

TIME SERIES CLUSTERING DENGAN MAHALANOBIS DISTANCE-BASED DYNAMIC TIME WARPING PADA PERAMALAN TINGKAT GEROMBOL HARGA BAWANG MERAH PROVINSI DI INDONESIA

PINGKAN FEBBE FIORELA KEREH



**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “*Time Series Clustering* dengan *Mahalanobis Distance-Based Dynamic Time Warping* pada Peramalan Tingkat Gerombol Harga Bawang Merah Provinsi di Indonesia” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Oktober 2025

Pingkan Febbe Fiorela Kereh
G1401211087

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

PINGKAN FEBBE FIORELA KEREH. *Time Series Clustering* dengan *Mahalanobis Distance-Based Dynamic Time Warping* pada Peramalan Tingkat Gerombol Harga Bawang Merah Provinsi di Indonesia. Dibimbing oleh ANWAR FITRIANTO dan AJI HAMIM WIGENA.

Bawang merah menjadi salah satu komoditas hortikultura strategis yang memiliki tingkat permintaan yang cukup tinggi di Indonesia dan produksinya dinilai sangat lambat karena sifatnya yang musiman. Hal tersebut berdampak pada terjadinya fluktuasi harga yang menyebabkan harga bawang merah menjadi tidak stabil. Penelitian ini bertujuan menggerombolkan provinsi berdasarkan pola pergerakan harga bawang merah yang serupa melalui pendekatan *time series clustering* dengan jarak *Mahalanobis distance-based dynamic time warping* (MDDTW) dan melakukan peramalan tingkat gerombol menggunakan metode *seasonal autoregressive integrated moving average* (SARIMA). Data yang digunakan dalam penelitian ini adalah harga bawang merah mingguan dari 34 provinsi di Indonesia yang bersumber dari Portal Satu Data Kementerian Perdagangan Republik Indonesia. Analisis data dilakukan dengan bantuan *software Rstudio*. Hasil penelitian menunjukkan pautan rataan sebagai pautan terbaik yang menggerombolkan data harga bawang merah dari 34 provinsi ke dalam tiga gerombol. Pemodelan deret waktu dilakukan pada data *prototype* setiap gerombol. Dari keseluruhan model tentatif setiap *prototype*, dipilih dua kandidat model terbaik berdasarkan nilai AIC terkecil. Seluruh model tersebut dievaluasi performanya dengan menerapkan skema *expanding window* dan diperoleh nilai rataan MAPE dari setiap model. Hasil peramalan setiap gerombol menunjukkan akurasi yang baik, dengan nilai rataan MAPE di bawah 20%.

Kata kunci: bawang merah, MDDTW, peramalan, SARIMA, *time series clustering*.

ABSTRACT

PINGKAN FEBBE FIORELA KEREH. Time Series Clustering with Mahalanobis Distance-Based Dynamic Time Warping in Forecasting the Cluster Level of Provinces Shallot Prices in Indonesia. Supervised by ANWAR FITRIANTO and AJI HAMIM WIGENA.

Shallot are one of the strategic horticultural commodities that have a fairly high demand in Indonesia, and their production is considered very slow due to their seasonal nature. This has an impact on price fluctuations, causing shallot prices to become unstable. This study aims to cluster provinces based on similar shallot prices movement patterns using a time series clustering approach with Mahalanobis distance-based dynamic time warping (MDDTW) and to forecast cluster level using the seasonal autoregressive integrated moving average (SARIMA) method. The data used in this study are weekly shallot prices from 34 provinces in Indonesia sourced from the One Data Portal The Ministry of Trade Republic of Indonesia. Data analysis was performed using Rstudio software. The results show that the average linkage is the best linkage that clusters shallot price data from 34 provinces into three clusters. Time series modeling was performed on the prototype data of each cluster. From all tentative models of each prototype, the two best model candidates were selected based on the smallest AIC value. All models were evaluated for performance by applying an expanding window scheme, and the average MAPE value of each model was obtained. The forecasting results for each cluster showed good accuracy, with an average MAPE value below 20%.

Keywords: forecasting, MDDTW, SARIMA, shallot, time series clustering.



@Hak cipta milik IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

TIME SERIES CLUSTERING DENGAN MAHALANOBIS DISTANCE-BASED DYNAMIC TIME WARPING PADA PERAMALAN TINGKAT GEROMBOL HARGA BAWANG MERAH PROVINSI DI INDONESIA

PINGKAN FEBBE FIORELA KEREH

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Statistika dan Sains Data

**PROGRAM STUDI SARJANA STATISTIKA DAN SAINS DATA
SEKOLAH SAINS DATA, MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
Ir. Mohammad Masjkur, M.S

Judul Skripsi : *Time Series Clustering dengan Mahalanobis Distance-Based Dynamic Time Warping* pada Peramalan Tingkat Gerombol Harga Bawang Merah Provinsi di Indonesia

Nama : Pingkan Febbe Fiorela Kereh

NIM : G1401211087

Disetujui oleh

Pembimbing 1:

Dr. Anwar Fitrianto, S.Si., M.Sc



Pembimbing 2:

Prof. Dr. Ir. Aji Hamim Wigena, M.Sc



Diketahui oleh

Ketua Program Studi:

Dr. Bagus Sartono, S.Si., M.Si.

