



TUGAS AKHIR - KS184822

**PENENTUAN ZONA MUSIM DI MOJOKERTO
MENURUT KARAKTERISTIK CURAH HUJAN DENGAN
METODE TIME SERIES BASED CLUSTERING**

**IFTITAH AYUNDARI
NRP 062115 4000 0039**

**Dosen Pembimbing
Dr. Sutikno, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



TUGAS AKHIR - KS184822

**PENENTUAN ZONA MUSIM DI MOJOKERTO
MENURUT KARAKTERISTIK CURAH HUJAN
DENGAN METODE *TIME SERIES BASED
CLUSTERING***

**IFTITAH AYUNDARI
NRP 062115 4000 0039**

**Dosen Pembimbing
Dr. Sutikno, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**

(Halaman ini sengaja dikosongkan)



FINAL PROJECT - KS184822

**DETERMINATION OF SEASONAL ZONE IN
MOJOKERTO BASED ON PRECIPITATION
CHARACTERISTICS USING TIME SERIES BASED
CLUSTERING**

**IFTITAH AYUNDARI
SN 062115 40000039**

**Supervisor
Dr. Sutikno, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**

(Halaman ini sengaja dikosongkan)

LEMBAR PENGESAHAN

**PENENTUAN ZONA MUSIM DI MOJOKERTO
MENURUT KARAKTERISTIK CURAH HUJAN
DENGAN METODE *TIME SERIES BASED
CLUSTERING***

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

Iftitah Ayundari

NRP. 062115 4000 0039

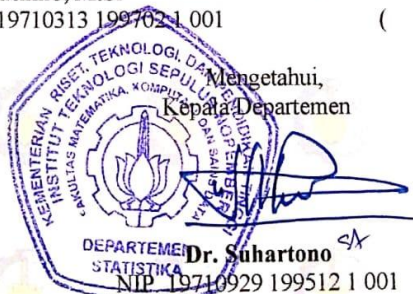
Disetujui oleh Pembimbing :

Dr. Sutikno, M.Si

NIP. 19710313 199702 1 001

()

Mengetahui,
Kepala Departemen



SURABAYA, JULI 2019

(Halaman ini sengaja dikosongkan)

PENENTUAN ZONA MUSIM DI MOJOKERTO MENURUT KARAKTERISTIK CURAH HUJAN DENGAN METODE *TIME SERIES BASED CLUSTERING*

Nama Mahasiswa : Iftitah Ayundari
NRP : 062115 4000 0039
Departemen : Statistika
Dosen Pembimbing : Dr. Sutikno, M.Si

Abstrak

Zona musim adalah daerah yang pos hujan rata-ratanya memiliki perbedaan yang jelas antara periode hujan dan musim kemarau. Metode analisis data yang diterapkan dalam pengelompokan ZOM ini adalah metode analisis cluster, yaitu mengelompokkan suatu objek yang memiliki kesamaan sifat tanpa menghilangkan struktur alami objek. Oleh karena kelompok yang dihasilkan memiliki makna, seperti pola atau klasifikasi. Data curah hujan merupakan data series sehingga dalam melakukan analisis cluster menggunakan jarak basis time series, yaitu pada penelitian ini menggunakan jarak Dynamic Time Warping (DTW) dan autocorrelation function (ACF). Tujuannya adalah mengetahui karakteristik zona musim di seluruh wilayah Mojokerto dan memetakan pola zona musim di Mojokerto berdasarkan data curah hujan menggunakan time series based clustering. Pengelompokan dilakukan dengan membentuk 2 hingga 5 kelompok, dan selanjutnya akan dilakukan pemilihan kelompok optimum dengan pseudo-f statistics dan kriteria silhouette. Data curah hujan yang digunakan adalah curah hujan bulanan. Evaluasi kinerja zona musim dilakukan dengan rasio simpangan baku dimana metode yang dibandingkan adalah pengelompokan dengan jarak ACF, DTW, dan hasil pengelompokan oleh BMKG. Berdasarkan hasil evaluasi kinerja zona musim dengan rasio simpangan baku metode terbaik yang dipilih adalah pengelompokan dengan jarak DTW

Kata Kunci : Curah Hujan, Clustering, Mojokerto, Zona Musim

(Halaman ini sengaja dikosongkan)

DETERMINATION OF SEASONAL ZONE IN MOJOKERTO BASED ON PRECIPITATION CHARACTERISTICS USING TIME SERIES BASED CLUSTERING

Name : Iffitah Ayundari
Student Number : 062115 4000 0039
Department : Statistics
Supervisor : Dr.Sutikno, M.Si

Abstract

Seasonal zones are areas where the average rainfall post has a clear difference between the period of the rainy season and the dry season. The data analysis method applied in ZOM grouping is cluster analysis method, which is grouping objects that have similar properties without removing the natural structure of the object. Because the groups produced have meaning, such as patterns or classifications. Rainfall data is a data series so that in conducting cluster analysis using time series base distance, that is in this study using the Dynamic Time Warping (DTW) distance and autocorrelation function (ACF). The aim is to find out the characteristics of the season zones in all Mojokerto regions and to map the pattern of the season zones in Mojokerto based on rainfall data using time series based clustering. Grouping is done by forming 2 to 5 groups, and then the optimum group selection will be carried out with pseudo-f statistics and silhouette criteria. The rainfall data used is monthly rainfall. Evaluation of the season zone performance is carried out with a standard deviation ratio where the method compared is grouping with distance ACF, DTW, and the results of grouping by BMKG. Based on the results of the evaluation of the season zone performance with the standard deviation ratio the best method chosen is grouping with distance DTW

Keywords: Rainfall, Clustering, Mojokerto, Seasonal Zone

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **”Penentuan Zona Musim Di Mojokerto Menurut Karakteristik Curah Hujan Dengan Metode *Time Series Based Clustering*”** dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada pihak-pihak sebagai berikut:

1. Bapak dan ibu, atas segala do’a, nasehat, kasih sayang, dan dukungan yang diberikan kepada penulis demi kesuksesan dan kebahagiaan penulis.
2. Dr. Sutikno, M.Si selaku dosen pembimbing Tugas Akhir dan sekaligus dosen wali, yang telah meluangkan waktu dan dengan sangat sabar memberikan bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir.
3. Dr. Bambang Widjanarko Otok, M.Si dan Dr. Santi Wulan Purnami, M.Si selaku dosen penguji yang telah banyak memberi masukan kepada penulis.
4. Dr. Suhartono selaku Ketua Departemen Statistika dan Dr. Santi Wulan Purnami, M.Si. selaku Ketua Program Studi Sarjana yang telah memberikan fasilitas, sarana, dan prasarana.
5. Seluruh dosen Statistika ITS yang telah memberikan ilmu dan pengetahuan yang tak ternilai harganya, serta segenap karyawan Departemen Statistika ITS.
6. Teman-teman Statistika ITS angkatan 2015 VIVACIOUS, yang selalu memberikan dukungan kepada penulis selama ini.
7. Teman-teman PSt 2017/2018 yang selalu memberikan dukungan dan motivasi dalam penyelesaian Tugas Akhir
8. Teman-teman seperjuangan bimbingan Tugas Akhir Pak Sutikno yang selalu memberikan informasi dan keceriaan.

9. Teruntuk teman teman Meyda, Dissa, Novita, Wulan, Cahya, Anin, Zikra, Azizah, Wanin yang selalu ada setiap saat untuk penulis.
10. Sahabat-sahabat perantauan di kos Pak Imam
11. Semua pihak yang turut membantu dalam pelaksanaan Tugas Akhir yang tidak bisa penulis sebutkan satu persatu.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Juli 2019

Penulis

DAFTAR ISI

JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Penelitian.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	4
1.5 Batasan Penelitian.....	4
BAB II TINJAUAN PUSTAKA	5
2.1 Analisis <i>Cluster</i>	5
2.1.1 Metode Hierarki.....	5
2.1.2 Metode Non-Hierarki.....	6
2.2 Analisis Faktor.....	7
2.3 <i>Autocorrelation Funcion</i> (ACF).....	8
2.4 Analisis <i>Cluster Data Time Series</i>	9
a <i>Dynamic Time Warping</i> (DTW) <i>Distance</i>	10
b <i>Autocorrelation Funcion</i> (ACF) <i>Distance</i>	11
2.5 <i>Pseudo-f Statistic</i>	12
2.6 <i>Koefisien Silhoutte</i>	13
2.7 Evaluasi Kinerja Terbaik.....	14
2.8 <i>Analsis of Variance</i> (ANOVA).....	14

2.9	Curah Hujan	15
2.10	Zona Musim	16
2.11	Zona Musim Mojokerto.....	16
2.12	Pengolahan Data Spasial	17
BAB III METODOLOGI PENELITIAN.....		19
3.1	Sumber Data	19
3.2	Variabel Penelitian	19
3.3	Struktur Data	19
3.4	Langkah Penelitian	22
BAB IV ANALISIS DAN PEMBAHASAN.....		27
4.1	Karakteristik Curah Hujan di Mojokerto.....	27
4.1.1	Rata Rata Curah Hujan di Mojokerto	27
4.1.2	Rata Rata Curah Hujan Per Bulan	32
4.1.3	Time Series Plot	29
4.2	Pengujian Autokorelosi Temporal.....	34
4.3	Analisis Faktor	35
4.4	<i>Clustering</i> dengan Jarak ACF	36
4.4.1	Perbandingan Hasil Pengelompokkan	42
4.5	<i>Clustering</i> dengan Jarak DTW	43
4.5.1	Perbandingan Hasil Pengelompokkan	49
4.6	Perbandingan Hasil <i>Cluster</i> Optimal dengan Zona Musim BMKG	50
BAB V KESIMPULAN DAN SARAN.....		53
5.1	Kesimpulan.....	53
5.2	Saran.....	54
DAFTAR PUSTAKA		55
LAMPIRAN-LAMPIRAN		59
BIODATA PENULIS		

DAFTAR TABEL

Tabel 2.1	Kriteria Koefisien <i>Silhouette</i>	14
Tabel 2.2	Perhitungan <i>Analysis of Variance</i>	15
Tabel 3.1	Struktur Data Curah Hujan (mm)	19
Tabel 3.2	Struktur Data Curah Hujan (mm) pada Analisis Faktor.....	20
Tabel 4.1	Nilai <i>Component Rotated Matrix</i>	36
Tabel 4.2	Analisis Ragam Pada Jarak ACF.....	43
Tabel 4.3	Analisis Ragam Pada Jarak DTW.....	50
Tabel 4.4	Perbandingan Hasil Cluster Optimal dengan Zona Musim BMKG	41

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 2.1	Ilustrasi Perhitungan Matriks Jarak	11
Gambar 2.2	Peta Zona Musim Mojokerto.....	17
Gambar 2.3	Lokasi Pos Hujan Mojokerto.....	17
Gambar 2.4	Contoh Ilustrasi Penentuan Batas ZOM dengan melakukan <i>overlay</i>	18
Gambar 3.1	Peta Pos Hujan Mojokerto.....	21
Gambar 3.2	Diagram Alir <i>Clustering Time Series</i>	24
Gambar 4.1	Rata – rata Curah Hujan di Mojokerto	27
Gambar 4.2	Rata- rata Curah Hujan di Mojokerto.....	28
Gambar 4.3	Rata-rata Curah Hujan Setiap Bulan di Mojokerto	29
Gambar 4.4	Time Series Plot Curah Hujan di Pacet	33
Gambar 4.5	Time Series Plot Curah Hujan di Mojosari	34
Gambar 4.6	ACF Data Curah Hujan Trowulan.....	35
Gambar 4.7	ACF Data Curah Hujan Sukosari	35
Gambar 4.8	Dendogram dengan Jarak ACF.....	37
Gambar 4.9	(a) Koefisien Silhoutte dan (b) Nilai Pseudo-f Statistics Hasil Pengelompokkan dengan Jarak ACF	38
Gambar 4.10	Pengelompokkan Zona Musim Menggunakan Jarak ACF Pada Faktor 1	39
Gambar 4.11	Pengelompokkan Zona Musim Menggunakan Jarak ACF Pada Faktor 3	40
Gambar 4.12	Pola Curah Hujan Zona Musim 1 dengan Jarak ACF	41
Gambar 4.13	Pola Curah Hujan Zona 2 Pada Klaster dengan Jarak ACF	41
Gambar 4.14	Perbandingan Rata Rata Curah Hujan Zona 1 dan Zona 2 dengan Jarak ACF	42
Gambar 4.15	Dendogram dengan Jarak DTW	46

Gambar 4.16 (a) Koefisien Silhoutte dan (b) Nilai Pseudo-f Statistics Hasil Pengelompokkan dengan Jarak DTW	45
Gambar 4.17 Pengelompokkan Zona Musim Menggunakan Jarak DTW dengan Faktor 1	46
Gambar 4.18 Pengelompokkan Zona Musim Menggunakan Jarak DTW dengan Faktor 3	47
Gambar 4.19 Pola Curah Hujan Zona Musim 1 dengan Jarak DTW	48
Gambar 4.20 Pola Curah Hujan Zona Musim 2 dengan Jarak DTW	48
Gambar 4.21 Perbandingan Rata Rata Curah Hujan Zona 1 dan Zona 2 dengan Jarak DTW	49

DAFTAR LAMPIRAN

Lampiran 1	Data Curah Hujan Bulanan (mm/bulan)	59
Lampiran 2	<i>Syntax Cluster</i> Jarak ACF	60
Lampiran 3	<i>Syntax Cluster</i> Jarak DTW	61
Lampiran 4	<i>Syntax</i> Menghitung Nilai <i>Pseudo-f Statistics</i>	62
Lampiran 5	Dendogram	63
Lampiran 6	Nilai Koefisien Silhoutte Jarak ACF 2-5 Cluster	64
Lampiran 7	Nilai Koefisien Silhoutte Jarak DTW 2-5 Cluster	65
Lampiran 8	<i>Time Series Plot</i> Pos Hujan di Mojokerto	66
Lampiran 9	Plot ACF Seluruh Pos Hujan Mojokerto	66
Lampiran 10	Zona Musim dengan Jarak ACF	67
Lampiran 11	Zona Musim dengan Jarak DTW	68
Lampiran 12	Data Pos Hujan	69
Lampiran 13	Output SPSS Analisis Faktor	70
Lampiran 14	ANOVA Pada Jarak ACF	71
Lampiran 15	ANOVA Pada Jarak DTW	73
Lampiran 16	Surat Pernyataan Data	75

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indikasi fenomena perubahan iklim di Indonesia dapat diamati dari perubahan pola curah hujan rata-rata di beberapa wilayah di Indonesia (BMKG, 2019). Iklim dapat didefinisikan sebagai ukuran statistik cuaca untuk jangka waktu tertentu dan cuaca menyatakan status atmosfer pada sembarang waktu tertentu. Dua unsur utama iklim adalah suhu dan curah hujan. Indonesia sebagai daerah tropis ekuatorial mempunyai variasi suhu yang kecil, sementara variasi curah hujannya cukup besar. Oleh karena itu curah hujan merupakan unsur iklim yang paling sering diamati dibandingkan dengan suhu. Dengan demikian iklim memiliki pengaruh terhadap pola curah hujan.

Berkaitan dengan perubahan iklim yang memiliki pengaruh juga terhadap pola curah hujan, Badan Meteorologi, Klimatologi, Geofisika (BMKG) melakukan *focus group discussion* pemutakhiran Zona Musim di Indonesia dengan tujuan *updating* normal curah hujan tiap sepuluh tahun, dimana standar normal yang selama ini digunakan adalah 1981-2010 sehingga BMKG perlu mempersiapkan diri menjelang normal baru 1991-2020. Zona musim adalah daerah yang pos hujan rata-ratanya memiliki perbedaan yang jelas antara periode musim hujan dan musim kemarau (BMKG, 2018). Penetapan normal curah hujan yang merupakan bagian dari normal iklim dengan menggunakan data selama 30 tahun sejalan yang telah di atur oleh *World Meteorological Organization* (WMO). Normal iklim dinyatakan sebagai rata-rata periode yang dihitung untuk jangka waktu yang seragam dan relatif panjang terdiri setidaknya tiga periode sepuluh tahun berturut-turut (WMO, 2007). Hasil evaluasi tersebut belum mutakhirnya prediksi zona musim yang dilakukan oleh BMKG dikarenakan karena zona musim yang tidak homogen dan metode yang digunakan belum sepenuhnya sesuai.

Penggunaan data kelompok stasiun lebih menguntungkan dari pada data masing-masing individu stasiun. Salah satu keuntungan data kelompok stasiun adalah data lebih homogen. Data pengamatan stasiun seringkali tidak homogen, karena adanya penggantian alat pengamatan (instrumentasi), perpindahan lokasi stasiun pengamatan, dan pengaruh urbanisasi (Bunkers et al., 1996). Dalam upaya mempelajari iklim, data yang digunakan harus lengkap. Namun problematika yang ada pada saat ini adalah tidak semua data di BMKG tersedia dan lengkap. Misalkan dalam melakukan penentuan zona musim menggunakan pos hujan terdapat kendala tidak semua pos hujan memiliki data yang lengkap. Oleh karena itu dalam melakukan pendugaan dapat menggunakan lokasi lokasi yang memiliki karakteristik yang sama. Untuk itu dilakukan penelitian pengelompokkan zona musim agar zona musim homogen. Data curah hujan bulanan yang akan digunakan dalam pengelompokkan zona musim. Karena adanya ketidaklengkapan data pada semua pos hujan maka metode analisis data yang sesuai adalah metode analisis *cluster*, yaitu mengelompokkan suatu objek yang memiliki kesamaan sifat tanpa menghilangkan struktur alami dari objek sehingga kelompok-kelompok yang dihasilkan memiliki makna, seperti pola atau klasifikasi. Badan Meteorologi Klimatologi dan Geofisika (BMKG) telah melakukan pengelompokkan iklim (pola curah hujan) dengan membuat Zona Musim (ZOM). Pembagian zona tidak dibatasi oleh wilayah administrasi kabupaten atau kota, sehingga dalam satu zona tertentu bisa mencakup beberapa kabupaten/ kota. Untuk Propinsi Jawa Timur terbagi atas 60 ZOM.

Data curah hujan merupakan data *series* sehingga dalam melakukan analisis *cluster* menggunakan jarak basis *time series*, yaitu pada penelitian ini menggunakan jarak *Dynamic Time Warping* (DTW) dan *Autocorrelation function* (ACF). *Cluster time series* adalah pendekatan yang paling banyak digunakan sebagai teknik eksplorasi dalam algoritma data mining yang lebih kompleks, seperti pengindeksan, klasifikasi, dan deteksi anomali. *Dynamic Time Warping* (DTW) telah diaplikasikan dalam *time*

series mining untuk menyelesaikan kesulitan dalam mengkluster data *time series*. *K-Means* adalah metode yang tercepat dari pada metode hierarki, tetapi jumlah *cluster* harus ditetapkan sebelumnya mengakibatkan ketidakpraktisan dalam mendapatkan hasil pengelompokan alami. Satu keuntungan dari pengelompokan hierarki adalah jumlah *cluster* tidak diperlukan sebagai parameter. Namun kelemahannya adalah persyaratan perhitungan jarak *Euclidean* tidak dapat menangani seri waktu yang lama karena skalabilitas yang buruk (Wang, Miles, & Hyndman, 2006). Skalabilitas merupakan kemampuan untuk menangani penambahan bobot yang diberikan.

