BEEF CATTLE FEED INTAKE AND GROWTH: EMPIRICAL BAYES
DERIVATION OF THE KALMAN FILTER APPLIED TO
A NONLINEAR DYNAMIC MODEL

J. W. Oltjen and F. N. Owens
Oklahoma State University
Stillwater 74078

ABSTRACT
Feed intake and growth rate of a single group of growing-finishing feedlot beef cattle are difficult to predict. Subsequent performance can be projected more precisely from past performance of a group of cattle. Using an adaptation of the statistical procedure called the empirical Bayes (EB) derivation of the Kalman filter, estimates from any dynamic model (M) can be adjusted based on past performance. The model may be either linear or nonlinear. With this procedure, predictions of intake and body weight gain are periodically updated by multiplying the estimates from M by statistically weighted factors. These factors are derived from the ratio of performance in each period to the performance predicted by M. For comparison to the EB adjustment, weighting of factors by least-squares (LS) adjustment also was tested to predict subsequent feed intake and gain. The test data base consisted of periodic feed intake and gain observations (usually 28 d) for 200 pens of feedlot steers. Bias of prediction was lower for EB than for M or LS for feed intake and (usually) gain. Intake and gain prediction errors averaged for the whole feeding period were .42 kg/d for intake and .14 kg/d for gain by EB, being .84 and .18 kg/d more precise than M and .12 and .33 kg/d more precise than LS predictions. More than two observations were needed before LS produced accurate prediction but after about 80 d, LS and EB estimates converged. Accuracy of both estimates continued to improve as days on feed increased. Over-adjustment of the prediction made LS clearly less accurate than M for the first few periods. In contrast, deviations early in a feeding period were properly weighted by EB, giving this procedure an advantage in early prediction of subsequent performance.

(Key Words: Bayesian Theory, Growth Models.)

Introduction
Predicting future biological behavior is one goal of models. Regression models are precise within a limited range, whereas models based on biological mechanisms have more general application (Oltjen, 1984). Though they aid in understanding, mechanistic models have generally not been used to predict animal performance.

Preliminary observations of a biological system, such as feed intake of a pen of cattle early in a finishing period, can be used to modify a model and thereby customize the prediction for a specific case. Model adjustment factors can be derived by least squares or a procedure based on the Bayes modification of the Kalman filter (Kalman, 1960; Kalman and Bucy, 1961). Using the Bayes procedures, information about prior behavior of a system can be integrated with current observations to predict future behavior (Allen and Jordan, 1982). These adjustments allow a mechanistic model to gain precision without sacrificing its generality.

Application of the Kalman filter to biological modeling has been limited (Meinhold and Singpurwalla, 1983). When applied to a dynamic linear system of bovine lactation (Goodall and Sprevak, 1985), it accurately estimated total milk yield after only a few observations early in lactation. It also has been used for recursive prediction of breeding values (Harville, 1979; Sallas and Harville, 1981; Hudson, 1984).

The objective of this research was to apply the empirical Bayes derivation of the Kalman filter to a nonlinear, dynamic system to predict feed intake and growth of beef cattle.

Experimental Procedure
To check the precision of predicting performance and gain, average daily body weight
gain (ADG) and daily feed dry matter intake (FI) for time periods within feeding trials with beef cattle were used. This data set was from feeding trials conducted at Oklahoma State University from 1975 to 1985 and contained observations from 200 pens of typical feedlot steers for which means are presented in table 1. Shrunken body weights and feed intakes were recorded for each period. A period is defined as the time interval from one weight to the next. Though these were frequently 28 d, in some instances periods lasted up to 57 d. Only 50 pens had observations for more than four periods. Three different predictions were used. First, the models of bovine growth and composition (Oltjen et al., 1986) and feed intake (Plegge et al., 1984) were employed without any adjustment (M). The growth model is based on mechanistic relationships between increasing cell numbers and cell size, genetic potential, energy expenditure and environmental influences (e.g., feed intake) during maturation of an animal. The feed intake model, in contrast to the growth model, is a regression equation based on the empirical relationships from a large data set which includes the effects of shrunken body weight (BW, kg), initial feeding period weight (IBW, kg), feed metabolizable energy (ME, mcal/kg) and implant treatment (IT, 0 if not implanted, 1 if implanted):

\[
\text{dry matter intake (kg/d)} = -43.045 - 0.004 \text{IBW} + 0.00003 \text{IBW}^2 \\
+ 0.07367 \text{BW} - 0.00008334 \text{BW}^2 \\
+ 24.5011 \text{ME} - 4.4019 \text{ME}^2 \\
+ 0.6 \text{IT}.
\]

