

# PEMODELAN KERUGIAN BENCANA BANJIR AKIBAT CURAH HUJAN EKSTREM MENGGUNAKAN EVT DAN COPULA

Ian Surya Prayoga<sup>1</sup>, Atina Ahdika<sup>2</sup>

<sup>1,2</sup> Program Studi Statistika, Fakultas MIPA, Universitas Islam Indonesia  
e-mail: <sup>1</sup>iansurya17@gmail.com, <sup>2</sup>atina.a@uii.ac.id

## Abstrak

Curah hujan ekstrem pada umumnya dapat mengakibatkan kerugian seperti banjir, tanah longsor, dan gagal panen. Pemodelan kerugian bencana banjir digunakan untuk memperkirakan seberapa parah kerusakan yang dialami ketika terjadi bencana banjir akibat curah hujan ekstrem. Penelitian ini bertujuan untuk memodelkan kerugian bencana banjir terhadap kerusakan rumah menggunakan metode *Extreme Value Theory (EVT)* dan copula. Metode EVT digunakan untuk memodelkan distribusi curah hujan dan rumah rusak, sedangkan copula digunakan untuk mengidentifikasi struktur dependensi antara curah hujan dengan kerusakan rumah yang ditimbulkan. Hasil dari penelitian ini didapatkan distribusi terbaik untuk kerusakan rumah di Jawa Barat, Jawa Tengah, dan Jawa Timur adalah distribusi *Generalized Extreme Value (GEV)* serta model copula terbaik yang menggambarkan dependensi antara curah hujan dan kerusakan rumah adalah copula Frank. Nilai parameter  $\hat{\theta}$  copula Frank Provinsi Jawa Barat sebesar 1.4999840280, Jawa Tengah sebesar -0.5816995330, dan Jawa Timur sebesar -0.8648329345. Parameter copula Frank bernilai positif (negatif) menunjukkan hubungan yang sifatnya positif (negatif) antara curah hujan ekstrem dan rumah rusak. Hasil pemodelan menunjukkan bahwa terdapat hubungan positif antara curah hujan dan rumah rusak di Provinsi Jawa Barat dan Jawa Tengah, serta hubungan negatif antara kedua variabel di Provinsi Jawa Timur. Hal tersebut menunjukkan bahwa terdapat kemungkinan penyebab lain yang lebih berpengaruh terhadap kerusakan rumah di Provinsi Jawa Timur dibandingkan curah hujan.

**Kata kunci:** Curah Hujan Ekstrem, Banjir, *Extreme Value Theory*, Copula

## Abstract

*In general, extreme rainfall can result in losses such as floods, landslides, and crop failure. Flood loss modeling is used to estimate how much damage is experienced when a flood occurs due to extreme rainfall. This study aims to model the losses of floods on house damage using the Extreme Value Theory (EVT) and copula methods. The EVT method is used to model the distribution of rainfall and damaged houses, while the copula is used to identify the dependency structure between the rainfall and the damage to the houses it causes. The results of this study show that the best distribution for house damage in West Java, Central Java and East Java is the Generalized Extreme Value (GEV) distribution and the best copula model that describes the dependency between rainfall and house damage is Frank copula. The value of the Frank copula parameter in West Java Province is 1.4999840280, in Central Java is -0.5816995330, and in East Java is -0.8648329345. The positive (negative) value of Frank copula parameter indicating a positive (negative) relationship between extreme rainfall and damaged houses. The modeling results show that there is a positive relationship between rainfall and damaged houses in West Java and Central Java Provinces, as well as a negative relationship between the two variables in East Java Province. This indicates that there are other possible causes that are more influencing the damage to houses in East Java Province than rainfall.*

**Keywords:** *Extreme rainfall, floods, Extreme Value Theory, Copula*

## PENDAHULUAN

Pada saat ini pola iklim global maupun regional telah mengalami banyak perubahan. Salah satu yang menyebabkan perubahan pola iklim ini yaitu efek rumah kaca. Efek rumah kaca menyebabkan gas karbondioksida menumpuk pada lapisan ozon sehingga menyebabkan panas bumi yang berasal dari cahaya matahari tidak dapat keluar dari permukaan bumi. Akibatnya, suhu rata – rata permukaan bumi pada berbagai wilayah mengalami kenaikan. Kenaikan suhu rerata harian ini dapat mengakibatkan perubahan pada pola curah hujan yang menjadi tinggi bahkan menjadi ekstrem. Berdasarkan Peraturan Kepala Badan Meteorologi Klimatologi, dan Geofisika Nomor: Kep. 009 Tahun 2010, curah hujan ekstrem berupa hujan lebat dan hujan es. Hujan lebat adalah curah hujan dengan intensitas 20 mm/jam, sedangkan hujan es adalah curah hujan yang berbentuk butiran es dengan garis tengah paling rendah 5 mm. Curah hujan yang ekstrem ini dapat menjadikan berbagai bencana alam salah satunya bencana banjir. Contohnya ketika curah hujan ekstrem terjadi pada daerah pertanian akan menyebabkan petani mengalami penurunan produktifitas panen atau bahkan mengalami gagal panen . Sedangkan jika curah hujan ekstrem terjadi pada daerah padat penduduk yang wilayahnya kurang memiliki daerah resapan air, akan menyebabkan banjir yang kemudian bisa menyebabkan kerusakan bangunan seperti rumah warga atau bangunan pertokoan. Oleh karena itu, perlu dilakukan pemodelan untuk digunakan meramalkan kejadian curah hujan ekstrem agar bisa meminimalisir dampak yang disebabkan dari curah hujan ekstrem tersebut. Hasil peramalan ini juga bisa digunakan untuk mengoptimalkan bantuan – bantuan yang akan diberikan kepada pihak terdampak jika terjadi bencana alam.

Dalam ilmu statistika, salah satu metode yang digunakan untuk mengidentifikasi kejadian ekstreme yaitu dengan *Extreme Value Theory* (EVT). Kejadian yang bersifat ekstrem pada data *heavytail* dapat diramalkan menggunakan

EVT. EVT juga dapat menjelaskan kerugian kejadian ekstrem yang kurang cocok dimodelkan dengan pendekatan biasa. *Extreme Value Theory* mengkaji perilaku stokastik variabel random secara minimum maupun maksimum (Kotz dan Nadarajah, 2000). Metode *extreme value theory* bertujuan untuk mengestimasi peluang suatu kejadian ekstrem dengan melihat ekor dari suatu distribusi berdasarkan nilai-nilai ekstrem yang diperoleh. Metode yang digunakan dalam mengidentifikasi suatu nilai ekstrem menggunakan *extreme value theory* yaitu metode *Block Maxima* (BM) dan metode *Peaks Over Threshold* (POT). Metode *Block Maxima* (BM) yaitu mengambil nilai maksimum dalam satu periode yang disebut sebagai blok dan metode *Peaks Over Threshold* (POT), yaitu mengambil nilai yang melewati suatu nilai threshold (McNeill, 1999).

