Available online at http://scik.org

J. Math. Comput. Sci. 11 (2021), No. 1, 577-584

https://doi.org/10.28919/jmcs/5154

ISSN: 1927-5307

MAX-STABLE PROCESSES WITH GEOMETRIC GAUSSIAN MODEL ON OCEAN WAVE HEIGHT DATA

ARIEF RACHMAN HAKIM*, BUDI WARSITO, HASBI YASIN

Department of Statistics, Diponegoro University, Semarang 50275, Indonesia

Abstract: In Spatial extreme value (SEV) is a method used to model an extreme event in several locations where there are dependencies between these locations. This method is a development of the Extreme Value Theory which is used in univariate cases. One of the approaches used in SEV is the max-stable process. Several models used in the max-stable process are the Smith, Schlater, Brown Resnick and Gaussian Geometric models. In general, these models have a Generalized Extreme Value (GEV) distribution. This study uses rainfall data in Central Java, then the extreme data is selected by the Maxima block method. The basic principle of maxima block is that extreme data is selected from the maximum data in each predefined block. The next step is modeling the extreme data with a geometric-Gaussian model. This method is a development of the Smith and Schlater models. The model obtained is then used to predict extreme rainfall using the return level. The result is that the maximum extreme rainfall in the next two periods is Pekalongan Station 1.16, Rembang station 0.8 and Semarang station 1.29 with an RMSE of 1.9.

Keywords: max-stable; spatial extreme value; height wave; geometric-Gaussian.

2010 AMS Subject Classification: 60E05.

*Corresponding author

E-mail address: arief.rachman@live.undip.ac.id

Received October 31, 2020

1. INTRODUCTION

Spatial extreme value (SEV) is a method used to model an extreme event in several locations where there are dependencies between these locations. This method is a development of the Extreme Value Theory which is used in univariate cases. One of the approaches used in SEV is the max-stable process. Max-stable processes are used to overcome SEV's main constraints regarding a flexible and inferential model. Several models used in the max-stable process are the Smith model developed by Smith (1990) [1] and applied to rainfall data in the UK, Schlater model was proposed by Schlater (2002) [2] who developed SEV modeling through the MSP approach based on Gaussian random field, Brown-Resnick model written by Kabluchko (2009) [3] by applying Brown-Resnick Process (Brown and Resnick, 1977) [4], Gaussian Geometric Model developed by Davison, Padoan and Ribatet (2010) [5]. The max-stable model has also been applied in Indonesia, one of which is carried out by Yasin, Warsito and Hakim (2019) [6], which has a good performance in modeling extreme rainfall. In general, MSP models have a Generalized Extreme Value (GEV) distribution. This research uses sea wave height data in Central Java, then the extreme data is selected by the Maxima block method. The basic principle of maxima block is that extreme data is selected from the maximum data in each predefined block. The next step is modeling extreme data and obtaining parameter estimates using the geometric-Gaussian model. This method is a development of the Smith and Schlater models. The model obtained is then used to predict extreme rainfall using the return level.

2. PRELIMINARIES

Max Stable process

Let X be a set of indexes and $\{Y_i(x)\}_{x\in X}, i=1,2,...,n$ where n is the independent replication of a continuous stochastic process. Assume that there is a continuous function where $a_n(x) > 0$ and $b_n(x) \in$ R such that :

$$Y(x)=\lim_{n\to\infty}\frac{\max_{i=1}^{n} Y_i(x)-b_n(x)}{a_n(x)}, \qquad n\to\infty, x\in X \tag{2.1}$$

where $Y_1,..,Y_n$ are independent replications of $Y,$ if the limit value exists (exist), then the process of limit *Y(x)* is called the max-stable process. There are two characteristics that follow the max-stable process. First, with one dimension, the marginal distribution follows the GEV distribution, namely $Y \sim GEV(\mu,\sigma,\xi)$ with the distribution function follow:

$$f(y;\mu,\sigma,\xi)=\exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]_+^{-\frac{1}{\xi}}\right\},-\infty<\mu,\xi<\infty,\sigma>0\,.$$

