

Komparasi Sarimax dan Prophet untuk Prediksi Tingkat Pemukonnan Kamar Hotel Jawa Tengah

*Comparison Of Sarimax and Prophet for Predicting The Hotel Room Occupancy Rate
In Central Java*

Rini Indriyati¹, Auliya Burhanuddin², Annisaa Utami³

^{1,2,3}Jurusan Teknik Informatika, Fakultas Informatika, Institut Teknologi Telkom Purwokerto,
Purwokerto

E-mail: ¹1910260@ittelkom-pwt.ac.id, ²auliia@ittelkom-pwt.ac.id, ³annisaa@ittelkom-
pwt.ac.id

Abstrak

Pemerintah Jawa Tengah memprioritaskan pemulihan pariwisata akibat Covid-19. Tingkat pemukonnan kamar (TPK) hotel dapat menjadi pertanda keberhasilan minat wisatawan terhadap daerah yang dikunjungi dan salah satu *leading indicator* dalam mengukur ekonomi pada sektor pariwisata. TPK hotel Jawa Tengah dipublikasikan oleh Badan Pusat Statistik (BPS) namun publikasi membutuhkan waktu satu atau dua bulan berikutnya. Pada penelitian ini melakukan komparasi model *Sarimax* dan *Prophet* untuk prediksi TPK hotel berbintang Jawa Tengah berdasarkan nilai *MAPE* dan waktu eksekusi program. Model dibuat menggunakan 2 kategori *dataset* yang telah di cek stasionaritasnya. *Dataset* 1 yaitu data 2005-2022 (menyertakan situasi pandemi covid 19 hingga pemberhentian PPKM) dan data set 2 yaitu 2005-2019 (tanpa menyertakan situasi pandemi covid 19). Hasilnya model *Sarimax* dan *prophet* terbaik adalah model yang menggunakan *dataset* 2 dengan nilai *MAPE* 10.43% dan 47,529 detik waktu eksekusi sedangkan model *prophet* dengan *dataset* 2 mendapatkan nilai *MAPE* 6.05% dan 0.892 detik waktu eksekusi. Selanjutnya model *Sarimax* dan *prophet* yang menggunakan *dataset* 2, dilakukan prediksi dari tahun 2022 hingga february 2023. Hasil prediksi kedua model dibandingkan dengan data aktual. Hasilnya, *Sarimax* lebih unggul daripada *prophet* dengan nilai *MAPE* 12.05% dan waktu eksekusi 0.014, *prophet* yaitu 14.07% dan 0.100 detik.

Kata kunci: Komparasi, Prediksi, *Sarimax*, *Prophet*, TPK.

Abstract

The Central Java government is prioritizing tourism recovery due to COVID-19. The hotel room occupancy rate (TPK) can be a sign of the success of tourist interest in the area visited and one of the leading indicators in measuring the economy in the tourism sector. Central Java hotel TPK is published by the Central Statistics Agency (BPS) one or two months later. This study compares the *Sarimax* and *Prophet* models for predicting the TPK of Central Java star hotels based on the *MAPE* value and program execution time. The model is made using two categories of datasets that have been checked for stationarity. Dataset 1 is 2005-2022 data (including the covid 19 pandemic situation until the termination of PPKM), and data set 2 is 2005-2019 (without including the covid 19 pandemic situation). The result is that the best *Sarimax* and *Prophet* models are models that use dataset 2 with a *MAPE* value of 10.43% and 47.529 seconds of execution time, while the *Prophet* model with dataset 2 gets a *MAPE* value of 6.05% and 0.892 seconds of execution time. Furthermore, *Sarimax* and *Prophet* models that use dataset 2 are predicted from 2022 to February 2023. The prediction results of both models are compared with the actual data. As a result, *Sarimax* is superior to *Prophet* with a *MAPE* value of 12.05% and an execution time of 0.014; *Prophet* is 14.07% and 0.100 seconds.

Keywords: Comparison, Prediction, *Sarimax*, *Prophet*, TPK.

1. PENDAHULUAN

Pandemi coronavirus disease 2019(Covid-19) telah mewabah di seluruh dunia[1]. Covid-19 mulai muncul pada akhir 2019, mulai merebak dan meledak secara lokal di Tiongkok pada akhir Januari 2020[2][3]. Kemudian meluas ke seluruh dunia dari Februari hingga akhir Mei 2020[4]. Chang Yong Rhee (2020) dari *International Monetary Fund* (IMF) mengatakan bahwa sektor yang mendapatkan tekanan paling kuat akibat penyebaran covid-19 salah satunya adalah sektor pariwisata[2]. Sektor pariwisata merupakan salah satu sumber pendapatan terbesar bagi pembangunan ekonomi negara[5]. Pada tahun 2020, Indonesia mengalami penurunan devisa di sektor pariwisata sebesar 20,7 juta rupiah dan sekitar 409.000 tenaga kerja di sektor tersebut menganggur[6]. Tahun 2021 adalah tahun kebangkitan bagi pariwisata Indonesia[7]. Pemerintah telah berjanji untuk mempromosikan pariwisata sebagai kontributor utama perekonomian negara[8].

Pariwisata di Jawa Tengah paling berdampak akibat covid-19. Tahun 2022, menjadi perhatian khusus bagi pemerintah Jawa Tengah untuk memprioritaskan pemulihan pariwisata[9]. Jawa Tengah tengah memiliki daya tarik wisata sebanyak 834 dengan jumlah wisatawan nusantara yang berkunjung adalah 57.900.863 orang dan tingkat penghunian kamar hotel berbintang mencapai 46.11%[10]. Tahun 2020, jumlah wisatawan nusantara yang datang ke Jawa tengah menurun, yaitu 22.629.085 orang dengan tingkat penghunian kamar hotel berbintang yaitu 28.8%. [11].

Tingkat penghunian kamar (TPK) hotel adalah persentase kamar hotel yang ditempati atau digunakan oleh tamu terhadap total jumlah kamar yang tersedia pada sebuah hotel. Semakin tinggi nilai TPK hotel maka menandakan pariwisata yang semakin baik karena TPK dapat menjadi pertanda keberhasilan sebuah hotel, minat wisatawan terhadap daerah yang dikunjungi[12] dan salah satu *leading indicator* dalam mengukur ekonomi pada sektor pariwisata[13].

