

Penerapan Metode Filter Kalman Dalam Perbaikan Hasil Prediksi Cuaca Dengan Metode ARIMA

Tomy Kurniawan, Lukman Hanafi, dan Erna Apriliani
 Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
 Institut Teknologi Sepuluh Nopember (ITS)
 Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia
e-mail: april@matematika.its.ac.id

Abstrak—*Time Series* merupakan salah satu metode analisis data. Salah satu metodenya adalah *ARIMA Box-jenkins* yang nantinya digunakan untuk meramalkan data selanjutnya. Penelitian ini didasarkan pada pengamatan dan sesuai hasil model peramalan analisis *time series* dari dua parameter meteorologi yaitu suhu udara dan kecepatan angin. Estimasi koefisien pada persamaan model akan dilakukan dengan menggunakan algoritma Filter Kalman. Setelah adanya penggunaan algoritma Filter Kalman akan dilakukan suatu pendekatan yang didasarkan pada koreksi linear dari bias prakiraan dalam penggunaan Filter Kalman. Selanjutnya akan lebih difokuskan pada studi parameter meteorologi satu waktu dimana diberikan m_i sebagai *direct output* dari model pada waktu t_i yang mengacu pada salah satu parameter (suhu udara atau kecepatan angin) serta y_i^0 sebagai bias dari prakiraan. Estimasi ini memungkinkan dalam bentuk linier pada polinomial dengan $n = 2$. Simulasi ini menggunakan *software* minitab16, matlab R2010a, dan *microsoft excel*. Hasil simulasi berupa grafik hasil estimasi *state* dalam real, ARIMA dengan minitab, dan Filter Kalman.

Kata Kunci—ARIMA Box-Jenkins, Filter Kalman

I. PENDAHULUAN

INDONESIA sebagai Negara kepulauan memiliki jumlah penduduk 242.968.342 jiwa di tahun 2012 menduduki peringkat ke-empat Negara dengan jumlah penduduk terbanyak di dunia [1]. Dan Surabaya, sebagai salah satu kota di Indonesia yang merupakan ibukota Jawa Timur memiliki jumlah penduduk hampir tiga juta jiwa di tahun 2012 atau menempati posisi tertinggi ke-dua jumlah penduduknya setelah kota Jakarta [2]. Seiring dengan jumlah penduduk yang banyak, maka aktivitas oleh warga Surabaya juga akan semakin padat dan bervariasi. Berbagai macam aktivitas di kehidupan sehari-hari ini baik secara langsung maupun tidak langsung nantinya akan sangat terpengaruhi oleh kondisi

cuaca yang terjadi pada hari tersebut. Misalnya pada penerbangan di Bandara Internasional Juanda, Kabupaten Sidoarjo, Jawa Timur, sempat dihentikan sementara waktu akibat cuaca buruk yang terjadi di kawasan Surabaya Selatan [3]. Ataupun dapat mengambil contoh sederhana di lingkungan sekitar seperti pada sektor pertanian, industri kerupuk, atau berbagai sektor lainnya yang jalannya berbagai aktivitas tersebut akan dipengaruhi oleh kondisi cuaca yang ada.

Pada umumnya cuaca terjadi karena suhu dan kelembaban yang berbeda antara satu tempat dengan tempat lainnya. Perbedaan ini bisa terjadi karena sudut pemanasan matahari yang berbeda dari satu tempat ke tempat lainnya karena perbedaan lintang bumi [4]. Dengan kondisi cuaca yang cenderung berbeda-beda antar tempat dan waktu pada suatu daerah, sehingga perlu adanya informasi tentang prakiraan cuaca agar mendukung segala aktivitas yang akan dikerjakan. Telah muncul berbagai pendekatan-pendekatan dalam membuat prakiraan cuaca yang dikembangkan oleh Negara maju yakni suatu metode dengan menginterpretasikan data hasil pengamatan dan data model prakiraan cuaca atau *Numerical Weather Prediction* (NWP). Dalam penggunaan metode ini masih menunjukkan kesalahan yang sistematis karena setiap prakirawan akan menghasilkan interpretasi yang berbeda-beda. Dengan kata lain, metode ini masih bersifat subjektif [5]. Untuk mengurangi pengaruh tersebut, maka digunakan suatu pendekatan yang menggunakan Filter Kalman. Namun dalam hal ini, perbaikan bias prediksi diimplementasikan pada hasil peramalan dari ARIMA (*Autoregressive Integrated Moving Average*) *Box-Jenkins* dikarenakan telah adanya data pengukuran dari BMKG stasiun Juanda Surabaya yang juga merupakan fungsi waktu.

