Bridging Livestock Survey Results to Published Estimates through State-Space Models: A Time Series Approach

Stephen Busselberg[1]
[1]National Agricultural Statistics Service, 3251 Old Lee Hwy, Fairfax, VA 22030

Abstract
The status quo in survey sampling is publication and dissemination of survey results which are directly calculated as a function of the sample design and the survey inclusion probabilities. This process is appropriate for population parameter estimates for the case in which the only information known about the population of interest comes from the sample. In other cases, however, components of a system related to the population may be available from sources outside the survey sample. This external information may allow a deterministic inferential solution to the population parameters of interest which are targeted by the survey. For this reason, the survey results may be incongruous with external data and publishing the survey results leads to contradictory information. Such is the case with livestock inventory surveys collected by the National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA). This paper details a solution to this issue of incongruity through the application of a State-Space model system.

Key Words: Time Series, State-Space, Kalman Filter, ARIMA

1. Introduction

The National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) publishes estimates of agricultural production, prices paid and received by farmers, farm labor, chemical use, and many other items of interest in the field of agriculture. These estimates of agricultural population parameters may also be available at the state, district, and often county levels. The main instrument with which NASS measures these agricultural population parameters is survey sampling. NASS has extensive survey programs which use all manner of data collection methods including but not limited to mailings, telephone, web entry, personal interviews, and field enumeration. After the data collection process is complete, the survey data is summarized according to the appropriate sampling design. The calculated survey results then go before an entity called the Agricultural Statistics Board for review.

1.1 The Agricultural Statistics Board (ASB)

The Agricultural Statistics Board (ASB) is a panel of individuals, often referred to as "subject matter" experts, who review and analyze the summarized survey results; compare these results to previously published estimates and survey results (what is published is not necessarily synonymous with the summarized survey results); and when available, evaluate this information against nonproprietary statistics available on the target commodity. If the panel finds the survey results incongruous with available data, the survey results are adjusted in such way that the adjustments make them compatible. The adjusted estimates are now considered appropriate for publishing. Thus, one way of interpreting an ASB estimate of a particular agricultural population parameter is decomposing it into the summarized survey results plus an ASB adjustment. Currently there is no operational method of estimating the variance of the ASB estimate of an agricultural population parameter. In addition, the method at which the ASB arrives at its adjustment is neither consistent, nor transparent, nor repeatable, nor based on accepted statistical methodology.

1.2 Office of Management and Budget (OMB) Guidelines

The Office of Management and Budget (OMB) Standards and Guidelines Standard 4.1 reads as follows:

> "Agencies must use accepted theory and methods when deriving direct survey-based estimates, as well as model-based estimates and projections that use survey data. Error estimates must be calculated and disseminated to support assessment of the appropriateness of the uses of the estimates or projections. Agencies must plan and implement evaluations to assess the quality of the estimates and projections."[1]

OMB Standard 4.1 emphasizes that survey-based estimates must use *accepted* theory and methods. In addition standard errors of the estimates must also be provided. This paper illustrates how livestock commodities and other autocorrelated agricultural commodities can be modeled as a State-Space system which, in conjunction with the Kalman Filter, can provide estimates and estimate variances consonant with available commodity data in order to be in compliance with OMB standards and guidelines.

2. State-Space Models and the Kalman Filter

State-space models, referred to in many cases as Dynamic Linear Models (DLM), were initially developed for use in aerospace research.[2] They are used in modern control theory to represent the state of a linear time-invariant system which can be applied to many physical processes. Wei (1994) defines the state of a system as "a minimum set of information from the present and past such that the future behavior of the system can be completely described by the knowledge of the present state and future input."[3] A state-space system holds the Markov property that the future state of a system is conditionally independent of the past given the present state of a system.[4] The Kalman Filter, an iterative algorithm, is used to estimate the state of the system in state-space representation at a specific point in time. Kalman Filters are commonly used in navigation systems

---

[1] http://www.whitehouse.gov/omb/assets/omb/inforeg/statpolicy/standards_stat_surveys.pdf
[2] Shumway & Stoffer (2006)
[3] Wei (1994)
[4] Gilks, Richardson, and Spiegelhalter (1996)

including GPS, sonar, and radar. The state, which can be expressed as either a scalar or a vector, is often an unobserved signal of which "noisy" measurements are taken, for example, by satellites or sensors. An illustration of this concept is the location of a target of interest. The state could be a vector of coordinates describing the target's location at a particular point in time. The measurements of this location would be satellite observations. The satellite estimates of location may not all be equal, as some satellites readings may be more accurate than others. The Kalman Filter would then estimate the true location vector, given multiple "noisy" satellite measurements. An example of a federal agency that employs state-space models is the Bureau of Labor Statistics (BLS). BLS estimates monthly employment and unemployment statistics through state-space models.[5]

