ACCEPTED MANUSCRIPT

Integrating deterministic lithostratigraphic models in stochastic realizations of subsurface heterogeneity. Impact on predictions of lithology, hydraulic heads and groundwater fluxes.

Marco Bianchi[1], Timothy Kearsey[2], Andrew Kingdon[1]

[1] British Geological Survey, Keyworth, Nottingham, UK.

[2] British Geological Survey, Edinburgh, UK.

Corresponding Author:

M. Bianchi, British Geological Survey, Environmental Science Centre, Keyworth, Nottingham, NG12 5GG, United Kingdom. Email: marcob@bgs.ac.uk.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract. Realistic representations of geological complexity are important to address several engineering and environmental challenges. The spatial distribution of properties controlling physical and geochemical processes can be effectively described by the geological structure of the subsurface. In this work, we present an approach to account for geological structure in geostatistical simulations of categorical variables. The approach is based on the extraction of information from a deterministic conceptualization of the subsurface, which is then used in the geostatistical analysis for the development of models of spatial correlation and as soft conditioning data. The approach was tested to simulate the distribution of four lithofacies in highly heterolithic Quaternary deposits. A transition probability-based stochastic model was implemented using hard borehole data and soft data extracted from a 3-D deterministic lithostratigraphic model. Simulated lithofacies distributions were also used as input in a flow model for numerical simulation of hydraulic head and groundwater flux. The outputs from these models were compared to corresponding values from models based exclusively on borehole data. Results show that soft lithostratigraphic information increases the accuracy and reduces the uncertainty of these predictions. The representation of the geological structure also allows a more precise definition of the spatial distribution of prediction uncertainty, here quantified with a metric based on Shannon information entropy. Correlations between prediction uncertainties for lithofacies, hydraulic heads and groundwater fluxes were also investigated. The results from this analysis provide useful insights about the incorporation of soft geological data into stochastic realizations of subsurface heterogeneity, and emphasize the critical importance of this type of information for reducing the uncertainty of simulations considering flux-dependent processes.

ACCEPTED MANUSCRIPT

Keywords: Subsurface heterogeneity, Geological modelling, Geostatistics, Uncertainty, Information entropy.

1 Introduction and background

Realistic representations of geological complexity are essential to address environmental challenges concerning utilisation, conservation, and management of natural resources, as well as long term geological disposal of CO_2 and radioactive waste. Architectural and textural characteristics of the facies across different scales control the distribution of physical (e.g., hydraulic conductivity, porosity, grain size, bulk density) and geochemical (e.g., mineralogical composition, effective diffusion coefficient, retardation factor, surface area) properties. These in turn influence fluid flow and solute transport processes. For example, in alluvial aquifers misrepresenting or ignoring connectivity of highly permeable sediments can lead to flawed interpretations of groundwater circulation and contaminant migration (Fogg, 1986; Webb and Anderson, 1996; Schiebe and Yabusaky, 1998; Labolle and Fogg, 2001; Proce et al., 2004; Bianchi et al., 2011; Bianchi and Zheng, 2015).

The distribution of geological units or facies can be modelled with numerous approaches (see Kolterman and Gorelick, 1996 and de Marsily et al., 2005 for comprehensive reviews), which generally fall into two categories: deterministic or stochastic approaches. Deterministic approaches combine direct and/or indirect geological observations with expert insight and interpretation to produce unique models of geological heterogeneity. These have the advantage of being consistent with known geological relationships (stratigraphic, chronological, lithological etc.) and/or with established conceptualizations of the geological system of interest (e.g., a certain depositional system). Because of the development of computer based 3-D geological modelling tools (Mallet,

ACCEPTED MANUSCRIPT

2002; Turner, 2006; Turner and Gable, 2007; Kessler et al., 2009; Zanchi et al., 2009; Royse, 2010), deterministic models have become increasingly popular not only in the oil and mining industry (e.g, Xu and Dowd, 2003; Perrin et al., 2005), but also to support water resources assessment and management (Artimo et al., 2003; Ross et al., 2005; Campbell et al., 2010; Nury et al., 2010; Gill et al., 2011; Raiber et al., 2012; Carreño et al., 2014), and numerical simulation of hydrogeological processes (D'Agnese et al., 1999; Robins et al., 2005; Wycisk et al., 2009; Bonomi, 2009; Blessent et al., 2009; Giambastiani et al., 2012; Turner et al., 2014). In flow and solute transport modelling, in particular, deterministic 3-D geological models are considered valuable in the preparatory stage of the numerical modelling implementation to establish a sound geological framework for the development of the conceptual model (Robins et al., 2005; Bredehoft, 2005). In later stages, hydraulic properties are assigned to convert geological units, which are most commonly identified on the basis of lithostratigraphic criteria, into hydrostratigraphic units. It is common, however, that initially assigned values will need to be calibrated to achieve a match between simulated and observed values of the state variables (i.e., hydraulic heads, discharge rates, fluxes, concentrations, etc.). A potential drawback of this process is a reduction in the ability of the numerical model to extrapolate beyond the data used for calibration. Another potential limitation for the use of deterministic 3-D geological models in flow and transport modelling applications is the fact that the parameterisation of hydrogeological properties in lithostratigraphic units for various reasons is not a trivial task (e.g., Watson et al., 2015). For instance, hydrostratigraphic and lithostratigraphic boundaries do not necessarily match. Difficulties arise when the deterministic conceptual model is based on direct or indirect geological observations that have different resolution, extent and/or support scale compared to hydrogeological properties. Parameter upscaling (e.g., Blöschl and Sivaplan, 1995; Wen

ACCEPTED MANUSCRIPT

and Gómez-Hernandez, 1996; Renard and de Marsily, 1997; Sanchez-Vila et al., 2006) then becomes a necessary and sometimes problematic step to transfer information between different scales.

Stochastic approaches for modelling geological heterogeneity, which generate models consisting of multiple, equally probable realizations of the subsurface, provide an alternative to deterministic models. For hydrogeological modelling, one obvious advantage of stochastic models is that they provide information about geological structure uncertainty which, together with parameter uncertainty, is considered a major source of uncertainty for flow and transport simulations (Neuman, 2003; Poeter and Anderson, 2005; Refsgaard et al., 2012; Gupta et al., 2012; Tsai and Elshall, 2013; He et al., 2014a; Chitsan et al., 2014). Here, as in several previous studies (e.g., Kolterman and Gorelick, 1996; Webb and Anderson, 1996; Ramanathan et al., 2010; Ronayne et al., 2010; Huang et al., 2012; Bianchi et al., 2015), the term "geological structure" refers to a spatial organization of geological properties that is consistent with arrangements or patterns created by geological processes. Although multiple geological models can be generated manually (e.g., Troldborg et al., 2010; Seifert et al., 2012; Courrioux et al., 2015; Lark et al., 2015), our focus in this study is on automated generation with geostatistical methods.

Traditional geostatistical methods for categorical variables (e.g., Journel, 1983) can produce realizations of facies assemblages honouring observations and a covariance model representing the spatial structure of the data. However, the inability of these methods to reproduce complex, curvilinear and interconnected structures, which are common especially in deltaic, fluvial and fluvio-glacial depositional systems, has motivated the development of alternative techniques such as multiple-point statistics (MPS) (Guardiano and Srivastava, 1993; Strebelle, 2002; Hu and Chugunova, 2008) and the transition probability approach (Carle and Fogg, 1996, 1997). Compared to traditional variogram-

ACCEPTED MANUSCRIPT

based methods, these methods also allow the inclusion of subjective geological interpretations in the geostatistical analysis. For MPS, this is achieved by selecting an appropriate training image as the representative conceptualization of geological heterogeneity for the system of interest. However, this selection is a critical and often challenging step especially for three-dimensional analyses (e.g., Le Coz et al., 2011; Comunian et al., 2012; Huysmans and Dassargues, 2012; Dell'Arciprete et al., 2012; Bluin et al., 2013; He et al., 2014b). With the transition probability approach, the spatial structure of geological data is modelled by mathematical functions (i.e., Markov chain models) relating transition probabilities to distance. Although this is a two-point geostatistical method, geological knowledge can be taken into account in the determination of the coefficients of these functions, which can be directly related to interpretable geological properties including proportions of each facies, mean lengths, connectivity, and juxtapositional tendencies (e.g., Carle et al., 1998; Weissmann and Fogg, 1999; Ritzi, 2000; Lee et al., 2007; He et al., 2014a; Bianchi et al., 2015).

Geostatistical models of subsurface heterogeneity are typically developed on the basis of "hard" data consisting of direct observations of lithology or other measureable properties. Since these data are usually collected in boreholes, the information required to characterize the geological system of interest is often adequate in the vertical direction, but not horizontally. The use of additional "soft" data consisting of indirect observations of geological properties, as well as qualitative and interpretative information (e.g., geophysical surveys or conceptualizations of the depositional system) has been shown to be effective to overcome limitations due to the lack of direct observations (Elfeki et al.,1995; Copty and Rubin, 1995; Hyndman and Gorelick, 1996; Liu et al, 2004; Elfeki, 2006; Emery and Robles, 2009; Ye and Khaleel, 2008; Engdal et al., 2010; He et al., 2014a). He et al. (2014a), for example, developed a stochastic model of the distribution of

ACCEPTED MANUSCRIPT

sand and clay in glacial deposits based on both soft geophysical data from airborne electromagnetic surveys and hard borehole observations. Validation analysis conducted on a subset of the borehole data showed that soft conditioning (i.e., conditioning to the soft data) significantly improved the accuracy of lithology predictions for the sand units, for which there was a scarcity of direct observations.

The majority of geostatistical methods, including the transition probability approach, assume that the probability distribution of the random variable is stationary, i.e., invariant under any translation in space. This assumption may not hold for geological properties due to the presence of trends, highly connected features, variations in the depositional conditions within the stratigraphic sequence, and/or discontinuities such as unconformities and faults. Non-stationarity can be effectively addressed with the incorporation of soft information into the geostatistical analysis. Weissmann and Fogg (1999), for example, present a stochastic model of facies distribution where sequence stratigraphic boundaries (i.e., soft interpretative data) are used to avoid unrealistic cross correlation across major unconformities in an alluvial fan. A similar approach has also been used in a point bar/channel system to simulate facies distributions into two distinct depositional units (Dell'Arciprete et al., 2002), and also in the recent study by Kearsey et al. (2015) to simulate lithological variations in glacial and post-glacial Quaternary deposits. In these previous studies, however, soft information is not directly incorporated in the geostatistical analysis, but it provides a conceptualization of the geological system for the identification of areas where stationarity may be assumed, thus allowing the appropriate application of geostatistical methods.

In this study we present an approach for full integration of deterministic geological models into geostatistical simulations of subsurface heterogeneity. The approach is applied to simulate the distribution of lithofacies in the complex glacial and post-glacial

ACCEPTED MANUSCRIPT

environment underlying the city of Glasgow (UK). One objective of the present study is to test the hypotheses that the geological structure inherited from a deterministic lithostratigraphic model can effectively improve the accuracy of the predictions of lithofacies distribution. Other important objectives of this work are: (1) to quantify variations in prediction uncertainty when the additional soft geological information is used in the stochastic simulations of lithofacies distributions; (2) to quantify correspondent variations in prediction uncertainty for hydraulic head and groundwater fluxes; (3) to analyse correlations between these variations. Although this analysis is focused on a specific site, it provides useful insights to understand the impact of geological uncertainty on groundwater flow modelling.

2 Materials and methods

2.1 Methodology outline

The methodology adopted in this study can be summarized as follows:

1) Two 3-D stochastic models of the distribution of four lithofacies in an area of 100 km^2 covering central Glasgow are generated. Both models are developed with the transition probability approach, but differ with respect to data used for transition probabilities estimations, fitted Markov chain models, and conditioning data. One model is solely based on hard lithological observations from boreholes, while the second model was developed using both hard data and soft data. Soft information was extracted from a 3-D lithostratigraphic model of the superficial deposits in the study area.

2) A split sample validation test is performed to assess the accuracy of both models.

ACCEPTED MANUSCRIPT

3) Prediction uncertainties of the two stochastic models are quantified with a normalized metric based on Shannon information entropy (Shannon, 1948) and then compared.

4) Stochastic realizations of lithofacies assemblages derived from the two models are used as inputs for a groundwater flow numerical model. The predicted uncertainties of calculated hydraulic heads and fluxes are quantified with the same normalized metric used for the lithological stochastic models.

5) Scatter plots are constructed to investigate correlations between prediction uncertainties for the different models.

2.2 Study area and geological setting

An area of 10 km × 10 km, covering central Glasgow (west central Scotland) alongside the River Clyde, was selected as the study area (Figure 1). Comprehensive descriptions of the geological and hydrogeological settings can be found in previous publications (e.g., Campbell et al., 2010; Finlayson et al., 2010; Finlayson, 2012; Turner et al., 2014; Kearsey et al., 2015) and reports published by the British Geological Survey (e.g., Browne and McMillan, 1989; Hall et al., 1998; McMillan et al., 2005; Merrit et al., 2007; Bonsor and Ó Dochartaigh, 2010; Merrit et al., 2012). Here, we present a synthetic description of main stratigraphic and hydrogeological features that are relevant for this study.

The stratigraphic setting of the study area is characterized by a sequence of Quaternary deposits. These were deposited during and after the last Late Devensian glaciation, and overlie a bedrock consisting of Carboniferous sedimentary rocks. Among the glacial lithostratigraphic units, the most widespread is the Wilderness Till Formation (Figure 1), a heavily compacted diamicton of glacial origin consisting of massive to locally stratified mixtures of pebbles, cobbles, and boulders in a sandy, silty, clayey matrix. This

ACCEPTED MANUSCRIPT

unit directly rests on top of the Carboniferous bedrock in most of the study area, and is particularly abundant across areas of higher ground on either side of the River Clyde valley. The Wildeness Till Formation is overlain by post-glacial lithostratigraphic units consisting of highly heterolithic deposits of glacio-fluvial, glacio-marine, glacio-lacustrine, and shallow, fluvially influenced, estuarine environments. The most common post-glacial units outcropping in the River Clyde valley are the Paisley Clay Member, generally consisting of laminated clay and silt, and the Gourock Sand Member, including fine- to coarse-grain sand, as well as silt, clay and gravel beds. Anthropogenic deposits of different lithology are also widespread in the study area, but they were not considered.

Groundwater circulation in the study area is poorly understood and still subject to investigation (Bonsor et al., 2010). Available data supported by modelling analysis (Turner et al., 2014) suggest a general flow direction from elevated areas toward the Clyde River valley where river stage oscillations are the main controlling factor for groundwater levels in the Quaternary deposits. Transmissivity values (50 m^2/d -100 m^2/d) also indicate a likely contribution of the Carboniferous bedrock to the regional hydrogeological system (Ball et al., 2006). However, the connection between the bedrock and the Quaternary deposits is still not fully understood (Turner et al., 2014). A limited number of hydraulic conductivity measurements in the unconsolidated deposits are reported by Bonsor et al. (2010). These data, derived mainly from permeameter tests, indicate high heterogeneity, which is to be expected given the highly heterolithic nature of the deposits. For example, variations up to 5 orders of magnitude are observed in deposits from the Gourock Sand Member. The geometric mean of the observations in samples from the Paisley Clay Member is equal to 0.11 m/d, although values range from a minimum of 0.07 m/d to a maximum value in excess of 150 m/d. This demonstrates a relatively more uniform

ACCEPTED MANUSCRIPT

distribution of hydraulic conductivity than is observed in the Wilderness Till Formation (geometric mean equal to 0.05 m/d).