Badan Pusat Statistik (BPS) melakukan penjadwalan tinjauan tiap bulannya yang disebut Survei Tingkat Penghunian Kamar Hotel (VHTS) yang terfokus pada hotel berbintang. BPS melakukan pencatatan hotel pada hotel yang dikunjungi oleh pencacah[14]. Hasil pengolahan VHTS dipublikasikan melalui website BPS. Namun publikasi tersebut membutuhkan waktu yang tidak instan. Publikasi biasanya dilakukan satu atau dua bulan berikutnya.

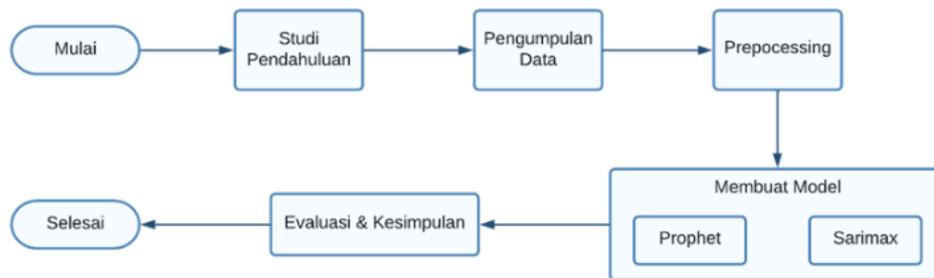
Perhitungan TPK hotel dimasa mendatang (prediksi) dapat dilakukan dengan bantuan teknologi sehingga dapat menghemat waktu dan sebagai alternatif menunggu publikasi pengolahan VHTS. Prediksi ialah proses memperkirakan apa yang akan terjadi di masa depan dengan menelaah informasi dimasa lalu[15]. Sebelumnya telah ada penelitian terkait prediksi TPK hotel yaitu penelitian [13] yang mengungkapkan bahwa metode *time series regression* tidak lebih unggul dibandingkan metode Arimax pada prediksi TPK hotel dan perlunya pengembangan metode prediksi dikarenakan pandemi COVID-19 menyebabkan pola data pada beberapa indikator penting mengalami perubahan yang ekstrim.

Prediksi dapat dilakukan dengan metode Arima, Arimax, *Sarima*, *Sarimax* dan *prophet* berdasarkan pada [16] mengungkapkan bahwa diantara algoritma *Arima*, *Arimax*, *Sarima* dan *Sarimax*. *Sarimax* merupakan model terbaik dengan nilai *MAPE* yaitu 28%. Selanjutnya penelitian [17] pada tahun 2020 oleh Setia Budi melakukan analisis perbandingan antara *Arima* dan *Prophet* dengan studi kasus pendaftaran mahasiswa baru yang menunjukkan bahwa *Prophet* lebih unggul dibandingkan Arima. Pada tahun yang sama, Penelitian [18] yang melakukan peramalan permintaan beban listrik rumah tangga menggunakan *machine learning* yaitu LSTM, Arimax, *Sarimax* dan *Prophet*. Hasil penelitian menyatakan LSTM merupakan model terbaik, disusul oleh *Sarimax*, *Prophet* dan yang terakhir adalah *Arimax* berdasarkan nilai *MAPE*. Namun, Penelitian [19] menyatakan bahwa model *Prophet* lebih baik dari pada *Sarimax* berdasarkan nilai *MAPE* yaitu *MAPE* 1,06%. Penelitian-penelitian tersebut, menggunakan *MAPE* dalam mengevaluasi model yang dibangun. *Mean Absolut Percentage error (MAPE)*

adalah kesalahan persentase rata-rata absolut. Semakin kecil nilai *MAPE* maka pemodelan yang dibuat semakin kecil kesalahannya dalam memprediksi data[20].

Berdasarkan penelitian sebelumnya, untuk memprediksi TPK dikemudian hari, perlu melakukan pengembangan model dan memperhatikan pola data karena data ekstrim akibat pandemi covid-19 sehingga penulis akan melakukan prediksi TPK hotel di Jawa Tengah berdasarkan data TPK hotel dari tahun 2005 hingga tahun 2022. Data TPK akan dibagi menjadi dua kategori yaitu *dataset 1* yang menyertakan kondisi pandemi covid-19 hingga pencabutan peraturan PPKM. *Dataset 2* adalah *dataset* yang menyertakan kondisi pandemi covid-19. Hal ini dilakukan untuk melihat seberapa besar data ekstrim pandemi covid-19 mempengaruhi pemodelan. Selanjutnya, melihat permasalahan pada penelitian [18] dan penelitian [19] maka pemodelan akan lakukan menggunakan *Sarimax* dan *Prophet* dan mencari model terbaik diantara keduanya berdasarkan nilai *MAPE* seperti pada penelitian terdahulu dan menambahkan evaluasi waktu eksekusi sehingga model terbaik berdasarkan nilai terkecil *MAPE* dan waktu tercepat eksekusi model.

2. METODE PENELITIAN



Gambar 1 Metodologi Penelitian

2.1 Studi Pendahuluan

Tahap awal ini dilakukan dilakukan studi pendahuluan terhadap rumusan masalah dengan melakukan studi literatur terhadap berbagai sumber yang valid.

2.2 Pengumpulan Data

Pengumpulan data pada penelitian ini berasal dari Badan Pusat Statistika (BPS) Jawa Tengah. Data penelitian termuat pada dokumen tahunan yaitu buku Statistik Tingkat Penghunian Kamar Hotel Provinsi Jawa Tengah 2005 – 2022 namun untuk data penghunian kamar hotel provinsi jawa tengah tahun 2010 didapatkan dari berita resmi statistik yang mana tingkat Penghunian Hotel Bintang Di Jawa Tengah pada tahun 2010 diunggah ke situs resmi BPS jawa tengah setiap bulannya.

2.3 Preprocessing

Proses ini melakukan memastikan bahwa *dataset* dapat digunakan, akan dilakukan uji stasionaritas dan melakukan Differencing *dataset* bila *dataset* belum stasioner lalu membuat *dataset*, dimana data akan menjadi 2 kategori *dataset 1* yang menyertakan kondisi pandemi covid-19 hingga pencabutan peraturan PPKM. *Dataset 2* adalah *dataset* yang menyertakan kondisi pandemi covid-19. Lalu kedua kategori *dataset* dibagi menjadi data Train dan data Test. data Train sebesar 80% dan data Test sebesar 20% dari keseluruhan *dataset*. Hal ini dilakukan untuk melihat seberapa besar data ekstrim pandemi covid-19 mempengaruhi pemodelan.