The second prediction used initial measurements of intake (\(y_1\)) and gain (\(y_2\)) as the set of observable system variables with the empirical Bayes procedure (EB) to modify the model's predictions (see Appendix). The third method used a multiplicative adjustment by a least-squares mean estimation procedure (LS) where the ratio of observed to model predicted performance each time period was averaged to establish adjustment factors which were multiplied by model estimates to predict later feed intakes and gains. The recursive EB and LS procedures were applied at the beginning of each period using the latest period's estimates. Initial values for \(\Theta_o\), \(\Sigma_o\), \(\alpha\) and \(\sigma\) were needed to apply the EB procedure. The term \(\Sigma\) is the estimate of the multiplicative adjustment factor (A) and its error (\(e\)) for feed intake and gain, respectively. The term \(\Sigma\) is the variance-covariance matrix of \(\Theta\). The term \(\alpha\) relates errors for the present period to those of the previous period, and \(\sigma\) is the variance-covariance matrix for the independent errors of intake and gain prediction, respectively, not accounted for by \(\alpha\). These were determined as follows:

\[
\Theta_o = \begin{bmatrix} 1.1039 \\ 1.0279 \\ 0 \\ 0 \end{bmatrix},
\]

\[
\Sigma_o = \begin{bmatrix} 0.008208 & -0.002017 & 0 & 0 \\
-0.002017 & 0.014805 & 0 & 0 \\
0 & 0 & 0.009972 & 0.002926 \\
0 & 0 & 0.002926 & 0.029841 \\
\end{bmatrix},
\]

\[
\alpha = \begin{bmatrix} 0.1147 \\ 0 \\
0 & -0.1078 \end{bmatrix},
\]

\[
\sigma = \begin{bmatrix} 0.009841 & 0.002962 \\
0.002962 & 0.029494 \end{bmatrix}.
\]

These were based on preliminary analysis of data from 18 additional independent pens not used in the empirical Bayes analysis. The values in \(\Theta_o\) result in an initial adjustment to the model predictions at time 0. Investigation revealed that the procedure was insensitive to values of \(\alpha\) and \(\sigma\), similar to the results of Goodall and Sprevak (1985). Feed intake and daily weight gain were predicted for each period or for the total feeding period. Predicted feed intake and weight gain for each period \(\hat{y}_m_{Tt}, \hat{F}_m\Theta_{p}(p)\), and \(\hat{L}_{S_{t-1}}\hat{V}_m{T}_t\) for M, EB and LS, respectively, see Appendix) were compared with each period's observed feed intake and daily weight gain, \(Y_T,\) and similar comparisons were made between observed and predicted total feed intake and daily weight gain. The EB prediction of total period intake and gain was estimated without using the period ahead correlated error correction (see Appendix). The differences between observed and predicted intakes and gains for each period were examined by analysis of variance with method (M, EB or LS) and pen as main effects. When significant effects were observed, means were compared by least significant difference.
TABLE 1. DESCRIPTION OF THE STEER DATA USED FOR ANALYSIS

