Leadership in Sustainable Infrastructure
Leadership en Infrastructures Durables

Vancouver, Canada
May 31 – June 3, 2017/ *Mai 31 – Juin 3, 2017*

EXTREME VALUES OF A GAMMA FIELD BY SPECTRALLY BASED MONTE CARLO SIMULATION

Adegbola, Adetola[1] and Yuan, Xian-Xun[2,3]
[1, 2] Ryerson University, Department of Civil Engineering, Toronto, ON, Canada M5B 2K3
[3] arnold.yuan@ryerson.ca

Abstract: Spatial uncertainty in structural degradation has been a topic of interest in time-dependent reliability. To account for this important uncertainty, a copula-based gamma distributed random field was developed by the authors. This paper presents systematic evaluation of extreme values of the gamma field by using a Monte Carlo simulation method that was developed recently based on Karhunen-Loève expansion. Built upon previous sensitivity studies, this paper focuses on the effects of different correlation functions on the extreme value of the gamma field. The results show that with the same correlation length the exponential correlation function gives a higher extreme value mean than the triangular function. More interestingly, the type of extreme value distribution obtained is dependent on the type of correlation function used for the gamma field. These results are useful for risk and reliability analysis of degraded structures with significant spatial uncertainty.

1 Introduction

Degradation or deterioration modeling is an important subject in time-dependent reliability analysis. It plays a central role in infrastructure asset management (Yuan 2016). Traditional stochastic degradation modeling has been focusing on the characterization and quantification of mean trend of deterioration and the modeling of effects of influential factors. Risk-informed decision making dictates that various uncertainties in deterioration be properly modeled and quantified. These uncertainties include variation of deterioration rates along time (temporal uncertainty), across different assets under the same design and working conditions (random effects), and in space (spatial uncertainty). While the first two uncertainties have attracted a plethora of research in literature, the modeling of spatial uncertainty in deterioration has been very limited. Integrated deterioration modeling with considerations of all three types of uncertainties is only sporadic in literature.

Gamma process model has gained a lot of traction in stochastic degradation modeling over the past decade, particularly after the review paper (van Noortwijk 2009) was published. More recently, this line of research has been expanded into multivariate domain. For example, Shemehsavar (2014) developed a bivariate gamma process to model the relationship between a latent degradation process and the observed degradation. The bivariate gamma process was constructed based on the Kibble-Wicksell bivariate gamma distribution. In another study, Wang et al. (2015) proposed a bivariate non-stationary gamma process of which the increment pairs, $\Delta X_1(t)$ and $\Delta X_2(t)$, follows a copula-based bivariate gamma distribution. They considered four different types of copula models including Gaussian, Frank, Gumbel, and Clayton. Inspired by those bivariate gamma process models, Adegbola and Yuan (2017) extended the idea into a copula-based gamma field and used it to model the spatial uncertainty in deterioration.

Among many interesting topics, the distribution of the extreme value of a gamma field at any given time is one of the most interesting quantities to be evaluated in system reliability and safety analysis. This issue is formulated as the extreme value distribution of the gamma field, which may be used to assess overall risk during structural integrity assessment. Extensive literature review indicates there is no analytical solution for the extreme value distribution. Therefore, numerical approach seems to be the last resort. As the random field is non-Gaussian, an efficient methodology to simulate the field was sought in the literature. There are two possible schemes: straightforward simulation by fine mesh with local averaging and spectrally based simulation. Preliminary studies have indicated that the straightforward simulation runs into a serious convergence problem, hence the need for the latter. To address this issue, Adegbola and Yuan (2017) developed a Monte Carlo simulation algorithm based upon Karhunen-Loève (K-L) expansion. The method has been validated to obtain reliable solution. With a maximum of 400 K-L expansion terms, the mean and standard deviation of the extreme value are converged with little computational error. Using the proposed simulation method, they studied how the extreme value statistics, namely the mean, standard deviation and distribution type, would change with the parameters of the gamma field. This paper extends the previous study by examining the effects of correlation function of the gamma field on the extreme value statistics.

