DISKUSSIONSBEITRÄGE

aus dem

Fachbereich

WIRTSCHAFTSWISSENSCHAFTEN

der

UNIVERSITÄT DUISBURG - ESSEN
Campus Essen

Nr. 149
Dezember 2005

Flexible Least Squares Estimation of State Space Models: An Alternative to Kalman-Filtering?

Andreas Kladroba

Andreas Kladroba

Flexible Least Squares Estimation of State Space Models: An Alternative to Kalman-Filtering?

1. Introduction

In 1990 Kalaba/Tesfatsion developed a Flexible Least Squares (FLS) approach for estimating state space models as an alternative to Kalman Filtering. In this paper we ask whether FLS is really an alternative. For answering this we use a simulation using FLS as a regression model with time varying parameters. We will estimate this model with Kalman Filtering and two variations of FLS. In a second step we will misspecify the so-called hyperstructure of the model and we will prove how the two ways of estimating (Kalman Filtering and FLS) react to this misspecification.

2. Kalman Filtering

We assume the following state space model with the transition equation

(1) $$\beta_{t+1} = A\beta_t + \varepsilon_t \qquad t = 1,\dots,T$$

and the measurement equation

(2) $$y_t = x_t\beta_t + \eta_t \qquad t = 1,\dots,T\,.$$

Here β_t is the unknown K x 1 vector of the regression coefficient, y_t is the endogenous variable, A is a known K x K transition matrix and x_t is the 1 x K vector of the exogenous variables. Moreover it is $E(\varepsilon) = E(\eta) = 0$, $E(\varepsilon\varepsilon') = \Sigma$ and $E(\eta\eta') = \Omega$.

Estimation of this model is caused by minimizing the variance of the estimation error

(3) $$\Gamma_t = E[(\beta_t - b_t)(\beta_t - b_t)]$$

The necessary and sufficient condition of this minimizing problem is given by the well-known Wiener-Hopf equation

(4) $$E[(\beta_t - b_t)y_i] = 0 \qquad \text{for } 1 \le i \le t$$

Because Kalman Filtering is well-known we can leave out a detailed derivation and we can confine ourself on the description of the resulting algorithm (Fig. 1).

