ENSEMBLE KALMAN FILTER PADA PEMULUSAN GANDA DUA PARAMETER DARI HOLT

Oleh :
Hery Andi Sitompul [1)]
Enzo W. B. Siahaan [2)]
Universitas Darma Agung, Medan [1,2)]
E-mail:
herystpl@gmail.com [1)]
Enzo.battra84@gmail.com [2)]

ABSTRACT

Forecasting will always produce errors as well as the model or forecasting method used, therefore it is very important to reduce these errors or errors. This study aims at applying the Ensemble Kalman Filter method to the forecasting model by smoothing two parameters from Holt and obtaining forecasting results with a smaller percentage of error than the results of trial and error. The Ensemble Kalman Filter method is offered as a way to reduce errors generated by Holt's Two-Parameter Double Exponential Smoothing model. In the study in this paper, it is proven that the combination of the Ensemble Kalman Filter on Holt's Two Parameter Exponential Smoothing can reduce errors significantly.
Keywords: Exponential Smoothing, MAPE, EnKF, Time Series

ABSTRAK

Peramalan akan selalu menghasilkan kesalahan atau galat sebaik apapun model atau metode peramalan yang dipergunakan, oleh karena itu penting sekali untuk mengurangi kesalahan atau galat tersebut. Tujan dari penelitian ini adalah untuk menerapkan metode Ensemble Kalman Filter pada model peramalan dengan pemulusan dua parameter dari Holt dan mendapatkan hasil peramalan dengan persentase kesalahan yang lebih kecil dari *hasil trial and error*. Metode Ensemble Kalman Filter ditawarkan sebagai satu cara untuk mengurangi persentase kesalahan yang dihasilkan oleh model Pemulusan Eksponensial Ganda Dua Paramneter dari Holt. Dalam kajian pada tulisan ini terbukti bahwa kombinasi Ensemble Kalman Filter pada Pemulusan Eksponensial Dua Parameter dari Holt dapat mengurangi persentase kesalahan secara signifikan.
Kata Kunci : **Pemulusan Eksponensial, MAPE, EnKF, Runtun Waktu**

1. PENDAHULUAN

1.1. Latar Belakang

Peramalan (*forecasting*) merupakan salah satu cabang matematika khususnya stastika yang digunakan untuk mengestimasi atau memprediksi suatu nilai variabel tak bebas dimasa mendatang berdasarkan deretan data masa lalu. Semakin akurat hasil estimasi dari suatu model peramalan maka model tersebut dapat dikatakan tangguh.

Terdapat berbagai metode peramalan yang sudah umum diketahui salah satunya adalah metode pemulusan. Pada umumnya metode ini menggunakan prinsip *trial and error* pada parameternya, dimana parameter yang menghasilkan nilai persentase kesalahan yang lebih kecil maka model tersebut yang akan menjadi model terbaik untuk prediksi.

Berdasarkan pengalaman tersebut, maka dalam artikel ini akan dicoba suatu formula baru apakah hasil *trial and error* tersebut dapat dikoreksi dengan metode tertentu. Dalam tulisan ini akan ditawarkan metode Ensemble Kalman Filter untuk mengoreksi hasil terbaik dari *trial and error* tersebut. Hal ini karena prinsip dari Ensemble Kalman Filter adalah mengasssimilasikan sebuah data pengukuran atau sebenarnya pada model dalam kurun waktu tertentu.

1.2. Maksud dan Tujuan

Maksud dari tulisan ini adalah untuk melakukan kajian apakah metode Ensemble Kalman Filter dapat mengoreksi hasil peramalan dengan metode pemulusan dua parameter dari Holt. Sedangkan tujuan dari tulisan ini adalah :

1. Menerapkan metode Ensemble Kalman Filter pada model peramalan dengan pemulusan dua parameter dari Holt.
2. Mendapatkan hasil peramalan dengan persentase kesalahan yang lebih kecil dari *hasil trial and error*.

