Stochastic investigation of long-term persistence in two-dimensional images of rocks

Panayiotis Dimitriadis[1*], Katerina Tzouka[1], Demetris Koutsoyiannis[1], Hristos Tyralis[1] Anna Kalamioti[1], Eleutherios Lerias[1] and Panagiotis Voudouris[2]

[1] Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Heroon Polytechniou 5, 157 80 Zografou, Greece

[2] Department of Mineralogy and Petrology, Faculty of Geology & Geoenvironment, National and Kapodistrian University of Athens, Panepistimioupolis-Ano Ilisia, 157 84 Zografou, Greece

[*] Corresponding author, pandim@itia.ntua.gr

Abstract: Determining the geophysical properties of rocks and geological formations is of high importance in many fields such as geotechnical engineering. In this study, we investigate the second-order dependence structure of spatial (two-dimensional) processes through the statistical perspective of variance vs. scale (else known as the climacogram) instead of covariance vs. lag (e.g. autocovariance, variogram etc.) or power vs. frequency (e.g. power spectrum, scaleogram, wavelet transform etc.) which traditionally are applied. In particular, we implement a two-dimensional (visual) estimator, adjusted for bias and for unknown process mean, through the (plot of) variance of the space-averaged process vs. the spatial scale. Additionally, we attempt to link the climacogram to the type of rocks and provide evidence on stochastic similarities in certain of their characteristics, such as mineralogical composition and resolution. To this end, we investigate two-dimensional spatial images of rocks in terms of their stochastic microstructure as estimated by the climacogram. The analysis is based both on microscale and macroscale data extracted from grayscale images of rocks. Interestingly, a power-law drop of variance vs. scale (or else known as long-term persistence) is detected in all scales presenting a similar power-exponent. Furthermore, the strengths and limitations of the climacogram as a stochastic tool are discussed and compared with the traditional tool in spatial statistics, the variogram. We show that the former has considerable strengths for detecting the long-range dependence in spatial statistics.

Keywords: rock image analysis; climacogram; variogram; Hurst-Kolmogorov behaviour; stochastic modelling.

1 Introduction

Extracting information from image analysis is very important in many fields of science. Diagnostic images are used for stochastic analysis of diseases in the field of medicine, e.g. X-ray images are used in bone disease (Jennane, 2001) or Magnetic Resonance Image (MRI) for better investigation and diagnosis of brain diseases (Vanitha, 2016). In geophysics, radar images are useful for the statistical analysis of geological structures, as for example to study the evolution of faults systems (Gloaguen, 2007), while air photos are examined for the investigation of the link between fault structure and earthquake rupture behaviour (Milliner et al., 2016). In hydrology and fluid mechanics the multi-scale investigation of an attribute, the inference of its statistical properties and its reconstruction through image processing, have been reported in many studies, e.g. in the reconstruction of a porous media from morphological information using 2D images of their microstructure (Talukdar et al. 2002), in the modelling of the pore space of rocks through three-dimensional micro-tomography images (Blunt et al., 2012; Rabbani et al., 2016), in the modelling of shale rock in multiple scales (Gerke et al., 2015) and others.

The typical tool that is used for the stochastic analysis of geostatistical fields is the variogram which is defined to be the half of the variance of the field difference at two points, as a function of the distance between these points. A comprehensive presentation of the variogram in geostatistics can be found in Chilès and Delfiner (2012). One particular issue of high importance is the detection of some scaling laws in 2d images of rocks that however, cannot be easily identified by the variogram. In section 3, we attempt to highlight the advantages of the climacogram for detecting scaling laws within spatial scales, such as the Long-Term-Persistence (LTP) behaviour or else known as Long-Term Change or Hurst-Kolmogorov behaviour (Hurst, 1951; Kolmogorov, 1941; Koutsoyiannis, 2002; 2016), where the autocovariance (or climacogram) of the stationary process decays as a power law function of lag (or scale). This is quite different from an exponential function of lag corresponding to the more well-known short-term persistence, or else Markov behaviour. Note that similar analyses have been applied in porous medium for the identification of the LTP behaviour but using the (auto)covariance or variogram (e.g. Hamzehpour et al., 2007) instead of the climacogram.

Here, we develop our model based on the climacogram at different spatial scales to detect such behaviours and to combine all scales to a single model (Stein et al., 2001). From this analysis, the identified LTP behaviour in the various examined rocks can perhaps explain a part of the large uncertainty intensively detected in the geological structures and soil formations (Heuvelink and Webster, 2001) and thus, help towards a better understanding of the related processes and the construction of corresponding prediction and generation algorithms. The uncertainty of a natural process can be quantified by, for example, its variability through second-order statistics, and it is highly correlated to the temporal or spatial window under which the prediction is being made for specified statistical error and confidence level. Within this window the process can be considered predictable and outside of this unpredictable in the sense that we can predict the process' confidence limits and expected value with the specified error. Evidently, the uncertainty or, equivalently, variability in natural processes depends on the length of predictability window for various time scales (Dimitriadis et al., 2016a) or of spatial scales as in this study. Naturally, as the prediction error increases, so will the length measured in time or space units of the predictability window.

The climacogram, that we use to investigate the stochastic properties of two-dimensional (2d) images of rock samples, has been extensively used in one-dimensional (1d) stochastic processes (for a review see, e.g., Koutsoyiannis, 2010; O'Connell et al., 2016; Dimitriadis, 2017), and in other 2d processes (Koutsoyiannis et al., 2011; Dimitriadis et al., 2013). It is defined to be the (plot of) variance of the space-averaged process vs. the spatial scale (Koutsoyiannis, 2016). It can provide a powerful option for process identification and estimation, alternative to more classical methods such as the method of moments, Bayesian methods, maximum likelihood and graphical methods (Elogne et al., 2008). Also, the climacogram can serve as an alternative way of viewing a natural process through the concept of scale as opposed to the more traditional ones, i.e., those of lag through the autocovariance and frequency through the power-spectrum. In fact, although the climacogram is mathematically equivalent to the aforementioned estimators of the second-order dependence structure, it exhibits smaller statistical uncertainty, and an easier way to handle the statistical bias and to generate synthetic timeseries (Dimitriadis and Koutsoyiannis, 2015). Recently (Dimitriadis and Koutsoyiannis, 2018), it has been implemented to higher-order structures exhibiting similar advantages as in the lower ones.

In our applications, we examine several images of rocks extracted from a Scanning Electron Microscope (SEM), from a polarising microscope and from field samples. Also, we compare the use of the climacogram for the LTP identification to that of the variogram through benchmark examples. Finally, we discuss the influence of the scale length and type of rock on the statistical estimation and we propose a stochastic process that adequately preserves the observed LTP behaviour in the examined 2d images of rocks.

2 Data

We use characteristic images of rocks in different scales obtained through open internet data bases, which are shown in Figs. 1 to 3 along with their source information. The coloured 8-bit images are first converted to grayscale shade (Fig. 4), with the black colour corresponding to zero intensity (minimum) and the white colour to one (maximum). Therefore, a number from zero to one is assigned to every single pixel of the image. In this way, we can measure the colour difference between pixels and use it as a rough estimation for the distinction of various groups of minerals, appearing with different colour intensity, that the rock is comprised of. For convenience, we use the upper left pixel of each image as the zero-initial point of the field in the Cartesian system. Also, all pictures have quite similar resolution (see Table 1) to enable a direct comparison of their stochastic properties and to avoid any errors introduced by the different content information (e.g. Gommes et al., 2012).