2.3 Hard borehole data

The borehole data set consists of 21,320 descriptions of lithology at specific depths in 4391 geotechnical boreholes (Figure 2). Since these data were collected over several decades by different investigators, the dataset required an extensive process of homogenization (Kearsey et al., 2015). At the end of this process six lithofacies were identified based on textural analysis of borehole log descriptions supported using grain size analysis data (Williams and Dobbs, 2012). In contrast to previous work we have further reduced the number of lithofacies to four, namely "soft clay" (sftC), "stiff clay diamicton" (stCD), "silt and sand" (SZ), and "sand and gravel" (SG). The "organic" lithofacies observed in the original dataset was not considered in the present study because it was observed in a very limited number of boreholes (see Figures 4 and 5 in Kearsey et al., 2015). Observations for lithofacies SZ in this work were defined by merging the originally distinct lithofacies "silt" and "sand".

The totality of these lithofacies observations was split into two subsets, one for model training and hard conditioning, while the second subset was used for model validation (Figure 2). Since the boreholes are highly clustered around specific areas, and in particular distributed along the road network, data selection for these subsets was not random. Instead, 52 clusters of boreholes were manually chosen to represent the model validation data, while the remaining boreholes were used for model training and as hard conditioning data. Validation clusters consist of a variable number of boreholes, ranging from a minimum of 2 (cluster #35) up to 14 (cluster #46). Before being used in the geostatistical analysis, both subsets were further processed by calculating the most frequent lithofacies that was found within each cell of the 3-D grid used for the

ACCEPTED MANUSCRIPT

geostatistical simulations. By the completion of this declustering process, hard conditioning data consist of 6684 lithofacies observations, while 2453 observations were used for model validation.

2.4 Soft lithostratigraphic information

A 3-D deterministic geological model of the lithostratigraphic setting of the superficial deposits has been made available for central Glasgow (Merrit et al., 2012). The model was developed using the GSI3D software (Kessler et al., 2009), and it is based on interpolation through triangulation of the boundaries between 13 lithostratigraphic units. These boundaries were defined using 101 interpreted cross-sections, which were constructed after considering several sources of geological data including borehole logs, digital elevation models, current and historical geological maps, and scanned geological cross-sections. The geotechnical boreholes considered in this work and the boreholes used for the development of the lithostratigraphic model constitute two different datasets, albeit a limited intersection between the two groups of data cannot be completely ruled out. A 3-D view of this model showing the lithostratigraphic units considered in this study is shown in Figure 1b.

Conversion of the 3-D deterministic lithostratigraphic model into usable information for the geostatistical simulations began with the estimation of the marginal probabilities of the four lithofacies within each lithostratigraphic unit. For this estimation, borehole data was initially assigned to groups according to the spatial distribution of the lithostratigraphic units in the deterministic model. Then, marginal probabilities were estimated with the following equation:

$$p_{i,j} = \frac{n_{i,j}}{\sum_i n_{i,j}} \qquad (1)$$

ACCEPTED MANUSCRIPT

where $n_{i,j}$ is number of occurrences of lithofacies i within each lithostratigraphic unit j. Soft lithostratigraphic data were then produced by sampling the distribution of lithostratigraphic units in the model with a 3-D grid, and by assigning correspondent marginal probabilities to the nodes of this grid. Several resolution values were tested before choosing a sampling grid with a resolution of four times larger than that of the grid used for geostatistical simulations. This value was chosen because it allows a sufficiently accurate representation of the geological structure in the geostatistical simulations with a reasonable number of soft conditioning points.

Estimated marginal probabilities for the considered lithofacies within each lithostratigraphic unit are presented in Table 1. Values indicate a unimodal distribution of lithofacies in most of units except for the Gourock Sand Member and the Cadder Sand and Gravel Formation, which are characterized by two equally frequent lithofacies. The occurrence of lithofacies stCD is highly likely to be found in the glacial units, in particular in the Wilderness Till Formation, but it is much less likely in the post-glacial units in the River Clyde valley where other lithofacies are significantly more frequent. For instance, lithofacies sftC is the most common lithofacies in the Paisley Clay Member and also in the Law Sand and Gravel Member. The probability of lithofacies SG is generally low in all the units except in the Broomhouse Sand and Gravel Member, while the probabilities of lithofacies SZ and sftC are very variable in the post-glacial units.

Values presented in Table 1 also confirm the inadequacy of the 3-D lithostratigraphic model in accounting for small scale lithological variability. This has also been previously demonstrated in the work of Kearsey et al. (2015) by comparisons between lithological observations in the boreholes and the main lithotypes (i.e., gravel, sand, etc.) in the published lithostratigraphic description of each unit. A similar assessment has been conducted in the present study by comparing the most probable lithofacies for

ACCEPTED MANUSCRIPT

each lithostratigraphic unit against the considered clusters of validation data. Results show that there is approximately one match in every three observations within each cluster. The poor performance of the deterministic lithostratigraphic model in predicting small scale lithological variability is not surprising given the highly heterogeneous nature of the deposits, and the fact that lithostratigraphic variability in the model and lithological variations in the borehole data are defined over two different scales. It is very likely that this inaccuracy is observed also for other hydrogeological properties, particularly for hydraulic conductivity. The question then is not if we can use this type of models to make inferences about properties and processes at scales that are smaller than those over which they are defined. Rather it is how we can address if it is possible to use their geological realism to inform smaller scale models of the subsurface, and how much such additional information affects the uncertainty of predictions of properties and processes.

2.5 Geostatistical modelling

Equally probable realizations of the distribution of the lithofacies were generated with the transition probability approach with the T-PROGS code (Carle and Fogg, 1996, 1997; Carle, 1999). With this approach, the spatial structure of the data is represented by transition probabilities rather than by the variogram or the covariance as in traditional geostatistical methods. The transition probability $t_{i,k}$ from a category (e.g., lithofacies) i to another category k is defined in terms of the conditional probability:

$$t_{i,k}(\mathbf{h}) = \Pr\left\{k(\mathbf{x}+\mathbf{h}) \mid i(\mathbf{x})\right\} \qquad (2)$$

where $\mathbf{x}$ and $\mathbf{h}$ are two vectors indicating spatial location and lag distance, respectively. From the definition of transition probability in Equation (2), this approach can be seen as a Markovian approach since the occurrence of category k at location $\mathbf{x} + \mathbf{h}$ is only dependent on the occurrence of category i at location $\mathbf{x}$. Therefore, a three-dimensional continuous-lag Markov Chain model can be developed to model discrete transition probabilities

ACCEPTED MANUSCRIPT

observed in the data. These models consist of linear combinations of mathematical functions, one for each direction in space and are generally in the form of exponentials, relating the transition probabilities to the lag **h**.

Markov chain models can be computed from knowledge of the entries in the embedded transition probability matrix, as well as of values of volumetric fraction, mean length and thickness of each category. The entries in the embedded transition probability matrix represent the conditional probabilities of a certain category to occur adjacent to the others along particular directions. As with a variogram-based geostatistical analysis, these input parameters are estimated by fitting a model to the transition probabilities observed in the data. A background category is also chosen such that its entries in the embedded transition probability matrix are calculated by difference from the entries of the other categories. The developed Markov chain model is then used as input for the generation of conditional realizations of the distribution of the categories. This is a two-step procedure that includes an initial step to generate a preliminary configuration based on Sequential Indicator Simulation (Deutsch and Journel, 1992), and a successive optimization step based on the simulated quenching algorithm, which is performed to improve the agreement between measured and modeled transition probabilities (Carle, 1999).

Two different transition probability-based stochastic models were implemented to generate three-dimensional conditional simulations of the spatial assemblage of the lithofacies in the study area. One model (thereafter referred to as M1) is entirely based on hard borehole data, which were used to calculate the transition probabilities and as conditioning data in the simulations. In particular, the Markov chain model, which was calibrated to match observed transition probabilities, assumes isotropic behavior in the horizontal plane (Kearsey et al., 2015) with lithofacies stCD as the background category. Volumetric proportions of the lithological categories in the model M1 are also consistent

ACCEPTED MANUSCRIPT

with correspondent values in the borehole hard data. The second model (thereafter referred to as M2) uses both hard borehole data and generated soft lithostratigraphic information in the transition probability calculation, Markov chain model development, as well as for conditioning. Therefore, the Markov chain model in model M2 differs from model M1 in terms of volumetric proportions, mean lengths, and embedded transition probabilities in the horizontal plane. However embedded transition probabilities in the vertical direction are the same for both models since the high resolution of the borehole data in the vertical direction allowed a sufficiently accurate calibration of the model without additional soft data. Lithofacies sfCD was also chosen as background category for model M2.

As shown in Figure 3, conditioning data for M1 and M2 are represented by the probability of occurrence of each lithofacies at points in the 3-D domain. Hard conditioning points representing to the boreholes are defined by indicators defined as (e.g., Weissmann and Fogg, 1999):

$$\mathbf{I}_k(\mathbf{x}) = \begin{cases} 1, \text{ lithofacies } k \text{ occurs at } \mathbf{x} \\ 0, \text{ otherwise} \end{cases} \quad (3)$$

These hard points are used by both models M1 and M2. For soft conditioning points, which were considered only in model M2, corresponding indicator values were replaced by the marginal probabilities p_i (Equation 1) of each lithofacies derived from the lithostratigraphic model.

Stochastic realizations of lithofacies distributions were generated using a regular 3-D grid with a resolution of 50 m in the horizontal directions (x and y) and 2 m in the vertical direction (z). On the horizontal plane, the grid covers an area of 9.5 km × 9.5 km. The top surface of the grid was defined by the NEXTMap® Britain Digital elevation model (50 m resolution), which had been modified to remove anthropogenic ground, whilst the

ACCEPTED MANUSCRIPT

bottom surface corresponds to the top surface of the pre-Quaternary bedrock in the lithostratigraphic model by Merritt et al. (2012).

2.6 Groundwater flow modelling

Groundwater flow was simulated by solving the equation of steady-state flow in porous media, which can be written as:

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y \frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right)+f=0 \qquad (4)$$

where K_x, K_y, and K_z are the principal components of the hydraulic conductivity tensor **K** in the respective coordinate directions, h is hydraulic head, and f is a volumetric sink/source term. For a volume element centered about any point (x, y, z), Equation (4) can be derived by combining a fluid conservation condition and Darcy's law, which defines the fluxes into and out of the element in directions perpendicular to its faces. Accordingly, these fluxes can be written as:

$$q_x=-K_x \frac{\partial h}{\partial x},\ q_y=-K_y \frac{\partial h}{\partial y},\ q_z=-K_z \frac{\partial h}{\partial z}, \qquad (5)$$

where q_x, q_y, and q_z are the scalar components of the Darcy flux (specific discharge) vector **q** in the respective coordinate directions. It is implied in Equations (4) and (5) that the principal components of **K** are aligned with the x, y, and z coordinate axes.

A numerical solution of Equation (4) was calculated with a finite-difference model implemented with MODFLOW-2005 (Harbaugh, 2005). The objective of this model is to provide a basic representation of groundwater system in order to understand the impact of soft lithostratigraphic conditioning on predictions of groundwater heads and fluxes. Accordingly, input parameters and boundary conditions were not calibrated against measured head values, but they were simply defined on the basis of available data or taken from literature (Turner et al., 2014). Therefore, model outputs should only be considered

ACCEPTED MANUSCRIPT

as generic and interpretative. Extensions of the finite-difference numerical grid in the *x* and *y* directions, as well as grid resolution and location of the cell centers, match the correspondent values of the grid used for geostatistical simulations. However, the base of the numerical grid corresponds to the horizontal plane at elevation *z* = -36.87 m above sea level (a.s.l.), which is the lowest elevation of the bedrock in the study area, while the top surface corresponds to the horizontal plane at elevation z = 1.13 m (a.s.l.). This value was chosen because measured groundwater levels indicate that most of the saturated zone within the Clyde valley sediments lies below sea level (Turner et al., 2014). With these assumptions, the numerical grid consists of 192 rows, 192 columns, and 19 layers. As shown in Figure 4, the modelled domain comprises both Quaternary deposits alongside the Clyde River and the bedrock. However, only the cells representing the deposits were considered in the analysis of prediction uncertainty.

Head-dependent flux boundary conditions were assigned to grid cells at the four lateral boundaries of the domain (GHB in Figure 4). Reference head and conductance values for the assignation of these boundary conditions were taken from the results of a previous groundwater flow model, whose domain encapsulates the whole Glasgow urban area (Turner et al. 2014). Non-uniform recharge rates assigned to the top layer of the numerical grid were also derived from estimated values of distributed recharge within this model. The bottom surface of the domain was considered a no-flux boundary. To simulate groundwater/surface water interactions, head-dependent flux boundary conditions were also assigned to the cells corresponding to the position of River Clyde in the numerical grid (RIV in Figure 4). To define these conditions, a river bed elevation of 2 m below ground surface and a uniform river bed conductance of 10 m^2/d were assumed, while river stage elevations were estimated from the NEXTMap® Britain Digital elevation model.

ACCEPTED MANUSCRIPT

Hydraulic conductivity fields in the flow simulations are directly linked to the stochastic realizations from models M1 and M2. For each realization, the link was established through the assignment of appropriate K values in the numerical grid in parallel with corresponding simulated lithofacies distribution. A uniform K value was assumed for each lithofacies, with values ranging from 0.01 m/d for lithofacies sftC to 150 m/d for lithofacies SG. Intermediate values were assumed for lithofacies sftCD and SZ (0.1 m/d and 15 m/d, respectively), while a value of 1 m/d was assigned to the bedrock. These values were chosen to be consistent both with experimental data (Ball et al., 2006; Bonsor et al., 2010) and also with calibrated values from the previous large-scale model (Turner et al., 2014).

For each cell of the numerical grid, the components q_x, q_y, and q_z of the Darcy flux vector were estimated with a finite-difference approximation of Equation (5). For instance, the component q_x at the interface between two cells $(i, j-1, k)$ and (i, j, k) along the x direction was initially calculated with the following:

$$q_x = -K_{i,j-1/2,k} \frac{h_{i,j,k} - h_{i,j-1,k}}{\Delta x} \qquad (6)$$

where $h_{i,j,k}$ and $h_{i,j-1,k}$ are the simulated hydraulic heads at the centers of the cells, Δx is the grid spacing along x, and $K_{i,j-1/2,k}$ is the harmonic mean of hydraulic conductivity between the two cells. A similar expression was then used to estimate the flux component at the other interface between between (i, j, k) and $(i, j+1, k)$. Finally the value of q_x at the center of the cell was calculated as the average between the fluxes at the two interfaces. This procedure was repeated for all the cells of the grid and for the other components q_y and q_z in the other directions.