2. TINJAUAN PUSTAKA

2.1. Peramalan Berbasis Runtun waktu

Runtun waktu merupakan kumpulan data yang tercatat berdasarkan periode waktu tertentu, misal : hari, minggu, bulan, dan tahun. Peramalan berbasis runtun waktu adalah suatu metode peramalan berdasarkan hasil analisa runtun waktu tersebut. Komponen dari analisa runtun waktu adalah tren, siklus, musiman dan gerak tak beraturan. Terdapat berbagai metode peramalan berbasis runtun waktu, diantaranya :

1. ARIMA (Autoregressive Integrated Moving Average), adalah model yang mengabaikan sepenuhnya variabel indepen dalam membuat peramalan. Model ini menggunakan data masa lalu dan sekarang dari variabel dependen untuk mengasilkan peramalan jangka pendek.
2. Regresi Deret Waktu, merupakan metode proyeksi tren dengan model regresi baik untuk jangka panjang dan pendek. Metode ini merupakan garis tren untuk persamaan matematis.
3. Metode Pemulusan. Metode ini bertujuan untuk mengurangi ketidakteraturan dari data masa lalu seperti faktor musiman, dan merupakan jenis peramalan untuk jangka pendek.

2.2. Metode Pemulusan

Metode pemulusan adalah sebuah metode peramalan kuantitatif yang melakukan pembobotan sederhana terhadap data masa lampau untuk memprediksi masa yang akan datang. Metode ini terdiri dari dua bagian yaitu :

1. Rata – rata bergerak (Moving Average/MA) merupakan metode yang sederhana dengan menggunakan rataan data deret waktu yang bergerak.
2. Pemulusan Eksopnensial, metode ini menggunakan pembobotan pada data masa lalu yang menurun secara

eksponensial dan dibutuhkan penentuan nilai parameter yang terbaik.

2.2.1. Pemulusan Eksponensial

Menurut Makridakis pemulusan eksponensial adalah suatu prosedur perbaikan terus menerus dalam peramalan sebuah objek oleh pengamatan terbaru dimana data yang lebih tua akan lebih kecil bobotnya dibandingkan dengan data yang lebih baru. Dalam pemulusan eksponensial terdapat satu atau lebih parameter yang dapat ditentukan secara eksplisit melalui sistem *trial and error*. Metode pemulusan eksponensial terdiri dari : Pemulusan eksponensial tunggal, ganda, dan tripel.

2.2.2. Pemulusan Eksponensial Tunggal

Metode ini biasanya digunakan untuk peramalan jangka pendek, misalnya satu bulan kedepan.Model ini mengasumsikan bahwa data bergerak disekitar nilai mean (rata – rata hitung) yang tetap tanpa adanya tren atau pola pertumbuhan konsisten.

Persamaan untuk Pemulusan Eksponensial Tunggal adalah :

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$

Dimana :

Y_t = Nilai aktual runtun waktu

$\hat{Y}_t$= Nilai peramalan pada waktu *t*

$\hat{Y}_{t+1}$ = Nilai peramalan pada waktu *t+1*

α = Parameter peramalan bernilai antara 0 dan 1

t = waktu aktual

2.2.3. Pemulusan Eksponensial Ganda

Metode ini menggunakan dua kali pemulusan yaitu pada periode level dan periode tren, metode ini digunakan pada data runtun waktu yang mengandung pola tren. Metode ini terbagi dua yakni Pemulusan Eksponensial Ganda satu parameter dari Brown dengan persamaan :

Inisasi Nilai Awal : $S'_t = S''_t = X_1$

$$F_{t+m} = a_t + b_t m$$

$$a_t = 2S''_t - S'_t$$

$$b_t = \frac{\alpha}{1-\alpha}(S'_t - S''_t)$$

$$S'_t = \alpha X_t + (1-\alpha)S'_{t-1}$$

$$S''_t = \alpha S'_t + (1-\alpha)S''_{t-1}$$

Dan Pemulusan Eksponensial Ganda Dua Parameter dari Holt, dengan persamaan sebagai berikut :

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + b_{t-1})$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}$$

$$F_{t+m} = a_t + b_t m$$

Dimana :

F_t = Nilai ramalan yang diinginkan.

S_t = Pemulusan pada waktu *t*

b_t = Pemulusan tren pada waktu *t*

m = periode waktu yang akan diramalkan

2.3. *Mean Absolute Percentage Error* (MAPE)

Mean Absolute Percentage Error (MAPE) merupakan salah satu ukuran kebaikan dari sebuah model peramalan, dimana perbedaan antara nilai peramalan dan nilai sebenarnya dihitung dalam bentuk nilai mutlak dan dijadikan dalam nilai persentase terhadap nilai sebenarnya, jika nilai MAPE dibawah 10% maka kinerja model sangat bagus dan bila nilai MAPE antara 10% dan 20% kinerja model dikatakan bagus. Nilai MAPE dapat ditentukan dengan :

