2003

Comment on "A reappraisal of the Kalman filtering technique as applied in river flow forecasting" by M. Ahsan and K. M. O'Connor

Jozsef Szilagyi
University of Nebraska-Lincoln, jszilagyi1@unl.edu

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 285 (2004) 286–289
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Discussion

Comment on "A reappraisal of the Kalman filtering technique, as applied in river flow forecasting" by Ashan, M., O'Connor, K.M., 1994. Journal of Hydrology 161, 197–226

Jozsef Szilagyi*

Conservation and Survey Division, University of Nebraska-Lincoln, Lincoln, NE 68588-0517, USA

Received 10 December 2002; revised 6 May 2003; accepted 23 May 2003

The authors reviewed the application of the Kalman filter (1960) in river flow forecasting. They correctly pointed out that application of the Kalman filter does not improve upon traditional, Box and Jenkins-type time series forecasting techniques when the stationary data (i.e. the difference between forecasts and observations) is considered perfectly free of measurement errors. It must be noted here, however, that there are no such things as error-free measurements. One can nonetheless generate such hypothetical measurements by a computer, as is done below, and in such a case the two techniques are indeed identical. The authors further asserted that, in case of measurement errors, the Kalman filter "results in reduced forecast efficiency" because during traditional optimization the measured flow values are implicitly assumed to be error-free, since model performance is inferred upon comparing the estimates to measured values, however, corrupted by measurement error, in a mean-squared-error sense (Ahsan and O'Connor, 1994). As the Kalman filter indeed assures that the estimates obtained by its application are optimal with regard to the unknown error-free signal, the above authors claim that the resulting estimates can/will not be optimal in comparison with the measured, error-laden values. The present author, however, will point out below, through theoretical considerations and through numerical experiments, that this latter claim is incorrect. In fact, the Kalman filter will always result in better estimates than traditional Box and Jenkins-type forecasts even if model performance is based on comparing model estimates with error-biased measurements.

The prerequisites of the original digital form of the Kalman filter to result in optimal estimates of the state vector are (Meditch, 1969): (a) the state vector (x) be a Gauss–Markov sequence; (b) the model (w) and measurement (v) errors be gaussian white sequences independent of each other, and; (c) the initial state vector be independent of w and v. If these prerequisites are met, then the following will also be true (Meditch, 1969): (1) $w(k \geq j)$ is independent of $x(j)$ and $z(j)$, $j = 0,1,\ldots$, where z denotes the measurement values, and; (2) $v(k)$ is independent of $x(k)$ for all k, and also independent of $z(j < k)$. Using these properties of the measurement error and knowing that the Kalman filter results in a minimum in the following model performance-test

$$J_1 = \mathrm{E}[(x(k) - \langle x(k) \rangle)^2], k = 1, 2, \ldots, \quad (1)$$

* Tel.: +402-472-9667; fax: +402-472-4608.
E-mail address: jszilagyil@unl.edu (J. Szilagyi).

doi:10.1016/S0022-1694(03)00223-3

where the sharp brackets designate a priori (i.e. forecasted) estimates of x and E denotes the expectation, it can quickly be shown that it must minimize the following traditionally employed model performance test as well

$$J_2 = \mathrm{E}[(z(k) - \langle z(k)\rangle)^2], k = 1, 2, \ldots \quad (2)$$

Since $z(k) = Hx(k) + v(k)$, where H is the measurement matrix, and $\langle z(k)\rangle = H\langle x(k)\rangle$, Eq. (2) can be written as

$$\mathrm{E}[(Hx(k) + v(k) - H\langle x(k)\rangle)^2]$$

$$= H\mathrm{E}[(x(k) - \langle x(k)\rangle)^2] + \mathrm{E}[(v(k))^2] \quad (3)$$

since the expected value of the terms that contain the product of x or $\langle x\rangle$ and v vanish due to independence, the latter because the a priori estimate, $\langle x\rangle$, of x can be expressed as a linear combination of the measurements prior to the time (k) of the measurement error (Meditch, 1969). Consequently, a minimization of Eq. (1) means the minimization of Eqs. (2) and (3), since the last term in Eq. (3) is given. It is immediately clear also, that in the absence of the measurement noise, Eqs. (1) and (2) are identical up to a multiplier of H, which is unity in case of a first-order scalar autoregressive, AR(1), sequence when written in a state-space form, so in this latter case the two performance statistics are strictly equal.