ACCEPTED MANUSCRIPT

2.7 Information entropy as a measure of prediction uncertainty

The concept of information entropy, introduced by Shannon (1948), was applied to quantify the prediction uncertainty of the implemented models. For a system with a discrete number of probable outcomes, information entropy is a measure of "missing information" i.e., the amount of information required for a complete probabilistic description of the system. The concept is appealing because it is based on a metric that is equal to 0 when no uncertainty exists (i.e., there is only one possible outcome) and to a maximum when there is the greatest uncertainty (i.e., all outcomes are equally likely). Moreover, it does not change when an additional outcome with null probability is added. This metric is the information entropy H, which is defined as:

$$H = -\sum_{i=1}^{N} p_i \log p_i \qquad (7)$$

where p_i is the probability of the outcome i out of N possible outcomes. The base of the logarithm in Equation (7), which is important for establishing units of H (e.g., bits, nats, etc.), is irrelevant here. As previously stated, H reaches the absolute maximum when $p_i = 1/N$ for every outcome i. Therefore, the maximum (H_{MAX}) is given by:

$$H_{MAX} = -\sum_{i=1}^{N} \frac{1}{N} \log \frac{1}{N} = \log N \qquad (8)$$

Accordingly, we can define a relative measure of uncertainty, with values between 0 (i.e., no uncertainty) and 1 (i.e., highest uncertainty), as the normalized metric (H_{NORM}):

$$H_{NORM} = \frac{-\sum_{i=1}^{N} p_i \log p_i}{H_{MAX}} \qquad (9)$$

Additional details on the theoretical aspects of information entropy can be found in the original work of Shannon and in several textbooks on Information Theory (e.g., Stone, 2015).

ACCEPTED MANUSCRIPT

While the concept of information entropy has been applied to various problems in hydrology (e.g., Singh, 2011), its potential as a measure of uncertainty for geological models as well as in groundwater flow and solute transport modelling has received less attention. Woodbury and Ulrych (1996), for instance, applied entropy concepts to recover the release history of a contaminant plume in a 1-D system with constant flow velocity. Mays et al. (2002) used a metric similar to the one defined in Equation (9) to evaluate the complexity of numerical simulations of infiltration through unsaturated heterogeneous soils. Information entropy has also been applied to quantify uncertainty in the context of structural geological models (Wellmann and Regenauer-Lieb, 2012) and geological maps (Wellmann, 2013; Stafleu et al., 2014), as well as to estimate spatial disorder in synthetic aquifers (Scheibe, 1993; Scheibe and Murray, 1998) and in three-dimensional realizations of distributions of sand and clay (Huang et al., 2012).

In this work, spatial distributions of the values of H_{NORM} for models M1 and M2 were generated by calculating the relative frequencies of the lithofacies at each node of the simulation grid. These frequencies were estimated on the basis of 90 realizations since it was observed a stabilization in the estimated values of H_{NORM} after about 80 realizations. Relative frequencies were then considered as p_i in Equation (9), where the maximum entropy H_{MAX} is equal to log(4) being four the number of lithofacies considered. A similar approach, based on the same number of realizations, was used to calculate H_{NORM} in the cells representing Quaternary deposits in the numerical grid of the flow model (Figure 4). In particular, two values of H_{NORM} were calculated, one for groundwater head predictions and another for predictions of the magnitude of Darcy flux vector ($\sqrt{q_x^2 + q_y^2 + q_z^2}$). In both cases, the distributions of simulated values were binned over 0.5 unit intervals, and their relative frequencies were considered for the calculation of H_{NORM}. Log transformed (base = 10) values were considered for the magnitude of Darcy flux. The information

ACCEPTED MANUSCRIPT

entropy of a uniform distribution of values binned over the same intervals was considered as H_{MAX} for normalization.

3 Results

3.1 Transition probability and Markov chain models

Values of volumetric fraction, mean length and mean thickness of each lithofacies are presented in Table 2 while transiograms showing observed transition probabilities and fitted Markov chain models for models M1 and M2 are shown in Figure 5. Horizontal transition probabilities were measured from data with a lag spacing of 100 m and a tolerance of 50 m, while a lag spacing of 1.5 m and a tolerance of 0.75 m were considered for calculating transition probabilities in the vertical direction. Developed Markov chain models honour transition probability data at the first lag. This condition, which is a suggested modelling practice given the large amount of training data (Carle, 1999), was also chosen because it allowed unbiased comparison between the two models, which would have been difficult if a manual calibration of the input parameters had been performed. With this approach, mean lengths for the lithofacies are also defined by the interception of the first lag transition rate (i.e. the slope of the Markov chain model as it approaches lag zero) to the horizontal axis of plots of auto-transition probabilities (diagonal elements in Figure 5).

The volumetric fraction for each lithofacies was estimated from hard borehole data for model M1 and from both hard and soft data for model M2. Values for the two models are in general agreement for all the lithofacies except for lithofacies stCD. In particular, the addition of soft conditioning data resulted in a about 10% increment in the volumetric fraction of this lithofacies compared to the value from the hard borehole data alone. For all the other lithofacies, the difference in estimated volumetric fractions in the two models is

ACCEPTED MANUSCRIPT

within a 6% to 9% range. The low volumetric fraction of lithofacies stCD in the hard borehole data is an effect of data clustering (Figure 2). This lithofacies is in fact particularly frequent in the Wilderness Till Formation (Table 1), which is the prevalent lithostratigraphic unit in the northern and southern sectors of the study area (Figure 1), where hard borehole data are more scattered. The addition of soft data in model M2 also allowed a more realistic estimation of the spatial continuity of this lithofacies as indicated by the value of the mean length for model M2, which is almost three times larger than the value estimated for model M1. Mean lengths for lithofacies sftC and SZ in the two models are practically identical, while the mean length of lithofacies SG in model M2 is about one third lower than the value estimated for M1. From the comparison of transiograms for this lithofacies in Figure 5a with those in Figure 5b, it is evident that the addition of soft information reduced the uncertainty in the transition probability estimations, as shown by a less scattered distribution of points in Figure 5b. This allowed a more accurate fitting of a Markov chain model.

3.2 Simulated lithofacies distributions

Statistical analysis of the stochastic realizations generated with model M1 indicate that the clayey lithofacies sftC is expected to be the most frequently occurring lithofacies in the studied area. However, as shown in Figure 6a and 6c, the spatial distribution of the all lithofacies in areas far from borehole data is very indistinct especially across sectors of higher ground on either side of the River Clyde valley. Here, according to the geological interpretation provided by the lithostratigraphic model and the estimated marginal probabilities (Table 1) we should expect a higher frequency of lithofacies stCD. The relatively frequent occurrence of this lithofacies at the ground surface in the River Clyde valley is also not consistent with the lithostratigraphic interpretation. These geological inconsistencies disappear with the addition of the soft lithostratigraphic information

ACCEPTED MANUSCRIPT

considered in model M2. In fact the distribution of lithofacies for model M2 shown in Figures 6c is in agreement with the distribution of the Wildeness Till Formation in the lithostratigraphic model, as well as the estimated frequency of lithofacies stCD with depth (Figure 6d) is in agreement with the interpreted lithostratigraphic unconformity between glacial and post-glacial deposits. Compared to model M1, the occurrence of lithofacies SZ and sftC in model M2 is also more consistent with a realistic conceptualization of a fluvial depositional system, with sand and silt deposited in point bars and natural levees along the river channel and finer sediments deposited on the floodplain.

The high uncertainty in the simulated probability distributions for model M1 is reflected in the distributions of H_{NORM} (Figure 6e-g), which is ubiquitously high (> 0.90) except in areas around hard conditioning points. The shape of histogram of the estimated values (Figure 7a) is in fact skewed toward the right with average and modal values equal to 0.883 and 0.947, respectively. The interpretation of these values according to the concept of information entropy indicates that two or more lithofacies are equally likely in the vast majority of the nodes of the domain. Conversely, in the spatial distribution of H_{NORM} for model M2 (Figure 6f-h) it is possible to identify areas where the occurrence of a particular lithofacies is significantly more likely than other. Although areas of high uncertainty remain, the average value of H_{NORM} for model M2 (0.705) is about 20% lower than the correspondent values for model M1.

Spatial distributions of estimated occurrence probabilities for the lithofacies (Figure 8) indicate that while in model M1 the location of areas with high (or low) probability is only controlled by the borehole hard data, in model M2 this location is also influenced by the geological structure imparted by the soft conditioning data. As a result, model M2 allows a more precise definition of the spatial distribution of occurrence probability for the lithofacies. For instance, the map of probability for lithofacies sftC

ACCEPTED MANUSCRIPT

calculated with model M2 (Figure 8b) indicates likely occurrence (> 0.40) in the valley deposits, especially in more distal areas from the River Clyde, as well as in depressions and tributary valleys on either side of the main valley. In all the other sectors of the studied area, probability is very low ($p_i < 0.10$). On the other hand, the probability distribution of this lithofacies from model M1 (Figure 8a) is characterized by a predominance of intermediate values ($0.25 < p_i < 0.35$) indicating high uncertainty except in proximity of the borehole locations. The same considerations apply for all the considered lithofacies.

Results of the validation test to estimate the accuracy of the two implemented models in predicting lithofacies occurrence at specific depths in 52 clusters of boreholes are presented in in Figure 9. The accuracy is expressed in terms of percentage of correct predictions in each cluster, while the uncertainty of the predictions is represented by the average of H_{NORM} values for each cluster. Results indicate that the addition of soft information increased the percentage of correct predictions in the clusters by 10% on average (47% for M1 vs. 57% for M2). This is a significant improvement given the high heterogeneity of the deposits in the study area and their complex depositional history. The number of clusters with a 100% of accurate predictions also increased (6 for model M2 vs. 1 in model M1). Four of these clusters (#43; #45; 50, #52) and, in general, the clusters where we observed a more significant improvement in accuracy are located in the north-western sector of the study area where the influence of soft conditioning on simulated lithofacies distributions is more evident (Figure 7b). A reduction in accuracy was observed in four clusters (#11, #23, #29 #35). However, three of these clusters have the lowest number of boreholes and therefore calculated accuracy values may not be significant. On the other hand, a 40% increment was measured in the percentage of accurate predictions in the cluster with the highest number of boreholes (#46). As expected, predictions from model M2 are also less uncertain than those from model M1, as indicated by lower values

ACCEPTED MANUSCRIPT

of H_{NORM} in all the clusters. The average normalized entropy of the predictions for model M2 is about 16% lower than the corresponding value for model M1.

3.3 Groundwater flow model results

Simulated hydraulic heads distributions indicate a convergent flow field with flow paths moving groundwater from both sides of the valley toward the River Clyde. Calculated magnitude values of the Darcy flux reflect the high heterogeneity of the K fields derived from the lithofacies distributions. For simulations considering realizations of the M1 model, the distribution of the ensemble mean values of the log-transformed magnitudes is characterized a mean of -2.07 and a variance of 0.25 (base = 10; values in m/d). Similar values were estimated for the simulations considering model M2. The variability of simulated values in a single cell of the modelled domain is much higher compared to variability of the distribution of the ensamble mean values. For example, the variance of the magnitudes of the Darcy flux estimated in a single cell is about 1.3 on average for both modelling scenarios.

Two-dimensional maps showing the distribution of calculated H_{NORM} values on the top surface of the numerical grid for model M1 and M2 are compared in Figure 10, while the histograms of all the calculated values for head and flux predictions are presented in Figure 7b and 7c. For both models, normalized entropy values for head predictions are about one third of the corresponding values for Darcy flux predictions. As expected, spatial distributions of H_{NORM} for head predictions are influenced by the boundary conditions in flow model in particular in the cells representing the River Clyde (Figures 10 and 10b). On the other hand, spatial distributions of low H_{NORM} values for flux predictions are only marginally influenced by boundary conditions in the flow model while they are more strongly affected by the positions of hard conditioning points in the lithofacies models (Figures 10c and 10d). As for the lithofacies models, the uncertainty of the

ACCEPTED MANUSCRIPT

predictions from the groundwater flow model is also generally lower when soft lithostratigraphic information is taken into account in the definition of the *K* fields. In fact the average value H_{NORM} for groundwater head predictions based on realizations of the M2 model is about 37% lower than the correspondent value for model M1 (0.238 vs. 0.327). Normalized entropy values are particularly low (< 0.25) in the western sector of the domain for the scenario considering model M2. Low values in both scenarios were also calculated in the eastern sector in an area between the meanders of the Clyde River, but these are the effect of the head-dependent boundary condition used to simulate the river. For flux predictions, the average H_{NORM} for the scenario considering model M2 is about 11% lower than the value estimated in simulations considering model M1 (0.753 vs. 0.688). However, as shown in the Figure 7c, more than 70% of the estimated values for the scenario considering model M1 are between 0.75 and 0.85, while less than 20% of values falls in that range for the scenario considering model M2. The higher prediction uncertainty for the simulations considering model M1 is also evident from the comparison of spatial distributions of H_{NORM} (Figure 10c and 10d). As for lithofacies predictions, the comparison also indicates that the incorporation of soft information in model M2 allows a more precise identification of areas with different levels of uncertainty.

4 Discussion

The comparisons of the results of the validation test and of spatial distributions of normalized entropy indicate an improved predictability and a reduction in uncertainty when soft information is integrated into stochastic simulations of lithofacies distribution. Similar reductions in prediction uncertainty of groundwater heads and fluxes were observed when the geological structure derived from the lithostratigraphic model was taken into account in the definition of the hydraulic conductivity fields in the groundwater

ACCEPTED MANUSCRIPT

flow model. Note that since the groundwater flow model was not calibrated against observations, our results regarding groundwater heads and fluxes predictions indicate an improvement in precision. This does not necessarily correspond to an improvement in prediction accuracy, which was not tested in the present work.

A more quantitative analysis of the impact of soft lithostratigraphic information on the uncertainty of the implemented models was conducted by calculating the following difference:

$$\Delta H_{NORM} = H_{NORM}^{M1} - H_{NORM}^{M2} \quad (10)$$

For both lithofacies and flow models, ΔH_{NORM} represent changes in normalized entropy between the modelling scenario considering only borehole hard data and the scenario considering both soft lithostratigraphic information and hard data. Values of ΔH_{NORM} were calculated for the cells of the groundwater flow model grid representing Quaternary deposits (Figure 5), and for the corresponding nodes of the simulation grid used in models M1 and M2. Although the spatial distributions of ΔH_{NORM} indicate a predominance of positive values, negative values were also observed. For lithofacies predictions (Figure 11a), these negative values are located around the boreholes indicating that the addition of soft conditioning points in the stochastic simulation have the effect of increasing entropy around the hard conditioning locations. This is a significant result because, on the basis of the principle of maximum entropy, it shows that model M2 provides the best representation of lithofacies distribution around the hard conditioning points. For the same principle, model M2 also provides the best predictions of groundwater heads in sectors of the numerical grid that are in proximity of boundary conditions, especially for areas between meanders of the Clyde River in the eastern sector of modelled domain where there is a concentration of negative ΔH_{NORM} values (Figure 11b). For flux predictions

ACCEPTED MANUSCRIPT

(Figure 11c), negative values of ΔH_{NORM} are also clustered around the hard conditioning points of lithofacies realizations. This result indicates that model M2 is the least biased for making predictions around model constraints (i.e., hard conditioning boreholes and/or boundary conditions), and therefore it best represents the current state of knowledge.