The paper is organized as follows. The fundamentals of gamma field and the Karhunen-Loève expansion based simulation method are reviewed in Section 2. Section 3 briefly reviews the generalized extreme value distribution and its three classical types. The results for different correlation functions are presented and discussed in Section 4. Section 5 concludes the paper.

2 The Gamma Random Field and Spectrally-Based Simulation Method

2.1 Gamma Field

Simply put, a gamma field is a random field whose marginal distribution of the field at any specific point of location follows a gamma distribution. Naturally, the joint distribution of the random vector at any given number of different points is then a multivariate gamma distribution. However, due to the complexity of the dependence structure of a multivariate gamma distribution, the technical definition is not so straightforward. Adegbola and Yuan (2017) developed the definition starting from the univariate gamma distribution and multivariate version, which were then extended into the infinite dimension case for a gamma field. The details can be found there. For completeness of the discussion, here only the key points are emphasized.

First of all, gamma distribution is a continuous probability distribution whose probability density function is expressed as

[1] $$g(x;\alpha,\,\beta)=\frac{x^{\alpha-1}\,e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)}\,,x\geq 0$$

where $\alpha>0$ and $\beta>0$ are the shape and scale parameters, respectively, and $\Gamma(p)=\Gamma(p,\infty)$ is the gamma function. In short hand, we write $X\sim Ga(\alpha,\beta)$. The mean and variance of the gamma random variable are $\alpha\beta$ and $\alpha\beta^2$, respectively. The coefficient of variation is $1/\sqrt{\alpha}$, independent of the scale parameter β.

Secondly, the multivariate gamma distribution is copula-based. Particularly, this study takes a Gaussian copula, which is expressed as

[2] $$C(u_1,\dots,u_n)=\Phi_n(\Phi^{-1}(u_1),\dots,\Phi^{-1}(u_n);\mathbf{R})$$

where $C(\cdot)$ is the n-dimensional copula function, $\Phi_n(z_1,\dots,z_n;\mathbf{R})$ denotes the joint cumulative distribution function of a standard multivariate normal distribution with a correlation matrix $\mathbf{R}=\{r_{ij}\}$; and $z_i=\Phi^{-1}(u_i)$ denotes the inverse of the standard normal CDF at probability u_i. For more details of copula and multivariate gamma distribution, refer to Joe (2001) and Kotz, Balakrishnan and Johnson (2000). With this, a random vector $(X_1,\dots,X_n)$ is said to follow a multivariate gamma distribution if they satisfy the following:

(1) $X_i\sim Ga(\alpha_i,\beta_i), i=1,\dots,n$

(2) The joint cumulative distribution function of $X_1,\dots,X_n$ is defined as

[3] $$H(\boldsymbol{x}) = \Phi_n(z_1, \dots, z_n; \mathbf{R})$$

where $z_i = \Phi^{-1}(u_i)$ and $u_i = G(x_i; \alpha_i, \beta_i)$. It is readily shown that the joint probability density function is expressed as

[4] $$h(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{R}^{-1}\mathbf{z}\right) \prod_{i=1}^{n} g(x_i; \alpha_i, \beta_i)$$

Finally, a random field $X(s)$ in a finite d-dimensional Euclid space Ω is called a homogeneous gamma field if it satisfies the following:

1) For any point $s \in \Omega$, the field value $X(s)$ is a random variable that follows a gamma distribution with shape α and scale β, i.e., $X(s) \sim Ga(\alpha, \beta)$.
2) For any n points $s_1, s_2, \dots, s_n \in \Omega$, the field values $X(s_1)$, $X(s_2), \dots, X(s_n)$ follow a multivariate gamma distribution defined previously with $\alpha_i = \alpha$ and $\beta_i = \beta$ for $i = 1, \dots, n$, and the correlation coefficient r_{ij} ($i, j = 1, \dots, n$) defined by $r_{ij} = r(\boldsymbol{h}_{ij})$, where $r(\cdot)$ is a correlation function, and $\boldsymbol{h}_{ij}$ a certain distance measure between the two points s_i and s_j, which may be a vector.