In order to examine the stochastic behaviour of the rock samples through climacogram we select samples of rocks at spatial resolution of μm, mm, cm and m and we analyze them based on the following samples:

a) Sample images from different rocks but in the same rock category are selected. In Fig. 1, we depict an image of limestone and one of sandstone. Both limestone and sandstone belong to the category of sedimentary rocks. Limestone is composed mainly of one mineral (calcite) while sandstone is composed of multiple minerals (e.g. quartz, feldspar, kaolinite, muscovite).

b) For a second application, we select images of rocks with similar mineral composition but from a different rock category. For example, in Fig. 1, we analyze a sample image from marble, which is a metamorphic rock consisting predominantly of calcite or dolomite and is formed when a sedimentary carbonate rock, such as limestone ($CaCO_3$) or dolomite $(Ca,Mg)(CO_3)_2$, is metamorphosed by natural rock-forming processes, so that the grains are recrystallized. Additionally, we analyze a sample image from limestone, which is a sedimentary rock composed of calcite ($CaCO_3$) that is converted to marble by the recrystallization of the calcite having the same mineralogical and chemical composition with marble.

c) Moreover, we select sample images from an igneous rock, i.e. rhyolite at moderate (mm) and meso (cm) scales (Fig. 2).

d) Finally, we compare sample images of a sandstone rock in four different scales (Fig. 3). Particularly, we compare images of sandstones in microscale (μm) using an image from the Scanning Electron Microscope (SEM), in moderate scale (mm) using an image from the polarizing microscope, in mesoscale using an image from a hand specimen (cm) and in macroscale using a field outcrop (m).

In Table 1, we estimate the marginal statistics of all sample images (Figs. 1 to 3) such as mean, standard deviation, and the coefficients of skewness and kurtosis, from which we can conclude that there is only a mild deviation from normality of the spatial data and therefore, no action of normalization is required. Note that a strong deviation from normality could impair the variogram structure (Varouchakis et al., 2016 and references therein).

Figure 1: Images (from left to right) of limestone, marble and sandstone, with dimensions between five to ten centimetres across.
Source: www.geo.auth.gr/106/theory/pet_sed_limestone_01.jpg
www.geo.auth.gr/courses/gmo/gmo106y_lab/photo/metamorphic/marble_2.jpg
http://geology.com/rocks/pictures/sandstone.jpg.

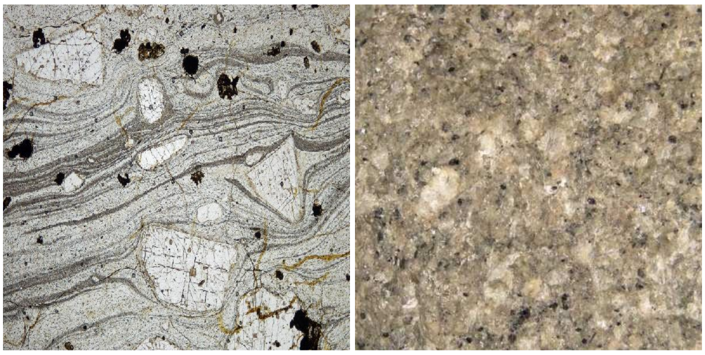

Figure 2: Images of rhyolite as seen from a thin section under polarizing microscope with 6 mm length (left) and from a hand specimen with 3 cm length (right).
Source: www.geo.auth.gr/317/photos_macro.htm
www.earth.ox.ac.uk/~oesis/micro/

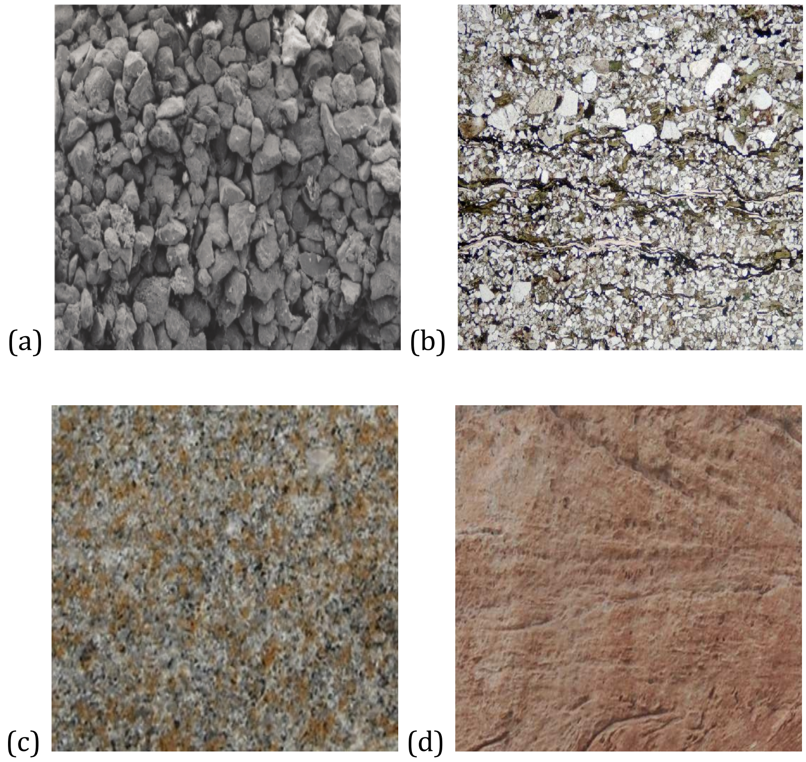

Figure 3: Images of sandstone as seen from (a) the SEM (50 μm), (b) a polarizing microscope, (3.5 mm), (c) a hand specimen (with length approximately 5 cm) and (d) a field outcrop (1 m). Source: http://sandia-exploration.com/high_porosity_photos.html www.earth.ox.ac.uk/~oesis/micro/ http://blogs.cedarville.edu/christian-geology/2015/02/two-new-papers-on-the-coconino-sandstone

3 Methodology

3.1 Climacogram

Assuming that $\underline{x}(\xi_1, \xi_2)$ is a 2d spatial stochastic process or field, the climacogram, as introduced by Koutsoyiannis (2010) for a one-dimensional process and expanded by Koutsoyiannis et al. (2011), is defined as the variance at a rectangular area $k_1 \times k_2$, i.e. (Dimitriadis et al., 2013):

$$\gamma(k_1, k_2) := \frac{Var\left[\int_0^{k_2}\int_0^{k_1} \underline{x}(\xi_1, \xi_2) d\xi_1\, d\xi_2\right]}{k^4} \tag{1}$$

where the underline is used to distinguish a random variable from a regular one, $k \coloneqq \sqrt{k_1 k_2}$ is the geometric mean of the continuous spatial scales k_1, k_2, each with dimensions of length, and Var[] denotes variance.

The climacogram is shown to have smaller statistical bias and variability (i.e. smaller standardized mean-square-error), zero discretization error as well as other properties more useful in stochastic model identification, building and generation than other stochastic tools such as (auto)covariance (or correlation) and power spectrum (Dimitriadis and Koutsoyiannis, 2015). As explained by Koutsoyiannis (2016) for 1d processes and Dimitriadis et al. (2013) for higher d dimensional processes, the d^{th} covariance is related to the $2d^{\text{th}}$ derivative of the d^{th} climacogram and since estimation of derivatives from data is too uncertain it makes a very rough graph. In addition, its estimation is highly biased compared to the climacogram, as explained in Koutsoyiannis (2003), where the expectation of the latter, i.e. E[γ], is much closer to its true value γ for large lags and LTP processes,. Furthermore, discretization (i.e. block averaging) of a process affects the covariance, which is different from that of the original process. The climacogram however is the same in both cases, and therefore, remains unaffected from the nugget effect. In practice discontinuities/jumps at scale zero can be avoided if a proper model for the climacogram is constructed and, hence, regularization becomes unnecessary, as opposed to the case of modelling based on the covariance; e.g. Chiles and Delfiner, (2012, ch. 2.4).