Pada tugas akhir ini, data hasil pengamatan oleh Badan Meteorologi Klimatologi dan Geofisika (BMKG) Stasiun Meteorologi Juanda Surabaya pada suhu udara dan kecepatan angin nantinya akan digambarkan pola yang ada dengan menggunakan Analisis Deret Waktu (*Time Series*). *Time series* adalah serangkaian pengamatan terhadap suatu

variabel yang diambil dari waktu ke waktu dan dicatat secara berurutan menurut urutan waktu kejadiannya dengan interval waktu yang tetap [6]. Setiap pengamatan dinyatakan sebagai variabel random Z_t yang diperoleh berdasarkan indeks waktu tertentu (t_i) dengan $i = 1, 2, \dots, n$ sehingga penulisan data *time series* adalah $Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}$. Model ARIMA yang didapat, akan digunakan untuk melakukan prediksi dan penggunaan Filter Kalman yang terdiri dari suatu himpunan atas persamaan matematika akan memberikan sebuah solusi komputasi yang efisien dari metode kuadrat terkecil. Sehingga nantinya metode Filter Kalman ini dapat diterapkan dalam perbaikan prediksi cuaca.

II. METODOLOGI PENELITIAN

A. Analisis Model Data dan Pembahasan

Pada tahapan ini akan dianalisis data yang didapat dari BMKG stasiun Juanda-Surabaya. Data yang didapat merupakan hasil pengamatan dari pihak terkait dengan menggunakan alat ukur suhu udara. Dan dalam hal ini data yang dianalisis adalah data suhu udara pada bulan Januari, Februari, Maret, dan April. Data yang ada akan diolah lebih lanjut dengan menggunakan *software* Minitab16. Dengan menggunakan metode analisis *time series* maka akan dicoba untuk mencari model data yang sesuai. Akan tetapi, kemungkinan model yang ada akan lebih dari satu tergantung dari hasil plot data yang ada. Berbagai kemungkinan model yang ada tersebut akan diuji kelayakannya sesuai dengan aturan-aturan yang ada pada metode analisis *time series* ini. Lalu akan diramalkan sebanyak 30 hari berikutnya. Dari data yang ada akan diuji kestasioneritasnya baik dalam *varians* dengan menggunakan plot *box-cox transformation* maupun dalam *mean* dengan menggunakan plot *time series* atau juga dapat dengan melihat nilai-nilai *autokorelasi*-nya akan turun secara cepat menuju nol. Setelah didapat data yang stasioner maka selanjutnya adalah plot *Autocorrelation Function* (ACF) dan plot *Partial Autocorrelation Function* (PACF) untuk mengetahui model yang memungkinkan. Ada berbagai kemungkinan model sementara yang terbentuk dengan melihat *lag* yang keluar dari garis merah putus-putus antara lain seperti pada tabel berikut ini

Tabel 1.