Penelitian sebelumnya mengenai *time series clustering* yang dilakukan oleh T. Warren Liao yang berjudul *Clustering of time series data-a survey* memberikan kesimpulan bahwa untuk aplikasi pengelompokan seri waktu tertentu, kuncinya adalah memahami karakteristik unik data dan kemudian merancang kesamaan yang sesuai. *International Journal of Advances in Intelligent Informatics* yang berjudul *clustering stationary and non-stationary time series based on autocorrelation distance of hierarchical and k-means algorithms* memberikan kesimpulan bahwa metode *k-means* adalah algoritma terbaik untuk mengklasifikasikan stasioner dan non-stasioner pada data *time series* (Riyadi, dkk, 2017).

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka didapatkan rumusan masalah yaitu sebagai berikut.

1. Bagaimana karakteristik curah hujan pada zona musim di seluruh wilayah Mojokerto?
2. Bagaimana penentuan zona musim di seluruh wilayah Mojokerto berdasarkan karakteristik curah hujan menggunakan metode *cluster* dengan jarak ACF?
3. Bagaimana penentuan zona musim di seluruh wilayah Mojokerto berdasarkan karakteristik curah hujan menggunakan metode *cluster* dengan jarak DTW?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, tujuan yang ingin

dicapai dalam penelitian ini adalah sebagai berikut.

1. Mengetahui karakteristik curah hujan zona musim di seluruh wilayah Mojokerto.
2. Mendapatkan zona musim yang optimal di wilayah Mojokerto berdasarkan karakteristik curah hujan menggunakan metode *cluster* dengan jarak ACF.
3. Mendapatkan zona musim yang optimal di wilayah Mojokerto berdasarkan karakteristik curah hujan menggunakan metode *cluster* dengan jarak ACF.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan beberapa manfaat bagi berbagai pihak, diantaranya sebagai berikut.

1. Memberikan informasi mengenai pembagian zona musim di wilayah Mojokerto berdasarkan data curah hujan.
2. Memberikan wawasan keilmuan statistika mengenai metode *cluster* dengan data *time series*.

1.5 Batasan Masalah

Pada analisis pengelompokan zona musim untuk wilayah Mojokerto dengan 25 pos hujan. Data yang digunakan memiliki periode 7 tahun terakhir. Pengelompokan dilakukan dengan membentuk 2 hingga 5 kelompok, dan selanjutnya akan dilakukan pemilihan kelompok optimum dengan *pseudo-f statistics* dan kriteria *silhouette*. Evaluasi kinerja zona musim dengan menggunakan rasio simpangan baku.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Cluster

Menurut Rencher & Alvin (2002) analisis *cluster* (analisis kelompok) adalah metode analisis yang digunakan untuk mengelompokkan obyek-obyek pengamatan menjadi beberapa kelompok sehingga akan diperoleh kelompok dimana obyek-obyek dalam satu kelompok tersebut mempunyai persamaan sedangkan dengan anggota kelompok lain mempunyai perbedaan. Terdapat dua metode *cluster* yaitu metode hirarki dan metode non-hirarki. Pada penelitian ini metode pengelompokan yang digunakan menggunakan metode hirarki (Rencher & Alvin, 2002).

2.1.1 Metode Hierarki

Analisis *cluster* metode hirarki diawali dengan mengelompokkan dua atau lebih obyek yang mempunyai kesamaan paling dekat. Selanjutnya dilanjutkan pada obyek lain yang mempunyai kedekatan kedua dan seterusnya. Prosedur yang perlu dilakukan adalah sebagai berikut (Johnson, 2007):

1. Mendapatkan matriks jarak $D = \{d_{ik}\}$ dan menetapkan akan ada sebanyak n *cluster* yang dapat terbentuk, dimana N adalah banyak obyek yang diamati.
2. Menentukan matriks jarak untuk kedekatan antar *cluster*.
3. Menggabungkan *cluster* yang terbukti memiliki kedekatan.
4. Mengulangi langkah 3 sebanyak $N-1$ kali hingga diperoleh banyaknya *cluster* yang terbentuk dan seberapa dekat jarak antar *cluster* tersebut.

Berikut ini merupakan beberapa macam metode pada analisis *cluster* metode hirarki:

a) *Single Linkage*

Metode dengan prinsip jarak minimum. Langkah pertama menemukan jarak terdekat pada $D = \{d_{ik}\}$, dimana $\{d_{ik}\}$ adalah nilai obyek ke- i pada variabel ke- k dengan $i=1,2,3,\dots,n$ dan

menggabungkan obyek yang berkorespondensi, dikatakan U dan V dan sembarang kelompok lain W adalah $d_{(UV)W} = \min\{d_{UW}, d_{VW}\}$, dan d_{UW}, d_{VW} adalah jarak tetangga terdekat dari kelompok U dan W, dan jarak kelompok V dan W (Johnson R.A, 2007).

b) Complete Linkage

Metode *complete linkage* menghasilkan banyak cara yang sama seperti metode *single linkage*, kecuali pada setiap tahapannya yaitu jarak (kesamaan) antara *cluster-cluster* ditentukan oleh jarak (kesamaan) antara dua obyek dimana satu dari tiap *cluster* yang paling jauh. *Complete linkage* menggunakan aturan jarak maksimum antar *cluster*. Dapat ditulis ke dalam rumus sebagai berikut (Johnson, 2007).

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \quad (2.1)$$

Dengan :

d_{UW} = jarak antara U dan W,

d_{VW} = jarak antara V dan W.

2.1.2 Metode Non-Hierarki

Metode non-hierarki digunakan apabila banyak kelompok yang akan dibentuk sudah diketahui terlebih dahulu. Salah satu contohnya adalah metode *K-means*. *K-means* pertama kali dikembangkan lebih dari tiga dekade yang lalu (J.Macqueen). Pada metode ini harus ditentukan terlebih dahulu besarnya *k*, yaitu banyaknya kelompok. Pemilihan *k* dapat ditentukan secara subjektif berdasarkan latar belakang bidang masing-masing.

K-Means merupakan salah satu metode data *clustering* non-hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih klaster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu klaster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lainnya. *K-Mean* juga

merupakan metode *clustering* berbasis jarak yang membagi data ke dalam sejumlah klaster dan algoritma ini hanya bekerja pada atribut numerik.

2.2 Analisis Faktor

Analisis faktor mengasumsikan suatu struktur spesifik tertentu dalam sebuah model untuk variabel-variabel random. Johnson & Wichern (1998) mengatakan bahwa tujuan dari analisis faktor adalah untuk menggambarkan hubungan-hubungan kovarian antara beberapa variabel yang mendasari tetapi tidak teramati, yang sering disebut faktor. Dengan kata lain analisis faktor berfungsi menerangkan variasi sejumlah variabel asal dengan menggunakan faktor yang lebih sedikit dan yang tidak teramati dengan anggapan bahwa variabel asal dapat dinyatakan sebagai kombinasi linier dari faktor-faktor itu ditambahkan dengan suku residual.

Variabel random X yang diamati dengan p buah variabel komponen, yang memiliki rata-rata μ dan matriks kovarian Σ , maka model faktor dari X yang merupakan kombinasi linier beberapa variabel saling bebas yang tidak teramati adalah F_1, F_2, \dots, F_m disebut sebagai *common factors* dan ditambahkan dengan $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ disebut *specific factor*, sehingga secara khusus dapat ditulis sebagai berikut :

$$\begin{aligned} X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \dots + l_{1m}F_m + \varepsilon_1 \\ X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \dots + l_{2m}F_m + \varepsilon_2 \\ &\vdots \\ X_p - \mu_p &= l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + \varepsilon_p \end{aligned} \quad (2.2)$$

Dengan : $F_j = \text{Common factor ke-}j$

$l_{ij} = \text{Loading factor ke-}j \text{ dan variabel ke-}i$

$\varepsilon_i = \text{Spesific factor ke-}i$

dimana : $i = 1, 2, \dots, p$ dan $j = 1, 2, \dots, m$

Dalam notasi matriks persamaan 2.2 dapat ditulis sebagai :

$$X_{(p \times 1)} - \mu_{(p \times 1)} = L_{(p \times m)} F_{(m \times 1)} + \varepsilon_{(p \times 1)} \quad (2.3)$$

2.3 Autocorrelation Function (ACF)

Autocorrelation Function (ACF) merupakan suatu fungsi yang digunakan untuk menjelaskan korelasi antar Q_t dan Q_{t+k} dari suatu proses yang sama dan hanya terpisah oleh lag waktu ke- k . Sehingga dengan menggunakan perhitungan ACF dapat diketahui apakah data pengamatan saling berhubungan atau disebut juga korelasi temporal. Perhitungan ACF untuk data *time series* Q_1, Q_2, \dots, Q_n dapat diperoleh melalui persamaan berikut (Wei, 2006)

$$\rho_k = \frac{\sum_{t=1}^{n-k} (Q_t - \bar{Q})(Q_{t+k} - \bar{Q})}{\sum_{i=1}^n (Q_t - \bar{Q})^2}, k = 0, 1, 2, \dots \quad (2.4)$$

Dimana,

ρ_k = merupakan koefisien autokorelasi pada *lag-k*

n = jumlah data observasi

k = selisih waktu

Q = data pengamatan

\bar{Q} = rata rata data pengamatan

Untuk mengetahui apakah koefisien autokorelasi signifikan atau tidak, perlu dilakukan uji dengan hipotesis sebagai berikut

$H_0: \rho_k = 0$ (koefisien autokorelasi tidak signifikan)

$H_1: \rho_k \neq 0$ (koefisien autokorelasi signifikan)

Statistik uji yang digunakan seperti pada persamaan 2.5 sebagai berikut (Wei, 2006)

$$t = \frac{\widehat{\rho}_k}{SE(\rho_k)} \quad (2.5)$$

dengan nilai $SE(\rho_k)$ diperoleh pada persamaan 2.6 berikut:

$$SE(\rho_k) = \sqrt{\frac{1}{n}(1 + 2\widehat{\rho}_1^2 + 2\widehat{\rho}_2^2 + \dots + 2\widehat{\rho}_{k-1}^2)} \quad (2.6)$$

Kriteria keputusan tolak H_0 jika $|t_{hit}| > t_{\alpha/2, n-1}$, dimana nilai α merupakan tingkat kesalahan. Selain itu dapat menggunakan batas signifikansi dengan persamaan sebagai berikut:

$$0 \pm t_{\alpha/2, n-1} SE(\rho_k) \quad (2.7)$$

Dikatakan signifikan ketika nilai $\widehat{\rho}_k$ berada di luar batas signifikansi dengan nilai α tertentu (Wei, 2006).

2.4 Analisis Cluster Data Time Series

Analisis *cluster* dapat diterapkan juga pada data *time series*, dimana terdapat prosedur dan algoritma pengelompokan tertentu bila dibanding dengan pengelompokan data *cross section*. Algoritma dan prosedur berbeda dilakukan karena data *time series* merupakan sekelompok nilai pengamatan dalam waktu yang sama dan observasi yang berbeda, dan tidak hanya pada satu selang waktu saja. Data *time series* pada dasarnya digolongkan sebagai data dinamis karena nilai fiturnya berubah sebagai fungsi waktu, yang berarti bahwa nilai setiap titik deret waktu adalah satu atau lebih pengamatan yang dilakukan secara kronologis (Rani & Sikka, 2012). *Cluster time series* adalah pendekatan yang paling banyak digunakan sebagai teknik eksplorasi, dalam algoritma data mining yang lebih kompleks, seperti penemuan, pengindeksan, klasifikasi, dan deteksi anomali (Chis, Baenerjee, & Hassanien, 2009).

Definisi 1: Time Series Clustering, diberikan dataset n data *time series* $D = \{F_1, F_2, F_3, \dots, F_n\}$ dilakukan partisis Secara *unsupervised* D terhadap $C = \{C_1, C_2, C_3, \dots, C_n\}$ sedemikian rupa sehingga dalam kelompok *time series* homogen, pengelompokan dilakukan berdasarkan ukuran kesamaan tertentu. Selanjutnya, C_i disebut sebagai *cluster*, dimana $D = \bigcup_{i=1}^k C_i$ dan $C_i \cap C_j = \emptyset$ untuk setiap $i \neq j$. (Aghabozorgi, Shirkhorshidi, & Wah,

2015)

Selama berkembangnya pengelompokan pada data deret waktu, banyak teknik yang dikembangkan diantaranya adalah penggunaan jarak-jarak pengelompokan yang sesuai dengan karakteristik data deret waktu (Liao, 2005). Jarak yang digunakan dalam mengukur kedekatan dua data *time series* pada penelitian ini adalah jarak *Dynamic Time Warping* (d_{DTW}) dan jarak *Autocorrelation Function* (d_{ACF}), berikut penjelasan untuk masing-masing jarak.

a. **Dynamic Time Warping (DTW) Distance**

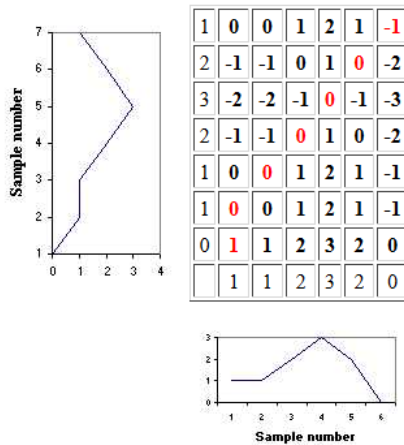
Dynamic time warping diperkenalkan pertama kali oleh Sakoe dan Chiba tahun 1978. Jarak DTW dapat digunakan untuk mengukur kedekatan untuk dua data *time series* dengan panjang yang berbeda. *Dynamic Time Warping* (DTW) menggunakan teknik pemrograman dinamis untuk menemukan semua jalur yang mungkin, dan memilihnya untuk menghasilkan jarak minimum antara dua deret waktu menggunakan matriks jarak, dimana masing-masing elemen dalam matriks adalah jarak kumulatif nilai minimal dari tiga tetangga sekitarnya. Jika diberikan dua data *time series* $Q = q_1, q_2, q_3, \dots, q_m$, dimana dengan ukuran m dan $C = c_1, c_2, c_3, \dots, c_n$ berukuran n , sehingga terbentuk matriks $m \times n$. Setiap elemen (i, j) merupakan jarak kumulatif dari jarak (i, j) dan nilai minimum dari tiga elemen yang berdekatan dengan elemen (i, j) , dimana $0 < i \leq n$ dan $0 < j \leq m$, sehingga dapat didefinisikan elemen (ij) sebagai berikut:

$$e_{ij} = d_{ij} + \min\{e_{(i-j)(j-1)}, e_{i(j-1)}\} \quad (2.8)$$

Dimana untuk e_{ij} merupakan elemen matriks (ij) . Nilai d_{ij} merupakan jarak antara q_i terhadap c_j . perhitungan nilai d_{ij} dapat dituliskan sebagai berikut:

$$d_{ij} = (q_i - c_j)^2 \quad (2.9)$$

Gambar 2.1 berikut ditampilkan ilustrasi matriks jarak, dimana untuk sumbu-y sebagai Q dan sumbu-x sebagai C . selanjutnya, untuk menemukan *optimal path*, harus menentukan *path* yang memberikan jarak kumulatif terkecil pada (m,n) . Pada Gambar 2.1 dapat digunakan untuk menentukan perhitungan jarak DTW. *Path* awal selalu $d_{11}, d_{12}, d_{23}, \dots, d_{ij}, \dots, d_{nm}$ dengan $j=i+1$. *Path* akan dimulai dari d_{11} sampai d_{nm} dan dilakukan penentuan semua jalur yang mungkin sampai menemukan jarak kumulatif terkecil. Berikut gambar untuk matriks jarak data Q dan C .



Gambar 2.1 Ilustrasi Perhitungan Matriks Jarak

Jarak kumulatif terkecil atau jarak DTW pada (m,n) didefinisikan sebagai berikut:

$$d_{DTW}(Q, C) = \min_{\forall w \in P} \left\{ \sqrt{\sum_{k=1}^K d_{wk}} \right\} \quad (2.10)$$

Dimana P adalah sekumpulan dari semua *warping path* yang mungkin, dan w_k adalah elemen (ij) ke- k pada *warping path* dan K adalah panjang dari *warping path* (Niennattrakul & Ratanamahatana, 2007)

b. *Autocorrelation Funcion (ACF) Distance*

Galeano dan Pena (2000) melakukan penelitian mengenai

hubungan data *time series* dengan menggunakan pendekatan *Autocorrelation Function* (ACF).

Ilustrasi untuk perhitungan jarak ACF adalah sebagai berikut, diberikan dua data *time series* $X_T = X_1, X_2, X_3, \dots, X_R$, dan $Y_T = Y_1, Y_2, Y_3, \dots, Y_R$.

$\hat{\rho}_{X_T} = (\hat{\rho}_{1,X_1}, \dots, \hat{\rho}_{L,X_R})'$ dan $\hat{\rho}_{Y_T} = (\hat{\rho}_{1,Y_1}, \dots, \hat{\rho}_{L,Y_R})'$ adalah vektor autokorelasi hasil pendugaan dari deret waktu X_T dan Y_T , untuk beberapa L seperti $\hat{\rho}_{X_i} \approx 0$ dan $\hat{\rho}_{Y_i} \approx 0$ dimana untuk $i > L$. Jarak antara dua deret waktu kemudian dapat dibentuk pada persamaan sebagai berikut

$$d_{ACF}(X_T, Y_T) = \sqrt{(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T})' \Omega (\hat{\rho}_{X_T} - \hat{\rho}_{Y_T})} \quad (2.11)$$

Dimana

$d_{ACF}(X_T, Y_T)$: jarak autokorelasi antara X_T dan Y_T .

$\hat{\rho}_{X_T}$: estimasi dari vector autokorelasi dari X_T ,

$\hat{\rho}_{Y_T}$: estimasi dari vector autokorelasi dari Y_T ,

Ω : merupakan matriks pembobot.