<table>
<thead>
<tr>
<th>Item</th>
<th>No. pens</th>
<th>Mean</th>
<th>SD(^a)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days on feed</td>
<td>200</td>
<td>141.5</td>
<td>26.2</td>
<td>111 - 201</td>
</tr>
<tr>
<td>Final day of period</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>36.7</td>
<td>11.9</td>
<td>27 - 57</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>74.6</td>
<td>23.9</td>
<td>55 - 121</td>
</tr>
<tr>
<td>3</td>
<td>176</td>
<td>101.3</td>
<td>21.1</td>
<td>83 - 140</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>121.1</td>
<td>11.2</td>
<td>111 - 147</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>140.0</td>
<td>.0</td>
<td>140 - 140</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>166.9</td>
<td>3.3</td>
<td>160 - 169</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>199.1</td>
<td>2.5</td>
<td>196 - 201</td>
</tr>
<tr>
<td>Shrunk wt, kg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial</td>
<td>200</td>
<td>290.0</td>
<td>41.6</td>
<td>193 - 357</td>
</tr>
<tr>
<td>Final</td>
<td>200</td>
<td>496.3</td>
<td>25.6</td>
<td>420 - 549</td>
</tr>
<tr>
<td>Daily gain</td>
<td>200</td>
<td>4.27</td>
<td>.13</td>
<td>1.18 - 1.79</td>
</tr>
<tr>
<td>Feed dry matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intake, kg/d</td>
<td>200</td>
<td>8.88</td>
<td>.85</td>
<td>7.07 - 10.80</td>
</tr>
<tr>
<td>ME, Mcal/kg</td>
<td>200</td>
<td>3.08</td>
<td>.12</td>
<td>2.90 - 3.57</td>
</tr>
</tbody>
</table>

\(^a\)Standard deviation.

Also, the effect of days on feed on these differences was analyzed by regression.

Results and Discussion

Precision of the various predictions of gain and intake for individual periods when predicted from the previous period are presented in table 2. Feed intakes and daily gains were under-estimated by the nonadjusted model (M) for most periods (table 2). For example, daily gain for the first period was under-predicted by .48 kg/d. This may, in part, be due to corresponding under-prediction in intake (—1.40 kg/d), or to additional gastrointestinal tract fill. However, the under-prediction of the EB procedure for period 1 is reduced by period 2, with mean feed intake and gain errors of —.09 and .21 kg/d, respectively. Feed intakes and gains for the total feeding period also were under-estimated by these models (table 3). Predictions by the EB-adjusted model for daily feed intake and gain in each feeding period and for the total time were much closer to observed values than the nonadjusted model predictions. This illustrates the advantage of an empirical Bayes procedure to alter a model to predict performance. The error in prediction became lower for EB than M or LS estimates for feed intake during each successive period, and for gains during most periods. Standard deviations generally were smaller for EB errors during the first few periods, particularly as compared with LS estimates. As the number of pens fed more than five periods is only one-fourth of the total pens fed, estimates for later periods become less reliable. The decrease in bias and the improvement in precision of prediction by applying the EB procedure after a single observation of performance demonstrates the value of the empirical Bayes procedure.

The LS estimates required several observations before performance was accurately predicted (tables 2 and 3). Regressions of inaccuracy of prediction expressed as a proportion of the observed value (residual/observed) decreased with time, as shown in figures 1 and 2. Precision was less accurate during the initial 80 d for LS than EB. Clearly, the use of the LS procedure was worse for predicting gain than using M itself during the first few periods. This is due to over-emphasis by LS of the deviation of the pen from the performance predicted by M. The EB procedure does not suffer from this limitation because the early deviations are more properly weighted by the Bayes statistics. After about 80 d, LS and EB estimates converged. Both EB and LS estimates for total period performance also continued to improve as time on feed progressed (figure 2). Precision of EB estimates of period performance, however, did not improve with time on feed but LS precision increased to equal the precision of EB as time passed (figure 1).

Coefficients of variation (R\(_t\)) for period ahead EB predictions of intake and gain de-
### TABLE 2. DIFFERENCE BETWEEN PREDICTED AND OBSERVED DAILY FEED INTAKE AND GAIN FOR INDIVIDUAL PERIODS FROM THE MODEL (M) AND EMPIRICAL BAYES (EB) OR LEAST-SQUARES (LS) ADJUSTED PREDICTIONS

