remote sensing is then the task of measuring from space the reflected light such as to *inversely* infer some properties of the material at the Earth's surface. This inverse process is the fundamental basis of remote sensing. Thus, it can be seen that remote sensing is a *tool* for measurement, much like a telescope or microscope. As a measurement tool, it is important to consider the measurement properties of remote sensing devices as this will have a bearing on subsequent spatial statistical analysis of the data.

Note: other types of remote sensing are common, including (i) in different wavelengths (e.g., microwave remote sensing, which focuses on extracting surface texture, di-electric properties of soil and polarization characters of various features using the microwave part of the electromagnetic spectrum) and (ii) with different platforms (e.g., the airborne and UAV platforms that can provide rich information about object shape, size, orientation, texture and contextual relationship). However, much of what is described here for optical remote sensing translates readily to these other types.

2.2 *Discretization and the measurement process*

In all cases of remote sensing, to measure one must discretize. This discretization occurs across the EMS, across space and across time, amongst others (e.g., numerical precision, angle of view). From the perspective of spatial statistics this amounts to an important sampling decision (and potentially a discretization of the actual space operated on) and so we review it briefly here.

In optical remote sensing, it is common to measure the EMS in broad segments called 'wavebands' (i.e., bounded integrals over the EMS). Satellite sensors such as Landsat operate a few broad wavebands, commonly referred to as multispectral remote sensing (Arvidson et al., 2006; Williams et al., 2006; Wulder et al., 2008, 2012; Yan and Roy, 2016). Variation between these wavebands can be used to infer properties of the material at the Earth's surface. For example, a high reflectance in the green and near-infrared wavebands is characteristic of vegetation, whereas a relatively low reflectance in all wavebands is characteristic of water, which tends to absorb light across the EMS (Curran, 1985). In contrast, hyperspectral optical sensors measure reflected light in many hundreds of wavebands, leading to the possibility to make more nuanced inferences about the Earth's surface materials.

Across space, discretization occurs through the array of cells that constitute the ‘sensor’. Modern optical sensing devices generally involve a rectangular array of cells onto which the reflected light is projected via a lens. This discretization decision allows the production of an image, which has great utility from a spatial statistical modelling perspective. However, it also necessarily invokes the three concepts of support, spatial resolution and pixel (Atkinson and Tate, 2000). The difference between the first two is subtle, but essentially hinges on the fact that the support is a first-order concept (being defined for a single measurement) while the spatial resolution is a second-order concept (depending on more than one observation). The pixel of each remotely sensed image is neither of these, being rather, simply an element or cell (of the data array or image) to which a measurement value is allocated. The support is a geostatistical concept representing the space on which a measurement is made, or observation is defined, and it has three parameters; size, geometry and orientation (Atkinson and Tate, 2000; Ge et al., 2019). It represents one element of the spatial (space-time) sampling strategy, with the other elements being the *pattern* of observations and the *extent*.

It is notable that the support of measurements in remote sensing is commonly represented by a 2D Gaussian function (or similar function) referred to as the ‘point spread function’ (PSF), with its tails extending far beyond the limits of a pixel (Wang et al., 2020). It is in this important regard that the support (PSF) is different to the way that most people imagine measurement on a pixel (which could be better described as having a ‘square wave response’). Far too little spatial statistics research in remote sensing has accommodated the spatial sampling effects of the PSF.

In time, discretization occurs through individual images which represent cross-sections through time. Time-series of remotely sensed images (i.e., space-time cubes) can be constructed readily because some of the world’s most popular and long-standing satellite sensor series (e.g., NOAA-AVHRR, Terra/Aqua-MODIS, Landsat-TM/ETM/OLI,) have been acquiring images for decades with a fixed revisit interval (e.g., 16 days for Landsat TM, 1-to-2 days for AVHRR/MODIS) (Arvidson et al., 2006; Williams et al., 2006; Wulder et al., 2008, 2012; Zhu and Woodcock, 2014; Yan and Roy, 2016). Unfortunately, the ubiquitous problem of cloud cover means that the frequency of useable images (or parts of images) is lower (i.e., longer) than the revisit intervals, but with appropriate statistical methods, complete times-series of images can be constructed (Song and Huang, 2012; Mondal et al., 2017; Wang and Atkinson, 2018; Belgiu and Stein, 2019; Guo et al., 2020) (see section 4).

It is interesting to consider that the above discretization processes in remote sensing determine to a large extent the nature of the spatial statistical models that might be applied to the data subsequently. The most obvious impact (i.e., constraint on subsequent statistical model choices) is that remotely sensed images are discretized across space into pixels. This means potentially that the Euclidean space itself is discretized into a regular grid of possible values. Operations that are made directly on that grid generally deny the underlying continuous space of the real world, an insight not dissimilar to that from aggregation in the so-called modifiable areal unit problem (MAUP; Openshaw, 1984; Fotheringham and Wong, 1991).

From a spatial statistical perspective it is useful to distinguish between continuous random fields or Random Functions (RFs) (also, in the specific case, termed Gaussian Processes, GPs) (that are stochastic in their attribute), and spatial objects (that could be stochastic in their geometry). Both stochastic models of the real world may be useful in different circumstances. In both cases, it is possible to fit such models to the image data directly, but this imposes the discretized space and constrains the solution to be on a a grid. For example, it is possible to define objects in remotely sensed images by grouping the labels of nearby clusters of pixels, but these objects will be blocky as a result and oriented in the same direction as the image overall (Aplin and Atkinson, 2001). Alternatives that escape the strictures of the pixel and image grid are possible and have gained much attention recently. The most obvious example of this comes from geostatistical change of support theory (Cressie, 1996; Kyriakidis, 2004; Atkinson, 2013). This is discussed further in section 4.

2.3 *Measurement error*

While describing the nature of remotely sensed data as a consequence of sampling decisions, it is worthwhile to make a philosophical statement about the nature of measurement error. Conceptually, we believe that all measurements about the real world are integrals over space and time; that is, they have a support in space-time, with a particular size, geometry and orientation. Measurement error is then added to this integral with a particular distribution (Atkinson and Tate, 2000). We have no evidence for this sequence of 'integral-then-error', but it is a useful, and rather natural, conceptual construct.

The measurement error can arise from many different sources, including sensor noise, atmospheric attenuation and uncertainty in the PSF definition. Measurement error can involve random error, but it can also involve systematic error. Not enough attention has been paid to accommodating this important source of uncertainty in spatial statistical models applied to remotely sensed data.

3. Information in Remotely Sensed Data

In this section, we develop some concepts related to the extraction of information from remotely sensed data. We do this in two parts: by considering continuous spatial variation and by considering spatial objects. First, we define information crudely.

3.1 *Definition of information*

For the purposes of this paper, we define information (or at least *potential* information) as the difference between data, values or things. In a single waveband remotely sensed image, therefore, potential information exists in the differences between pixels. Since, mathematically, the difference between two pixels A and C separated by, and joined by, a third pixel B in-between them is already represented in the two relations A-B and B-C, it is clear that potential information in an image is local, existing only between neighbouring pixels that share a common boundary (edge) (specifically the King's neighbourhood case).

Differences also can exist with data *outside* the image, and differences can exist between the image and the expectations of the viewer, which are amassed as a function of experiences, and generalizations of these experiences, over time (to summarize the complex cognitive process of the human brain). Nevertheless, the definition of (potential) information as difference holds in all these three cases; the concept is general (Wang et al., 2019).

It is in this context that the variance parameter of a Gaussian statistical distribution is useful as a diagnostic since it is based on difference and describes what is expected *on average*. It is particularly useful when extended spatially into the spatial covariance function that parameterizes a RF (or GP).

3.2 *Spatial variation on a grid*

Accepting that remotely sensed measurements are made generally on an image grid, let us start by considering the spatial variation that exists amongst the pixel values on that grid. We focus on the RF model for illustrative purposes; in particular, its parameterization through the spatial covariance, but many other approaches could be used in its place.

It is possible to calculate the empirical spatial covariance and to fit a model to it using an appropriate method of inference. Common permissible (authorized) functions include the Matérn family of models, including the popular exponential covariance model. The exponential model has two parameters, the so-called sill variance and the range (or pseudo-range). To a certain extent, the sill variance of the exponential model can be thought of as the spatial equivalent of the point variance, although strictly it is the *a priori* variance and not the sample variance (Journel and Huijbregts, 1978). The range on the other hand has no equivalent in the point distribution.

The range parameter, as a simplifying representation, tells us a lot about the potential information content of the image. For example, if the range is long (large) relative to the extent of the image then there is much redundancy in the image (more data for little information); conversely if it is short (small) there is much potential information relative to the number of data. In a related sense, the range informs about the scale(s) of spatial variation present in the image. Useful references on scale in remote sensing and geography have been provided elsewhere (Atkinson and Tate, 2000; Wu and Li, 2009; Goodchild, 2011; Lloyd, 2014; Zhang et al., 2014a; Jiang and Brandt, 2016; Jiang, 2018).

Beyond the range, which represents an upper limit on the extent of any correlation, the *shape* of the spatial covariance function also is informative. For example, one can think of the exponential model, which is asymptotic towards the sill variance, as representing a *set* of scales of variation, each with its own information-to-redundancy ratio. Put differently, and invoking briefly the object-based view of the world, if the image were comprised of objects, the objects would be of different sizes.