Apabila jarak ACF tidak memiliki pembobot maka matriks pembobot adalah matriks identitas. Sehingga, jarak autokorelasi menjadi seperti pada persamaan 2.12.

$$d_{ACFU}(X_T, Y_T) = \sqrt{(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T})' (\hat{\rho}_{X_T} - \hat{\rho}_{Y_T})} \quad (2.12)$$

2.5 *Pseudo-f* Statistic

Penentuan jumlah kluster optimum pada pengelompokan data berskala numerik digunakan nilai *pseudo-f* (Orpin & Kostylev, 2006). Nilai *pseudo-f* tertinggi menunjukkan bahwa jumlah kelompok telah optimal. Kelompok dikatakan optimal apabila homogen antar anggota dalam kelompok yang sama dan heterogen antar anggota dalam kelompok lain. y

$$f - Pseudo = \frac{(R^2 / (c-1))}{((1-R^2) / (n-c))} \quad (2.13)$$

Dengan,

$$R^2 = \left(\frac{SST - SSW}{SST} \right) \quad (2.14)$$

$$SST = \sum_{i=1}^c \sum_{j=1}^{n_c} \sum_{k=1}^p (x_{ijk} - \bar{x}_k)^2 \quad (2.15)$$

$$SSW = \sum_{i=1}^c \sum_{j=1}^{n_c} \sum_{k=1}^p (x_{ijk} - \bar{x}_{ik})^2 \quad (2.16)$$

Dimana

SST : total jumlah dari kuadrat jarak terhadap rata rata keseluruhan

SSW : total jumlah dari kuadrat jarak objek terhadap rata rata kelompoknya

n : banyak sampel

c : banyak *cluster*/kelompok

n_c : banyak data pada kelompok ke- c

p : banyak variabel

x_{ijk} : kelompok ke- i pada sampel ke- j dan variabel ke- k

\bar{x}_k : rata rata seluruh sampel pada variabel ke- k

\bar{x}_{ik} : rata rata kelompok ke- i pada variabel ke- k

2.6 Koefisien Silhoutte

Menurut Hair, dkk., (2010) validasi internal adalah cara paling sederhana untuk mengevaluasi algoritma pengelompokan data yang didasarkan pada asumsi bahwa cluster yang anggotanya dekat satu sama lain dan jauh dari anggota kelompok lainnya. Ketepatan sebuah pengelompokan menunjukkan seberapa baik proses pengelompokkan dan kualitas kelompok yang terbentuk. Terdapat beberapa macam ukuran ketepatan untuk mengetahui kualitas suatu pengelompokkan. Salah satu ukuran ketepatan yang

dapat digunakan dalam menentukan ketepatan pengelompokkan deret waktu adalah koefisien silhouette yang digunakan untuk mengetahui kualitas pengelompokkan (Hair, dkk, 2010).

$$S(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \quad (2.17)$$

dimana $a(i)$ adalah rata-rata jarak antar anggota dalam kelompok, $b(i)$ adalah nilai minimum dari rata-rata jarak dari objek i dengan objek yang berada di *cluster* lainnya, sedangkan $S(i)$ merupakan nilai koefisien *silhouette* pada objek ke- i . Nilai koefisien *silhouette* secara keseluruhan merupakan rata-rata $S(i)$ setiap objek. Pada Tabel 2.1 ditampilkan kriteria koefisien *silhouette* yang ditetapkan oleh Kauffman dan Rousseeuw (1990) (Kaufman & Rousseeuw, 1990).

Tabel 2.1 Kriteria Koefisien *Silhouette*

Nilai Koefisien <i>Silhouette</i>	Keterangan
0,71 – 1,00	<i>Strong</i>
0,51 – 0,70	<i>Good</i>
0,26 – 0,50	<i>Weak</i>
0,00 – 0,25	<i>Bad</i>

2.7 Evaluasi Kinerja Terbaik

Untuk melihat kinerja terbaik digunakan kriteria dua nilai simpangan baku, yaitu dalam kelompok (S_W) dan antar kelompok (S_B) (Bunkerset, dkk, 1996). Rata rata simpangan baku dalam kelompok dinyatakan dengan:

$$S_W = \frac{1}{K} \sum_{k=1}^K S_k \quad (2.18)$$

Dimana, K adalah banyaknya kelompok yang terbentuk dan S_k merupakan simpangan baku kelompok ke- k . Sedangkan simpangan baku antar kelompok dinyatakan sebagai berikut:

$$S_B = \left[\frac{1}{c-1} \sum_{k=1}^c (\bar{X}_k - \bar{X})^2 \right]^{1/2} \quad (2.19)$$

Dimana, \bar{X}_k adalah rata-rata kelompok ke- k dan \bar{X} adalah rata-rata seluruh kelompok. Semakin kecil nilai S_W dan semakin besar nilai S_B maka metode tersebut memiliki kinerja yang baik, artinya mempunyai homogenitas yang tinggi.

2.8 Analisis of Variance (ANOVA)

Analysis of Variance (ANOVA) merupakan teknik analisis yang digunakan untuk menguji tentang perbedaan rata-rata populasi secara univariat (Johnson dan Wichern, 2002). Pengujian kesamaan rata-rata secara univariat dilakukan dengan hipotesis:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_g$$

$$H_1: \text{minimal terdapat satu } \mu_i \text{ yang tidak sama, } i = 1, 2, \dots, g$$

Tabel ANOVA untuk perbandingan rata-rata populasi secara univariat disajikan pada Tabel 2.2.

Tabel 2.2 Analisis Ragam (ANOVA)

Sumber variasi	Jumlah Kuadrat (SS)	Derajat bebas
Perlakuan	$SS_{tr} = \sum_{i=1}^g n_i (x_i - \bar{x})^2$	$g-1$
Residual	$SS_{res} = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$	$\sum_{i=1}^g n_i - g$
Total (Terkoreksi)	$SS_{cor} = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$	$\sum_{i=1}^g n_i - 1$

(Sumber : Drapper & Smith, 1992)

Statistik uji yang digunakan yaitu sebagai berikut:

$$F = \frac{SS_{tr}/g-1}{SS_{res} / \left(\sum_{i=1}^g n_i - g \right)} \quad (2.20)$$

Tolak H_0 jika $F_{hitung} > F_{\alpha, (g-1), \sum_{i=1}^g n_i - g}$

2.9 Curah Hujan

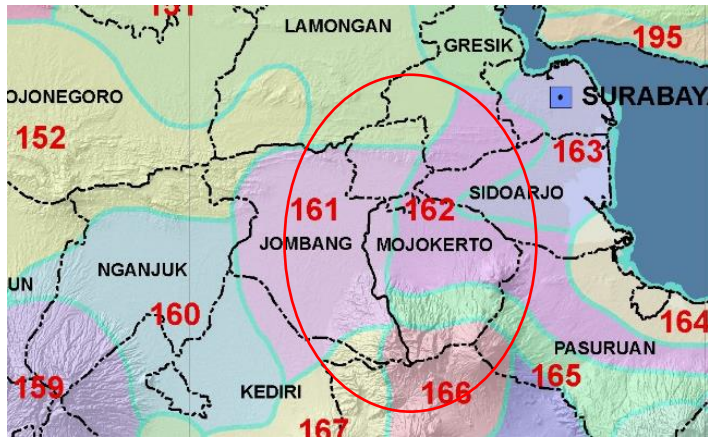
Curah hujan merupakan jumlah air yang jatuh di permukaan tanah datar selama periode tertentu yang diukur dengan satuan tinggi (mm) di atas permukaan horizontal bila tidak terjadi evaporasi, *run off*, dan infiltrasi. Jadi, jumlah curah hujan yang diukur, sebenarnya adalah tebalnya atau tingginya permukaan air hujan yang menutupi suatu daerah luasan di permukaan bumi/tanah. Satuan curah hujan yang umumnya dipakai oleh BMKG adalah millimeter (mm). Curah hujan 1 (satu) millimeter, artinya dalam luasan satu meter persegi pada tempat yang datar tertampung air setinggi 1 (satu) millimeter atau tertampung air sebanyak 1 (satu) liter atau 1000 ml (BMKG, 2012).

2.10 Zona Musim

Zona Musim (ZOM) adalah daerah yang pola hujan rata-ratanya memiliki perbedaan yang jelas antara periode musim kemarau dan periode musim hujan. Wilayah ZOM tidak selalu sama dengan luas daerah administrasi pemerintahan. Dengan demikian satu kabupaten/ kota dapat saja terdiri dari beberapa ZOM dan sebaliknya satu ZOM dapat terdiri dari beberapa kabupaten. Untuk wilayah Jawa Timur merupakan daerah yang memiliki pola monsoon, dan dibagi dalam 60 daerah zona musim (BMKG, 2012).

2.11 Zona Musim Mojokerto

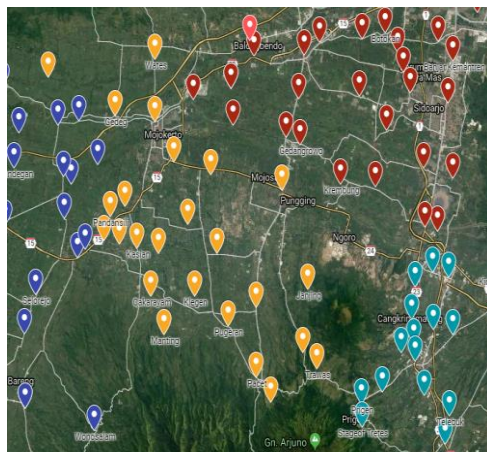
Secara administrasi Mojokerto terletak di Provinsi Jawa Timur, Indonesia tepatnya 50 km barat daya dari ibukota provinsi Jawa Timur yaitu Surabaya. Berikut merupakan peta zona musim Mojokerto.



Gambar 2.2 Peta Zona Musim Mojokerto

(Sumber: <https://karangploso.jatim.bmkg.go.id/index.php/informasiiklim/peta-zona-musim-di-Provinsi-Jawa-Timur>)

Lokasi dari pos hujan di Mojokerto ditunjukkan pada Gambar 2.3. Di Mojokerto terdapat 25 pos hujan dengan 3 zona musim.

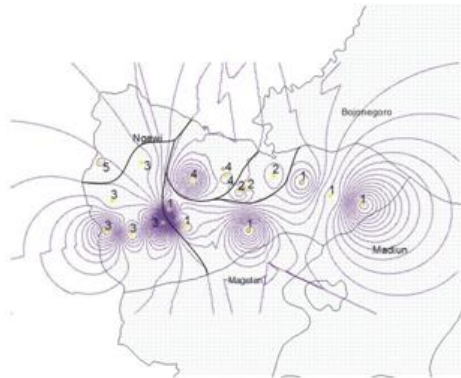


Gambar 2.3 Lokasi Pos Hujan Mojokerto

(Sumber: <https://karangploso.jatim.bmkg.go.id/index.php/profil/sitemap/82-peralatan-observasi-klimatologi/55555575-lokasi-penakar-hujan-manual-ombrometer-di-jawa-timur>)

2.12 Pengolahan Data Spasial

Overlay adalah proses tumpang – susun beberapa buah peta tematik dalam rangkaian kegiatan pengambilan kesimpulan secara spasial (Eko Budiyanto,2010). *Overlay* menampilkan suatu peta digital pada peta digital yang lain beserta atribut atributnya dan menghasilkan peta gabungan keduanya yang memiliki informasi atribut dari kedua peta tersebut. Berikut merupakan ilustrasi dari proses tumpang susun yang diperoleh dari penelitian yang dilakukan oleh Febrina Dinar Palupi Utami tahun 2010 dengan judul Pengelompokkan Zona Musim dengan *Fuzzy K-Means Clustering*.



Gambar 2.4 Contoh Ilustrasi Penentuan Batas ZOM dengan melakukan *overlay*

BAB III

METODOLOGI PENELITIAN

Pada bab ini dijelaskan mengenai sumber serta data yang digunakan dalam penelitian. Data yang digunakan akan ditampilkan untuk struktur datanya serta diilustrasikan dalam bentuk Peta. Selain itu, pada bab ini juga jelaskan langkah penelitian serta diagram alir penelitian.

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Dinas Pekerjaan Umum dan Sumber Daya Air Provinsi Jawa Timur.

3.2 Variabel Penelitian

Variabel penelitian yang digunakan adalah curah hujan bulan Januari 2011 hingga Desember 2017

3.3 Struktur Data

Struktur data yang akan dianalisis ditunjukkan pada Tabel 3.1 sebagai berikut.

Tabel 3.1 Struktur Data Curah Hujan (mm/bulan)

Pos Hujan (S_i)	Curah Hujan (mm/bulan) ($Y_{i,p}$)				
	Jan 2011	Feb 2011	..	Nov 2017	Des 2017
S_1	$Y_{1,1}$	$Y_{2,1}$..	$Y_{83,1}$	$Y_{84,1}$
S_2	$Y_{1,2}$	$Y_{2,2}$..	$Y_{83,2}$	$Y_{84,2}$
S_3	$Y_{1,3}$	$Y_{2,3}$..	$Y_{83,3}$	$Y_{84,3}$
...	⋮
S_{25}	$Y_{1,25}$	$Y_{2,25}$..	$Y_{83,25}$	$Y_{84,25}$

Keterangan:

S_i : nama lokasi pos hujan di mojosuro ke- i , $i = 1, 2, \dots, 25$;
 $Y_{i,p}$: variabel curah hujan pada pos hujan ke- i bulan ke- p , $i = 1, 2, \dots, 25$; $p = 1, 2, \dots, 84$

Sedangkan struktur data yang digunakan pada analisis

faktor terdapat dalam Tabel 3.2.

Tabel 3.2 Struktur Data Curah Hujan Untuk Analisis Faktor (mm/bulan)

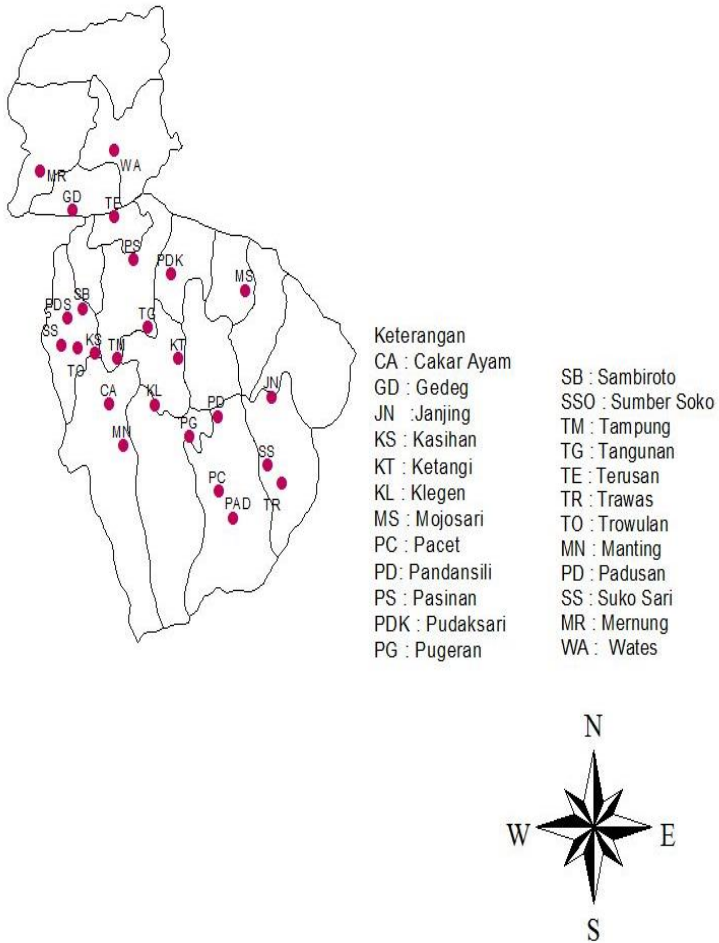
Pos Hujan	Tahun	Jan	Feb	...	Nov	Des
S ₁	2011	X _{1,1,1}	X _{1,1,2}	...	X _{1,1,11}	X _{1,1,12}
S ₂		X _{2,1,1}	X _{2,1,2}	...	X _{2,1,11}	X _{2,1,12}
⋮		⋮	⋮	⋮	⋮	⋮
S ₂₅		X _{25,1,1}	X _{25,1,2}	...	X _{25,1,11}	X _{25,1,12}
S ₁		2012	X _{1,2,1}	X _{1,2,2}	...	X _{1,2,11}
S ₂	X _{2,2,1}		X _{2,2,2}	...	X _{2,2,11}	X _{2,2,12}
⋮	⋮		⋮	⋮	⋮	⋮
S ₂₅	X _{25,2,1}		X _{25,2,2}	...	X _{25,2,11}	X _{25,2,12}
⋮	⋮		⋮	⋮	⋮	⋮
S ₁	2017	X _{1,7,1}	X _{1,7,2}	...	X _{1,7,11}	X _{1,7,12}
S ₂		X _{2,7,1}	X _{2,7,2}	...	X _{2,7,11}	X _{2,7,12}
⋮		⋮	⋮	⋮	⋮	⋮
S ₂₅		X _{25,7,1}	X _{25,7,2}	...	X _{25,7,11}	X _{25,7,12}
⋮		⋮	⋮	⋮	⋮	⋮

Keterangan:

S_i : nama lokasi pos hujan di Mojokerto ke-i, i = 1,2,...,25;

X_{i,j,k} : variabel curah hujan pada pos hujan ke-i, tahun ke-k, dan bulan ke-k, i = 1,2,...,25; j = 1,2,...,7; k=1,2,...,12.

Gambar 3.1 merupakan peta dari pos hujan di Mojokerto dengan nama dan singkatan.



Gambar 3.1 Peta Pos Hujan Mojokerto

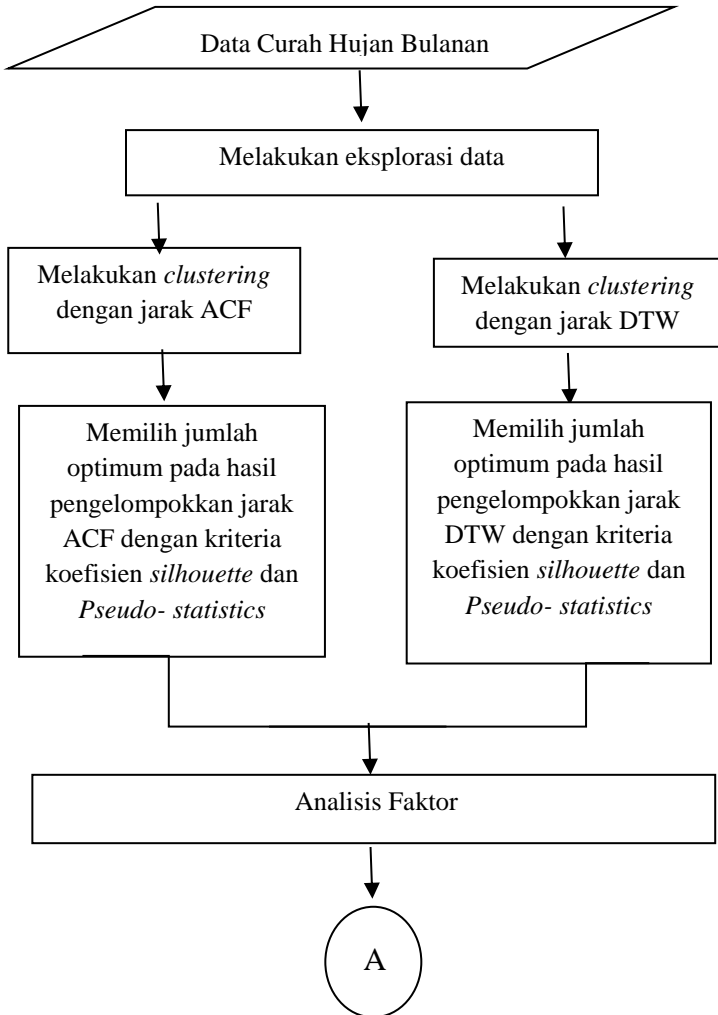
3.4 Langkah Analisis Data

Melakukan *clustering* untuk pemetaan zona musim di Mojokerto menggunakan data curah hujan, dengan langkah sebagai berikut:

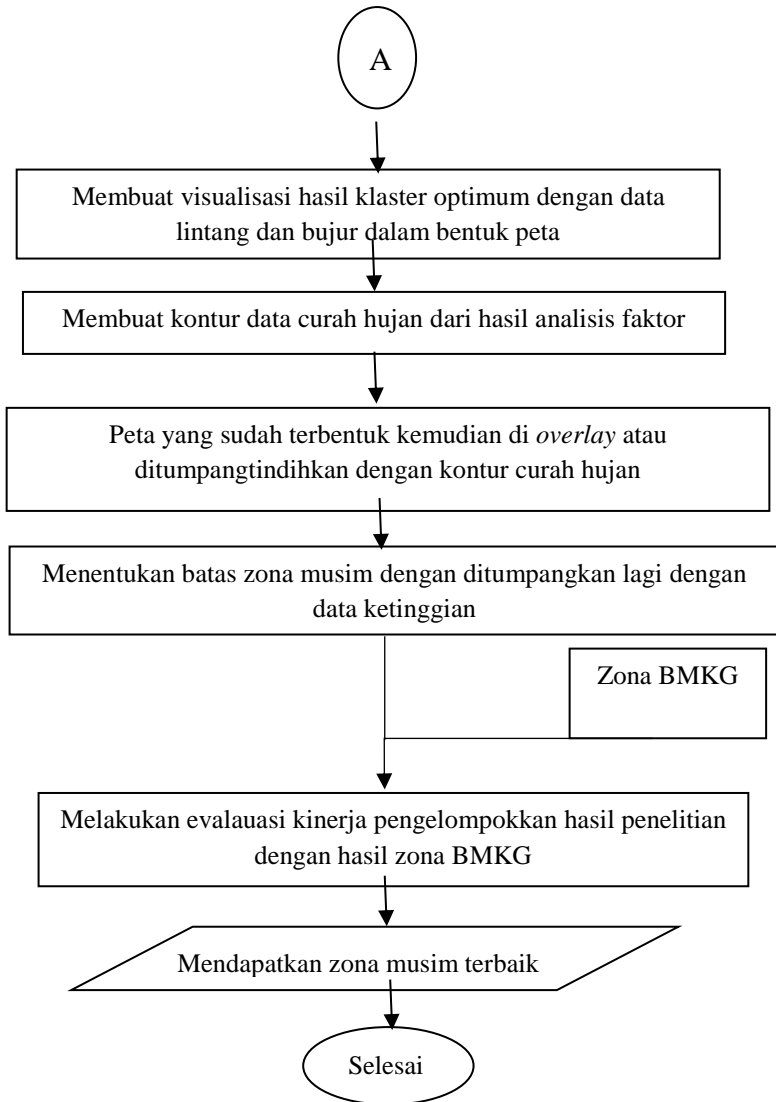
- a. Data yang digunakan merupakan data curah hujan setiap wilayah di Mojokerto pada jarak ACF dan DTW. Jumlah *cluster* ditentukan sebanyak 2 hingga 5 zona musim.
- b. Melakukan *clustering* dengan menggunakan jarak *dynamic time warping* (DTW) pada data curah hujan dengan langkah sebagai berikut:
 - i. Mendapatkan matriks jarak DTW dan menetapkan akan ada sebanyak n *cluster* yang dapat terbentuk, dimana N adalah banyak obyek yang diamati.
 - ii. Menentukan matriks jarak untuk kedekatan antar *cluster*.
 - iii. Menggabungkan *cluster* yang terbukti memiliki kedekatan menggunakan metode *complete linkage*.
 - iv. Mengulang langkah 3 sebanyak $n-1$ kali hingga diperoleh banyaknya *cluster* yang terbentuk dan seberapa dekat jarak antar *cluster* tersebut.
 - v. Melakukan *clustering* dengan jumlah *cluster* sebanyak 2 hingga 5
 - vi. Memilih jumlah *cluster* terbentuk menggunakan kriteria *pseudo f statistics* dan koefisien *silhouette*.
- c. Melakukan *clustering* dengan menggunakan jarak *autocorrelation function* (ACF) pada data curah hujan dengan langkah sebagai berikut:
 - i. Mendapatkan matriks jarak ACF dan menetapkan akan ada sebanyak n *cluster* yang dapat terbentuk, dimana N adalah banyak obyek yang diamati.
 - ii. Menentukan matriks jarak untuk kedekatan antar *cluster*.
 - iii. Menggabungkan *cluster* yang terbukti memiliki kedekatan menggunakan metode *complete linkage*.
 - iv. Mengulang langkah 3 sebanyak $n-1$ kali hingga diperoleh banyaknya *cluster* yang terbentuk dan seberapa dekat jarak antar *cluster* tersebut.

- v. Melakukan *clustering* dengan jumlah *cluster* sebanyak 2 hingga 5
- vi. Memilih jumlah *cluster* optimum yang terbentuk menggunakan kriteria *pseudo f statistics* dan koefisien *silhouette*
- d. Melakukan analisis faktor dari data curah hujan.
- e. Membuat visualisasi hasil kluster dalam bentuk peta dengan data lintang dan bujur
- f. Membuat kontur data curah hujan dari hasil analisis faktor
- g. Peta yang sudah terbentuk kemudian di *overlay* atau ditumpangtindihkan dengan kontur curah hujan.
- h. Menentukan batas zona musim pada hasil (g) dengan cara di *overlay* dengan data ketinggian
- i. Melakukan uji perbedaan rata rata dari pengelompokkan zona musim yang terbentuk.
- j. Menghitung evaluasi kinerja pengelompokkan hasil penelitian dan pengelompokkan dari BMKG dengan rasio simpangan baku S_w dan S_B
- k. Membandingkan ketepatan pengelompokkan hasil penelitian dengan hasil zona musim BMKG.

Tahapan penelitian dapat dituliskan dalam diagram alir sebagai berikut



Gambar 3.2 Diagram Alir *Clustering Time Series*



Gambar 3.2 Diagram Alir *Clustering Time Series*(lanjutan)

(Halaman ini sengaja dikosongkan)

BAB IV

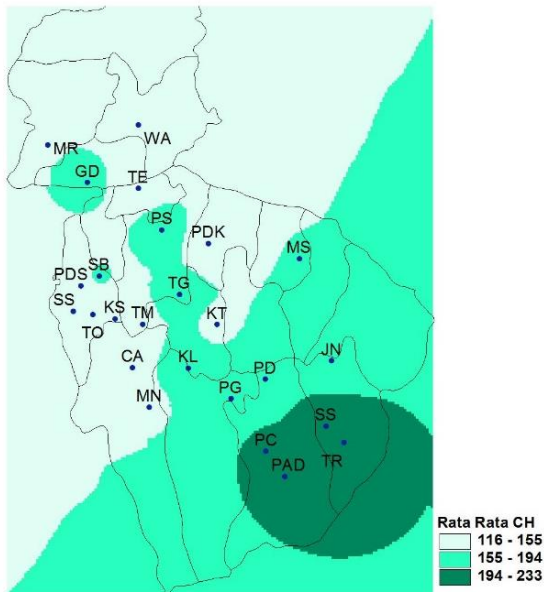
ANALISIS DAN PEMBAHASAN

4.1 Karakteristik Curah Hujan di Mojokerto

Karakteristik curah hujan di Mojokerto meliputi rata rata curah hujan di Mojokerto pada bulan Januari 2011 hingga Desember 2017 dan pola curah hujan tiap bulan di Mojokerto. Terdapat 25 pos hujan yang terletak di Mojokerto.

4.1.1 Rata Rata Curah Hujan di Mojokerto

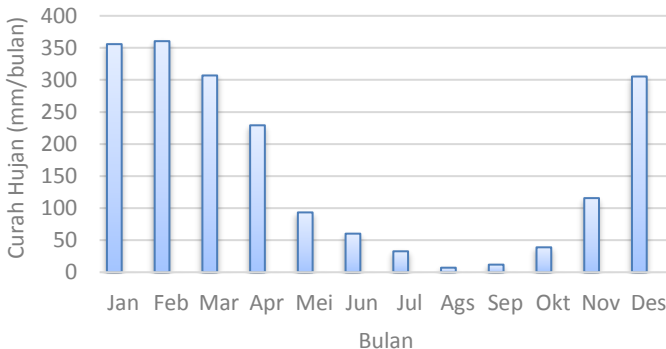
Rata rata curah hujan di Mojokerto pada bulan Januari 2011 hingga Desember 2017 dapat dilihat pada Gambar 4.1. Dari Gambar 4.1 dapat diperoleh informasi bahwa daerah yang memiliki curah hujan rendah di Mojokerto berwarna hijau muda dan curah hujan tertinggi berwarna hijau tua.



Gambar 4.1 Rata – rata Curah Hujan di Mojokerto

Wilayah Mojokerto yang memiliki curah hujan tertinggi dengan rata-rata 194 hingga 233 mm per tahun di sekitar Sukosari, Trawas, dan Pacet, dan Padusan, dimana disimbolkan dengan warna hijau tua. Pada wilayah tersebut rata-rata curah hujan cenderung tinggi mengingat wilayah tersebut merupakan daerah dataran tinggi. Sedangkan untuk curah hujan paling rendah memiliki warna hijau dimana sebagian berada di sekitar Mojokerto bagian utara.

BMKG mengelompokkan normal curah hujan terbagi dalam 3 kategori, yaitu rendah dengan curah hujan 0-100 mm, menengah dengan curah hujan (100-300 mm) dan tinggi dengan curah hujan 300-500 mm. Menurut nilai *range* yaitu selisih nilai minimum dan maksimum dari rata-rata curah hujan di Mojokerto termasuk pada kategori rendah hingga menengah.



Gambar 4.2 Rata-rata Curah Hujan di Mojokerto

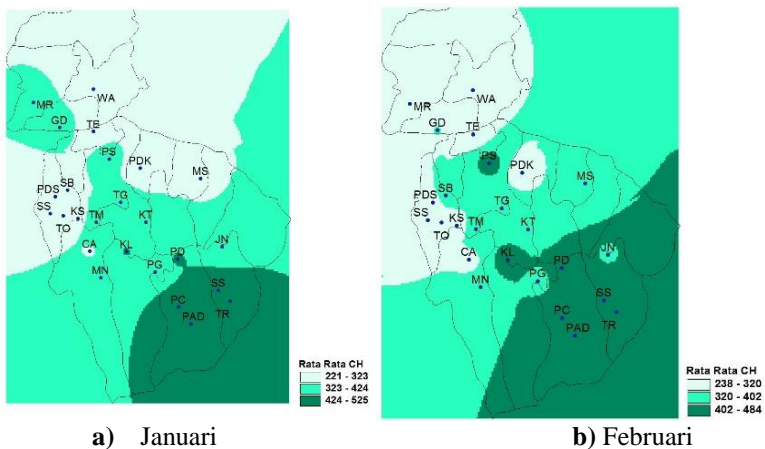
Gambar 4.2 memberikan informasi bahwa tipe curah hujan di Mojokerto berbentuk U atau dapat disebut tipe muson yang mengalami musim hujan di awal dan akhir tahun. Kriteria musim kemarau menurut BMKG ditetapkan berdasarkan curah hujan dalam dasarian (10 hari) yang kurang dari 50 milimeter. Puncak curah hujan tertinggi terjadi pada bulan Desember, Januari, dan Februari. Hal ini dapat dimungkinkan karena pada bulan tersebut terjadi musim penghujan. Sedangkan pada bulan Juli hingga

Oktober memiliki rata rata curah hujan terendah, hal ini dapat mengindikasikan bahwa pada pada saat itu terjadi musim kemarau.

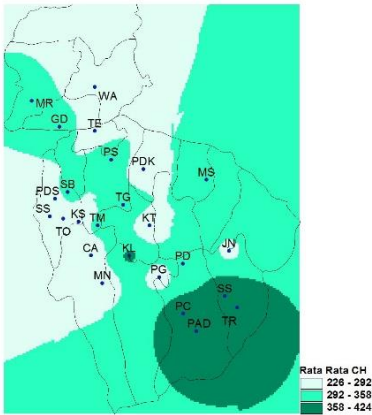
4.1.2 Rata Rata Curah Hujan Per Bulan

Rata rata curah hujan setiap bulan dapat memeberikan informasi mengenai pergerakan curah hujan di Mojokerto pada musim penghujan dan musim kemarau. Rata rata curah hujan setiap bulan di Mojokerto ditampilkan pada Gambar 4.3.

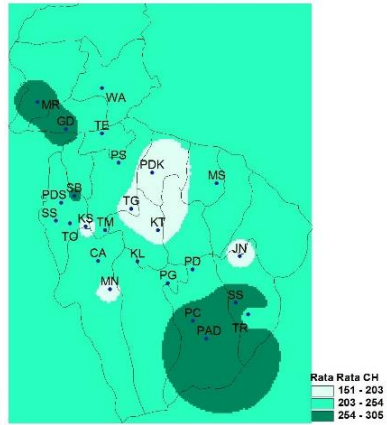
Dapat diketahui bahwa pada bulan Desember, Januari, dan Februari memiliki intensitas curah hujan paling tinggi. Hal ini mengindikasikan bahwa puncak dari musim hujan terjadi pada bulan tersebut. Intensitas curah hujan mulai menurun pada bulan Maret hingga Mei. Pada bulan Mei terdapat beberapa pos hujan yang memiliki curah hujan cukup tinggi dan sebagian lainnya sudah memasuki musim kemarau, sehingga pada bulan Mei terjadi musim peralihan dari musim hujan ke musim kemarau atau musim transisi. Mulai memasuki musim kemarau pada bulan Juni hingga Oktober. Pada bulan November terdapat beberapa pos hujan di daerah Trawas dan Pacet sudah memasuki musim hujan, sehingga pada bulan November merupakan bulan peralihan dari musim kemarau ke musim penghujan.



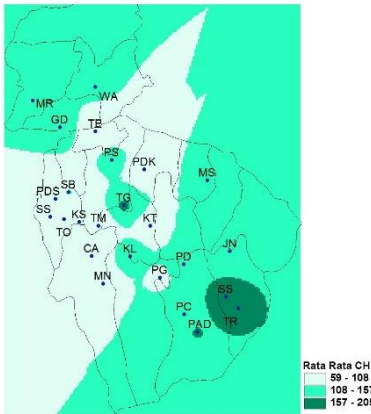
Gambar 4.3 Rata-rata Curah Hujan Setiap Bulan di Mojokerto



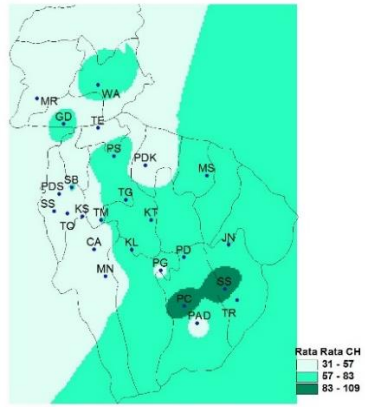
c) Maret



d) April

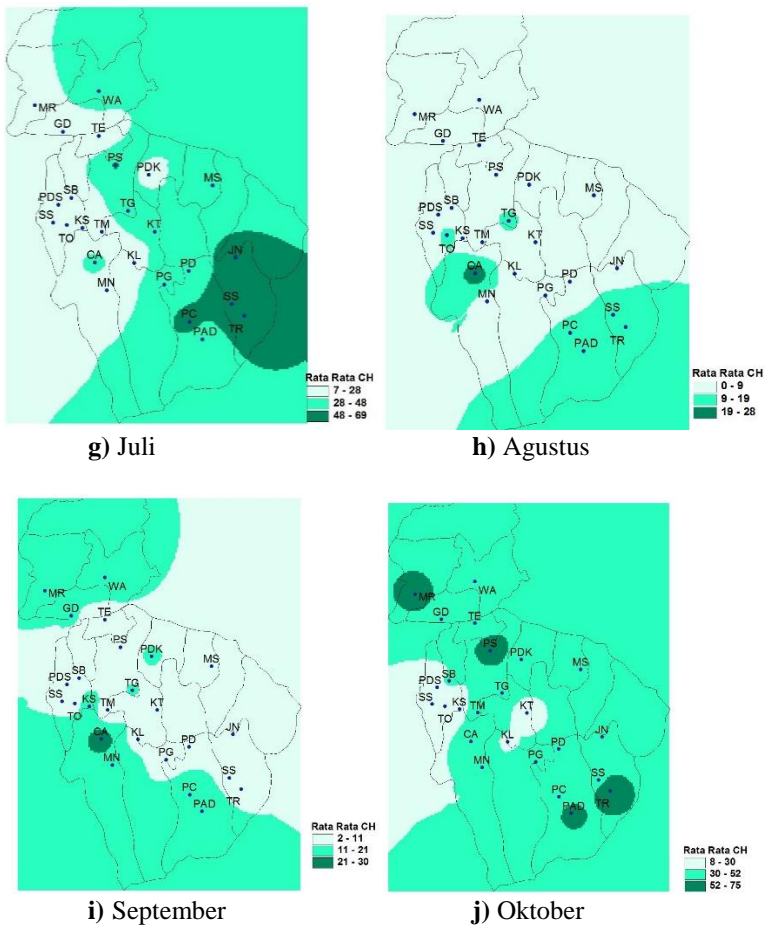


e) Mei

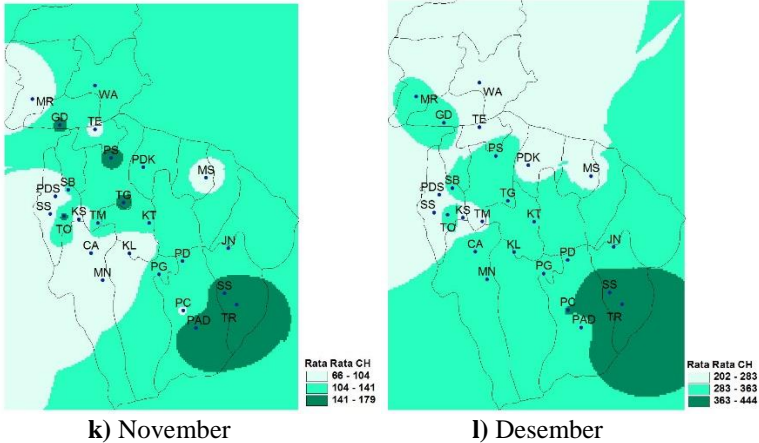


f) Juni

Gambar 4.3 Rata-rata Curah Hujan Setiap Bulan di Mojokerto (lanjutan)



Gambar 4.3 Rata-rata Curah Hujan Setiap Bulan di Mojokerto (lanjutan)



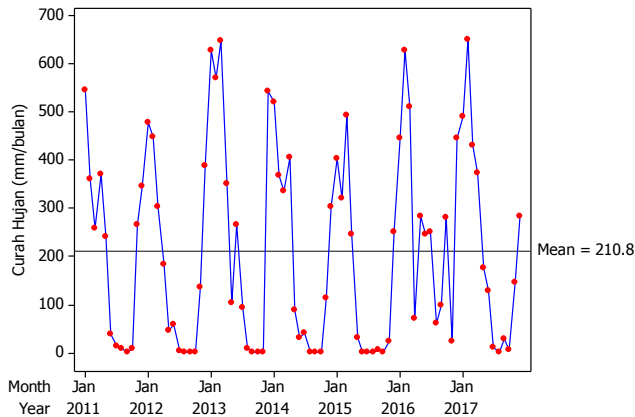
k) November
l) Desember
Gambar 4.3 Rata-rata Curah Hujan Setiap Bulan di Mojokerto(lanjutan)

Berdasarkan Gambar 4.3 pada bagian (a) atau pada bulan Januari secara umum wilayah Mojokerto memiliki curah hujan rata rata pada kategori menengah (100-300 mm). Namun terdapat daerah yang memiliki curah hujan rata rata pada kategori tinggi (300-500 mm) yaitu Pacet, Sukosari, Padusan, dan Trawas. Pada bulan Juli hingga Oktober secara umum wilayah Mojokerto memiliki curah hujan rata rata pada kategori rendah. Hal ini dapat mengindikasikan bahwa pada bulan tersebut terjadi musim kemarau. Pada bulan November wilayah Mojokerto memiliki curah hujan rata rata pada kategori rendah (0-100 mm) yaitu pada daerah yang berwarna hijau muda. Namun sebagian besar lainnya berada pada kategori menengah (100-300 mm). Hal ini dapat memberikan indikasi bahwa pada bulan November terjadi musim transisi.