Scatter plots of ΔH_{NORM} values for lithofacies predictions vs. the correspondent values for groundwater head and flux predictions provide indications about the relationship between prediction uncertainties from the different models (Figure 12). Since these plots can be seen as a sensitivity analysis of the uncertainties in the outputs of the groundwater flow model with respect to the uncertainty in the lithological model, the slopes of the linear least-square models fitted to the data provide an indication of the sensitivity coefficients. Note that boundary-condition cells in groundwater model, as well as hard conditioning points in the lithofacies models, were not considered in this analysis. The comparison of these scatter plots indicates that, with respect to the uncertainty in lithofacies predictions, the uncertainty in groundwater flux predictions is almost four times more sensitive than the uncertainty in groundwater head predictions. This result has important practical implications for groundwater modelling because it shows that adding geological information to better characterize subsurface heterogeneity may be a very effective way to reduce the uncertainty of predictions based on groundwater fluxes. However, if the objective of the model is a simple reconstruction of the hydraulic head distribution in a certain area, adding geological information may not be as an effective approach in reducing the model prediction uncertainty as the calibration of input parameters and boundary conditions, or the collection of additional head observations.

ACCEPTED MANUSCRIPT

5 Summary and conclusions

We have proposed an approach to generate conditional stochastic realizations of the spatial distribution of geological categories that account for geological structure. The approach is based on the extraction of information from a deterministic conceptualization of the subsurface to be used in geostatistical analysis for establishing models of spatial correlation, as well as for conditioning stochastic simulations. This information is provided by a grid of soft data points representing marginal probabilities of the categories in the units of the deterministic model. As a result, the realistic geological structure of the deterministic model is imprinted in the realizations of the stochastic model. Although this study focused on the transition probability method, the proposed approach can also be applied to variogram-based indicator algorithms, and generally to any other geostatistical method capable of taking into account soft information as conditioning data.

The approach was tested to simulate the distribution of four lithofacies in highly heterolithic Quaternary deposits. A transition probability-based stochastic model (M2) was implemented using both hard borehole data and soft data extracted from a 3-D deterministic lithostratigraphic model. Another model (M1) considering only hard borehole data was also implemented. Lithofacies distributions for the two models were then used to define hydraulic conductivity fields for prediction groundwater head distribution and Darcy fluxes. Comparisons between predicted results from the two models permit the following conclusions regarding the proposed methodology to be made. These also provide general insights about the incorporation of soft geological information into stochastic realizations of subsurface heterogeneity and its impact on groundwater flow modelling.

1) Soft lithostratigraphic information increased the predictability of the stochastic lithofacies model. The number of correct predictions in 52 clusters of validation boreholes

ACCEPTED MANUSCRIPT

increased by 10% on average, with values up to 80%. A potential limitation of the approach is the risk of introducing systematic errors in the definition of the geological structure due to flaws and uncertainty in the deterministic model. However, in our case study we measured a reduction in accuracy in a very limited number of validation clusters notwithstanding a very low accuracy of the lithostratigraphic model and high associated uncertainty.

2) An overall reduction (about 20% on average) in prediction uncertainty measured with a metric based on Shannon information entropy was observed in the stochastic model that considered geological structure. This reduction in spatial entropy confirms the conclusions of two previous studies (Scheibe, 1993; Huang et al. 2012). These studies however considered synthetic representations of geological systems characterized by a binomial distribution of facies, while in this work we have investigated spatial entropy in a real site with four lithofacies.

3) Compared to the results of the model based on hard data exclusively, increased entropy around the hard conditioning points was observed in the model considering soft conditioning data. According to the principle of maximum entropy, this indicates that incorporation of soft information allows a better representation of the stochastic distribution of the lithofacies in those areas. When these realizations were used as input for groundwater flow simulations, a similarly positive increment in entropy was observed in areas where values of predicted hydraulic heads were mostly influenced by boundary conditions.

4) The representation of the geological structure in the spatial distribution of lithofacies allows a more precise definition of spatial uncertainty. This can be particularly useful to support the design of geological investigations because it provides a geological basis for the identifications of areas where further exploration is required to further reduce

ACCEPTED MANUSCRIPT

uncertainty, as well as of unsampled areas where this is not necessary because there is sufficient indirect information for making informed predictions about the property of interest.

5) A general reduction in prediction uncertainty for heads and fluxes was observed when soft lithostratigraphic information was taken into account into the definition of the *K* fields in the flow model. In particular, a 37% reduction was observed in the average normalized entropy for head predictions and about 11% for predicted fluxes.

6) Scatter plots of variations in normalized entropy for the implemented models indicate that there is a correlation between variations in prediction uncertainties. For the particular case considered in this study, flux predictions are about 4 times more sensitive to variations in lithofacies uncertainty than the head predictions. This result emphasizes the critical importance of geological information for reducing prediction uncertainty in models that simulate flux-dependent processes such as advective transport, multiphase flow, and groundwater recharge/discharge.

In statistical sciences, the concept of information entropy provides an established framework for the analysis of the uncertainty of categorical data (e.g., Wilcox, 1967). However, before the present study, a relatively small number of studies applied this concept to quantify uncertainty in the context of geological and groundwater flow modeling. Our results indicate that information entropy is an ideal metric to quantify uncertainty in spatially distributed stochastic models of properties and processes, and to compare and correlate uncertainty or variations in uncertainty between different models.

Acknowledgments. This work was undertaken as part of the "Research Fellowship Programme" funded by the British Geological Survey (Natural Environment Research Council). We are grateful to the constructive comments from Murrey Lark and Keith

ACCEPTED MANUSCRIPT

Turner, which have significantly improved this work. Authors would like to thank Majdi Mansour for providing a regional groundwater flow model of the study area. Three anonymous reviewers are grateful acknowledged for their comments, which helped improving the final version of the manuscript. The authors publish with the permission of the Executive Director of the British Geological Survey.

References

Amorocho, J., and B. Espildora (1973). Entropy in the assessment of uncertainty in hydrologic systems and models, Water Resour. Res., 9(6), 1511–1522, doi:10.1029/WR009i006p01511.

Artimo, A., J. Makinen, R.C. Berg, C.C. Abert, and V.P. Salonen (2003). Three-dimensional geologic modelling and visualisation of the Virttaankangus aquifer, southwestern Finland. Hydrogeology Journal, 11, 378–386.

Ball, D., M. Graham, B.É. Ó Dochartaigh, K. Irving K, and E. Simpson (2006). Scottish aquifer properties: 2006 interim report. CR/06/073N, British Geological Survey, Keyworth, UK.

Bianchi, M., C. Zheng, C. Wilson, G. R. Tick, G. Liu, and S. M. Gorelick (2011b). Spatial connectivity in a highly heterogeneous aquifer: From cores to preferential flow paths, Water Resour. Res., 47, W05524, doi:10.1029/2009WR008966.

Bianchi, M., and C. Zheng (2015). Explaining “anomalous” solute transport at the Macrodispersion Experiment (MADE) site from a geological perspective, The MADE Challenge for Groundwater Transport in Highly Heterogeneous Aquifers: Insights from 30 Years of Modeling and Characterization at the Field Scale and Promising Future Directions, 5-8 October, Valencia, Spain. .

ACCEPTED MANUSCRIPT

Blessent, D., R. Therrien R, and K. MacQuarrie (2009). Coupling geological and numerical models to simulate groundwater flow and contaminant transport in fractured media. Computer & Geosciences, 35 (9), 1897–1906.

Blouin M, Martel R, Gloaguen E (2013). Accounting for aquifer heterogeneity from geological data to management tools. Ground Water 51:421–431.

Blöschl, G., and Sivapalan (1995). Scale issues in hydrological modelling — a review. Hydrol. Processes, 9, 251–290.

Bonomi, T. (2009). Database development and 3D modeling of textural variations in heterogeneous, unconsolidated aquifer media: Application to the Milan plain. Computer & Geosciences, 35, 134–145.

Bonsor, H.C., and B.É. Ó Dochartaigh (2010). Groundwater monitoring in urban areas – a pilot investigation in Glasgow, UK. British Geological Survey Internal Report IR/10/020.

Bonsor, H.C., S.H. Bricker, B.É. Ó Dochartaigh and K.I.G. Lawrie (2010). Project progress report 2010-11: Groundwater monitoring in urban areas – a pilot study in Glasgow, UK. British Geological Survey Open Report IR/10/087, 63 pp.

Bredehoeft, J. (2005). The conceptualization model problem-surprise. Hydrogeology Journal 13 (1), 37 – 46.

Browne, M.A.E., and A.A.A. McMillan (1989). Quaternary geology of the Clyde valley. Research Report SA\89\1. British Geological Survey (63 pp.).

Campbell, S.D.G., J.E. Merritt, B.É. Ó Dochartaigh, M. Mansour, A.G. Hughes, F,M. Fordyce, D.C. Entwisle, A.A. Monaghan, and S.C. Loughlin (2010). 3D geological models and their hydrogeological applications: supporting urban development – a case study in Glasgow-Clyde, UK. Zeitschrift der Deutschen Gesellschaft für Geowissenschaften, 161(2), 251-262.

ACCEPTED MANUSCRIPT

Carle, S.F. (1999). T-PROGS: Transition Probability Geostatistical Software, version 2.1. Davis, California: University of California.

Carle, S.F., and G.E. Fogg (1996). Transition probability-based indicator Geostatistics. Mathematical Geology, 28(4), 453–476.

Carle S.F., and G.E. Fogg (1997). Modeling spatial variability with one and multidimensional continuous-lag Markov chains. Mathematical Geology, 29(7), 891–918.

Carle S.F., E.M. LaBolle, G.S. Weissmann, D. VanBrocklin, and G.E. Fogg (1998). Geostatistical simulation of hydrostratigraphic architecture: a transition probability / Markov approach. In: Concepts in Hydrogeology and Environmental Geology No. 2, SEPM Special Publication, p. 147–170

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, L.J. Slooten (2005). Inverse problem in hydrogeology. Hydrogeology Journal (13), 206–222.

Carreño, F., S. García Martínez, J.L. Ramos, R. Fernández Martínez, and A. Mabeth-Montoya (2014). Building a 3D geomodel for water resources management: case study in the Regional Park of the lower courses of Manzanares and Jarama Rivers (Madrid, Spain). Environmental Earth Sciences, 1(1), 1866-6280.

Chitsazan, N., Pham, H. V. and Tsai, F. T.-C. (2014). Bayesian Chance-Constrained Hydraulic Barrier Design under Geological Structure Uncertainty. Groundwater. doi: 10.1111/gwat.12304

Comunian, A, P. Renard, and J. Straubhaar (2012). 3D multiple-point statistics simulation using 2D training images. Computer & Geosciences, 40, 49–65, doi:10.1016/j.cageo.2011.07.009.

Courrioux G., C. Allanic, B. Bourgine, A. Guillen, T. Baudin, F. Lacquement, S. Gabalda, F. Cagnard, B. Le Bayon, J. Besse, D. Marquer, P. Trap, P.-H. Leloup, D. Schreiber

ACCEPTED MANUSCRIPT

(2015). Comparisons from multiple realizations of a geological model. Implication for uncertainty factors identification. IAMG 2015 : The 17th annual conference of the International Association for Mathematical Geosciences, September 2015, Freiberg, Germany.

Copty, N., and Y. Rubin (1995), A stochastic approach to the characterization of lithofacies from surface seismic and well data, Water Resour. Res., 31(7), 1673–1686, doi:10.1029/95WR00947.

D'Agnese, F.A., C.C. Faunt, M.C. Hill, and A.K. Turner (1999). Death valley regional ground-water flow model calibration using optimal parameter estimation methods and geoscientific information systems, Advances in Water Resources 22(8), 777-790.

de Marsily, G., F. Delay, J. Goncalves, P. Renard, V. Teles, and S. Violette (2005). Dealing with spatial heterogeneity. Hydrogeology Journal, 13 (1), 161–183.

Dell'Arciprete, D., R. Bersezio, F. Felletti, M. Giudici, A. Comunian, and Ph. Renard (2012). Comparison of three geostatistical methods for hydro-facies simulation: a test on alluvial sediments. Hydrogeology Journal, 20, 299-311.

Deutsch, C.V., and A.G. Journel (1992). GSLIB: Geostatistical Software Library and User's Guide.369 pp. Oxford University Press, London, 369 pp.

Elfeki, A.M.M. (2006). Reducing concentration uncertainty using the coupled Markov Chain approach. Journal of hydrology 317 (1), 1-16

Elfeki, A.M.M, G.J.M. Uffink, and F.B.J. Barends (1995). Stochastic simulation of heterogeneous geological formations using soft information, with an application to groundwater. In Stochastic simulation of heterogeneous geological formations using soft information, with an application to groundwater, in: Groundwater Quality: Remediation and Protection, QG'95, eds. K. Kovar and Krasny, IAHS Publication 225.

ACCEPTED MANUSCRIPT

Emery, X., and L.N. Robles (2009). Simulation of mineral grades with hard and soft conditioning data: application to a porphyry copper deposit. Computational Geosciences, 13(1), 79-89.

Engdahl, N. B., G. S. Weissmann, and N. D. Bonal (2010b), An integrated approach to shallow aquifer characterization: Combining geophysics and Geostatistics. Computer & Geosciences, 14(2), 217–229.

Finlayson, A. (2012). Ice dynamics and sediment movement: last glacial cycle, Clyde basin, Scotland. Journal of Glaciology, 58(209), 487-500.

Finlayson, A., J. Merritt, M. Browne, J. Merritt, A. McMillan, and K. Whitbread (2010). Ice sheet advance, dynamics, and decay configurations: evidence from west central Scotland. Quaternary Science Reviews, 29 (7–8), 969–988.

Fogg, G. E. (1986), Groundwater Flow and Sand Body Interconnectedness in a Thick, Multiple-Aquifer System, Water Resour. Res., 22(5), 679–694, doi:10.1029/WR022i005p00679.

Giambastiani, B.M.S., A.M. McCallum, M.S. Anderson, B.F.J. Kelly, and R.I. Acworth (2012). Understanding groundwater processes by representing aquifer heterogeneity in the Maules Creek Catchment, Namoi Valley (New South Wales, Australia). Hydrogeology Journal, 20 (6), 1027–044.

Gill, B., D. Cherry, M. Adelana, X. Cheng, and M. Reid (2011). Using threedimensional geological mapping methods to inform sustainable groundwater development in a volcanic landscape, Victoria, Australia. Hydrogeology Journal, 19 (7), 1349–1365.

Guardiano, F., and M. Srivastava (1993). Multivariate geostatistics: Beyond bivariate moments. In Geostatistics Troia'92, edited by A. Soares, pp. 133–144, Springer, Dordrecht, Netherlands.

ACCEPTED MANUSCRIPT

Gupta, H. V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye (2012), Towards a comprehensive assessment of model structural adequacy, Water Resour. Res., 48, W08301, doi:10.1029/2011WR011044.

Hall, I.H.S., M.A.E. Browne, I.H. Forsyth, (1998). Geology of the Glasgow District: Memoir for 1:50 000 Geological Sheet 30E (Scotland): 117 S, Nottingham (British Geological Survey). pp. 117.

He, X., J. Koch, T. O. Sonnenborg, F. Jørgensen, C. Schamper, and J. Christian Refsgaard (2014a). Transition probability-based stochastic geological modeling using airborne geophysical data and borehole data, Water Resour. Res., 50, 3147–3169, doi:10.1002/2013WR014593.