<table>
<thead>
<tr>
<th>Period</th>
<th>No. pens</th>
<th>Feed intake (kg/d)</th>
<th></th>
<th>Daily gain (kg/d)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD(^a)</td>
<td>EB</td>
<td>SD</td>
<td>LS</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>-1.46b</td>
<td>1.14</td>
<td>-.63c</td>
<td>1.25</td>
<td></td>
<td></td>
<td>-.48b</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>-1.46b</td>
<td>1.23</td>
<td>-.09c</td>
<td>1.05</td>
<td>.76d</td>
<td>2.02</td>
<td>-.48b</td>
</tr>
<tr>
<td>3</td>
<td>176</td>
<td>-1.15b</td>
<td>1.03</td>
<td>.42c</td>
<td>.69</td>
<td>.56d</td>
<td>.76</td>
<td>-.37b</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>-1.03b</td>
<td>.89</td>
<td>.35c</td>
<td>.93</td>
<td>.42c</td>
<td>1.03</td>
<td>-.29b</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>-.71b</td>
<td>.91</td>
<td>.79c</td>
<td>.56</td>
<td>.92c</td>
<td>.51</td>
<td>-.24b</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>-.71b</td>
<td>1.00</td>
<td>.96c</td>
<td>.71</td>
<td>1.06c</td>
<td>.75</td>
<td>.04b</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>-.32b</td>
<td>.66</td>
<td>1.57c</td>
<td>.73</td>
<td>1.54c</td>
<td>.78</td>
<td>.29b</td>
</tr>
</tbody>
</table>

\(^{a}\)Standard deviation.
\(^{b},^{c},^{d}\)Means without a common superscript in each row within feed intake or daily gain differ (P<.05).

### TABLE 3. DIFFERENCE BETWEEN PREDICTED AND OBSERVED DAILY FEED INTAKE AND GAIN FOR THE TOTAL FEEDING PERIOD FROM THE MODEL (M) AND EMPIRICAL BAYES (EB) OR LEAST-SQUARES (LS) ADJUSTED PREDICTIONS

<table>
<thead>
<tr>
<th>Period</th>
<th>No. pens</th>
<th>Feed intake (kg/d)</th>
<th></th>
<th>Daily gain (kg/d)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD(^a)</td>
<td>EB</td>
<td>SD</td>
<td>LS</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>-1.13b</td>
<td>.84</td>
<td>-.23c</td>
<td>.82</td>
<td>.31d</td>
<td>.89</td>
<td>-.25b</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>-1.12b</td>
<td>.84</td>
<td>.16c</td>
<td>.66</td>
<td>.43d</td>
<td>.58</td>
<td>-.25b</td>
</tr>
<tr>
<td>3</td>
<td>176</td>
<td>-1.23b</td>
<td>.81</td>
<td>.20c</td>
<td>.50</td>
<td>.25c</td>
<td>.43</td>
<td>-.31b</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>-1.10b</td>
<td>.79</td>
<td>.19c</td>
<td>.40</td>
<td>.33c</td>
<td>.19</td>
<td>-.24b</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>-.51b</td>
<td>.86</td>
<td>.31c</td>
<td>.29</td>
<td>.33c</td>
<td>.19</td>
<td>-.25b</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>-.51b</td>
<td>.86</td>
<td>.16c</td>
<td>.23</td>
<td>.11c</td>
<td>.17</td>
<td>-.25b</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>-.13b</td>
<td>.94</td>
<td>.15c</td>
<td>.11</td>
<td>-.02c</td>
<td>.18</td>
<td>-.23b</td>
</tr>
</tbody>
</table>

\(^{a}\)Standard deviation.
\(^{b},^{c},^{d}\)Means without a common superscript in each row within feed intake or daily gain differ (P<.05).
creased from 13.5 and 21.1% for period 1 to 10.5 and 18.3% for period 7, respectively (table 4). Coefficients of variation for intake and gain predictions for the total feeding period decreased from 9.1 and 12.2% for predictions made at the beginning of period 1 to 4.0 and 5.6% for those made before period 7. Thus, as time increased, total feeding period projections became more precise despite little improvement in period predictions. When one considers variation in biological data, achieving a coefficient of variation under 10% would be considered excellent. The EB filter works by making most of its adjustment early in the feeding period, after period 1 (range 27 to 57 d). This suggests that variation between pens of cattle on feed can be detected based on measurements during the first 28 to 56 d on feed. The EB filter makes the correct adjustment at this time. Later adjustments are rather small by comparison because period-to-period variation dominates. Total period performance continues to be fine-tuned after the first period, but the filter does little to improve period ahead predictions after the first 56 d. No attempt to weight the present pen’s variation with that of the original 18 chosen to estimate initial values was made. Thus, EB precision should improve even more in actual practice where both the model and prior independent error variance (σ) are continually updated.