Second, the k-dimensional marginal distribution follows the multivariate extreme value distribution. In equation (2.1) the stochastic process $\{Y(x)\}_{x\in X}$ is the Max Stable Process (de Haan, 1984) [7]. if $a_n(x)= n$ and $b_n(x)= 0$, then Y(x) is also a simple Max Stable Process. $\{Z(x)\}_{x\in X}$ is a max-stable that has a Frechet unit margin with its distribution function $F(z)=\exp\left(-\frac{1}{z}\right),z>0$. This process can be obtained by standardizing $\{Y(x)\}_{x\in X}$ in order to obtain:

$$\{Z(x)\}_{x\in X}=\left\{1+\frac{\xi(x)\left(Y(x)-\mu(x)\right)}{\lambda(x)}\right\}_{+\quad x\in X}^{1/\xi(x)}, \tag{2.2}$$

Where $\mu(x),\xi(x)$ and $\lambda(x)>0$ is a continuous function. The Z process in equation (2.2) still includes the max-stable process (Padoan, Ribatet, Sisson, 2010) [5].

Gaussian Geometric Model

The Gaussian Geometric Model was developed by Davison, Padoan and Ribatet (2010) [5] by defining a dependency structure, which is $W_i(x)=\exp(\sigma\varepsilon(x)-\sigma^2/2)$ to form the following equation:

$$Z(x)=\max_i \xi_i \exp(\sigma\varepsilon(x)-\sigma^2/2), x\in X \tag{2.3}$$

ε_i is a standard Gaussian process (Ribatet) with zero mean with a correlation function component $\rho(h)$, Where $\varepsilon(0)=0$ and the Bivariate Comulative Distribution Function (CDF) model refers to the bivariate smith model with the equation:

$$P_r\left[Z\left(x_1\right)\le z_1, Z\left(x_1\right)\le z_2\right]=\exp\left[-\frac{1}{z_1}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_2}{z_1}\right)-\frac{1}{z_2}\Phi\left(\frac{a}{2}+\frac{1}{a}\log\frac{z_1}{z_2}\right)\right] \tag{2.4}$$

where Φ is Comulative Distribution Function normal standard and $a=\sigma\sqrt{2(1-\rho(h))}$. So that equation 2.4 turns into equation 2.5:

$$\begin{aligned}P_r\left[Z\left(x_1\right)\le z_1, Z\left(x_1\right)\le z_2\right]=\exp\Bigg[&-\frac{1}{z_1}\Phi\left(\frac{\sigma\sqrt{2(1-\rho(h))}}{2}+\frac{1}{\sigma\sqrt{2(1-\rho(h))}}\log\frac{z_2}{z_1}\right)\\&-\frac{1}{z_2}\Phi\left(\frac{\sigma\sqrt{2(1-\rho(h))}}{2}+\frac{1}{\sigma\sqrt{2(1-\rho(h))}}\log\frac{z_1}{z_2}\right)\Bigg]\end{aligned} \tag{2.5}$$

In general the Gaussian Geometric CDF is

$$\begin{aligned}F\left(z(x_i),z(x_j)\right)=\exp\Bigg[&-\frac{1}{z(x_i)}\Phi\left(\frac{\sigma\sqrt{2(1-\rho(h))}}{2}+\frac{1}{\sigma\sqrt{2(1-\rho(h))}}\log\frac{z(x_j)}{z(x_i)}\right)\\&-\frac{1}{z(x_j)}\Phi\left(\frac{\sigma\sqrt{2(1-\rho(h))}}{2}+\frac{1}{\sigma\sqrt{2(1-\rho(h))}}\log\frac{z(x_i)}{z(x_j)}\right)\Bigg]\end{aligned} \tag{2.6}$$

with $z\left(x_i\right)$ and $z\left(x_j\right)$ is a max-stable process with Frechet margins at location-i and location-j.. Whereas Φ is the normal standard cumulative distribution function and $\rho(h)$ is the correlation function as in the Schlather model which consists of Whittle-Matern, Cauchy, Powered Exponential and Bessel, while h is the euclid distance between location 1