2.1 State-Space Representation of a System

The state-space representation of a system consists of two equations – a state equation which describes the behavior of the state process, and an observation equation which relates the state to the measurements or observations of the process. One state-space representation of the state equation given in Shumway & Stoffer (2006) is written

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$

The parameter $\boldsymbol{\Phi}$ is referred to as the transition matrix, as it "transitions" the past state at time $t-1$ to the current state of the system at time t. The vector $\boldsymbol{w}_t$ is a vector of Gaussian white noise with mean vector $\boldsymbol{0}$ and covariance matrix $\boldsymbol{Q}$. The observation equation that relates the state vector $\boldsymbol{x}_t$ to a vector of measurements $\boldsymbol{y}_t$ is

$$\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{v}_t$$

The parameter $\boldsymbol{A}$ is referred to as the measurement matrix. The vector $\boldsymbol{v}_t$ is a vector of Gaussian white observational noise with mean vector $\boldsymbol{0}$ and covariance matrix $\boldsymbol{R}$. Although the measurements or observations $\boldsymbol{y}_t$ can be predicted as in a "linear regression" problem, the traditional intent of state-space representation is to estimate the unobserved true system state $\boldsymbol{x}_t$ for all $\boldsymbol{t}$. This is accomplished using the Kalman Filter given noisy measurements $\boldsymbol{y}_t$ of the state $\boldsymbol{x}_t$.

2.2 The Kalman Filter Estimate of the System State

Given a system in state-space representation, the Kalman Filter is used to estimate the state of the system at any point in time. Defining the Kalman Filter estimate of the system state vector $\boldsymbol{x}_t$ as

$$\boldsymbol{x}_{t|s} = E[\boldsymbol{x}_t | \boldsymbol{y}_s, \boldsymbol{y}_{s-1}, \ldots, \boldsymbol{y}_1, \boldsymbol{\Phi}, \boldsymbol{Q}, \boldsymbol{A}, \boldsymbol{R}]$$

and the covariance matrix of the Kalman Filter estimate of the system state vector as

[5] Pfeffermann and Tiller (2006)

$$P_{t|s} = E\left[\left(x_t - x_{t|s}\right)\left(x_t - x_{t|s}\right)'\right]$$

with measurement vectors $\boldsymbol{y}_t$ and $t \in \{1, \dots, n\}$; and given an initial state estimate $\boldsymbol{x}_{0|0}$ and variance $\boldsymbol{P}_{0|0}$, the Kalman Filter estimates of the state for the system equations given in section 2.1 can be calculated as follows:

$$\boldsymbol{x}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{x}_{t-1|t-1}$$
$$\boldsymbol{x}_{t|t} = \boldsymbol{x}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{A}'\left[\boldsymbol{A}\boldsymbol{P}_{t|t-1}\boldsymbol{A}' + \boldsymbol{R}\right]^{-1}\left[\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{x}_{t|t-1}\right]$$
$$\boldsymbol{x}_{t-1|n} = \boldsymbol{x}_{t|t} + \boldsymbol{P}_{t-1|t-1}\boldsymbol{\Phi}'\left[\boldsymbol{P}_{t|t-1}\right]^{-1}\left[\boldsymbol{x}_{t|n} - \boldsymbol{x}_{t|t-1}\right]$$

The first equation is the Kalman Forecast, which provides a forecast of the system state, given observations up to the preceding point in time. The second equation is the Kalman Filter which describes the system state at a particular time given observations up to that point in time. The third and final equation is called the Kalman Smoother which provides an estimate of the system state at a point in time given all system observations. The Kalman Smoother is a function of both the Kalman Forecast and Kalman Filter. For clarification purposes, the “umbrella” term “Kalman Filter” refers to all three state estimates, namely the Kalman Forecast, Filter, and Smoother as defined. The corresponding variance estimates are

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{P}_{t-1|t-1}\boldsymbol{\Phi}' + \boldsymbol{Q}$$
$$\boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{A}'\left[\boldsymbol{A}\boldsymbol{P}_{t|t-1}\boldsymbol{A}' + \boldsymbol{R}\right]^{-1}\boldsymbol{A}\boldsymbol{P}_{t|t-1}$$
$$\boldsymbol{P}_{t-1|n} = \boldsymbol{P}_{t|t} + \boldsymbol{P}_{t-1|t-1}\boldsymbol{\Phi}'\left[\boldsymbol{P}_{t|t-1}\right]^{-1}\left[\boldsymbol{P}_{t|n} - \boldsymbol{P}_{t|t-1}\right]\left[\boldsymbol{P}_{t|t-1}\right]^{-1}\boldsymbol{\Phi}\boldsymbol{P}_{t-1|t-1}$$

Proofs of the above Kalman Filter equations can be found in Shumway & Stoffer (2006). These Kalman Filter equations are relevant only to the state-space representation given in this paper. Variants of the system equations require variants of the Filter. The parameters $\boldsymbol{\Phi}, \boldsymbol{Q}, \boldsymbol{A}, \boldsymbol{R}$ can be estimated using the Expectation Maximization (EM) algorithm which is also illustrated in Shumway & Stoffer (2006).