Figures 3, 4 and 5 illustrate how the EB procedure can be applied to one pen of cattle fed for 140 d. The model alone predicted feed intake at 6.4 kg over the total feeding period. Least-squares and EB total feed intakes were estimated for the next period and for the entire feeding period after each 28-d interval, as shown in figures 3 and 4, respectively. Next period final weight was estimated by adding next period gain to observed weight at the beginning of the period (figure 5). Initial feed

---

**TABLE 4. ESTIMATED COEFFICIENT OF VARIATION OF EMPIRICAL BAYES DAILY FEED INTAKE AND WEIGHT GAIN PREDICTIONS FOR EACH PERIOD**

<table>
<thead>
<tr>
<th>Period</th>
<th>Period ahead prediction</th>
<th>Total feeding period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intake</td>
<td>Gain</td>
</tr>
<tr>
<td>1</td>
<td>.135</td>
<td>.211</td>
</tr>
<tr>
<td>2</td>
<td>.115</td>
<td>.203</td>
</tr>
<tr>
<td>3</td>
<td>.111</td>
<td>.194</td>
</tr>
<tr>
<td>4</td>
<td>.109</td>
<td>.190</td>
</tr>
<tr>
<td>5</td>
<td>.107</td>
<td>.187</td>
</tr>
<tr>
<td>6</td>
<td>.106</td>
<td>.184</td>
</tr>
<tr>
<td>7</td>
<td>.105</td>
<td>.183</td>
</tr>
</tbody>
</table>
intake and gain were underpredicted, so LS and EB adjusted the projection upward. Empirical Bayes predictions were closer than LS or M estimates for period 2, but EB and LS predictions converged by the end of the fourth period. Although this pen had a prediction inaccuracy larger than that of typical pens, EB prediction of total period performance continued to improve with time. Large under-estimation of intake and performance during the first period caused over-compensation in subsequent estimates of performance. This bias continued until late in the feeding period. Despite these over-adjustments by the LS and EB procedures, predictions of intake and gain were more precise than from M alone most periods and for the total feeding period.

The algorithm to adjust model predictions by LS or EB methods is an attempt to make any mechanistic biological model that accurately predicts performance of a population useful in management to predict performance of a subset of the population based on preliminary data from that subset. Most complex simulation models of animal performance over a wide range of conditions are quite general. They yield precise estimates but require a number of inputs which are infrequently known. Their predictions in a given situation thus may be biased. On the other hand, simpler empirical models based on past observation usually are quite accurate and precise for a given set of management conditions, but their application under other management conditions yields imprecise estimates. To gain both empirical and mechanistic modeling advantages, one can capture the generality of the mechanistic models and improve its accuracy by altering predictions based on past observations. The use of EB is an example of this approach. General models based on fundamental principles can be integrated with a producer's database to improve precision of forecasting future performance. As part of an integrated management information system, this procedure allows a manager to project future performance of cattle already on feed and to answer "what-if" questions about changes in his management strategy. With increasing emphasis on efficient use of information for managing modern beef cattle enterprises, this procedure should enhance planning and decision making.

Both gain and feed intake were being predicted in this paper although in practice the lat-
The coefficient of variation in predicted gain can be reduced from over 10% (this study) to less than 5% if feed intake is known (Oltjen et al., 1986), knowledge of feed intake limits the precision of prediction. Rather than developing better models or adjustments to predict gain directly, it may prove more fruitful to develop more complete models of feed intake and employ widely accepted models (Lofgreen and Garrett, 1968; Fox and Black, 1984; NRC, 1984; Oltjen et al., 1986) to predict gain. Many factors that alter efficiency of utilization of feed energy have been incorporated into those models already. Until similar models are acceptable for feed intake prediction, adaptive procedures are needed to adjust intake equations as demonstrated by the use of the empirical Bayes methods in this paper. Thus, besides making mechanistic biological models more precise, the EB procedure is also useful for adapting empirical equations to new situations.