He, X. L., T.O. Sonnenborg T.O., F. Jørgensen, K.H. Jensen (2014b). The effect of training image and secondary data integration with multiple-point geostatistics in groundwater modelling. Hydrology and Earth System Sciences, 18(8), 2943-2954.

Hu, L. Y., and T. Chugunova (2008). Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review. Water Resour. Res., 44, W11413, doi:10.1029/2008WR006993.

Huang, L., R.W. Ritzi, and R. Ramanathan (2012). Conservative models: parametric entropy vs. temporal entropy in outcomes. Ground Water 50, no. 2: 199–206. DOI: 10.1111/j.1745-6584.2011.00832.x

Huysmans, M., and A. Dassargues (2012). Modeling the effect of clay drapes on pumping test response in a cross-bedded aquifer using multiple-point Geostatistics. Journal of Hydrology, 450–451, 159–167.

Hyndman, D. W., and S. M. Gorelick (1996). Estimating lithologic and transport properties in three dimensions using seismic and tracer data: The Kesterson aquifer. Water Resour. Res., 32(9), 2659–2670, doi:10.1029/96WR01269.

ACCEPTED MANUSCRIPT

Journel, A.G. (1983). Nonparametric estimation of spatial distributions. Mathematical Geology 15 (3), 445–468.

Kearsey T., J. Williams, A. Finlayson, P. Williamson, M. Dobbs, B. Marchant, A. Kingdon, D. Campbell (2015). Testing the application and limitation of stochastic simulations to predict the lithology of glacial and fluvial deposits in Central Glasgow, UK. Engineering Geology, 187(17), 98-112.

Kessler, H., S. Mathers, H.-G. Sobisch (2009).The capture and dissemination of integrated 3D geospatial knowledge at the British Geological Survey using GSI3D software and methodology. Computers & Geosciences, 35, 1311–1321.

Koltermann, C. E., and S. M. Gorelick (1996). Heterogeneity in Sedimentary Deposits: A Review of Structure-Imitating, Process-Imitating, and Descriptive Approaches. Water Resour. Res., 32(9), 2617–2658, doi:10.1029/96WR00025.

Konikow, L. F. (2011). The Secret to Successful Solute-Transport Modeling. Groundwater. 49: 144–159. doi: 10.1111/j.1745-6584.2010.00764.x

LaBolle E.M., and G.E. Fogg (2001), Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, Transport Porous Media, 42 (1-2), 155–179.

Lark, R.M., R.S Lawley, A.J.M. Barron, D.T. Aldiss, K. Ambrose, A.H. Cooper, J.R. Lee,and C.N. Waters (2015), Uncertainty in mapped geological boundaries held by a national geological survey: eliciting the geologists' tacit error model, Solid Earth, 6, 727–745.

Le Coz, M., Genthon P., and Adler P.M. (2011). Multiple-Point Statistics for Modeling Facies Heterogeneities in a Porous Medium: The Komadugu-Yobe Alluvium, Lake Chad Basin. Mathematical Geosciences, 43(7), 861-878.

ACCEPTED MANUSCRIPT

Lee S-Y., S.F. Carle, and G.E. Fogg (2007). Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation. Adv. Water Resour., 30, 1914–1932.

Liu, Y., A. Harding, W. Abriel, and S. Strebelle (2004). Multiple-point simulation integrating wells, three-dimensional seismic data, and geology. AAPG Bull., 88(7), 905–921, doi:10.1306/02170403078.

Mallet, J.L. (2002). Geomodeling. Oxford University Press, New York, 599pp.

Mays, D. C., B.A. Faybishenko, and S. Finsterle, Information entropy to measure temporal and spatial complexity of unsaturated flow in heterogeneous media, Water Resour. Res., 38(12), 1313, doi:10.1029/2001WR001185, 2002.

McMillan, A.A., R.J.O. Hamblin, and J.W. Merritt (2005). An overview of the lithostratigraphical framework for the Quaternary and Neogene deposits of Great Britain (onshore). Research Report RR\04\04 British Geological Survey, 38 pp.

Merritt, J.E., A.A. Monaghan, D.C. Entwisle, A.G. Hughes, S.D.G. Campbell, and M.A.E. Browne (2007). 3D attributed models for addressing environmental and engineering geoscience problems in areas of urban regeneration: a case study in Glasgow, UK. First Break, 25, 79–84.

Merrit, J., S. Loughlin, K. Whitbread, R. Terrington (2012). Model metadata report for Central Glasgow GSI3D model. British Geological Survey Internal Report, IR/12/032. 29pp.

Neuman, S. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. Stochastic Environ. Res. Risk Assess., 17(5), 291–305, doi:10.1007/s00477-003-0151-7.

Nury, S.N., X. Zhu, I. Cartwright, and L. Ailleres (2010). Aquifer visualization for sustainable water management. Management of Environmental Quality, 21(2), 253–274.

ACCEPTED MANUSCRIPT

Perrin, M., B. Zhu, J.Rainaud, J., and S. Schneider (2005). Knowledge-driven applications for geological modelling. Journal of Petroleum Science and Engineering, 47, 89–104.

Poeter, E., and D. Anderson (2005), Multimodel ranking and inference in ground water modelling, Ground Water, 43(4), 597–605, doi:10.1111/j.1745-6584.2005.0061.x.

Proce, C.J., R.W. Ritzi, D.F. Dominic and Z. Dai (2004), Modeling multi-scale heterogeneity and aquifer interconnectivity, Ground Water, Vol.42(5), 658-670.

Raiber M., P.A. White, C.J. Daughney, C. Tschritter, P. Davidson, and S.E. Bainbridge (2012). Three-dimensional geological modeling and multi-variate statistical analysis of water chemistry data to analyse and visualize aquifer structure and groundwater composition in the Wairau Plain, Marlborough District, New Zealand. Journal of Hydrology, 436, 13–34.

Ramanathan, R., A. Guin, R. W. Ritzi Jr., D. F. Dominic, V. L. Freedman, T. D. Scheibe, and I. A. Lunt (2010), Simulating the heterogeneity in braided channel belt deposits: 1. A geometric-based methodology and code, Water Resources Research, 46, W04515, doi:10.1029/2009WR008111.

Renard P, de Marsily G (1997) Calculating equivalent permeability: a review. Advances in Water Resources 20, 253–278.

Refsgaard, J.C., S. Christensen, T.O. Sonnenborg, D. Seifert, A.L. Højberg, and L. Troldborg (2012). Review of strategies for handling geological uncertainty in groundwater flow and transport modeling. Advances in Water Resources 36: 36–50.

Ritzi, R.W. (2000). Behavior of indicator variograms and transition probabilities in relation to the variance in lengths of hydrofacies. Water Resour. Res., 36(11), 3375-3381.

ACCEPTED MANUSCRIPT

Robins, N.S., H.K. Rutter, S. Dumpleton, and D.W. Peach. (2004). The role of 3D visualization as an analytical tool preparatory to numerical modelling. Journal of Hydrology, 301 (1–4), 287–295.

Ronayne, M. J., S. M. Gorelick, and C. Zheng (2010). Geological modeling of submeter scale heterogeneity and its influence on tracer transport in a fluvial aquifer. Water Resour. Res., 46, W10519, doi:10.1029/2010WR009348.

Ross, M., M. Parent, and R. Lefebvre (2005). 3D geologic framework models for hydrogeology and land-use management: a case study from a Quaternary basin of southwestern Quebec, Canada. Hydrogeology Journal 13 (5–6), 690–707.

Royse, K.R. (2010). Combining numerical and cognitive 3D modeling approaches in order to determine the structure of the Chalk in the London Basin. Computer & Geosciences, 36, 500 –511.

Sanchez-Vila X, Guadagnini A, Carrera J (2006), Representative hydraulic conductivities in saturated groundwater flow. Reviews of Geophysics 44, RG3002, doi:10.1029/2005RG000169.

Scheibe, T.D. (1993),. Characterization of the spatial structuring of natural porous media and its impacts on subsurface flow and transport. Ph.D. dissertation, Stanford University, Stanford, California.

Scheibe, T.D., and C.J. Murray (1998), A comparison of stochastic simulation techniques for groundwater transport modeling. In Uses of Sedimentologic and Stratigraphic Information in Predicting Reservoir Heterogeneity, ed. G.S. Fraser, J.M. Davis, 107–118. Tulsa, Oklahoma: Special Publication of the SEPM (Society for Sedimentary Geology).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Seifert, D., T. O. Sonnenborg, J. C. Refsgaard, A. L. Hojberg, and L. Troldborg (2012), Assessment of hydrological model predictive ability given multiple conceptual geological models, Water Resour. Res., 48, W06503, doi:10.1029/2011WR011149.

Shannon, E.C.(1948). A mathematical theory of communication. Bell System Technical Journal 27, 379–423.

Singh, V. P. (2011). Hydrologic synthesis using entropy theory: Review, J. Hydrol. Eng., 16(5), 421–433.

Stafleu, J, D. Maljers, J. Hummelman, J.L. Gunnink (2014). Visualization of uncertainty in 3D stochastic voxel models of the Netherlands, EGU General Assembly 2014, held 27 April - 2 May, 2014 in Vienna, Austria.Stone, J.V. (2015). Information Theory: A Tutorial Introduction. Sebtel Press, 260 pp.

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geolology, 34,1–21, doi:10.1023/A:1014009426274

Troldborg, L., J. C. Refsgaard, K. H. Jensen, and P. Engesgaard (2007), The importance of alternative conceptual models for simulation of concentrations in a multi-aquifer system, Hydrogeol. J., 15(5), 843–860.

Tsai, F.T.-C., and A.S. Elshall (2013). Hierarchical Bayesian model averaging for hydrostratigraphic modeling: Uncertainty segregation and comparative evaluation. Water Resources Research, 49(9), 5520–5536.

Turner, A.K. (2006). Challenges and trends for geological modelling and visualization. Bulletin of Engineering Geology and the Environment, 65(2),109–127.

Turner, A.K., and C.W. Gable (2007). A review of geological modeling. Three-dimensional geologic mapping for groundwater applications. Minnesota Geological Survey Open-file Report07-4, pp.75–79.

ACCEPTED MANUSCRIPT

Turner, R.J., M.M. Mansour, R. Dearden, B.É. Ó Dochartaigh, and A.G. Hughes (2014), Improved understanding of groundwater flow in complex superficial deposits using three-dimensional geological-framework and groundwater models: an example from Glasgow, Scotland (UK). Hydrogeology Journal, 23(3), 493-506.

Vassena, C., L. Cattaneo, and M. Giudici (2010). Assessment of the role of facies heterogeneity at the fine scale by numerical transport experiments and connectivity indicators. Hydrogeology Journal, 18(3), 651-668.

Watson, C., J. Richardson, B. Wood, C. Jackson, A.G. Hughes (2015). Improving geological and process model integration through TIN to 3D grid conversion. Computers & Geosciences, 82, 45-54.

Webb, E. K., and M. P. Anderson (1996). Simulation of Preferential Flow in Three-Dimensional, Heterogeneous Conductivity Fields with Realistic Internal Architecture. Water Resour. Res., 32(3), 533–545, doi:10.1029/95WR03399.

Weissmann, G.S., S.F. Carle and G.E. Fogg (1999). Three-dimensional hydrofacies modeling based on soil surveys and transition probability Geostatistics. Water Resour. Res., 35(6), 1761-1770.

Wellmann, J.F. (2013). Information Theory for Correlation Analysis and Estimation of Uncertainty Reduction in Maps and Models. Entropy, 15, 1464-1485.

Wellmann, J.F., and K. Regenauer-Lieb (2012). Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. Tectonophysics, 526-529, 207-216.

Wen, X-H., and J. Gómez-Hernández (1996). Upscaling hydraulic conductivities in heterogeneous media: An overview. Journal of Hydrology 183 (1–2), ix-xxxii.

ACCEPTED MANUSCRIPT

Williams, J.D.O., and M.R. Dobbs (2012). Particle size distribution analysis of central Glasgow in relation to lithology classification used in stochastic modelling. British Geological Survey, Internal Report, IR/12/039, 28 pp.

Wilcox, A.R. (1967). Indices of qualitative variation. ORNL-TM-1919, Oak Ridge National Laboratory, Oak Ridge, Tennessee (USA).

Wycisk, P., T. Hubert, W. Gossel, and C. Neumann (2009). High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites. Computer & Geosciences, 35, 165–182

Xu, C., and P.A. Dowd (2003). Optimal construction and visualization of geological structures. Computers & Geosciences, 29, 761–773.

Ye, M., and R. Khaleel (2008), A Markov chain model for characterizing medium heterogeneity and sediment layering structure, Water Resour. Res., 44, W09427, doi:10.1029/2008WR006924.

Zanchi, A., M. De Donatis, A. Gibbs, J.-L. Mallet (2009). Imaging geology in 3D. Computers & Geosciences, 35(1), 1-3.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

List of Tables

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

List of Figures

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

white dashed lines. Correspondent normalized entropy distributions for model M1 (e, g) and model M2 (f, h). Black dots indicate borehole data locations. Light grey areas indicate where the bedrock is at the surface.

Figure 7. Histograms of normalized entropy values for models M1 and M2. (a) lithofacies predictions; (b) groundwater head predictions; (c) flux predictions.

Figure 8. Occurrence probability maps for the different lithofacies calculated with model M1 (a, c, e, g) and model M2 (b, d, f, h). Black dots indicate borehole data locations. Light grey areas indicate where the bedrock is at the surface.

Figure 9. Validation test results for models M1 (a) and M2 (b). Red lines indicate the average percentage on correct predictions.

Figure 10. Two-dimensional normalized entropy distributions for groundwater flow model predictions over the top surface of the 3-D grid. The black line indicates the River Clyde. (a) Groundwater head predictions based on model M1. (b) Groundwater head predictions based on model M2. (c) Groundwater flux predictions based on model M1. (d) Groundwater flux predictions based on model M2.

Figure 11. Spatial distributions of variations in normalized entropy between models M1 and M2 for lithofacies predictions (a), groundwater head predictions (b), and flux predictions (c). The black line indicates the River Clyde. Black dots indicate the hard conditioning data.