Model Sementara Berdasarkan Plot ACF dan PACF pada Data Suhu Udara

model (ARIMA)	<i>white noise</i>	<i>significant</i>	keterangan
3,1,1	x	x	-
3,1,0	x	x	-
0,1,1	x	v	-
1,1,1	v	v	tanpa konstanta

2,1,1	x	x	-
2,1,0	x	v	-
1,1,0	x	v	-

Dari tabel 1 di atas tampak bahwa model ARIMA (1,1,1) telah memenuhi kelayakan model pada *time series* yaitu data yang *white noise* dan *significant*. Berdasarkan Model ARIMA (p, d, q) yang dikenalkan oleh Box dan Jenkins dengan orde p sebagai operator dari AR (*Autoregressive*), orde d merupakan *differencing*, dan orde q sebagai operator dari MA (*Moving Average*):[6]

$$\varphi_p(B)(1 - B)^d Z_t = \theta_0 + \theta_q(B)a_t \tag{1}$$

sehingga untuk ARIMA (1,1,1) adalah

$$\varphi_1(B)(1 - B)^1 Y_t = \theta_1(B)a_t \tag{2}$$

dengan $\varphi_1 = 0,5977$; $\theta_1 = 0,9806$, sehingga

$$Y_t = 1,5977 Y_{t-1} - 0,5977 Y_{t-2} + a_t - 0,9806 a_{t-1} \tag{3}$$

Hal yang sama dilakukan pada data kecepatan angin, dan berbagai kemungkinan model sementara yang terbentuk dengan melihat *lag* yang keluar dari garis merah putus-putus yang ditunjukkan pada plot ACF dan PACF terhadap data kecepatan angin yang telah stasioner antara lain seperti pada tabel berikut ini dengan terdapat empat kemungkinan model yaitu ARIMA (2,1,1), ARIMA (1,1,0), ARIMA (1,1,0), dan ARIMA (0,1,1).

Tabel 2.

Model Sementara Berdasarkan Plot ACF dan PACF pada Data Kecepatan Angin

model (ARIMA)	<i>white noise</i>	<i>significant</i>	keterangan
2,1,1	v	x	-
1,1,0	x	v	-
1,1,1	v	x	-
0,1,1	v	v	tanpa konstanta

Pada tabel 2 di atas ,menunjukkan bahwa model ARIMA (0,1,1) telah memenuhi kelayakan model pada *time series* yaitu data yang *white noise* dan *significant*. Dari persamaan (1), maka untuk bentuk model ARIMA (0,1,1) adalah

$$\varphi_0(B)(1 - B)^1 Y_t = \theta_1(B)a_t \tag{4}$$

dengan $\theta_1 = 0,6183$, sehingga

$$Y_t = Y_{t-1} + a_t - 0,6183 a_{t-1} \tag{5}$$

B. Penerapan Metode Filter Kalman

Akan dilakukan simulasi terhadap hasil perhitungan sebelumnya ke dalam algoritma Filter Kalman dengan menggunakan *software* Matlab R2010a. Pada kasus ini akan

mengestimasi nilai koefisien polinomial a_0 dan a_1 yang ada pada persamaan model berikut:[9]

$$y_i^0 = a_{0,i} + a_{1,i}m_i + \dots + a_{n-1,i}m_i^{n-1} + \varepsilon_i \quad (6)$$

Karena dalam estimasi ini akan mengambil nilai $n = 2$ sehingga persamaan (6) berubah menjadi

$$y_i^0 = a_{0,i} + a_{1,i}m_i \quad (7)$$

dengan $x(t_i) = \begin{bmatrix} a_{0,i} \\ a_{1,i} \end{bmatrix}$ dan $H_i = [1 \ m_i]$ dimana m_i

merupakan data ke- i . Selain itu juga akan ditentukan beberapa sebagai nilai awal yaitu

A sebagai matriks sistem.

N sebagai masukkan banyak iterasi yang diinginkan.

Q sebagai matriks kovarians.

$R_k = R$ sebagai matriks kovarians.

a_{00} sebagai masukkan nilai awal a_{00} .

a_{10} sebagai masukkan nilai awal a_{10} .