3. U.S. Hog Inventory as a State-Space System

NASS estimates quarterly hog inventory in the United States using the results from two survey statistics in addition to import, export, and hog slaughter estimates obtained from external sources. Detailed information about the survey process can be accessed from the official NASS website.[6] The remainder of this paper demonstrates how these livestock inventories, subject to various constraints relating to these external statistics, can be represented as a state-space system and therefore estimated using the Kalman Filter. Parameterization of a system using a model conveys *and* relies on various assumptions.

[6] http://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Hog_Inventory/index.asp

Consider the system state to be a vector of the true total U.S. hog and pig inventory $\boldsymbol{H}$, and pig crop inventory $\boldsymbol{P}$ (total pigs birthed and weaned) at a point in time, or $\boldsymbol{x}_t = [H_t \quad P_t]'$. The objective is then to define a state equation that shows the state of the system or true inventory at a given time t is reached by some transition from the inventory at time $t-1$, and to define an observation equation that relates this true inventory at a given time t to various measurements of that true inventory.

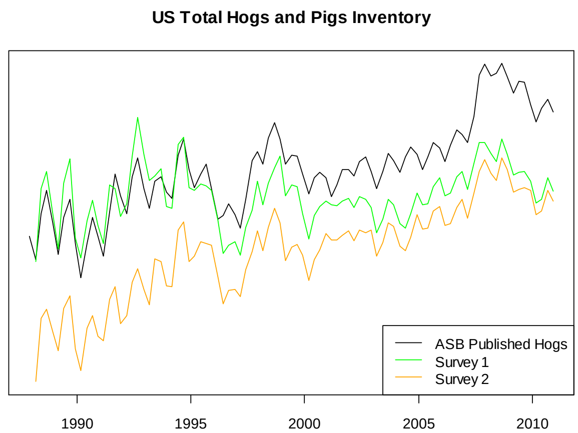

Figure 1: Total U.S. Hog inventory

Figure 1 shows the total quarterly U.S. hog inventory as determined by the ASB (black line) and two quarterly survey estimates (green and orange lines) of the true total hog inventory. The ASB published inventory is an estimate or a *measurement* of the true US hog inventory. In order to determine a specific parameterization of the system equations, we must make some assumptions about the ASB estimate relative to the true inventory. Three possible assumptions are

1. The ASB published inventory is the true inventory.
2. The ASB published inventory is an unbiased estimate of the true inventory.
3. The ASB published inventory is a biased estimate of the true inventory.

If the true inventory were known, we could compare the Kalman Filter estimate of the true inventory and the ASB estimate of the true inventory with the true inventory. However, the true inventory is not known. For the purposes of assessing the performance of state-space representation and estimation, it is more convenient to assume (1). In the case of (2), the observation variance in the observation equation relationship between the ASB estimate and the true state would be a measure associated variance of the board process, composed of varying individuals and varying emphasis on relevent commodity relationships with external data. Case (3) requires additional assumptions about the nature of bias between the ASB estimate and the true inventory. For example, is the bias time invariant, or constant over time; or is the bias time variant, i.e. changing with time? From Figure 1 we can only infer that the survey results have a time variant bias relative to the ASB published estimates. In order to feasibly assess the performance of the state-space representation of livestock inventory with respect to hogs, we will assume (1), that the ASB published inventory is the true inventory.

3.1 Defining the System State Equation

The assumption that the ASB published estimate is the true system state implies that we have observed the true system state for quarters of available published inventory. The state equation is therefore any model that transitions past values of published inventory to the present.