The bivariate PDF function for the Gaussian Geometric model is obtained from the derivative of the bivariate CDF function by simplifying equation 2.6 to equation 2.7 as follows:

$$F\left(z(x_i),z(x_j)\right)=\exp\left[-\frac{\Phi(w(h))}{z(x_i)}-\frac{\Phi(v(h))}{z(x_j)}\right] \tag{2.7}$$

where

$$w(h)=\frac{a(h)}{2}+\frac{1}{a(h)}\log\left(\frac{z(x_j)}{z(x_i)}\right)$$

$$w(h)=\frac{a(h)}{2}+\frac{1}{a(h)}\log\left(\frac{z(x_i)}{z(x_j)}\right)$$

$$a(h)=\sigma\sqrt{2(1-\rho(h))}$$

So the bivariate PDF form for the Gaussian Geometric model is

$$\begin{aligned} F\left(z(x_i), z(x_j)\right) &= \exp\left[-\frac{\Phi(w(h))}{z(x_i)} - \frac{\Phi(v(h))}{z(x_j)}\right] \\ &\times\left\{\left(\frac{\Phi(w(h))}{z^2(x_i)} + \frac{\varphi(w(h))}{a(h)z^2(x_i)} - \frac{\varphi(v(h))}{a(h)z(x_i)z(x_j)}\right)\right. \\ &\times\left(\frac{\Phi(w(h))}{z^2(x_j)} + \frac{\varphi(w(h))}{a(h)z^2(x_j)} - \frac{\varphi(v(h))}{a(h)z(x_i)z(x_j)}\right) \\ &\left.+\left(\frac{v\varphi(w(h))}{a^2(h)z^2(x_i)z(x_j)} + \frac{v\varphi(w(h))}{a^2(h)z(x_i)z^2(x_j)}\right)\right\} \end{aligned} \tag{2.8}$$

Parameter Estimation for Gaussian Geometric using Maximum Pairwise Likelihood Estimation, this estimation method is a parameter that uses the pairwise density function of two variables (PDF). The basic principle in MPLE is to create the first derivative of each parameter and then equalize to zero. In the context of the spatial model the GEV distribution is defined as follows:

$$GEV\left(\mu(s), \sigma(s), \xi(s)\right) \tag{2.9}$$

where the GEV distribution parameters follow the trend surface model which has the form of multiple regression model equations with coordinates (latitude) and longitude (longitude) as explanatory variables. The trend surface model is stated as follows [8]:

$$\begin{aligned} \hat{\mu}(s) &= \beta_{\mu,0} + \beta_{\mu,1}\, longitude(s) + \beta_{\mu,2}\, latitude(s) \\ \hat{\sigma}(s) &= \beta_{\sigma,0} + \beta_{\sigma,1}\, longitude(s) + \beta_{\sigma,2}\, latitude(s) \\ \hat{\xi}(s) &= \beta_{\xi,0} \end{aligned} \tag{2.10}$$

parameter estimation process β_{μ}, β_{σ}, and β_{ξ} based on PDF of each Max-Stable model using MPLE.

3. MAIN RESULTS

Parameter estimates for $\hat{\mu}, \hat{\sigma}, \hat{\xi}$ univariate each model using MLE and solved with the numeric method using BFGS (Broyden Fletcher Goldfarb Shanno). Count process using R software.

Table 1 Estimates Parameter GEV Univariate

Station	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$
Pekalongan	0,826	0,44	0,012
Rembang	0,92	0,78	-0,098
Semarang	0,435	0,358	0,53

Then, let transformation extreme Height wave data which obtained from block maxima into Frechet distribution using transformation Z:

$$Z(s_i) = \left(1 + \xi \frac{y - \mu}{\sigma}\right)^{1/\xi}$$

where y is extreme value sample, s_i shows the location of the Height wave observation stations and $\hat{\mu}, \hat{\sigma}, \hat{\xi}$ are GEV parameter.