Setelah mempunyai nilai awal maka untuk selanjutnya akan mencari nilai dari *noise* dengan random yang berdistribusi normal. Dari menjalankan perangkat lunak akan ditentukan yakni

Model Sistem[11]

$$X_{k+1} = A_k X_k + B_k U_k + G_k w_k \quad (8)$$

$$\begin{bmatrix} a_{0,i} \\ a_{1,i} \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{0,i} \\ a_{1,i} \end{bmatrix}_k + w_k \quad (9)$$

Model Pengukuran[11]

$$z_k = H_k X_k + v_k \quad (10)$$

$$y_k^0 = [1 \ m_i] \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}_k + v_k \quad (11)$$

Pada Tahap Prediksi ini akan dicari nilai *kovarians error* P_k^- dan nilai estimasi \hat{X}_k^- sebagai berikut[11]

$$\hat{X}_k^- = A \hat{X}_{k-1} + w_k \quad (12)$$

dengan diberikan nilai $\hat{X}_0 = \begin{bmatrix} -8 \\ 8 \end{bmatrix}$; $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ dan

w_k yang didapat dengan random berdistribusi normal.

$$P_k^- = A P_{k-1}^- A^T + Q_k \quad (13)$$

dengan $P_0 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ dan $Q_k = \begin{bmatrix} 0,001 & 0 \\ 0 & 0,001 \end{bmatrix}$

Pada Tahap Koreksi akan mencari nilai dari *Kalman Gain* K_k dengan menggunakan hasil perhitungan pada tahap prediksi sebagai berikut[11]

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (14)$$

dengan $R = 1$ dan untuk mendapatkan nilai koreksi dari \hat{X}_k^- akan juga menggunakan nilai \hat{X}_k^- yang telah didapatkan pada tahap sebelumnya sebagai berikut[11]

$$\hat{X}_k = \hat{X}_k^- + K_k (z_k - H \hat{X}_k^-) \quad (15)$$

dengan z_k merupakan identik dengan y_k^0 yang didapatkan dari bias atau selisih antara data dengan hasil *forecast* pada minitab ke-k. Lalu untuk menghitung nilai *kovarians error* P_k pada tahap koreksi ini akan menggunakan nilai dari yang didapatkan dari perhitungan di tahap prediksi sebagai berikut[11]

$$P_k = (I - K_k H) P_k^- \quad (16)$$

dengan I merupakan matriks identitas.

III. ANALISA DAN PEMBAHASAN

Simulasi dilakukan dengan menerapkan algoritma Filter Kalman pada persamaan. Hasil simulasi akan dievaluasi dengan cara membandingkan keadaan *real* dengan hasil estimasi Filter Kalman dan ARIMA. Dan pada tahapan ini, akan dilakukan 3 kali simulasi dengan banyak menjalankan perangkat lunak yang berbeda-beda yaitu 3, 5, dan 7 kali menjalankan perangkat lunak baik pada data suhu udara maupun kecepatan angin. Dan akan ditentukan beberapa nilai awalan untuk variabel yang ada yaitu diantaranya adalah

A sebagai matriks sistem adalah $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Q sebagai matriks kovarians adalah 0,001 dengan

$$Q_k = \begin{bmatrix} 0,001 & 0 \\ 0 & 0,001 \end{bmatrix}$$

$R_k = R$ sebagai matriks kovarians adalah 1.

a_{00} sebagai masukkan nilai awal a_{00} adalah -8.

a_{10} sebagai masukkan nilai awal a_{10} adalah 8.

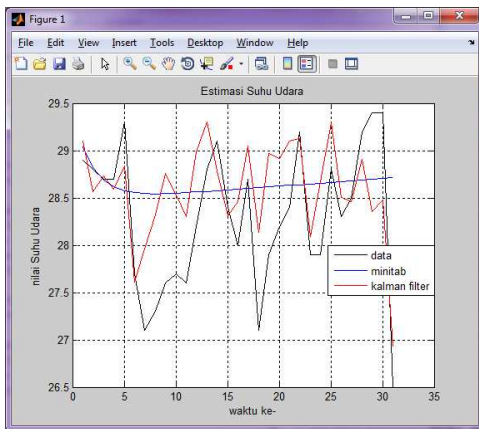
$$\text{nilai awal estimasi } \hat{X}_0 = \begin{bmatrix} a_{00} \\ a_{10} \end{bmatrix}$$

$$\text{nilai awal kovarians error } P_0 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