2.4 Membuat Model

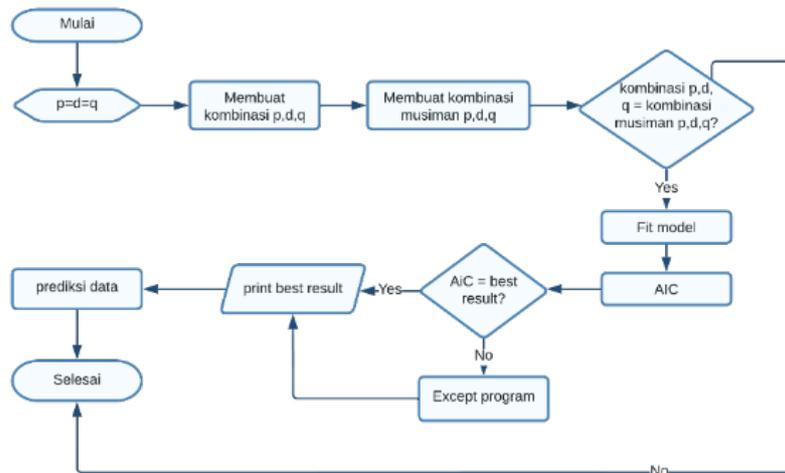
Tahap ini dilakukan perancangan arsitektur model *Sarimax* dan *Prophet* dengan melakukan konfigurasi terhadap parameter-parameter yang akan digunakan. Pada model *Sarimax* parameter yang digunakan adalah parameter *seasonal* musiman yang terdiri dari kombinasi variabel p (orde AR), d (*differencing*), q (MA non musiman). Parameter *seasonal* dengan variabel p,d,q akan melakukan *fitting* dan dihitung nilai kriteria informasi Akaike (AIC) yang merupakan matriks perbandingan kecocokan model[21].

$$AIC = e^{\frac{2k}{n} \sum_{i=1}^n \hat{u}_i^2} \text{ atau } \ln AIC = \frac{2k}{n} + \frac{\sum_{i=1}^n \hat{u}_i^2}{n} \quad (1)$$

Keterangan:

- k = jumlah parameter yang diestimasi dalam model regresi
- n = jumlah observasi
- \hat{u} = sisa (residual)

Jika AIC dengan model terbaik telah didapatkan maka nilai AIC untuk parameter *seasonal* akan digunakan untuk memprediksi data.



Gambar 2 Diagram Alir Model Sarimax

Rumus model *Sarimax*[32] adalah:

$$Y_t = \beta_0 + \beta_1 V_{1,t} + \dots + \beta_i V_{i,t} + \omega_r \quad (2)$$

$$Y_t = \beta_0 + \beta_1 V_{1,t} + \dots + \beta_i V_{i,t} + \frac{\theta_q(B)\theta_Q(B^S)e_t}{\Phi_P(B^S)\phi_P(B)(1-B)^d(1-B^S)^D} \quad (3)$$

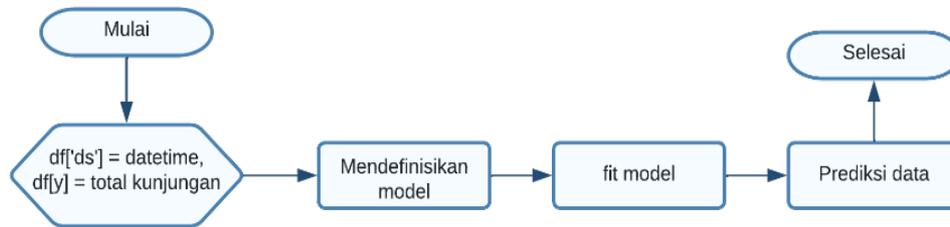
$$\Phi_P(B^S)\phi_P(B)(1-B)^d(1-B^S)^D Y_t = \theta_q(B)\theta_Q(B^S)e_t + \beta_0 + \beta_1 V_{1,t} + \dots + \beta_i V_{i,t} \quad (4)$$

Keterangan:

- Y_t = Data pada waktu ke- t
- $V_{i,t}$ = Variabel ke- i pada saat t dengan $i = 1,2,3$ dan seterusnya
- β_0, β_k = Koesifien regresi
- ω_r = Residual pada waktu ke- t dari proses runtun waktu
- p,d,q = Orde AR, *differencing*, MA non musiman
- P,D,Q = Orde AR, *differencing*, MA musiman
- $\Phi_P(B^S)$ = Koefisien komponen AR musiman S dengan orde P
- $\Phi_Q(B^S)$ = Koefisien komponen MA musiman S dengan orde Q

e_t = Nilai *error* pada waktu ke- t
 B = Operator *backward shift*

Model *Prophet* menggunakan parameter variabel ds untuk *datetime* dan variabel df ['y'] untuk presentase tingkat penghunian kamar hotel. Selanjutnya, model dilakukan pendefinisian dan *fitting* untuk memprediksi data.



Gambar 3 Diagram Alir Model Prophet

Rumus *Prophet* adalah sebagai berikut[17].

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t) \quad (5)$$

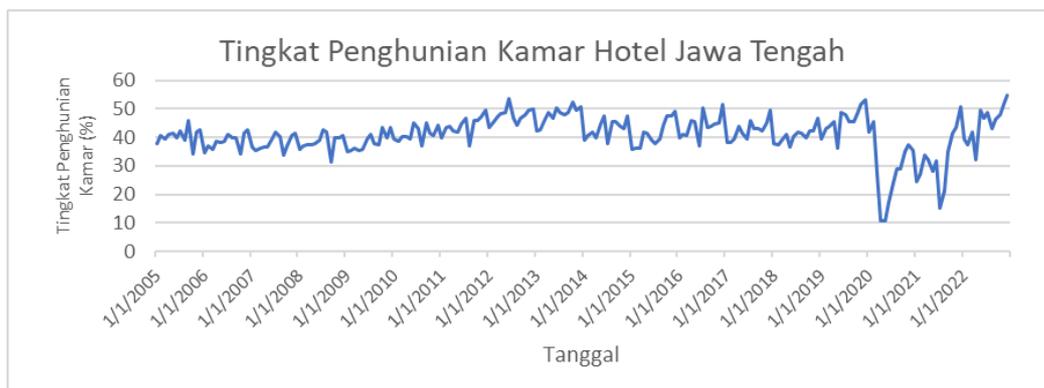
Keterangan:

- $g(t)$ adalah adalah kurva pertumbuhan linier atau logistic untuk memodelkan perubahan non-periodik dalam deret waktu
- $s(t)$ adalah variasi periodik (contohnya musiman mingguan/tahunan)
- $h(t)$ adalah efek liburan (waktu libur yang diatur oleh pengguna)
- $\varepsilon(t)$ adalah istilah kesalahan yang memperhitungkan setiap perubahan tidak teratur yang biasanya tidak dapat diakomodasi oleh mode

2.5 Evaluasi dan Kesimpulan

Tahap evaluasi menjelaskan bahwa proses model yang telah dilakukan dari awal sampai akhir akan di lakukan analisis dengan melihat data hasil prediksi, nilai error dari model menggunakan *MAPE* dan kecepatan waktu eksekusi kedua model, kemudian dilakukan penyimpulan berdasarkan hasil yang telah di analisis tersebut.

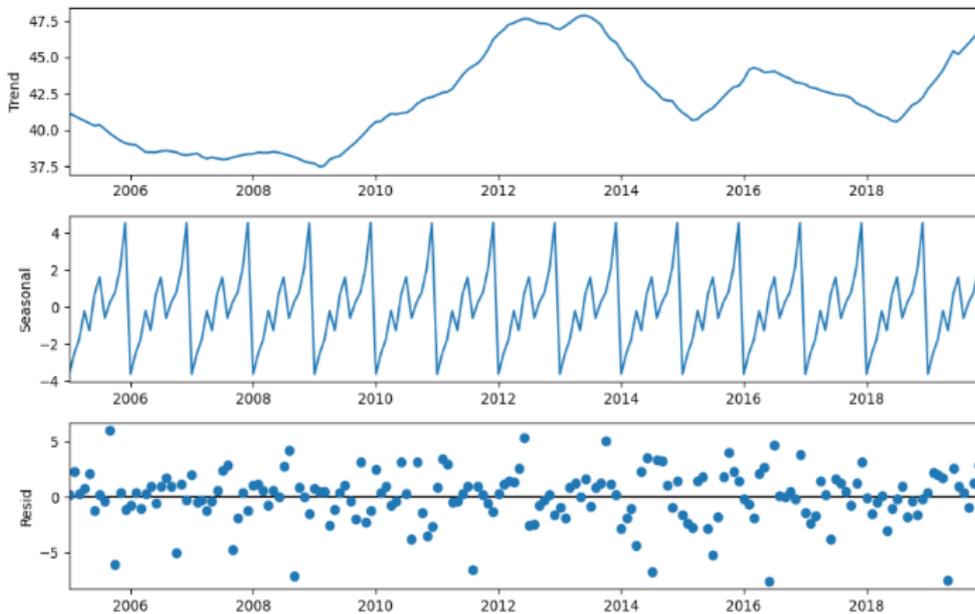
3. HASIL DAN PEMBAHASAN



Gambar 4 Penghunian Kamar Hotel Berbintang Jawa Tengah

Tingkat Penghunian Kamar Hotel Jawa Tengah dinyatakan dalam persen. Pada tahun 2005 hingga tahun 2019, presentase penghunian kamar hotel berbintang di Jawa Tengah berada di atas 30% namun pada bulan mei 2020 terjadi penurunan yang sangat drastis, hal ini disebabkan oleh situasi pandemi covid-19. Pandemi covid-19 terjadi di Indonesia pada Januari

2020, hal tersebut menyebabkan tingkat penghunian kamar hotel berbintang di Jawa tengah menjadi tak menentu yakni turun dan naik secara drastis diatas rata rata. Rata rata tingkat penghunian kamar hotel berbintang di Jawa tengah adalah 40.95% dengan standar deviasi yaitu 6.75%.



Gambar 5 Tren dan Seasonal *Dataset*

Gambar 2 menunjukkan karakteristik *dataset* memiliki pertumbuhan trend yang kecil sebelum pandemi covid-19. Pemberlakuan pembatasan kegiatan masyarakat (PPKM) dimana masyarakat sebisa mungkin melakukan kegiatan di rumah termasuk dalam kegiatan berkerja yang dilakukan secara daring membuat sebagian besar bisnis perhotelan ditutup dan saat pencabutan peraturan PPKM dilarang, masyarakat mulai mendatangi pariwisata dan menginap di hotel. Hal ini membuat trend menurun pada awal tahun 2020 dan melonjak naik di pertengahan 2020. Namun, ternyata terdapat pola musiman yang berulang dimana setiap akhir tahun, masyarakat akan berlibur dan mengunjungi hotel berbintang. *Dataset* juga memiliki residual di berbagai titik sehingga perlu mencari tahu apakah *dataset* berdistribusi normal namun karena *dataset* memiliki panjang lebih dari 30 maka distribusi dapat dianggap normal[22]. Selanjutnya, *dataset* dibagi menjadi 2 kategori yaitu *dataset* dengan tahun dimana covid-19 masuk ke jawa tengah sampai dengan pemberhentian peraturan PPKM yaitu januari 2020 hingga desember 2022 dan *dataset* sebelum 29 situasi pandemi covid-19 hal ini dilakukan untuk menganalisa model yang akan dibuat, seberapa besar pandemi covid-19 akan sangat mempengaruhi prediksi TPK Jawa Tengah.

Tabel 1 Kategori *Dataset*

Kondisi	Nama data	Periode	Jumlah
Pandemi Covid 19	<i>Dataset 1</i>	2005 - 2022	215
Sebelum pandemi Covid 19	<i>Dataset 2</i>	2005 - 2019	180

Stasioneritas di cek menggunakan test Augmented Dickey-Fuller (ADF).