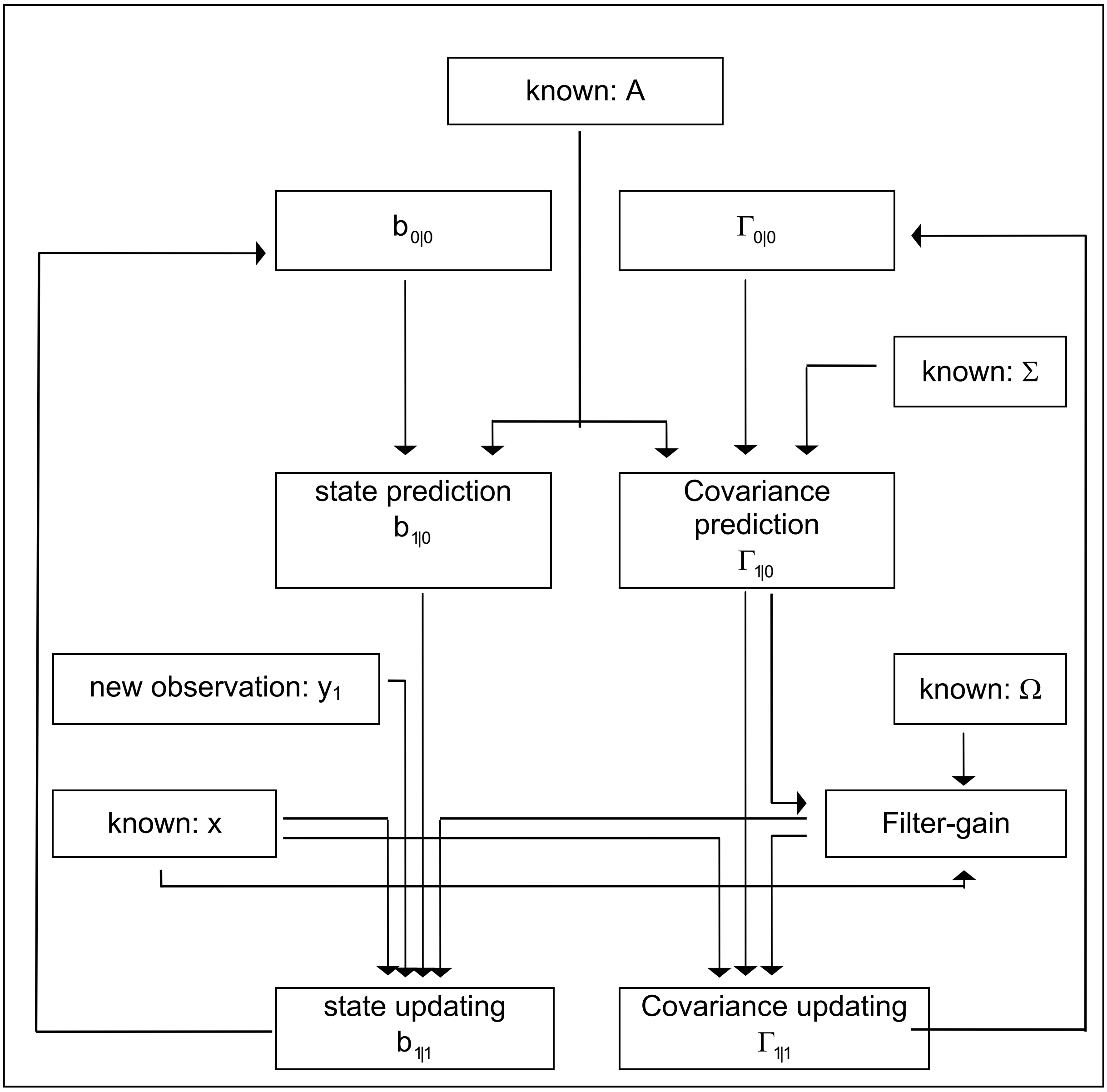

Fig. 2 Kalman-Filtering Algorithm

Proceeding from the starting values of $b_{0|0}$ and $\Gamma_{0|0}$ with the help of the prediction equations we can calculate

(5) $$b_{t+1|t} = Ab_{t|t}$$

(6) $$\Gamma_{t+1|t} = A\Gamma_{t|t}A' + \Sigma$$

the values of $b_{1|0}$ and $\Gamma_{1|0}$. Getting the observations of $t=1$ the predictions of $b_{1|0}$ and $\Gamma_{1|0}$ were updated by

(7) $$b_{t+1|t+1} = b_{t+1|t} + K_{t+1}\left(y_{t+1} - x'_{t+1} b_{t+1|t}\right)$$

(8) $$\Gamma_{t+1|t+1} = \left(I - K_{t+1} x'_{t+1}\right)\Gamma_{t+1|t}$$

with: $K_{t+1} = \Gamma_{t+1|t} x_{t+1}\left(x'_{t+1}\Gamma_{t+1|t} x_{t+1} + \Omega\right)^{-1}$ Kalman-gain

I = identity matrix

The so derived values $b_{1|1}$ and $\Gamma_{1|1}$ were used as the starting values of the next step of the algorithm.

2. Flexible Least Squares

For estimating the model (1) and (2) Kalaba/Tesfatsion (1988) distinguish

1. the „measurement error"

(9) $$r_M^2 = \sum_{t=1}^{T} \left[y_t - x_t b_t\right]' \left[y_t - x_t b_t\right]$$ and

2. the dynamic error

(10) $$r_D^2 = \sum_{t=1}^{T-1} \left[b_{t+1} - A b_t\right]' \left[b_{t+1} - A b_t\right].$$

The rule of adaption is to minimize the sum of the two errors

(11) $$C = \sum_{t=1}^{T} \left[y_t - x_t b_t\right]' \left[y_t - x_t b_t\right] + \lambda \sum_{t=1}^{T-1} \left[b_{t+1} - A b_t\right]' \left[b_{t+1} - A b_t\right]$$

Here λ stands for the „trade of" between the two errors. This means that in the case of nearly constant parameters r_M^2 must be rather large compared to r_D^2. Assuming a minimized C, a reduction of r_M^2 is only possible if r_D^2 increases by a large step. So in this case λ must be rather large. For really constant parameters λ must be infinite. λ can be illustrated by a line of a constant C with the gradient $\frac{dr_M^2}{dr_D^2} = -\lambda$ (Fig. 2).