Appendix

A brief empirical Bayes derivation of the Kalman filter for the usual linear case follows; the notation is that of Meinhold and Singpurwalla (1983). Let $Y_t$ be a vector of observable variables at time $t$; $Y_t$ depends on some non-observable quantity $\Theta_t$, known as the “state of nature.” The linear observation equation relates $Y_t$ and $\Theta_t$:

$$Y_t = F_t \Theta_t + v_t$$  \[1\]

where $F_t$ is known and $v_t$, the observation error, is assumed to be normally distributed with mean zero and variance $V_t$. The system equation incorporates the dynamic nature:

$$\Theta_t = G_t \Theta_{t-1} + w_t$$  \[2\]

where $G_t$ is known and $w_t$, the system equation error, is assumed to be normally distributed with mean zero and variance $W_t$. Note also that the derivation allows $F$ and $G$ to be time dependent.

Harrison and Stevens (1976) showed that Bayesian formulation results in the following estimates of $\Theta$ and its variance-covariance matrix $\Sigma$. First set $\Theta_0$ and $\Sigma_0$ as the initial prior estimates ($t = 0$). The state predictor is then

$$\hat{\Theta}_t(p) = G_t \hat{\Theta}_{t-1}$$  \[3\]

with its variance-covariance matrix

$$R_t = G_t \Sigma_{t-1} G_t^T + W_t.$$  \[4\]

The gain in accuracy due to the filter is

$$K_t = R_t F_t^T [F_t R_t F_t^T + V_t]^{-1}$$  \[5\]

so the posterior estimate at time $t$ becomes

$$\hat{\Theta}_t = \hat{\Theta}_t(p) + K_t [Y_t - F_t \hat{\Theta}_t(p)]$$  \[6\]

with variance-covariance

$$\Sigma_t = R_t - K_t F_t R_t.$$  \[7\]

The problems associated with models where observational errors ($e_t = Y_t - \hat{Y}_t(p)$, where $\hat{Y}_t(p) = F_t \hat{\Theta}_t(p)$) are a correlated time-series may be corrected if expressed as

$$e_t = a_t e_{t-1} + e'_t$$  \[8\]

where $a_t$ is an autoregressive model and $e'_t$ is an independent sequence normally distributed with mean zero and variance $\sigma$. The formulation can be arranged so that $\Theta_t$, the state of nature, includes $e_t$, and that $G_t$, the system equation term, includes $a_t$. Therefore, $V_t$ is zero and the independent error and its variance, $e'_t$ and $\sigma$, are the non-zero parts of $w_t$ and $W_t$, respectively. In many situations, $F_t$ and $G_t$ (including $a_t$) do not change with time, as is also true for the variances $V_t$ and $W_t$; however, the state of the system $\Theta_t$ does change because the error is time dependent.

The general approach for nonlinear systems is similar to that outlined above. A set of variables from the dynamic model which also may be observed in the biological system are identified. For each variable, a linear function relating model and actual values is assumed. Parameters of this function are estimated using the Bayesian Kalman filter as observations are recorded. The updated function is used to modify model estimates to predict future system behavior.