Note: the above is a coarse statement to illustrate the concept only (e.g., the range is independent of the number of times that a pattern is repeated). Nevertheless, we contend that such insights are potentially useful. A recent trend in spatial statistics applied to remote sensing has been to no longer analyze spatial statistical functions such as the spatial covariance for what they tell us about the nature of the property of interest. This is discussed again in section 4.

Despite the above insights about the scales of spatial variation and potential information (and redundancy), what is intelligible or interpretable (and pleasing) to a human being is not the same as 'potential information'. An image that is rich with potential information can be difficult to 'read' by a human being. This is for two reasons: (i) human beings tend to naturally identify functional objects and invoke the simplifying object-based view of the world and (ii) it is easier to identify only a few things than many things. This is why we differentiated between potential information (with the underlying variation characterized on average by the spatial covariance) and what we would think of more naturally as information (e.g., the underlying variation collapsed down further into semantically meaningful, functional, object-based representations).

3.3 *Spatial variation on a continuous space*

If one considers again the discretization process described in section 2.2, it is not difficult to see that the values in pixels are integrals (plus some measurement error). As such, the discussion in section 3.2 above relates to the differences *between* pixels in an image and nothing is said about the differences *within* pixels that have been obscured through the measurement process. Through measurement, all the variation (potential information) within the support (i.e., PSF) is reduced down to a single value and all that remains in terms of potential information lies in the differences between the pixel values (and differences with other data, and with the interpreter's expectations). This statement is obvious, but it also has profound implications for the principled statistical handling of remotely sensed imagery.

It is the intersection of the sampling strategy (spectral, spatial, temporal) implicit in the imaging sensor with the real world that determines the spatial variation and potential information content of remotely sensed imagery (noting, importantly, that spatial variation exists only in data after measurement and not before it). A major parameter of the sampling framework is the support, with its three sub-parameters. In terms of information (i.e., neglecting uncertainty momentarily), it is not whether the support is large or small that matters; it is whether the support is large or small relative to the spatial range (also frequency) of the variation that is produced in the data, and especially the variation due to the features of potential interest to the investigator. If the support is too large, the variation may not be resolvable. If the support is too small, there may be too much redundancy in the image. Interestingly, as alluded to above, the sweet spot for human interpretation generally involves a lot of redundancy. The human brain requires some redundancy in order to resolve structure, or to 'see' functional objects.

The consequence of acknowledging that within-support variation is lost through measurement is to reconsider the nature of the data that spatial statisticians can operate on and the specification of the models that are appropriate to fit to the data. For example, downscaling and image fusion have become a very popular topics in geostatistics and spatial statistics applied to remote sensing (Song and Huang, 2012; Sales et al., 2013; Wang and Atkinson, 2018; Belgiu and Stein, 2019; Guo et al., 2020; He and Yokoya, 2020; Dadras Javan et al., 2021; Li et al., 2021). It is possible, conceptually, to define the stochastic model at the point support scale and to fit such point support models to data observed on a positive finite support. This insight is crucial and it is leading to spatial statistical models that try to escape the strictures of the measurement processes that created the data in the first place. This is important because, after all, in environmental and related sciences, our interest is not generally in the data; it is (or should be) in the real world.

Geostatistical change of support (CoS) models do this in some respects (e.g., in that the RF is spatially continuous), but not in others (e.g., the spatial covariance is defined initially on a positive measurement support, and subsequent inference at a finer support is ill-posed) (Kyriakidis, 2004; Yoo and Kyriakidis, 2006; Liu et al., 2008; Yoo et al., 2010; Wang et al., 2015; Jin et al., 2018). Despite their success and widespread adoption in recent years CoS models represent a ‘step along the way’. Thus, this paper makes a call to spatial statisticians to reconsider the remotely sensed image, not as the object of study, but as a partial window on the real world, and to design spatial statistical models that acknowledge this deficit. CoS models are described further in section 4.

3.4 *Spatial objects*

It is important to view the consequence of discretization across space in relation to the object-based model. As discussed above, humans naturally identify and label functional objects in their surroundings. They do this primarily to survive; an evolved ability. However, commonly these functional objects (e.g., car, telephone, desk) are human constructs only; strictly they do not exist in the real world. It can be reasonably argued that animals and plants are singular objects in the real world (in the sense that they are singularly integrated collections of biochemical processes), a view common in ecology (Forman, 1995), but it is also true that they are simultaneously collections of physical particles (i.e., not objects at all). Such a philosophical discussion is important, but beyond the present scope.

Despite the above, if we can accept the legitimacy of the existence of spatial objects in the *real* world, then their intersection with a regular grid of measurement cells with a particular support creates spatial data on those objects. From these data, inversely, the objects may be identified and labelled. However, the ability to do this depends on the interaction between the support and those spatial objects in the real world, specifically the size of the support relative to the size of the objects. Too large a support and the object may not be sufficiently resolvable. Too small a support and the object may be identified, but at large data redundancy cost. If the objective is to resolve the variation in the geometry (boundary) of the object in detail then an even smaller support may be required.

Lying between measurement and the ability to identify and label the original objects of interest is the concept of the ‘mixed pixel’ (Peng et al., in press). Mixed pixels occur when more than one object class contributes to the overall signal measured and allocated to a pixel. For example, if the interest is in identifying cars in a car park, the intersection of the support with the scene may lead to many pixels within the image that are partially car and partially car park. These so-called mixed pixels occur along the boundaries of the spatial objects of interest. The existence of mixed pixels demands attention to the spatial support issue and the selection of statistical methods that address this problem head on. Since remotely sensed images commonly cover scenes that comprise multiple objects, the mixed pixel problem is fairly ubiquitous.

The goal of spatial statistical analysis is not always focused on the segmentation and labelling of objects, of course (see section 3.2 and 3.3 above). However, it is important to consider that scenes comprised of objects (i.e., phenomena that humans would readily identify as functional objects with semantic meaning) are the norm in remote sensing. For example, in an urban area, an image may include data relating to buildings, gardens, garages, retail outlets, industrial buildings, car parks, roads, rivers, train lines, and so on. In this context, it is interesting that the focus of much spatial statistical analysis in remote sensing has been based on the analysis of spatial continua (e.g., through application of regression models and geostatistical RFs) (Moran, 1950), and less so stochastic objects and their boundaries (e.g., Mandelbrot, 1967). We feel that application of RFs in remote sensing should be done acknowledging the spatial object-based nature of the variation in images, with attention also paid to the stochastic analysis of objects and the graphs that connect them. We discuss this problem further in Section 4.

3.5 A note on sampling

Developments in spatial sampling design have received much benefit from remote sensing (Wang et al. 2012). Spatial sampling design methods can be categorized into model-based and design-based sampling. Model based sampling requires the use of an optimization function, such as equal spreading or obtaining the minimized geostatistical Kriging variance, while design-based sampling requires a random component in the sampling design. A relatively straightforward procedure is to implement a design on the discrete pixels in an image. In such a case, it is straightforward to allocate a spatial statistical sampling design such as random or grid sampling, where the pixels to be sampled are identified. Similarly, a model-based optimal sampling strategy can be implemented to optimize the classification of an image, for example, using the Kappa statistic.

Of some interest in relation to sampling design is the variability *within* a pixel. Such variability is commonly ignored by averaging the within-block variability to create and allocate a single reflectance value to the pixel, the support of which is governed by the point spread function. Rulinda et al. (2011) undertook sampling within a pixel: field data were collected within single pixels of the MSG-Seviri NDVI product. Its spatial resolution of approximately 5 x 5 km at the latitude of the study area (Rwanda) was too coarse to provide reliable information for the purpose of properly studying NDVI variability, and a statistical design was implemented within five individual pixels. Two transects in the EW direction were allocated at random positions on the NS-axis within the area projected on the ground; similarly two transects were positioned at random positions on the EW axis to investigate the variability in the NS direction. It was, thus, possible to characterize the within-pixel spatial variability.

More recently, Wang et al. (2020) suggested the spatial statistical trinity. In this generic framework a relation is presented between universe, sampling and inference. This conceptual integration is useful because it points to the utility of designing model-based approaches that are connected fully through this trinity, and which identify the best choice amongst various estimators for a universe or population under study. We suggest that more attention is required to develop this trinity further, for example, to generalize sampling to escape the strictures of spatial discretization and the spatial support, amongst others.

Having introduced some fundamental concepts in relation to measurement and spatial information, we now review some recent developments in spatial statistics (Section 4) and in machine learning (Section 5) in remote sensing.

4. Characterizing Imagery Using Spatial Statistics

As introduced above, remote sensing images consist of data, commonly represented as 'digital numbers' (Section 2), while the interest is generally in extracting information from these data (Section 3). The spatial statistical models that *characterize* the variation in spatial data, and which allow us to predict or forecast (predict in the future) some property of interest are, thus, critical.

For spatial statistical modeling, we see major recent contributions as threefold: (i) developments in geostatistical change of support theory and multiple point geostatistics, that depend on higher-order moments; (ii) the development and application of explicitly *spatial* statistical regression models, which supersede traditional linear regression models by introducing a spatial dependence term between pixels, which plays a critical role, and which requires the *Bayesian inference* paradigm; and (iii) the handling of *objects*, such as by random sets and fuzzy objects, where the spatial variation of the content, and random delineations of object boundaries, are the major uncertainties to address.