To illustrate the above claim, stationary scalar AR(1) sequence values were generated by the computer

$$x(k + 1) = \Phi x(k) + w(k), \qquad k = 0, 1, \ldots \quad (4)$$

where Φ is the prescribed AR(1) sequence parameter. The $x(k)$ values were disturbed to simulate the effect of the measurement process such as

$$z(k) = x(k) + v(k), \qquad k = 1, 2, \ldots \quad (5)$$

with w and v being gaussian white sequences with zero means and prescribed variances.

The parameter Φ was then estimated by the Yule-Walker equations (Bras and Rodriguez-Iturbe, 1993)

$$\rho_1 = \langle\Phi\rangle\rho_0 = \langle\Phi\rangle \quad (6)$$

where $\rho_0 = 1$, and ρ_I are the lag-zero and lag-one sample autocorrelation coefficients, respectively. With the help of the estimated Φ value, $\langle\Phi\rangle$, lag-one forecast values were created by (Bras and Rodriguez-Iturbe, 1993)

$$\langle z(k+1)\rangle = \mu + \langle\Phi\rangle(z(k) - \mu) \qquad k = 1, 2, \ldots \quad (7)$$

where μ denotes the sample mean. For long sequences ($k_{\max}$ is in the order of 10^4) the so derived estimates indeed minimize Eq. (2), meaning that no other systematically chosen trial value of Φ when used with Eq. (7) results in better model performance as calculated by Eq. (2).

To see if the Kalman filter can improve upon these estimates, systematically chosen trial values of Φ were created (simply increasing the trial value of Φ from a minimum value up to a maximum value [i.e. 1] with a prescribed increment, in the order of 10^{-4}) and applied in the filter equations as described by Meditch (1969) and Szollosi-Nagy (1989). A summary of the filter algorithm used is such

Initial values:

Q = variance of w (prescribed in the random number generator); R = variance of v (prescribed in the random number generator); P = sample variance of x; X = sample mean of x;

Start with $k = 1$:

$$X = \langle\Phi\rangle X; \langle x(k)\rangle = X; P = \Phi^2 P + Q; K = P(P + R)^{-1};$$
$$X = X + K(z(k) - X); P = (1 - K)P;$$

Increase k by one, return to start until $k < k_{\max}$.

For each trial value of Φ, Eq. (2) was evaluated with the filter calculated $\langle x(k)\rangle$ values being substituted for $\langle z(k)\rangle$, and finally the Φ value (and the corresponding Kalman filter obtained forecasts) declared as optimal that resulted in the minimum value of Eq. (2). The so obtained Φ values and lag-one forecasts or a priori estimates of $z(k)$ were always superior to the one obtained by the Box and Jenkins-type forecasts of Eq. (7). See Table 1 for a comparison. Table 1 also displays how uncertainties in the values of Q, R, and of P (since these are generally not known with real data) affect the filter estimates. From Table 1 it can be seen that the Kalman filter results in robust forecasts with stationary data which means that it is not very sensitive to errors in its initial input. As expected, its superiority over traditional time-series forecasts

Table 1
Performance statistics of the Box and Jenkins-type (BJ) and the Kalman filter (K) optimized forecasts