The correlation function is a very important second-order characteristic of a homogeneous random field. For one-dimensional fields, commonly used correlation functions are exponential, Gaussian, spherical, and triangular. Particularly, an exponential correlation function is expressed as

[5] $$C(s_1, s_2) = e^{-3h/\theta}$$

for $h = |s_1 - s_2| \geq 0$, and a triangular function as

[6] $$C(s_1, s_2) = \begin{cases} 1 - h/\theta, & 0 \leq h \leq \theta \\ 0, & h > \theta \end{cases}$$

In both models, θ is called correlation length. Altogether, the shape parameter α, the scale parameter β, and the correlation length θ define the one-dimensional homogeneous gamma field. In stochastic degradation modeling, the three parameters can be estimated from inspection data by using, e.g., maximum likelihood method and Bayesian method.

2.2 K-L Expansion-Based Monte Carlo Simulation

Simple mesh-based Monte Carlo simulation using conventional Cholesky decomposition of the multivariate gamma distributions will not work for the gamma field, particularly when the focus of study in on the extreme value (in our study, the maximum of deterioration). The Karhunen-Loève (K-L) expansion was used to generate homogeneous Gaussian field (Ghanem & Spanos 1991), and recently extended to simulate the gamma field (Adegbola & Yuan 2017).

The idea of the K-L expansion for a Gaussian field is very simple. Basically, a zero-mean Gaussian field $Z(s, \theta)$ defined on a finite one-dimensional domain $\Omega = [-L, L]$ can be represented by the K-L expansion with infinite number of terms as

[7] $$Z(s, \theta) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\,\xi_i(\theta) f_i(s)$$

where $\xi_i(\theta)$ are independent standard normal variables, and λ_i and $f_i(s)$ are eigenvalues and eigenfunctions, respectively, of the covariance function. For an exponential correlation function as defined in [5], the eigenfunctions are shown in pair as

[8] $$f_n(s) = \frac{\cos(\omega_n s)}{\sqrt{L + \sin(2\omega_n L)/2\omega_n}}; \; f_n^*(s) = \frac{\sin(\omega_n^* s)}{\sqrt{L - \sin(2\omega_n^* L)/2\omega_n^*}}$$

and the corresponding eigenvalues as

[9] $$\lambda_n = \frac{2c}{\omega_n^2 + c^2}; \quad \lambda_n^* = \frac{2c}{\omega^{*2}_n + c^2}$$

where $c = 3/\theta$, and ω_n and ω_n^* are the solutions of the following transcendental equations, respectively:

[10a] $c - \omega \tan(\omega L) = 0$

[10b] $\omega + c \tan(\omega L) = 0$

For the triangular covariance function given as [6], the eigenfunctions are

[11] $$f_n(s) = \frac{\cos(\omega_n s) + \tan\left(\frac{\omega_n L}{2}\right)\sin(\omega_n s)}{\sqrt{L + (\tan^2\left(\frac{\omega_n L}{2}\right) - 1)(\frac{L}{2} - \frac{\sin(2\omega_n L)}{4\omega_n}) + \frac{\sin^2(\omega_n L)}{\omega_n}\tan\left(\frac{\omega_n L}{2}\right)}}; \; f_n^*(s) = \frac{\cos(\omega_n^* s)}{\sqrt{\frac{L}{2} + \frac{\sin(2\omega_n^* L)}{2\omega_n^*}}}$$

for odd and even n, respectively, and the eigenvalues are

[12] $$\lambda_n = \frac{2c}{\omega_n^2}; \quad \lambda_n^* = \frac{2c}{\omega_n^{*2}}$$

where $\omega_n = \frac{n\pi}{a}$ and $\omega_n^* = \frac{n\pi}{a}$ again are for odd and even n, respectively. These results are well established in the literature of stochastic field, e.g., Ghanem and Spanos (1991).

In reality, often only the truncated K-L expansion is used, resulting in the following approximate expansion:

[13] $$Z(s,\theta) = \sum_{i=1}^{n}\left[\xi_i\sqrt{\lambda_i}f_i(s) + \xi_i^*\sqrt{\lambda_i^*}f_i^*(s)\right]$$

where ξ and ξ^* are independent standard normal variates.