Assuming that our sample is an area $n\Delta \times n\Delta$, where n is the number of intervals (e.g. pixels) along each spatial direction and Δ is the discretization unit (determined by the image resolution, e.g. pixel length), the empirical classical estimator of the climacogram for a 2d process can be expressed as:

$$\underline{\hat{\gamma}}(\kappa_1, \kappa_2) = \frac{1}{n^2/\kappa^2 - 1} \sum_{i=1}^{n/\kappa_1} \sum_{j=1}^{n/\kappa_2} \left(\underline{x}_{i,j}^{(\kappa)} - \underline{\overline{x}}\right)^2 \tag{2}$$

where the "^" over $\underline{\gamma}$ denotes estimation, $\kappa \coloneqq \sqrt{\kappa_1 \kappa_2}$ is the geometric mean of the discrete scales κ_1, κ_2, with $\kappa_1 = k_1/\Delta$ and $\kappa_2 = k_2/\Delta$ the dimensionless spatial scales, $\underline{x}_{i,j}^{(\kappa)} = \frac{1}{\kappa^2} \sum_{\psi=\kappa_1(j-1)+1}^{\kappa_1 j} \sum_{\xi=\kappa_2(i-1)+1}^{\kappa_2 i} \underline{x}_{\xi,\psi}$ is the sample average of the space-averaged process at scale κ, and $\underline{\overline{x}} = \sum_{i,j=1}^{n} \underline{x}_{i,j} / n^2$ is the sample average. Note that the maximum available scale for this estimator is $n/2$.

A variety of processes exhibit LTP behaviour (e.g. Dimitriadis, 2017), the simplest one being the isotropic Hurst-Kolmogorov (HK) process, i.e. power-law decay of variance as a function of scale, and defined for a 1d or 2d process as:

$$\gamma(k) = \frac{\lambda}{(k/\Delta)^{2d(1-H)}} \tag{3}$$

where λ is the variance at scale $k = \kappa\Delta$ ($k_1 = k_2 = \kappa\Delta$), d is the dimension of the process/field (i.e., for a 1d process d = 1, for a 2d field d = 2, etc.), and H is the Hurst parameter (0 < H < 1).

The HK behaviour can be easily identified through the log-log slope (e.g. Dimitriadis et al., 2016b) $\gamma^{\#}(k) \coloneqq \mathrm{d}(\log\gamma(k))/\mathrm{d}(\log k)$ of the climacogram at large scales k, which is also linked to the Hurst parameter by $\lim_{k\to\infty} \gamma^{\#}(k) = 2d(H-1)$. Particularly, the HK behaviour corresponds to a slope milder than $-d$, where equality, i.e. $\lim_{k\to\infty} \gamma^{\#}(k) = -d$, indicates a Markov or a white noise process, (the proof for a 1d field can be seen Dimitriadis and Koutsoyiannis, 2015), and can be similarly expanded to an isotropic field of any dimension. In other words, if the slope is smaller (milder) than $-d$ then the physical process is more likely to behave as a positively correlated process (or else persistent), whereas for slopes steeper than $-d$ as an anti-correlated process (or else anti-persistent). For example, in Fig. 4, an example of a positively correlated 2d process is depicted and compared to a white noise process and to an anti-correlated one (for $H \to 0$).

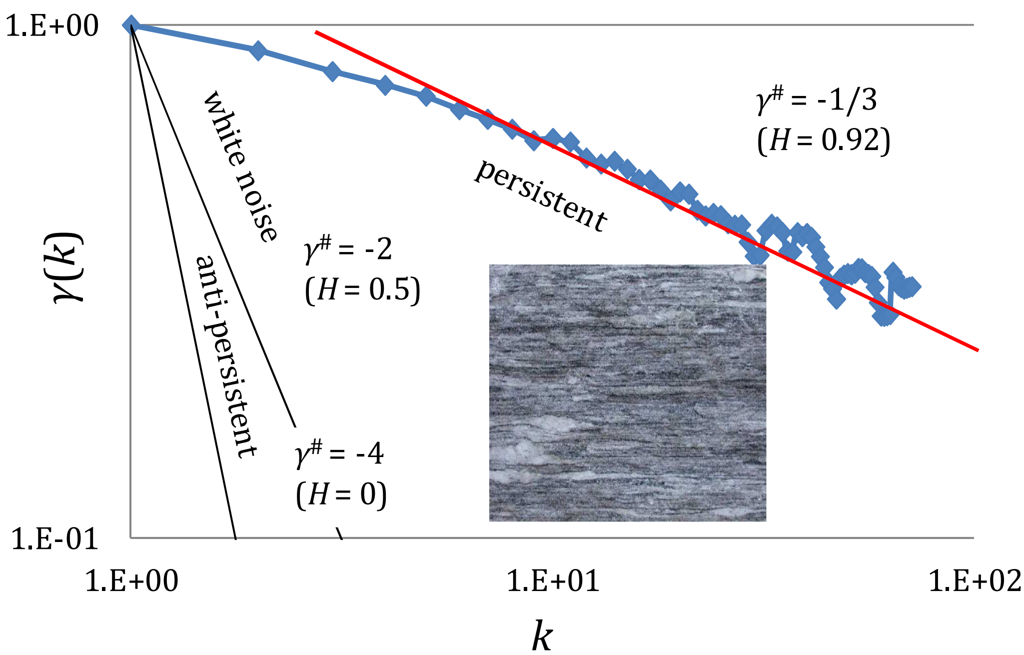

Figure 4: Climacograms of a gneiss shown in grayscale (2d HK process with H = 0.92), a white-noise process (H = 0.5) and the lower limit of an anti-persistent processes ($H \to 0$).
Source of image within the figure: http://www.geo.auth.gr/106/theory/pet_met_gneiss_01.jpg

An important remark is that our analysis depends only on the investigation of the second-order statistics (i.e. variance of the averaged process vs. scale with an unknown mean of the process) and therefore, since it is generic, it can be applied to any type of marginal distribution. For example, let us consider the isotropic d-dimensional fractional-Gaussian-noise process (fGn) based on scale, i.e. $(\underline{x}^{(k)} - \mu) =_{\mathrm{d}} (k/l)^{d(H-1)}(\underline{x}^{(l)} - \mu)$, where $=_{\mathrm{d}}$ denotes equality in distribution, μ is the mean of the process and l, k are the d-dimensional scales defined through their geometric mean, i.e. $k = (k_1 k_2 \dots k_d)^{1/d}$ and similarly for l (Dimitriadis et al., 2013; for the 1d and 2d cases see also Mandelbrot and Van Ness, 1968; Qian et al., 1998; Koutsoyiannis et al., 2011). While the process marginal distribution is an isotropic Gaussian one $\mathrm{N}(\mu, \sqrt{\lambda})$, its dependence structure can be (separately to the marginal distribution) described by Eqn. (3), without loss of generality.

3.2 Variogram

The variogram is one of the basic tools in the field of geostatistics since it describes the spatiotemporal correlation of a process. The original term semi-variogram is coined by Matheron (1963) who expanded D.G. Krige's theory of regionalized variables and incorporated them into the theoretical framework of geostatistics. The variogram is introduced by Kolmogorov (1941), as the first-order structure function, in the study of the atmospheric turbulence and weather. Later, Jowett (1952) used the term mean-squared difference (Cressie 1989, pp. 197-202; Cressie and Wikle, 2011, p. 588). Earlier studies using the variogram are presented in the field of agriculture and particularly, in the yields of crops by Mercer and Hall (1911), in the soil survey by Youden and Mehlich (1937), in the field of the meteorology by Gandin (1965), in the forestry field by Matérn (1960) and in mine valuation by Krige (1966); further information can be found in Webster and Oliver (2007).

Modelling and estimation of the variogram is one of the most crucial steps for the kriging interpolation method (Boogaart, 2003). Beyond the numerous applications of the variogram in spatial modelling in mining engineering, it is also extensively used in geology and especially in hydrogeology, e.g. in spatial modelling of geological attributes for groundwater modelling, for the selection of the optimum grid size of the model size of an aquifer (Mohammadi, 2012), for estimating the groundwater quality parameters (Tirzo, 2014), for detecting discontinuous faults (Mohammad et al., 2015), and for detecting periods of change in a river flow time series, (Chiverton et al, 2015). For a stationary and isotropic 2d random field $\underline{x}_{\boldsymbol{s}}$ where $\boldsymbol{s}$ is any point in the process domain, the 2d (semi) variogram in continuous space is defined as (e.g. Witt and Malamud, 2013):

$$V(\boldsymbol{u}) \coloneqq \frac{1}{2} E\left[\left(\underline{x}_{s} - \underline{x}_{s+\boldsymbol{u}}\right)^2\right] \tag{4}$$

where $\boldsymbol{s} = (s_1, s_2)$ is the continuous spatial vector of the process, with s_1 and s_2 the distances from origin in each direction with units of length, $\boldsymbol{u} = (u_1, u_2)$ is the continuous spatial lag vector, with u_1 and u_2 corresponding to the lag in each direction with units of length, and E[] denotes expectation.

