Detecting discontinuities in time series of upper air data: Demonstration of an adaptive filter technique

I. Zurbenko¹
P.S. Porter²
J. Chen¹
S.T. Rao¹
J.Y. Ku³
R. Gui²

DISSEMINATION OF THIS DOCUMENT IS UNLIMITED

¹ State University of New York at Albany
Albany, New York 12222

² University of Idaho
Idaho Falls, Idaho 83402

³ New York State Department of Environmental Conservation
Albany, New York 12233

MASTER

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

**Portions of this document may be illegible
in electronic image products. Images are
produced from the best available original
document.**

ABSTRACT

The issue of global climate change due to increased anthropogenic emissions of greenhouse gases in the atmosphere has gained considerable attention and importance. Climate change studies require the interpretation of weather data collected in numerous locations and/or over the span of several decades. Unfortunately, these data contain biases caused by changes in instruments and data acquisition procedures. It is essential that biases are identified and/or removed before these data can be used confidently in the context of climate change research.

The purpose of this paper is to illustrate the use of an adaptive moving average filter and compare it with traditional parametric methods. The advantage of the adaptive filter over traditional parametric methods is that it is less effected by seasonal patterns and trends. The filter has been applied to upper air relative humidity and temperature data. Applied to generated data, the filter has a root mean squared error accuracy of about 600 days when locating changes of 0.1 standard deviations and about 20 days for changes of 0.5 standard deviations. In some circumstances, the accuracy of location estimation can be improved through parametric techniques used in conjunction with the adaptive filter.

I. Introduction

In recent years the issue of global climate change due to increased anthropogenic emissions of greenhouse gases in the atmosphere has gained considerable attention and importance. Recognizing the need for a long-term data base to address global climate changes, the National Climatic Data Center (NCDC) has embarked on a project to create an upper air data base consisting of radiosondes, pibals, surface reports, and station histories for the northern and southern hemispheres. Unfortunately, these data contain biases caused by changes in instruments and data acquisition procedures (Elliott and Gaffen, 1991; Schwartz and Doswell, 1991; Garand, et al. 1992). It is essential that biases are identified and/or removed before these data can be used confidently in the context of greenhouse-gas induced climate modification.

The primary object of this investigation is to develop and apply purely statistical methods of detection and adjustment for biases in the upper air temperature and humidity data. A purely statistical approach is needed because detailed knowledge of radiosonde instrumentation, data reduction methods and changes in data acquisition procedures are often unavailable. Even when there are good station histories, purely statistical methods can provide independent confirmation of and information about suspected biases.

In addressing this problem, it is presumed that changes in data due to instrument change are characterized by their abrupt and persistent nature in contrast to changes due to long term climate (slow but persistent modification of data characteristics) or due to short term weather systems (abrupt but ephemeral modification of data characteristics).

Detection of abrupt changes due to instrumentation or data reduction is complicated by the nature of daily temperature and humidity data. This data can be characterized as high frequency noise combined with low frequency seasonal and trend components. The seasonal components vary from year to year, limiting the ability of traditional parametric techniques, which rely on their removal. In this paper we describe a procedure which locates suspected discontinuities using a low-pass filter (KZA) and compare it with better known parametric techniques.

The KZA filter is a modification of the low pass KZ filter (Zurbenko, 1986, Rao and Zurbenko, 1994) designed to remove short term fluctuations attributable to seasonal and short term climate change, but which tends to smooth abrupt discontinuities, making estimation of their time of introduction difficult. The KZA filter is an adaptive form of the KZ filter which more precisely defines abrupt discontinuities in time series.

The KZ and KZA filters have the spectral characteristics of a rectangular filter, but do not require the latter's complicated parameter estimation. As such, the KZ and KZA filters remove some frequencies leaving others untouched, permitting physical scientists to apply their expertise and intuition to filtered data. In other words, data filtered by the KZ or KZA reflect physical phenomena, unlike data treated by other techniques, which may remove unwanted components of the data but distort phenomena of interest.

Parametric methods used include those of Alexandersson (Hawkins, 1977; Alexandersson, 1986) and Schwarz' criterion (Schwarz, 1978). These methods assume normally distributed, independent data and are suitable when applied to seasonally adjusted monthly averages of upper air data having relatively large, not too closely spaced, breaks. As the number of breaks increase these methods lose power, in part because of the difficulty of seasonal adjustment in the presence of large numbers of discontinuities.

