ISBN: 978-602-14387-0-1

PREDIKSI CURAH HUJAN DI KOTA SEMARANG DENGAN METODE KALMAN FILTER

Tika Dhiyani Mirawati[1], Hasbi Yasin[2], Agus Rusgiyono[3]
[1]Mahasiswa Jurusan Statistika FSM UNDIP
[2,3]Staff Pengajar Jurusan Statistika FSM UNDIP

Abstract

The rainfall data is very interesting to be studied because it is constitutes one of the biggest factor that influence the climate on a region and human life sector. In this studies, the rainfall prediction is utilized by Kalman Filter method. The implementation of Kalman Filter analysis in this research is used for modelling and forecasting rainfall in Semarang city. This method provide a recursive solution to minimize error. Kalman Filter consists of state equation and observation equation. The maximum rainfall in 2013 is 406 mm happening in February and the minimum rainfall is 35 mm happening in July. Overall, the average rainfall in 2013 at Semarang city is 196,25 mm.

Keywords : Rainfall, Kalman Filter, Time Series

1. Pendahuluan

Menurut Dewantara (2012), cuaca dan iklim merupakan sebuah proses fenomena di atmosfer yang keberadaannya sangat penting dalam berbagai aktivitas kehidupan. Perhatian mengenai informasi cuaca dan iklim semakin meningkat seiring dengan meningkatnya fenomena alam yang tidak lazim terjadi atau biasa disebut dengan cuaca ekstrim yang sulit untuk dikendalikan dan dimodifikasi. Dampak yang ditimbulkan oleh cuaca ekstrim tersebut dapat diminimalisir dengan penyediaan informasi mengenai peluang terjadinya cuaca ekstrim seperti prediksi curah hujan di suatu daerah dalam jangka waktu tertentu, prediksi terjadinya gempa, angin kencang dan gelombang laut yang berpotensi mengakibatkan bencana alam. Data curah hujan sangat menarik untuk dikaji sebab curah hujan merupakan salah satu faktor terbesar yang mempengaruhi iklim suatu wilayah dan mempengaruhi berbagai sektor kehidupan manusia. Prediksi curah hujan dapat dilakukan dengan metode Kalman Filter. Metode ini digunakan untuk menyatakan suatu model runtun waktu yang ditampilkan dalam bentuk linier *state space* (Brockwell and Davis, 1991). Menurut Meinhold dan Singpurwala (1983), model, teknik, dan notasi dari Kalman Filter hampir sama dengan model regresi linier dan analisis runtun waktu. Perbedaannya terletak pada sifat rekursif yang ada pada Kalman Filter (Welch and Gary, 2001).

ISBN: 978-602-14387-0-1

Kalman Filter memiliki beberapa keunggulan dibandingkan dengan metode lain seperti yang dijelaskan berikut

1. Proses estimasi menggunakan bentuk dari kontrol umpan balik (rekursif) yang dapat memperkecil nilai *Mean Square Error* (MSE) dan *noise* (Tresnawati.R., dkk , 2010).
2. Dapat terus diperbaharui dengan data terbaru sehingga nilai prediksi selalu *update* (Welch and Gary, 2001).
3. Mudah diterapkan dalam berbagai disiplin ilmu karena sifatnya yang rekursif (Meinhold and Singpurwala, 1983).

Selain memiliki berbagai keunggulan Kalman Filter juga memiliki kelemahan menurut Wei (2006) yaitu keberhasilan dalam mendapatkan hasil prediksi optimal bergantung pada ketepatan estimasi keadaan (*state*) awal pada data observasi terbaru.