4.1 *Geostatistics*

Geostatistics, emerging since the 1960s in mining (Cressie, 1993; Journel, 1993; Cressie and Wikle, 2015), has since the 1980s served as a spatial prediction engine in remote sensing. Geostatistics serves as a useful and important statistical model for (i) handling *missing data* (e.g., generated by the presence of clouds or cloud shadow, or by a failing sensor component); (ii) *upscaling* from a set of pixels to a homogeneous object; (iii) *downscaling* from a coarse pixel that covers a certain area on the ground to a finer spatial resolution; and (iv) *fusion* of images of a certain spatial resolution with other images of a different spatial resolution. Major contributions came from Atkinson et al. (1992, 1994) and Addink and Stein (1999). These concerned the absence of pixels and filled in the empty pixels by geostatistical interpolation. Many examples exist of the development of spatial statistical models for filling gaps due to clouds and cloud shadows (Chen et al., 2014, 2020, 2021) and due to sensor failures (dropped pixels) (Chen et al., 2011; Chen et al., 2012; Wang et al., 2021).

At the beginning of the century, scaling issues became more prominent, and several groups advanced the field in the search to address the challenging issue of downscaling spatial continua (Cressie, 1996; Kyriakidis, 2004; Pardo-Igúzquiza et al., 2006; Goovaerts, 2006, 2007; Atkinson et al, 2008; Atkinson, 2013; Huhtengs and Vohland, 2016; Wang et al., 2016; Jeganathan and Mondal, 2017). Even greater effort was paid to the challenge of downscaling reflectance to categories (referred to as sub-pixel mapping) (Atkinson, 1997; Tatem et al., 2001, 2002; Atkinson, 2005; Khasetkasem et al., 2005; Thornton et al., 2007; Tolpekin and Stein, 2009; Ardila et al., 2011; Nguyen et al., 2011; Su et al., 2012; Ling et al., 2013; Ai et al., 2014; Wang et al., 2014; Hu et al., 2015; Ge et al., 2016; Chen et al., 2018). Both change of support goals aim to escape the strictures of the pixel in remote sensing. As introduced above, the pixel is seen commonly as the average reflectance of light from a limited support on the terrain, and disentangling the reflectance into a set of finer supports representing the original reflectance or mapping to (e.g., land cover) categories is challenging (and ill-posed mathematically). This can be done if additional information is available. For example, a support of, say 30 m x 30 m (a common unit in Landsat images) in an agricultural area may consist of a building, pavement and agricultural fields. If we have some prior expectation of where the infrastructure elements may be located, than it is essentially possible to downscale the aggregated signal into more specific spatially resolved information.

As introduced in Section 2, a key concept in geostatistical analysis is the spatial covariance function, or its related function, the semivariogram (Rossi et al., 1992). Both characterize the spatial variation and parameterize a RF. Under the condition of second-order stationarity of the RF, the two are related by a simple expression. With this in mind, remotely sensed images can be conceptualized as a realization from a RF. However, the appropriateness of this stationarity decision may be somewhat hard to maintain from a geographical point of view. It is more natural to consider that the pixel values as generated by the reflectance from crisply defined spatial objects on the ground, such as agricultural fields, buildings and water bodies. More reasonably, the spatial variation within the objects, can be considered as homogeneous and generated by a stochastic RF, while the variation *between* objects less so.

A typical way ahead is as follows. One first identifies spatial units, usually related to land cover, that capture the major distinguishing elements in the scene. Such spatial units are obtained by segmentation and classification of, for example, multi-band images. Often, these units have a clear physical meaning. Next, the seemingly homogeneous units are further considered: variability exists within these units and such variability, traditionally expressed by the standard deviation, is currently better described by RFs. Such a stratified approach distinguishes the between-strata variability from the within-strata variability. While this approach seems comprehensive, we note that such integrated modelling is relatively rare and, further, there exists a scarcity of attempts to do this incorporating change of support theory.

The last decade has seen the emergence of multiple-point geostatistics (MPG; Guardiano and Srivastava, 1993; Strebelle, 2002; Liu, 2006; Mariethoz et al., 2010; Ge and Bai, 2011; Straubhaar et al., 2011; Tahmessabi et al., 2012; Bai et al., 2013; Ge et al., 2013; Tang et al., 2015). Geostatistics based on the stationary covariance-based RF model is limited because the spatial covariance function is a two-point statistic (Atkinson, 2004). This means that moments are limited to first and second-order and, thus, by definition such RF models can simulate only very simple images with variation that is smooth and continuous and that lack detail and information. MPG replaces the spatial covariance with a training image from which rich, higher-order moments can be obtained (Strebelle, 2002). There exist many concerns over the MPG approach, but it does bring the convincing advantage of being able to simulate more realistic remotely sensed images.

In the above context, we argue two things: (i) researchers developing and applying stochastic RF models in remote sensing should pay more attention to the stochastic modelling of objects and their boundaries (as well as the combination of the object-based and RF models) and (ii) the geostatistical RF (or GP) is surprisingly limited in its ability to characterize data and it is curious to us that multiple-point geostatistics and related higher-order moment approaches do not seem to have found wide application outside of a few key research groups. Put differently, and notwithstanding their specific utilities, given that it is so obvious that stationary covariance-based RFs (GPs) are unrealistic and inappropriate representations, why are they still so ubiquitous in remote sensing?

4.2 *Mixed models and the Bayesian inference paradigm*

Spatial statistical modeling of remote sensing images is based commonly upon the linear model. Such models can be applied at the individual pixel level or at the object level. Often, the linear model falls short because of its assumptions of i.i.d. residuals. More commonly, the presence of spatial dependence in the residuals from the regression model needs to be acknowledged and taken into account. This has led to the development of mixed regression models or spatial regression models.

An early example of a mixed regression model was the autologistic regression model (Augustin et al., 1996). This combined a generalized linear model (GLM; a linear regression model predicting continua augmented with a link function on the predictand mapping the prediction to some other data type; in the autologistic case a binary outcome). The autologistic model was fitted using the Gibbs sampler. Augustin et al. (1996) presented an interesting and relevant study on biodiversity where they developed the autologistic model based on the Gibbs sampler in a remote sensing context.

A spatial regression model can be conceptualized as being an additive model that in the most simple case combines two effects; a linear fixed effect and a spatial random effect, plus an error term. The linear fixed effect term is the usual linear (or GLM) model, while the spatial random effect can be a geostatistical model as, for example, in regression kriging or a conditionally autoregressive (CAR) or simultaneous autoregressive (SAR) term, as in Augustin et al. (1996), amongst others.

A major difference between the geostatistical and CAR/SAR spatial random effect that usually guides the choice of the most appropriate approach is the aggregation level of the data. Geostatistics commonly is used to deal with data on a grid or on a quasi-point support, while SAR and CAR models are most suited for data represented on irregular supports such as census Wards.

Mixed regression models generally require inference using the Bayesian inference paradigm. In the seminal paper of Diggle et al. (1998) for the first time the term “model-based geostatistics” was coined as a major step forward to integrate geostatistics (and the mixed regression model approach) with Bayesian inference. This paper attracted major attention and put modern and computer-intensive geostatistical modeling of spatial data into a wider statistical context. In relation to classical geostatistics, the key gain of the “model-based geostatistics” approach was the important step to admit and model the uncertainty in the parameters of the RF. At the same time, the approach emphasized explanation through covariates over geostatistical prediction, which was relegated to operating on the linear model residuals. This shift in emphasis can be argued strongly from an inference and prediction perspective, but it has led to a reduction in attention on the characterization of spatial variation.

4.3 *Objects*

Spatial statistics is well developed when it comes to the identification of segments and allocation of classes. Based on image analysis, homogeneous spatial objects can be identified that are then assigned a class label. A key goal in remote sensing is to spatially segment imagery, and various methods have been developed. *K*-nearest neighbor classification and maximum likelihood classification are archetypal tools for classifying images consisting of multiple wavebands. Problems emerging in hyperspectral images where the number of bands (typically some hundreds) can be prohibitive for applying these image classification methods have largely been overcome, while segmentation and classification of single band images is of a simpler nature. However, the objects and their classes emerging from segmentation and classification have inherent uncertainties. On the one hand, objects are rarely homogeneous internally, and thresholds or processing adjustments have to be applied to overcome over-segmentation and the emergence of anomaly classes. On the other hand, the spatial boundaries between classes are often far form crisp, even in the natural world. Fisher et al. (2004) recognized this most clearly when posing the question: Where is Helvellyn? The mountain clearly is somewhere, but the edges of the mountain are gradually there, and it is impossible to state with full confidence when, during a hike, one steps for the first time on the mountain.

In a range of papers, attention was given to random sets as a methodology to represent uncertain objects. Zhao et al. (2010) studied the uncertainty of lake boundaries, Zhou et al. (2013) focused on traffic objects derived from LiDAR data, Sidiropoulou Velidou et al. (2015) investigated the occurrence of linear geological objects, while Kohli et al. (2016) studied the delineation of slum areas in different cities around the globe. Random sets are based on probabilistic functions, and specify the probability that an object is present at a location. Their application leads to identification of the core of an object (i.e., the area where the object exists with certainty, the support of an object), and intermediate areas where an object is present with different degrees of probability. As Figure 1 shows, the core set (C) is most certainly the object, and the region $U \supset C$ represents an area where possibly the object exists as well, while the white area outside U, i.e. U^C is outside the support, and does not include the object.