Pada penelitian ini penulis akan melakukan pemodelan kerugian bencana banjir akibat curah hujan ekstrem di Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur menggunakan *Extreme Value Theory* (EVT) dan pendekatan copula. EVT dalam penelitian ini digunakan untuk memodelkan curah hujan ekstrem. Curah hujan memiliki ekor distribusi yang gemuk (*heavytail*) sehingga perlu menggunakan EVT dalam pemodelannya. Sedangkan copula digunakan untuk menggambarkan ketergantungan antar variabel curah hujan dan kerugian bencana banjir. Dalam penelitian ini kerugian akibat banjir yang dimaksud yaitu rumah rusak. Tiga Provinsi tersebut merupakan salah satu daerah yang memiliki tingkat kasus bencana banjir yang tinggi, sehingga apabila terjadi hujan ekstrem yang berkesinambungan tentu saja berpengaruh terhadap kerusakan atau kerugian.

Berdasarkan rumusan masalah diatas, penelitian ini memiliki tujuan sebagai berikut:

- a. Mendapat estimasi parameter pada pemodelan *Spatial Extreme Value* dengan pendekatan Copula spasial untuk curah hujan antara Jawa Barat, Jawa Tengah, dan Jawa Timur.

- b. Mengetahui dependensi spasial curah hujan antara Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur.
- c. Mendapat model kerugian akibat bencana banjir di Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur berdasarkan pemodelan *Spatial Extreme Value* dengan pendekatan Copula.

**METODE**

**Curah Hujan**

Menurut BMKG (BMKG, 2020), pengertian curah hujan adalah banyaknya curah hujan yang jatuh ke tanah dalam kurun waktu tertentu, dalam satuan milimeter (mm) di atas horizontal. Curah hujan juga dapat diartikan sebagai ketinggian air hujan yang tertampung dalam alat pengukur hujan, yang mempunyai permukaan datar, tidak menyerap, tidak menyerap, dan tidak mengalir. Nilai curah hujan 1 (satu) milimeter artinya satu milimeter air hujan atau satu liter air hujan ditampung di area seluas satu meter persegi di daerah datar. (BMKG, 2020).

**Banjir**

Banjir dapat diartikan sebagai tempat yang tergenang karena luapan yang meluap melebihi kapasitas pengolahan air suatu daerah atau tempat tertentu, dan menimbulkan banyak kerugian, seperti kerugian material, sosial dan ekonomi. Banjir juga dapat dijelaskan sebagai kejadian: lahan kering normal terendam karena curah hujan yang tinggi, dan biasanya area di wilayah tersebut rendah hingga cekung. Pada saat yang sama, banjir juga dapat disebabkan oleh limpasan limpasan, yang limpasannya melebihi kapasitas sistem drainase atau sistem aliran sungai. Selain itu, rendahnya kemampuan air untuk masuk ke dalam tanah juga dapat menyebabkan banjir yang menyebabkan tanah tidak dapat lagi menyerap air.

**EVT (*Extreme Value Theory*)**

EVT (*Extreme Value Theory*) adalah salah satu ilmu dalam statistika yang digunakan untuk melihat pola atau perilaku ekor (*tail*) dari suatu distribusi data ekstrem

(Amalia, 2017). EVT digunakan untuk memodelkan peristiwa - peristiwa yang bersifat ekstrem, dimana peristiwa tersebut jarang terjadi dan berlangsung dalam waktu singkat akan tetapi peristiwa tersebut memberikan dampak kerugian yang cukup besar. EVT biasanya diterapkan pada kejadian yang besar dalam peristiwa alam seperti curah hujan, banjir, dan polusi udara.

Terdapat dua pendekatan yang digunakan untuk melakukan analisis menggunakan EVT yaitu dengan metode BM (*Block Maxima*) dan metode POT (*Peaks Over Threshold*). Metode BM (*Block Maxima*) yaitu pendekatan yang digunakan untuk mengidentifikasi nilai ekstrem dengan mengambil data atau nilai maksimum dari suatu observasi dalam satu periode tertentu yang disebut sebagai blok. Pendekatan ini hanya menghasilkan satu nilai ekstrem pada setiap blok. Sedangkan metode *Peaks Over Threshold* (POT), yaitu pendekatan yang mengidentifikasi nilai ekstrem dengan mengambil nilai yang melewati suatu nilai batas (*threshold*) tertentu (Rinaldi, 2016).

**Distribusi *Generalized Extreme Value***

Metode *block maxima* mengaplikasikan teorema Fisher dan Tippet (1928) dalam Gilli dan Kellezi (2006), dimana dalam teorema tersebut menyatakan bahwa data sampel nilai ekstrem yang diambil dengan metode BM akan berdistribusi Gumble, Frechet atau Weibull. Kombinasi dari ketiga distribusi ini masuk ke dalam satu keluarga disebut sebagai distribusi *Generalized Extreme Value* (GEV). Menurut Mallor, Nualart, dan Omei (2009) *Generalized Extreme Value* (GEV) memiliki *cumulative distribution function* (CDF) seperti persamaan (1) berikut :

$$F(x; \mu, \sigma; \xi) = \begin{cases} \exp\left\{-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{\frac{1}{\xi}}\right\} & , -\infty < x < \infty, \xi \neq 0, -\infty < \mu < \infty, \sigma > 0 \\ \exp\left\{-\exp\left(\frac{x-\mu}{\sigma}\right)\right\} & , -\infty < x < \infty, \xi = 0, -\infty < \mu < \infty, \sigma > 0 \end{cases} \tag{1}$$

Probability distribution function (pdf) untuk distribusi GEV seperti persamaan (2).

$$f(x; \mu, \sigma, \xi) = \begin{cases} \frac{1}{\sigma} \left\{ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right\}^{\frac{1}{\xi}} \exp \left\{ - \left( 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right)^{\frac{1}{\xi}} \right\}, & \xi \neq 0, 1 + \xi \left( \frac{x - \mu}{\sigma} \right) > 0 \\ \frac{1}{\sigma} \exp \left\{ \frac{x - \mu}{\sigma} \right\} \exp \left\{ - \exp \left( \frac{x - \mu}{\sigma} \right) \right\}, & \xi = 0 \end{cases} \quad (2)$$

Dimana  $x$  adalah nilai ekstrem yang diperoleh dari *block maxima* dengan  $-\infty < x < \infty$ ,  $\mu$  adalah parameter lokasi (location) dengan  $-\infty < \mu < \infty$ ,  $\sigma$  adalah parameter skala (scale) dengan  $\sigma > 0$  dan  $\xi$  adalah parameter bentuk (shape) dengan  $-\infty < \xi < \infty$ .