Pada penulisan tugas akhir ini akan diambil studi kasus untuk memprediksi jumlah curah hujan perbulan di kota Semarang dengan metode Kalman Filter. Sebelumnya data curah hujan diidentifikasi model ARIMA (p,d,q) untuk pembentukan model Kalman Filter. Setelah model pada Kalman Filter terbentuk, dilakukan peramalan pada data curah hujan perbulan di kota Semarang beberapa langkah ke depan. Data yang digunakan adalah data curah hujan di kota Semarang setiap bulannya dari tahun 2005 sampai dengan 2012 yang diambil dari Badan Meteorologi Klimatologi dan Geofisika (BMKG) Stasiun Klimatologi Semarang.

2. Tinjauan Pustaka

2.1 Analisis Runtun Waktu

Berdasarkan Makridakis *et.al* (1999), analisis runtun waktu mulai dikenalkan oleh George E.P.Box dan Gwilym M. Jenkins (1976). Dasar pemikiran runtun waktu adalah pengamatan sekarang tergantung pada satu atau beberapa pengamatan sebelumnya. Model runtun waktu dibuat karena secara statistik ada korelasi (dependensi) antar deret pengamatan. Untuk melihat adanya deret dependensi antar pengamatan, dapat melakukan uji korelasi antar pengamatan yang dikenal dengan *Autocorrelation Function* (ACF). Tujuan analisis runtun waktu antara lain memahami dan menjelaskan mekanisme tertentu, meramalkan suatu nilai di masa depan, dan

ISBN: 978-602-14387-0-1

mengoptimalkan sistem kendali. Analisis runtun waktu dapat diterapkan pada bidang ekonomi, bisnis, industri, teknik dan ilmu-ilmu sosial.

Menurut Soejoeti (1987), misalkan Z_1, Z_2, ... , Z_t merupakan proses stokastik untuk runtun waktu diskrit. Proses tersebut disebut stasioner jika mean dan variansinya konstan untuk setiap titik *t* dan kovarian yang konstan untuk setiap selang waktu *k*

1. $E(Z_t) = \mu$ konstan untuk semua *t*
2. $Var(Z_t) = \sigma^2$ konstan untuk semua *t*
3. $Cov(Z_t, Z_{t+k}) = \gamma_k$ konstan untuk semua *t* dan semua *k≠0* γ_k adalah autokovariansi pada lag *k*.

Ketidakstasioneran dalam runtun waktu dapat meliputi *mean* yang tidak konstan, varian yang tidak konstan ataupun keduanya (*mean* dan varian tidak konstan). Proses ketidakstasioneran dalam *mean* dapat diubah menjadi stasioner dengan melakukan pembedaan (differensi) yang tepat pada data. Jika proses runtun waktu tidak stasioner dalam varian maka dilakukan transformasi stabilitas varian.

Persamaan umum model ARIMA (p,d,q) :

$$\phi_p(B)(1-B)^d \dot{Z}_t = \theta_0 + \theta_q(B)\, a_t$$

dimana $\phi_p(B) = \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)$ adalah operator stasioner AR dan $\theta_q(B) = \left(1 - \theta_1 B - \cdots - \theta_q B^q\right)$ adalah operator invertibel MA. Saat d = 0 proses dalam keadaan stasioner. Orde p,d,q didapatkan dari lag yang signifikan pada plot pada plot *Autocorrelation Funtion*(ACF) dan *Partial Autocorrelation Funtion* (PACF).

2.2 Kalman Filter

Kalman Filter adalah sebuah metode bagian dari *state space* (ruang keadaan) yang dapat diterapkan dalam model prakiraan statistik. Sesuai dengan Wei (2006), metode ini menggunakan teknik rekursif dalam mengintegrasikan data pengamatan terbaru ke model untuk mengoreksi prediksi sebelumnya dan melakukan prediksi selanjutnya secara optimal berdasarkan informasi data di masa lalu maupun berdasarkan informasi data saat ini.