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left\|\frac{Y_t - \widehat{Y_t}}{Y_t}\right\| \times 100\%$$

2.3. Kalman Filter dan Ensemble Kalman Filter

Kalman Filter merupakan metode pendekatan estimasi fungsi parameter dalam peramalan deret waktu (*time series*), berupa algoritma yang menggabungkan model dan pengukuran (observasi). Algoritma ini mengestimasi keadaan dari suatu proses dengan cara meminimumkan rataan dari galat (*Mean square Error*) secara rekursif dan efisien, sehingga sering disebut penduga rekursif. Data pegukuran terbaru menjadi bagian penting dari algoritma ini, karena data terbaru ini akan mengoreksi hasil prediksi, sehingga hasil estimasi selalu mendekati kondisi sebenarnya. Sedangkan Ensemble Kalman Filter adalah kelanjutan dari Kalman Filter untuk masalah nonlinier, yaitu merupakan pendekatan metode Monte Carlo pada algoritma Kalman Filter.

2.3.1.Kalman Filter

Misalkan $x_n \in \mathbb{R}$ adalah suatu proses stokastik dan asumsikan bahwa x_n memenuhi persamaan linier berikut :

$x_n = Ax_{n-1} + Bu_{n-1} + w_{n-1}, n \geq 1$ ….(1)

$z_n = Hx_n + v_n$…………………………..(2)

Dimana :

A adalah matriks model linier

B adalah matriks kontrol masukan

u_n adalah variabel kontrol

w_n adalah gangguan pada sistem

x_n adalah variabel *state*

z_n adalah pengukuran atau observasi

v_n adalah gangguan pada pengukuran

n adalah indeks integer waktu

Dengan keadaan awal x_0 diberikan berdistribusi normal :

$x_0 \sim N(0, P_0^*)$ ……………………………(3)

Dimana P_0^* adalah matriks kovariansi untuk proses stokastik x_0 dan $w_n \sim N(0, Q_n), v_n \sim N(0, R_n)$ dimana Q_n adalah matriks kovariansi gangguan pada sistem, R_n adalah matriks kovariansi untuk gangguan pada pengukuran.

Misalkan x_n didefenisikan sebagai keadaan (*state*) pada waktu t_n dan x_{n+1} menyatakan keadaan pada saat t_{n+1}, dan misalkan $\hat{x}_{n+1}^-$ menyatakan estimasi prior dari proses *state* berdasarkan informasi yang diperoleh hingga waktu t_{n+1} , dan misalkan $\hat{x}_n$ menyatakan estimasi terbaik dari proses *state* (posterior) sistem yang telah difilter hingga waktu t_n, maka dapat dirumuskan bahwa :

$\hat{x}_{n+1}^- = E(x_{n+1}|z^*)$

$\hat{x}_n = E(x_n|z^*)$

Dimana z^* adalah hasil pengukuran hingga waktu t_n. Dengan demikian model posterior untuk satu tahap waktu kedepan dapat dicari dengan :

$\hat{x}_{n+1}^- = A\hat{x}_n + Bu_n$ ……………………..(4)

Dan matriks kovariansi dapat dihitung dengan persamaan berikut :

$P_{n+1}^- = AP_nA' + Q$

Dimana

$P_{n+1}^- = E[(\hat{x}_{n+1}^- - \hat{x}_n)(\hat{x}_{n+1}^- - \hat{x}_n)'|z^*]$

$P_n = E[(\hat{x}_n - x_n)(\hat{x}_n - x_n)'|z^*$

tahap pemfilteran adalah tahap dimana estimasi posterior dihitung, yang dapat ditentukan melalui persamaan berikut :

$\hat{x}_n = \hat{x}_n^- + K_n(z_n - H\hat{x}_n^-)$

Dengan K_n adalah kalman Gain yang dapat dihitung dengan :

$K_n = P_n^- H'[HP_n^- H' + R_n]^{-1}$ …………...(5)

dan matriks kovariansi posterior dapat diperbaharui melalui persamaan berikut :

$P_n^* = (I - K_n H)P_n^-$

2.3.2. Ensemble Kalman Filter

Ensemble Kalman Filter (EnKF) dikembangkan dan diperkenalkan oleh Geir Evensen tahun 1994, untuk mengatasi masalah nonlinier pada model oceanografi dan menunjukkan bahwa EnKF adalah suatu metode yang menjanjikan.