We describe and demonstrate procedures using synthetic data with known characteristics, and actual data from Oakland, California and Hilo, Hawaii, and Hong Kong, stations with well known histories. Simulations with synthetic data provide operating characteristics of statistical methods under controlled conditions. Usually, the performance of a method using normally distributed independent data will reflect its most optimistic expectation. Actual data are needed to design appropriate conditions for simulations and, in the case of stations with well documented histories, may also provide some insight into method capability.

II. Methods

A. Parametric methods

Simulation studies by Easterling and Peterson (1992) suggested that the method of Alexandersson's (AM) (Hawkins, 1977; Alexandersson, 1986) was the most effective of several techniques for locating single discontinuities in time series of independent, normally distributed random variables.

AM is the maximum likelihood estimator for a single change point in a time series of independent, normally distributed observations. It considers the null hypothesis that there are no discontinuities against the alternative that there is one. The power of AM method applied to normal, independent data is described elsewhere (Porter et.al 1994). The test statistic for AM is T_0 , computed as follows:

$$T_0 = \max_{5 \leq r \leq n-5} (T_r) = \max_{5 \leq r \leq n-5} [r\bar{X}_1^2 + (n-r)\bar{X}_2^2] \quad (1)$$

where n is the sample size, r is the break location, $\bar{X}_1$ and $\bar{X}_2$ are the sample means occurring before and after the break, respectively, and X represents data standardized to 0 mean and unit variance. Under H_0 , T_r is distributed as χ^2 . The value of T_0 is compared with the appropriate critical value. Critical values depend on the minimum sample size that is considered acceptable, in this study, 5.

The original single break point (BP) AM performs well for single breaks of relatively large size (say 0.5 standard deviations of a process) but poorly when multiple breaks are present (Porter, et.al, 1994). When two consecutive breaks occur in the same direction, they may be interpreted as one large break, distorting the break time estimate. Two breaks occurring in opposite direction tend to be averaged, avoiding detection.

A modification of AM that looks for multiple breaks performs better than single BP methods when multiple breaks are actually present. However, the single BP AM is most powerful when there is actually a single BP.

The principle problem with AM is that it requires independent, normally distributed observations, necessitating adjustment for trend and seasonality. Daily observations of seasonally adjusted upper air data are serially correlated, reducing the effective

sample size, and changing the test's critical values. If critical values are not adjusted for serial correlation, too many false positives result.

The Schwarz' criterion (SC) (Schwarz, 1978; Yao, 1988) was also used for multiple break point problems. SC is given by:

$$SC(R') = \frac{n}{2} \ln \sigma_{R'}^2 + R' \ln n \quad (4)$$

where R' is the number of break points minimizing $SC(R')$, n is sample size, and $\sigma_{R'}^2$ is the variance of the residuals created by subtracting the $R'+1$ means. $SC(R)$ is computed for all possible values of R and break locations, the minimum being $SC(R')$. (SC is closely related to the Aikaike information criterion).

B. Seasonal adjustment

When parametric methods were applied to actual data or synthetic data having seasonal patterns, the original series of daily observations data were first averaged by month and deseasonalized using the program PEST (Brockwell and Davis, 1991). PEST uses a 12 point moving average to compute the seasonal component, which is then subtracted from each datum. PEST assumes that the seasonal component is invariant, meaning that there is a single value for the January seasonal component that is subtracted from all January averages. The invariance assumption is approximate since the January seasonal effect changes from year to year.

C. Adaptive moving average filter

Since a long term moving average which filters out seasonal variation and high frequency noise dampens sharp breaks in time series data, an adaptive filter was developed that dynamically adjusts the length of the moving average according to the rate of change of the process. As the rate of change increases, the length of the filter changes.

The adaptive filter must be preceded by a simple moving average filter (KZ) (Zurbenko, 1986) run iteratively until a series that is absent of high frequency noise is obtained. The simple moving average is computed as follows:

$$y_i = \frac{1}{2q+1} \sum_{j=-q}^q x_{i+j} \quad (5)$$

where $2q+1$ is the filter window. The moving average process is iterative in that the process is repeated, with y_i replacing x_i in equation (5). Termed the "KZ" filter, this iterative simple moving average is an efficient low pass linear filter discussed in detail in Zurbenko (1986).