Berdasarkan Welch and Bishop (2001), konsep Kalman Filter terdiri dari dua tahapan yakni peramalan dan pembaharuan. Pada tahap peramalan, dihasilkan nilai estimasi untuk keadaan (*state*) di waktu sekarang dan nilai kovarian error yang digunakan sebagai informasi estimasi awal untuk langkah selanjutnya. Tahap pembaharuan berfungsi sebagai korektor. Pada tahap ini dihasilkan pengukuran baru

ISBN: 978-602-14387-0-1

yang didapat dari nilai estimasi awal. Setelah kedua tahap terpenuhi, proses tersebut akan berulang kembali dengan nilai estimasi yang didapat dari tahap pengukuran digunakan sebagai informasi awal tahap peramalan yang kedua, begitu seterusnya hingga didapat nilai yang konvergen.

Persamaan umum dari model Kalman Filter sesuai dengan Hamilton (1994) jika diberikan y_t yang dinotasikan sebagai vektor berukuran (n x 1) variabel terobservasi pada waktu t, model dinamis untuk y_t yang mungkin tidak teramati dapat dijelaskan pada vektor ξ_{t+1} yang berukuran (r x 1). Representasi ruang keadaan untuk y diberikan pada persamaan berikut :

Persamaan keadaan : $\xi_{t+1} = F\xi_t + v_{t+1}$

Persamaan observasi : $y_t = H'\xi_t + w_t$

dimana :

ξ_t : vektor keadaan berukuran (r x 1).

F : matriks parameter berukuran (r x r) pada persamaan keadaan.

H' : matriks parameter berukuran (n x r) pada persamaan observasi.

v : vektor *noise* berukuran (r x 1) pada persamaan keadaan.

w : vektor *noise* berukuran (n x 1) pada persamaan observasi.

y : vektor dari variabel terobservasi berukuran (n x 1).

Q : matriks varian kovarian berukuran (r x r) dari *noise* persamaan keadaan

R : matriks varian kovarian berukuran (n x n) dari *noise* persamaan observasi.

r : orde ARIMA dari data pengamatan.

n : jumlah variabel bebas pada data pengamatan

Vektor v_t dan w_t harus mengikuti asumsi *white noise* dengan :

$$E(v_t\ v'_\tau) = \begin{cases} Q, & \text{untuk } t = \tau \\ 0, & \text{lainnya} \end{cases}$$

dimana Q merupakan matriks berukuran (r x r).

$$E(w_t\ w'_\tau) = \begin{cases} R, & \text{untuk } t = \tau \\ 0, & \text{lainnya} \end{cases}$$

dimana R adalah matriks berukuran (n x n).

ISBN: 978-602-14387-0-1

Vektor *noise* v_t dan w_t diasumsikan tidak berkorelasi dengan semua lag sehingga untuk semua t dan τ nilai $E(v_t\ w'_\tau) = 0$.

3. Metodologi Penelitian

Data yang digunakan berupa data sekunder yang diambil dari Badan Metereologi Klimatologi dan Geofisika (BMKG) Stasiun Klimatologi Semarang mengenai data curah hujan perbulan di kota Semarang dari tahun 2005-2012 yang digunakan sebagai variabel penelitian.

Langkah analisis data yang akan dilakukan adalah :

1. Melakukan analisis stasioneritas pada data curah hujan dari tahun 2005-2012 dengan membuat plot runtun waktu, ACF, dan PACF.
2. Membentuk model ARIMA terbaik pada data curah hujan meliputi:
   a. Identifikasi model berdasarkan plot ACF dan PACF.
   b. Melakukan uji signifikansi parameter dan uji diagnostik meliputi normalitas residual dan uji asumsi independensi residual.
   c. Melakukan verifikasi model.
3. Membuat model ruang keadaan Kalman Filter meliputi :
   a. Membentuk sebuah fungsi pada ruang keadaan yang mewakili proses ARIMA terbaik dari data curah hujan Kota Semarang.
   b. Mengestimasi nilai awal parameter dari fungsi yang telah dibentuk.
   c. Melakukan pemfilteran dari data curah hujan dengan output parameter dari hasil fungsi yang telah diestimasi.
   d. Menghitung nilai residual, membuat plot hasil filtering dan normalitas residual.
   e. Melakukan prediksi curah hujan untuk tahun 2013 dari data yang telah difilter kemudian mencari batas atas dan bawah hasil prediksi.