4.1.3 Time Series Plot

Salah satu karakteristik curah hujan bisa dilakukan dengan menggunakan *time series plot* untuk melihat bagaimana kondisi curah hujan di Mojokerto. Berikut merupakan pola curah hujan dari beberapa pos hujan di Mojokerto. Untuk pos hujan lainnya terdapat pada Lampiran 8.

a. Time Series Plot Pos Hujan Pacet



Gambar 4.4 Time Series Plot Curah Hujan di Pacet

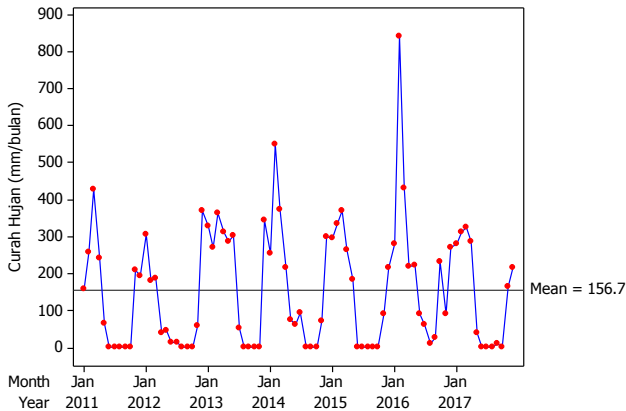
Selama periode tahun 2011 hingga 2017 curah hujan di Pacet memiliki rata rata 210,8 mm/bulan. Curah hujan tertinggi selalu berada di awal dan akhir tahun. Hal ini berarti bahwa musim penghujan terjadi pada setiap awal dan akhir tahun. Pada pertengahan tahun 2016 curah hujan di pos hujan Pacet mengalami peningkatan fluktuasi dibandingkan tahun tahun sebelumnya. Dimana terdapat rata rata curah hujan pada pertengahan tahun 2016 berada di atas rata rata.

b. Time Series Plot Pos Hujan Mojosari

Berikut merupakan *time series plot* curah hujan bulanan dari pos hujan di Mojosari dari tahun 2011 hingga 2017 atau selama 7 tahun.

Dibandingkan dengan *time series plot* dari pos hujan Pacet, rata rata pada pos hujan Mojosari memiliki curah hujan lebih rendah dibandingkan dengan di Pacet. Pada pos hujan Mojosari memiliki rata rata curah hujan 156,7 mm/bulan. Selama periode 7 tahun yaitu dari tahun 2011 hingga 2017 curah hujan tertinggi terjadi pada tahun 2016 dimana mengalami kenaikan curah hujan

yang cukup signifikan pada puncak curah hujan tertinggi dari tahun tahun sebelumnya.



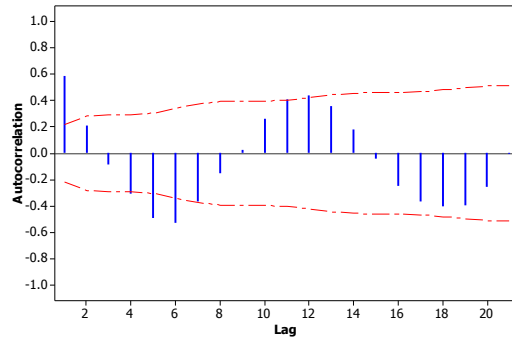
Gambar 4.5 Time Series Plot Curah Hujan di Mojosari

4.2 Pengujian Autokorelasi Temporal

Pengujian autokorelasi temporal dilakukan untuk mengetahui data pengamatan curah hujan tiap bulan saling berhubungan. Pengujian ini menggunakan nilai ACF (*Autocorrelation Function*). Adanya autokorelasi saat terdapat *lag* (nilai ACF) keluar dari batas signifikansi. Pengujian ini merupakan hal yang utama, karena sebagai landasan untuk melakukan *cluster* dengan jarak *time series*. Berikut merupakan pengujian yang dilakukan pada beberapa pos hujan, untuk pos hujan yang lainnya terdapat pada Lampiran 9.

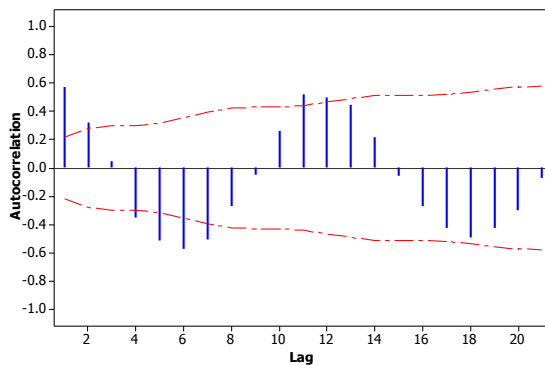
a. ACF Pos Hujan Trowulan

Plot ACF dari pos hujan Trowulan disajikan dalam Gambar 4.6. Pada Gambar 4.6 terdapat *lag* signifikan yang ditandai dengan keluarnya nilai ACF dari batas signifikan. Taraf signifikan yang digunakan sebesar 5%. Salah satu *lag* yang keluar adalah pada *lag* 1. Apabila terdapat minimal 1 *lag* yang keluar dari batas signifikansi atau *cut off* maka data curah hujan di pos hujan Trowulan memiliki autokorelasi temporal.



Gambar 4.6 ACF Data Curah Hujan Trowulan

b. ACF Pos Hujan Sukosari



Gambar 4.7 ACF Data Curah Hujan Sukosari

Plot ACF dari pos hujan Sukosari ditampilkan pada Gambar 4.7. Pada Gambar 4.7 terdapat *lag* signifikan yang ditandai dengan keluarnya nilai ACF dari batas signifikan. Taraf signifikan yang digunakan sebesar 5%. Salah satu *lag* yang keluar adalah pada *lag* 1. Apabila terdapat minimal terdapat 1 *lag* yang keluar dari batas signifikansi atau *cut off* maka data curah hujan di pos hujan Sukosari memiliki autokorelasi temporal.

4.3 Analisis Faktor

Sebelum melakukan pengelompokan stasiun curah hujan dilakukan pereduksian variabel dengan analisis faktor. Tujuannya

untuk menggambarkan karakteristik curah hujan di Mojokerto.

Tabel 4.1 Nilai *Component Rotated Matrix*

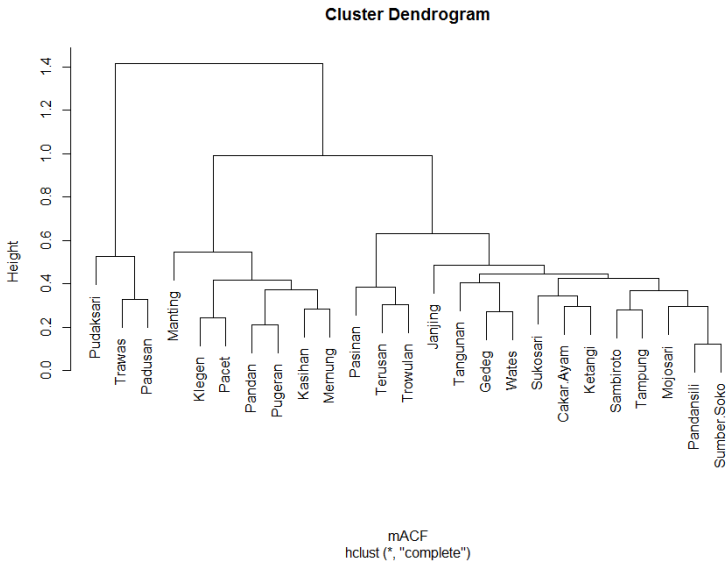
Bulan	F 1	F 2	F 3	F 4
Juni	0,818	0,272	0,176	-0,005
Mei	0,783	0,057	0,311	0,066
Juli	0,780	0,383	0,002	0,223
November	-0,597	0,094	0,393	0,487
Agustus	0,509	0,492	-0,198	0,257
Oktober	0,249	0,890	-0,068	0,088
September	0,144	0,860	-0,050	-0,013
Februari	0,121	0,663	0,516	-0,098
Maret	0,032	0,200	0,822	-0,001
Januari	0,190	-0,053	0,783	0,170
April	0,004	-0,230	0,685	0,169
Desember	0,168	0,094	0,213	0,869

Tabel 4.1 menunjukkan bahwa karakteristik curah hujan di Mojokerto apabila dilakukan analisis factor dihasilkan 4 faktor. Faktor 1 yang beranggotakan bulan Mei, Juni, Juli, Agustus, dan November menggambarkan musim transisi, Faktor 2 yang meliputi bulan Oktober, September, dan Februari. Faktor 3 yang menggambarkan musim penghujan beranggotakan bulan Januari, Maret, dan April. Dan faktor 4 yang merupakan bulan transisi dari musim kemarau ke musim penghujan beranggotakan bulan Desember.

Pengelompokkan curah hujan di Mojokerto menggunakan analisis factor dengan menghasilkan 4 faktor memiliki keragaman 73,37%.

4.4 Clustering dengan Jarak ACF

Sub bab ini merupakan analisis pembagian zona musim berdasarkan hasil *cluster* dengan jarak ACF. Pada tahap *clustering* digunakan jumlah *cluster* sebanyak 2 hingga 5 *cluster*. Berikut merupakan dendrogram hasil pengelompokkan dengan menggunakan jarak ACF



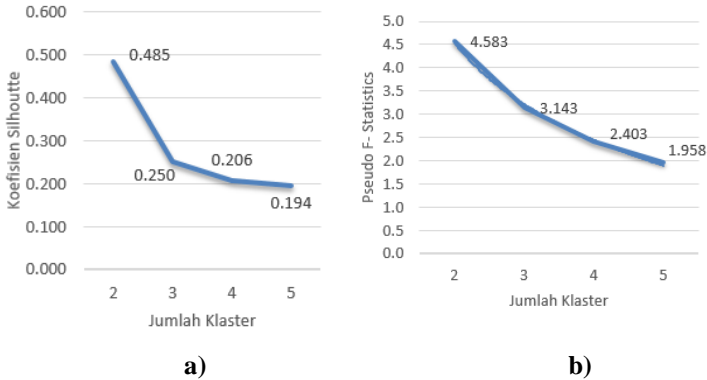
Gambar 4.8 Dendrogram dengan Jarak ACF

Hasil dendrogram menunjukkan adanya kedekatan antar objek yang dapat dijadikan panduan objek mana yang memiliki kemiripan karakteristik. Dua objek dengan karakteristik sama akan digambarkan sebagai dua titik yang posisinya berdekatan, semakin dekat posisi dua buah objek maka semakin mirip dan sebaliknya. Dari Gambar 4.8 dapat dijelaskan apabila dihasilkan 2 pengelompokan atau 2 kluster maka anggota dari kluster 1 adalah Pudaksari, Trawas, dan Padusan. Sedangkan anggota kluster 2 adalah pos hujan selain di kluster 1. Begitu pula dengan pengelompokan dengan 3,4,dan 5 kluster.

Untuk menentukan *cluster* optimal yang selanjutnya menjadi penentuan zona musim menggunakan kriteria koefisien *silhouette* dan *pseudo-f statistics* yang ditampilkan pada Gambar 4.9.

Gambar 4.9 memberikan informasi bahwa hasil kluster optimal berdasarkan kriteria *silhouette* dan *pseudo-f statistics* adalah 2 kluster atau menghasilkan 2 zona musim di Mojokerto. Dapat diketahui bahwa semakin banyak kluster yang terbentuk nilai

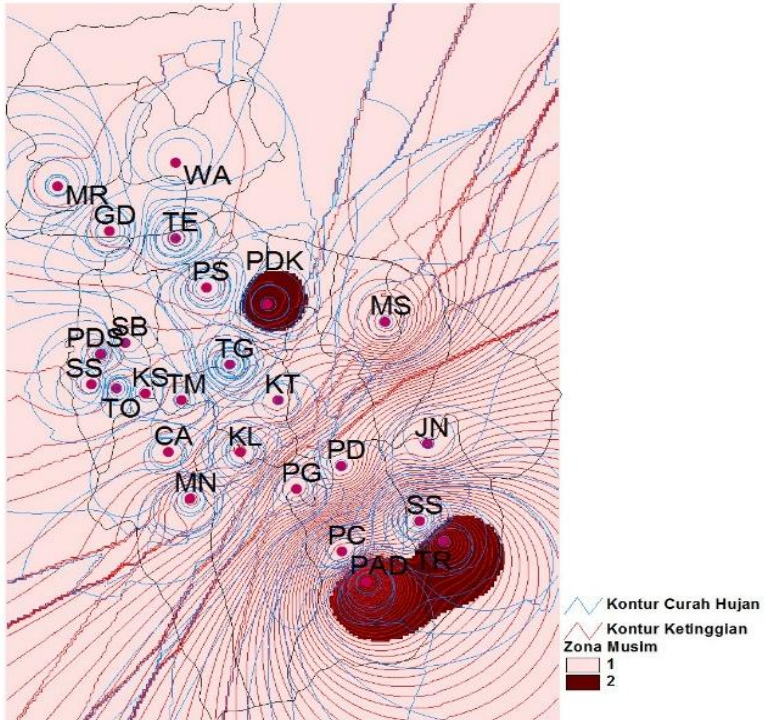
koefisien *silhoutte* dan *pseudo-f statistics* semakin kecil sehingga semakin tidak optimal dalam pengelompokkan zona musim.



Gambar 4.9 (a) Koefisien *Silhoutte* dan (b) Nilai *Pseudo-f Statistics* Hasil Pengelompokkan dengan Jarak ACF

Pembagian wilayah zona musim dengan ditumpang tindihkan dengan data ketinggian dan curah hujan ditampilkan pada Gambar 4.10. Anggota kelompok masing masing zona musim adalah sebagai berikut

- Zona Musim 1 : Cakar Ayam, Gedeg, Janjing, Kasihan, Ketangi, Klegen, Mojosari, Pacet, Pandan, Pandansili, Pasinan, Pugeran, Sambiroto, Sumber Soko, Tampung, Tangunan, Trowulan, Manting, Sukosari, Mernung, dan Wates
- Zona Musim 2 : Pudaksari, Trawas, dan Padusan

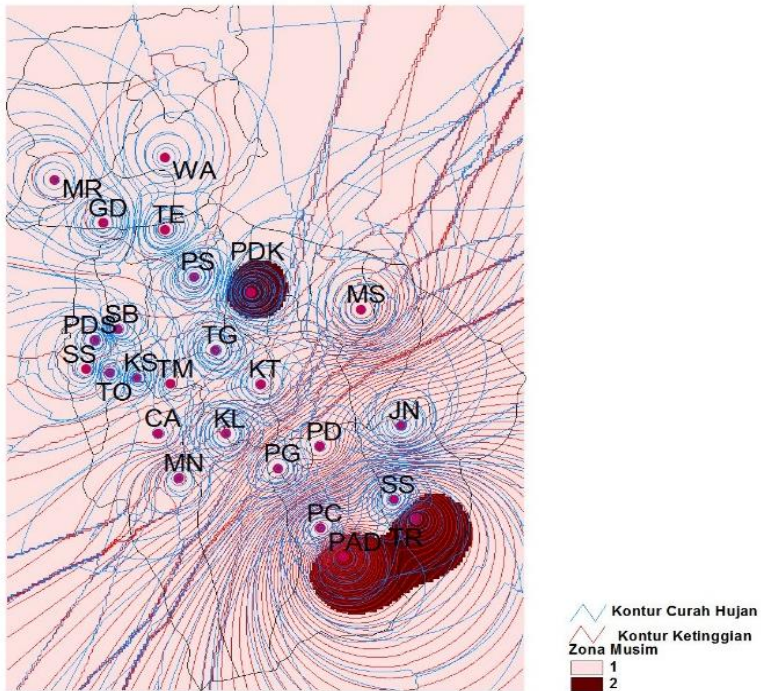


Gambar 4.10 Pengelompokan Zona Musim Menggunakan Jarak ACF Pada Faktor 1

Kontur yang berwarna biru menunjukkan kontur curah hujan dari faktor 1. Kontur curah hujan terlihat sedikit merenggang menunjukkan bahwa pada bulan tersebut sedang terjadi peralihan musim atau musim transisi. Kontur ketinggian ditunjukkan dengan garis kontur berwarna merah. Pada kontur ketinggian menunjukkan bahwa semakin banyak garis kontur yang berdekatan atau berkumpul menunjukkan semakin tinggi suatu daerah. Pada Gambar 4.10 dapat dilihat bahwa pos hujan yang berada di dataran tinggi diantaranya adalah pos hujan Pacet, Trawas, Sukosari, dan Padusan.

Berikut merupakan visualisasi pengelompokan zona musim dengan jarak ACF dimana dilakukan *overlay* pada ketinggian dan

curah hujan dari faktor 3

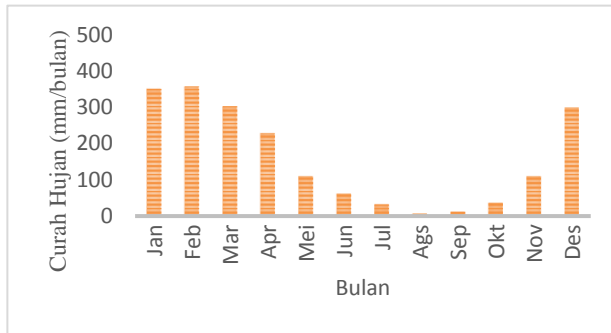


Gambar 4.11 Pengelompokan Zona Musim Menggunakan Jarak ACF Pada Faktor 3

Dibandingkan dengan Gambar 4.10, kontur curah hujan pada faktor 3 lebih rapat dari faktor 1. Hal ini dikarenakan pada factor 3 beranggotakan bulan bulan musim penghujan.

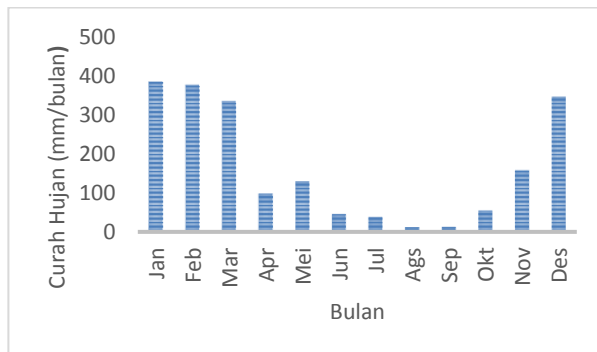
Sebagian besar sebaran pos hujan di Mojokerto termasuk ke dalam zona 1. Sedangkan pada zona 2 terdiri dari 3 pos hujan yaitu Trawas, Padusan, dan Pudaksari. Secara visual, pada zona musim 2 terdapat anggota klaster yang tidak berdekatan. Dari 3 pos hujan yang berada di zona musim, 2 pos hujan terletak di dataran tinggi dan sisanya pos hujan yang terletak di dataran rendah. Dengan demikian secara visual pengelompokan dengan jarak ACF tidak begitu terlihat pola pengelompokan zona musim yang terbentuk.

Pola curah hujan pada setiap zona dilihat berdasarkan rata rata curah hujan bulanan setahun pada Gambar 4.12 dan Gambar 4.13.