ACCEPTED MANUSCRIPT

Figure 12. Scatter plots comparing variations in normalized entropy for lithofacies predictions with correspondent variations for groundwater head predictions (a) and flux predictions (b). Red dashed lines indicate linear least squares models fitted to the data points. The equation of the line in (a) is $y = 0.078 + (0.082 \pm 0.005)x$. The equation of the line in (b) is $y = 0.020 + (0.305 \pm 0.003)x$.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

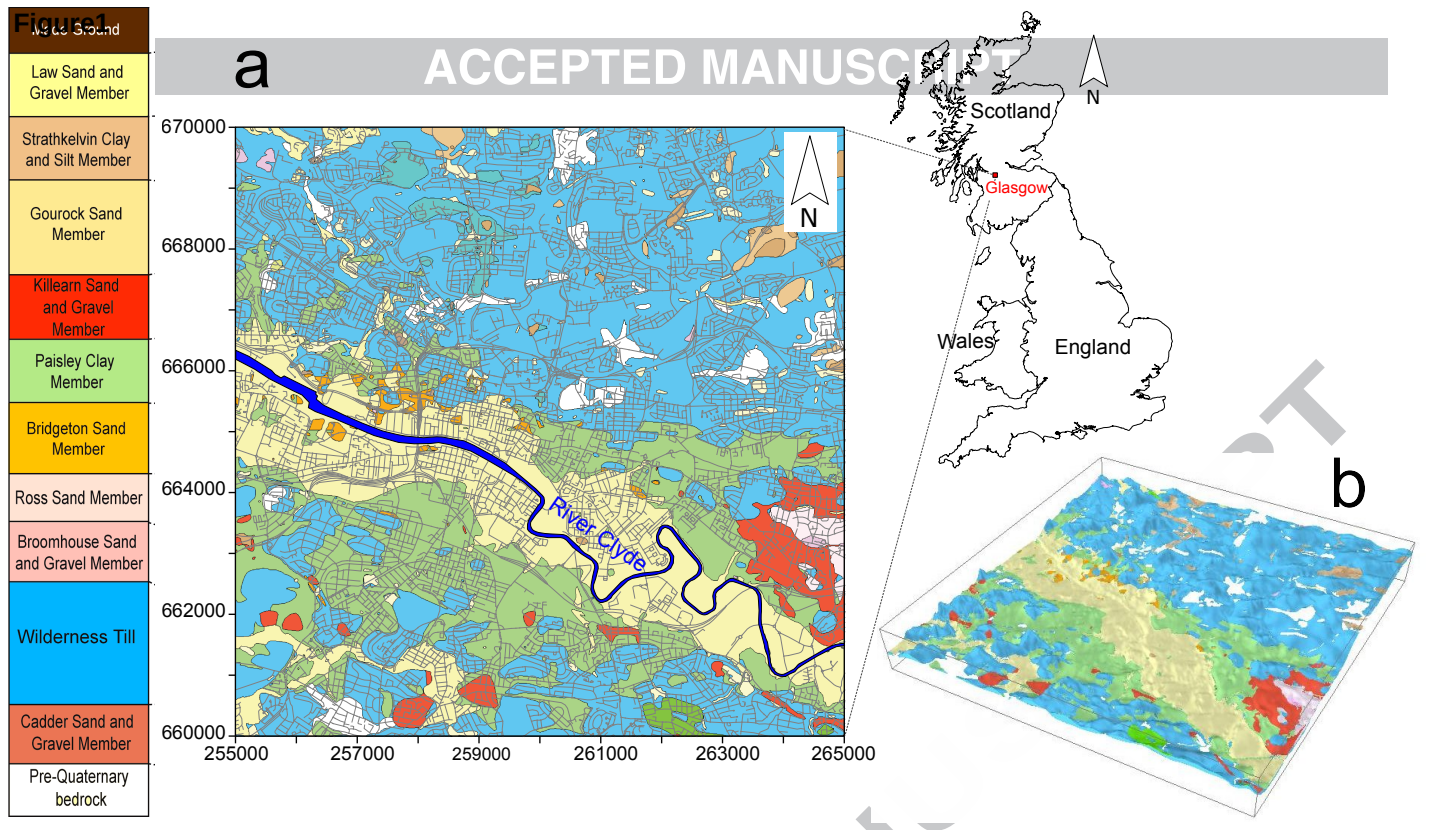
Lithostratigraphic unit	s ftC	s tfCD	S Z	S G
Law Sand and Gravel Member	0.70	0.07	0.00	0.23
Lacustrine	0.72	0.16	0.11	0.01
Gourock Sand Member	0.35	0.04	0.34	0.27
Paisley Clay Member	0.62	0.04	0.30	0.04
Killearn Sand and Gravel Member	0.19	0.10	0.62	0.08
Bridgeton Sand member	0.10	0.05	0.76	0.09
Ross Sand Member	0.08	0.04	0.81	0.07
Broomhouse Sand and Gravel Member	0.08	0.09	0.25	0.59
Wilderness Till Formation	0.15	0.65	0.14	0.06
Cadder Sand and Gravel Formation	0.00	0.56	0.00	0.44

ACCEPTED MANUSCRIPT

Parameter	sftC		stCD		SZ		SG	
	1	2	1	2	1	2	1	2
Vol. fraction	.31	.29	.31	.35	.27	.26	.11	.10
Mean length (m)	20	18	10	94	0	3	00	7
Mean thickness (m)	.7	.7	.9	.9	.6	.6	.1	.1

ACCEPTED MANUSCRIPT

Figure1



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

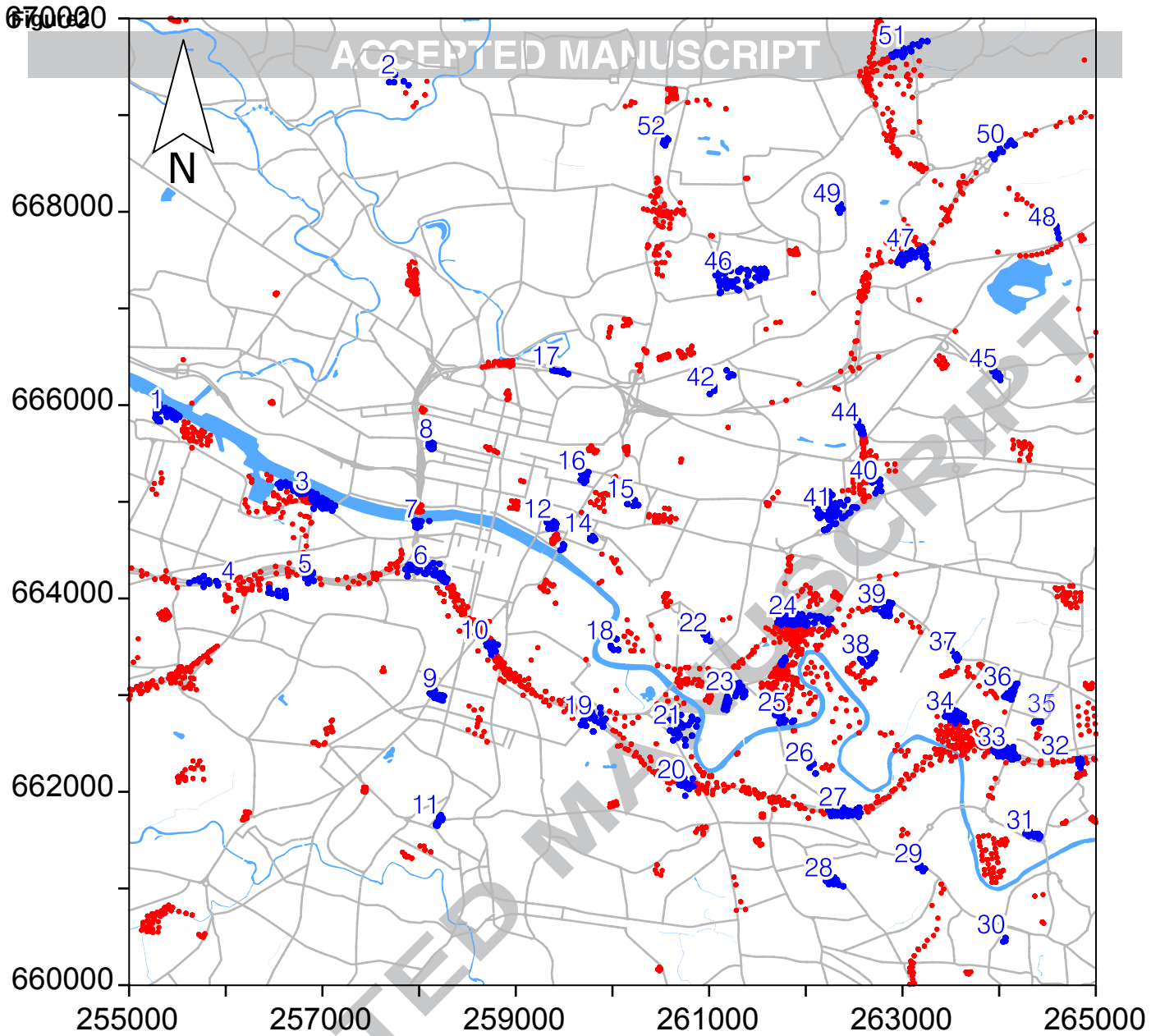


Figure3

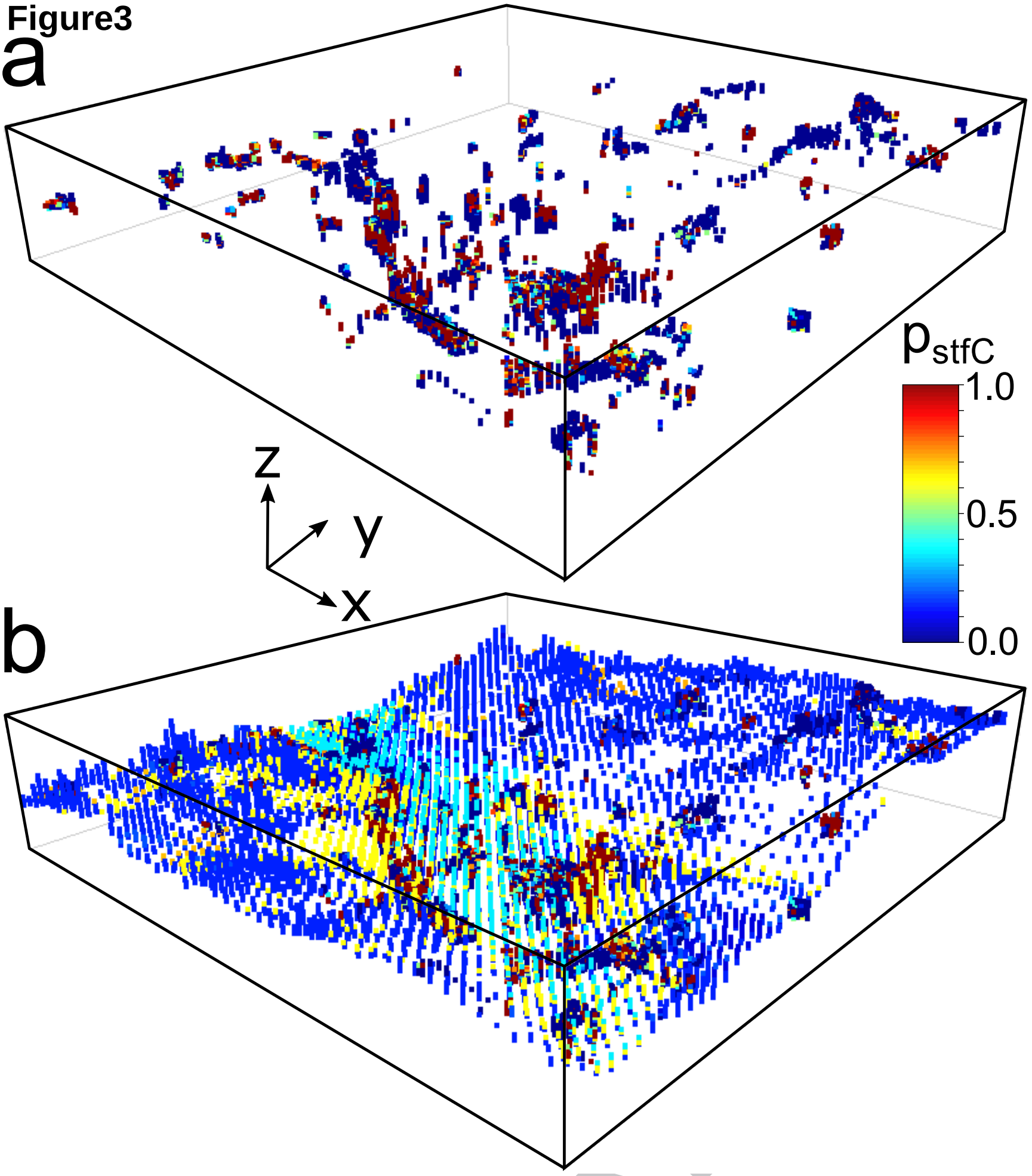
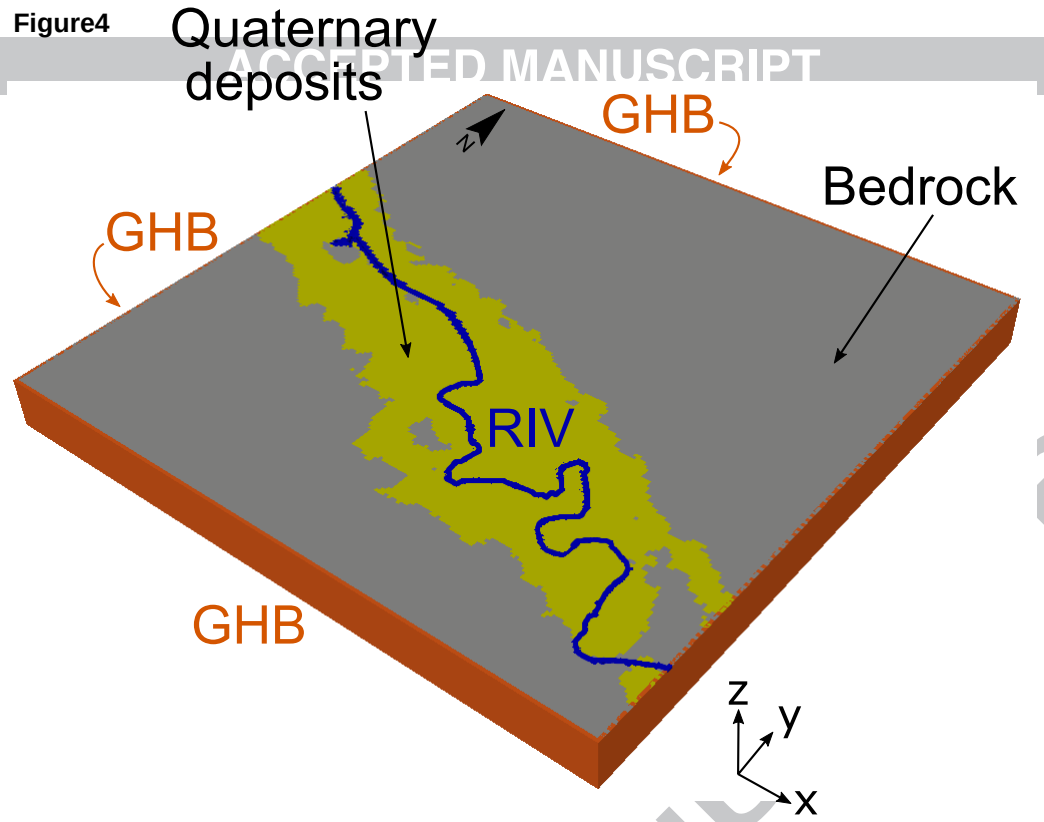