### Uji Kesesuaian Distribusi

Tes Anderson Darling adalah tes yang digunakan untuk menentukan apakah data mengikuti distribusi tertentu. Tes Anderson Darling dengan sebuah program dapat digunakan untuk menguji kesesuaian distribusi GEV dengan data yang ekstrem (Engmann & Cousineau, 2011).

Berikut merupakan hipotesis pada Uji Anderson Darling

#### 1. Uji Hipotesis :

$H_0: F(x) = F^*(x)$  (Data mengikuti distribusi teoritis  $F^*(x)$ )

$H_1: F(x) \neq F^*(x)$  (Data mengikuti distribusi teoritis  $F^*(x)$ )

#### 2. Statistik Uji :

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) (\ln(F^*(x_i))) + \ln(1 - (F^*(x_{n+1-i}))) \quad (3)$$

(3)

Keterangan:

$F(x)$  : fungsi distribusi kumulatif data sampel

$F^*(x)$  : fungsi distribusi kumulatif

$n$  : ukuran sampel

#### 3. Daerah Kritis

Tolak  $H_0$  jika  $p\text{-value} < \alpha$  atau jika nilai  $AD_{hitung} > AD_{tabel}$

#### 4. Kesimpulan :

Kesimpulan didapatkan dengan cara membandingkan nilai  $AD_{hitung}$  dengan nilai  $AD_{tabel}$  atau dengan membandingkan nilai  $p\text{-value}$  dengan tingkat signifikansi ( $\alpha$ ).

### Max Stable Process

Max Stable Processes (MSP) merupakan perluasan dari distribusi *multivariate extreme value* ke dimensi tak hingga (*infinite dimension*). Suatu fungsi distribusi G dikatakan max stable jika dan hanya jika G berdistribusi GEV (Ramadhani I. R., 2015). Dalam metode MSP, sampel diambil dari nilai maksimum (*Maxima*) pada setiap lokasi (proses spasial) (Cooley, Nyckah, & Naveau, 2007). Dalam konsep spasial ekstrem terdapat dua metode pendekatan yaitu *max-stable* dan *copula* (Davison, Padoan, & Ribatet, 2012).

### Copula

Menurut teorema sklar, copula merupakan suatu fungsi yang menghubungkan fungsi distribusi multivariat dengan distribusi marginalnya (Nelsen & Flores, 2005). Copula terbagi menjadi dua macam families, yaitu elliptical copula dan archimedian copula. Untuk kasus *Spatial Extreme Model* copula yang dapat digunakan adalah elliptical copula karena keluarga copula elliptical ini mampu menggambarkan kekuatan ketergantungan antara pasangan variabel spasial (AghaKouchak, Easterling, Hsu, Schubert, & Sorooshian, 2013). Copula yang termasuk dalam elliptical copula adalah *Gaussian copula* dan *Student's t-copula*. Sedangkan Copula yang termasuk dalam archimedian yaitu Copula Joe, Copula Frank, Copula Gumble dan Copula Clayton.

Selain copula dengan satu parameter tersebut, terdapat juga copula dengan dua parameter. Copula dengan dua parameter, merupakan gabungan dua copula satu parameter. Strukturnya yang lebih fleksibel memungkinkan koefisien ketergantungan ekor atas dan bawah berbeda (Ramadhani I., 2019). Copula dengan dua parameter yaitu Copula BB1, BB6, BB7 dan BB8. Selain copula yang disebutkan di atas, terdapat juga copula yang dirotasi sebesar  $90^\circ$ ,  $180^\circ$  dan  $270^\circ$  dari copula Clayton, Gumbel, Joe, BB1, BB6, BB7 dan BB8. Ketika memutarnya  $180^\circ$ , hal tersebut berarti diperoleh copula yang sesuai dengan copula

aslinya, sedangkan rotasi 90° dan 270° memungkinkan untuk pemodelan ketergantungan negatif yang tidak mungkin dilakukan dengan copula standar yang tidak dirotasi.

Kelas copula yang memungkinkan untuk berbagai macam struktur ketergantungan diberikan oleh keluarga Archimedean (Schöolzel & Friederichs, 2008). Copula ini dapat dibangun dengan generator  $\phi$  sebagai berikut:

$$C_X(u, v) = \phi^{-1}(\phi(u) + \phi(v)) \quad , 0 \leq u, v \leq 1 \quad (4)$$

### Pemilihan Model Terbaik

Salah satu cara menentukan model terbaik yaitu menggunakan AIC (Akaike's Information Criterion). Saat memilih model terbaik, pertimbangkan jumlah parameter dalam model. Semakin kecil nilai AIC, semakin baik dan layak untuk digunakan model tersebut. Menurut Ligas dan Banasick (2012) Akaike Information Criterion (AIC) didefinisikan dengan persamaan (5) sebagai berikut:

$$AIC = -2\ell_p(\hat{\beta}) + 2q \quad (5)$$

Dimana  $\ell_p(\hat{\beta})$  adalah fungsi *ln pairwise likelihood*. Maximum Pairwise Likelihood Estimation (MPLE) adalah metode estimasi parameter yang menggunakan fungsi densitas pairwise/berpasangan dari dua variabel. Prinsip dasar dalam MPLE adalah membuat turunan pertama terhadap masing-masing parameter lalu menyamakan dengan nol. Fungsi *ln pairwise likelihood* didefinisikan melalui persamaan (6) berikut ini:

$$\ell_p(\hat{\beta}) = \text{fungsi } \ln \text{ pairwise likelihood}$$

$$\ell_p(\hat{\beta}) = \sum_{i=1}^k \sum_{j=1}^{m-1} \sum_{k=j+1}^m \ln(f(u_{ji}, u_{ki}; \hat{\beta})) \quad (6)$$