The adaptive filter is then accomplished as follows:

- The original time series, $X(t)$, is subjected to the $KZ_{q,p}$ filter, where q is the half length of the simple moving average and p is the number of iterations of the same filter.

$$Z(t) = KZ_{q,p}[X(t)] \quad (6)$$

- The absolute values ($D(t)$) of the differenced $Z(t)$ are defined:

$$D(t) = |Z(t+q) - Z(t-q)| \quad (7)$$

- The slope of $D(t)$ is defined:

$$D'(t) = D(t+1) - D(t) \quad (8)$$

When a data point is located on an upslope area of $D(t)$ the half length of the moving average before the data point (tail, q_T) remains the same q as in the original filter, while the length ahead of the data point (head, q_H), is shortened as a function of $D(t)$. In a downslope area (with respect to $D(t)$), only the half length behind the data point (head) will be reduced. The reduced length in the vicinity of a break point "sharpens" the moving average.

The adaptive filter is then defined by :

$$Y_t = \frac{1}{q(t)_H + q(t)_T} \sum_{i=-q_T(t)}^{q_H(t)} X_{t+i} \quad (9)$$

where,

$$q_H(t) \begin{cases} q, & \text{if } D'(t) < 0 \\ f(D(t)) \cdot q, & \text{if } D'(t) \geq 0 \end{cases}$$

$$q_T(t) \begin{cases} q, & \text{if } D'(t) > 0 \\ f(D(t)) \cdot q, & \text{if } D'(t) \leq 0 \end{cases}$$

q is the half length in the initial $KZ_{p,q}$ filter and $f(D(t))$ corresponds to:

$$f(D(t)) = 1 - \frac{D(t)}{\max[D(t)]} \quad (10)$$

Plots of the filtered data (Y_t) and the first differences of Y_t provide visual evidence of discontinuities in time series.

Probabilistic statements about suspected discontinuities are made by comparing the filter relative standard deviation (RSD) with the standard deviations occurring near suspected breaks. Values of the RSD much larger than 1 indicate abrupt changes in the mean level.

The adaptive filter alone, the adaptive filter used in combination with AM, and the Schwarz' criterion were evaluated under various conditions described below. Previously reported results (Porter et al. 1994) indicated that AM applied to data having a single break point had better accuracy than the KZA, but that KZA was more powerful, suggesting that in combination the two techniques might reflect the strengths of both methods and overcome their weaknesses. In combining the KZA and AM, the AM was applied to a window of size $4 \cdot \text{RMSE}$ centered on the location identified by the KZA, where RMSE refers to the root mean squared error of the break location estimate. In the case of seasonal or actual data, when the window size was larger than 60 days, AM was applied to seasonally adjusted data. When windows were 60 days or smaller, AM was applied to raw data. For such small windows, the effect of seasonality or trends is quite limited.

D. Simulation studies

Simulations were carried out using standard normal random numbers with breaks inserted. Break sizes ranged from 0.1 to 1.0 standard deviation of the random numbers, denoted " σ ". Simulation studies were also conducted with data resembling relative humidity data by summing three sin functions together with white noise. The period of the function was 360 days. The cycle did not repeat exactly from year to year in order to simulate the variable seasonal effects observed in actual data. The white noise represented 50% of the total variability of the seasonal data, and break sizes ranged from 0.1 to 1.0 standard deviations of the white noise (Fig. 1).

Unless otherwise noted, for each combination of break pattern, break size, noise level, ten repetitions were carried out. This number is quite adequate for comparing methods and for indicating whether a procedure has merit, the principle intent of this paper.

D. Data Base

The actual data used in the study included relative humidity data collected at 1000mb for Hilo, HI and Oakland, CA, and temperature data collected at 850mb at Hong Kong. Data collected at 0:00 h Greenwich time were used. The Hilo series runs from 1950 to 1990, the Oakland series runs from 1948 to 1990, and the Hong Kong data from 1956 to 1990. All of these stations have relatively well known station histories.

The Oakland 1000 mb data has a total variance of $18.0 (\%RH)^2$ and a noise variance of $10.3 (\%RH)^2$. The same figures for the Hilo data are $11.2 (\%RH)^2$ and $5.8 (\%RH)^2$, while for Hong Kong, the values are 4.0 and $0.7 ^\circ C^2$, meaning that high frequency noise accounts for 57%, 52%, and 18% of total variability in the Oakland, Hilo, and Hong Kong data, respectively. On visual observation, seasonal variation is much stronger in the Hong Kong temperature data than in the relative humidity data of Oakland and Hilo.