Misalkan $x_n \in \mathbb{R}$ adalah suatu proses stokastik, dan diasumsikan bahwa x_n memenuhi persamaan nonlinier berikut :

$x_n = f(x_{n-1}) + w_n$ …………………….(6)

$z_n = Hx_n + v_n$ …………………………(7)

Gagasan dari ensemble adalah simulasi Monte Carlo pada model *state* (x_n), dimana satu kumpulan ensemble dibangkitkan untuk model *state* (x_n) kemudian setiap anggota ensemble dievaluasi melalui persamaan (6) untuk mendapatkan sekelompok ensemble baru yang disebut dengan prior yang selanjutnya dilakukan estimasi dengan proses Kalman Filter biasa. Hal yang sama juga dilakukan terhadap model observasi (7) dengan membangkitkan sekelompok melalui observasi sebenarnya, dengan ensemble ini proses updating/filter dilakukan dengan Kalman Filter.

Berikut ini adalah tahapan dalam proses Ensemble Kalman Filter (EnKF) :

- $X := [x_1, x_2, \dots, x_n]' \in \mathbb{R}^{N \times N_e}$ adalah matriks ensemble untuk *state*.
- $\overline{X} \coloneqq [\overline{x}_1, \overline{x}_2, \dots, \overline{x}_n]' \in \mathbb{R}^{N \times 1}$ adalah vektor rata – rata ensemble.
- $z \in \mathbb{R}^{m \times 1}$ adalah vektor observasi
- $P_e := \frac{(X-\overline{X})(X-\overline{X})'}{N_e - 1}$ adalah matriks kovariansi ensemble.
- $Z \coloneqq z + \gamma \in \mathbb{R}^{m \times N_e}$adalah matriks observasi dengan gangguan.
- $\gamma: (\in_1, \in_2, \dots, \in_n)$ adalah vektor gangguan dimana setiap $\in_i \sim N(0, \sum_i^e)$
- $Re := \frac{\gamma\gamma'}{N_e - 1}$ adalah matriks observasi dengan gangguan.

3. METODE PENELITIAN

Tahapan kerja yang dilakukan dalam tulisan ini akan dimulai dengan mengambil sebuah data runtun waktu yang memuat pola tren. Contoh kasus yang digunakan dalam tulisan ini adalah data sekunder dari pihak kedua yaitu dari Nila Yuwida, Lukman Hanafi, Nuri Wahyuningsih[] berupa data pengunjung Kusuma Agrowisata Batu, Malang tahun 2009 sampai 2010.

Periode	Jumlah Pengunjung	Periode	Jumlah Pengunjung
1	1430	13	2060
2	1520	14	1930
3	1610	15	2070
4	1390	16	2180
5	1370	17	2290
6	1740	18	2250
7	1420	19	2040
8	1410	20	2270
9	1620	21	2230
10	1800	22	2420
11	1640	23	2390
12	1710	24	2660

Dari data diatas akan dilakukan peramalan dengan model Pemulusan Eksponensial Ganda Dua Parameter dari Holt dengan cara *trial and error* pada nilai parameternya sampai dieperoleh parameter yang optimal.

Model Pemulusan Eksponensial Dua Parameter dari Holt terbaik yang diperoleh selanjutnya akan dikoreksi dengan metode Ensemble Kalman Filter dengan mengikuti algoritma berikut :