In 2d discrete space the variogram is similarly defined as:

$$V(h_1, h_2) \coloneqq \frac{1}{2} \mathrm{E}\left[\left(\underline{x}_{i,j} - \underline{x}_{i+h_1, j+h_2}\right)^2\right] \tag{5}$$

where $h_1 = u_1/\Delta$ and $h_2 = u_2/\Delta$ are the dimensionless spatial lags, and $\underline{x}_{i,j}$ is the space-discretized process.

It can be shown that the 2d variogram is directly linked to the 2d autocovariance function:

$$V(h_1, h_2) = \frac{1}{2}\left(\mathrm{E}\left[\underline{x}_{i,j}{}^2\right] + \mathrm{E}\left[\underline{x}_{i+h_1,j+h_2}{}^2\right]\right) - \mathrm{E}\left[\underline{x}_{i,j}\underline{x}_{i+h_1,j+h_2}\right] = c(0,0) - c(h_1, h_2) \tag{6}$$

where $c(h_1, h_2)$ is the 2d discrete autocovariance function and $c(0,0)$ the discrete variance of the 2d process with grid resolution $\Delta\times\Delta$. Note that the 2d climacogram is directly linked to the 2d autocovariance and thus, the 2d variogram, as (Dimitriadis et al., 2013):

$$c(h_1, h_2) \coloneqq \partial^4 \left(h_1{}^2 h_2{}^2 \gamma(h_1, h_2)\right) / \left(4\partial h_1{}^2 \partial h_2{}^2\right) \tag{7}$$

A common classical unbiased estimator of the 2d variogram can be expressed as (e.g., Witt and Malamud, 2013):

$$\underline{\hat{V}}(h_1, h_2) = \frac{1}{2n} \sum_{i,j=1}^{n} \left(\underline{x}_{i,j} - \underline{x}_{i+h_1,j+h_2}\right)^2 \tag{8}$$

Note that the maximum available lag for this estimator is n-1 (as in the autocovariance function).

Despite the extensive use of the variogram in many fields several of its limitations are often disregarded. As shown above the variogram is directly linked to the autocovariance function and therefore it carries along some of the autocovariance strengths, such as providing estimations for a large range of lags, as well as limitations, such as discretization error (Dimitriadis et al., 2016b). Other difficulties related to the variogram include the estimation of the sill, the kriging error for non-Gaussian processes, erratic behaviours of computed variograms when data are skewed or contain extremely high or low values and are discussed by Boogaart (2003) and Gringarten and Deutsch (2001). To this end, many solutions and transformations are recommended, such as to transform the data to the Gaussian space through implicit (or transformation-based) schemes before performing variogram calculations. However, it is noted that when the preservation of the LTP behaviour is of interest the selection of the appropriate implicit scheme should be done in caution and the choice of an explicit scheme is often preferable (see discussion of explicit vs. implicit schemes in Dimitriadis and Koutsoyiannis, 2018).

3.3 Climacogram vs. variogram for LTP identification

3.3.1 Background information

As explained in previous sections, the variogram, i.e. $V(h)$, is based on the covariance as a function of spatiotemporal lag and it is the arithmetic distance (or else separation or residual) between two positions in the two-dimensional spatiotemporal field, whereas the climacogram, i.e. $\gamma(k)$, is the variance of the averaged process as a function of spatiotemporal scale k (or else the block covariance as a function of support size; Stein et al., 2001). Variance and covariance are not identical except for $\gamma(0)$ = $c(0)$ in continuous time/space or $\gamma(\Delta)$ = $c_\Delta(0)$ in discrete time/space where discretization/regularization is at time/space scale equal to Δ. In general, the concepts of climacogram, variogram (or autocovariance) and power spectrum are all mathematically equivalent since they all contain the same information of the second-order dependence structure but expressed as a function of scale, lag and frequency, respectively (Dimitriadis et al., 2016b). In other words, they can be all constructed and express the second-order dependence structure provided that the mathematical expression of either one is given (Koutsoyiannis, 2016).

Here, we compare the climacogram and the variogram estimators in terms of identification of LTP processes. For comparison between additional methods and benchmark investigations of LTP processes see also Witt and Malamud (2013). A major advantage of the climacogram is that both Markov and white noise processes exhibit the same behaviour in terms of their climacogram at large scales, i.e., H = 0.5, whereas the variogram is bounded by $c(0)$ at large lags, a characteristic originating from its definition, i.e. $\lim_{h\to\infty}\left(c(0)-c(h)\right)=c(0)$. Additionally, it can be easily shown that the log-log derivative of the (1d, 2d, etc.) variogram always tends to zero for an LTP process:

$$\lim_{h\to\infty}\frac{\mathrm{d}\left(\log\left(V(h)\right)\right)}{\mathrm{d}(\log h)}=-\lim_{h\to\infty}\frac{h}{c(0)}\frac{\mathrm{d}c(h)}{dh}\sim\lim_{h\to\infty}\frac{1}{c(0)h^{2d(1-H)}}=0 \quad (9)$$

where $c(h)$ is the continuous-space autocovariance of the isotropic HK process, with $h=\sqrt{h_1^2+h_2^2}$ the isotropic lag, and h_1 and h_2 the spatial lags).

3.3.2 Methodology

As explained above, the LTP behaviour cannot be easily estimated from the variogram. For illustration, we estimate the climacogram and variogram and assess the difference in estimation uncertainty for LTP processes through the variance of the estimator. Particularly, we apply a Monte-Carlo analysis for a HK process with various Hurst parameters by generating 10^2 spatial fields with n = 10^2 each and estimate their climacograms and variograms. For the generation scheme we use the Symmetric-Moving-Average (SMA) algorithm introduced by Koutsoyiannis (2000) and applied in 2d spatial precipitation fields by Koutsoyiannis et al. (2011) and in various other 2d processes (Dimitriadis et al., 2013). In the SMA scheme, the simulated process is expressed through the sum of products of coefficients a_j and white noise terms $\underline{v}_i$,:

$$\underline{x}_i=\sum_{j=-l}^{l}a_{|j|}\underline{v}_{i+j} \quad (10)$$

where the summation bound l theoretically equals infinity but a finite number can be used for preserving the dependence structure up to lag l. Also, for simplicity and without loss of generality we assume that $\mathrm{E}[\underline{x}] = \mathrm{E}[\underline{v}] = 0$ and $\mathrm{E}[\underline{v}^2] = \mathrm{Var}[\underline{v}] = 1$. This scheme can be used for stochastic generation of any type of second-order process structure represented by functions such as the climacogram, power spectrum or variogram, and it exhibits several advantages over other widely used schemes (Dimitriadis and Koutsoyiannis, 2018).

For an HK process with $H > 0.5$ the SMA coefficients can be estimated analytically (Koutsoyiannis, 2016):

$$a_j = C\left(\frac{|j+1|^{H+\frac{1}{2}} + |j-1|^{H+\frac{1}{2}}}{2} - |j|^{H+\frac{1}{2}}\right) \tag{11}$$

where $C = \sqrt{2\Gamma(2H+1)\sin(\pi H)\gamma(\Delta)/\Gamma^2\left(H+\frac{1}{2}\right)\left(1+\sin(\pi H)\right)}$, $\Gamma(\,)$ is the gamma function and Δ the spatial resolution.

The employment of an uncertainty analysis in this task of spatial model identification and building is rather important (Heuvelink, 1998). Here, we perform a sensitivity analysis on the variogram and climacogram estimator to highlight each one's pros and cons, while a similar analysis for the same estimators in 1d processes can be seen in Dimitriadis et al. (2016b).

3.3.3 Results

In Fig. 5, we show the results from this analysis by focusing on the variance of each estimator θ, i.e. $\mathrm{Var}[\underline{\hat{\theta}}]$, for each process and for each scale and lag.