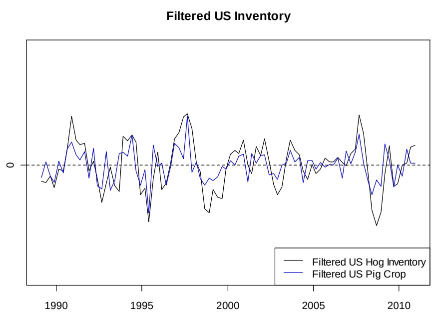

Figure 2: U.S. Filtered Hogs and Pig Crop

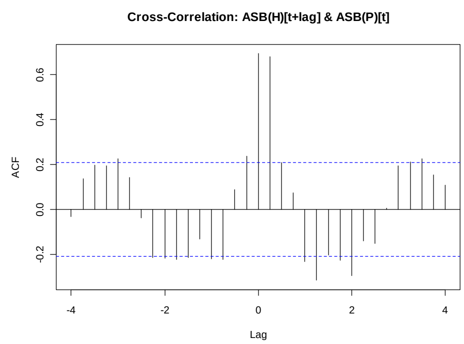

Figure 3: Cross-Correlation U.S. Hogs and Pig Crop

Figure 2 shows U.S. hog and pig crop inventories after applying a causal linear filter to $\boldsymbol{x}_t$. In the frequency domain, let $\boldsymbol{X}^*(z) = (1-z)(1-z^4)\boldsymbol{X}(z)$, where $\boldsymbol{X}(z)$ is the Z-Transform of $\boldsymbol{x}_t$, and $\boldsymbol{X}^*(z)$ is the Z-Transform of $\boldsymbol{x}_t^*$**.** This transformation filters out nonstationary trends and quarterly seasonality. The cross-correlation (Figure 3) structure between filtered U.S. hogs and filtered pig crop suggests that a vector autoregressive representation is a reasonable candidate for a state process model. We can model the transformation in the time domain as

$$\boldsymbol{x}_t^* = \begin{bmatrix} H_t^* \\ P_t^* \end{bmatrix} = \sum_{k=1}^{P} \begin{bmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{bmatrix}_k \begin{bmatrix} H_{t-k}^* \\ P_{t-k}^* \end{bmatrix} + \begin{bmatrix} w_{1_t} \\ w_{2_t} \end{bmatrix}$$

where $\boldsymbol{x}_t = \boldsymbol{x}_t^* + \boldsymbol{x}_{t-1} + \boldsymbol{x}_{t-4} - \boldsymbol{x}_{t-5}$. Therefore we will define our state equation as

$$\boldsymbol{x}_t = \begin{bmatrix} H_t \\ P_t \end{bmatrix} = \sum_{k=1}^{P} \begin{bmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{bmatrix}_k \begin{bmatrix} H_{t-k}^* \\ P_{t-k}^* \end{bmatrix} + \begin{bmatrix} H_{t-1} \\ P_{t-1} \end{bmatrix} + \begin{bmatrix} H_{t-4} \\ P_{t-4} \end{bmatrix} - \begin{bmatrix} H_{t-5} \\ P_{t-5} \end{bmatrix} + \begin{bmatrix} w_{1_t} \\ w_{2_t} \end{bmatrix}$$

3.2 Defining the System Observation Equation

The observation equation relates the state of the system as defined by the state equation to the measurements or observations of the state. The specific information available to the Agricultural Statistics Board during the hog inventory estimation is the survey results for surveys 1 and 2 which include total inventory and death loss. External data consists of hog imports and exports from Canada, in addition to commercial and farm slaughter counts. Slaughter is a federally inspected item and consequently considered by the ASB to be very accurate. When hog and pig crop inventory survey results are not consistent with slaughter counts, the accuracy of the survey results is questioned by the ASB. A complete list of observations or measurements related to the system state (hogs and pig crop) is as follows:

1. ASB official published inventories
2. Survey estimates of inventory
3. Death loss
4. Imports
5. Exports
6. Slaughter

3.2.1 ASB Official Published Inventories

For this paper, we assume that the ASB published inventories for hog and pig crop are the true inventories. Therefore the observation noise vector is zero with zero variance/covariance and the measurement matrix is equivalent to the identity matrix, which implies we are observing the true state. Defining $ASB(x)$ as the ASB inventory of x, our system of ASB observation equations are

$$\begin{bmatrix} ASB(H_t) \\ ASB(P_t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} H_t \\ P_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} H_t \\ P_t \end{bmatrix}$$

3.2.2 Survey Estimates of Inventories

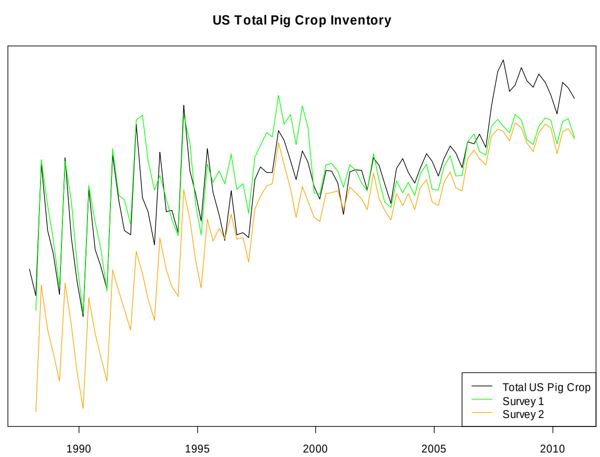

Figure 4: Total U.S. Pig Crop

The ASB published estimates for total U.S. hog (pig crop) inventory are shown with the survey results in Figure 1 (Figure 4). Assuming that the ASB estimate is truth, true inventory could be represented by the following relationship:

$$Survey(H_t) = \alpha_1 H_t + v_{3_t}$$

$$Survey(P_t) = \alpha_2 P_t + v_{4_t}$$

This relationship implies, however, that the bias is constant over time (α_1 and α_2 are time-invariant). Figure 1 and Figure 4 show this is not the case, as it appears the relationship is changing with time. For example, from 1988 to 1995, survey 1 for total hogs (Figure 1) is very close to the board, but after 1995, the gap between them grows with time. If we apply the causal linear filter $(1-z)(1-z^4)X(z)$ discussed in section 3.1 to both hogs and pig crop and their corresponding survey results, we can see in Figure 5 and Figure 6 that we can represent the survey results as

$$Svy(H_t) - Svy(H_{t-1}) - Svy(H_{t-4}) + Svy(H_{t-5}) = H_t - H_{t-1} - H_{t-4} + H_{t-5} + v_{3_t}$$
$$Svy(P_t) - Svy(P_{t-1}) - Svey(P_{t-4}) + Svy(P_{t-5}) = P_t - P_{t-1} - P_{t-4} + P_{t-5} + v_{4_t}$$

or concisely

$$Survey^*(H_t) = H_t^* + v_{3_t}$$
$$Survey^*(P_t) = P_t^* + v_{4_t}$$

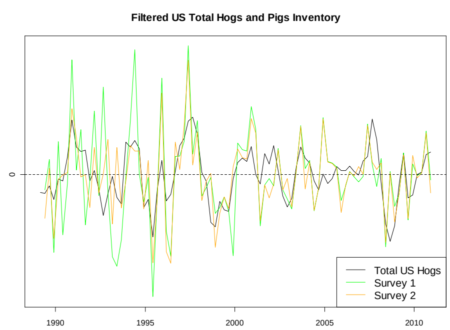

Figure 5: Filtered U.S. Hog Inventory

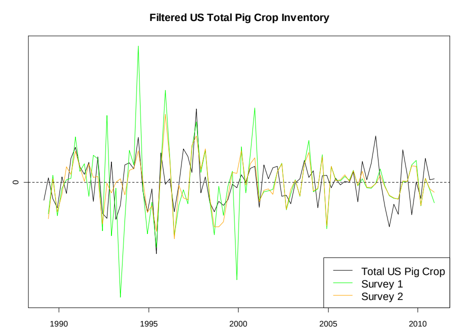

Figure 6: Filtered U.S. Pig Crop Inventory

The advantage of the filtered representation of the survey results is that the transformed results appear to equate to the system state by the *fixed* coefficient row vector $[1 \quad -1 \quad -1 \quad 1]$ in $\boldsymbol{A}$. This representation has two advantages. It eliminates the need for estimation of those parameters, and at the same time eliminates the need for estimating changing bias over time. State-Space representation is a time-invariant model.

3.2.3 "Balance Equation" Relationship

Defining E for exports, I for imports, D for death loss, and S for the sum of farm and commercial slaughter, the relationship between external data and U.S. hog and pig crop inventories can be expressed through the following relationship constraint:

$$H_t = H_{t-1} + P_t + I_t - E_t - D_t - S_t$$

If we subtract the right side of the equation from both sides we can define the residual as

$$r_t = H_t - (H_{t-1} + P_t + I_t - E_t - D_t - S_t)$$

If the residual is zero, then the inventories are in balance with the supply and deposition of hogs to the system. We can replace H_t, H_{t-1}, P_t with ASB values and survey values to examine the residuals from the various estimators $ASB(H_t), Survey(H_t)$ to get a sense of how "in balance" the system is over time according to the various estimators.

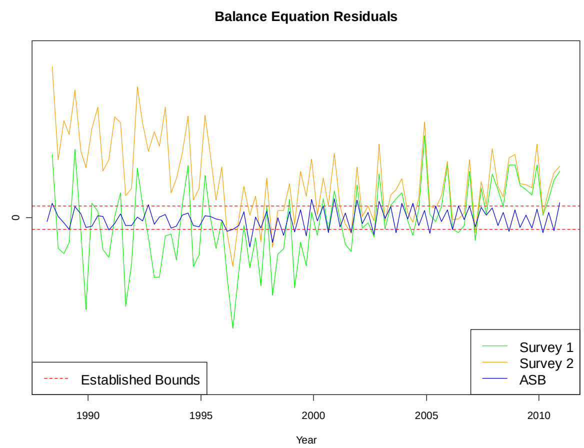

Figure 7: Balance Equation Residuals

Figure 7 graphs the residuals together with a dashed red line that represents a threshold established as a reasonable bound for the inventory to be in balance. From Figure 7 it is clear that the ASB made adjustments to the survey results so that the residuals of the balance equation would fall within the acceptable threshold. This constraint can be included in the observation vector by collecting the terms that correspond to external data or "external observations" on the left side of the equation, and collecting the terms for hog and pig crop inventories with the residual on the right side of the equation.