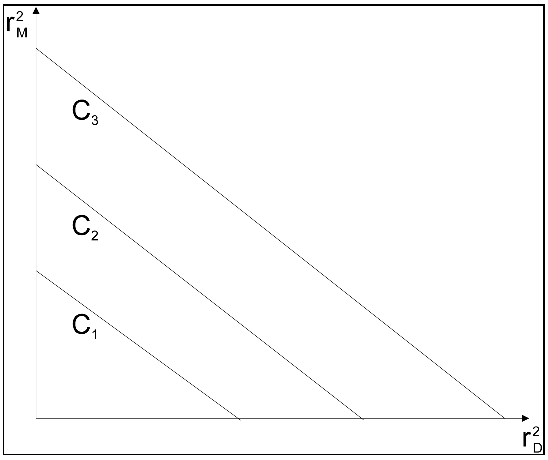

Fig. 2 Illustration of λ

Assuming that $t \geq 2$ then the minimum problem of equation (11) can be transformed into a recursive form:[1]

(12) $$\phi_t = \inf\{[y_t - x_t b_t]' [y_t - x_t b_t] + \lambda[b_{t+1} - Ab_t]' [b_{t+1} - Ab_t] + \phi_{t-1}\}$$

with: $\phi_{t-1} = \inf C_{t\text{-}1}$

Depending on a later to be determined K x K matrix Q, a K x 1 vector p and a scalar r the term in brackets can be written as follows:

(13) $$b_t'(\lambda A' A + x_t' x_t + Q_{t-1})b_t - (2\lambda b_{t+1}' A + 2y_t x_t + 2p_{t-1}')b_t + \lambda b_{t+1}' b_{t+1} + y_t^2 + r_{t-1}$$

The first order necessary condition for minimizing is:

(14) $$[\lambda A' A + x_t' x_t + Q_{t-1}]b_t - (\lambda b_{t+1}' A + y_t x_t + p_{t-1}')' = 0$$

Transforming to b_t we get:

(15) $$b_t^* = [\lambda A' A + x_t' x_t + Q_{t-1}]^{-1}(\lambda A' b_{t+1} + x_t' y_t + p_{t-1})$$

or in another form:

(116) $$b_t^{FLS} = U_t^{-1} z_t$$

with: $$U_t = x_t' x_t + Q_{t-1}$$

$$z_t = x_t' y_t + p_{t-1}$$

$$Q_t = \lambda[I - AG_t]$$

[1] Kalaba/Tesfatsion (1989)

$$p_t = G_t' \left[x_t' y_t + p_{t-1} \right]$$

$$G_t' = \lambda V_t A_t$$

$$V_t = \left[\lambda A_t' A_t + x_t' x_t + Q_{t-1} \right]^{-1}$$

As starting values Kalaba/Tesfatsion (1989) recommend:

- for Q_0 a K x K zero matrix
- for p_0 a K x 1 zero vector and
- for $r_0 = 0$.

4. FLS vs. Kalman-Filtering

In the following simulation we want to compare FLS and Kalman Filtering. Here we have to distinguish between the estimation with a correct specified hyperstructure and the estimation with a misspecified hyperstructure. The term „hyperstructure“ stands for all variables which must be known before the beginning of the parameter estimation. This means that the hyperstructure of Kalman Filtering includes

- the transition matrix A
- the covariance matrices Σ and Ω
- the starting values $b_{0|0}$ and $\Gamma_{0|0}$.

The hyperstructure of the FLS estimation includes

- the transition matrix A
- the starting values p_0 and Q_0.

Here the problem arises that p_0 and Q_0 are also unknown in the simulation but we can at least calculate Q_0 with the help of equation (15). For this calculation, we need assumptions relative to the qualities of p_0 and Q_0. In the following we assume that p_0 is a zero vector and Q_0 is a diagonal matrix.