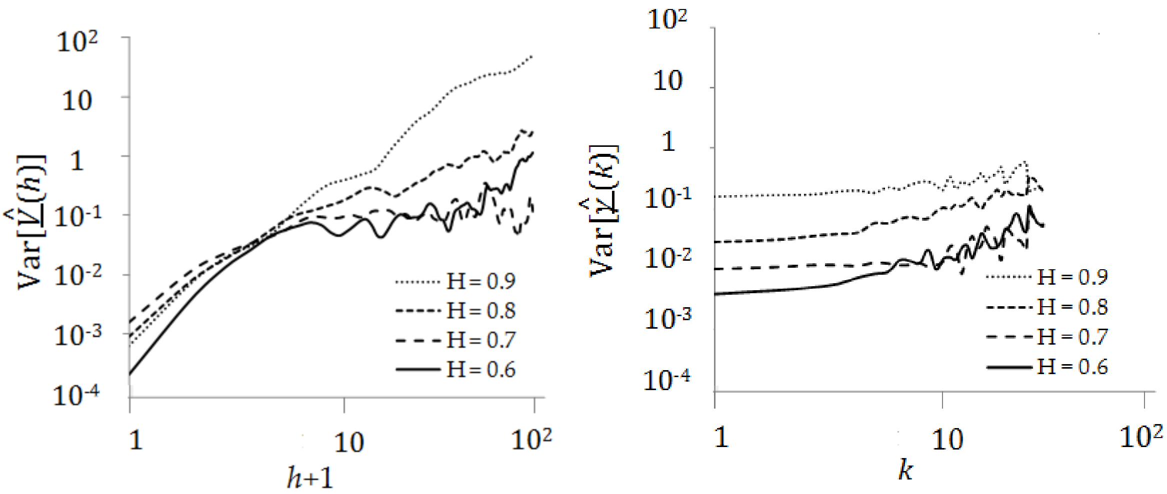

Figure 5: [left] Variance of the variogram estimator (i.e. $\mathrm{Var}[\underline{\hat{V}}(h)]$) and [right] of the climacogram estimator, i.e. $\mathrm{Var}[\underline{\hat{\gamma}}(k)]$, for various synthetic 2d spatial fields corresponding to HK processes. Note that we plot the variogram vs. h+1 so that all lags (including zero lag) is depicted in the graph as in the case of scales in the climacogram.

In Fig. 5, we observe that although the variance of both variogram and climacogram respectively increase monotonically due to the increasing uncertainty at higher lags/scales, they exhibit a different behaviour at small and large lags/scales. Particularly, the variogram has a smaller variance at small lags, whereas the climacogram has smaller variance at large scales. Therefore, an important conclusion is that the variogram can more validly identify process behaviour at small lags, i.e. estimation of local properties, such as fractal dimension as described in Gneiting and Schlather (2004), while the behaviour at large scales, such as in case of an LTP process, can be better identified and quantified by the climacogram. Similar results are also expected for generalized HK processes as well as for higher dimensions (e.g. Dimitriadis et al., 2013).

4 Application of the climacogram at different types of rock and scales

In this section, we estimate the climacogram and the variogram for each rock sample. For the current analysis we estimate solely the isotropic stochastic properties of each rock and we do not take into consideration any anisotropic and/or inhomogeneous characteristics. For this latter type, investigations should apply the climacogram (or the autocovariance, variogram etc.) by testing different rotation angles into the anisotropic sample/image or by identifying several homogeneous regions in the inhomogeneous sample/image (Gerke et al., 2014; Karsanina et al., 2015; Dimitriadis et al., 2017). We can then apply climacogram-based methods to adjust for the statistical bias but also to identify other properties of the process (Dimitriadis and Koutsoyiannis, 2016b). Finally, in case an HK behaviour is identified, we can estimate the Hurst parameter by several algorithms with a variety of such algorithms including two versions of the 1d climacogram (named LSSD and LSV method) is presented in Tyralis and Koutsoyiannis (2011, and references therein). Note that here we use a version of the LSV method but for two-dimensions.

4.1 Comparison among categories of rocks

4.1.1 Comparison of climacograms among rocks of different category

We compare the climacograms of two rocks which comprised of the same minerals but from different category (Fig.1), namely a limestone and a marble (i.e. metamorphosed limestone). In Fig. 6, we observe that the climacograms of these two rocks behave quite similar, mostly due to the fact that limestone and marble have the same mineral composition, i.e. calcite and recrystallized calcite, both consist of one mineral (calcite) and are both light coloured rocks. The statistical characteristics of their minerals indicate an LTP behaviour, since the log-log slope of the climacogram for both rocks lies within the interval (-2, 0), indicating a Hurst parameter within the interval (0.5, 1). The characteristics of the dependence structure of limestone are approximately (see also in Table 1): σ = 0.05 and H = 0.85, and of marble are: σ = 0.04, H = 0.82, where σ is the sample standard deviation. Note that the climacogram of a grayscale image is dimensionless.

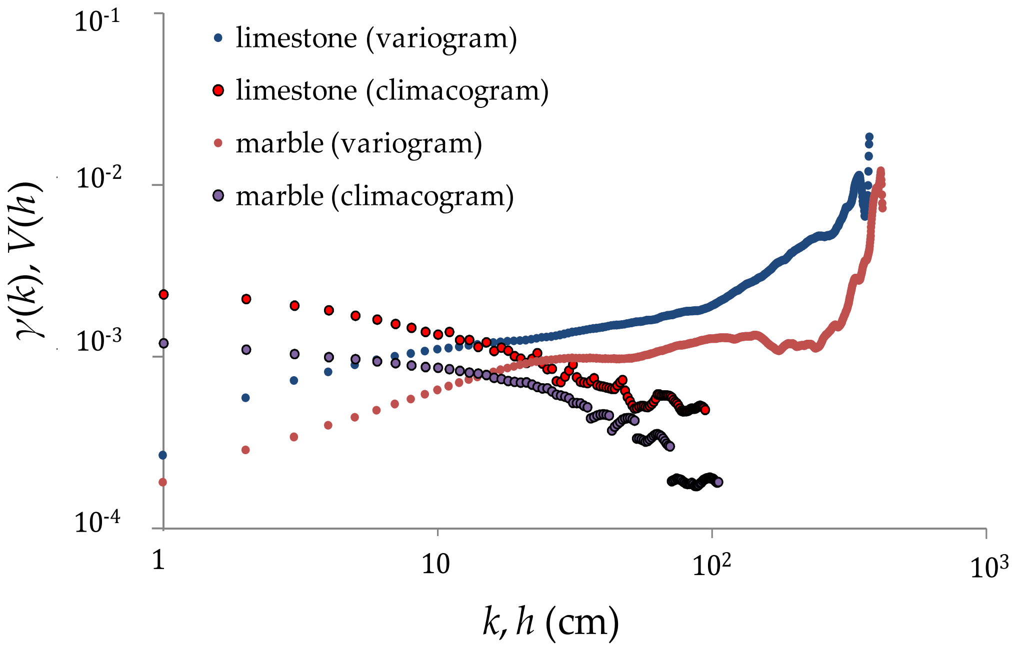

Figure 6: Climacograms of sample images from limestone and marble and the corresponding variograms for illustration purposes.

4.1.2 Comparison of climacograms among different rocks of same category

A comparison of climacograms for rocks of the same category, namely a limestone and a sandstone (Fig. 1), at the same spatial scale (hand specimen, i.e., in cm) is shown in Fig.7. Comparing the two estimated climacograms, we notice that the range of the variance at scale 1 varies significantly since the limestone is a mono-mineralic rock compared to sandstone. Again, the statistical characteristics of their components indicate LTP behaviour. The characteristics of the dependence structure of sandstone are approximately (see also in Table 1): σ = 0.14 and H = 0.61.