**Theory.** The theory is similar to that used by Goodall and Sprevak (1985) for linear systems but it must be generalized for the use of nonlinear models here. For a system where $n$ observable variables in the biological system $Y_t = [y_{1t}, \ldots, y_{nt}]^T$ have been identified, let $\hat{Y}_{mt}$ be the vector of dynamic model predicted variables at time $t$ and let $Y_t$ be the observed val-
ues. Now let \( \Theta_t \) be a vector describing the Kalman filter "state of the system" consisting of \( A = [a_1, \ldots, a_n]^T \) with \( n \) associated errors at time \( t \), \( e_t = [e_{1t}, \ldots, e_{nt}]^T \):

\[
\Theta_t = \begin{bmatrix} A \\ e_t \end{bmatrix}
\]

[9]

Arrange a matrix \( F_t \) so that

\[
Y_{it} = Y_{mit}(a_i + e_{it})
\]

[10]

where \( M_n \) is a matrix:

\[
M_n = \begin{bmatrix}
y_{m1} & 0 & 0 \\
0 & y_{m2} & 0 \\
0 & 0 & \cdots \\
0 & 0 & y_{mn}
\end{bmatrix}
\]

[11]

This results in the notation of equation 1. Each \( a_i \) may be considered a multiplicative adjustment of dynamic model output to better describe the system's actual behavior.

The relationship between errors also may be incorporated using equation 8. If \( \alpha \) is a known transformation matrix relating time-series errors, then

\[
G = \begin{bmatrix} I_n & 0 \\
0 & \alpha \end{bmatrix}
\]

[12]

where \( I_n \) is the identity matrix of size \( n \) and \( 0 \) are \( n \times n \) zero matrices. Also,

\[
W = \begin{bmatrix} 0 & 0 \\
0 & \sigma \end{bmatrix}
\]

[13]

where \( \sigma \) is the variance-covariance matrix of the independent errors, \( e_t' \). The elements of \( \alpha \) and \( \sigma \) may be estimated from previous data, or they may be estimated recursively by a bootstrapping approach (Sprevak and Newmann, 1980).

To apply the outlined procedure, initial values for the state and its variance-covariance matrix must be specified. If no previous information is available, \( \Theta_0 \) is set to represent no deviation between the model predicted and observed values, and \( \Sigma_0 \) is the identity matrix. This allows incoming observations to dictate its convergence. Where historical data are available or the procedure has been applied previously, the initial values are used to weight the Bayes estimates.

Equation 3 gives the prior prediction of the state of the system, \( \hat{\Theta}_t(p) \), for time \( t \), the next period. This is used to predict the observable variable value adjustment prior to its observation:

\[
\hat{Y}_{mt}(p) = F_t \hat{\Theta}_t(p)
\]

[14]

with prior variance \( F_t R_t F_t^T \), where \( R_t \) is the prior variance-covariance of \( \hat{\Theta}_t(p) \) from equation 4.

At time \( t \), \( Y_t \) is observed, \( \hat{Y}_{mt} \) is calculated, and the Kalman filter gain is computed following equation 5. Applying \( K_t \), we have the new state of the system \( \hat{\Theta}_t \) and its variance \( \Sigma_t \) by using equations 6 and 7, respectively. These posterior estimates are then used in the next step of the recursive procedure.

Application. How the theory may be used to convert a dynamic, nonlinear simulation model into a forecasting system must be explained. Let \( S = [s_1, \ldots, s_q]^T \) be \( q \) state variables that define a system at time \( t \); also let \( P \) be the system model parameters and \( E \) be the external environmental variables (inputs). The model may be stated as:

\[
\frac{ds_i}{dt} = z_i (S, P, E)
\]

[15]

where \( z_i \) (\( i = 1 \) to \( q \)) are functions of the state variables, the parameters and the environment (France and Thornley, 1984). In practice, only a few of the state variables, parameters and environmental variables would be used in each function. This model may be solved by integration.

Now let \( Y = [y_1, \ldots, y_n]^T \) be the \( n \) observable system variables used in the forecasting procedure above. These are also functions of \( S, P \) and \( E \). The elements of \( Y \) may be state variables themselves, but are more likely to be auxiliary variables, i.e., functions of state variables or rates. Many equations in the model are written as a function of one or more of the \( y_i \):

\[
\frac{ds_i}{dt} = z_i' (Y, S, P, E)
\]

[16]
The forecasting algorithm is then applied by converting the dynamic model so that all \( y_i \) appearing as arguments in functions are replaced by \( a_i (\theta) y_i \) for the multiplicative adjustment. This is necessary so that adjustment for one variable less confounds adjustment for other variables.

**Literature Cited**