Nilai dari *Akaike Information Criterion* (AIC) digunakan untuk menentukan model *trend surface* terbaik. Persamaan model *trend surface* sebagai berikut:

$$\begin{aligned} \hat{\mu}(j) &= \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1} \text{longitude}(j) + \hat{\beta}_{\mu,2} \text{latitude}(j) \\ \hat{\sigma}(j) &= \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1} \text{longitude}(j) + \hat{\beta}_{\sigma,2} \text{latitude}(j) \\ \hat{\xi}(j) &= \hat{\beta}_{\xi,0} \end{aligned} \quad (7)$$

Model *trend surface* parameter  $\hat{\mu}(p), \hat{\sigma}(p), \hat{\xi}(p)$  adalah semua kombinasi model dengan komponen spasial garis lintang (*latitude*) dan bujur (*longitude*). Mulai dari model parameter hanya dengan satu variabel penjelas sampai model parameter dengan model kuadratik. Keseluruhan kombinasi model berlaku untuk semua parameter  $\hat{\mu}(p)$  dan  $\hat{\sigma}(p)$ , sedangkan untuk parameter  $\hat{\xi}(p)$  diasumsikan konstan (Hakim, 2016).

### Distribusi Poisson

Distribusi Poisson adalah distribusi diskrit, yang merupakan probabilitas keluaran sistematis dalam satuan standar tertentu x kali. Contoh aplikasi termasuk jumlah dering telepon per jam di kantor, jumlah goresan atau cacat pada permukaan produk, jumlah bakteri dalam kultur, dan penetapan nomor telepon yang salah. (Shrader, 1991). Fungsi kepadatan peluang untuk distribusi poisson dengan parameter  $\lambda$  adalah:

$$f(x; \lambda) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!}, & x = 0, 1, 2, \dots \\ 0, & x \text{ lainnya} \end{cases} \quad (7)$$

### Distribusi Binomial Negatif

Distribusi Binomial Negatif adalah distribusi hasil percobaan bernoulli yang diulang sampai mendapatkan sukses ke-k. Fungsi Probabilitas distribusi Binomial Negatif adalah sebagai berikut:

$$\Pr(X = x) = \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k}, \quad x = k, k+1, k+2, \dots \quad (8)$$

Dengan  $\Pr(X=x)$  adalah probabilitas terjadi sukses ke-k pada percobaan ke x dan  $\theta$  merupakan probabilitas sukses dari setiap percobaan konstan.

### MAPE (Mean Absolute Percentage Error)

MAPE digunakan untuk mengukur keakuratan nilai estimasi model, yang dinyatakan dalam bentuk persentase rata-rata kesalahan absolut, dan lebih banyak digunakan untuk membandingkan data

dengan rasio interval waktu yang berbeda (Robial, 2018). Skala pengkategorian nilai MAPE yang digunakan pada penelitian dan formula perhitungannya dijelaskan pada Tabel 1 dan persamaan (9) berikut:

$$MAPE = \left( \frac{1}{n} \right) \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (9)$$

Keterangan:

$A_t$  : Data aktual pada periode ke-t

$F_t$  : Data peramalan pada periode ke-t

$n$  : Jumlah periode peramalan

Seperti dijelaskan pada Tabel 1 suatu

Tabel 1. Ukuran Kesalahan

MAPE	Hasil Peramalan
<10%	Sangat Baik
10-20%	Baik
20-50%	Layak/Cukup
>50%	Buruk

model dikatakan memiliki kinerja sangat baik apabila memiliki nilai MAPE dibawah 10% dan memiliki kinerja baik apabila nilai MAPE berkisar antara 10% - 20% dan dikatakan layak apabila nilai MAPE berkisar antara 20% - 50% dan apabila lebih dari itu dikatakan berkinerja buruk (Razak, 2017). Akan tetapi MAPE memiliki kelemahan menjadi tak terbatas atau undefined jika ada nilai nol dalam seri.

### MASE (Mean Absolute Scaled Error)

MASE dapat digunakan untuk membandingkan peramalan pada satu seri ataupun untuk membandingkan hasil perkiraan beberapa seri. Nilai tipikal untuk

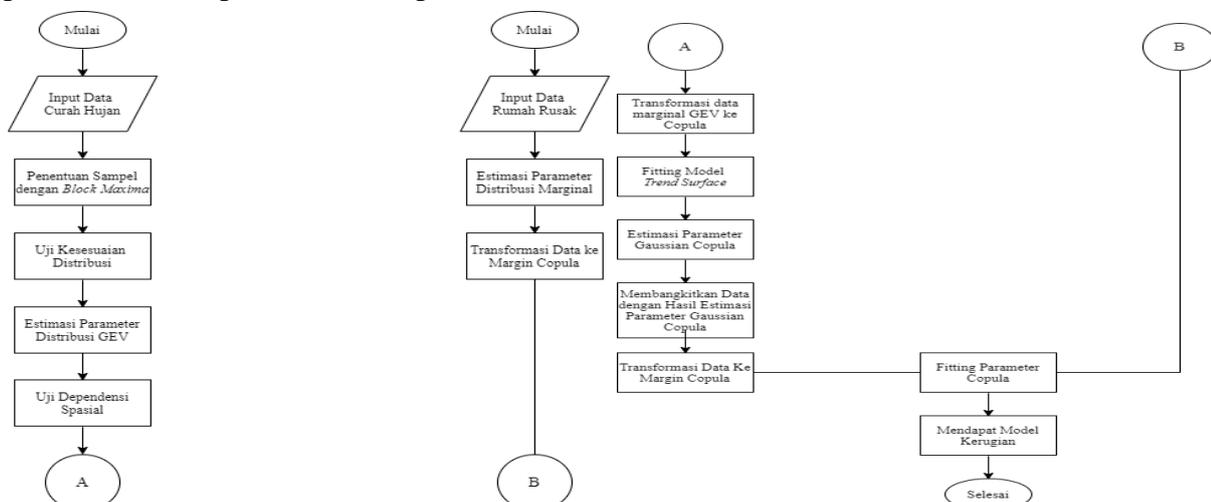
nilai satu-langkah MASE kurang dari satu, karena biasanya mungkin untuk mendapatkan perkiraan lebih akurat daripada metode naive.

$$MASE = \frac{mean|e_t|}{\frac{1}{n-1} \sum_{i=2}^n |Y_i - Y_{i-1}|} \quad (10)$$

Nilai MASE multistep sering lebih besar dari satu, karena menjadi lebih sulit untuk diperkirakan (Hyndman, 2006).