In the case of a correct specification of the hyperstructure as well as in the case of a misspecification of the hyperstructure we carry out three simulations:

1. Multiple regression with constant parameters: Here we have three independent variables x_1, x_2 and x_3 with the parameters $\beta_1 = 10$, $\beta_2 = 6$ and $\beta_3 = -12$.
2. Multiple regression with structural breaks: The parameters start similar to simulation 1 but in period 17 β_2 changes into –6, in period 26 β_1 changes to –10 and in period 35 β_3 changes to +12.
3. Multiple regression with time varying parameters: Here, we have only two independent variables β_1 starting at 10 and decreasing by 1% per period and β_2 starting at 6 and decreasing by 1% per period. Moreover, both parameters are randomly disturbed.

Finally, it is to mention that the FLS estimation is given for a large as well as for a small λ.

In the following, only the estimation of parameter β_1 is observed for reasons of clearness. The results of the estimation of the other parameters are similar.

4.1 Estimation with a correct specified hyperstructur

Figure 3[2] shows the results of simulation 1. It can be seen that FLS with a small λ estimates the real coefficient very well. However, FLS with a large λ and Kalman Filtering need a rather long phase of adaption. After that, the estimation is rather well, too.

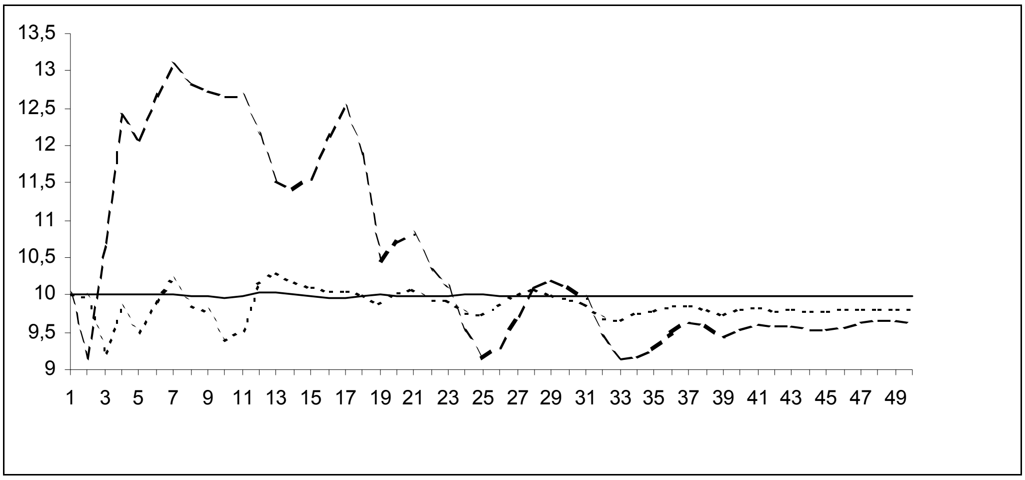

Figure 3: Simulation 1

Figure 4 shows that Kalman Filtering and FLS with a small λ react similar. Until the first break in period 17 we have a very good adaption to the real coefficient. In period 17 we have the break of parameter β_2. The estimation of β_1 differs from the true value for just a short period of time and it arrives a good adaption again quickly. We get a similar reaction after period 35 (break of parameter β_3). After the break of parameter β_1 the new value is reached after a short period of time.

The FLS estimation with a large λ reacts very sensitive at the breaks and the adaption to the real value again proceeds very slowly.