4. Hasil dan Pembahasan

4.1 Statistik Deskriptif Data Curah Hujan Kota Semarang

Data yang dianalisis dalam penelitian tugas akhir ini adalah data curah hujan di kota Semarang berupa data bulanan mulai dari Januari 2005 hingga Desember 2011 dengan deskripsi sebagai berikut.

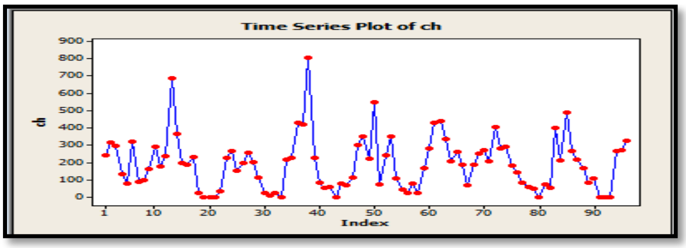
ISBN: 978-602-14387-0-1

Tabel 1. Statistik Deskriptif Jumlah Curah Hujan Kota Semarang Januari 2005 – Desember 2012

Tahun	Minimum	Maksimum	Mean	Standar Deviasi
2005	81	324	205,33	91,83
2006	0	689	186,33	201,75
2007	1	689	155,50	127,44
2008	2	806	214,67	229,73
2009	25	552	182,58	158,80
2010	70	443	272,50	112,94
2011	0	402	153,67	123,16
2012	0	490	184,33	151,34

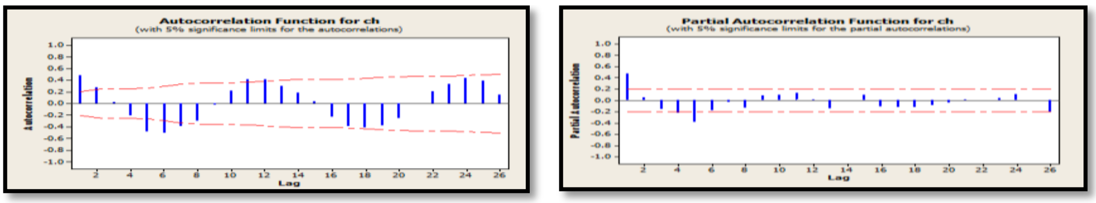
4.2 Analisis Runtun Waktu Data Curah Hujan

Langkah awal dalam analisis runtun waktu adalah membuat plot data yang terdiri dari plot runtun waktu, plot ACF dan plot PACF data curah hujan Kota Semarang dari Januari tahun 2005 – Desember 2011. Berikut adalah plot plot runtun waktu, plot ACF dan plot PACF data curah hujan Kota Semarang dari Januari tahun 2005 – Desember 2012:



Gambar 1. Plot Runtun Waktu Data Curah Hujan Kota Semarang Januari 2005 – Desember 2012

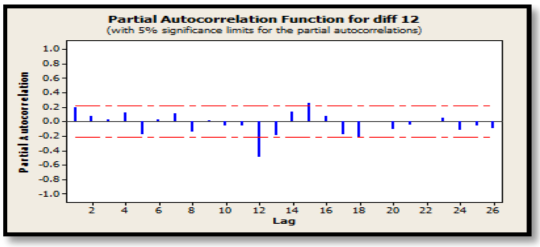
Gambar 1 menunjukkan data curah hujan secara visual berfluktuasi tajam dan terdapat perubahan nilai di *mean*. Plot ACF pada Gambar 2 memperlihatkan adanya pola musiman yang berulang pada kelipatan 12. Hal ini menyebabkan data curah hujan harus didifferensi pada lag ke 12. Plot ACF dan PACF setelah differensi di lag 12 adalah berikut :