NIP 197804112005011002



Tanggal Ujian:
15 Oktober 2025

Tanggal Lulus:



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Februari 2025 sampai bulan Oktober 2025 ini ialah peramalan data deret waktu tingkat gerombol, dengan judul “*Time Series Clustering dengan Mahalanobis Distance-Based Dynamic Time Warping* pada Peramalan Tingkat Gerombol Harga Bawang Merah Provinsi di Indonesia”.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa tidak akan dapat terselesaikan tanpa adanya dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Moudy Eldy Kereh, Ibu Eunike Rinewati Suhardy, Hizkia Rinaldy Kereh, dan Laudia Febby Fioreta Kereh selaku keluarga penulis yang tiada henti memberikan doa, dukungan, dan motivasi.
2. Bapak Dr. Anwar Fitrianto, S.Si., M.Sc. dan Bapak Prof. Dr. Ir. Aji Hamim Wigena, M.Sc. selaku dosen pembimbing yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan arahan, bimbingan, serta motivasi dalam proses penyusunan karya ilmiah.
3. Bapak Ir. Mohammad Masjkur, M.S. selaku dosen penguji luar sekaligus dosen moderator seminar hasil, dan Bapak Dr. Agus Mohamad Soleh, S.Si., M.T. selaku dosen moderator kolokium atas saran dan masukan yang diberikan dalam rangka perbaikan penulisan karya ilmiah.
4. Seluruh dosen dan staf akademik di lingkungan Program Studi Statistika dan Sains Data SSMI IPB University yang telah memberikan ilmu dan pengalaman berharga selama masa studi penulis.
5. George Imanuel Heumasse yang selalu memberikan doa, dukungan, semangat, motivasi, serta kesabaran dalam mendampingi penulis selama proses penyusunan karya ilmiah.
6. Dindana, Rani, Nazuwa, Amalia, Zafira, Hanifa dan seluruh teman-teman Statistika IPB angkatan 58 yang telah berjuang bersama dan saling menguatkan selama berlangsungnya kehidupan perkuliahan penulis.
7. Teman-teman seperjuangan dan semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan bantuan, semangat, serta kebersamaan yang berarti dalam penyelesaian karya ilmiah ini.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Oktober 2025

Pingkan Febbe Fiorela Kereh



— Bogor Indonesia —

IPB University

DAFTAR ISI

DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiii
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	3
II TINJAUAN PUSTAKA	4
2.1 <i>Time Series Clustering</i>	4
2.2 <i>Mahalanobis Distance-Based Dynamic Time Warping</i>	6
2.3 Evaluasi Ukuran Kemiripan	8
2.4 Ukuran Ketepatan Jumlah Gerombol	8
2.5 Pemodelan Deret Waktu Tingkat Gerombol	10
2.6 <i>Time Series K-Fold Cross Validation</i>	13
2.7 Ukuran Evaluasi Performa Model	14
III METODE	15
3.1 Data	15
3.2 Prosedur Analisis Data	15
IV HASIL DAN PEMBAHASAN	17
4.1 Praproses dan Eksplorasi Data	17
4.2 Penggerombolan Deret Waktu	19
4.3 Pemodelan Deret Waktu Tingkat Gerombol	25
4.4 Peramalan Periode Mendatang	30
V SIMPULAN DAN SARAN	32
5.1 Simpulan	32
5.2 Saran	32
DAFTAR PUSTAKA	33
LAMPIRAN	36
RIWAYAT HIDUP	43

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR TABEL

1	Kategori penggerombolan berdasarkan nilai koefisien <i>silhouette</i>	9
2	Kriteria identifikasi model SARIMA non musiman	12
3	Kriteria identifikasi model SARIMA musiman	12
4	Kategori nilai MAPE	14
5	Matriks hasil perhitungan jarak MDDTW	19
6	Hasil perhitungan koefisien korelasi <i>cophenetic</i> setiap pautan	20
7	Rincian anggota setiap gerombol	22
8	Hasil uji ADF data <i>prototype</i> setiap gerombol	27
9	Model tentatif setiap <i>prototype</i> berdasarkan plot ACF dan PACF	27
10	Nilai MAPE hasil evaluasi model gerombol A	28
11	Nilai MAPE hasil evaluasi model gerombol B	29
12	Nilai MAPE hasil evaluasi model gerombol C	29
13	Hasil uji diagnostik keseluruhan model	29

DAFTAR GAMBAR

1	Jarak MDDTW antara dua deret waktu x dan y	8
2	Ilustrasi skema <i>expanding window</i>	14
3	Diagram alir prosedur analisis data	15
4	Plot deret waktu harga bawang merah mingguan	17
5	Plot kotak garis harga bawang merah mingguan	18
6	Plot penentuan jumlah gerombol optimal	21
7	Peta hasil penggerombolan pautan rata-rata dengan jarak MDDTW sebanyak tiga gerombol	22
8	Plot deret waktu gerombol A	23
9	Plot deret waktu gerombol B	24
10	Plot deret waktu gerombol C	24
11	Plot deret waktu <i>prototype</i> gerombol A	25
12	Plot deret waktu data latih <i>prototype</i> gerombol A setelah dilakukan transformasi <i>Box-Cox</i>	26
13	Skema <i>expanding window</i> pada data latih setiap <i>prototype</i>	28
14	Plot hasil peramalan	30

DAFTAR LAMPIRAN

1	Lampiran 1 Matriks hasil perhitungan jarak MDDTW	37
2	Lampiran 2 Plot deret waktu <i>prototype</i> setiap gerombol	38
3	Lampiran 3 Plot <i>Box-Cox</i> data <i>prototype</i> setiap gerombol	39
4	Lampiran 4 Plot ACF sebelum dan setelah dilakukan <i>differencing</i>	40
5	Lampiran 5 Plot ACF dan PACF dari <i>prototype</i> yang sudah stasioner	41
6	Lampiran 6 Histogram sisaan model terbaik setiap gerombol	42