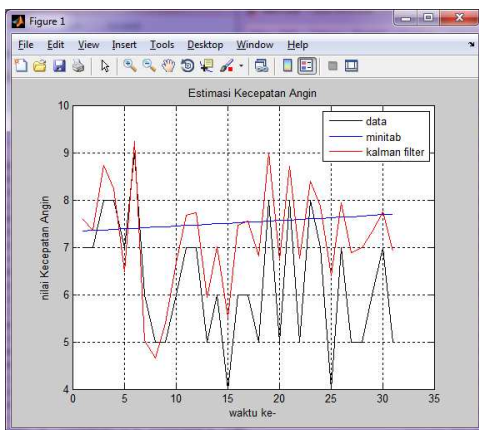
Selanjutnya akan dilakukan dengan menerapkan algoritma Filter Kalman pada persamaan. Hasil simulasi akan dievaluasi dengan cara membandingkan keadaan *real* dengan hasil estimasi Filter Kalman dan ARIMA. Dan pada tahapan ini, akan dilakukan simulasi sebanyak 3 kali menjalankan perangkat lunak baik pada data suhu udara maupun kecepatan angin.

ditunjukkan pada Gambar 1. Dari Gambar 1 tampak baik simulasi pada data suhu udara maupun data kecepatan angin memiliki rata-ran yang lebih mendekati data setelah menggunakan algoritma Filter Kalman jika dibandingkan dengan hasil prediksi dari ARIMA.

Melalui cara yang sama maka hasil simulasi akan dievaluasi dengan cara membandingkan keadaan *real* dengan hasil estimasi Filter Kalman dan ARIMA. Dan pada tahapan ini, akan dilakukan simulasi sebanyak 5 kali menjalankan perangkat lunak baik pada data suhu udara maupun kecepatan angin.



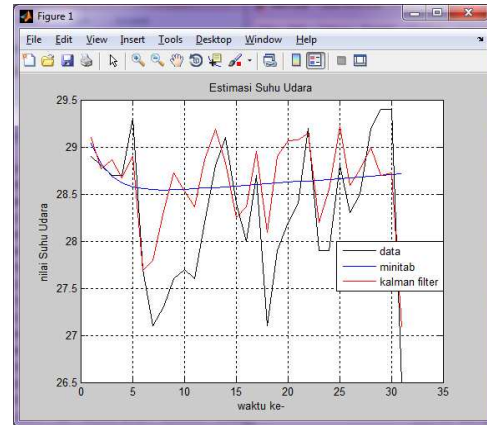
(a)



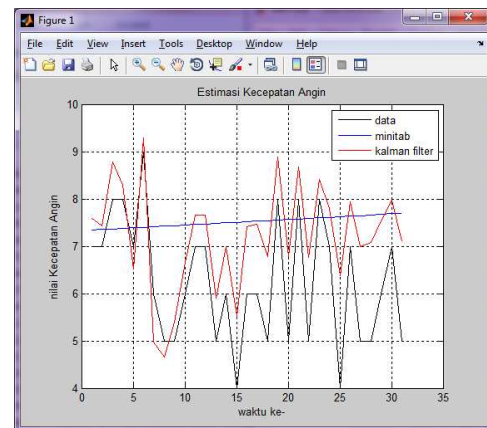
(b)

Gambar 1. Hasil Simulasi I pada Data (a) Suhu Udara dan (b) Kecepatan Angin

Dengan 3 kali menjalankan perangkat lunak akan dicari nilai rata-rata yang kemudian dilakukan plot dengan menggunakan *software* MatlabR2010 seperti yang



(a)