Once the K-L expansion of a Gaussian field is established, the simulation of the one-dimensional gamma field is based on memoryless transformation. This involves evaluating the standard normal CDF of the Gaussian variates and then finding the inverse gamma transform with known shape and scale parameters. The step-by-step simulation procedure is as follow:

1) Set the parameters L, α, β and θ.

2) Prepare the eigenvalues and eigenfunctions for the triangular correlation model.

3) Generate a number of standard normal random variates. The number of variates should be equal to the number of eigenvalues. Use Eq. 13 to generate a realization of the Gaussian field. The field can be discretized and represented by a finite number of sampling points with grid size equal to $0.01L$. For each grid point s, calculate the value of the Gaussian field $z(s)$, and evaluate the standard normal CDF at $u(s) = z(s)$.

4) Perform the inverse transform of the gamma CDF for each $u(s)$, resulting in a random realization of the gamma field $x(s)$.

5) Find the maximum value of the generated gamma field. The maximum value represents a random realization of the EV of the gamma field.

6) Repeat steps 3–5 a total of 100 000 times.

7) Obtain the statistics of the maximum values.

Note that if obtaining the extreme value is the sole purpose of the simulation, then steps (4) and (5) can be further simplified into one step. That is, one can first find the extreme value of the generated Gaussian field, and then take the inverse transform of the gamma CDF to find the extreme value of the gamma field. This will drastically reduce the computation time.

A slightly different approach from the procedure above without using a mesh can be employed by noting that any random realization of the Gaussian field is a deterministic function of s in a finite region, and therefore its maximum value can be evaluated numerically by using, for example, a genetic algorithm, which is readily available in, e.g., MATLAB. This approach results in slightly higher estimate of the extreme value. However, the price is that it takes more time to maximize the sampled random field using MATLAB's genetic algorithm. In this study, the mesh method is used nevertheless.

3 Extreme Value Distribution

An extreme value (EV) can either be the maximum or minimum value in a sample. The extreme value is a random variable that has been found to follow one of three possible types of asymptotic probability distributions, namely Gumbel, Frechet and Weibull distributions. Combining all three types of extreme value distributions gives the generalized extreme value (GEV) distribution with the probability density function given as

[14] $$f(x) = \sigma^{-1}\left(1+\frac{k(x-\mu)}{\sigma}\right)^{-1-1/k} e^{-\left(1+\frac{k(x-\mu)}{\sigma}\right)^{-1/k}}$$

for $1 + k(x-\mu)/\sigma > 0$, where k, σ and μ represents the shape parameter, scale parameter and location parameter, respectively. The shape parameter k determines the distribution type of the extreme value. Precisely, for $k < 0$, the distribution reduces to a Weibull distribution (type III); for $k > 0$ it reduces to Frechet distribution (type II), and when $k = 0$ it is a limit case which represents the Gumbel distribution (type 1).

4 Sensitivity Analysis

The sensitivity study concerns how the field parameters of the gamma field affect the statistics of the extreme value. Furthermore, it aims to compare and contrast the results from the exponential and triangular correlation functions. Initial studies (Adegbola and Yuan 2017) have indicated that the scale parameter β of the univariate gamma distribution has only a multiplying effect on the extreme value mean of gamma fields, for example the EV distribution mean of a gamma field with $\beta = 2$ is 2 times that of EV distribution of a gamma field with $\beta = 1$, provided the other parameters α & θ of the field remain the same. Therefore, $\beta = 1$ is used for all cases in the sensitivity analyses. The sensitivity study focuses on the evaluation of the relationship between the mean and standard deviation of the extreme value of the gamma field with the shape parameter α and the correlation length θ for two different correlation functions.

Preliminary studies show that the number of terms required to achieve convergence in terms of mean to 3 significant figures in all cases of correlation length θ varies from 50 to 300 terms. This is because the less correlated the gamma random field is, the higher the number of terms required to achieve convergence. Therefore for each case, the study uses 300 KL expansion terms and 100 000 simulation runs to study the relationship between the gamma field parameters.