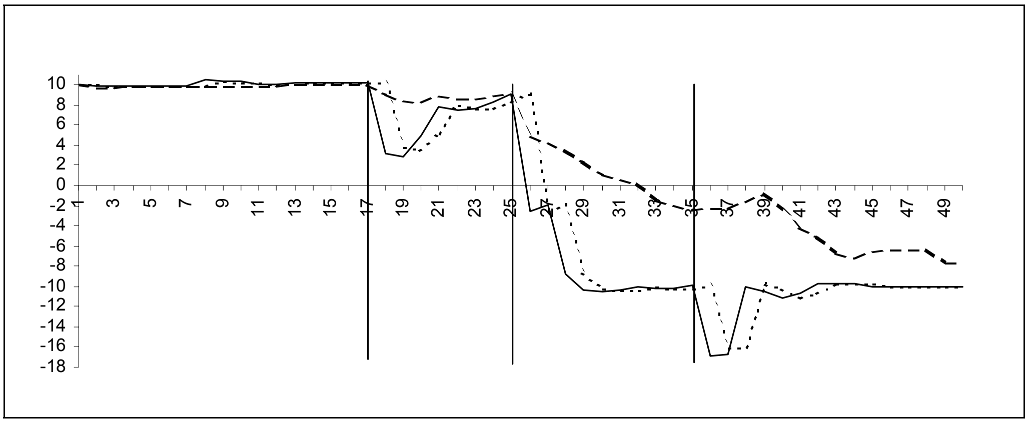

Fig. 4: Simulation 2

Figure 5 shows that Kalman Filtering has a very good adaption to the real coefficient. After a phase of adaption we get a good result with FLS with a large λ, too. The FLS estimation with a small λ rolls around the real value. An adaption does not appear to take place.

[2] In Fig. 3 – 6 the full line represents FLS with a small λ, the coarse pointed line represents FLS with a large λ and the fine pointed line represents the Kalman Filtering estimation. Moreover in fig. 5 the line with the rhombus‘ represents the real coefficients.

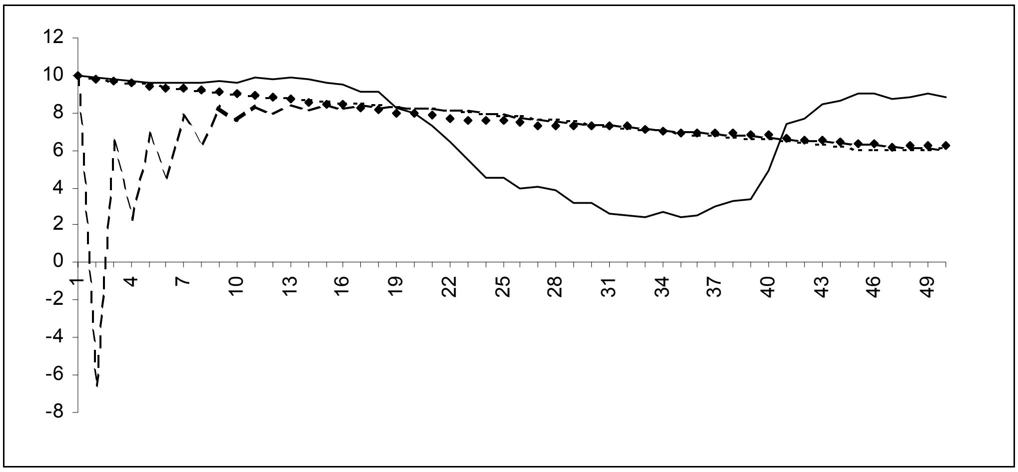

Fig. 5: Simulation 3

Here we should stop shortly for summing up.

- In the cases of constant coefficients and constant coefficients with a structural break the FLS estimation with a small λ seems superior to the FLS estimation with a large λ. The result had to be expected and it can be explained with the example of simulation 1 as follows. The real coefficients are not time varying like the estimated coefficients because FLS and Kalman Filtering estimate time varying coefficients in all cases. Therefore, on the one hand (real coefficients) we do not have a dynamic error but on the other hand (estimated coefficients) we do have one. So the importance of the dynamic error within the scope of the estimation has to be much to large if we use a large λ. In the same way, we can explain the superiority of FLS with a large λ over FLS with a small λ.
- A superiority of Kalman Filtering over FLS or vice versa is not to be recognized. So Kalman-Filtering is superior in simulation 3 but it is inferior to FLS with a small λ in simulation 1.

4.2 Estimation with a misspecified hyperstructure

In this chapter we want to ask which influence a misspecified hyperstructure has on the estimations with the help of Kalman Filtering and FLS. We have to distinguish the case of a wrong transition matrix A from the case of a wrong starting value of b.

1. We should start with the second case. Here we must disregard that in the case of FLS we cannot vary the starting value of b directly but only with the help of varying p_0 or Q_0. In the following we use a variation of Q_0.