The best model selected based on Takeuchi Information Criterion (TIC). The smallest TIC value is 1213.43 from the combination Trend Surface model that is

The trend surface model obtained from the max-stable smith model is as follows:

$$\hat{\mu}(s) = -27{,}176 - 4{,}158\,\text{latitude}(s)$$
$$\hat{\sigma}(s) = -49{,}23 - 7.423\,\text{latitude}(s)$$
$$\hat{\xi}(s) = 1{,}365$$

Then obtained parameter estimates that differ for each location, where parameter $\hat{\xi}(s)$ are constant in each location, shown in Table 2:

Table 2. Smith Model Parameter Estimates

Rain Monitoring Post	$\hat{\mu}(s)$	$\hat{\sigma}(s)$	$\hat{\xi}(s)$
Pekalongan Station	1,34	1,67	1,365
Rembang Station	0,70	0,53	1,365
Semarang Station	1,72	2,36	1,365

To calculate estimates of extreme height wave predictions based on a certain time period, the Return level is used. The estimation results of parameter $\hat{\mu}(s), \hat{\sigma}(s), \hat{\xi}(s)$ on each model are then used to calculate the predicted return level of extreme height wave at the three locations monitoring stations. When calculating the return level, it is stipulated that the return period has a value that is inversely proportional to the probability value, $T = 1/p$ same as $p = 1/T$, where the maximum probability is $p = 1$. Prediction of Return level is not possible to calculate the first period and it is possible to calculate it from the second period. The last Step is computed to predict Return Level of extreme height wave for each observation station, it will predict for 1, 2 and 3 years. The results of parameter estimation for each location in table 2 are then used to calculate the return level with the following formula:

$$z_p(s) = \hat{\mu}(s) - \frac{\hat{\sigma}(s)}{\hat{\xi}(s)}\left(1 - \left[-\ln\left(1 - \frac{1}{T}\right)\right]^{-\xi(s)}\right).$$

where $T = 1\,\text{year} \times 12\,(\text{number of block}) = 12$. Return level calculation in return periods 2 month for Pekalongan Station 1,16 meter, Rembang station 0,8 meter and semarang station 1,29 meter.

For model validation, the Root Mean Square Error (RMSE) benchmark was used to measure the performance of Brown-Resnick models. The RMSE formula for testing data is as follows:

$$RMSE = \sqrt{\frac{1}{S}\sum_{i=1}^{S}\left(Y_i - \hat{Y}_i\right)^2},$$

where S is the number of locations, which Y_i is the actual observed value obtained from testing data and $\hat{Y}_i$ is the estimated or predicted value on the return period (T). The RMSE value is 1.9.

CONCLUSION

Further research is needed by including the element of time in the extreme spatial modeling of value so that a space-time model will be formed and may produce more accurate predictions. The conclusion of this research is return level prediction with the next two periods obtained results for each station, namely Pekalongan Station 1.16 meters, Rembang station 0.8 meters and Semarang station 1.29 meters. This model is good enough if it is used to predict with a short period of time, if it is used to predict with a large enough T (long period) then the results tend to be even greater.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] R.L. Smith, Max-stable Processes and spatial extremes, https://www.researchgate.net/profile/Stilian_Stoev2/publication/271095588_Upper_bounds_on_value-at-risk_for_the_maximum_portfolio_loss/links/564b824b08aeab8ed5e7694a.pdf, 1990.

[2] M. Schlather, Models for Stationary Max-Stable Random Fields. Extremes 5 (2002), 33–44.

[3] Z. Kabluchko, M. Schlather, L. De Haan, Stationary max-stable fields associated to negative definite functions. Ann. Probab. 37(5) (2009), 2042-2065.

[4] B.M. Brown, S.I. Resnick, Extreme values of independent stochastic processes, J. Appl. Probab. 14 (1977), 732–739.

[5] A.C. Davison, S. Padoan, M. Ribatet, Statistical Modeling of Spatial Extremes, Stat. Sci. 27(2) (2012), 161-186.

[6] H. Yasin, B. Warsito, A.R. Hakim, Prediksi curah hujan ekstrem di kota semarang menggunakan spatial extreme value dengan pendekatan max stable process (MSP), Media Stat. 12 (2019), 39-49.

[7] L. De Haan, A spectral representation for max-stable processes. Ann. Probab. 12(4) (1984), 1194-1204.

[8] H. Yasin, A.R. Hakim, B. Warsito, R. Santoso, Extreme rainfall prediction using spatial extreme value by Max Stable Process (MSP) Smith model approach, J. Phys.: Conf. Ser. 1217 (2019), 012110.