$$I_t - E_t - D_t - S_t = H_t - H_{t-1} - P_t - r_t$$

If we define ${v_5}_t = -r_t$ and $N_t = I_t - E_t - D_t - S_t$ then our observation equation for the external balance constraint items becomes

$$N_t = H_t - H_{t-1} - P_t + {v_5}_t$$

3.2.4 Slaughter and Pig Crop

An additional constraint which the Agricultural Statistics Board uses when making adjustments to survey results is the relationship that current slaughter corresponds approximately to those pigs born two quarters back in time. This suggests the relationship

$$S_t = \alpha P_{t-2} + {v_6}_t$$

where $\alpha = 1$. However, when we graph S_t and $ASB(P_{t-2})$ as shown in Figure 8, we see that before the year 2000, $\alpha < 1$ and after 2000, $\alpha > 1$. A second time we run into the problem of a time-variant measurement matrix. As an alternative, we can apply the linear filter transformation to slaughter and pig crop and obtain time-invariant fixed elements for the measurement matrix, justified by the graph of the transformation in Figure 9. Figure 9 suggests the observation equation

$$S_t - S_{t-1} - S_{t-4} + S_{t-5} = P_{t-2} - P_{t-3} - P_{t-6} + P_{t-7} + v_{6_t}$$
$$S_t^* = P_{t-2}^* + v_{6_t}$$

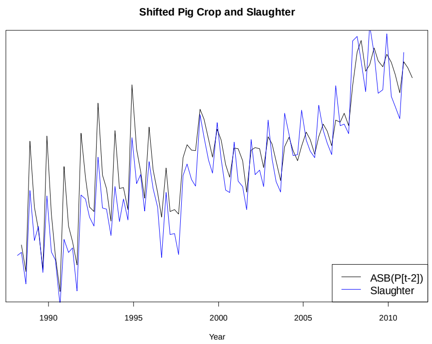

Figure 8: Slaughter and Pig Crop

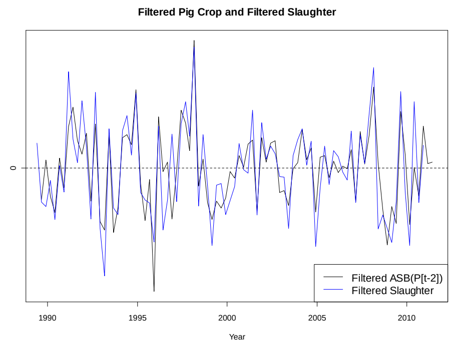

Figure 9: Filtered Slaughter and Pig Crop

3.2.5 Slaughter and Hogs

A final relationship used by the Agricultural Statistics Board that ties slaughter to hogs is nonlinear. The relationship is based on the assumption that the annual proportional change in hogs is slaughtered over the following two quarters. This assumption is mathematically expressed by the relationship

$$\frac{S_t + S_{t-1}}{S_{t-4} + S_{t-5}} = \frac{H_{t-2}}{H_{t-6}} + v_{7_t}$$

The fit is demonstrated in Figure 10. The introduction of this nonlinear constraint requires the use of the Extended Kalman Filter[7] which is a variant of the Kalman Filter presented in section 2.2.

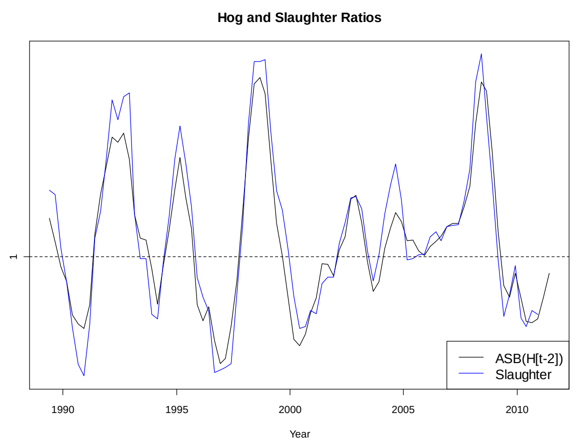

Figure 10: Slaughter and Hog Ratio Fit

[7] Anderson and Moore (1979)