In Kohli et al. (2016) the knowledge of 19 experts was used to delineate a slum area. In certain parts of the image, all experts agreed on the presence of a slum resulting in identifying C. In some other parts the experts agreed that there was no slum, hence identifying U^C. In the remaining parts there was no consensus, resulting in intermediate interpretations. The fraction of agreeing experts then served as the probability of the presence of a slum. In Zhao et al. (2010) a spatio-temporal analysis was used, where the probability of a lake was identified as its presence in a time-series of 12 monthly periods over a time span of 10 years.

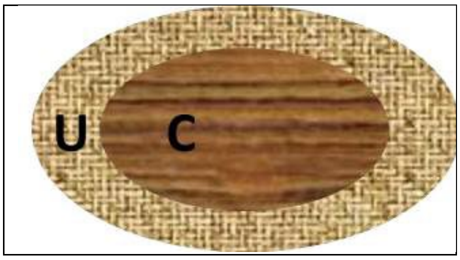

Figure 1. Random set representation of an uncertain object.

Research and thinking on fuzzy set objects extends the conceptual model that we build in this paper, and the statements that we make in relation to this model, such as calls for more research on spatial objects, should include considerations of their often fuzzy definition in reality.

5. The Rise and Rise of Machine Learning

Machine learning and deep learning have seen considerable success and widespread application in remote sensing in recent years. This resurgence is interesting, not least because earlier in the careers of the authors of this paper, we saw the “rise-then-decline” of artificial neural networks (ANNs) in remote sensing in the 1990s (Atkinson and Tatnall, 1997). That decline over two decades ago occurred primarily as a result of the criticism of statisticians and scientists that ANNs were ‘black box’ models. That is, one could not easily interpret (or control) what was going on inside the box because the ANN comprised so many parameters (e.g., weights between nodes in the multiple layers of the ANN). This lack of interpretability remains a concern with modern machine learning approaches although progress is being made through ‘explainable AI’ (XAI). Nevertheless, the vast numbers of data associated with remotely sensed images, and with time-series of remotely sensed images, in particular, have meant that the data-led approaches of machine learning were destined to find their niche in remote sensing.

The recent resurgence we see as occurring primarily in three phases, the first relating to machine learning, the second relating to deep learning and the third to explainable AI (XAI).

5.1 *Machine learning*

The first resurgence occurred through the development and adoption of specific techniques that were ‘game changers’ (i.e., which brought sufficient novelty to update the community’s thinking about what was possible). The archetypal example of machine learning in remote sensing was the feed-forward, back-propagation ANN. Essentially a flexible, nonlinear regression model, the ANN was applied widely with high accuracy in the 1990s (Atkinson and Tatnall, 1997).

Focusing primarily on classification (as opposed to regression) of remotely sensed images, the techniques that brought paradigm shifts were, for example, the support vector machine (SVM; Yang et al., 2006; Zhang et al., 2014b) which demonstrated that training data (support vectors) near the (non-linear) boundaries in feature space between classes are more important (indeed all that is needed) relative to those further away, and they could be identified through local kernels, and (ii) the Random Forest, a tree-based classifier that when fitted has the advantage of having an expression spatially as a non-stationary model, allowing generalization of parameters to local conditions, and which brings the added advantage of identifying the importance of each input feature (Brieman, 2001; Rodriguez-Galliano et al., 2012; Huhtengs and Vohland, 2016). Such machine learning algorithms shifted the balance of attention from the model to the data and, thus, were ideal for taking advantage of the massive numbers of data produced by remote sensing satellites.

At the same time, a disadvantage of machine learning algorithms arises for precisely the same reason as above; they generally focus on the data and, thus, miss the opportunity to focus on reality. This problem is irrelevant in relation to human choices expressed through the internet since such human choices lack a spatial support, but they matter in the environmental and related sciences. How can spatial statisticians integrate conceptually rich understandings, as presented in this paper, into such models?

5.2 *Deep learning*

More recently, deep learning has produced a sharp rise in interest from the remote sensing community, although that interest is now plateauing. The story of deep learning in remote sensing is interesting and we *precis* it here, focusing on how it works, what it can do and what it cannot do. The cardinal example of deep learning applied to remote sensing is the convolutional neural network (CNN), with many examples in the literature (Masi et al., 2016; Song et al., 2018; Zhang et al., 2018). Thus, we focus on the CNN, acknowledging that it is but one of many deep learning approaches (Das and Ghosh, 2016; Shao and Cai, 2018; Zhang et al., 2018; Yeh et al., 2019). The CNN was designed for the task of identifying or classifying "higher-order representations" that are generally (but not uniquely) object-based. This statement requires some unpacking.

Prime among the concepts introduced above is that the task of the CNN is different than for standard classifiers. Standard classifiers in remote sensing are targeted on labelling low-order representations, primarily land cover; the first-order *state* of the land surface. In contrast, land use is a higher-order representation that relates to *function*. CNNs are able to predict such higher-order representations, whereas standard classifiers cannot. Through the research of the 1980s and 1990s, it is well known that while land cover can be inferred directly from remotely sensed reflectance on a per-pixel basis, land use cannot. In contrast, land use must be inferred through the relations between pixels or through the relations between objects defined on those pixels (or objects defined on a continuous space mapped to those pixels). It is for this reason that so-called texture classifiers of the 1980s and 1990s were applied successfully to classify land use. Texture classifiers first created texture "bands" by applying texture filters to the original broadband imagery and then discriminated between the classes of interest in the higher dimensional feature space created by the original-plus-texture wavebands. The CNN exploits a similar principle through convolution (texture) and pooling layers *within* a deep neural network such that the features to extract and utilize are determined automatically based on processing of the input.

This brings us to a second concept that defines CNNs in remote sensing. CNNs take as their input an image patch instead of an image pixel. Indeed, CNNs were designed originally for the identification of single representations within images (e.g. it is an agricultural field, it is a forest). Put simply, the CNN exploits second-order and higher-order relations (e.g., texture) in the input image patch to target the classification of higher-order representations such as (to give a simple case) land use. This focus on higher-order representations will be expanded on in section 6.

A third interesting concept related to CNNs is that the representations that are targeted are commonly (albeit not ubiquitously) readily expressed as spatial objects. For example, an agricultural field or a forest patch can be thought of as a functional object. Hence, the land use classes of interest, such as field and forest, can be conceptualized as functional objects in a scene.

The above conceptualization of what a CNN is can help us to determine what it can and cannot do. First, since CNNs are targeted on higher-order representations their utility is primarily in doing what standard classifiers cannot (e.g., identifying higher-order features such as 'it's a train station', 'it's a golf course'), and they generally are not required for classifying land cover, even if they can do that accurately. A recent example concerns the sensitivity of areas for bush fires to start off (Bergado et al., 2021). Second, the fact that the CNN takes a patch as input means that the prediction has a coarse spatial support, even if the result is allocated to a central pixel artificially. This is an unfortunate consequence when one is interested not in identifying something within an image, but labelling the *multiple* 'objects' that exist across an image (as is commonly the case in remote sensing). Third, the fact that CNNs generally target spatial *objects* is completely missed by the algorithm that is focused only on identification or labelling a feature and not at all on its geometry. These factors should guide application of CNNs in remote sensing, and also give clues as to where gaps exist for further development.

There are many other exciting ANN developments in remote sensing presently, including in relation to generative adversarial networks (GANs; Bermudez et al., 2019; Fuentes Reyes et al., 2019) and U-Nets which aim to resolve the above issues to some extent. U-nets have contributed widely in remote sensing research. Their use requires careful selection of the involved parameters. See Persello and Stein (2017) for a general presentation in the remote sensing domain. The applications are useful in image segmentation, where clear advances have been gained in building outline detection (Zhao et al., 2021) and in detection of informal settlements (Mboga et al., 2017). This also includes the rather technical polarimetric SAR data that require complex arithmetic (Mullissa et al. 2019). The purpose of this paper is not to review these methods, but rather to draw attention to the underlying concepts.

It is interesting to note that whereas much change of support research has focused on increasing the spatial resolution, deep learning methods such as the CNN decrease it implicitly by abstracting higher-order representations from patches. This gap should provide tangible *foci* for future methodological development.

5.3 *Explainable AI*

Explainable AI is currently a very hot topic in remote sensing. Whereas publications using deep learning have plateaued, explainable AI is on the rise. The goal of explainable AI is to render the inner mappings of AI approaches, predominantly machine and deep learning algorithms in the present remote sensing context, amenable to interpretation. There exist several different levels to this including access, intelligibility and so on. For recent reviews see Angelov et al. (2021) and Linardatos et al. (2021). An interesting example of the development of an XAI approach in remote sensing is Gu et al. (2020). In this approach, IF-THEN rules are encoded within the algorithm and the result of classification is presented to the user not only as a class allocation per image patch, but in terms of the IF-THEN rules that led to the allocation. This means that the user is easily able to understand why the decision was made and whether the decision makes sense. This closes the loop between prediction and user-based validation and allows the investigator the opportunity to understand, most crucially, how to improve the model. We expect to see much research attention being paid to XAI approaches in remote sensing over the next few years.

6. Semantic and Ontological Considerations

The concepts introduced through the sections of this paper lead us to a more refined understanding of the nature of remotely sensed data and, thereby, the appropriateness of spatial statistical methods for application to these data. Perhaps most importantly it can reveal gaps in the capability of some methods that point to the need for model development. It is for this reason that we were motivated to write this paper, because we feel that it may motivate other researchers, in particular, spatial statisticians, to restate problems in remote sensing, and rethink the spatial statistical solutions that are appropriate for them. In this section we develop this conceptual model further by considering the choice of method for particular goals, and by introducing semantic and ontological considerations (Wang et al., 2020).