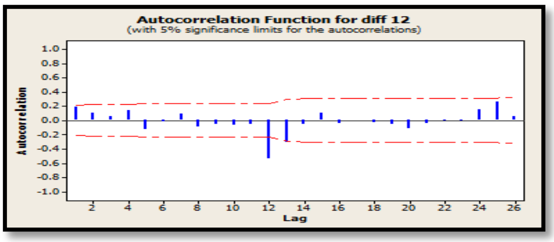


Gambar 2. Plot ACF dan PACF Data Curah Hujan Kota Semarang

ISBN: 978-602-14387-0-1

Januari 2005 – Desember 2012





Gambar 3. Plot ACF dan PACF Data Curah Hujan Kota Semarang Januari 2005 – Desember 2012 Setelah Differensi Lag 12

Berdasarkan Gambar 3 plot ACF dan PACF menunjukkan data curah hujan telah stasioner setelah differensi di lag 12. Langkah selanjutnya adalah pemilihan model ARIMA terbaik seperti terlihat pada Tabel 2.

Tabel 2. Estimasi dan Uji Signifikansi Parameter Model

Model	Parameter	T	P-*Value*	Keputusan
ARIMA $(0,1,1)^{12}$	Θ_1 = 0,8781	9,71	0,0000	Signifikan
ARIMA (0,0,1) $(0,1,1)^{12}$	θ_1 = -0,0631	-0,58	0,566	Tidak Signifikan
	Φ_1 = 0,8778	9,49	0,0000	Signifikan
ARIMA (1,0,1) $(0,1,1)^{12}$	ϕ_1 = -0,6083	-0,21	0,833	Tidak Signifikan
	θ_1 = -0,6246	-0,22	0,825	Tidak Signifikan
	Θ_1 = 0,9055	10,15	0,0000	Signifikan
ARIMA (0,0,1) $(1,1,1)^{12}$	Φ_1 = -0,3960	-3,29	0,001	Signifikan
	θ_1 = -0,0007	-0,01	0,995	Tidak Signifikan
	Θ_1 = 0,8676	9,31	0,0000	Signifikan

Hasil estimasi dan pengujian signifikansi parameter data curah hujan pada Tabel 2 menunjukkan ARIMA (0,0,0)$(0,1,1)^{12}$ merupakan model yang seluruh parameternya signifikan. Model tersebut memiliki juga memenuhi asumsi normalitas karena memiliki nilai *p-value* lebih besar dari α dan memenuhi uji independensi residual sehingga ARIMA (0,0,0)$(0,1,1)^{12}$ merupakan model terbaik dan memiliki bentuk

$$Z_t = a_t - 0{,}8781\ a_{t-12}$$

dengan nilai *Mean Square Error* (MSE) 12082.

ISBN: 978-602-14387-0-1

4.3 Kalman Filter

Persamaan Kalman Filter dibuat berdasarkan model terbaik ARIMA $(0,0,0)(0,1,1)^{12}$ yang direpresentasikan dalam bentuk ruang keadaan (*state space*) sehingga didapat persamaan Kalman Filter sebagai berikut

Persamaan keadaan

$$\xi_{t+1} = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xi_t + \begin{bmatrix} \varepsilon_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Persamaan observasi

$$y_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \xi_t + w_t$$

Hasil prediksi curah hujan kota Semarang tahun 2013 dengan Kalman Filter diberikan pada Tabel 3. Hasil prediksi curah hujan kota Semarang di tahun 2013 menunjukkan bahwa Kota Semarang memiliki curah hujan maksimum pada bulan Februari dengan jumlah curah hujan sebesar 406 mm sedangkan jumlah curah hujan minimum terjadi pada bulan Juli yakni sebesar 37 mm. Rata-rata curah hujan di tahun 2013 adalah sebesar 196,25mm.