Gambar 4.12 Pola Curah Hujan Zona Musim 1 dengan Jarak ACF

Pola curah hujan di zona musim 1 berbentuk U atau memiliki tipe Monsunal. Tipe Monsunal mengalami musim penghujan cenderung pada awal dan akhir tahun. Pada zona 1 musim kemarau dimulai pada bulan Mei hingga November dengan puncak musim kemarau pada bulan Agustus dimana memiliki rata rata curah hujan paling rendah. Sedangkan untuk puncak penghujan dimana memiliki rata rata curah hujan tertinggi terjadi pada bulan Februari.

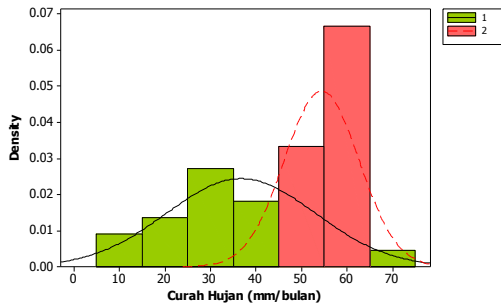


Gambar 4.13 Pola Curah Hujan Zona 2 Pada Klaster dengan Jarak ACF

Pada zona 2 musim kemarau dimulai lebih awal dari zona 1 yaitu pada bulan April hingga Oktober dengan puncak musim

kemarau pada bulan Agustus dan september dimana memiliki rata rata curah hujan paling rendah. Sedangkan puncak dari musim penghujan terjadi pada bulan Januari.

Perbandingan rata rata curah hujan zona 1 dan zona 2 menggunakan jarak ACF ditampilkan pada Gambar 4.14. dari Gambar 4.14 diketahui bahwa rata rata curah hujan pada zona 2 lebih tinggi dibandingkan dengan zona 1.



Gambar 4.14 Perbandingan Rata Rata Curah Hujan Zona 1 dan Zona 2 dengan Jarak ACF

Pada Gambar 4.14 diketahui bahwa pola curah hujan di zona 1 memiliki perbedaan dengan zona 2. Pada zona 1 curah hujan cenderung lebih stabil dibandingkan dengan zona 2. Dapat terlihat pada zona 2 pola curah hujan cenderung lebih tinggi. Artinya daerah di zona musim 2 memiliki musim penghujan yang lebih sering dibandingkan dengan daerah lain di zona musim 1.

Untuk melihat apakah ada perbedaan antara kedua zona musim maka dilakukan uji t. Dengan taraf signifikansi 5%, perhitungan uji t menghasilkan nilai *p-value* 0,04. Karena 0,04 kurang dari 0,05 maka dapat disimpulkan adanya perbedaan rata rata curah hujan pada pengelompokkan zona musim menggunakan jarak ACF.

4.4.1 Perbandingan Hasil Pengelompokkan

Perbandingan hasil pengelompokkan dari zona musim yang terbentuk dengan menggunakan analisis ragam (ANOVA). ANOVA dilakukan untuk menganalisis variabilitas data menjadi

dua sumber variasi yaitu variasi di dalam kelompok (*within*) dan variasi antar kelompok (*between*). Variabel yang digunakan merupakan rata rata curah hujan tiap bulan.

Pengujian ANOVA dari variabel curah hujan ditampilkan pada Tabel 4.2 berikut, indikator lain dari uji ANOVA terdapat pada Lampiran 13.

Tabel 4.2 ANOVA pada Jarak ACF

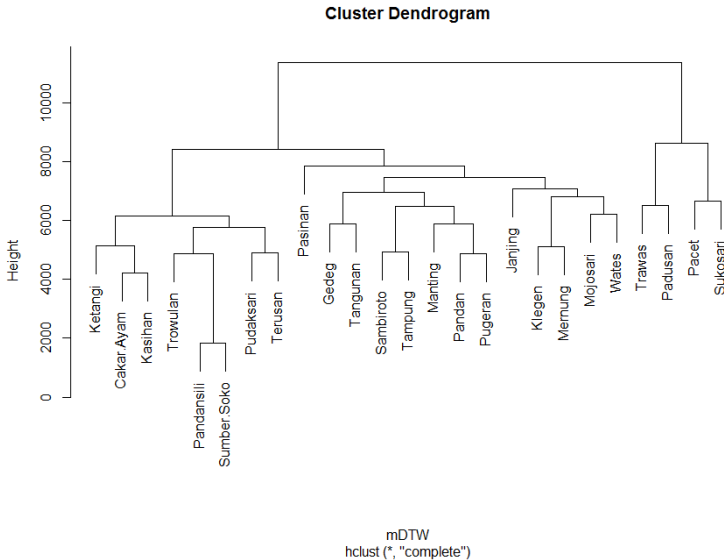
Variabel	F	Signifikansi	Keterangan
Jan	0,414	0,526	Gagal tolak
Feb	0,113	0,74	Gagal tolak
Mar	0,641	0,431	Gagal tolak
Apr	0	0,986	Gagal tolak
Mei	0,592	0,45	Gagal tolak
Jun	1,79	0,194	Gagal tolak
Jul	0,301	0,589	Gagal tolak
Ags	1,475	0,237	Gagal tolak
Sep	0,038	0,847	Gagal tolak
Okt	3,309	0,082	Gagal tolak
Nov	7,086	0,014	Tolak
Des	1,681	0,208	Gagal tolak

Dengan menggunakan taraf signifikansi 5% variabel rata rata curah hujan pada bulan November memiliki keputusan tolak H_0 . Artinya rata rata curah hujan bulan November untuk zona musim 1 dan zona musim 2 dengan menggunakan jarak ACF memiliki perbedaan. Sedangkan untuk rata rata curah hujan di bulan yang lain selain bulan November tidak memiliki perbedaan yang ditunjukkan dengan keputusan gagal tolak hipotesis nol.

4.5 Clustering dengan Jarak DTW

Pada tahap *clustering* dengan jarak DTW digunakan jumlah *cluster* sebanyak 2 hingga 5 *cluster*. Berikut merupakan

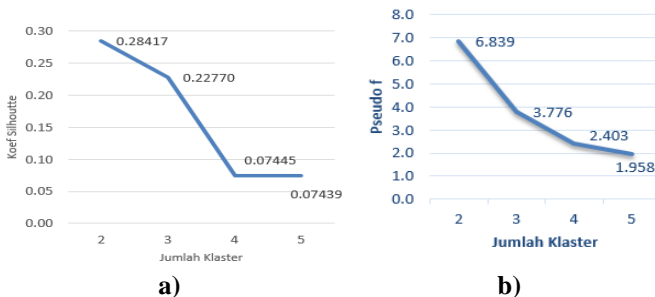
dendrogram dari pengelompokkan dengan menggunakan jarak DTW



Gambar 4.15 Dendrogram dengan Jarak DTW

Hasil dendrogram menunjukkan adanya kedekatan antar objek yang dapat dijadikan panduan objek mana yang memiliki kemiripan karakteristik. Dua objek dengan karakteristik sama akan digambarkan sebagai dua titik yang posisinya berdekatan, semakin dekat posisi dua buah objek maka semakin mirip dan sebaliknya. Dari Gambar 4.15 dapat dijelaskan apabila terbentuk 2 pengelompokkan atau 2 klaster maka anggota dari klaster 1 adalah Pacet, Trawas, Sukosari, dan Padusan. Sedangkan anggota klaster 2 adalah pos hujan selain di klaster 1. Begitu pula dengan pengelompokkan dengan 3,4,dan 5 klaster.

Untuk menentukan *cluster* optimal yang selanjutnya menjadi penentuan zona musim menggunakan kriteria koefisien *silhouette* dan *pseudo-f statistics* yang ditampilkan pada Gambar 4.16.



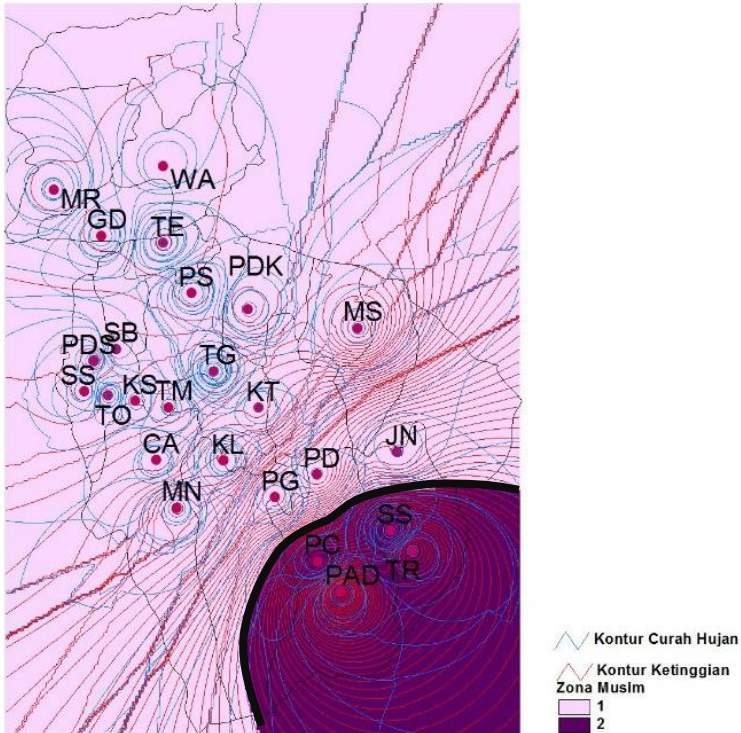
Gambar 4.16 (a) Koefisien *Silhouette* dan (b) Nilai *Pseudo-f Statistics* Hasil Pengelompokkan dengan Jarak DTW

Dari Gambar 4.16 dapat diketahui bahwa semakin banyak kelompok yang terbentuk maka nilai Koefisien *Silhouette* dan *pseudo-f statistics* semakin kecil. Hal ini dapat berarti bahwa semakin banyak kelompok yang terbentuk semakin tidak optimal. Sehingga yang dipilih adalah 2 kluster karena memiliki nilai koefisien *silhouette* dan *pseudo-f statistics* yang lebih tinggi dibandingkan 3,4,dan 5 kluster.

Anggota dari masing masing zona musim adalah sebagai berikut

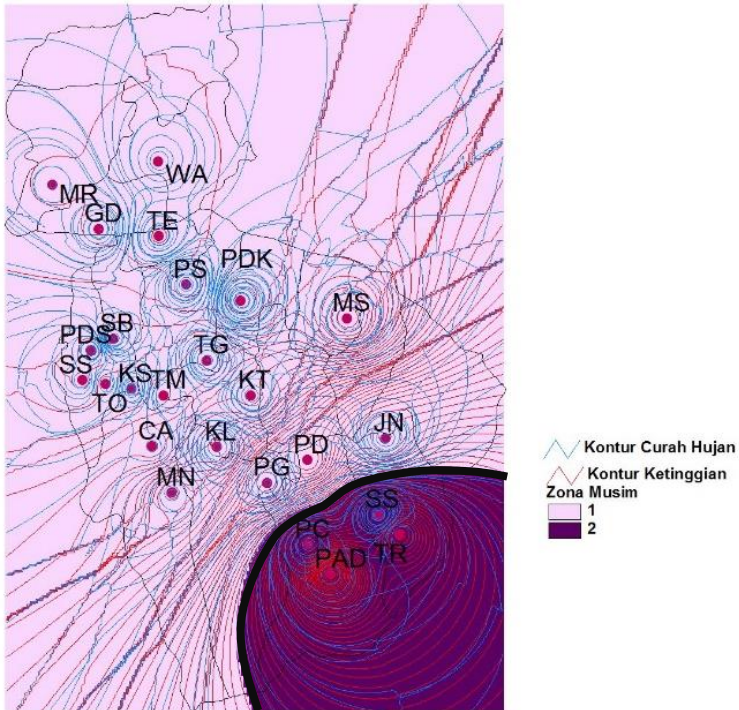
- Zona Musim 1 : Cakar Ayam, Gedeg, Janjing, Kasihan, Ketangi, Klegen, Mojosari, Pudaksari, Pandan, Pandansili, Pasinan, Pugeran, Sambiroto, Sumber Soko, Tampung, Tangunan, Trowulan, Manting, Mernung, dan Wates
- Zona Musim 2 : Pacet, Trawas, Sukosari, dan Padusan

Berikut merupakan visualisasi dari pembentukan zona musim dengan jarak DTW yang ditampilkan pada Gambar 4.17.



Gambar 4.17 Pengelompokan Zona Musim Menggunakan Jarak DTW dengan Faktor 1

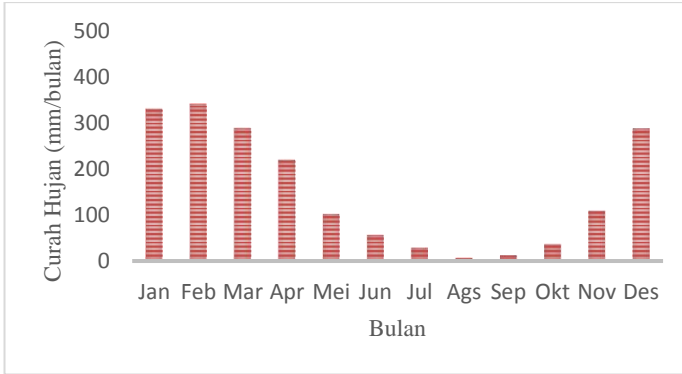
Sebagian dari pos hujan di Mojokerto berada di zona 1, sedangkan hanya ada 4 pos hujan saja yang berada di zona 2 yaitu Sukosari, Pacet, Padusan, dan Trawas. Ke empat pos hujan tersebut memiliki ketinggian yang lebih tinggi dibandingkan dengan pos hujan yang berada di zona 1. Secara visual pengelompokan dengan jarak DTW dapat menangkap pola pengelompokan dari zona musim yang terbentuk. Zona musim yang terbentuk dibedakan dari ketinggian tempat pos hujan. Artinya zona musim 1 meliputi pos hujan di daerah dataran rendah sedangkan zona musim 2 beranggotakan pos hujan yang berada di kawasan dataran tinggi.



Gambar 4.18 Pengelompokan Zona Musim Menggunakan Jarak DTW dengan Faktor 3

Gambar 4.17 dan Gambar 4.18 menunjukkan perbedaan kontur curah hujan dari faktor 1 dan faktor 3. Factor 1 memiliki kategori musim penghujan sedangkan faktor 3 merupakan kategori musim transisi dari penghujan ke musim kemarau.

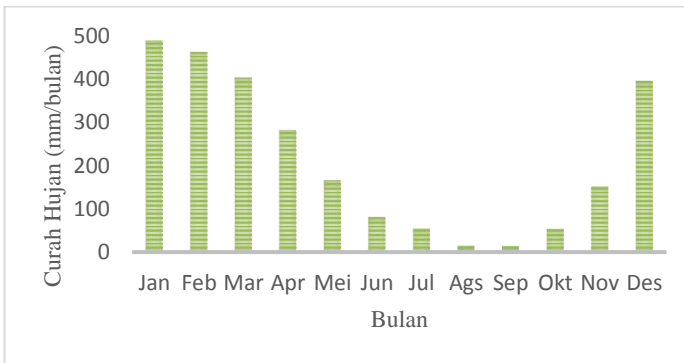
Pola curah hujan di zona 1 dengan menggunakan jarak DTW ditampilkan pada Gambar 4.19 . Zona musim 1 memiliki pola curah hujan monsunal, dimana musim hujan terjadi pada awal dan akhir tahun. Puncak musim penghujan terjadi di bulan Februari karena memiliki curah hujan tertinggi. Musim kemarau di zona 1 dimulai pada bulan Mei hingga November.



Gambar 4.19 Pola Curah Hujan Zona Musim 1 dengan Jarak DTW

Gambar 4.19 menunjukkan pola curah hujan pada zona 1, diaman pada zona 1 memiliki tipe U atau tipe monsonal. Tipe zona tersebut akan mengalami musim hujan di awal dan di akhir tahun. Pada zona 1 memiliki puncak curah hujan tertinggi atau puncak musim hujan pada bulan Januari dan Februari. Sedangkan musim kemarau diprediksi pada pertengahan bulan dimulai pada bulan Mei hingga Oktober.

Pola curah hujan di zona musim 2 dengan jarak DTW ditampilkan pada Gambar 4.20

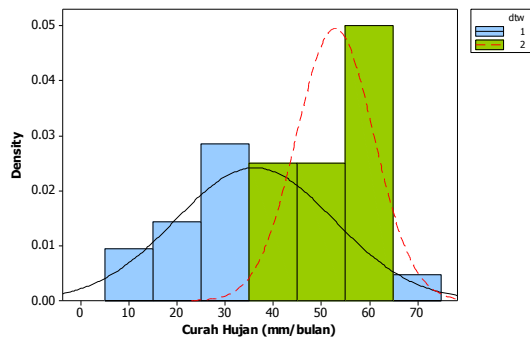


Gambar 4.20 Pola Curah Hujan Zona Musim 2 dengan Jarak DTW

Di zona musim 2 yang meliputi Pacet, Trawas, Padusan, dan

Sukosari memiliki pola curah hujan monsun. Puncak dari musim penghujan terjadi pada bulan Januari dan puncak dari musim kemarau pada bulan Agustus dan September.

Perbandingan rata rata curah hujan zona 1 dan zona 2 menggunakan jarak DTW ditampilkan pada Gambar 4.19. dari Gambar 4.21 diketahui bahwa rata rata curah hujan pada zona 2 lebih tinggi dibandingkan dengan zona 1.



Gambar 4.21 Perbandingan Rata Rata Curah Hujan Zona 1 dan Zona 2 dengan Jarak DTW

Dari histogram pada Gambar 4.21 dapat diketahui bahwa curah hujan di zona 2 lebih tinggi dibandingkan dengan zona 1. Hal ini dimungkinkan karena di zona 2 meliputi daerah yang termasuk ke dalam dataran tinggi.

Untuk melihat apakah ada perbedaan antara kedua zona musim maka dilakukan uji t. Dengan taraf signifikansi 5%, perhitungan uji t menghasilkan nilai *p-value* 0,015. Karena nilai *p-value* kurang dari 0,05 maka dapat disimpulkan adanya perbedaan rata rata curah hujan pada pengelompokkan zona musim menggunakan jarak DTW.

4.5.1 Perbandingan Hasil Pengelompokkan

Perbandingan hasil pengelompokkan dari zona musim yang terbentuk dengan menggunakan analisis ragam (ANOVA). Berikut merupakan hasil dari uji ANOVA dari variabel curah hujan pada

pengelompokkan menggunakan jarak DTW dengan taraf signifikansi 5%.

Tabel 4.3 ANOVA Pada Jarak DTW

Variabel	F	Signifikansi	Keterangan
Jan	28,995	0,000	Tolak
Feb	10,323	0,004	Tolak
Mar	21,078	0,000	Tolak
Apr	14,365	0,001	Tolak
Mei	12,292	0,002	Tolak
Jun	5,206	0,032	Tolak
Jul	11,514	0,002	Tolak
Ags	6,902	0,015	Tolak
Sep	0,371	0,548	Gagal Tolak
Okt	3,83	0,063	Gagal Tolak
Nov	7,208	0,013	Tolak
Des	21,313	0,000	Tolak

Dengan menggunakan taraf signifikansi 5% didapatkan keterangan gagal tolak hipotesis nol pada bulan September dan Oktober, hal ini memberikan kesimpulan rata rata curah hujan pada bulan September dan Oktober di zona 1 maupun zona 2 tidak memiliki perbedaan. Hal ini dapat terjadi karena pada bulan tersebut terjadi musim kemarau sehingga tidak ada perbedaan rata rata curah hujan.