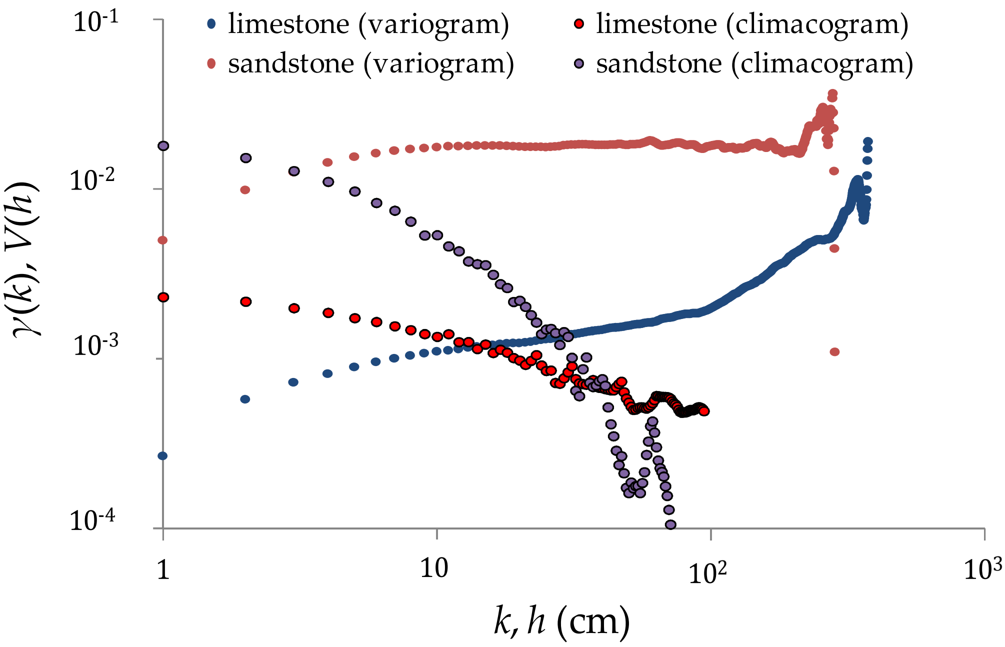

Figure 7: Climacograms of a limestone and a sandstone sampe image, and the corresponding variograms for illustration purposes.

4.2 Comparison among scales of rocks

4.2.1 Comparison of climacograms among different scales

We analyze sample images from an igneous rock, i.e. rhyolite (Fig.2), at moderate scale (mm) and mesoscale (cm). In Fig. 8, both climacograms exhibit approximately the same LTP behaviour (Table 1) with a spatial displacement of the climacogram in mesoscale one order of magnitude as much as the difference in image resolution of the two rocks. The characteristics of the dependence structure of rhyolite at moderate scale are approximately (Table 1): σ = 0.21 and H = 0.77, and of rhyolite at mesoscale: σ = 0.12, H = 0.77. Note that the Hurst parameter is the same in both cases.

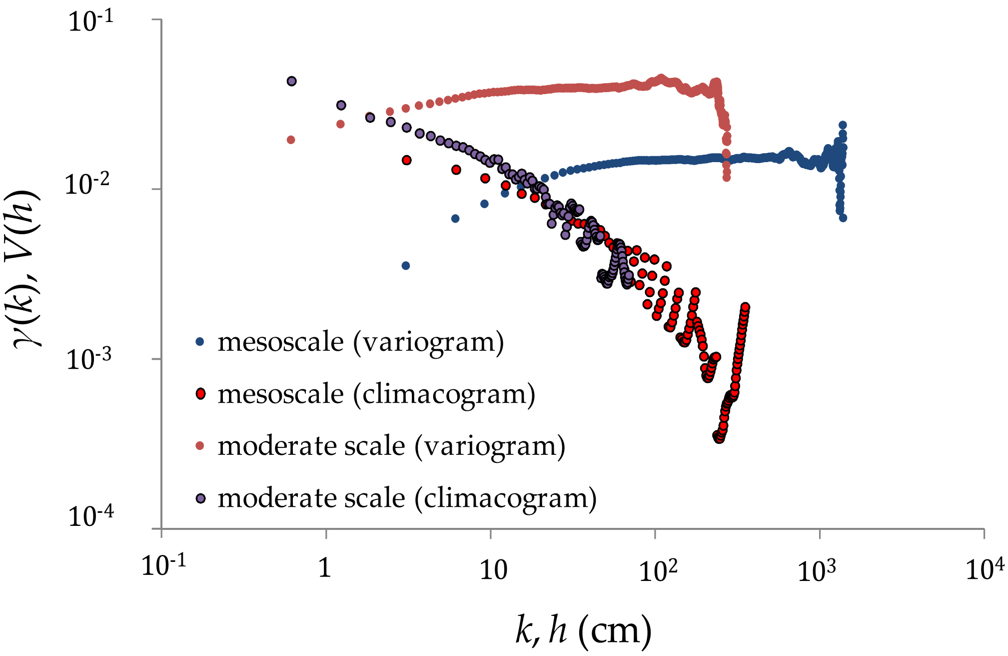

Figure 8: Climacograms of a rhyolite rock at moderate scale (mm) and mesoscale (cm), and the corresponding variograms for illustration purposes.

4.2.2 Comparison of climacograms at multiple scales

In Fig. 9, we combine climacograms from images of sandstone at four different scales. Particularly, we analyze sample images (Fig. 3) with resolution of microscale (μm and cm), mesoscale (cm) and macroscale (m). Here, LTP behaviour is more evident and the overall Hurst parameter is estimated approximately equal to 0.85, when the bias is taken into account through the unbiased estimator of the 2d climacogram for HK processes, i.e. (Dimitriadis et al., 2013) $\underline{\hat{\gamma}}(\kappa_1, \kappa_2)(1 - \kappa_1\kappa_2/n^2) + \gamma(n)$, based on Eqn. 2 and 3.

Note that the quick drop of each climacogram at large scales is due to low statistical sampling at large scales (observe that on average the estimated Hurst parameter is increasing with sample length.) This can be roughly removed by following the rule of thumb of fitting the climacogram to a stochastic model up to the 10% of the extent of available scales (Dimitriadis and Koutsoyiannis, 2015).

It is interesting to see that all examined rock formations exhibit LTP behaviour, with Hurst parameters ranging from 0.6 to 0.85 (not adjusted for bias) and an overall 0.85 (adjusted for bias). Therefore, the uncertainty/variability of these rocks seems to be much larger than that emerging from a white noise or a Markov process.

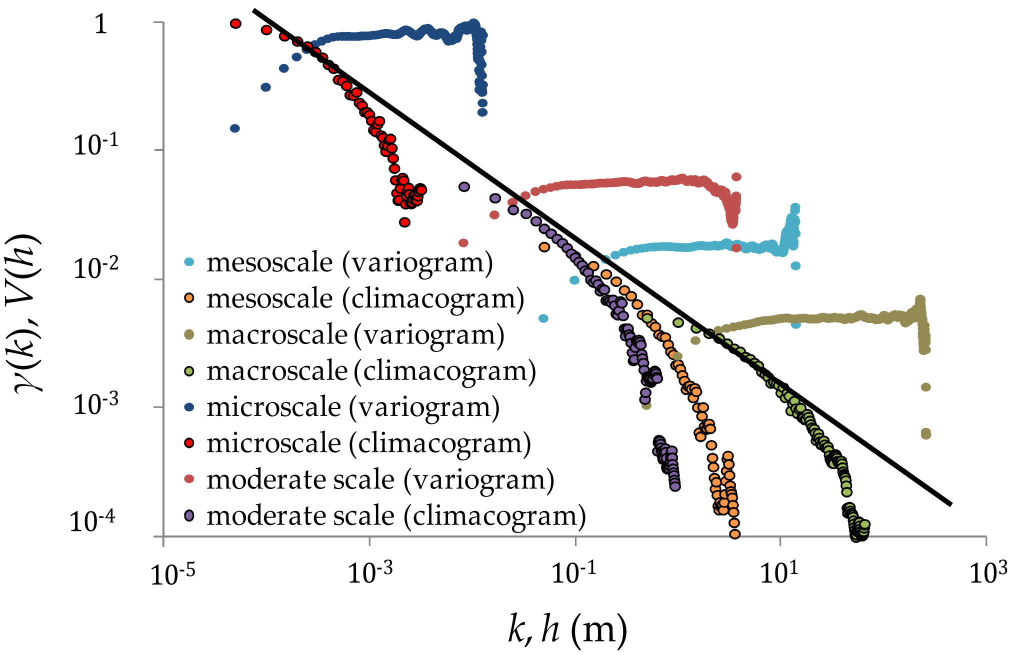

Figure 9: Climacograms of images from sandstone at four different ranges of scales and the corresponding variograms for illustration purposes.