The system of observation equations derived in sections 3.2.1-3.2.5 is summarized as follows:

$$[\boldsymbol{y}_t] = [\boldsymbol{A}\boldsymbol{x}_t] \qquad\qquad + [\boldsymbol{v}_t]$$

$$\begin{bmatrix} ASB(H_t) \\ ASB(P_t) \\ Survey^*(H_t) \\ Survey^*(P_t) \\ N_t \\ S_t^* \\ \dfrac{S_t + S_{t-1}}{S_{t-4} + S_{t-5}} \end{bmatrix} = \begin{bmatrix} H_t \\ P_t \\ H_t^* \\ P_t^* \\ H_t - H_{t-1} - P_t \\ P_{t-2}^* \\ \dfrac{H_{t-2}}{H_{t-6}} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ v_{3_t} \\ v_{4_t} \\ v_{5_t} \\ v_{6_t} \\ v_{7_t} \end{bmatrix}$$

4. Performance

It has been shown that hog and pig crop inventories can be represented as a state-space system which includes time dependent relationship constraints which the Agricultural Statistics Board examines in adjusting hog and pig crop inventory to be congruous with available external economic data. Both hogs' and pig crop's relationship with slaughter suggests that current slaughter data provides relevant information for the estimation of those commodities two quarters prior. The Agricultural Statistics Board uses current slaughter to make revisions to estimates published up to four quarters back in time. This can be accomplished in a state-space model via a Kalman Smoother estimator window of bandwidth 5, or

$$\boldsymbol{x}_{t-i|t} = E[\boldsymbol{x}_{t-i}|\boldsymbol{y}_t]$$

$$\text{where } i \in \{0,1,\ldots,4\}$$

The state-space representation of hog and pig crop inventories developed in this paper assumes that the ASB published estimates are the true inventories. In order to compare the model results with the published ASB inventories, we can select a time s such that $1 < s < n$, where 1 indexes the first vector of observations in time and n indexes the most recent quarter available (last observations). Defining $\boldsymbol{\Theta}_k = \{\boldsymbol{\phi}_k, \boldsymbol{Q}_k, \boldsymbol{R}_k\}$ as the parameter space, we can successively calculate $\boldsymbol{\Theta}_{\mathrm{k|k-5}} = \boldsymbol{\Theta}_{\mathrm{k}}|\mathbf{y}_{\mathrm{k-5}}$ and $\boldsymbol{x}_{k-4|k}|\boldsymbol{\Theta}_{k|k-5}$ for $k \in \{s, s+1, \ldots, n\}$ by defining $A\hat{S}B(\boldsymbol{x}_{k-4}) = \boldsymbol{x}_{k-4|k}|\boldsymbol{\Theta}_{k|k-5}$. We can then compare the actual ASB published estimates $ASB(\boldsymbol{x}_{k-4})$ to the Kalman Smoother estimates $A\hat{S}B(\boldsymbol{x}_{k-4})$ for $k \in \{s, s+1, \ldots, n\}$ to show what would have hypothetically occurred had the model *replaced* the ASB from time $s-4$ to time $n-4$. In this scenario the ASB observation would be

$$\boldsymbol{y}_t = \begin{cases} ASB(\boldsymbol{x}_t) & 1 \leq t < s-4 \\ A\hat{S}B(\boldsymbol{x}_t) & s-4 \leq t \leq n \end{cases}$$

Choosing $s \sim \frac{1}{2}n + 5$ we can produce estimates of this fourth and final revision for the last half of the data as if the ASB had never existed from $t \in \{s-4, s-3, \ldots, n\}$. Essentially the model estimate becomes "truth" and is used in parameter estimation after the fourth revision. This is estimable because of the information provided by the external data and accomplished by setting the corresponding ASB rows of the measurement matrix to zero (this is how missing observations are commonly treated) for $\boldsymbol{y}_{\boldsymbol{k-4}}$ through $\boldsymbol{y}_{\boldsymbol{k}}$ for each iteration $k \in \{s, s+1, \ldots, n\}$. Figure 11 shows the scenario results for U.S. hog inventory and Figure 12 shows the scenario results for U.S. pig crop inventory. Figure 13 demonstrates the balance equation residuals calculated from the hog and pig crop model estimates staying will within the established threshold.