Figure4

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

Figure5

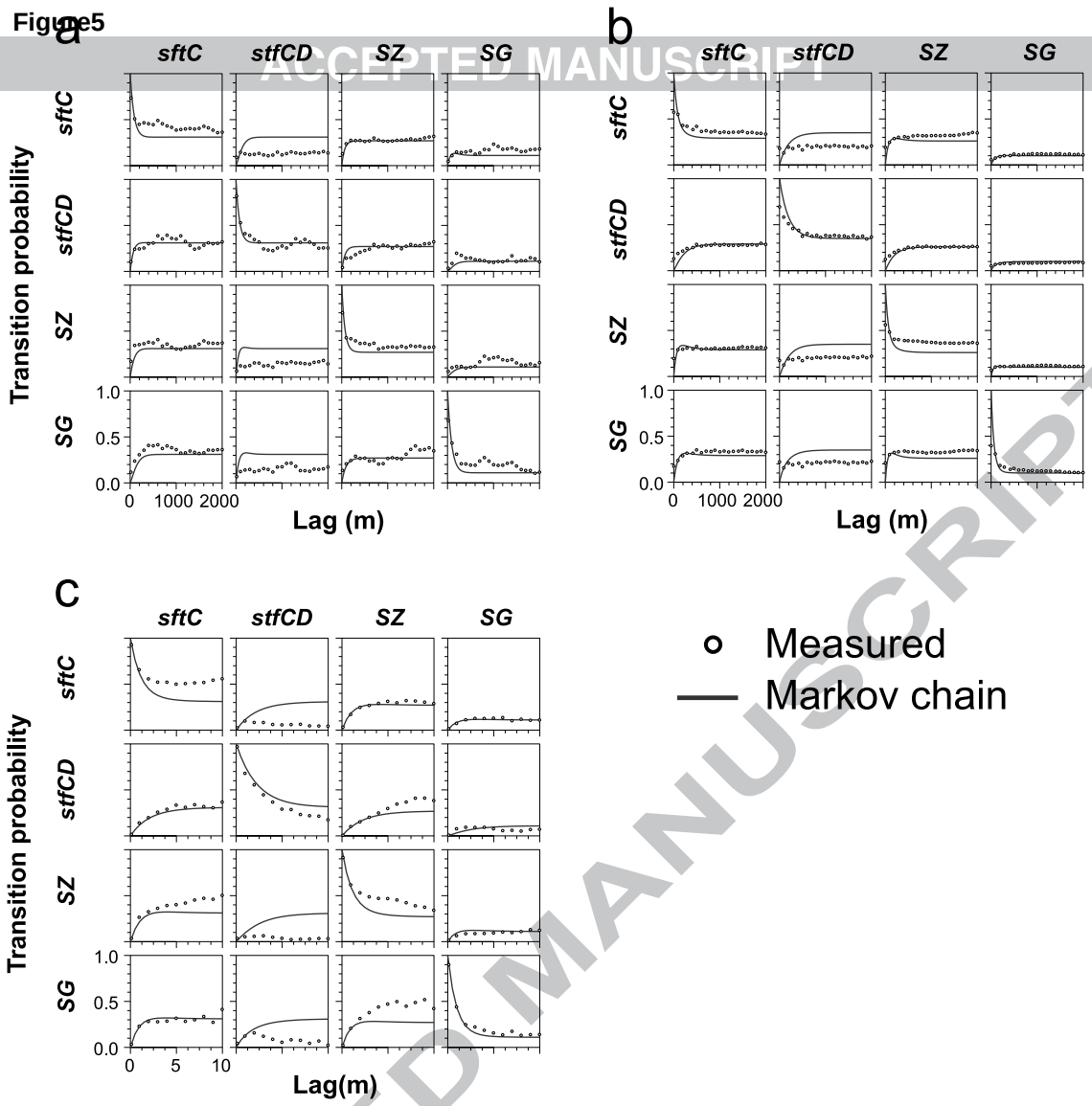


Figure6

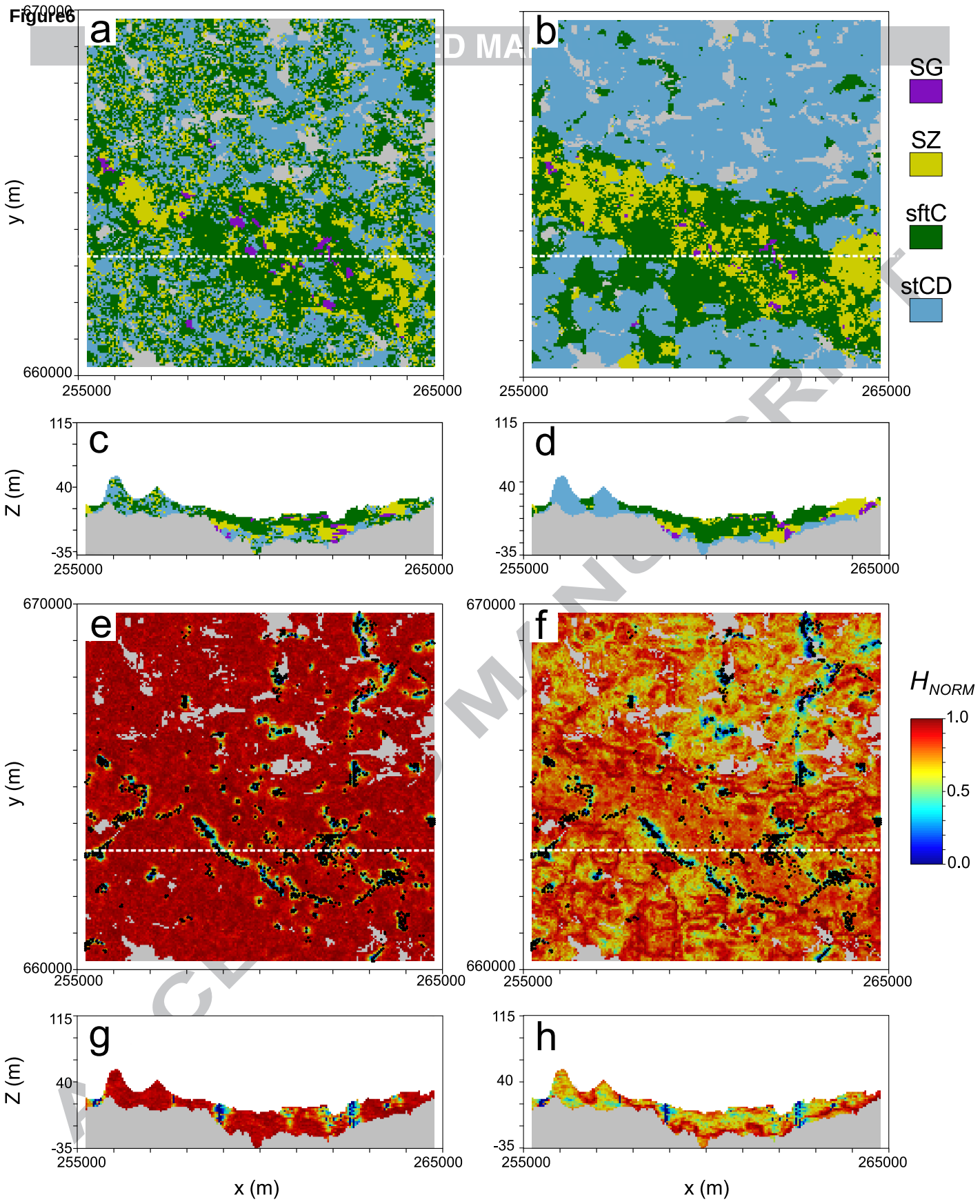
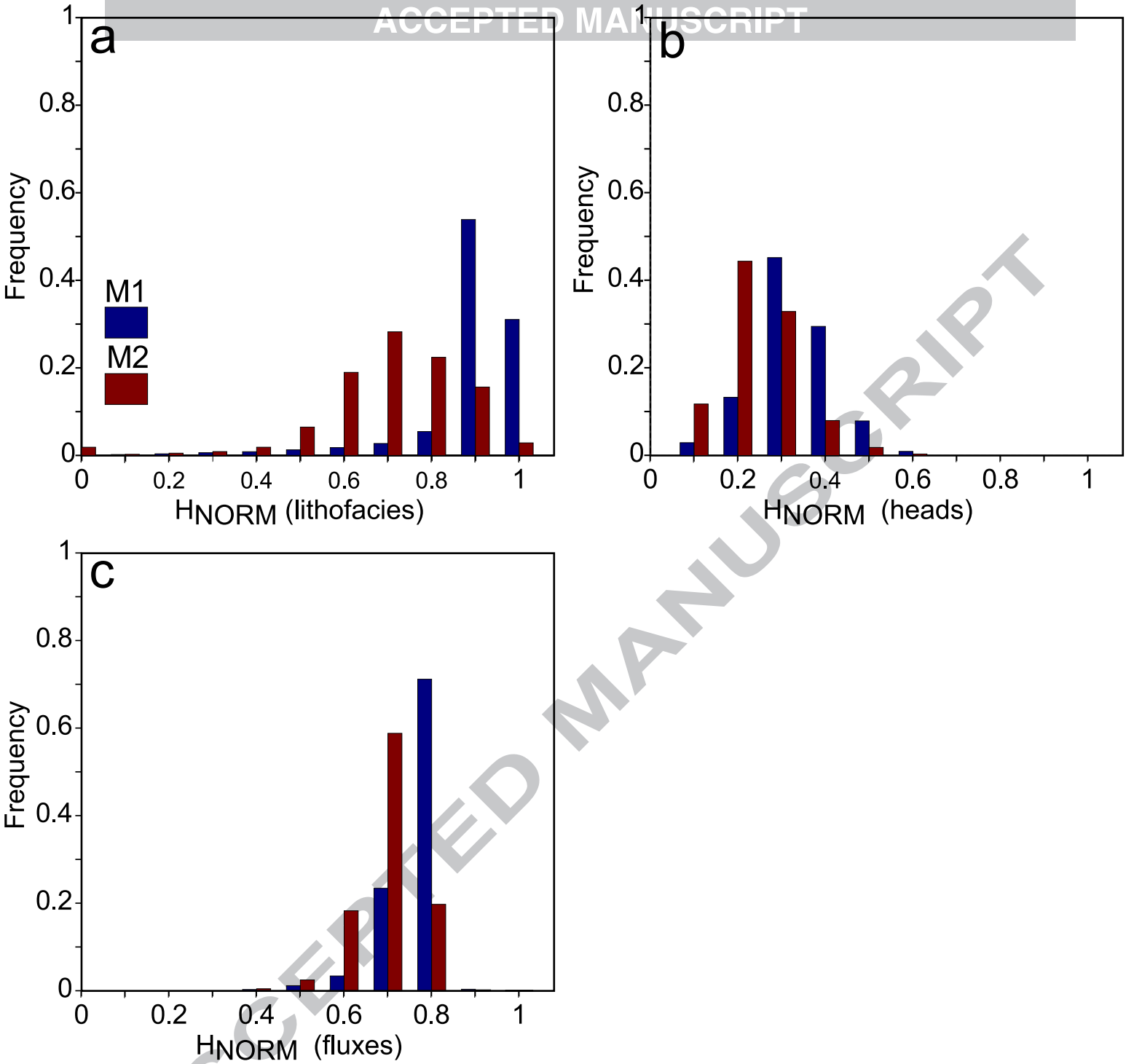

a
b
c
d
e
f
g
h
SG
SZ
sftC
stCD
H_{NORM}
1.0
0.5
0.0
y (m)
Z (m)
x (m)
670000
660000
255000
265000
115
40
-35