Tabel 2 Test *ADF*

Nama	Nilai ADF	p-value	Critical Value			Kesimpulan
			1%	5%	10%	
<i>Dataset 1</i>	-1.607	0.479	-3.470	-2.879	-2.576	Tidak stasioner
<i>Dataset 2</i>	-2.490	0.115	-3.462	-2.875	-2.574	Tidak stasioner

Pada *dataset 1* dan *dataset 2* nilai $ADF > Critical Value$ dan nilai $p-value > 0.05$ artinya kedua *dataset* tidak stasioner. *dataset* dapat menjadi stasioner dengan melakukan Differencing time series. Differencing time series dilakukan dengan menstabilkan variance dengan cara log transformation pada *dataset*. *Dataset* yang telah diubah menjadi log dikurangi dengan rolling rata rata *dataset* log. Efek dari hal tersebut akan ada data Nan atau data yang bernilai kosong. Data kosong tersebut dapat dilakukan `dropna()` selanjutnya dites Kembali menggunakan tes ADF.

Tabel 3 Test ADF Setelah Differencing

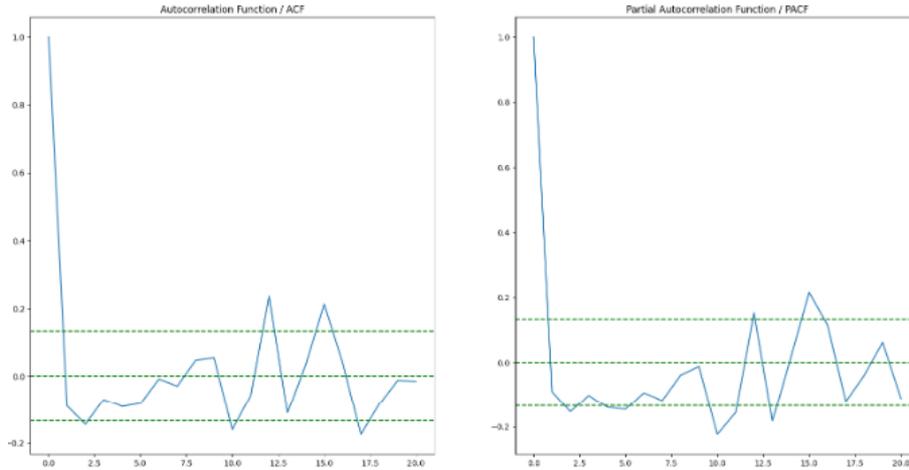
Nama	Nilai ADF	p-value	Critical Value			Kesimpulan
			1%	5%	10%	
<i>Dataset 1</i>	-2.463	0.124	-3.470	-2.879	-2.576	Tidak stasioner
<i>Dataset 2</i>	-3.290	0.015	-3.462	-2.875	-2.574	Tidak stasioner pada nilai kepercayaan 99% atau Critical Value 1%

Terdapat perubahan pada nilai ADF dan $p-value$ setelah dilakukan *Differencing time series* sebelum dilakukan *Differencing* nilai ADF adalah -1.607 dan nilai $p-value$ 0.479 menjadi nilai ADF -2.463 dan $p-value$ 0.124 pada *dataset 1* sedangkan pada *dataset 2* nilai ADF dari -2.490 menjadi -3.290. nilai ADF lebih besar dari pada *Critical Value* 5% dan 10% sedangkan nilai $p-value$ dari 0.115 menjadi 0.015 yang mana $p-value < 0.05$ telah sesuai. Selanjutnya *Differencing* dapat dilakukan dengan operator *backshift*. Langkah awal *Differencing* dengan operator *backshift* adalah menerapkan `shift()` kepada transformasi log. Data log dikurangi data log yang telah dilakukan *shift*. Selanjutnya membuang data yang bernilai kosong dengan `dropna()`. Cek Kembali dengan tes ADF

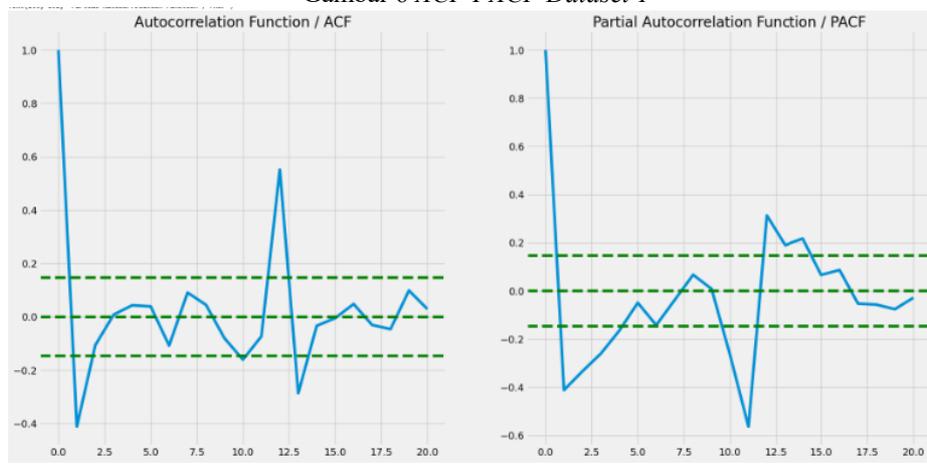
Tabel 4 Test ADF Setelah Differencing Menggunakan Backshift

Nama	Nilai ADF	p-value	Critical Value			Kesimpulan
			1%	5%	10%	
<i>Dataset 1</i>	-3.078	0.02	-3.470	-2.879	-2.576	Tidak stasioner pada nilai kepercayaan 99% atau Critical Value 1%
<i>Dataset 2</i>	-3.343	0.01	-3.462	-2.875	-2.574	Tidak stasioner pada nilai kepercayaan 99% atau Critical Value 1%

Berdasarkan Tabel 4 yaitu *dataset 1* dan *dataset 2* pada nilai $p-value$ telah berhasil menunjukkan bahwa $p-value < 0.05$ sedangkan nilai ADF gagal pada kepercayaan 1% tidak stasioner namun karena nilai ADF lebih kecil daripada *Critical Value* 5% dan 10% artinya *dataset* cukup stasioner. mengecek stasioner juga dapat dilakukan menggunakan *Autocorrelation Function (ACF)* dan *Partial autocorrelation function (PACF)*. Fungsi *ACF* dan *PACF* membantu dalam melakukan plot korelasi antara *time series* dan *lag series* dengan 95% tingkat kepercayaan,