Pada bulan lainnya atau selain bulan September dan Oktober memiliki keterangan tolak hipotesis nol yang berarti bahwa pada zona 1 dan zona 2 terdapat perbedaan rata rata curah hujan.

4.6 Perbandingan Hasil *Cluster Optimal* dengan Zona Musim BMKG

Terdapat 3 zona musim di Mojokerto berdasarkan pengelompokkan yang dilakukan oleh Badan Meteorologi,

Klimatologi, dan Geofisika (BMKG). Dari pengelompokan tersebut akan dibandingkan evaluasi kinerja zona musim milik BMKG dengan zona musim hasil pengelompokan menggunakan *time series based clustering*. Untuk melihat evaluasi kinerja zona musim hasil *time series based clustering* dengan zona musim milik BMKG digunakan kriteria dua nilai simpangan baku, yaitu dalam kelompok (S_W) dan antar kelompok (S_B). Metode yang dipilih adalah yang memberikan rasio nilai S_W/S_B terkecil.

Tabel 4.4 Perbandingan Hasil *Cluster* Optimal dengan Zona Musim BMKG

Metode	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Ags	Sep	Okt	Nov	Des	
ACF	SW	106.51	99.83	73.16	53.11	46.47	17.98	17.19	6.81	4.33	12.29	29.94	67.11
	SB	400.17	384.32	349.93	229.78	139.16	37.39	40.47	13.98	12.85	63.90	183.12	369.38
	Rasio	0.27	0.26	0.21	0.23	0.33	0.48	0.42	0.49	0.34	0.19	0.16	0.18
DTW	SW	43.16	49.74	36.75	28.75	31.31	25.08	14.85	4.22	5.32	12.28	31.71	42.48
	SB	574.06	526.74	464.60	315.12	200.36	93.69	67.68	18.84	14.59	61.83	174.06	453.43
	Rasio	0.08	0.09	0.08	0.09	0.16	0.27	0.22	0.22	0.36	0.20	0.18	0.09
BMKG	SW	41.93	47.20	39.79	27.72	32.59	20.83	13.20	4.68	5.53	13.79	31.06	39.78
	SB	312.05	335.35	256.17	158.18	100.23	57.06	34.12	8.21	8.05	35.50	93.78	261.26
	Rasio	0.13	0.14	0.16	0.18	0.33	0.37	0.39	0.57	0.69	0.39	0.33	0.15

Telah diketahui bahwa semakin kecil nilai S_W dan semakin besar nilai S_B maka metode tersebut memiliki kinerja yang baik. Dari Tabel 4.4 diketahui bahwa metode pengelompokan dengan jarak DTW memiliki nilai Rasio S_W/S_B paling kecil diantara pengelompokan dengan jarak ACF dan milik BMKG. Hal ini berarti metode dengan jarak DTW merupakan metode terbaik untuk mengelompokkan zona musim di Mojokerto. Karena simpangan baku dalam kelompok untuk metode pengelompokan dengan jarak DTW lebih kecil berarti keragaman antar objek semakin homogen. Sedangkan simpangan baku antar kelompok lebih besar yang berarti keragaman antar kelompok semakin heterogen.

(Halaman ini sengaja dikosongkan

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil analisis yang telah dilakukan, diperoleh hasil sebagai berikut:

1. Karakteristik curah hujan di Mojokerto dapat dilihat melalui rata-rata curah hujan. Dari rata-rata curah hujan per bulan selama 7 tahun dapat diketahui bahwa curah hujan di Mojokerto memiliki pola bahwa daerah dataran tinggi memiliki intensitas curah hujan lebih tinggi dibandingkan dengan daerah dataran rendah. Di Mojokerto puncak curah hujan tertinggi berada di bulan Januari dan Februari. Dan puncak curah hujan terendah pada bulan Agustus dan September. Pola curah hujan di Mojokerto memiliki tipe musonal.
2. Pengelompokan zona musim di Mojokerto menggunakan metode pengelompokan jarak ACF dan DTW. Penentuan kluster optimum menggunakan koefisien *silhouette* dan *pseudo-statistics*. Pada jarak ACF kluster optimum yang dihasilkan sebanyak 2 zona musim dan pada jarak DTW kluster optimum yang dihasilkan juga 2 zona musim.
3. Evaluasi kinerja zona musim dilakukan dengan rasio simpangan baku dimana metode yang dibandingkan adalah pengelompokan dengan jarak ACF, DTW, dan hasil pengelompokan oleh BMKG. Berdasarkan hasil evaluasi kinerja zona musim dengan rasio simpangan baku metode terbaik yang dipilih adalah pengelompokan dengan jarak DTW karena memiliki rasio simpangan bakun paling kecil yang berarti keragaman antar objek semakin homogen dan keragaman antar kelompok semakin heterogen. Berikut hasil pengelompokan dengan metode terbaik yaitu dengan jarak DTW
 - Zona Musim 1: Cakar Ayam, Gedeg, Janjing, Kasihan,

Ketangi, Klegen, Mojosari, Pandan, Pandansili, Pasinan, Pudaksari, Pugeran, Sambiroto, Tampung, Tangunan, Terusan, Trowulan, Manting, Mernung, dan Wates

- Zona Musim 2 : Pacet, Trawas, Padusan, dan Sukosari.

5.2 Saran

Berdasarkan kesimpulan yang diperoleh dapat diberikan saran dalam penelitian ini penggunaan ACF secara visualisasi kurang menangkap pola pengelompokan dari zona musim, oleh karena itu pada penelitian selanjutnya bisa menggunakan jarak yang lain dalam menganalisis pengelompokan zona musim dengan metode *time series based clustering*. Agar lebih akurat dalam pengelompokan zona musim dapat digunakan seluruh wilayah Jawa Timur karena pengelompokan zona musim tidak mengenal batas administrasi.

DAFTAR PUSTAKA

- Aghabozorgi, S., Shirkhorshidi, A., & Wah, T. (2015). Time Series Clustering - A decade review. *Information System* ,53 , 16-38.
- Ahrens, B. (2006). Distance in Spatial Interpolation of Daily Rain Gauge Data. *Journal of Hydrologi Earth System Sciences* , 10, 197-208.
- Aldrian, E., & R.D, S. (2003). Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *International Journal of Climatology* , 23, 1435-1452.
- BMKG. (2017, Mei 7). Diakses Juli 18, 2018, from <http://juanda.jatim.bmkg.go.id>
- BMKG. (2019). Diakses Feb 14, 2019, from <https://www.bmkg.go.id/iklim/perubahan-normal-curah-hujan.bmkg>.
- BMKG. (2019). Diakses Feb 14, 2018, from <https://www.bmkg.go.id/iklim/perubahan-normal-curah-hujan.bmkg>
- BMKG. (2019). Diakses February 13, 2019, from <https://karangploso.jatim.bmkg.go.id/index.php/informasi-iklim/peta-zona-musim-di-provinsi-jawa-timur>
- BMKG. (2012). *Denpasar.bmkg.go.id*. Diakses October 19, 2018, from <http://balai3.denpasar.bmkg.go.id/daftar-istilah-musim>
- Budyanto, Eko.(2010).*Sistem Informasi Geografis dengan ArcView GIS*. Yogyakarta: Penerbit ANDI
- Bunkers W.J., Miller J.R., DeGaetano A.T. 1996. Definition of Climate Regions in the Northern Plains Using an Objective

- Cluster Modification Technique. *Journal of Climate* 9:130-146.
- Chis, M., Baenerjee, S., & Hassanien, A. (2009). Clustering Time Series Data: An Evolutionary Approach. *Foundations of Computational* , 6, 193-207.
- Draper, N.R., & Smith, H. (1992). *Analisis Regresi Terapan Edisi Kedua*. Jakarta: PT. Gramedia Pustaka.
- D.T, Kusma., & N.A.-U.B, Luhur. (2015). Prototipe Komparasi Model Clustering Menggunakan Metode K-Means dan FCM untuk menentukan Strategi Promosi. *Jurnal TI Com* , 3 (Sekolah Tinggi Teknik - PLN Jakarta).
- Galeano, P., & Pena, D. (2000). *Multivariate Analysis in Vector Time Series*. Resenhas,383-403
- Hair, J., Anderson, R., Tatham, R., & Black, W. (2010). *Multivariate Data Analysis (7 ed.)*. USA: Pearson College Div.
- Hasanah, N., Ugiarto, M., & Puspitasari, N. (2017). Sistem Pengelompokan Curah Hujan Menggunakan Metode K-Means di Wilayah Kalimantan Timur. *Prosiding Seminar Nasional Ilmu Komputer dan Teknologi Informasi* , 2, 2.
- Haynes, B. (1947). *techniques of Observing Weather*. New York: John Wiley & Sons.
- Hung, C. M., Chang, J., & Yang, D. (2005). An Efficient K-Means Clustering Algorithm Using Simple Partitioning. *Journal of Information Science and Engineering* , XXI (1), 57-77.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. United States of America: Pearson Education, Inc.

- J. MacQueen, Somemethod for classification and analysis of multivariate observations, in: L.M. LeCam, J. Neyman (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol 1, pp. 281-297
- Kaufman, L., & Rousseeuw, P. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley.
- Liao, T. (2005). Clustering of time series data-a survey. *Pattern Recognition Society* 38 , 1857-1874.
- Liu, Y., & Liu, L. (2016). Rainfall Feature Extraction Using Cluster Analysis and Its Application on Displacement Prediction for a Cleavage-Parallel Landslide in the Three - Gorges Reservoir Area. *Natural Hazards and Earth System Sciences* .
- Niennattrakul, V., & Ratanamahatana, C. (2007). On Clustering Multimedia Time and Dynamic Time Warping. In *Multimedia and Ubiquitous Engineering, 2007. MUE '07. International Conference on, 2007*,pp. 733-738.
- Orpin, A., & Kostylev, V. (2006). Towards a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats.
- Prawiwardoyo, S. (2007). *Meteorologi*. Bandung: Institut Teknologi Bandung.
- Rani, S., & Sikka, G. (2012). Recent Techniques of Clustering of Time Series Data: A survey. *International Journal of Computer Applications* , 52, pp. 1-9
- Rencher, & Alvin, C. (2002). *Methods of Multivariate Analysis Second Edition* (2nd penyunt ed.). (Wiley-Interscience, Ed.) United States of America.
- Riyadi, M.A.A, Pratiwi, D. S., Irawan, A. R., & Fithriasari, K. (2017, November). Clustering stationary and non-stationary

time series based on autocorrelation of hierarchical and k-means algorithms. *International Journal of Advances in Intelligent Informatics*, 3, 154-160.

Singh, W. (2008). *Robustness of Three Hierarchical Agglomerative Clustering Techniques for Ecological Data*. Iceland: University of Iceland.

Suryanto, J. (2017). Analisa Perbandingan Pengelompokkan Curah Hujan 15 Harian Provinsi DIY Menggunakan Fuzzy Clustering dan K-Means Clustering. *Jurnal AGRIFOR*, XVI, 2.

Sutanto, H. (2009). Cluster Analysis. *Jurnal Pendidikan Matematika, FMIPA UNY* 3, pp. 6-8.

Tjasjono, B. (2004). *Klimatologi Umum*. Bandung: ITB.

Wang,X., Smith,K., Hyndman, R. (2006). Characteristic-Based Clustering for Time Series Data. *Data Mining and Knowledge Discovery (13)*, 335-364. Springer Science + Bussines Media,LLC Manufactures in the United States.

WMO.2007. *The Role Of Climatological Normals In A Changing Climate, World Climate Data and Monitoring Program World Meteorology Organization*. Geneva

Wei, W. W., 2006. *Time Series Analysis Univariate and Multivariate Methods*. 2nd Edition. United States of America: Pearson Education, Inc

LAMPIRAN

Lampiran 1. Data Curah Hujan Bulanan (mm/bulan)

Pos Hujan	Jan 2011	Feb 2011	Mar 2011	...	Oct 2017	Nov 2017	Des 2017
Cakar Ayam	282	226	184	...	24	252	247
Gedeg	363	210	399	...	52	523	445
Janjing	373	279	257	...	34	339	159
Kasihani	339	189	221	...	10	135	239
Ketangi	273	222	179	...	43	178	311
Klegen	300	395	160	...	0	93	289
Mojosari	159	258	427	...	0	163	216
Tampung	247	228	180	...	110	328	298
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Tangunan	338	139	330	...	47	411	272
Terusan	335	244	446	...	25	348	480
Trawas	611	365	268	...	96	444	457
Trowulan	281	240	334	...	47	596	758
Manting	374	188	148	...	12	217	369
Padusan	661	439	304	...	50	664	632
Sukosari	601	403	245	...	41	421	582
Mernung	299	270	342	...	0	0	0
Wates	314	375	328	...	108	346	145

Lampiran 2. *Syntax Cluster Jarak ACF*

```

#Memanggil Data
library(TSclust)
data=read.csv("E:/SMT8/Laporan+OlahData/cobata/fix.csv",header=TRUE)

#Menghitung Jarak ACF
mACF=diss(data,"ACF")

#Membuat Dendogram
clusterACF=hclust(mACF, "complete")
plot(clusterACF)

#Mencari Anggota Setiap Jumlah Cluster
clusterCut2 <- cutree(clusterACF, 2)
clusterCut3 <- cutree(clusterACF, 3)
clusterCut4 <- cutree(clusterACF, 4)
clusterCut5 <- cutree(clusterACF, 5)

#Mencari Koefisien Silhoutte
sil2 = silhouette (clusterCut2,mACF)
sil3 = silhouette (clusterCut3,mACF)
sil4 = silhouette (clusterCut4,mACF)
sil5 = silhouette (clusterCut5,mACF)

#Plot Koefisien Silhoutte
hasil=data.frame(mean(sil2[,3]),mean(sil3[,3]),mean(sil4[,3]),
mean(sil5[,3]))
hasil=unlist(hasil[1,], use.names=FALSE)
plot(c(0,hasil), type="b", xlab="Jumlah Cluster", ylab="Koefisien Silhouette",col = "Blue")

#Mengeluarkan Anggota Cluster
cluster2ACF=data.frame(clusterCut2, sil2[,1], sil2[,2], sil2[,3])
cluster3ACF =data.frame(clusterCut3, sil3[,1], sil3[,2], sil3[,3])
cluster4ACF =data.frame(clusterCut4, sil4[,1], sil4[,2], sil4[,3])
cluster5ACF =data.frame(clusterCut5, sil5[,1], sil5[,2], sil5[,3])
Hasil=data.frame(cluster2ACF, cluster3ACF, cluster4ACF, cluster5ACF)
write.csv(Hasil, "E:/SMT 8/Laporan+Olah Data/cobata/acffix.csv")

```

Lampiran 3. Syntax Cluster Jarak DTW

```

#Memanggil Data
library(TSclust)
data=read.csv("E:/SMT8/Laporan+OlahData/cobata/fix.csv",
header=TRUE)

#Menghitung Jarak DTW
mDTW=diss(data, "DTW")

#Membuat Dendrogram
clusterDTW=hclust(mDTW, "complete")
plot(clusterDTW)

#Mencari Anggota Setiap Jumlah Cluster
clusterCut2 <- cutree(clusterDTW, 2)
clusterCut3 <- cutree(clusterDTW, 3)
clusterCut4 <- cutree(clusterDTW, 4)
clusterCut5 <- cutree(clusterDTW, 5)

#Mencari Koefisien Silhoutte
sil2 = silhouette (clusterCut2,mDTW)
sil3 = silhouette (clusterCut3,mDTW)
sil4 = silhouette (clusterCut4,mDTW)
sil5 = silhouette (clusterCut5,mDTW)

#Plot Koefisien Silhoutte
hasil=data.frame(mean(sil2[,3]),mean(sil3[,3]),mean(sil4[,3]),
mean(sil5[,3]))
hasil=unlist(hasil[1,], use.names=FALSE)
plot(c(0,hasil), type="b", xlab="Jumlah Cluster", ylab="Koefisien Silhouette",col = "Red")

#Mengeluarkan Anggota Cluster
cluster2DTW=data.frame(clusterCut2, sil2[,1], sil2[,2], sil2[,3])
cluster3DTW=data.frame(clusterCut3, sil3[,1], sil3[,2], sil3[,3])
cluster4DTW=data.frame(clusterCut4, sil4[,1], sil4[,2], sil4[,3])
cluster5DTW=data.frame(clusterCut5, sil5[,1], sil5[,2], sil5[,3])
Hasil=data.frame(cluster2DTW, cluster3DTW, cluster4DTW, cluster5DTW)
write.csv(Hasil, "E:/SMT 8/Laporan+Olah Data/cobata/dtwfix.csv")

```

Lampiran 4. Syntax Menghitung Nilai *Pseudo-f Statistics*

```

icdrate = function(Data, nc,c)
{
  n = dim(Data)[1]
  p = dim(Data)[2]
  X = Data[,1:(p-1)]
  Group = Data[,p]

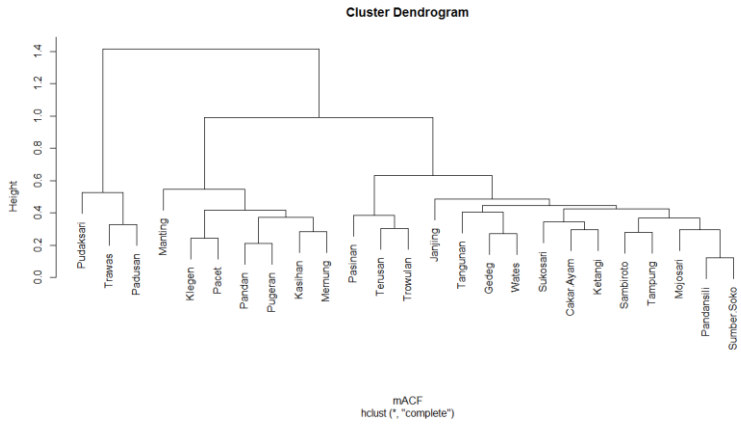
  p = dim(X)[2]
  Mean.X = matrix(ncol = p, nrow = (nc+1))
  for (i in 1:nc)
  {
    for (j in 1:p)
    {
      Mean.X[i,j] = mean(X[which(Group==i),j])
      Mean.X[(nc+1),j] = mean(X[,j])
    }
  }

  SST = matrix(ncol=p, nrow=n)
  for (i in 1:n)
  {
    for (j in 1:p)
    {
      SST[i,j] = (X[i,j] - Mean.X[(nc+1),j])^2
    }
  }
  SST = sum(sum(SST))

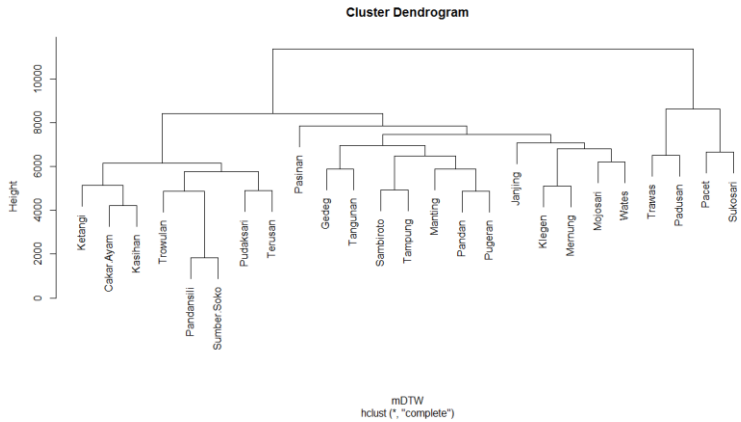
  SSE = matrix(ncol=p, nrow=n)
  for (i in 1:n)
  {
    for (j in 1:p)
    {
      for (k in 1:nc)
      {
        if (Group[i]==k)
        {
          SSE[i,j] = (X[i,j] - Mean.X[k,j])^2
        }
      }
    }
  }
  SSE = sum(sum(SSE))
  Rsq = (SST-SSE)/SST
  icdrate = 1-Rsq
  Pseudof = (Rsq/(c-1))/((icdrate)/(nc-c))
  list(Rsq=Rsq, icdrate=icdrate,Pseudof=Pseudof)}