6.1 *Choice of goal, method and spatial resolution*

Woodcock and Strahler (1987) first identified that the choice of spatial resolution in remote sensing is conditional not only on the goal of the analysis (which is fairly obvious), but also on the method and the frequency of spatial variation in the scene. In fact, as Woodcock and Strahler explain, more commonly, it is the *choice of method* and choice of spatial resolution, that is conditional on the goal and the interaction of the spatial resolution with the frequency of spatial variation. The same holds true today. Common goals in remote sensing for handling continuous spatial variation are the statistical prediction of continua (e.g., biomass per ha) based on regression-type models and the classification of land cover, both of which can be achieved operating on pixels directly. Invoking the object-based data model, a common goal is to segment (i.e., identify) and classify (label) objects in an image, again operating at the pixel level (on local connected groups of pixels). Whether these pixel-level goals are appropriate depends to a certain extent on the data, and more specifically the interaction of the spatial resolution with the frequency of spatial variation in the scene (in either continuous variation or implicit objects).

Generally, but especially where the pixel size is large relative to the scales of variation of interest, the goal can be restated to focus on the punctual support (or quasi-point support) scene of interest (i.e., ultimately reality) rather than the image itself. The image is, after all, generally not the researcher's interest. Indeed, the image is limited precisely by the sampling strategy decisions that were taken to achieve measurement, and it is in this sense an *extremely* partial representation of reality. Refocusing the goal outside of this limited sampled view of reality offers the possibility to escape the strictures of the pixel, and measurement and sampling processes. It offers the possibility to fit models defined on a punctual support that map to the data on a positive support, such that other mappings can be generated readily (e.g., to a coarser support, to a finer support or to a support with complex geometry such as a census Ward). This is important, because one of the most significant challenges in environmental data science today is the inability to allow datasets to speak to each other that were obtained with different supports (and other measurement characteristics) (Gotway and Young, 2002, 2007; Young and Gotway, 2007). We are thinking of data fusion as one example, but more generally we are thinking of interoperability as another. Such interoperability requires principled mappings that can transform data on one support to data on another support exchangeably. We are some way off from such a vision. Nevertheless, it is an important vision because environmental data come in a wide variety of shapes and sizes (property definitions, sampling frameworks, measurement characteristics, error characteristics), and we draw the attention of spatial statisticians to it.

6.2 *A hierarchical ontology for remote sensing*

This paper has highlighted that remotely sensed scenes are comprised of land covers (states) (e.g., Hansen and Loveland, 2012) and land uses (functional objects) (e.g., Chen et al., 2021). We now extend this thinking to develop a conceptual ontological model that is potentially of great use in considering the goals of spatial statistical modelling in remote sensing. We contend that land cover and land use are intimately linked in a coupled ontology, with land use sitting at a higher-order of representation above land cover (Zhang et al., 2019; Hong et al., 2019; Wang et al., 2020). This is fairly clear when one considers the classic case of the land use 'urban' which we know to be comprised of constituent land covers (grass, tarmac, concrete, roof tiles, water and so on) arranged in particular spatial patterns. In this sense, land use is 'built on' the constituent land cover states. Not only this, but we contend that whereas land cover exists as a pixel-based concept (it is meaningful to describe the land cover state in a pixel, e.g., grass), land use exists more meaningfully as an object-based concept (e.g., residential buildings, car park, roads). Note that the land use 'urban' is slightly different in that it is effectively higher-order than, say, its constituent buildings.

Zhang et al. (2019) realized the above and developed a statistical joint distribution modelling approach that capitalized on the ontological connectedness between land cover and land use as a higher-order representation. To predict land use it was necessary to use a CNN (see above), making this joint distribution model unique in that it the coupled a low-order classifier (multi-layer perceptron) with a higher-order classifier (CNN); referred to as 'joint deep learning' (JDL). Moreover, the joint distribution model was fitted between land cover defined at the pixel level and land use defined as *objects*. Prediction of one was used to inform the other. The consequence of joint modelling was that the accuracy of classification of *both* land cover and land use increased greatly through iterative fitting, exploiting the joint dependence.

In fact, as alluded to above in relation to the urban land use, a *hierarchical* ontology can be defined for land cover and land use, where functional higher-order representations sit *above* land use. Concepts such as ‘train station’ and ‘golf course’ are complex higher-order representations that lie at a higher level in a hierarchical ontology than land covers (lowest level) and land uses (the level above land cover). For example, in the case of a train station, the concept is predicated on both land covers (tarmac, roof tiles, gravel tracks) and land uses (long thin buildings, roads, car park, railway lines) arranged in specific identifiable patterns. Indeed, ‘urban’ is also a complex construct and sits above some more fundamental land covers and land uses, as suggested above. This conceptualization is important, not least because it helps to direct the application of techniques such as deep learning CNNs to appropriate goals, but also because it suggests new possibilities for the development of spatial statistics.

What is the value, at least in a research context, of yet another pixel-based, standard classification method once the semantically and ontologically rich, and sampling framework-free, conceptualization of reality presented here is considered?

7. Summary

The contribution of this paper is not to review the methods of recent years, but to review the conceptualizations and representations underlying these methods, and to offer some common themes. The *foci* of this paper are necessarily a small and biased (partial) sample of what may be important conceptually in relation to the development and application of spatial statistics in remote sensing. Nevertheless, it is hoped that readers of *Spatial Statistics* will find some inspiration for their future research ideas.

The common themes drawn out through this paper focus on conceptualizations and representations, and can be summarized as follows:

1. *Measurement, sampling and data*: remotely sensed images are a function of what is out there in reality and spatial (space-time) sampling processes. As a result of the harsh razor of the sampling framework, and particularly the spatial support, remotely sensed data represent an *extremely* limited window on the world. Methods that are applied directly to the data are hostage to the sampling framework, as though the image were reality. Spatial statistical methods that obviate the strictures of the support and pixel should be the focus of future research.
2. *Information and variation*: geostatistical RF characterizations can be helpful in revealing the scales of spatial variation in spatial data and, thus, the potential information content and, conversely, redundancy in data. They can provide some insights about the data (and the scene) to experienced interpreters, and these insights have a range of uses, including guiding the researcher as to whether a particular method is appropriate for a task. The recent trend towards prediction (through mixed regression models) in place of characterization and interpretation is

worthy of reflection as we feel that something has been lost. At the same time, it should be acknowledged that the geostatistical RF (GP) model is extremely limited as it is based on two-point statistics. In addition, we introduced the object-based data model as an alternative to spatial continua, noting that such representations are human constructs. More attention should be paid in remote sensing to object-based conceptualizations by those applying RF models, and to stochastic models of object geometry.

3. *Spatial statistical models*: We discussed major recent developments in spatial statistics applied to remote sensing as geostatistical change of support theory and multiple point geostatistics, mixed (spatial) regression models using the Bayesian inference paradigm and fuzzy spatial objects. These are just a few of the key developments, but they serve to illustrate a trend towards increasingly complex modelling taking advantage of computer power, and the development of our conceptual understanding of both principled statistical models and the landscape to which those models are applied.
4. *Machine and deep learning*: the key advantage underlying the recent success of deep learning is that it offers the possibility to predict something that was hardly achievable before. Unfortunately, not all applications of deep learning (primarily CNNs) in remote sensing have targeted higher-order representations. The message is: use the right model for the risk task and consider your goals carefully. We also showed that standard CNNs suffer the drawbacks of an induced patch-sized support and an inability to represent object boundaries directly (although alternative deep learning approaches do aim to tackle these).
5. *Ontologies and graphs*: We suggested a *hierarchical* ontology of land cover and land use, coupled with yet higher-order representations, for remote sensing. It seems to us that defining the appropriate conceptual ontology should come first, the selection of goals second and the selection of appropriate methods third. This is a gross over-simplification, of course, but it is salutary to ask how often do we actually consider the ontological landscape on which we are operating? Probably rarely or at least not often enough. We certainly argue that a firm conceptual understanding from measurement and sampling through data models to statistical model characterizations is key to useful inference and prediction. Indeed, it is through careful construction of an appropriate conceptualization that inference can be made meaningful.

It is interesting to reflect that the broad conceptual view of the landscape of remote sensing as introduced in this paper was developed over decades (see Atkinson and Tate, 2000). New spatial statistical methods and data science methods with novel capabilities were introduced over this time, and users of those methods were educated through their study and application, allowing them to enrich their own conceptual model and understanding of reality and remotely sensed data, as well as what a good spatial statistical model should be. This was certainly the case in our experiences.

As remote sensing is a mature subject, we believe that the broad conceptual model that we have presented here for remote sensing is fairly advanced. It is possible that other subjects would benefit from similar, explicitly spatial, conceptual constructions.

Acknowledgements

We thank the Editor and Referees for their useful comments which greatly improved this paper. PMA is funded by the EPSRC New Approaches to Data Science Programme grant EP/R01860X/1.

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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 7, 4887-4896.

Addink, E., Stein, A., 1999. A comparison of conventional and geostatistical methods to replace clouded pixels in NOAA-AVHRR images. *International Journal of Remote Sensing*. 20, 961-977.