4.1 The Mean and Standard Deviation of the Extreme Value

With $L = 1$ and $\beta = 1$, the mean and standard deviation of a univariate gamma distribution reduces to α and $\sqrt{\alpha}$, respectively. Therefore, α and $\sqrt{\alpha}$ are used to normalize the mean and standard deviation of the EV, respectively.

Figures 1 and 2 show the change in the normalized extreme value mean with respect to the shape parameter and the correlation length for both exponential and triangular correlation functions, respectively. As α increases, the normalized mean of the extreme value decreases for a given correlation length (Figure 1). This is because the rate of increase in the absolute value of the extreme value mean is lower than the linear increase in α. Consequently, there is a decrease in the normalized extreme value mean. An alternative interpretation for this is provided by the fact that a gamma distribution with a bigger shape parameter has a lighter right tail than one with a smaller shape parameter. This means there is a relatively lower probability of getting more extreme values from a gamma distribution with a larger α. The normalized mean for a given θ continues to decrease and eventually converges to 1 when α becomes very large. It is an established fact that a Gaussian distribution can be used to approximate a gamma distribution with a large shape parameter. On the other hand, Figure 1 also shows that the exponential correlation function gives a slightly higher normalized extreme value mean than the triangular correlation function. This is explained by the different rates of decay of both correlation functions.

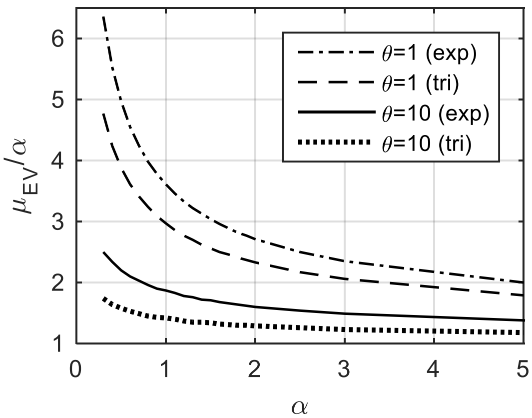

Figure 1: Normalized extreme value mean against the shape parameter for exponential and triangular correlation functions.

Figure 2 shows that the decrease trend is also dependent upon the correlation length. As the correlation length increases for a given α, the normalized extreme value decreases. For a fixed α in Figure 2, increasing θ implies less randomness and decreased statistical distance between any two points in the field. In other words, decreasing correlation length of a random field increases the effective number of independent random variables in the field. As the correlation length becomes very small, the correlation between any two given points reduces further. Hence there is a higher probability of getting more extreme values. Similar to the trend in Figure 1, the exponential correlation function gives a higher normalized extreme value mean than the triangular correlation function.

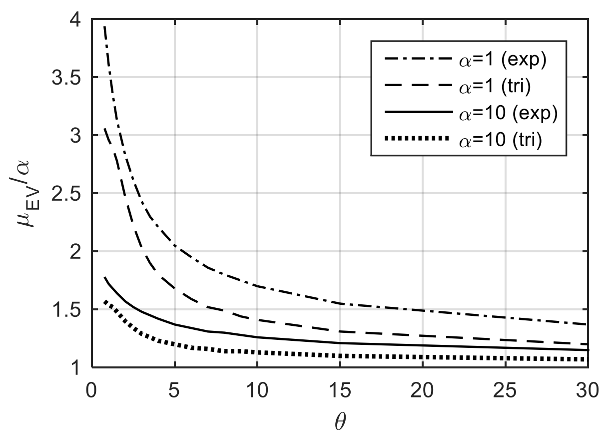

Figure 2: Normalized extreme value mean against correlation length for exponential and triangular correlation functions.

From Figures 3 and 4, the trends of the normalized standard deviation of the extreme value are slightly different than that of the normalized mean. As shown in Figure 3, for a given θ, as the shape parameter increases, the normalized standard deviation decreases. However, in Figure 4, as the correlation length increases, the normalized standard deviation for the triangular correlation function increases at first when the correlation length is small, peaks at a certain value and then starts to decline as the correlation length keeps increasing. The normalized standard deviation for the exponential correlation function tends to decrease or remain constant with correlation length, depending on α.