III. Results and Discussion

A. Independent, normally distributed random numbers

The accuracy of break location estimation depends on break size and pattern. For the KZA, breaks are located more accurately when consecutive breaks occur in opposite direction than when they occur in the same direction. For the former pattern, the KZA locates discontinuities with an accuracy of about 190 for discontinuities of 0.1σ to about 70 for discontinuities of 0.5σ (root mean squared error of break location) (Table 1). For the latter pattern, the range is 600 to 180 RMSE. When breaks are 0.1σ or smaller, the KZA detects them, but the accuracy is poor and the number of false positives increases. Figure 2 is a typical result with KZA, showing also the result obtained with the original KZ filter.

Table 1. Accuracy of break location estimation, applied to normal, independent, random numbers. Values are root mean squared error of the break location estimate.

break pattern	sample size	break size	RMSE		
			filter alone	filter + AM	SC alone
two breaks opposite direction	3000	0.1	190	*	**
		0.2	140	*	90
		0.5	70	70	20
two breaks same direction	3000	0.1	600	*	***
		0.2	385	*	110
		0.5	180	30	30

* filter error too large

** break not detected

*** large numbers of false positives

Break pattern does not much affect the SC method and for the case of normal random numbers, estimates breaks more accurately than the KZA when breaks are 0.2σ or larger. When breaks are on the order of 0.1σ , the SC method produces a large number of false positives.

For normal random numbers (white noise), it is not useful to use parametric methods in conjunction with the filter. For small

breaks, the window identified by the KZA and searched by AM is quite large (estimated location ± 2 RMSE). When break sizes reach 0.5σ , the SC method alone is better than the KZA.

B. Random numbers with a linear trend

The presence of a trend does not significantly impair the performance of the KZA. The RMSE for data having three breaks in the same direction and a trend magnitude of 0.0005% was 700, 650, 400, and 20 for break sizes of 0.1, 0.2, 0.5, and 1.0σ , respectively. Figure 3 shows typical results for the KZA. For the data of Figure 3, SC estimated break locations midway between actual breaks. The trend cannot be removed from this data using parametric methods because for the trend is negated by the breaks. First differences would remove the trend, but would reduce the amount of information available about the break.

C. Seasonal patterns

Seasonal patterns decrease the accuracy of estimation of discontinuity location and, as with white noise, the accuracy of break location depends on break pattern and size. Again, for the KZA, breaks are located more accurately when consecutive breaks occur in opposite direction than when they occur in the same direction (Table 2). A typical result for the KZA is shown in Fig. 4.

Table 2. Accuracy of break location estimation, applied to simulated seasonal pattern. Values given are root mean squared error of break location estimate.

break pattern	sample size	break size	RMSE		
			filter alone	filter + AM	SC alone
single break in center	5400	0.1	525	*	**
		0.2	240	160	90
		0.5	125	30	40
two breaks in opposite direction	5400	0.1	230	*	**
		0.2	130	160	25
		0.5	70	10	20
two breaks same direction	5400	0.1	270	*	300
		0.2	220	*	230
		0.5	110	40	55

* filter error too large

** break not detected

When break sizes are small (0.1σ), window sizes are quite large (± 2 RMSE, or about 2000 days), and the KZA used with AM is not more accurate than KZA alone. At break sizes larger than 0.2σ , the accuracy of KZ used with AM is better than KZA alone (Table 2).

Data were also tested which had seasonal patterns and multiple breaks occurring in opposite directions. The KZA is able to detect multiple breaks where SC cannot. Figure 5 is a typical result for such data, showing seasonal data having breaks about every 400 days and the KZA result. SC was unable to detect any of the breaks in data with this break pattern because the multiple breaks disrupt attempts at seasonal adjustment.

D. Hilo and Oakland relative humidity data

Hilo relative humidity, t00, 1000 mb:

Breaks found by the KZA in the Hilo data are summarized in Table 3. The raw and filtered data are shown in Figs. 6a and 6b. Break times were also estimated by AM applied to a window

identified by the KZA method, and by SC applied to monthly, seasonally adjusted averages. For each of the five breaks identified by the KZA, there is a possible explanation in the station history.