Anderson, M.C., Kustas, W.P., Norman, J.M., 2003. Upscaling and downscaling - a regional view of the soil-plant-atmosphere continuum. *Agronomy Journal*. 95, 1408-1423.

Aplin, P., Atkinson, P.M., 2001. Sub-pixel land cover mapping for per-field classification. *International Journal of Remote Sensing*. 22, 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 Journal of Photogrammetry and Remote Sensing*. 66, 762-775.

Arvidson, T., Goward, S., Gasch, J., Williams, D., 2006. Landsat-7 long-term acquisition plan, development and validation. *Photogrammetric Engineering and Remote Sensing*. 72, 1137-1146.

Atkinson, P.M., 1997. Mapping sub-pixel boundaries from remotely sensed images. In: *Proceedings of GISRUK '96.* Canterbury, UK.

Atkinson, P.M., Webster, R., Curran, P.J., 1992. Cokriging with ground-based radiometry. *Remote Sensing of Environment*. 41, 45-60.

Atkinson, P.M., Webster, R., Curran, P.J., 1994. Cokriging with airborne MSS imagery. *Remote Sensing of Environment*. 50, 335-345.

Atkinson, P.M., Tatnall, A.R., 1997. Introduction: neural networks in remote sensing, *International Journal of Remote Sensing*. 18, 699-709.

Atkinson, P.M., Tate, N. J., 2000. Spatial scale problems and geostatistical solutions: a review. *Professional Geographer*. 52, 607-623.

Atkinson, P.M., 2004. Super-resolution land cover classification using the two-point histogram. In: *GeoENV IV-Geostatistics for Environmental Applications*. Springer, the Netherlands.

Atkinson, P.M., 2005. Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*. 71, 839-846.

Atkinson, P.M., 2013. Downscaling in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*. 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 Transactions on Geoscience and Remote Sensin*g. 46, 573-580.

Augustin, N.H., Mugglestone M.A., Buckland, S.T., 1994. An autologistic model for the spatial distribution of wildlife, *Journal of Applied Ecology*. 33, 339-347.

Bai, H.X, Ge, Y., Mariethoz, G., 2016. Utilizing spatial association analysis to determine the number of multiple grids for multiple-point statistics. *Spatial Statistics*. 17, 83–104.

Belgiu, M., Stein, A., 2019. Spatiotemporal image fusion in remote sensing. *Remote Sensing*. 11, 818.

Bergado, J.R., Persello, C., Reinke, K., Stein, A., 2021. Predicting Wildfire Burns from Big Geodata using Deep Learning. *Safety Science*. 140, 105276.

Bergin, E., Buytaert, W., Onof, C., Wheater, H., 2012. Downscaling of rainfall in Peru using generalised linear models. In: EGU General Assembly Conference (Vol.14). EGU General Assembly Conference Abstracts.

Bermudez, J.D., Happ, P.N., Feitosa, R.Q., Oliveira, D.A.B., 2019. Synthesis of multispectral optical images from SAR/optical multitemporal data using conditional generative adversarial networks. IEEE Geoscience and Remote Sensing Letters. 16, 1220-1224.

Berrocal, V.J., Craigmile, P.F., Guttorp, P., 2012. Regional climate model assessment using statistical upscaling and downscaling techniques. *Environmetrics*, 23, 482–492.

Breiman, L., 2001. Random forests. *Machine Learning*. 45, 5−32.

Chen, F., Zhao, X., Ye, H., 2012. Making use of the Landsat 7 SLC-off ETM+ image through different recovering approaches. *Data Acquisition Applications*.

Chen, J., Zhu, X., Vogelmann, J.E., Gao, F., Jin, S., 2011. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sensing of Environment*. 115, 1053-1064.

Chen, S., Wang, X., Guo, H., Xie, P., Sirelkhatim, A.M., 2020. Spatial and temporal adaptive gap-filling method producing daily cloud-free NDSI time series. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 13, 2251-2263.

Chen, Y., Cao, R., Chen, J., Liu, L., Matsushita, B., 2021. A practical approach to reconstruct high-quality Landsat NDVI time-series data by gap filling and the Savitzky–Golay filter. *ISPRS Journal of Photogrammetry and Remote Sensing*. 180, 174-190.

Chen, Y.H., Ge, Y., Heuvelink, G.B.M., Hu, J., Jiang, Y., 2015. Hybrid constraints of pure and mixed pixels for soft-then-hard super-resolution mapping with multiple shifted images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 8, 2040-2052.

Chen, Y.H., Ge, Y., Heuvelink, G. B. M., An, R., Chen, Y., 2018. Object-based super-resolution land cover mapping from remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 56, 328-340.

Chen, Z., Pu, H., Wang, B., Jiang, G.-M., 2014. Fusion of hyperspectral and multispectral images: a novel framework based on generalization of pan-sharpening methods. *IEEE Geoscience and Remote Sensing Letters*. 11, 1418-1422.

Cressie, N.A., 1993. *Statistics for Spatial Data*. John Wiley & Sons, New York.

Cressie, N.A., 1996. Change of support and the modifiable unit problem. *Geographical Systems.* 3, 159-180.

Cressie, N.A., Wikle, C.K., 2015. *Statistics for Spatio-Temporal Data*. John Wiley & Sons, New York.

Curran, P.J., 1985. *Principles of Remote Sensing*, Taylor and Francis, London.

Dadrass Javan, F., Samadzadegan, F., Mehravar, S., Toosi, A., Khatami, R., Stein, A., 2021. A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*. 171, 101-117.

Das, M., Ghosh, S.K., 2016. Deep-STEP: a deep Learning approach for spatiotemporal prediction of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*. 13, 1984-1988.

Diggle, P.J., Tawn, J.A., Moyeed, R.A., 1998. Model-based geostatistics, *Applied Statistics*. 47, 299-350.

Duffaut, L.E., Posadas, A.N., Carbajal, M., Quiroz, R., 2017. Multifractal downscaling of rainfall using normalized difference vegetation index (NDVI) in the Andes plateau. *PLOS ONE*. 12, e0168982.

Fisher, P.A., Wood, J., Cheng, T. 2004. Where Is Helvellyn? Fuzziness of multi-scale landscape morphometry. *Transactions of the Institute of British Geographers*, *New Series*. 29, 106-128

Fotheringham, A.S., Wong, D.W.S., 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment & Planning A.* 23, 1025-1044.

Forman, R.T.T., 1995. *Land Mosaics: the Ecology of Landscapes and Regions*. Cambridge University Press; Cambridge.

Fuentes Reyes, M., Auer, S., Merkle, N., Henry, C., Schmitt, M., 2019. SAR-to-optical image translation based on conditional generative adversarial networks-optimization, opportunities and limits. *Remote Sensing*. 11, 2067.

Ge, Y., 2013. Sub-pixel land-cover mapping with improved fraction images upon multiple-point simulation. *International Journal of Applied Earth Observation and Geoinformation*. 22, 115-126.

Ge, Y., Bai, H.X., 2011. Multiple-point simulation-based method for extraction of objects with spatial structure from remotely sensed imagery, *International Journal of Remote Sensing*. 32, 2311- 2335.

Ge, Y., Liang, Y., Wang, J., Zhao, Q., Liu, S., 2015. Upscaling Sensible Heat Fluxes with Area-to-Area Regression Kriging. *IEEE Geoscience and Remote Sensing Letters*. 12, 656-660

Ge, Y., Chen, Y.H., Stein, A., Li, S., Hu, J., 2016. Enhanced sub-pixel mapping with spatial distribution patterns of geographical objects. *IEEE Transactions on Geoscience and Remote Sensing*,. 54, 2356-2370.

Ge, Y., Chen, Y.H., Li, S., Jiang, Y., 2014. Vectoral boundary-based sub-pixel mapping method for remote-sensing imagery. *International Journal of Remote Sensing*. 35, 1756-1768.

Ge, Y., Wang, J., Wang, J., Jin, Y., Liu, M., Wang, J., Chen, Y., Cheng, Q., Stein, A., Atkinson, P.M., 2019. Principles and methods of scaling geospatial Earth science data. *Earth Science Reviews*. 197, 102897.

Goodchild, M.F., 2004. The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers*. 94, 300-303.

Goodchild, M.F., 2011. Scale in GIS: an overview. *Geomorphology.* 130, 5-9.

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. *International Journal of Health Geographics*. 5, 52.

Goovaerts, P., 2007. Kriging and semi-variogram deconvolution in the presence of irregular geographical units. *Mathematical Geosciences*. 40, 101-128.

Gotway, C.A., Young, L.J., 2002. Combining incompatible spatial data. *Publications of the American Statistical Association*. 97, 632-648.

Gotway, C.A., Young, L.J., 2007. A geostatistical approach to linking geographically aggregated data from different sources. *Journal of Computational and Graphical Statistics*. 16, 115-135.

Gu, X., P.P. Angelov, C. Zhang, P.M. Atkinson, in press, A semi-supervised deep rule-based approach for complex satellite sensor image analysis. *IEEE Pattern Analysis and Machine Intelligence*.

Guardiano, F., Srivastava, R.,, 1993. Multivariate geostatistics: beyond bivariate moments. In: Soares, A. (Ed.), *Geostatistics-Troia.* Kluwer Academic, Dordrecht, pp. 133–144.