Table 1: Marginal statistical characteristics of 2d rock samples.

Type of rock	$n \times n$	σ	σ/μ	C_s	C_k	H
limestone (Fig. 1)	141 376	0.048	0.092	-0.527	3.016	0.847
marble (Fig. 1)	176 400	0.035	0.051	0.241	3.538	0.818
sandstone (Fig. 1)	81 225	0.135	0.246	-0.278	3.061	0.612
rhyolite thin section (Fig. 2)	202 500	0.210	0.322	-0.846	3.313	0.766
rhyolite hand-specimen (Fig. 2)	207 936	0.123	0.208	-0.146	3.607	0.773
sandstone microscale (Fig. 3)	67 081	1.000*	0.355	0.128	2.418	0.765
sandstone moderate scale (Fig. 3)	202 500	0.232	0.404	-0.398	2.115	0.772
sandstone mesoscale (Fig. 3)	81 225	0.135	0.246	-0.278	3.061	0.713
sandstone macroscale (Fig. 3)	272 484	0.072	0.155	-0.419	3.237	0.754

*the variance of the SEM sample is arbitrarily set to 1 since it cannot be directly compared to the other samples due to the completely different sampling method

5 Summary and discussion

The aim of this study is to examine the stochastic similarities of rocks in terms of second-order dependence structure expressed through the climacogram and in particular, whether they exhibit long-term persistence for a wide range of scales and rock formations. The presented analysis may be useful for gaining insight and making inference at scales in which data acquisition is difficult or costly.

A common characteristic drawn from the current research and the analysis in all the rock formations and scales is the power-law decay of climacogram, i.e. variance of the scaled averaged process. This structure signifies a long-term-persistent, or else known as Hurst-Kolmogorov (HK) behaviour, as the Hurst parameter ranges within 0.5 and 1, signifying a difference from a white noise process (i.e. absence of autocorrelation) or Markov behaviour (i.e. exponential decay of autocorrelation). This result can be useful towards a more realistic reconstruction of rock images through appropriate stochastic models that take into account the long-term-persistence, such as the proposed 2d HK one (Eqn. 3), which is a two parameter model entirely based on the climacogram. Also, this large variability introduced by a rock formation may give insight on how a low variability often observed in precipitation at large scales (e.g. Tyralis et al, 2017 and references therein) is translated, through a non-linear rainfall-runoff system (e.g. Manfreda, 2008), to sometimes larger variability for the same range of scales in river stage/discharges (e.g. Hurst, 1951; Koutsoyiannis et al., 2008).

An additional result is that images of the same rock type at different scales, from micro to macro, suggest similar type of clustering, i.e. with a similar scaling parameter. In particular, the Hurst parameter is estimated (on average) around 0.75 in most cases (Table 1) when the bias is not taken into account and 0.85 from the combination of all climacograms adjusted for bias (Fig. 9). This result suggests that the examined rock formations and range of scales exhibit a similar power-law decay of the second-order dependence structure, with a similar Hurst parameter $0.5 < H < 1$. In other words, this behaviour is characterized by high statistical uncertainty (here quantified through variability) which, for the examined range of scales, is larger than the one corresponding to a white noise or a Markov process Interestingly, similar Hurst parameters have been estimated in various other processes (Dimitriadis, 2017) of completely different nature from the ones analyzed here. For example, for an isotropic turbulence timeseries of massive length, H is estimated at 0.83 (Dimitriadis and Koutsoyiannis, 2018), while a global analysis from thousand of stations of atmospheric wind and temperature also indicated similar values (Koutsoyiannis et al., 2018).

A final remark is that while the variogram seems to be more appropriate for investigating the local behaviour in small-scale structures of a process, the climacogram is shown to perform more robustly in estimating large-scale properties, especially when a possible HK behaviour is of interest. This result is based on the variability quantification of both in several benchmark tests on HK processes using Monte-Carlo techniques.

Acknowledgement

The Authors are grateful to the Editor A. Stein, the anonymous Guest Editor and the three anonymous Reviewers for the effort put in handling the review and for their fruitful comments and suggestions that helped us improve the paper.

Code availability

The calculations for the 2d climacogram and 2d variogram estimators are implemented in Matlab and the corresponding scripts are available in contact with the corresponding author.

References

Blunt, M.J., B. Bijeljic, H. Dong, O. Gharbi, S. Iglauer, P. Mostaghimi, A. Paluszny, and C. Pentland, Pore-scale imaging and modelling, *Advances in Water Resources*, 51, 197–216, 2013.

Boogaart, K.G., Odds and ends of variogram modelling, *Proceedings of the annual conference of the International Association for Mathematical Geology IAMG 2003*, 7-12 September 2002, Portsmouth, 2002.

Chilès, J.P., and P. Delfiner, *Geostatistics, Modelling Spatial Uncertainty*, Wiley Series in Probability and Statistics, Wiley, 2012.

Chiverton, A., J. Hannaford, I.P. Holman, R. Corstanje, C. Prudhomme, T.M. Hess, and J.P. Bloomfield, Using variograms to detect and attribute hydrological change, *Hydrology and Earth System Science*, 19, 2395–2408, doi:10.5194/hess-19-2395-201, 2015.

Cressie, N., Geostatistics, *The American Statistician*, 43(4), 1989.

Cressie, N., and C.K. Wikle, *Statistics for Spatio-Temporal Data*, Wiley, Hoboken, N.J., 2011.

Dimitriadis, P., Hurst-Kolmogorov dynamics in hydrometeorological processes and in the microscale of turbulence, PhD thesis, Department of Water Resources and Environmental Engineering – *National Technical University of Athens*, 2017.

Dimitriadis, P., D. Koutsoyiannis, and C. Onof, N-Dimensional generalized Hurst-Kolmogorov process and its application to wind fields, *Facets of Uncertainty: 5th EGU Leonardo Conference – Hydrofractals 2013 – STAHY 2013*, Kos Island, Greece, European Geosciences Union, International Association of Hydrological Sciences, International Union of Geodesy and Geophysics, doi:10.13140/RG.2.2.15642.64963, 2013.

Dimitriadis, P., and D. Koutsoyiannis, Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes, *Stochastic Environmental Research & Risk Assessment*, 29 (6), 1649–1669, doi:10.1007/s00477-015-1023-7, 2015.

Dimitriadis, P., D. Koutsoyiannis, and K. Tzouka, Predictability in dice motion: how does it differ from hydrometeorological processes?, *Hydrological Sciences Journal*, 61 (9), 1611–1622, doi:10.1080/02626667.2015.1034128, 2016a.

Dimitriadis, P., D. Koutsoyiannis, and P. Papanicolaou, Stochastic similarities between the microscale of turbulence and hydrometeorological processes, *Hydrological Sciences Journal*, doi:10.1080/02626667.2015.1085988, 2016b.

Dimitriadis, P., K. Tzouka, H. Tyralis, and D. Koutsoyiannis, Stochastic investigation of rock anisotropy based on the climacogram, *European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19*, Vienna, EGU2017-10632-1, European Geosciences Union, 2017.

Dimitriadis, P., and D. Koutsoyiannis, Stochastic synthesis approximating any process dependence and distribution, *Stochastic Environmental Research & Risk Assessment*, doi:10.1007/s00477-018-1540-2, 2018.

Elogne S, D.T. Hristopulos, and M. Varouchakis, An application of Spartan spatial random fields in environmental mapping: focus on automatic mapping capabilities, *Stochastic Environmental Research & Risk Assessment*,22(5):633–46, 2008.