## METODOLOGI PENELITIAN

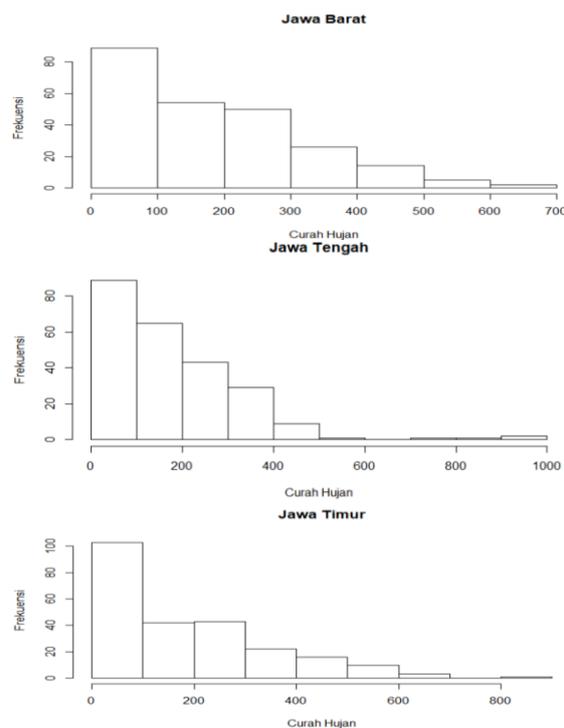
Penelitian ini menggunakan data sekunder yang diambil melalui 2 (dua) website. Data curah hujan diperoleh dari website dataonline.bmkg.go.id dan untuk data kerusakan rumah diperoleh dari website dibi.bnpb.go.id. Proses pengolahan data dilakukan dengan menggunakan software R 3.4.3. Sampel penelitian ini adalah data curah hujan bulanan dan data kerusakan rumah di Jawa Barat, Jawa Tengah dan Jawa Timur dari Maret 2000 hingga Desember 2019. Jumlah sampel yang diperoleh yaitu 242 data untuk masing-masing variabel. Setelah itu untuk menganalisis menggunakan extreme value theory, data sampel dikelompokkan menggunakan metode block maxima yaitu dengan membagi data ke dalam blok periode tertentu. Berdasarkan pengelompokan yang diacu dari BMKG, mengklasifikasikan pola hujan pada sebagian besar wilayah dipulau jawa merupakan pola hujan monsun, data dibagi ke dalam blok periode tiga bulanan, block



Gambar 1. Flowchart dari analisis penelitian ini

yang terbentuk yaitu Desember – Januari - Februari (DJF), Maret – April - Mei (MAM), Juni – Juli - Agustus (JJA) dan September - Oktober - November (SON). Selama periode sampel (2000-2019) terbentuk 79 blok. Dari satu blok diambil satu nilai ekstrem, nilai ekstrem yang diambil merupakan nilai maksimum dari masing-masing blok.

## HASIL DAN PEMBAHASAN



Gambar 2. Distribusi Curah Hujan di Jawa Barat, Jawa Tengah, dan Jawa Timur

Melalui analisis visual menggunakan histogram terlihat adanya data ekstrem dan pola data *heavytail* pada ketiga Provinsi tersebut yang ditunjukkan pada Gambar 2. Pada Gambar 2 terlihat bahwa pola dari distribusi data cenderung miring ke kiri, selain itu juga terlihat bahwa terdapat data yang memiliki frekuensi tinggi disekitar nilai nol, disisi lain juga terdapat data yang bernilai jauh lebih besar dari nilai nol akan tetapi memiliki frekuensi yang sangat kecil, sehingga mengindikasikan adanya pola data *heavytail*. Karena terdapat indikasi nilai ekstrem dan data *heavytail* ini menunjukkan bahwa data curah hujan tidak berdistribusi normal dan mengakibatkan penelitian ini menggunakan metode *extreme value theory*.

## Data Sampel dengan *Block Maxima*

Penentuan sampel dilakukan menggunakan metode *block maxima* yaitu dengan membagi data kedalam blok periode tertentu. Berdasarkan pengelompokan yang diacu dari BMKG, mengklasifikasikan pola hujan pada sebagian besar wilayah dipulau jawa merupakan pola hujan monsun, data dibagi ke dalam blok periode tiga bulanan, *block* yang terbentuk yaitu Desember – Januari - Februari (DJF), Maret – April - Mei (MAM), Juni – Juli - Agustus (JJA) dan September - Oktober - November (SON). Selama periode sampel (2000-2019) terbentuk 79 blok. Dari satu blok diambil satu nilai ekstrem, nilai ekstrem yang diambil merupakan nilai maksimum dari masing-masing blok.

## Uji Kesesuaian Distribusi

Selanjutnya dilakukan pengujian *Anderson Darling* untuk mengetahui kesesuaian distribusi. Pengujian hipotesis sebagai berikut:

$H_0: F_x = F^* x$  (Data mengikuti distribusi teoritis  $F^* x$ )

$H_1: F_x \neq F^* x$  (Data mengikuti distribusi teoritis  $F^* x$ )

Statistik uji yang digunakan yaitu pada persamaan (3) dengan menggunakan tingkat signifikansi  $\alpha = 5\%$ , tolak  $H_0$  jika

$AD_{hitung}$  lebih besar dari  $AD_{tabel}$ .

Tabel 2. Hasil Uji *Anderson Darling*

No	Provinsi	$AD_{hitung}$	P-value	Keputusan
1.	Jawa Barat	0.5653532	0.9344986	Gagal Tolak $H_0$
2.	Jawa Tengah	0.3106373	0.5876178	Gagal Tolak $H_0$
3.	Jawa Timur	0.9652988	0.9972561	Gagal Tolak $H_0$

Berdasarkan hasil pengujian menggunakan *Anderson Darling*, terlihat bahwa nilai p-value ketika provinsi tersebut lebih besar dari  $\alpha$  sehingga menghasilkan keputusan gagal tolak  $H_0$ . Hal ini berarti bahwa data sampel ekstrem blok tiga bulanan sudah mengikuti distribusi GEV.