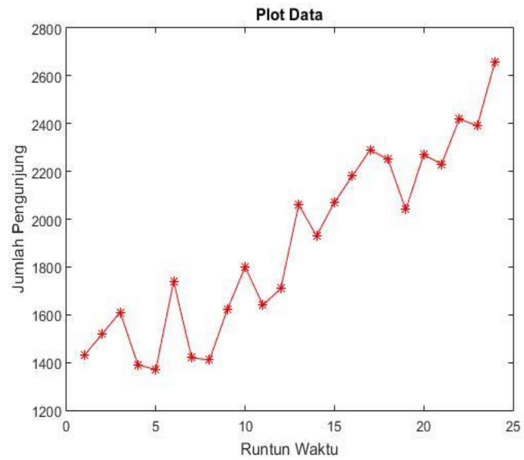
Pada saat $n = 0$

- Masukkan ensemble variabel awal : $x_0 = X \sim N(\mu_0, \sum_0)$
- Langkah prediksi : bangkitkan gangguan pada model $w_n \sim N(0, \sum_n^w)$, prediksi satu tahap kedepan $\hat{x}_n^- = f(\hat{x}_{n-1}^-) + w_n$, $f(\hat{x}_{n-1}^-)$. Tahapan prediksi disini menggunakan model optimal dari Pemulusan Eksponensial Ganda Dua Parameter dari Holt yang diperoleh pada tahap sebelumnya.
- Tahap *filter/updating* : bangkitkan gangguan pada observasi $\in_n \sim N(0, \sum_n^\in)$, $z_n = H\hat{x}_n^- + \in_n$
- Tentukan Kalman Gain : $K_n = PeH[HPeH' + Re]^{-1}$
- Perbaharui *state* dengan pengukuran (assimilasi data/data solusi eksak) : $\hat{x}_n = \hat{x}_n^- + K_n(z_n - H\hat{x}_n^-)$.

Tahapan selanjutnya adalah membandingkan hasil prediksi dari kedua model peramalan tersebut, apakah metode Ensemble Kalman Filter yang ditawarkan dapat memberikan perbaikan hasil.

4. HASIL DAN PEMBAHASAN

Data runtun waktu yang disajikan dalam tabel diatas diplot untuk melihat kecenderungan data.



Gambar 1. Plot data aktual

Hasil plotting data dalam gambar terlihat dengan jelas bahwa adanya pola tren dengan kecenderungan naik seiring dengan berjalannya waktu meskipun sedikit fluktuatif, dengan demikian model yang tepat untuk menganalisa data tersebut adalah model pemulusan ganda, yaitu Dua Parameter Holt.

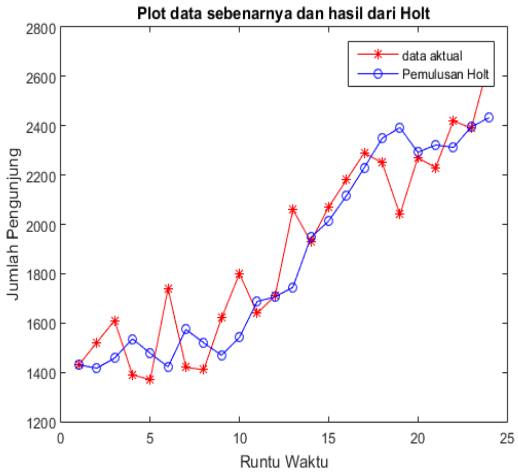
Langkah perhitungan akan dimulai dengan inisiasi nilai awal peramalan dengan ketentuan sebagai berikut :

$S_1 = X_1 = 1430$

$b_1 = \frac{(X_2-X_1)+(X_3-X_2)+(X_4-X_3)}{3} = -13.3333$

Dari hasil running program matlab perhitungan peramalan Pemulusan Eksponensial Ganda Dua Parameter dari Holt, dengan melakukan *trial and error* pada kombinasi nilai parameter α dan γ antara 0 dan 1 diperoleh model parameter terbaik adalah $\alpha = 0.4$ dan $\gamma = 0.3$ dengan MAPE terkecil yakni 6.7839%. Sehingga model Pemulusan Eksponensial Dua Parameter dari holt adalah :

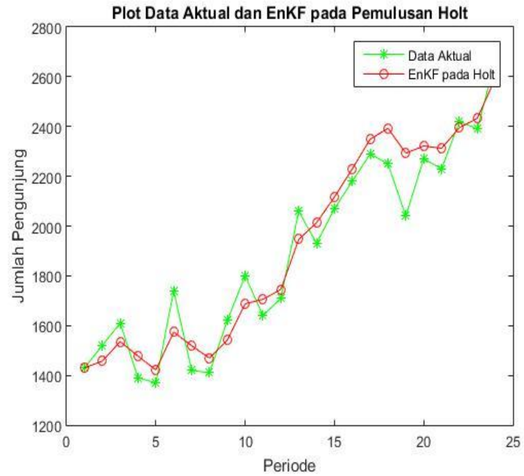
$$S_t = 0.4X_t + (0.6)(S_{t-1} + b_{t-1})$$
$$b_t = 0.3(S_t - S_{t-1}) + (0.7)b_{t-1}$$
$$F_{t+m} = a_t + b_t m$$