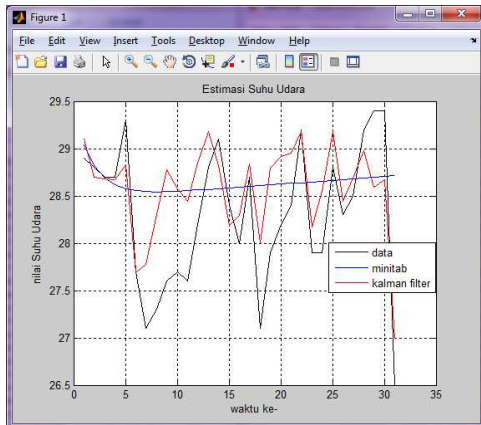
(b)

Gambar 2. Hasil Simulasi II pada Data (a) Suhu Udara dan (b) Kecepatan Angin

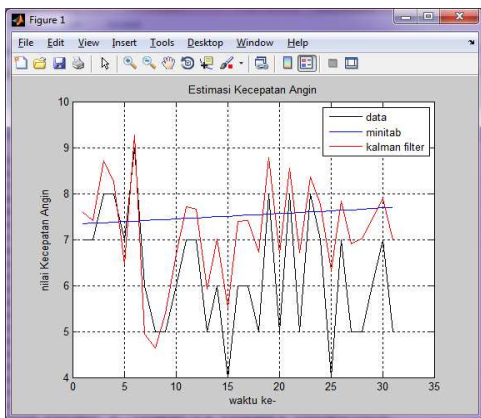
Setelah 5 kali menjalankan perangkat lunak maka akan dicari nilai rata-rata yang kemudian dilakukan plot dengan menggunakan *software* MatlabR2010 seperti yang ditunjukkan pada Gambar 2. Dari Gambar 2 tampak baik simulasi pada data suhu udara maupun data kecepatan angin memiliki rata-ran yang lebih mendekati data setelah menggunakan algoritma Filter Kalman jika dibandingkan dengan hasil prediksi dari ARIMA.

Kemudian dari cara yang sama maka hasil simulasi akan dievaluasi dengan cara membandingkan keadaan *real* dengan

hasil estimasi Filter Kalman dan ARIMA. Dan pada tahapan ini, akan dilakukan simulasi sebanyak 7,kali menjalankan perangkat lunak baik pada data suhu udara maupun kecepatan angin.



(a)



(b)

Gambar 3. Hasil Simulasi III pada Data (a) Suhu Udara dan (b) Kecepatan Angin

Dengan 7 kali menjalankan perangkat lunak akan dicari nilai rata-rata yang kemudian dilakukan plot dengan menggunakan software MatlabR2010 seperti yang ditunjukkan pada Gambar 3. Dari Gambar 3 tampak baik simulasi pada data suhu udara maupun data kecepatan angin memiliki ratahan yang lebih mendekati data setelah menggunakan algoritma Filter Kalman jika dibandingkan dengan hasil prediksi dari ARIMA.

Dapat diamati bersama dari Gambar 1-3 bahwa kondisi yang paling sesuai untuk mengestimasi nilai polinomial yang nantinya berguna sebagai perbaikan prediksi dari ARIMA melalui minitab dengan Filter Kalman adalah disaat menggunakan kondisi pada simulasi III pada data suhu udara dan kondisi simulasi III pada data kecepatan angin. Dapat dibandingkan dari gambar simulasi I, simulasi II, dan simulasi III pada data suhu udara maka yang simulasi III menunjukkan hasil prediksi yang lebih akurat dan mendekati data aslinya seperti yang ditunjukkan pada Gambar 3 Bagian (a).

Dalam hal ini suatu peramalan tidak lepas dari suatu kesalahan perhitungan peramalan. Ada beberapa metode untuk mengevaluasi hasil peramalan yang nantinya dapat digunakan untuk mengetahui keakuratan hasil peramalan yang telah dilakukan terhadap data yang sebenarnya. Metode yang digunakan dalam penelitian ini yaitu *Mean Absolute Percentage Error* (MAPE) yang merupakan nilai tengah kesalahan persentase nilai mutlak dari suatu peramalan.