Tabel 3. Hasil Prediksi Curah Hujan Kota Semarang Tahun 2013

Bulan	Peramalan (mm)	Batas Bawah	Batas Atas
JANUARI	385	161	609
FEBRUARI	406	184	629
MARET	224	2	447
APRIL	172	0	394
MEI	159	0	381
JUNI	113	0	335
JULI	37	0	257
AGUSTUS	53	0	276
SEPTEMBER	81	0	303
OKTOBER	161	0	383
NOVEMBER	249	26	471
DESEMBER	315	92	537

5. Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan, maka dapat disimpulkan bahwa model ARIMA terbaik untuk data curah hujan di kota Semarang periode Januari 2005 sampai dengan Desember 2012 digunakan untuk pembuatan model Kalman Filter terbaik adalah ARIMA (0,0,0) $(0,1,1)^{12}$.

Hasil prediksi curah hujan kota Semarang di tahun 2013 menunjukkan bahwa Kota Semarang memiliki curah hujan tertinggi pada bulan Februari dengan jumlah curah hujan sebesar 385 mm sedangkan jumlah curah hujan terendah terjadi pada bulan Juli yakni sebesar 37 mm. Rata-rata curah hujan di tahun 2013 adalah sebesar 196,25 mm.

DAFTAR PUSTAKA

Bain, L.J and Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics*. Second Edition. Duxbury Press. California.

Brockwell,P.J. and Davis, R.A. 1991. *Time Series :Theory and Methods*.Second Edition. Springer-Verlag, Inc. New York.

Conover, W.J. 1980. *Practical Nonparametric Statistics*. Second Edition. John Wiley. New York.

Dewantara, B.S. 2012. *Peramalan Cuaca Tradisional*. http://www.bagassdsite.blogspot.com/2012/10/peramalan-cuaca-tradisional.html [25 Oktober 2012]

Hamilton, J.D. 1994. *Time Series Analysis*. Princeton University Press. New Jersey.

Makridakis, S. *et.al*. 1999. *Metode dan Aplikasi Peramalan*. Jilid 1 Ed ke-2. Suminto,H., penerjemah. Bina Rupa Aksara. Jakarta.

Meinhold, R.J. and Singpurwala, N. D. 1983. Understanding The Kalman Filter. *The American Statistician.* Volume 37 No. 2 : 123 – 127. American Statistical Association.

Petris, G. 2010. An R Package for Dynamic Linear Models. *Journal of Statistical Software*. Volume 36 Issue 12. American Statistical Association. USA.

Petris,G. and Petrone, S. 2011. State Space Models in R. *Journal of Statistical Software*. Volume 41 Issue 4. American Statistical Association. USA.

Soejoeti, Z. 1987. *Materi Pokok Analisis Runtun Waktu.* Cetakan Pertama. Karunika. Jakarta.

Tresnawati, R. dan Komalasari, K.E. 2011. Skenario Tenggang Waktu SST Nino 3.4 Terhadap Curah Hujan untuk Meningkatkan Akurasi Prediksi Kalman Filter. *Jurnal Meteorologi dan Geofisika*. Volume 12 No. 3 : 243 – 251. Puslitbang BMKG. Jakarta.

Tresnawati, R., Nuraini, T.A, dan Hanggoro, W. 2010. Prediksi Curah Hujan Bulanan Menggunakan Metode Kalman Filter dengan Prediktor SST Nino 3.4 Diprediksi. *Jurnal Meteorologi dan Geofisika*. Volume 11 No. 2 : 106 – 115. Puslitbang BMKG. Jakarta.

Wei, W.W.S. 2006. *Time Series Analysis Univariate and Multivariate Methods*. Second Edition. Pearson Education, Inc. US.

Welch, G. and Bishop, G. 2001. *An Introduction to the Kalman Filter*. ACM In