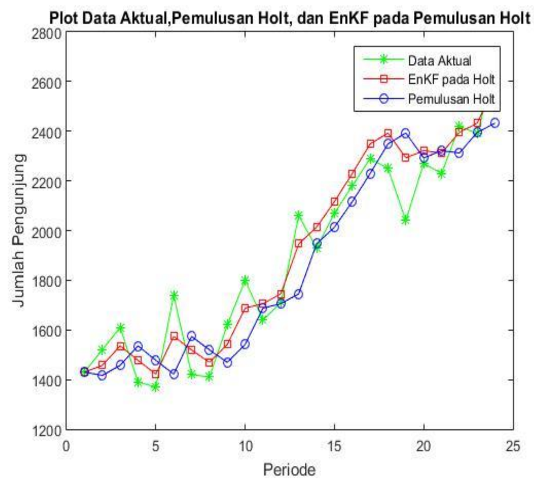
Gambar 2. Plot hasil peramalan dan data aktual

Model peramalan ini terbilang cukup bagus karena nilai MAPE<10%. Tetapi akan lebih baik lagi jika kita dapat memperkecil persentase kesalahan dari peramalan tersebut.

Selanjutnya model terbaik yang diperoleh akan dikombinasikan dengan metode Ensemble Kalman Filter dengan tujuan untuk memperkecil nilai MAPE sehingga model dapat dikatakan lebih akurat. Proses ini dilakukan mengikuti algoritma yang disajikan diawal dengan mengambil jumlah ensemble 50. Hasil running program matlab untuk tahap ini dapat dilihat dari gambar grafik berikut :



Gambar 3. Plot data aktual dan peramalan dengan EnKF Pada Pemulusan Holt



Gambar 4. Plot data aktual, peramalan dengan pemulusan Holt, dan peramalan dengan EnKF pada pemulusan Holt

Terlihat bahwa grafik Ensemble Kalman Filter (EnKF) pada Pemulusan Holt sangat dekat dengan data aktual jika dibandingkan dengan Pemulusan Ekspnensial dari Holt saja. Dari perhitungan nilai MAPE juga menunjukkan hal yang sama, dapat dilihat dari tabel berikut :

Model	MAPE
Pemulusan Dua Parameter Holt	6.7839%
Ensemble Kalman Filter pada Pemulusan Holt	4.3206%

Tabel 2. Perbandingan akurasi model

Dapat dilihat bahwa terjadi koreksi persentase kesalahan sebesar 2.4633% oleh karena kombinasi metode Ensemble Kalman Filter dan Pemulusan Eksponensial Dua Parameter dari Holt.

5. SIMPULAN DAN SARAN

5.1. Simpulan

Dari hasil dan pembahasan pada contoh kasus yang sudah dilakukan diatas dapat ditarik kesimpulan yaitu :

1. Terjadi koreksi hasil peramalan model Pemulusan Eksponensial dua Parameter dari Holt oleh kombinasi Ensemble Kalman Filter pada Pemulusan Eksponensial Ganda Dua Parameter daari Holt sebesar 2.4633% .
2. Metode Ensemble Kalman Filter dapat dikombinasikan dengan Pemulusan Eksponensial Ganda Dua Parameter dari Holt dengan sangat baik.

5.2. Saran

Dalam kajian ini model terbaik Pemulusan Eksponensial Ganda Dua Parameter dari Holt dilakukan berdasarkan *trial and error* pada nilai parameternya, oleh karena itu perlu suatu cara untuk mendapat parameter yang optimum tanpa melakukan perhitungan yang berulang – ulang.

6. DAFTAR PUSTAKA

Geir Evensen (2003), *The Ensemble Kalman Filter : Theorical Formulation and Practical Implementation*. Ocean Dynamics Vol.53

Hery Andi Sitompul (2011), *Estimasi Parameter Reservoir Komposit dengan Teknik Ensemble Kalman Filter dan Algoritma Genetika*, Thesis ITB.

Makridakis S.,Wheelwright,S.C. (1989), *Forecasting Methods for Management*, *5 ed*.John Wiley & Sons.Inc.

Nila Yuwida,Lukman Hanafi,Nuri Wahyuningsih (2012), *Estimasi Parameter α dan γ Dalam Pemulusan Ganda Dua Parameter dengan Modifikasi Golden section*,Jurnal Sains dan Seni ITS vol I,No.1.

Wayne A.Fuller (1996), *Introduction to Statistical Time Series. Second edition*,John Wiley & Sons.Inc.

Yihua Yang (2020), *Ensemble Kalman Filtering with Perturbed Observation in Weather forcasting and Data Assimilation*. University of Reading.