### Estimasi Parameter GEV

Selanjutnya data sampel ekstrem curah hujan, digunakan untuk mengestimasi parameter distribusi GEV. Hasil estimasi parameter GEV disajikan pada Tabel 3.

Tabel 3. Hasil Estimasi Parameter Distribusi GEV

No	Provinsi	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\zeta}$
1.	Jawa Barat	226.021	141.5225	-0.2187897
	Jawa Tengah	0656	902	
2.	Jawa Tengah	179.299	128.5039	0.08802087
	Jawa Timur	86206	3438	
3.	Jawa Barat	181.837	162.3772	-
	Jawa Timur	45676	1168	0.04369635

Selanjutnya untuk melakukan analisis *spatial* dengan pendekatan copula, data perlu dilakukan transformasi ke distribusi frechet karena distribusi frechet memiliki ekor yang paling *heavytail* dibandingkan distribusi GEV yang lain dan copula lebih tepat diterapkan untuk kasus *heavytail*. Setelah data ditransformasi ke frechet data perlu ditransformasi lagi ke copula.

### Analisis Dependensi Spasial

Dalam penelitian ini sebanyak tiga lokasi atau wilayah yang diteliti. Hasil perhitungan koefisien ekstermal dapat dilihat pada tabel 4.

Tabel 4. Hasil Analisis Dependensi Spasial

No	Pasangan dua Provinsi	Jarak Euclid	Koefisien Ekstermal
1.	Jawa Barat - Jawa Tengah	2.823	1.424
2.	Jawa Tengah - Jawa Timur	2.403	1.027
3.	Jawa Barat - Jawa Timur	5.234	1.424

Nilai dari jarak euclid menggambarkan jarak antar lokasi dalam satuan derajat desimal. Satu derajat desimal sama dengan 111.319 km. Sehingga jarak euclid bernilai 2.83 sama dengan 2.83 x 111.319 yaitu 315.03277 km. Dari tiga pasang ini, nilai koefisien ekstermal berada pada rentang nilai 1.03- 1.42 sehingga dapat disimpulkan bahwa terdapat dependensi

*spatial* antar lokasi. Dari hasil tersebut bisa disimpulkan bahwa ketiga Provinsi tersebut memiliki ketergantungan spasial.

### Fitting Model Trend Surface

Pada penelitian ini, terdapat 9 kombinasi model *trend surface* dengan kombinasi variabel *longitude* dan *latitude*. Estimasi curah hujan ekstrem dilakukan menggunakan kombinasi model terbaik dari 9 kombinasi model yang ada. Hasil dari kombinasi model *trend surface* terdapat pada Tabel 5.

Tabel 5. Hasil Kombinasi Model Trend Surface

Kombinasi Ke	Kombinasi Model	AIC
1.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}u(j) + \hat{\beta}_{\mu,2}v(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}u(j) + \hat{\beta}_{\sigma,2}v(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-90.1559
2.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}u(j) + \hat{\beta}_{\mu,2}v(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}u(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-89.2062
3.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}u(j) + \hat{\beta}_{\mu,2}v(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}v(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-92.2881
4.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}u(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}u(j) + \hat{\beta}_{\sigma,2}v(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-97.6887
5.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}v(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}u(j) + \hat{\beta}_{\sigma,2}v(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-92.7640
6.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}u(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}u(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-92.5209
7.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}u(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}v(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-93.2563
8.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}v(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}u(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-93.0347
9.	$\hat{\mu}(j) = \hat{\beta}_{\mu,0} + \hat{\beta}_{\mu,1}v(j)$ $\hat{\sigma}(j) = \hat{\beta}_{\sigma,0} + \hat{\beta}_{\sigma,1}v(j)$ $\zeta(j) = \hat{\beta}_{\zeta,0}$	-91.4700

Berdasarkan tabel 5 diatas diketahui model yang terbaik adalah model ke-4 dengan AIC sebesar -97.6887. Dari model GEV terbaik dihitung nilai estimasi parameternya. Hasil estimasi patameter tersebut kemudian dimasukkan kedalam model sehingga diperoleh persamaan model *trend surface* terbaik sebagai berikut:

$$\hat{\mu}(j)=0.3145207+0.0006628u(j)$$

$$\hat{\sigma}(j)=2.41652+0.11434v-0.01206u(j)$$

$$\hat{\xi}(j)=-0.1449$$

Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur disajikan dalam Tabel 6.

Setelah mendapatkan parameter copula gaussian dari hasil model *trend surface* terbaik, kemudian parameter itu digunakan untuk membangkitkan data random. Data random yang diperoleh seperti terdapat pada Tabel 7.

### Estimasi Parameter Gaussian Copula

Nilai estimasi parameter copula untuk masing-masing lokasi curah hujan di

Tabel 6. Hasil Estimasi Parameter Gaussian Copula dengan Model *Trend Surface*

Provinsi	Latitude	Longitude	Lokasi ( $\hat{\mu}$ )	Skala ( $\hat{\sigma}$ )	Bentuk ( $\hat{\xi}$ )
Jawa Barat	-6.8836	107.5973	0.386	0.332	-0.145
Jawa Tengah	-6.9486	110.4199	0.388	0.290	-0.145
Jawa Timur	-7.3846	112.7833	0.389	0.212	-0.145

Tabel 7. Data Random Menggunakan Gaussian Copula

No.	Jawa Barat	Jawa Tengah	Jawa Timur
1	0.606009757	0.55245563	0.31616379
2	2.274362693	1.51697878	1.58185776
3	1.056833761	2.31653719	3.13887328
4	295.1140986	119.768470	27.7540633
5	0.547848582	0.70725176	0.58033621
⋮	⋮	⋮	⋮
79.	0.524510522	1.49878442	1.683195015