The Schwarz criterion (SC) located a break in 1987 that was not considered significant using the KZA. The SC method provides a single result for the number of breaks and their locations. The KZA, on the other hand, provides probability levels and one can rank the breaks in order of probability level. Only those breaks considered significant at the 95% probability level are listed in the table under the KZ method.

Oakland relative humidity, t00, 1000 mb:

The KZA of the Oakland data indicates the breaks listed in Table 4 (Figs. 7a and 7b). Station histories may explain the 3 largest breaks. SC detected a break in addition to those of KZA. As with the Hilo data, the additional SC break is evident in the filtered data, but not considered significant (Figs. 7a and 7b).

Hong Kong Temperature, t00, 850 mb:

[AS OF THE TIME OF PREPARATION OF THIS DRAFT, THE STATION HISTORY FOR HONG KONG HAD NOT BEEN REVIEWED BY THE AUTHORS.]

The KZA estimated 5 breaks in the Hong Kong data ranging in size from 0.7 to 2.8 °C. SC alone did not find the break of size 0.7 °C in 1985 (Table 5 and Figs. 8a and 8b).

Table 3. Estimated break locations and sizes for Hilo, Hawaii, relative humidity at 1000 mb and t00

date			break size		station history
filter alone	filter + AM	SC alone	relative humidity (%)	standard deviations of noise	
57/02/2 8		56/11			a
65/01/2 8	65/03/3 0	64/12	3.81	0.66	b
69/06/0 9	69/01/2 9	69/03	1.08	0.19	c
72/04/2 6	72/07/0 6	72/02	10.73	1.87	d
81/10/0 8	81/07/2 0	81/12	1.28	0.22	e
		87/03			f

(a) Station move August, 1957

(b) Carbon hygistor was introduced on February 27, 1965

(c) Computerization of data reduction took place in 1969

(d) A new sensor was put into operation on May 2, 1972

(e) A modified carbon hygistor was introduced

(f) No station history supporting this break

Fig.3 Measure of errors for some approximations of vapor pressure over water with respect to Wexler (1976) formulations.

Table 4. Estimated break locations and sizes for Oakland, California, relative humidity at 1000 mb and t00

date			break size		station history
filter alone	filter + AM	Schwarz criteria	relative humidity (%)	standard deviations of noise	
		56/06			a
63/07/0 3	63/03/23		6.38	0.62	b
69/09/2 .1	69/09/22	70/02	3.97	0.38	c
72/04/1 9	72/08/09	72/06	10.84	1.05	d
77/03/2 4	76/11/20		2.13	0.21	e
79/07/1 0	79/06/08		1.14	0.11	e

(a) New wind equipment October, 1954

(b) Carbon hygistor was introduced on April 3, 1963

(c) A quick fix: manual to computerized data reduction

(d) A new sensor was introduced on May 12, 1972

(e) No station histories to support these changes

Table 5. Estimated break locations and sizes for Hong Kong temperature data at 850 mb and t00

date			break size		station history
filter alone	filter + AM	Schwarz criteria	temperature °C	standard deviation of noise	
66/03/2 1	65/12/25	66/11	0.7	1.0	+
67/11/1 6	66/10/16	67/04	0.8	1.1	+-
72/8/30	72/09/21	72/11	0.5	0.7	+
73/12/2 2	74/01/01	73/4	2.8	2.8	+

85/12/3 1	86/02/05		0.7	1.0	+
--------------	----------	--	-----	-----	---

+ Station history not reviewed at time of preparation

IV. Summary

The KZA is a non-parametric approach to locating discontinuities in time series of upper air data that is capable of detecting breaks of size 0.1σ in the presence of seasonal patterns and trends. When breaks are larger than about 0.2σ of process noise, parametric methods may more precisely locate breaks. Multiple breaks prevent meaningful seasonal adjustment, hampering the use of parametric methods. Even when seasonal adjustment is effective, parametric methods lose power as the number of breaks increase.

More accurate location of relatively small breaks may be possible by using the KZA to approximately locate them, followed by a parametric method like AM. AM is not much affected by seasonality when applied to windows of raw data smaller than about 60 days. In this application, it would be assumed that a break existed, and AM would be used to search for its precise location. Detecting and precisely locating breaks smaller than 0.1σ may be possible with a regional filter that utilizes information from neighboring stations and several elevations at the same time.