$k_{max} = 10{,}000$	$\langle Q\rangle = Q = 1$, $\langle R\rangle = R = 1$	$\langle Q\rangle = Q = 1$, $\langle R\rangle = R = 0.5$	$\langle Q\rangle = Q = 0.5$, $\langle R\rangle = R = 1$	$\langle Q\rangle = 2Q = 2$, $\langle R\rangle = 0.5R = 0.5$	$\langle Q\rangle = 0.5Q = 0.5$, $\langle R\rangle = 2R = 2$	$\langle Q\rangle = Q = 1$, $\langle R\rangle = R = 0$
$\Phi = 0.8$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$
$\langle\Phi\rangle^{BJ} =$	0.59	0.69	0.48	0.59	0.60	0.80
$\langle\Phi\rangle^{K} =$	0.80	0.81	0.80	0.69	0.90	0.80
$J_1^K/J_1^{BJ} =$	0.93	0.96	0.90	0.96	0.99	1
$J_2^k/J_2^{BJ} =$	0.96	0.97	0.95	0.97	0.99	1
	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$
$\langle\Phi\rangle^{BJ} =$	0.59	0.68	0.47	0.59	0.60	0.79
$\langle\Phi\rangle^{K} =$	0.80	0.81	0.80	0.69	0.90	0.79
$J_1^K/J_1^{BJ} =$	0.93	0.96	0.91	0.96	0.98	1
$J_2^k/J_2^{BJ} =$	0.96	0.97	0.95	0.98	0.99	1
	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$
$\langle\Phi\rangle^{BJ} =$	0.58	0.68	0.47	0.56	0.59	0.80
$\langle\Phi\rangle^{K} =$	0.80	0.80	0.80	0.66	0.89	0.80
$J_1^K/J_1^{BJ} =$	0.93	0.97	0.90	0.96	0.98	1
$J_2^k/J_2^{BJ} =$	0.96	0.98	0.95	0.98	0.99	1
$\Phi = 0.95$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$	$\langle P(0)\rangle = V(x)$
$\langle\Phi\rangle^{BJ} =$	0.87	0.91	0.81	0.85	0.87	0.96
$\langle\Phi\rangle^{K} =$	0.95	0.95	0.95	0.90	0.98	0.96
$J_1^K/J_1^{BJ} =$	0.85	0.92	0.75	0.90	0.93	1
$J_2^k/J_2^{BJ} =$	0.90	0.94	0.85	0.94	0.96	1
	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$	$\langle P(0)\rangle = 2V(x)$
$\langle\Phi\rangle^{BJ} =$	0.88	0.91	0.81	0.87	0.85	0.95
$\langle\Phi\rangle^{K} =$	0.95	0.95	0.95	0.92	0.97	0.95
$J_1^K/J_1^{BJ} =$	0.85	0.93	0.75	0.90	0.97	1
$J_2^k/J_2^{BJ} =$	0.91	0.95	0.86	0.93	0.99	1
	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$	$\langle P(0)\rangle = 0.5V(x)$
$\langle\Phi\rangle^{BJ} =$	0.87	0.90	0.79	0.86	0.86	0.95
$\langle\Phi\rangle^{K} =$	0.95	0.95	0.95	0.91	0.98	0.95
$J_1^K/J_1^{BJ} =$	0.85	0.93	0.75	0.90	0.91	1
$J_2^k/J_2^{BJ} =$	0.90	0.94	0.87	0.94	0.94	1

The sharp brackets denote estimates, and V designates the sample variance. Note that altogether $6 \times 6 = 36$ pieces of Gauss-Markov sequences were generated, each consisting of 10,000 values.

decreases with diminishing measurement errors, to give identical forecasts with the latter in the complete absence of measurement errors, which again, it must be stressed, is never the case with any kind of measured data.

It is believed by the present author that the Kalman filter, as has been in the past, will continue to be an important tool for the practicing hydrologist even though, as was shown above, its performance is routinely inferred (by default) from comparing the filter forecasts to error-laden measurements. Caution has to be taken though during parameter optimization in general. The relatively large difference in the optimized value of the parameter, Φ, in Table 1 between the two techniques underlines the importance of performing model optimization *together* with the Kalman filter and not separately in order to get truly optimal estimates.

References

Ahsan, M., O'Connor, K.M., 1994. A reappraisal of the Kalman filtering technique, as applied in river flow forecasting. J. Hydrol. 161, 197–226.

Bras, R.L., Rodriguez-Iturbe, I., 1993. Random Functions and Hydrology, Dover, New York.

Kalman, R.E., 1960. A new approach to linear filtering and prediction theory. Trans. ASME, J. Basic Engng. Ser. D 82, 35–45.

Meditch, J.S., 1969. Stochastic Optimal Linear Estimation and Control, McGraw-Hill, New York.

Szollosi-Nagy, A., 1989. A Mederbeli Lefolyas Real-Time Elorejelzese Dinamikus Strukturalis-Sztochasztikus Modellekkel (Real-Time Streamflow Forecasting using Dynamic Structural-Stochastic Models), VITUKI, Budapest.