Gerke K. M., M.V. Karsanina, R.V. Vasilyev, and D. Mallants, Improving pattern reconstruction using correlation functions computed in directions, *EPL (Europhysics Letters)*, 106(**6**), 66002 2014.

Gerke K.M., M.V. Karsanina, and D. Mallants, Universal Stochastic Multiscale Image Fusion: An Example Application for Shale Rock, *Scientific Reports*, 5, 15800-15880, doi: 10.1038/srep15880, 2015.

Gloaguen R., P.R. Marpu, and I. Niemeyer, Automatic extraction of faults and fractal analysis from remote sensing data, *Nonlinear Processes in Geophysics*, 14, 131–138, 2007.

Gneiting, T., and M. Schlather, Stochastic models that separate fractal dimension and the Hurst effect, *Society for Industrial and Applied Mathematics Review*, 46(2), 269-282, 2004.

Gommes C.J., Y. Jiao, and S. Torquato, Microstructural degeneracy associated with a two-point correlation function and its information content, *Physical Review E*, 85, 051140, 2012.

Gringarten E., and C.V. Deutsch, Variogram interpretation and modeling, *Mathematical Geology*, 33(4), 507-534, 2001.

Hamzehpour, H., M.R. Rasaei, and M. Sahimi, Development of optimal models of porous media by combining static and dynamic data: the permeability and porosity distributions, *Physical Review E.*, 75, 056311/1–056311/17, 2007.

Heuvelink, G.B.M, Uncertainty analysis in environmental modelling under a change of spatial scale, *Nutrient Cycling in Agroecosystems*, 50, 255–264, 1998.

Heuvelink, G.B.M., and R. Webster, Modelling soil variation: past, present, and future, *Geoderma*, 100, 269–301, 2001.

Hurst, H.E., Long-term storage capacity of reservoirs, *Trans. Amer. Soc. Civ. Eng.*, 116, 770–808, 1951.

Jennane, R.W., J. Ohley, S. Majumdar, and G. Lemineur, Fractal Analysis of Bone X-Ray Tomographic Microscopy Projections, *IEEE Transactions on Medical Imaging*, 20(5), 2001.

Karsanina M.V., K.M. Gerke, E.B. Skvortsova, and D. Mallants, Universal spatial correlation functions for describing and reconstructing soil microstructure, *PLoS ONE*, 10 (5), 2015.

Kolmogorov, A.N., Dissipation energy in locally isotropic turbulence, *Doklady Akademii Nauk. SSSR*, 32, 16-18, 1941.

Koutsoyiannis, D., The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal,* 47(4):573-595. doi:10.1080/02626660209492961, 2002.

Koutsoyiannis, D., Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48 (1), 3-24, doi:10.1623/hysj.48.1.3.43481, 2003.

Koutsoyiannis, D., H. Yao, and A. Georgakakos, Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods, *Hydrological Sciences Journal*, 53 (1), 142–164, doi:10.1623/hysj.53.1.142, 2008.

Koutsoyiannis, D., A random walk on water, *Hydrology and Earth System Sciences*, 14, 585–601, 2010.

Koutsoyiannis, D., A. Paschalis, and N. Theodoratos, Two-dimensional Hurst-Kolmogorov process and its application to rainfall fields, *Journal of Hydrology*, 398 (1-2), 91–100, 2011.

Koutsoyiannis, D., P. Dimitriadis, F. Lombardo, and S. Stevens, From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data, *Advances in Nonlinear Geosciences*, edited by A.A. Tsonis, 237–278, doi:10.1007/978-3-319-58895-7_14, Springer, 2018.

Koutsoyiannis D., Generic and parsimonious stochastic modelling for hydrology and beyond, *Hydrological Sciences Journal*, 61 (2), 225–244, doi:10.1080/02626667.2015.1016950, 2016.

Mandelbrot, B.B., van Ness, J.W., Fractional Brownian motions, fractional noises and applications, *SIAM Review*, 10 (4), 422–437, 1968.

Manfreda, S., Runoff generation dynamics within a humid river basin, *Nat. Hazards Earth Syst. Sci.*, 8, 1349–1357, doi:10.5194/nhess-8-1349-2008, 2008.

Matheron, G., Principles of geostatistics, *Economic Geology*, 58, 1246-1266, 1963.

Milliner, C.W.D, C. Sammis, A.A. Allam, J.F. Dolan, J. Hollingsworth, S. Leprince, and F. Ayoub, Resolving Fine-Scale Heterogeneity of Co-seismic Slip and the Relation to Fault Structure, *Scientific Reports*, 6, 27201, doi: 10.1038/srep27201, 2016.

Mohammadi, Z., Variogram-based approach for selection of grid size in groundwater modelling, *Applied Water Science*, 3(3), 597–602, doi: 10.1007/s13201-013-0107-0, 2013

Mohammad, S.F., R. Hamidreza, and K. Milad, Variography, an Effective Tool to Detect Geological Discontinuities among Geo-Electrical Data: A Case Study of Hamyj Copper Deposit, *Geodynamics Research International Bulletin,* 2 (5), 9-13, 2015.

O'Connell, PE, D. Koutsoyiannis, H. F. Lins, Y. Markonis, A. Montanari, and T.A. Cohn, The scientific legacy of Harold Edwin Hurst (1880 – 1978), *Hydrological Sciences Journal*, 61 (9), 1571–1590, doi:10.1080/02626667.2015.1125998, 2016.

Qian, H., G. M. Raymond, and J. B. Bassingthwaighte, On two-dimensional fractional Brownian motion and fractional Brownian random field, *J. Phys. A: Math. Gen.,* 31, L527, 1998.

Rabbani A., S. Ayatollahi, R. Kharrat, and N. Dashti, Estimation of 3-D pore network coordination number of rocks from watershed segmentation of a single 2-D image, *Advances in Water Resources*, 94, 264-277, 2016.

Stein A., J. Riley, and N. Halberg, Issues of scale for environmental indicators, *Agriculture, Ecosystems and Environment*, 87, 215–232, 2001.

Talukdar, M.S., O. Torsaeter, and M.A. Ioannidis, Stochastic Reconstruction of Particulate Media from Two-Dimensional Images, *Journal of Colloid and Interface Science*, 248, 419–428, 2002.

Tizro, A.T., K. Voudouris, and S. Vahedi, Spatial Variation of Groundwater Quality Parameters: A Case Study of a Semiarid Region of Iran, *International Bulletin of Water Resources & Development*, 1(3), 1-14, doi:10.1006/jcis.2001.8064, 2014.

Tyralis, H, and D. Koutsoyiannis, Simultaneous estimation of the parameters of the Hurst-Kolmogorov stochastic process, *Stochastic Environmental Research & Risk Assessment,* 25(1), 21-33, doi:10.1007/s00477-010-0408-x, 2011.

Tyralis, H., P. Dimitriadis, D. Koutsoyiannis, P.E. O'Connell, K. Tzouka, and T. Iliopoulou, On the long-range dependence properties of annual precipitation using a global network of instrumental measurements, *Advances in Water Resources*, 2017.

Vanitha, S., R. Rajeswari, and D. Ebenezer, An Analysis and Reduction of Fractional Brownian motion Noise in Biomedical Images Using Curvelet Transform and Various Filtering and Thresholding Techniques, *International Refereed Journal of Engineering and Science*, 5(10), 18-27, 2016.

Varouchakis, E.A., K. Kolosionis, and G.P. Karatzas, Spatial variability estimation and risk assessment of the aquifer level at sparsely gauged basins using geostatistical methodologies, *Earth Science Informatics,* 1–12, 2016.

Webster, R., and M.A. Oliver, *Geostatistics for Environmental Scientists,* Second Edition, Wiley, Chichester, 2007.

Witt, A., and B. Malamud, Quantification of Long-Range Persistence in Geophysical Time Series: Conventional and Benchmark-Based Improvement Techniques, *Surveys in Geophysics*, 34:541–651, doi: 10.1007/s10712-012-9217-8, 2013.