References

- Alexandersson, H. 1986. A homogeneity test applied to precipitation data. *J. Climatology* 6: 661-675
- Brockwell, P. and R. Davis. 1991. *Time Series Analysis*. Springer-Verlaag
- Easterling, D.R. and Peterson, T.C. 1992. Techniques for detecting and adjusting for artificial discontinuities in climatological time series: A review. 12th Conference on Probability and Statistics in the Atmospheric Sciences. June 22-26. Toronto, Ontario, Canada
- Elliott W.P. and Gaffen D.J., 1991: On the utility of radiosonde humidity archives for climate studies. *Bull. Amer. Meteor. Soc.*, 72, 1507-1520.
- Garand, L., Grassotti, C. Halle, J., and G.L. Klein. 1992. On Differences in Radiosonde Humidity-Reporting Practices and Their Implications for Numerical Weather Prediction and Remote Sensing. *Bulletin of the American Meteorological Society* 73: 1417-1423
- Hawkins, D.M. 1977. Testing a sequence of observations for a shift in location. *JASA* 72: 180-186
- Hinkley, D.V. 1970. Inference about the change-point in a sequence of random variables. *Biometrika* 57: 1-17
- Porter, P.S., Rao, S.T., Zurbenko, I., Chen, J., and Ku, J.Y. 1993. Statistical methods for detecting discontinuities in time series of upper air data. Quarterly report, NOAA contract #50EANE2000078
- Rao, S.T. and Zurbenko, I. 1994. Detecting and tracking changes in ozone air quality. *AWMA Journal*, accepted
- Schwarz, G. (1978). Estimating the dimension of a model, *Ann. Statist.* 6: 461-464
- Schwarz, B.E. and Doswell, C.A. 1991. North American Rawinsonde Observations: Problems, concerns, and a call to action. *Bulletin of the American Meteorological Society* 72: 1885-1896
- Yao, Y.C. 1988. Estimating the number of change-points via Schwarz' criterion. *Statistics and Probability Letters* 6: 181-189
- Zurbenko I.G., 1986: *The spectral analysis of time series*. North-Holland.

Figure titles

Figure 1. Synthetic seasonal data (a) Sum of sin curves (seasonal pattern), (b) normal noise (0,1), (c) sum of noise and seasonal pattern

Figure 2. KZA(100,3) result for 3,000 $N(0,1)$ random numbers with breaks of size 0.5 at 1,000, 2000, 3000 and 4000.
(A) R

Figure 3. KZA(100,3) applied to 3,000 $N(0,1)$ random numbers with breaks of size 0.6 at 1,000 and 1.2 at 2,000 and a trend of magnitude 0.0006/day. (A) Raw data, (B) data after KZ filter, (C) data after KZA filter, (D) trend line and break pattern

Figure 4. KZA(100,3) applied to 5,400 seasonal data and $N(0,1)$ random numbers with breaks of size 0.5 at 1,000, 2,000, 3,000 and 4,000 (A) Raw data, (B) data after KZ filter, (C) data after KZA filter, (D) trend line and break pattern

Figure 5. KZA(100,3) applied to 5,400 seasonal data and $N(0,1)$ random numbers with multiple breaks of size 0.5 about every 300 days (A) Raw data, (B) data after KZ filter, (C) data after KZA filter

Figure 6. KZA(365,3) for Hilo, Hawaii (a) Raw data; (b) Filtered data

Figure 7. KZA(365,3) for Oakland, California (a) Raw data; (b) Filtered data

Figure 8. KZA(365,3) for Hong Kong (a) Raw data; (b) Filtered data

A

Time Series

B

Moving Average, $p=201, q=3$

C

Adaptive Moving Average, $\alpha=0.5$

D

level 0.7

A

B

C

D

A

Time Series

B

Moving Average, $p=365, q=3$

C

Adaptive Moving Average, $\alpha=0.5$

none, t00, p = 365, q = 3

Adaptive smooth, r1 = 0.1, r2 = 1, Dt = 1, Dc = 1

none, t00, p = 365, q = 3

Adaptive smooth, r1 = 0.1, r2 = 1, Dt = 1, Dc = 1

Fig.1 Relative errors of basic formulations of vapor pressure with respect to Goff-Gratch formulation (1946).

Fig.3 Measure of errors for some approximations of vapor pressure over water with respect to Wexler (1976) formulations.