Guo, Y., Wang, C., Lei, S., Yang, J., Zhao, Y., 2020. A framework of spatio-temporal fusion algorithm selection for Landsat NDVI time series construction. *ISPRS International Journal of Geo-Information*. 9, 665.

Hansen, M., Loveland, T., 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*. 122, 66–74.

He, W., Yokoya, N., 2018. Multi-temporal Sentinel-1 and -2 data fusion for optical image simulation. *ISPRS International Journal of Geo-Information*. 7, 10.

Hong, D., Yokoya, N., Ge, N., Chanussot, J., Zhu, X.X., 2019. Learnable manifold alignment (LeMA): A semi-supervised cross-modality learning framework for land cover and land use classification. *ISPRS Journal of Photogrammetry and Remote Sensing*. 147, 193-205.

Hu, J., Ge, Y., Chen, Y.H., Li, D., 2015. Super-resolution land cover mapping based on multiscale spatial regularization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 8, 2031-2039.

Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*. 178, 127-141.

Jeganathan, C., Mondal, S., 2017. Fractal-based pattern extraction from time-series NDVI data for feature identification. *IEEE Journal of Selected topics in Applied Earth Observations and Remote Sensing*. 10, 5258-5264.

Jiang, B., 2018. Spatial heterogeneity, scale, data character, and sustainable transport in the big data era. *ISPRS International Journal of Geo-Information*. 7, 167.

Jiang, B., Brandt, S.A., 2016. A fractal perspective on scale in geography. *ISPRS International Journal of Geo-Information*. 5, 95.

Jin, Y., Ge, Y., Wang, J.H., Chen, Y.H., et al., 2018b. Downscaling AMSR-2 soil moisture data with geographically weighted area-to-area regression kriging. *IEEE Transactions on Geoscience and Remote Sensing*. 56, 2362-2376.

Journel, A.G., 1993. Geostatistics: roadblocks and challenges. *Mathematical Geosciences.* 92, 213–224.

Journel, A.G., and Huijbregts, Ch.J., 1978. *Mining Geostatistics*. Academic Press, London and New York.

Kasetkasem, T., Arora, M.K., Varshney, P.K., 2005. Super-resolution land cover mapping using a Markov random field-based approach. *Remote Sensing of Environment*. 96, 302-314.

Kohli. D., Stein, A., Sliuzas, R. 2016. Uncertainty analysis for image interpretations of urban slums. *Computers, Environment and Urban Systems*. 60, 37-49.

Kyriakidis, P.C., 2004. A geostatistical framework for area-to-point spatial interpolation. *Geographical Analysis*. 36, 259–289.

Kyriakidis, P.C., Journel, A.G., 1999. Geostatistical space–time models: a review. *Mathematical Geology*. 31, 651-684.

Li, A., Bo, Y.C., Zhu, Y., Guo, P. et al., 2013. Blending multi-resolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method. *Remote Sensing of Environment*. 135, 52-63.

Li, X., Ling, F., Du, Y., Zhang, Y., 2014. Spatially adaptive super-resolution land cover mapping with multispectral and panchromatic images. *IEEE Transactions on Geoscience and Remote Sensing*. 52, 2810-2823.

Li, J., Li, Y., Cai, R., He, L., Chen, J., Plaza, A., 2021. Enhanced spatiotemporal fusion via MODIS-like images. *IEEE Transactions on Geoscience and Remote Sensing*. 1-17.

Linardatos, P., Papastefanopoulos, V., Kotsiantis, S., 2021. Explainable AI: a review of machine learning interpretability methods. *Entropy*. 23, 18.

Ling, F., Li, X., Du, Y., Xiao, F., 2013. Sub-pixel mapping of remotely sensed imagery with hybrid intra- and inter-pixel dependence. *International Journal of Remote Sensing*. 34, 341-357.

Liu, Y., 2006. Using the SNESIM program for multiple-point statistical simulation. *Computers and Geosciences.* 32, 1544–1563.

Liu, X.H., Kyriakidis, P.C., Goodchild, M. F., 2008. Population-density estimation using regression and area-to-point residual kriging. *International Journal of Geographical Information Science*. 22, 431–447.

Lloyd, C.D., 2014. *Exploring Spatial Scale in Geography*. John Wiley & Sons, London.

Mandelbrot, B., 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*. 156, 636.

Masi, G., Cozzolino, D., Verdoliva, L., Scarpa, G., 2016. Pansharpening by convolutional neural networks. *Remote Sensing*. 8, 594.

Mariethoz, G., Renard, P., Straubhaar, J., 2010. The direct sampling method to perform multiple-point geostatistical simulations. *Water Resources Research*. 46, W11536.

Mboga, N., Persello, C., Bergado, J.R., Stein, A., 2017. Detection of informal settlements from VHR images using convolutional neural networks. *Remote Sensing*. 9, 1106.

Miller, H.J., 2004 Tobler's first law and spatial analysis. *Annals of the Association of American Geographers*. 94, 284-289.

Mondal, S., Jeganathan, C., Giriraj, A., Peejush P., 2017. Cloud noise mapping and reduction for enhancing the time series MODIS NDVI for improved vegetation monitoring. *GIScience and Remote Sensing*. 54, 202-229.

Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika*. 37, 17-23.

Mullissa, A.G., Persello, C., Stein, A., 2019. PolSARNet: A deep fully convolutional network for polarimetric SAR image classification. *IEEE Journal of Selected topics in Applied Earth Observations and Remote Sensing*, DOI: 10.1109/JSTARS.2019.2956650.

Nguyen, Q.M., Atkinson, P.M., Lewis, H.G., 2011. Super-resolution mapping using Hopfield Neural Network with panchromatic imagery. *International Journal of Remote Sensing*. 32, 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. *International Journal of Remote Sensing*. 32, 7203-7226.

Openshaw, S., 1984. The modifiable areal unit problem, *Concepts and Techniques in Modern Geography (Number 38)*. Geo Books, Norwich, England.

Pardo-Igúzquiza, E., Chica-Olmo, M., Atkinson, P. M., 2006. Downscaling cokriging for image sharpening. *Remote Sensing of Environment*. 102, 86-98.

Peng, K., Wang, Q., Tang, Y., Tong, X., Atkinson, P.M., in press. Geographically weighted spatial unmixing for spatio-temporal fusion. *IEEE Transactions on Geoscience and Remote Sensing*.

Persello, C., Stein, A., 2017. Deep fully convolutional networks for the detection of informal settlements in VHR images. *IEEE Geoscience and Remote Sensing Letters*. 14, 2325-2329.

Quattrochi, D.A., Goodchild, M.F., 1997. *Scale in Remote Sensing and GIS*. Lewis Publishers, New York.

Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M., Jeganathan, C., 2012. Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment*. 121, 93-107.

Rossi, R., Mulla, D., Journel, A., Franz, E., 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*. 62, 277-341.

Rulinda, C., Bijker, W., Stein, A., 2011. The chlorophyll variability in Meteosat derived NDVI in a context of drought monitoring. *Procedia Environmental Sciences*. 3, 32-37.

Sales, M.H.R., Souza, C.M., Kyriakidis, P.C., 2013. Fusion of MODIS images using kriging with external drift. *IEEE Transactions on Geoscience and Remote Sensing*. 51, 2250-2259.

Sidiropoulou Velidou, D., Tolpekin, V., Stein, A., Woldai, T. 2015. Use of Gestalt theory and random sets for automatic detection of geological features. *Mathematical Geosciences*. 47, 249-276

Shao, Z., Cai, J., 2018. Remote sensing image fusion with deep convolutional neural network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 11, 1656-1669.

Song, H., Huang, B., 2012. Spatiotemporal reflectance fusion via sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*. 50, 3707-3716.

Song, H., Liu, Q., Wang, G., Hang, R., Huang, B., 2018. Spatiotemporal satellite image fusion using deep convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 11, 1-9.

Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R., Besson, O., 2011. An improved parallel multiple-point algorithm using a list approach. *Mathematical Geosciences.* 43, 305–328.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*. 34, 1-21.

Stein, A., Van der Meer, F., Gorte, B.G.H., 1999. *Spatial Statistics and Remote Sensing*. Kluwer, Dordrecht.

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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 5, 1403-1417.

Tahmasebi, P., Hezarkhani, A., Sahimi, M., 2012. Multiple-point geostatistical modeling based on the cross-correlation functions. *Computers and Geosciences.* 16, 779–797.

Tang, Y., Atkinson, P.M., Zhang, J., 2015. Downscaling remotely sensed imagery using area-to-point cokriging and multiple-point geostatistical simulation. *ISPRS Journal of Photogrammetry and Remote Sensing*. 101, 174-185.

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. *International Journal of Applied Earth Observation and Geoinformation*. 3, 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 Sensing of Environment*. 79, 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. *Computers and Geosciences*. 33, 1261-1272.

Tolpekin, V., Stein, A., 2009. Scale effects and parameter optimization for super resolution mapping from satellite images with Markov Random Fields. *IEEE Transactions in Geoscience and Remote Sensing*. 47, 3283-3297

Wang, C., Zhuo, X., Li, P., Chen, N., Wang, W., Chen, Z., 2020. An ontology-based framework for integrating remote sensing imagery, image products and in-situ observations. *Journal of Sensors*, ID6912820.

Wang, J., Stein, A., Gao, B., Yong, G., 2012. A review of spatial sampling. *Spatial Statistics.* 2, 1-14.