$$\text{Didefinisikan : } MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{X_t - F_t}{X_t} (100) \right|$$

dengan X_t adalah nilai data ke- t , F_t adalah nilai peramalan ke- t , dan N adalah banyaknya data[10]. Sehingga untuk nilai MAPE dari hasil peramalan yang telah dilakukan ditunjukkan oleh tabel berikut:

Tabel 3. Nilai MAPE pada Hasil Peramalan Data Suhu Udara

No.	Model	Nilai MAPE
1	ARIMA	2,20288828
2	KF Simulasi I	2,100066667
3	KF Simulasi II	1,98968
4	KF Simulasi III	1,935357143

Sedangkan dapat juga dibandingkan dari gambar simulasi I, simulasi II, dan simulasi III pada data kecepatan angin maka yang simulasi III menunjukkan hasil prediksi yang lebih akurat dan mendekati data aslinya seperti yang ditunjukkan pada Gambar 3 Bagian (b). Di samping itu juga dari gambar di bawah ini dapat dibandingkan nilai mutlak kesalahan peramalan baik untuk ARIMA dengan minitab dan Filter Kalman dengan hasil simulasi III pada data suhu udara menunjukkan bahwa nilai mutlak kesalahan peramalan untuk Filter Kalman simulasi III cenderung lebih kecil jika dibandingkan dengan yang lain.

Selain itu, dapat diketahui juga akurasi peramalan yang ada dengan melihat nilai MAPE dari hasil peramalan yang telah dilakukan seperti yang ditunjukkan oleh tabel berikut:

Tabel 4. Nilai MAPE pada Hasil Peramalan Data Kecepatan Angin

No.	Model	Nilai MAPE
1	ARIMA	28,90889689
2	KF Simulasi I	19,0153
3	KF Simulasi II	19,30076
4	KF Simulasi III	18,73821429

IV. KESIMPULAN

Dari seluruh proses yang telah dibahas di atas dapat disimpulkan bahwa model data adalah ARIMA (1,1,1) untuk suhu udara dan ARIMA (0,1,1) untuk kecepatan angin. Selain itu, penggunaan algoritma Filter Kalman mempunyai pengaruh baik terhadap perbaikan hasil prediksi terbukti dari nilai MAPE pada data suhu udara dengan ARIMA adalah 2,20288828 sedangkan nilai MAPE dengan Filter Kalman mempunyai nilai yang lebih kecil yaitu 1,935357143. Hal yang sama juga berlaku pada data kecepatan angin dimana penggunaan algoritma Filter Kalman mempunyai pengaruh baik terhadap perbaikan hasil prediksi terbukti dari nilai MAPE dengan ARIMA adalah 28,90889689 sedangkan nilai MAPE dengan Filter Kalman mempunyai nilai yang lebih kecil yaitu 18,73821429.

DAFTAR PUSTAKA

- [1] Anonim. (2013). 10 Negara Dengan Jumlah Penduduk Terbanyak di Dunia. <http://suhadahanum.blogspot.com/2013/09/10-negara-dengan-jumlah-penduduk.html>. diakses tanggal 27 Agustus 2013.
- [2] Anonim. (2013). 10 Kota Terpadat di Indonesia. <http://pengetahuanpintars.com/2013/11/10-kota-terpadat-di-indonesia.html>. diakses tanggal 29 Agustus 2013.
- [3] Anonim. (2013). Penerbangan Juanda dihentikan akibat cuaca buruk. <http://antaranews.com/2013/12/penerbangan-juanda-dihentikan-akibat-cuaca-buruk.html>. diakses tanggal 12 September 2013.
- [4] Anonim. (2013). Cuaca. <http://wikipedia.com/2013/09/cuaca.html>. diakses tanggal 29 Agustus 2013.
- [5] Zakir, A. Analisis dan Pengembangan Sistem Metode Prakiraan Cuaca di Bidang Informasi Meteorologi. Tugas Akhir, Universitas Gunadarma : Jakarta.
- [6] Wei, W. S. (1990). Time Analysis Univariate and Multivariate Methods. Addison Wesley Publishing Company, Inc. : Amerika.