Tabel 8. Prediksi Curah Hujan dengan Gaussian Copula

No	Jawa Barat		Jawa Tengah		Jawa Timur	
	Aktual	Prediksi	Aktual	Prediksi	Aktual	Prediksi
1	317.1	233.8	337.9	291.3	273.2	-9.9
2	526.4	219.0	326.7	944.1	589	92.4
3	563.8	379.1	454.6	211.9	571.2	169.7
4	457.7	415.1	593.6	426.5	640.7	296.7
5	304.8	82.8	381.5	128.8	556.2	117.9
6	416.7	200.4	270.7	205.2	318.4	388.1
7	299.9	467.1	809.5	456.1	886	434.6
8	499.8	-3.0	324.4	361.4	516.7	101.0
9	410.5	350.6	905	248.2	255.6	317.1
10	365.7	434.7	483.4	276.7	652.5	281.8
11	557.1	320.3	429.5	224.7	581.7	493.4
12	381.5	256.4	382.1	177.9	398.5	420.7
13	537	232.4	437.4	75.8	445.9	37.9
14	637	364.9	469	453.8	461.1	144.9
15	418.7	177.8	991.9	125.9	455.1	33.9
16	455	-2.0	261.8	382.4	479.8	214.7
17	311.5	271.9	324.3	204.4	589.6	165.0
18	389.3	642.0	404.5	117.8	400.7	106.1
19	483.2	467.9	708.6	205.9	528.5	128.9
20	322.9	187.1	372.8	232.2	487.8	69.8

Untuk membandingkan hasil prediksi curah hujan dengan data curah hujan yang asli, data random dari gaussian copula pada tabel 7 harus ditransformasikan kembali menjadi bentuk distribusi GEV terlebih dahulu. Hasil dari prediksi dapat dilihat pada Tabel 8.

Selanjutnya besaran error dari hasil prediksi curah hujan pada Tabel 8 dapat dilihat pada Tabel 9.

Tabel 9. Besar Error hasil Prediksi Curah Hujan dengan Gaussian Copula

Nomor	Provinsi	MAPE
1.	Jawa Barat	0.4484322
2.	Jawa Tengah	0.5410097
3.	Jawa Timur	0.6301803

Pada Tabel 9 terlihat bahwa nilai MAPE untuk Provinsi Jawa Barat sebesar 0.4484322, Provinsi Jawa Tengah sebesar 0.5410097 dan untuk Provinsi Jawa Timur yaitu 0.6301803. Jadi bisa disimpulkan bahwa model curah hujan untuk Provinsi Jawa Barat layak digunakan karena memiliki nilai MAPE 44.8 % dan model untuk Provinsi Jawa Tengah dan Jawa Timur dapat dikatakan buruk karena memiliki nilai MAPE 54% dan 63%.

### Estimasi Parameter Data Rumah Rusak

Oleh karena data rumah rusak yang digunakan memiliki banyak nilai 0, maka distribusi yang bisa digunakan untuk melakukan estimasi parameter yaitu distribusi poisson dan negative binomial.

Tabel 10. Hasil Estimasi Parameter Menggunakan Distribusi Poisson

Variabel	Distribusi	Parameter		AIC
		Parameter 1	Parameter 2	
Jawa Barat	Poisson	120.6329	-	34360.99
Jawa Tengah	Poisson	121.3544	-	35592.09
Jawa Timur	Poisson	257.8861	-	94807.33
Jawa Barat	Binomial Negatif	0.1060129	120.642237	605.7986
Jawa Tengah	Binomial Negatif	0.128063	38.2205640	638.5495
Jawa Timur	Binomial Negatif	0.06679072	258.07516612	526.7250

Dari Tabel 10 tersebut terlihat bahwa nilai AIC dari distribusi binomial negatif lebih kecil dari pada distribusi poisson, sehingga untuk data rumah rusak distribusi yang digunakan yaitu distribusi binomial negatif.

### Fitting Copula

Dari Hasil *Fitting* keluarga copula untuk Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur diperoleh hasil copula terbaik yang dapat dilihat Tabel 11.

Tabel 11. Besar Error hasil Prediksi Curah Hujan dengan Gaussian Copula

Provinsi	Copula	Parameter	AIC
Jawa Barat	Frank	1.49998	-1.50126
Jawa Tengah	Frank	-0.58169	1.23643
Jawa Timur	Frank	-0.86483	1.045856

Berdasarkan Tabel 11 diperoleh model ketergantungan antara curah hujan ekstrem dengan rumah rusak akibat banjir pada Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur yaitu copula Frank. Bila variabel tersebut mengikuti copula Frank artinya memiliki hubungan yang erat ketika kedua variabel rendah atau kuat dilihat dari nilai parameternya.

Berdasarkan model yang terpilih, model masing – masing copula dapat ditulis seperti berikut ini:

1. Model copula frank pada Jawa Barat yaitu:

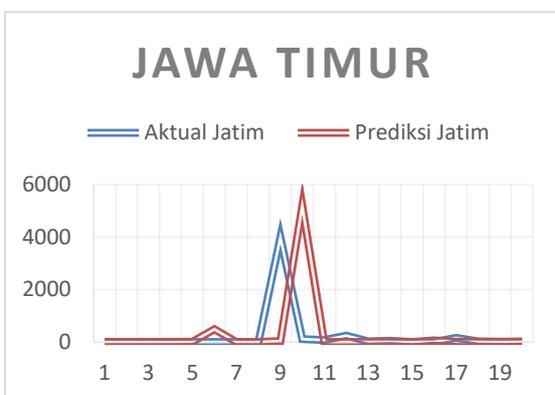
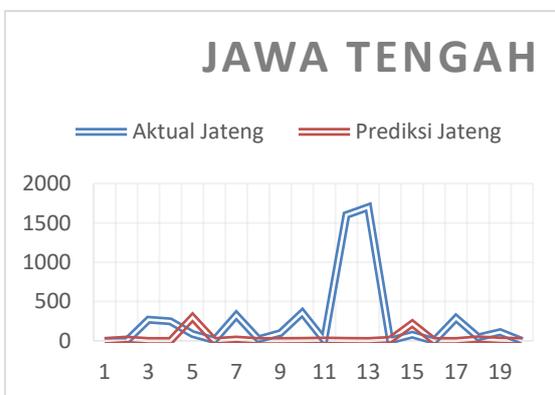
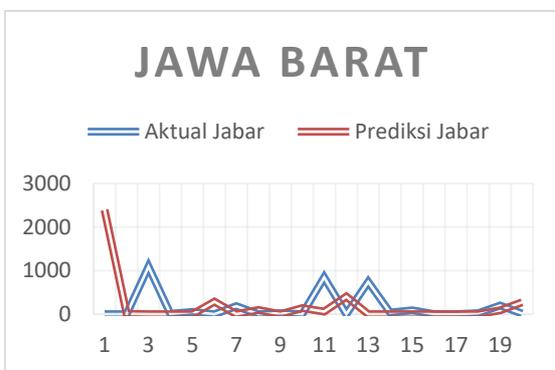
$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left( 1 + \frac{(e^{\theta u} - 1)(e^{\theta v} - 1)}{e^{\theta} - 1} \right) \\ = -\frac{1}{1.49} \ln \left( 1 + \frac{(e^{1.49u} - 1)(e^{1.49v} - 1)}{e^{1.49} - 1} \right)$$