Wang, J., Gao, B., Stein, A., 2020. The spatial statistic trinity: A generic framework for spatial sampling and inference. *Environmental Modelling & Software*. 134, 104835.

Wang, J.F., Wang, Y., Zeng, H., 2016. A geostatistical approach to the change-of-support problem and variable-support data fusion in spatial analysis. *Journal of Geographical Systems*. 18, 45-66.

Wang, Q., Shi, W., Atkinson, P.M., Zhao, Y., 2015c. Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*. 166, 191-204.

Wang, Q., Shi, W., Atkinson, P.M., 2014. Sub-pixel mapping of remote sensing images based on radial basis function interpolation. *ISPRS Journal of Photogrammetry and Remote Sensing*. 92, 1-15.

Wang, Q., Shi, W., Atkinson, P.M., Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*. 166, 191-204.

Wang, Q., Shi, W., Li, Z., Atkinson, P.M., 2016. Fusion of Sentinel-2 images. *Remote Sensing of Environment*. 187, 241-252.

Wang, Q., Atkinson, P.M., 2018. Spatio-temporal fusion for daily Sentinel-2 images. *Remote Sensing of Environment*. 204, 31-42.

Wang, Q., Shi, W., Atkinson, P.M., 2019. Information loss-guided multi-resolution image fusion. *IEEE Transactions on Geoscience and Remote Sensing*. 58, 45-57.

Wang, Q., Tang, Y., Atkinson, P.M., 2020. The effect of the point spread function on downscaling continua. *ISPRS Journal of Photogrammetry and Remote Sensing*. 168, 251-267.

Wang, Q., Wang, L., Wei, C., Jin, Y., Li, Z., Tong, X., Atkinson, P.M., 2021. Filling gaps in Landsat ETM+ SLC-off images with Sentinel-2 MSI images. *International Journal of Applied Earth Observation and Geoinformation*. 101, 102365.

Williams, D., Goward, S., Arvidson, T., 2006. Landsat: yesterday, today, and tomorrow. *Photogrammetric Engineering and Remote Sensing*. 72, 1171−1178.

Woodcock C. E., Strahler, A. H., 1987. The factor of scale in remote sensing. *Remote Sensing of Environment*. 21, 311–332.

Wu, H., Li, Z. L., 2009. Scale issues in remote sensing: a review on analysis, processing and modeling. *Sensors*. 9, 1768.

Wulder, M.A., Masek, J.G., Cohen, W.B., 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*. 122, 2−10.

Wulder, M.A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2008. Landsat continuity: issues and opportunities for land cover monitoring. *Remote Sensing of Environment*. 112, 955−969.

Xu, X., Zhong, Y., Zhang, L., 2014. Adaptive subpixel mapping based on a multiagent system for remote-sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 52, 787-804.

Yan, L., Roy, D. P., 2016. Conterminous United States crop field size quantification from multi-temporal Landsat data. *Remote Sensing of Environment*. 172, 67–86.

Yang, F., White, M.A., Michaelis, A.R., Ichii, K., Hashimoto, H., Votava, P., Nemani, R.R., 2006. Prediction of continental-scale evapotranspiration by combining MODIS and AmeriFlux data through support vector machine. *IEEE Transactions on Geoscience and Remote Sensing*. 44, 3452–3461.

Yeh, C.H., Huang, C.H., Kang, L.W., 2019. Multi-scale deep residual learning-based single image haze removal via image decomposition. *IEEE Transactions on Image Processing*. 29, 3153-3167.

Yoo, E.H., Kyriakidis, P.C., 2006. Area-to-point kriging with inequality-type data. *Journal of Geographical Systems*. 8, 357-390.

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. *Geographical Analysis*. 42, 78-98.

Young, L.J., Gotway, C.A., 2007. Linking spatial data from different sources: the effects of change of support. *Stochastic Environmental Research and Risk Assessment*. 21, 589-600.

Zhang, J.X., Atkinson, P.M., Goodchild, M. F., 2014a. *Scale in Spatial Information and Analysis*. Taylor & Francis, London.

Zhang, Y., Du, Y., Ling, F., Fang, S., Li, X., 2014b. Example-based super-resolution land cover mapping using support vector regression. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 7, 1271-1283.

Zhang, Q., Yuan, Q., Zeng, C., Li, X., Wei, Y., 2018. Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*. 56, 4274-4288.

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P.M., 2019, Joint deep learning for land cover and land use classification, *Remote Sensing of Environment*. 221, 173-187.

Zhao, W., Persello, C., Stein, A., 2021. Building outline delineation: from aerial images to polygons with an improved end-to-end learning framework. *ISPRS Journal of Photogrammetry and Remote Sensing. ****

Zhao, X., Stein, A., Chen, X., 2010. Application of random sets to model uncertainties of natural entities extracted from remote sensing images. *Stochastic Environmental Research and Risk Assessment*. 24, 713–723.

Zhou L., Stein A., 2013. Application of random sets to model uncertainty of road polygons extracted from airborne laser points. *Computers, Environment and Urban Systems*. 41, 289-298

Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*. 144, 152-171.

Appendix 1

Table A1. Classification of references cited in this paper, provided to support further search and review.

Model Type	Data Model	Application Goal	Reference
Geostatistics	Random Field (RF) stochastic model representing spatial continua applied to images	Spatial prediction in mining	Cressie, 1993; Journel, 1993; Cressie and Wikle, 2015.
		Spatial prediction in remote sensing	Atkinson et al., 1992, 1994; Addink and Stein, 1999.
		Gap filling due to cloud and cloud shadow in remote sensing	Chen et al., 2014, 2020, 2021.
		Error removal due to sensor failure in remote sensing	Chen et al., 2011; Chen et al., 2012; Wang et al., 2021.
		Time-series image construction in remote sensing	Song and Huang, 2012; Mondal et al., 2017; Wang and Atkinson, 2018; Belgiu and Stein, 2019; Guo et al., 2020.
Geostatistical change of support and downscaling continua	Random Field (RF) stochastic model representing spatial continua applied to images	Increase in spatial resolution in remote sensing above that of the input image	Cressie, 1996; Kyriakidis, 2004; Pardo-Igúzquiza et al., 2006; Goovaerts, 2006, 2007; Atkinson et al, 2008; Atkinson, 2013; Huhtengs and Vohland, 2016; Wang et al., 2015, 2016; Yoo and Kyriakidis, 2006; Liu et al., 2008; Yoo et al., 2010; Jin et al., 2018.
Sub-pixel mapping	Various solutions, but generally admits a solution space not limited to the discretized image	Classification of land cover in in remote sensing at a finer spatial resolution than the input image	Atkinson, 1997; Tatem et al., 2001, 2002; Atkinson, 2005; Khasetkasem et al., 2005; Thornton et al., 2007; Tolpekin and Stein, 2009; Ardila et al., 2011; Nguyen et al., 2011; Su et al., 2012; Ling et al., 2013; Ai et al., 2014; Wang et al., 2014; Hu et al., 2015; Ge et al., 2016; Chen et al., 2018.
Multiple Point Geostatistics	Data-based approach applied directly to the	Image pattern recreation and simulation	Guardiano and Srivastava, 1993; Strebelle, 2002; Liu, 2006; Mariethoz et al., 2010; Ge and Bai, 2011; Straubhaar et al., 2011;

	discretized space of the image		Tahmessabi et al., 2012; Bai et al., 2013; Ge et al., 2013; Tang et al., 2015.
Model-Based Geostatistics using Bayesian Inference	Random Field (RF) stochastic model representing spatial continua applied to images	Prediction of continua based on covariates. Models the uncertainty in RF estimation	Augustin et al. 1996; Diggle et al. 1998.
Random Sets	Object-based stochastic model commonly applied directly to images	Segmentation and classification of images into objects, including object boundary delineation and uncertainty therein	Zhao et al. 2010; Zhou et al. 2013; Sidiropoulou Velidou et al. 2015; Kohli et al. 2016.
Non-linear learning (i.e., **Machine Learning** methods, including ANNs, SVMs, Random Forest, GANs, etc.) for classification	Data-based approach applied directly to the discretized space of the image	Classification of remote sensing images, commonly to land cover	Atkinson and Tatnall, 1997; Yang et al., 2006; Zhang et al., 2014b; Brieman, 2001; Rodriguez-Galliano et al., 2012; Huhtengs and Vohland, 2016; Bermudez et al., 2019; Fuentes Reyes et al., 2019; Persello and Stein, 2017; Zhao et al., 2021; Mboga et al., 2017; Mullissa et al. 2019.
Higher-order, non-linear learning (i.e, **Deep Learning** methods, including CNNs, U-Nets) for classification	Data-based approach applied directly to the discretized space of the image (although see U-Nets)	Higher-order (spatial and functional) classification of remote sensing images, commonly to land use	Masi et al., 2016; Song et al., 2018; Zhang et al., 2018; Das and Ghosh, 2016; Shao and Cai, 2018; Zhang et al., 2018; Yeh et al., 2019; Bergado et al., 2021.
Interpretable, non-linear learning (i.e., **XAI** - Explainable Artificial Intelligence) for classification	Data-based approach applied directly to the discretized space of the image	Identification of the decisions underlying the classification of images	Gu et al. 2020; Angelov et al., 2021; Linardatos et al., 2021.
Ontological Models	Commonly defined for spatial objects on a continuous space	Conceptualization of remote sensing scenes such as to aid the design of stochastic models	Hong et al., 2019; Wang et al., 2020.