A variation of the starting value of b is comparable to the structural breaks in simulation 2. So we can say that wrong starting values are compensated by FLS as well as by Kalman Filtering without any problems.

2. A variation of the transition matrix A results into large problems for both estimation methods. Even relatively small misspecifications (<5%) cause a fast exponential break out in the Kalman Filtering case (graphic representation is not possible) respectively no adaption to the real coefficient in the case of FLS (Fig. 6).

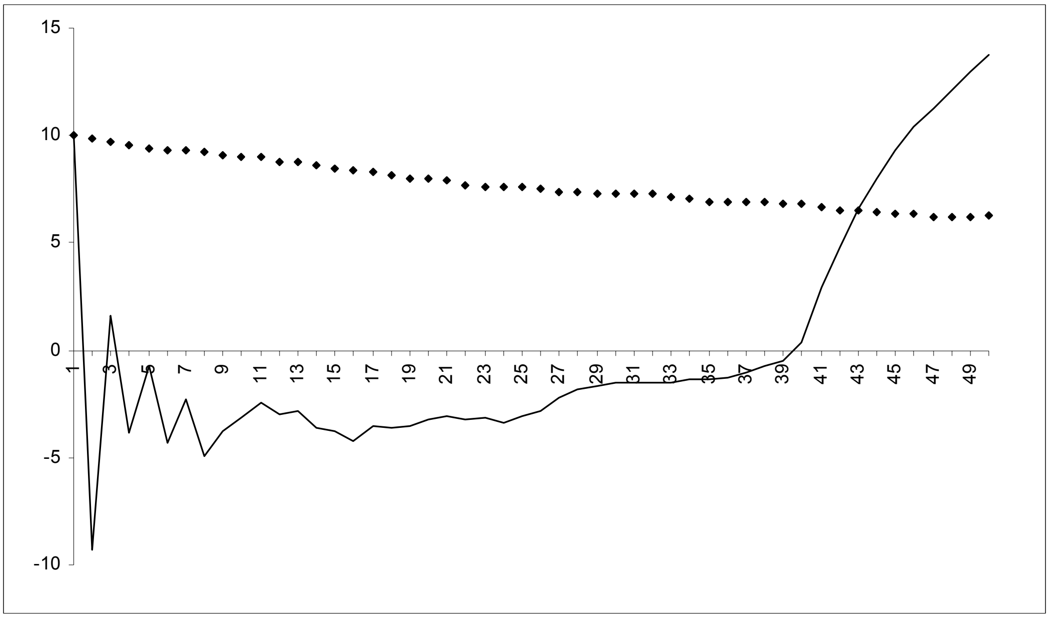

Fig. 6 FLS estimation with misspecified transition matrix

Finally, it must be mentioned that in the case of Kalman Filtering moreover the covariance matrices Σ, Ω and Γ can be misspecified. Here we can say that this has no effects on the parameter estimation.

5. Conclusion

The question in this paper was whether FLS is a real alternative to Kalman Filtering. This question must be answered with „yes“ and with „no“. The answer „yes“ is to be given if we are looking for a method which achieves the same or better estimation results like Kalman Filtering and which does not cause more problems than Kalman Filtering. This can be found here. We must say „no“ if we are looking for a method with which we can go around the problems of Kalman Filtering. This cannot be achieved by FLS. So we can say that as a matter of principle we have two equally good methods for estimating state space models. Kalman Filtering has got the advantages that

1. it is implemented into many statistical software packages (SAS/IML, RATS etc.)
2. methods for estimating the hyperstructure (for example ML-estimating with the help of the EM algorithm) are better prepared as in the FLS case.

References:

[1] Kalaba, R./L. Tesfatsion (1988), The Flexible Least Squares Approach to Time Varying Linear Regression, in Journal of Economic Dynamics and Control 12, pp. 43 - 48

[2] Kalaba, R./L. Tesfatsion (1989), Time-Varying Linear Regression via Flexible Least Squares, in: Computers Math. Applic. 17, pp. 1215 – 1245

[3] Kalaba, R./L. Tesfatsion (1990), Flexible Least Squares for Approximately Linear Systems, in: IEEE: Transactions on Systems, Man, and Cybernatics 20/5, S. 978 - 989