Gambar 6 ACF PACF Dataset 1



Gambar 7 ACF PACF Dataset 2

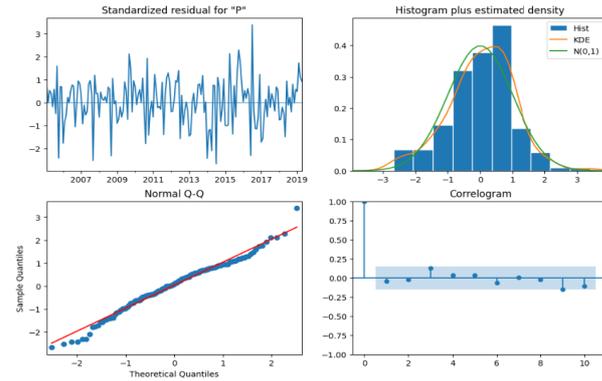
Grafik diatas menunjukkan bahwa garis berwarna biru/time series berpotongan dengan garis hijau putus putus/shift time series dengan keyakinan 95% (dalam hal ini dengan lag=20 baik dalam batas atas maupun batas atas) sehingga grafik time series stasioner dengan keyakinan 95% dan mengalami efek auto correlation. *Dataset* telah stasioner makan selanjutnya setiap kategori *dataset* dipecah menjadi data latih dan data tes dengan perbandingan 80%:20% dari panjang data.

Tabel 5 Pembagian *Dataset*

Kondisi	Kategori data	Nama data	Tanggal	Jumlah data
Covid 19	Dataset 1	Train	01-01-2005 sampai 01-04-2019	184
		Test	01-05-2019 sampai 01-12-2022	43
Sebelum Covid 19	Dataset 2	Train	01-01-2005 sampai 01-12-2016	132
		Test	01-01-2017 sampai 01-12-2019	36

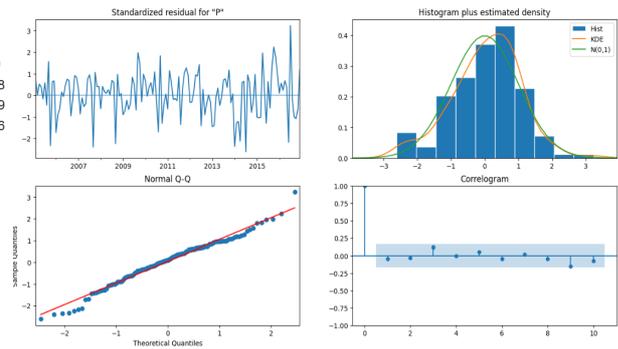
Model sarimax dengan kombinasi terbaik pada *dataset* 1 adalah (0,1,1) dan seasonal (1,0,0,12) dengan nilai AIC yaitu -591.6979246702509 sedangkan model dengan kombinasi terbaik pada *dataset* 2 adalah (0, 1, 1) dan seasonal (1, 0, 0, 12) dengan nilai AIC -720.4298713582441. *Testing* model menggunakan variabel *steps*=200 dan *alpha*=0.05

SARIMAX Results
 Dep. Variable: Presentase No. Observations: 172
 Model: SARIMAX(0, 1, 1)x(1, 0, [], 12) Log Likelihood 363.215
 Date: Tue, 04 Apr 2023 AIC -720.430
 Time: 23:29:15 BIC -711.005
 Sample: 01-01-2005 HQIC -716.606
 -04-01-2019
 Covariance Type: opg
 coef std err z P>|z| [0.025 0.975]
 ma.L1 -0.7040 0.055 -12.906 0.000 -0.811 -0.597
 ar.S.L12 0.6309 0.054 11.702 0.000 0.525 0.737
 sigma2 0.0008 7.97e-05 10.096 0.000 0.001 0.001
 Ljung-Box (L1) (Q): 0.38 Jarque-Bera (JB): 3.81
 Prob(Q): 0.54 Prob(JB): 0.15
 Heteroskedasticity (H): 1.35 Skew: -0.25
 Prob(H) (two-sided): 0.26 Kurtosis: 3.54



Gambar 8 Model Sarimax Dataset 1

SARIMAX Results
 Dep. Variable: Presentase No. Observations: 144
 Model: SARIMAX(0, 1, 1)x(1, 0, [], 12) Log Likelihood 298.849
 Date: Wed, 05 Apr 2023 AIC -591.698
 Time: 01:51:32 BIC -582.809
 Sample: 01-01-2005 HQIC -588.086
 -12-01-2016
 Covariance Type: opg
 coef std err z P>|z| [0.025 0.975]
 ma.L1 -0.7376 0.056 -13.144 0.000 -0.848 -0.628
 ar.S.L12 0.6136 0.063 9.691 0.000 0.489 0.738
 sigma2 0.0009 9.29e-05 9.216 0.000 0.001 0.001
 Ljung-Box (L1) (Q): 0.29 Jarque-Bera (JB): 3.12
 Prob(Q): 0.59 Prob(JB): 0.21
 Heteroskedasticity (H): 1.94 Skew: -0.27
 Prob(H) (two-sided): 0.02 Kurtosis: 3.48



Gambar 9 Model Sarimax Dataset 2

Fitur model prophet yang digunakan untuk kedua dataset yaitu menggunakan $change_point_range = 0.8$, $n_change_points = 5$, $change_point_prior_scale = 0.05$, $yearly_seasonality = 'auto'$.

Evaluasi model terdiri atas evaluasi waktu eksekusi model dan nilai MAPE. Evaluasi waktu eksekusi dihitung dari tahapan pembuatan model hingga model dapat melakukan test dataset. Berikut adalah tabel evaluasi model sarimax.