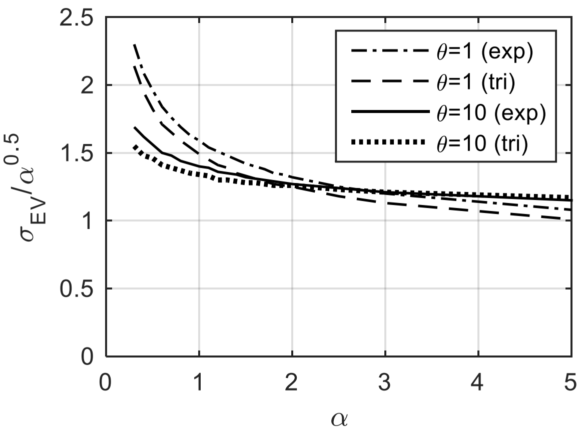

Figure 3: Normalized extreme value standard deviation against the shape parameter for exponential and triangular correlation functions.

In practice, the parameters of the field can be estimated from deterioration data. Once they are estimated, the extreme value statistics of the gamma field can be evaluated by using the algorithm proposed in this paper and such statistics may be subsequently used in reliability analysis.

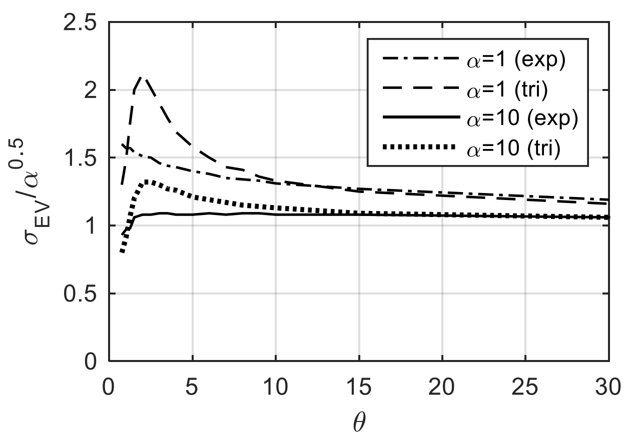

Figure 4: Normalized extreme value standard deviation against the correlation length for exponential and triangular correlation functions.

4.2 Trends of Generalized Extreme Value (GEV) Distribution Shape Parameter

The simulated extreme values allow one to study the EV distribution types with different field parameters. In this section specifically, the relationship between GEV shape parameter and parameters of the gamma field are studied for both exponential and triangular correlation functions. Results are presented in Figures 5 and 6.

Figure 5 shows that as α increases for a highly correlated field e.g. $\theta = 100$, the extreme value shape parameter k decreases and actually changes signs from a positive value to a negative value. At smaller values of α, the extreme value distribution obtained is type II, and transitions to type I and eventually turns to type III as α becomes much larger. Therefore, the extreme value distribution type is a function of α. For a less correlated gamma field $\theta = 1$, the triangular correlation function results in a type II GEV distribution while the exponential function gives a type III distribution, except at lower values of shape parameter.

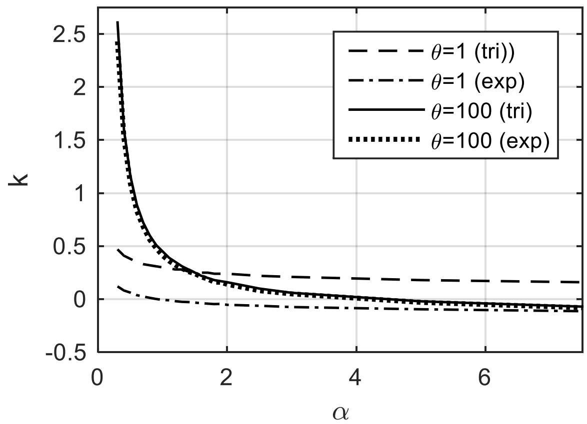

Figure 5: Shape parameter of the generalized extreme value distribution against univariate shape parameter for exponential and triangular correlation functions.