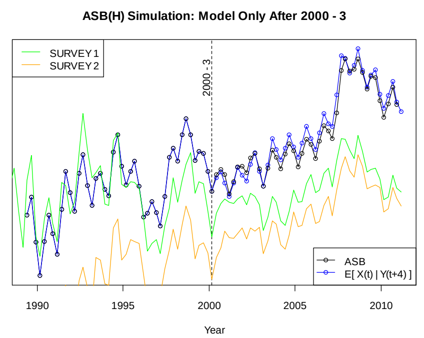

Figure 11: Window Smoother Estimate for U.S. Hogs

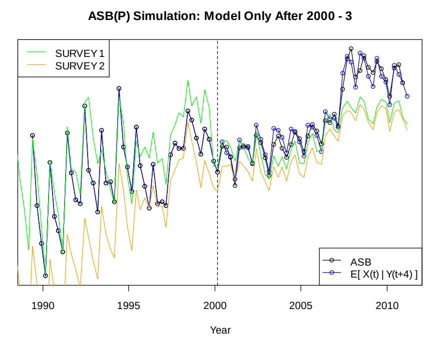

Figure 12: Window Smoother Estimate for Pig Crop

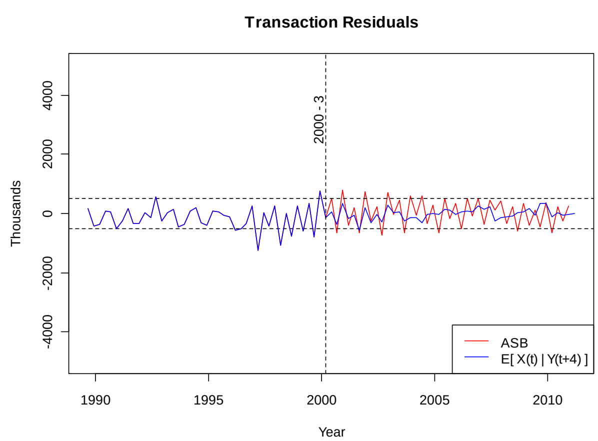

Figure 13: Balance Sheet Residuals

5. Discussion

The state-space representation of U.S. hog inventories presented in this paper is parameterized to estimate the *ASB estimate* of true inventories. This parameterization assumes that the *ASB estimate* of true inventories is not an estimate, but the actual inventories. This assumption is not realistic as true inventories are an unknown quantity. The ASB contains variability due to the individual composition of the board members and inconsistent emphasis on relationship constraints. Operationally, it is more realistic to assume that the ASB published estimate of inventories is a *measurement* of the truth. Not only is this a more realistic approach, but it allows for inclusion of a measurement of "expert opinion" in the Kalman Filter estimate. Some possible parameterizations and their implications are demonstrated in the following equations:

$$\begin{bmatrix} ASB(H_t) \\ ASB(P_t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} H_t \\ P_t \end{bmatrix}$$

Equation 1: The ASB is the True Inventory

$$\begin{bmatrix} ASB(H_t) \\ ASB(P_t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} H_t \\ P_t \end{bmatrix} + \begin{bmatrix} v_{1_t} \\ v_{2_t} \end{bmatrix}$$

Equation 2: The ASB is an Unbiased Estimate of True Inventory

$$\begin{bmatrix} ASB(H_t) \\ ASB(P_t) \end{bmatrix} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \begin{bmatrix} H_t \\ P_t \end{bmatrix} + \begin{bmatrix} v_{1_t} \\ v_{2_t} \end{bmatrix}$$

Equation 3: The ASB is a Time-Invariant Biased Estimate of True Inventory

Equation 1 has already been demonstrated in this paper. Equation 3 requires additional assumptions about the nature of the ASB bias relative to the true inventory. It would also require reassessing the nature of the bias of the survey results relative to the true inventory. Equation 2 would imply a robust believe that the Agricultural Statistics Board approaches the right level of inventory, but is still a variable estimate of truth. Given any of the three assumptions, state-space representation of livestock inventory utilizes the Kalman Filter to make use of all available measurements of a commodity and their relationships to exogenous data. Additionally, state-space representation allows those measurements to be emphasized according to their fit to the parametric representation of constraints that implies those assumptions.

Acknowledgements

I would like to thank Dr. Kristine Bell for productive discussions during the model development and for suggesting how to incorporate the balance sheet equation constraint into the observation equation.

References

Anderson, D.O. and Moore, John B. *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1979.

Box, G.E.P. and Jenkins G.M., and Reinsel G.C. *Time Series Analysis: Forecasting and Control*. Hoboken, N.J.: John Wiley & Sons, Inc., 2008.

Gentle, J.E. *Numerical Linear Algebra for Applications in Statistics*. Berlin: Springer-Verlag, 1998

Gilks, W.R.; Richardson S.; and Spiegelhalter, D.J. *Morkov Chain Monte Carlo in Practice.* London, UK: Chapman & Hall, 1996.

Pfeffermann, Danny, and Tiller, Richard. *Small-Area Estimation with State-Space Models Subject to Benchmark Constraints: Journal of the American Statistical Association*, December 2006, Vol. 101, No. 476.

Rencher, Alvin C. *Methods of Multivariate Analysis*. Danvers, MA: Wiley-Interscience, 2002.

Shumway, R.H. and Stoffer, D.S. *Time Series Analysis and its Applications*. New York: Springer Science+Business Media, LLC., 2006.

Wei, William S. Time *Series Analysis: Univariate and Multivariate Methods*. Redwood City, CA: Addison-Wesley Publishing Co., 1990.