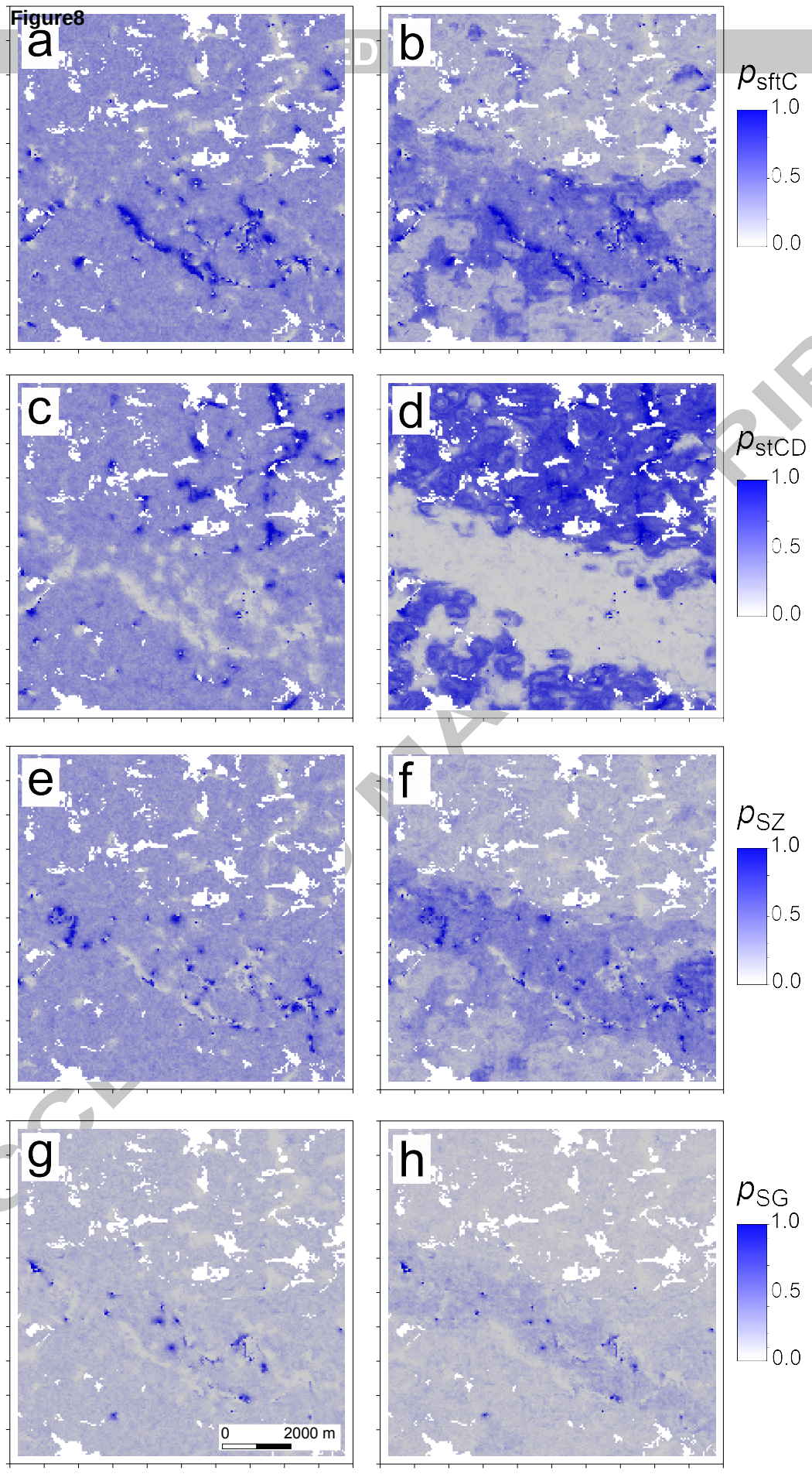
Figure7

ACCEPTED MANUSCRIPT

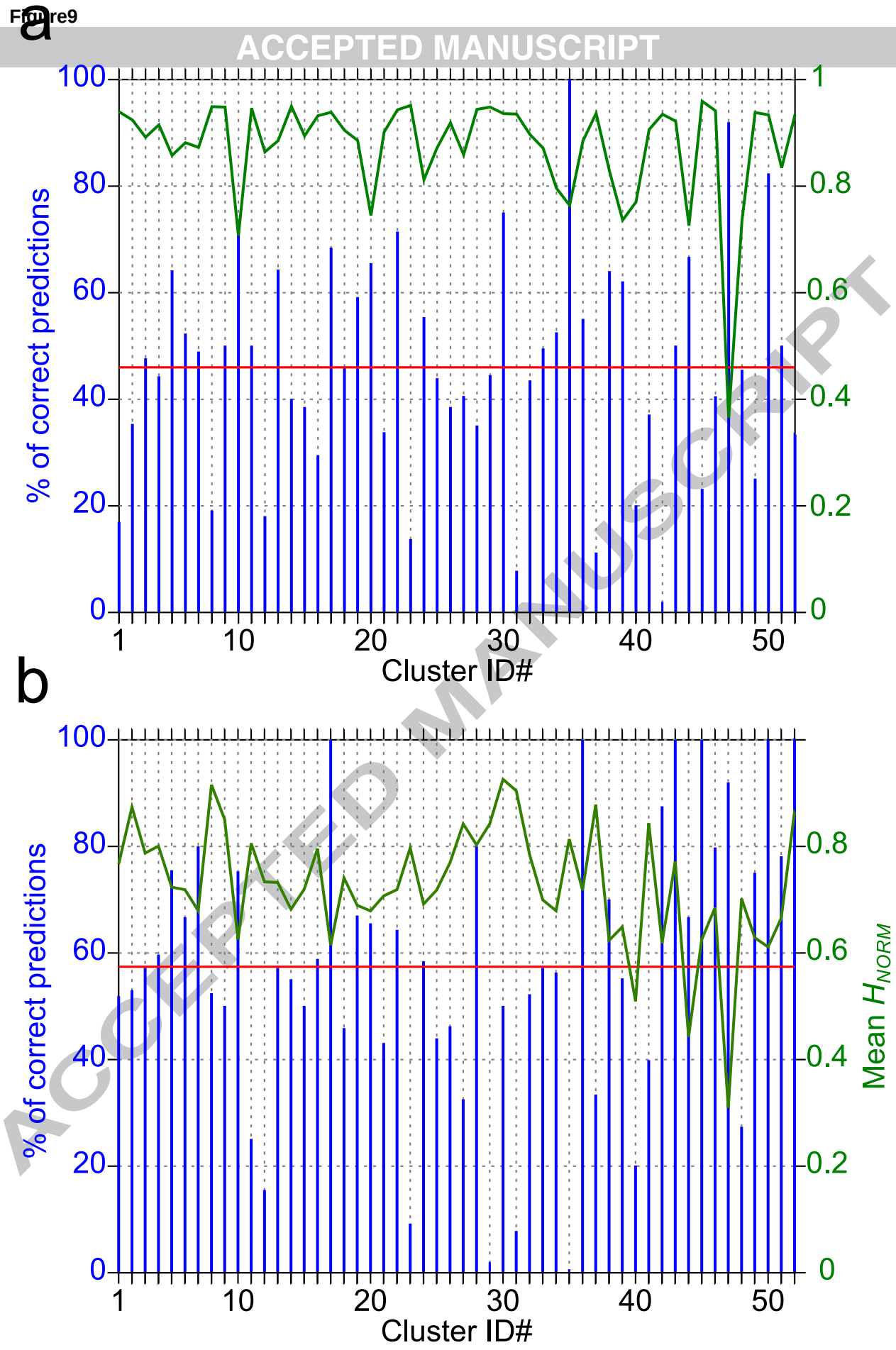


ACCEPTED MANUSCRIPT

Figure8


a
b
p_{sftC}
1.0
0.5
0.0
c
d
p_{stCD}
1.0
0.5
0.0
e
f
p_{SZ}
1.0
0.5
0.0
g
h
p_{SG}
1.0
0.5
0.0
0
2000 m

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

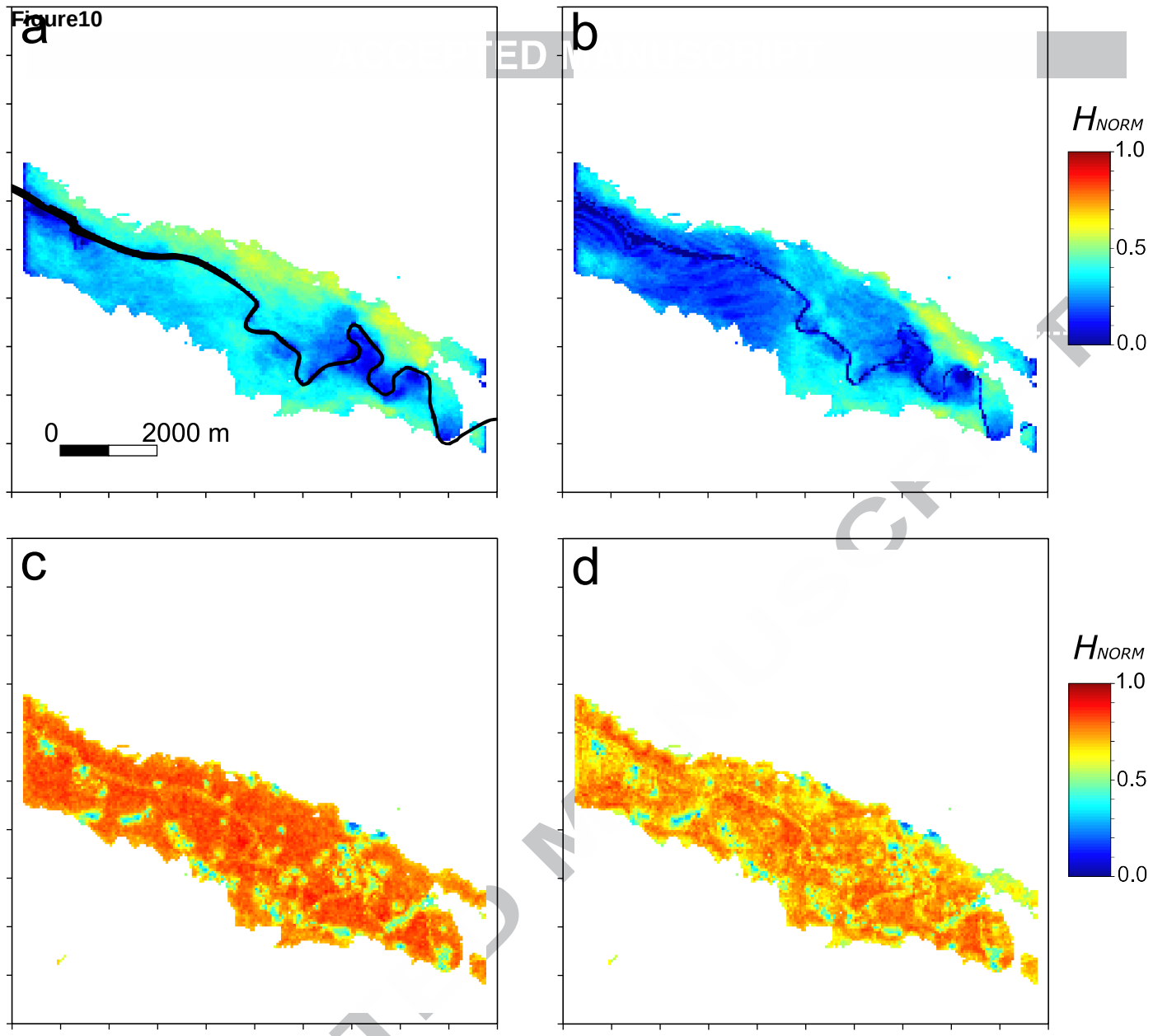
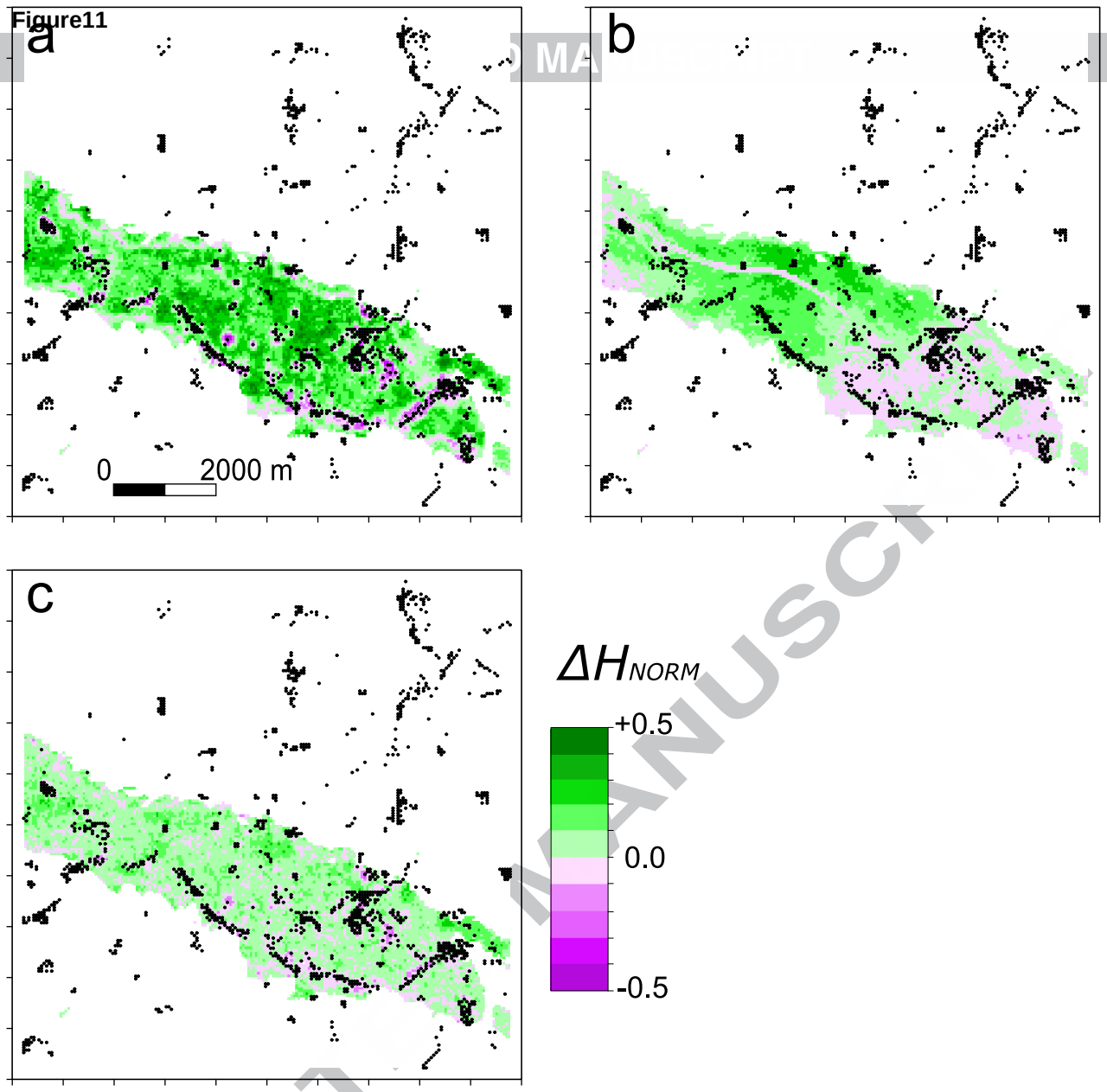
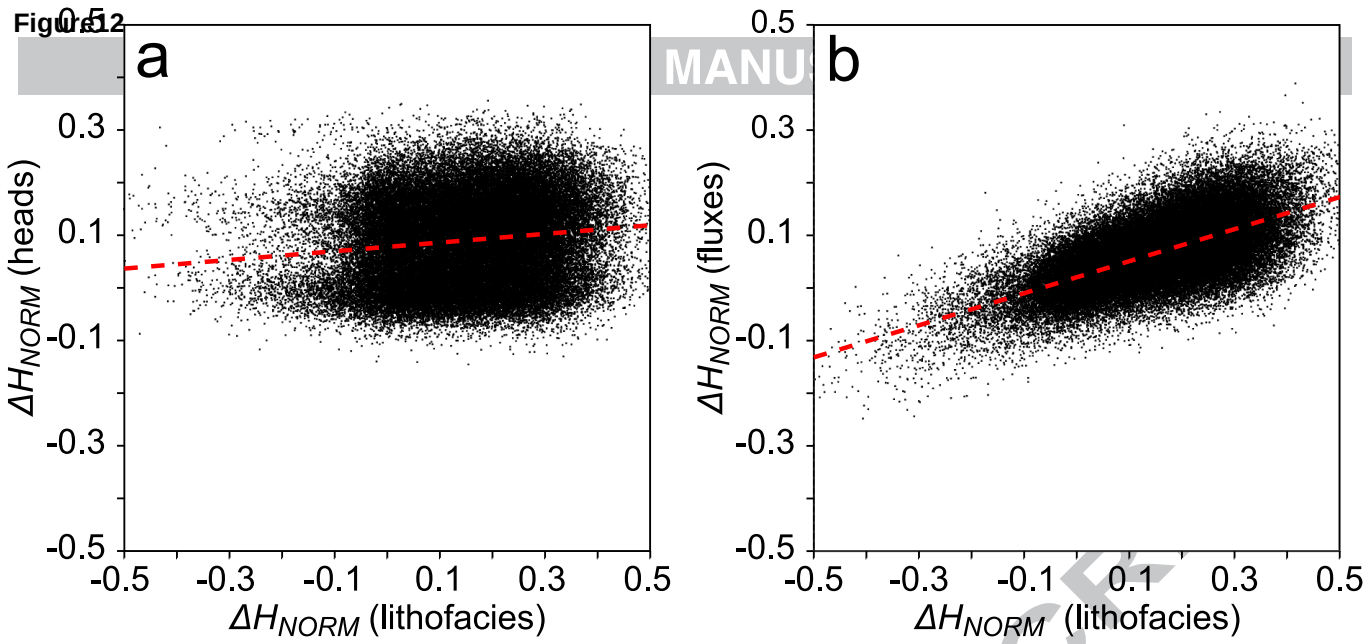


Figure10
a
b
c
d
0
2000 m
H_{NORM}
1.0
0.5
0.0
H_{NORM}
1.0
0.5
0.0

Figure11




a
b
0.5
0.3
0.1
-0.1
-0.3
-0.5
ΔH_{NORM} (heads)
ΔH_{NORM} (fluxes)
ΔH_{NORM} (lithofacies)
ΔH_{NORM} (lithofacies)

ACCEPTED MANUSCRIPT

Highlights

Novel approach to account for geological structure in stochastic models of subsurface heterogeneity**.**

Approach tested in a highly heterogeneous environment.

Accounting for geological structure reduces model uncertainty, here quantified with a metric based on Shannon information entropy.

Correlations found between prediction uncertainties for lithofacies, hydraulic heads and groundwater fluxes.

Insights regarding the impact of geological information in fluid flow and solute transport models.

ACCEPTED MANUSCRIPT