2. Model copula frank pada Jawa Tengah yaitu:

$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left( 1 + \frac{(e^{\theta u} - 1)(e^{\theta v} - 1)}{e^{\theta} - 1} \right) \\ = \frac{1}{0.58} \ln \left( 1 + \frac{(e^{-0.58u} - 1)(e^{-0.58v} - 1)}{e^{0.58} - 1} \right)$$

3. Model copula frank pada Jawa Timur yaitu:

$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left( 1 + \frac{(e^{\theta u} - 1)(e^{\theta v} - 1)}{e^{\theta} - 1} \right) \\ = \frac{1}{0.86} \ln \left( 1 + \frac{(e^{-0.86u} - 1)(e^{-0.86v} - 1)}{e^{0.86} - 1} \right)$$



Gambar 3. Prediksi Rumah Rusak dengan Copula Terbaik

Selanjutnya dari model yang telah diperoleh dilakukan random data menggunakan model copula terbaik yang kemudian hasilnya akan dibandingkan dengan data asli. Untuk membandingkan hasil prediksi rumah dengan data rumah rusak yang asli, data random dari copula terbaik harus ditransformasikan kembali menjadi bentuk distribusi binomial negatif terlebih dahulu. Hasil peramalan data rumah rusak dapat dilihat pada Gambar 3.

Selanjutnya dari hasil prediksi tersebut, dilakukan perhitungan error untuk mengetahui seberapa baik hasil prediksi.

Tabel 12. Besar Error Hasil Prediksi Rumah Rusak dengan Copula Terbaik

Nomor	Provinsi	MASE
1.	Jawa Barat	1.009497
2.	Jawa Tengah	0.84625
3.	Jawa Timur	1.112388

Dari Tabel 12 diperoleh nilai MASE untuk dari hasil prediksi pada Provinsi Jawa Barat yaitu 1.009497, untuk provinsi Jawa Tengah yaitu 0.84625 dan untuk Jawa Timur yaitu 1.112388.

## KESIMPULAN

Beberapa simpulan sesuai dengan masalah yang dirumuskan, yakni antara lain:

1. Model curah hujan ekstreme pada Provinsi Jawa Barat, Jawa Tengah dan Jawa Timur menggunakan estimasi *Spatial Extreme Value* dengan pendekatan copula menghasilkan model *trend surface* sebagai berikut:  
 $\hat{\mu}(j) = 0.3145207 + 0.0006628u(j)$   
 $\hat{\sigma}(j) = 2.41652 + 0.11434v - 0.01206u(j)$   
 $\hat{\xi}(j) = -0.1449$
2. Nilai koefisien eksternal pada ketiga pasang Provinsi berada pada rentang nilai 1.03- 1.42 sehingga dapat disimpulkan bahwa terdapat dependensi *spatial* antar Provinsi, artinya ketiga Provinsi tersebut memiliki ketergantungan spasial.
3. Model Copula untuk Provinsi Jawa Barat, Provinsi Jawa Tengah dan Jawa Timur mengikuti copula Frank. Nilai

parameter  $\hat{\theta}$  copula Frank pada Provinsi Jawa Barat sebesar 1.4999840280, Provinsi Jawa Tengah sebesar -0.5816995330 dan Provinsi Jawa Timur sebesar -0.8648329345.

## DAFTAR PUSTAKA

- AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., & Sorooshian, S. (2013). *Extremes in a Changing Climate: Detection, Analysis and Uncertainty*. New York: Springer.
- Amalia, L. F. (2017). Estimasi Parameter pada Pemodelan Spatial Extreme Value Dengan Pendekatan Copula. *BMKG*. (2020). Diambil kembali dari Daftar Istilah Klimatologi: <http://balai3.denpasar.bmkg.go.id/daf-tar-istilah-musim>
- Coles, S. (2001). Improving the Analysis of Extreme Wind Speed with Information-sharing Models. *de l'Institut Pierre Simon Laplace*, no. 11, p. 12, 284.
- Cooley, D., Nyckah, D., & Naveau, P. (2007). A dependence measure for multivariate and spatial extremes: Properties and inference. *Journal of the American Statistical Association*, 824-840.
- Davison, A., Padoan, S., & Ribatet, M. (2012). Statistical Modeling of Spatial Extremes. *Statistical Science*, 161-186.
- Engmann, S., & Cousineau, D. (2011). *Jurnal of Applied Quantitative Methods*, Vol 6, No 3.
- Hakim, A. R. (2016). *Pemodelan Spatial Extreme Value dengan Pendekatan Max-Stable Process*. Surabaya: Institute Teknologi Sepuluh Nopember.
- Hyndman, R. J. (2006). Another Look At Forecast-Accuracy Metrics For Intermittent Demand. *FORESIGHT*, 43-46.
- Nelsen, R. B., & Flores, M. Ú. (2005). The lattice-theoretic structure of sets of bivariate copulas and quasi-copulas. *Comptes Rendus Mathematique*, vol.341, no.9, hal.583-586.
- Ramadhani, I. (2019). Identifikasi Struktur Dependensi dan Prediksi Indeks Harga Saham Gabungan Menggunakan Regresi Berbasis D-Vine Copula. Yogyakarta: Universitas Islam Indonesia.
- Ramadhani, I. R. (2015). *TESIS - SS 142501*.
- Razak, M. A. (2017). Peramalan Jumlah Produksi Ikan Dengan Menggunakan Backproagation Neural Network (Studi Kasus: UPTD Pelabuhan Perikanan Banjarmasin). Surabaya: Intitute Teknologi Sepuluh Nopember.
- Rinaldi, A. (2016). Sebaran Generalized Extreme Value (GEV) dan Generalized Pareto (GP) untuk Pendugaan Curah Hujan Ekstrem di Wilayah DKI Jakarta. *Al-Jabar: Jurnal Pendidikan Matematika Vol. 7, No. 1*, Hal 75 - 84.
- Schöölzel, C., & Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research – introduction to the copula approach. *Nonlinear Processes in Geophysics*.