Tabel 6 Evaluasi Model Sarimax

Tahap	Dataset 1	Dataset 2
Mencari nilai kombinasi	49.188 detik	47.290 detik
Fitting data	0.291 detik	0.207 detik
Validasi dan Train dataset	0.044 detik	0.032 detik
Jumlah waktu	49.523 detik	47.529 detik
MAPE	48.6%	10.43%

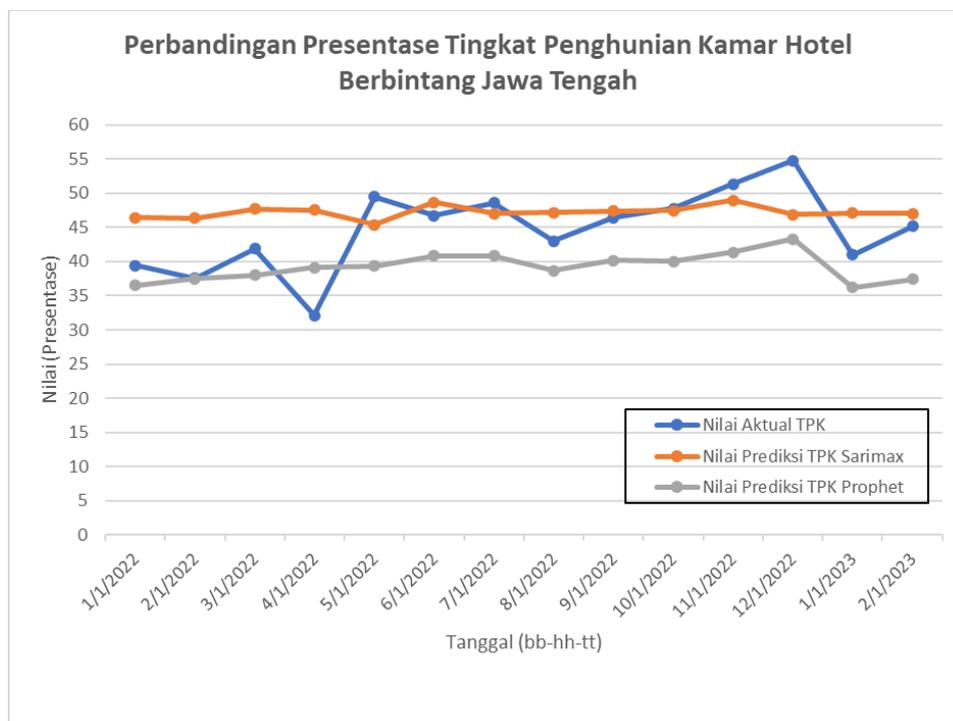
Berdasarkan tabel diatas, dapat disimpulkan bahwa model dengan dataset 2 lebih baik dalam segi waktu eksekusi dan nilai MAPE. Dataset yang semakin panjang membutuhkan waktu eksekusi yang lebih lama.

Tabel 7 Evaluasi Model Prophet

Tahap	Dataset 1	Dataset 2
Fitting data	14.10 detik	0.441detik
Validasi dan Train dataset	0.077 detik	0.121detik
Penyesuaian fitting, validasi dan training dataset	0.260 detik	0.333 detik
Jumlah waktu	14.437 detik	0.892 detik
MAPE	12.68%	6.05%

Sama seperti model sarimax, dapat disimpulkan bahwa model dengan dataset 2 lebih baik dalam segi waktu eksekusi dan nilai MAPE. Dataset yang semakin panjang membutuhkan

waktu eksekusi yang lebih lama. Tahapan evaluasi berikutnya adalah melakukan pengujian model terhadap hasil peramalan. Model yang diujikan adalah model dengan nilai *MAPE* terbaik berdasarkan kategori *dataset*. Model *sarimax* terbaik yaitu model *sarimax* dengan pelatihan menggunakan *dataset 1* bergitupula dengan model *prophet* yang mana model *prophet* terbaik yaitu model *prophet* dengan pelatihan menggunakan *dataset 1*. Prediksi dilakukan dari tahun 2020 hingga februari 2023. Hasil prediksi nilai presentase TPK dibandingkan dengan nilai aktual presentase TPK dan dihitung nilai *MAPE* serta dihitung waktu eksekusi prediksinya. Waktu eksekusi model *sarimax* dengan *dataset 1* adalah 0.014 detik sedangkan waktu eksekusi model *prophet* dengan *dataset 1* adalah 0.100 detik. Model *sarimax* lebih cepat dalam mengeksekusi peramalan untuk masa depan. Selanjutnya, berdasarkan nilai *MAPE*, *Sarimax* dan *prophet* dapat memprediksi nilai presentase TPK dengan baik namun model *sarimax* lebih unggul dibandingkan dengan model *prophet*. *MAPE* model *sarimax* yaitu 12.05% sedangkan *MAPE* model *prophet* yaitu 14.07%. Berikut merupakan visualisasi perbandingan hasil prediksi nilai presentase TPK dengan nilai atual presentase TPK.



Gambar 10 Visualisasi Perbandingan Nilai *MAPE* *Sarimax* Dan *Prophet*

4. KESIMPULAN DAN SARAN

Kesimpulan pada penelitian komparasi *Sarimax* dan *prophet* untuk prediksi presentase tingkat penghunian kamar (TPK) hotel berbintang jawa tengah adalah sebagai berikut:

1. Komparasi *Sarimax* dan *Prophet* untuk mencari model terbaik dihitung berdasarkan nilai *MAPE* dan waktu eksekusi program. Model dibuat menggunakan 2 kategori *dataset* yang telah dicek stasionaritasnya. *Dataset 1* yaitu data 2005 hingga 2022 (menyertakan situasi pandemi covid 19 hingga pemberhentian PPKM). *Dataset 2* yaitu data 2005 hingga 2019 (tanpa menyertakan situasi pandemi covid 19). Hasilnya model *Sarimax* dan *prophet* terbaik adalah model yang menggunakan *dataset 2* dengan nilai *MAPE* 10.43% dan 47,529 detik waktu eksekusi sedangkan model *prophet* yang menggunakan *dataset 2* mendapatkan nilai *MAPE* 6.05% dan 0.892 detik waktu eksekusi

2. Berdasarkan model *Sarimax* dan *prophet* yang menggunakan *dataset 2*, dilakukan prediksi dari tahun 2022 hingga februari 2023. Hasil prediksi kedua model dibandingkan dengan data aktual dan di hitung nilai *MAPE* berserta waktu eksekusi program. Hasilnya, model *Sarimax* lebih unggul daripada model *Prophet* dengan nilai *MAPE* 12.05% dengan waktu eksekusi 0.014 sedangkan model *Prophet* mendapatkan nilai *MAPE* 14.07% dengan waktu eksekusi program 0.100 detik.

Saran pada penelitian komparasi *Sarimax* dan *prophet* untuk prediksi presentase tingkat penghunian kamar (TPK) hotel berbintang jawa tengah sebagai pertimbangan untuk penelitian berikutnya adalah menambahkan variabel exogenous dan membagi data dalam kategori yang lebih banyak berdasarkan panjang data.