```

Lampiran 5. Dendrogram Dendrogram Jarak ACF



Dendrogram Jarak DTW



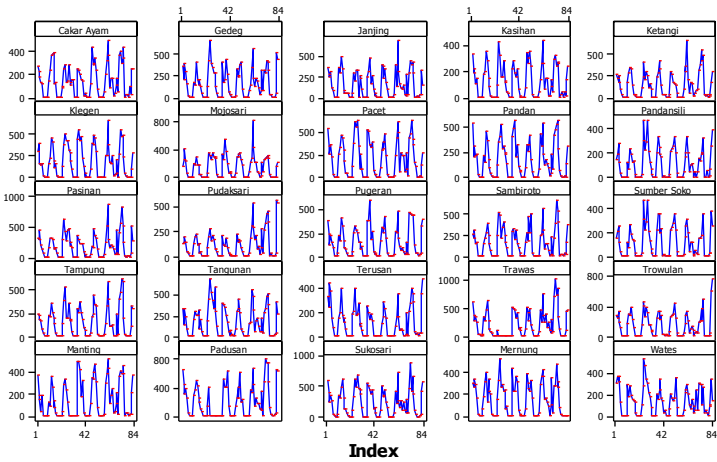
Lampiran 6. Nilai Koefisien Silhoutte Jarak ACF 2-5 Cluster

Pos Hujan	Cluster 2	sil2...3.	Cluster Cut3	sil3...3.	Cluster Cut4	sil4...3.	Cluster Cut5	sil5...3.
Cakar.Ayam	1	0.4997	1	0.199	1	0.236	1	0.236
Gedeg	1	0.4937	1	0.324	1	0.134	1	0.134
Janjing	1	0.3566	1	0.242	1	0.025	1	0.025
Kasihlan	1	0.5910	2	0.344	2	0.268	2	0.140
Ketangi	1	0.5630	1	-0.003	1	0.103	1	0.083
Klegen	1	0.5505	2	0.440	2	0.383	2	0.348
Mojosari	1	0.5894	1	-0.091	1	0.028	1	-0.010
Pacet	1	0.5727	2	0.446	2	0.384	2	0.318
...
Tampung	1	0.5738	1	0.032	1	0.111	1	0.083
Tangunan	1	0.4621	1	0.282	1	0.124	1	0.124
Terusan	1	-0.0743	1	0.092	3	0.387	3	0.387
Trawas	2	0.5487	3	0.504	4	0.377	4	0.377
Trowulan	1	0.2607	1	0.394	3	0.179	3	0.179
Manting	1	0.4382	2	0.176	2	0.121	5	0.000
Padusan	2	0.5932	3	0.546	4	0.381	4	0.381
Sukosari	1	0.5313	1	0.231	1	0.281	1	0.262
Mernung	1	0.5796	2	0.275	2	0.191	2	0.212
Wates	1	0.4388	1	0.320	1	0.153	1	0.153

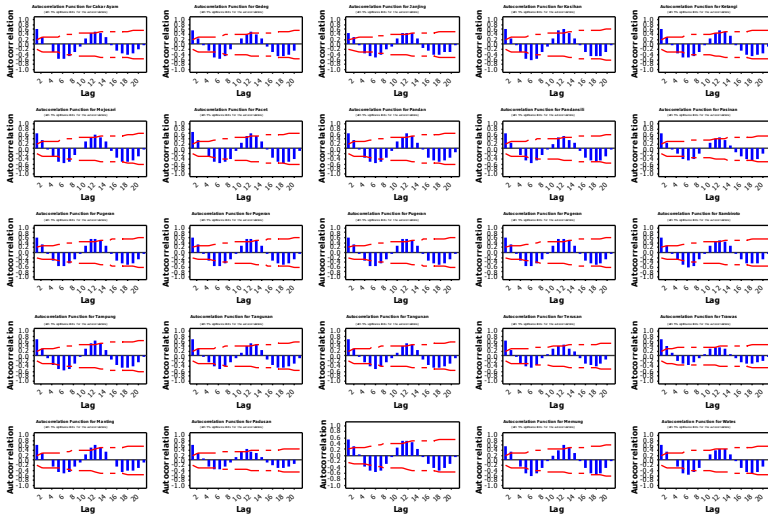
Lampiran 7. Nilai Koefisien Silhoutte Jarak DTW 2-5 Cluster

Pos Hujan	Cluster Cut2	sil2...3.	Cluster Cut3	sil3...3.	Cluster Cut4	sil4...3.	Cluster Cut5	sil5...3.
Cakar.Ayam	1	0.3871	1	0.371	1	0.169	1	0.163
Gedeg	1	0.2499	1	0.183	2	-0.033	2	-0.022
Janjing	1	0.2086	1	0.164	2	-0.005	2	0.003
Kasihani	1	0.3907	1	0.354	1	0.173	1	0.162
Ketangi	1	0.3265	1	0.280	1	0.105	1	0.091
Klegen	1	0.1772	1	0.105	2	0.099	2	0.105
Mojosari	1	0.3071	1	0.267	2	-0.070	2	-0.056
Pacet	2	0.0757	2	0.148	3	0.115	3	0.104
Pandan	1	0.1558	1	0.050	2	0.090	2	0.107
Pandansili	1	0.4344	1	0.421	1	0.298	1	0.287
Pasinan	1	0.1616	1	0.127	2	0.026	4	0.000
Pudaksari	1	0.3772	1	0.376	1	0.223	1	0.219
Trawas	2	0.2403	3	0.165	4	0.165	5	0.165
Trowulan	1	0.3044	1	0.302	1	0.216	1	0.204
Manting	1	0.2454	1	0.184	2	-0.020	2	-0.007
Padusan	2	0.2871	3	0.061	4	0.061	5	0.061
Sukosari	2	0.2518	2	0.031	3	0.031	3	0.031
Mernung	1	0.2900	1	0.226	2	0.027	2	0.043
Wates	1	0.3451	1	0.280	2	-0.132	2	-0.122

Lampiran 8. Time Series Plot Pos Hujan di Mojokerto

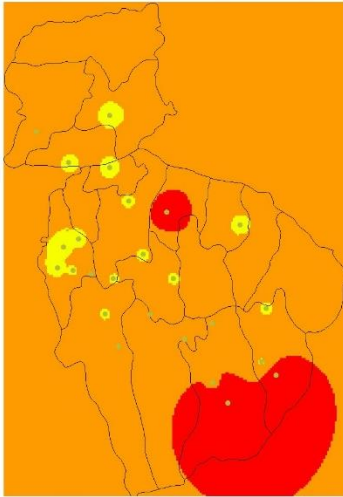


Lampiran 9. Plot ACF Seluruh Pos Hujan Mojokerto

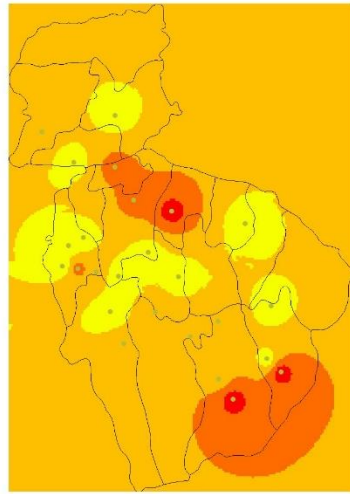


Lampiran 10. Zona Musim dengan Jarak ACF

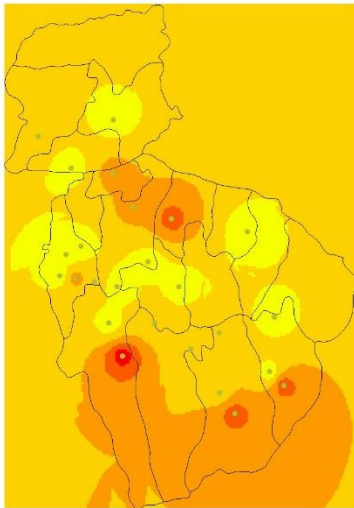
3 Zona Musim



4 Zona Musim

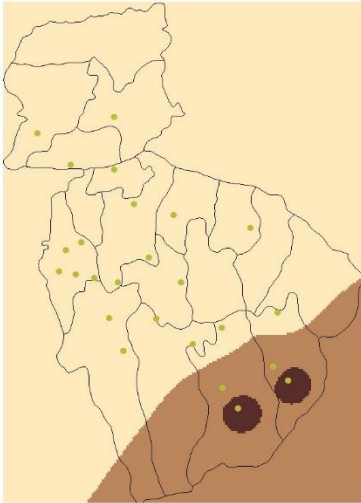


5 Zona Musim

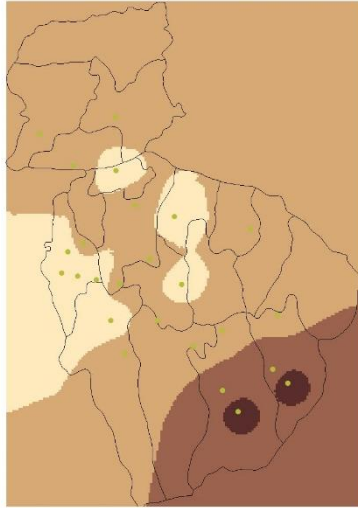


Lampiran 11. Zona Musim dengan Jarak DTW

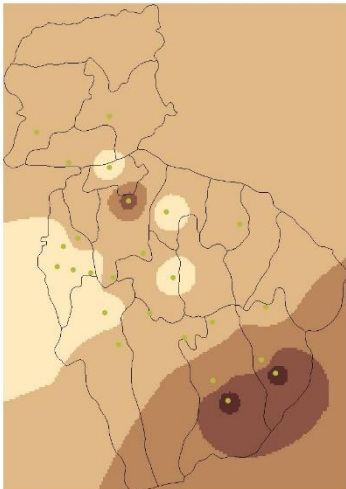
3 Zona Musim



4 Zona Musim



5 Zona Musim



Lampiran 12. Data Pos Hujan

Pos Hujan	Singkatan	Longitude	Latitude	Elevasi (m)
Cakar Ayam	CA	112.4279	-7.59991	105
Gedeg	GD	112.3929	-7.45635	39
Janjing	JN	112.5821	-7.59464	379
Kasihani	KS	112.4143	-7.56205	72
Ketangi	KT	112.4934	-7.56637	116
Klegen	KL	112.4709	-7.60029	149
Mojosari	MS	112.557	-7.51542	33
Pacet	PC	112.5313	-7.66478	594
Pandan	PD	112.5307	-7.60895	299
Pandansili	PDS	112.3879	-7.53644	43
Pasinan	PS	112.4506	-7.4931	36
Pudaksari	PDK	112.4868	-7.50362	34
Pugeran	PG	112.5041	-7.62392	306
Sambiroto	SB	112.4022	-7.52888	63
Sumber Soko	SS	112.3821	-7.5561	57
Tampung	TM	112.4357	-7.56618	83
Tangunan	TG	112.4643	-7.5432	85
Terusan	TE	112.4324	-7.461	29
Trawas	TR	112.5915	-7.6581	707
Trowulan	TO	112.3973	-7.5587	74
Manting	MN	112.4409	-7.63058	202
Padusan	PAD	112.5457	-7.68431	850
Sukosari	SS	112.5777	-7.6452	600
Mernung	MR	112.3622	-7.42713	39
Wates	WA	112.4323	-7.4117	33

Lampiran 13. Output SPSS Analisis Faktor**KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.753
Bartlett's Test of Sphericity	Approx. Chi-Square	969.062
	df	66
	Sig.	.000

Communalities

	Initial	Extraction
Jan	1.000	.680
Feb	1.000	.731
Mar	1.000	.717
Apr	1.000	.550
Mei	1.000	.718
Jun	1.000	.774
Jul	1.000	.805
Ags	1.000	.607
Sep	1.000	.763
Okt	1.000	.866
Nov	1.000	.756
Des	1.000	.838

Extraction Method: Principal
Component Analysis.

Lampiran 14. ANOVA Pada Jarak ACF

		<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
Jan	<i>Between Groups</i>	2684.012	1	2684.012	0.414	0.526
	<i>Within Groups</i>	149122.76	23	6483.598		
	<i>Total</i>	151806.77	24			
Feb	<i>Between Groups</i>	781.332	1	781.332	0.113	0.74
	<i>Within Groups</i>	159209.4	23	6922.148		
	<i>Total</i>	159990.73	24			
Mar	<i>Between Groups</i>	2520.063	1	2520.063	0.641	0.431
	<i>Within Groups</i>	90367.558	23	3929.024		
	<i>Total</i>	92887.621	24			
Apr	<i>Between Groups</i>	0.47	1	0.47	0	0.986
	<i>Within Groups</i>	34623.856	23	1505.385		
	<i>Total</i>	34624.327	24			
Mei	<i>Between Groups</i>	989.869	1	989.869	0.592	0.45
	<i>Within Groups</i>	38462.724	23	1672.292		
	<i>Total</i>	39452.593	24			
Jun	<i>Between Groups</i>	784.087	1	784.087	1.79	0.194
	<i>Within Groups</i>	10072.471	23	437.934		
	<i>Total</i>	10856.558	24			
Jul	<i>Between Groups</i>	82.906	1	82.906	0.301	0.589
	<i>Within Groups</i>	6342.773	23	275.773		

	<i>Total</i>	6425.678	24			
	<i>Between</i>					
	<i>Groups</i>	62.505	1	62.505	1.475	0.237
Ags	<i>Within</i>					
	<i>Groups</i>	974.96	23	42.39		
	<i>Total</i>	1037.464	24			
	<i>Between</i>					
	<i>Groups</i>	1.455	1	1.455	0.038	0.847
Sep	<i>Within</i>					
	<i>Groups</i>	878.434	23	38.193		
	<i>Total</i>	879.889	24			
	<i>Between</i>					
	<i>Groups</i>	828.421	1	828.421	3.309	0.082
Okt	<i>Within</i>					
	<i>Groups</i>	5758.816	23	250.383		
	<i>Total</i>	6587.238	24			
	<i>Between</i>					
	<i>Groups</i>	6027.827	1	6027.827	7.086	0.014
Nov	<i>Within</i>					
	<i>Groups</i>	19564.978	23	850.651		
	<i>Total</i>	25592.805	24			
	<i>Between</i>					
	<i>Groups</i>	5561.034	1	5561.034	1.681	0.208
Des	<i>Within</i>					
	<i>Groups</i>	76098.212	23	3308.618		
	<i>Total</i>	81659.246	24			

Lampiran 15. ANOVA Pada Jarak DTW

		Sum of Squares	df	Mean Square	F	Sig.
JAN	Between Groups	84655.48	1	84655.48	28.995	0
	Within Groups	67151.292	23	2919.621		
	Total	151806.77	24			
FEB	Between Groups	49562.288	1	49562.288	10.323	0.004
	Within Groups	110428.44	23	4801.237		
	Total	159990.73	24			
MAR	Between Groups	44418.916	1	44418.916	21.078	0
	Within Groups	48468.705	23	2107.335		
	Total	92887.621	24			
APR	Between Groups	13311.371	1	13311.371	14.365	0.001
	Within Groups	21312.955	23	926.65		
	Total	34624.327	24			
MEI	Between Groups	13741.357	1	13741.357	12.292	0.002
	Within Groups	25711.236	23	1117.88		
	Total	39452.593	24			
JN	Between Groups	2003.701	1	2003.701	5.206	0.032
	Within Groups	8852.857	23	384.907		
	Total	10856.558	24			
JL	Between Groups	2143.653	1	2143.653	11.514	0.002
	Within Groups	4282.025	23	186.175		

	Total	6425.678	24			
AGS	Between Groups	239.468	1	239.468	6.902	0.015
	Within Groups	797.996	23	34.695		
	Total	1037.464	24			
SEP	Between Groups	13.971	1	13.971	0.371	0.548
	Within Groups	865.918	23	37.649		
	Total	879.889	24			
OKT	Between Groups	940.394	1	940.394	3.83	0.063
	Within Groups	5646.843	23	245.515		
	Total	6587.238	24			
NOV	Between Groups	6106.945	1	6106.945	7.208	0.013
	Within Groups	19485.86	23	847.211		
	Total	25592.805	24			
DES	Between Groups	39275.016	1	39275.016	21.313	0
	Within Groups	42384.229	23	1842.793		
	Total	81659.246	24			

LAMPIRAN 16. Surat Pernyataan Data

SURAT KETERANGAN


Saya yang bertanda tangan di bawah ini menerangkan bahwa :

- Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : Ifitah Ayundari
NRP : 062115 4000 0039

Telah mengambil data di instansi/perusahaan kami :
Nama Instansi : Dinas Pekerjaan Umum Sumber Daya Air
Provinsi Jawa Timur
Divisi/ bagian : Bidang Perencanaan Sumber Daya Air
sejak tanggal 18 April sampai dengan 18 Mei untuk
keperluan Tugas Akhir/ ~~Thesis~~ Semester ~~Gasal~~/Genap* 2018/2019.
- Tidak Keberatan/~~Keberatan~~* nama perusahaan dicantumkan dalam Tugas Akhir/ Thesis mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
- Tidak Keberatan/~~Keberatan~~* bahwa hasil analisis data dari perusahaan dipublikasikan dalam E journal ITS yaitu Jurnal Sains dan Seni ITS.

Surabaya, 2 Juli 2019
Kasie Hidrologi dan Pemantauan
Kualitas Air
Bidang Perencanaan SDA


 (Catur Arik / K., S.T, M.Eng)
 NIP. 10790325 200501 2 009

*(coret yang tidak perlu)

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Penulis bernama lengkap Iftitah Ayundari ini dilahirkan di Jombang, 25 Januari 1997. Merupakan anak pertama dari dua bersaudara. Penulis menempuh pendidikan formal di SDN Megaluh, SMPN 2 Megaluh, dan SMAN 2 Jombang. Kemudian penulis diterima sebagai mahasiswa Departemen Statistika ITS melalui jalur SNMPTN pada tahun 2015. Selama masa perkuliahan penulis aktif dalam organisasi lingkup jurusan sebagai staff *Professional Statistics* (2016/2017) dan bendahara PSt (2017/2018). Selain itu, penulis juga aktif dalam kepanitian diantaranya Sie Acara SPORTIF (2015), Sie Perkap (PRS 2017), Mentor Gerigi (2017), dan Sie Kesehatan ISCO (2017). Penulis juga cukup aktif dalam beberapa kegiatan pengaplikasian ilmu statistik diantaranya survey motor oleh MPM, survey pelanggan Trans 7, dan survey perilaku konsumen yang diselenggarakan oleh Bappeko. Apabila pembaca ingin memberi kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir dapat menghubungi penulis melalui nomor telepon 085330502725 atau iayundari@gmail.com