The GEV shape parameter is found to increase with correlation function for $\alpha = 0.5$ and it stays in the positive zone all through for both triangular and exponential correlation functions and therefore, the extreme value distribution type is type II (Figure 6). However, for $\alpha = 10$, the GEV shape parameter for the triangular correlation function increases, peaks at a certain value and then decreases as the gamma field becomes more spatially correlated. Therefore, the extreme value distribution type obtained in this case changes from type II to type I and finally ends with type III. The same trend is observed for $\alpha = 100$, which is close enough to a Gaussian field. For the exponential function, the GEV shape parameter remains negative for $\alpha = 10$ and $\alpha = 100$, revealing a Type III, or Weibull distribution.

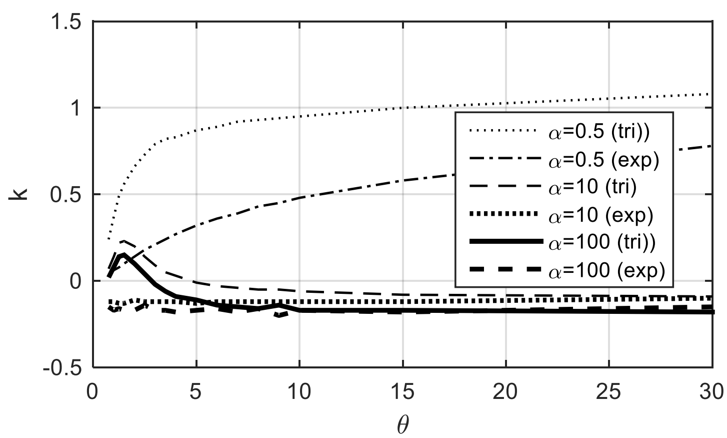

Figure 6: Shape parameter of the generalized extreme value distribution against correlation length for exponential and triangular correlation functions.

5 Conclusions

This paper presents a copula-based gamma field to model the spatial dependence. To study the extreme value behaviour, a Karhunen-Loève expansion based simulation algorithm is proposed. Based on this method, it is found that the normalized extreme value mean decreases as the shape parameter and correlation length of the gamma field increases for both exponential and triangular correlation functions, whereas the normalized standard deviation sometimes shows non-monotonic trend with the correlation length. The extreme value distribution type obtained is a function of the parameters of the gamma field. Moreover, the distribution type is also sensitive to the correlation function of the gamma field. These results are useful for risk and reliability analysis of degraded structures with significant spatial uncertainty.

References

Adegbola, A., Yuan, X.-X. 2017. Karhunen-Loève Expansion for Extreme Values of a Homogeneous Copula-Based Gamma Field. The 2017 Annual European Safety and Reliability (ESREL) Conference. June 18-22, 2017, Portoroz, Slovenia.

Ghanem, R. & Spanos, P. 1991. *Stochastic finite element: A spectral approach*. Dover Publications, New York, USA.

Joe, H. 2001. *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC.

Kotz, S., Balakrishnan, N. & Johnson, N. L. 2000. *Continuous Multivariate Distributions*, Vol. 1. John Wiley & Sons.

Shemehsavar S. 2014. A bivariate gamma model for a latent degradation process. *Communications in Statistics – Theory and Methods* 43: 1924-1938.

Van Noortwijk, J. 2009, A Survey of the application of Gamma processes in maintenance. *Reliability Engineering and System Safety*. 94(1):2-21.

Wang, X., Balakrishnan, N., Guo, B. & Jiang, P. 2015, Residual life estimation based on bivariate non-stationary gamma degradation process *J. Stat. Comp. Sim.*, **85**(2): 405-421.

Yuan, X.-X. 2016. Principles and Guidelines of Deterioration Modeling for Water and Wastewater Assets. *Infrastructure Asset Management*. DOI: 10.1680/jinam.16.00017, in press.