DAFTAR PUSTAKA

- [1] D. R. A. U. Khasanah, H. Pramudibyanto, and B. Widuroyekti, "Pendidikan Dalam Masa Pandemi Covid-19," *J. Sinestesia*, vol. 10, no. 1, pp. 41–48, 2020, [Online]. Available: <https://sinestesia.pustaka.my.id/journal/article/view/44>.
- [2] K. Perencanaan *et al.*, "Covid-19, New Normal, dan Perencanaan Pembangunan di Indonesia," *J. Perenc. Pembang. Indones. J. Dev. Plan.*, vol. 4, no. 2, pp. 240–252, 2020, doi: 10.36574/jpp.v4i2.118.
- [3] N. A. Suriadi, V. Evi, Y. Deril, and W. S. Sintia, "Stategi Penataan Infrastruktur Kota Majene Yang Adaptif Terhadap Pandemi," vol. 3, no. 1, pp. 24–30, 2021.
- [4] A. A. Hidayanti, B. D. Prathama, and S. Wardah, "Analisis Korelasi Pearson Dalam Menentukan Hubungan Kualitas Produk, Pelayanan, Lokasi Dan Kepuasan Terhadap Loyalitas Pada Pelanggan Rumah Nutrisi Herbalife Mataram," *J. Innov. Knowl.*, vol. 1, no. 2, pp. 185–198, 2021.
- [5] Y. benony Walakula, "Analisis Eksistensi Pariwisata Indonesia di Tengah Situasi Pandemi Corona Virus Disease (Covid19)," *NOUMENA Ilmu Sos. Keagamaan*, vol. 1, no. 1, pp. 47–52, 2020, [Online]. Available: <https://e-journal.iaknambon.ac.id/index.php/N/article/view/165>.
- [6] Kemenparekraf, "Tren Pariwisata Indonesia di Tengah Pandemi," *Kemenparekraf.Go.Id.* 2021, [Online]. Available: <https://kemenparekraf.go.id/ragam-pariwisata/Tren-Pariwisata-Indonesia-di-Tengah-Pandemi>.
- [7] M. RI, "Gus Jazil_ Jadikan Tahun 2021 Sebagai Tahun Kebangkitan Indonesia Lepas Dari Pandemi." 2021, [Online]. Available: <https://www.mpr.go.id/berita/gus-jazil-jadikan-tahun-2021-sebagai-tahun-kebangkitan-indonesia-lepas-dari-pandemi>.
- [8] H. Kemensetneg, "Pariwisata, Lokomotif Baru Penggerak Ekonomi Indonesia | Sekretariat Negara." 2019, [Online]. Available: https://setneg.go.id/baca/index/pariwisata_lokomotif_baru_penggerak_ekonomi_indonesia.
- [9] H. P. J. Tengah, "Jateng Prioritaskan Pemulihan UKM dan Pariwisata Pascawabah Covid-19," 2020. https://humas.jatengprov.go.id/detail_berita_gubernur?id=4174 (accessed May 16, 2023).
- [10] *Buku Pariwisata Jawa Tengah dalam Angka*. Dinas Kepemudaan, Olahraga dan Pariwisata Provinsi Jawa Tengah, 2019.
- [11] O. dan P. P. J. T. Dinas Kepemudaan, *Buku Pariwisata Jawa Tengah dalam Angka 2020*. Dinas Kepemudaan, Olahraga dan Pariwisata Provinsi Jawa Tengah, 2021.
- [12] "BADAN PUSAT STATISTIK KABUPATEN OKU SELATAN." 2023.
- [13] I. Ayuningtyas and I. Wirawati, "Nowcasting Tingkat Penghunian Kamar Hotel Menggunakan Google Trends," *Semin. Nas. Off. Stat.*, vol. 2020, no. 1, pp. 338–343, 2021, doi: 10.34123/semnasoffstat.v2020i1.636.
- [14] BPS, "BPS Kabupaten Brebes," *BPS Kabupaten Brebes*. 2022, [Online]. Available:

- <https://brebeskab.bps.go.id/statictable/2022/03/23/2127/produksi-tanaman-sayuran-bawang-merah-menurut-kecamatan-di-kabupaten-brebes-kuintal-2020-dan-2021.html>.
- [15] A. D. Yulianto, "Implementasi Deep Learning Dengan Menggunakan Pemodelan Gated Recurrent Unit (Gru) Untuk Prediksi Harga Saham Di Indeks ...," Universitas Muhammadiyah Semarang, 2021.
- [16] N. Permatasari, "Penggunaan Indeks Google Trend Dalam Peramalan Jumlah Pengunjung Taman Rekreasi Selecta Tahun 2020," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 1019–1024, Nov. 2021, doi: 10.34123/semnasoffstat.v2021i1.993.
- [17] C. Chandra and S. Budi, "Analisis Komparatif ARIMA dan Prophet dengan Studi Kasus Dataset Pendaftaran Mahasiswa Baru," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 278–287, 2020, doi: 10.28932/jutisi.v6i2.2676.
- [18] A. K. Panigrahi, "Forecasting Residential Electricity Load Demand using Machine Learning," 2020.
- [19] S. P. Sharma, J. R., and K. Deepa, "Forecasting India S&P BSE SENSEX and USA S&P-500 Benchmark Indices Using SARIMAX and Facebook Prophet Library," in *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2022, pp. 1523–1530, doi: 10.1109/ICICCS53718.2022.9788127.
- [20] I. R. Julianto, Indwiarti, and A. A. Rohmawati, "Prediksi Jumlah Kunjungan Wisatawan Di Jawa Barat Dengan Model ARIMAX Dan SARIMAX Menggunakan Data Google Trends," *e-Proceeding Eng.*, vol. 8, no. 4, pp. 4229–4241, 2021.
- [21] M. Y. Darsyah and M. N. Ramadhan, "Pemodelan Jumlah Kasus Penyakit Kusta Di Provinsi Sulawesi Tenggara Menggunakan Metode Regresi Poisson Inverse Gaussian," *J. Litbang ...*, pp. 135–140, 2022, doi: 10.51402/jle.v3i1.9.
- [22] Faradiba, "Penggunaan Aplikasi Spss Untuk Analisis Statistika Program," *SEJ (School Educ. J.*, vol. 10, no. 1, pp. 65–73, 2020, [Online]. Available: <https://jurnal.unimed.ac.id/2012/index.php/school/article/view/18067>.