

Paper SAS6447-2016
Linear State Space Models in Retail and Hospitality

Beth Cabbage, SAS Institute Inc.

ABSTRACT

Retailers need critical information about the expected inventory pattern over the life of a product to make pricing, replenishment, and staffing decisions. Hotels rely on booking curves to set rates and staffing levels for future dates. This paper explores a linear state space approach to understanding these industry challenges, applying the SAS/ETS[®] SSM procedure. We also use the SAS/ETS SIMILARITY procedure to provide additional insight. These advanced techniques help us quantify the relationship between the current inventory level and all previous inventory levels (in the retail case). In the hospitality example, we can evaluate how current total bookings relate to historical booking levels. Applying these procedures can produce valuable new insights about the nature of the retail inventory cycle and the hotel booking curve.

INTRODUCTION

Customer behavior can be a mystery. Why does my hotel have more bookings on the second Tuesday in June than on the third Tuesday in June? This plaid shirt is very popular in Virginia, but much less so in Oregon. I need to know how these variations will affect my planning. The answers to these questions are often difficult to pinpoint, but can be critical to business operations. SAS offers many analytical tools and solutions to help answer such questions. PROC SSM is a new procedure in SAS/ETS 9.4, and it has robust capabilities in the modeling space. The procedure has an extremely flexible framework, so it can accommodate a variety of different models. The syntax is easily modified, and allows the user to quickly test many different models. PROC SSM provides output that is the same across models, regardless of whether a simple regression model or a complex steady state model is estimated. Finally, it offers significant run-time advantages. Hospitality and retail inventory models showcase these features of PROC SSM.

STATE SPACE MODELS OF BOOKING CURVES

Hotels use booking curves to predict the final number of rooms that will be booked on any given day. This also allows hotels to have a view of when to expect bookings to be made. For a property that is primarily leisure travel, bookings might occur significantly before the stay date. Properties that see a large amount of business travel likely have a higher percentage of bookings occur much closer to the stay date. The booking curve is a visual representation of this process. Figure 1 shows the shape of a typical booking curve:

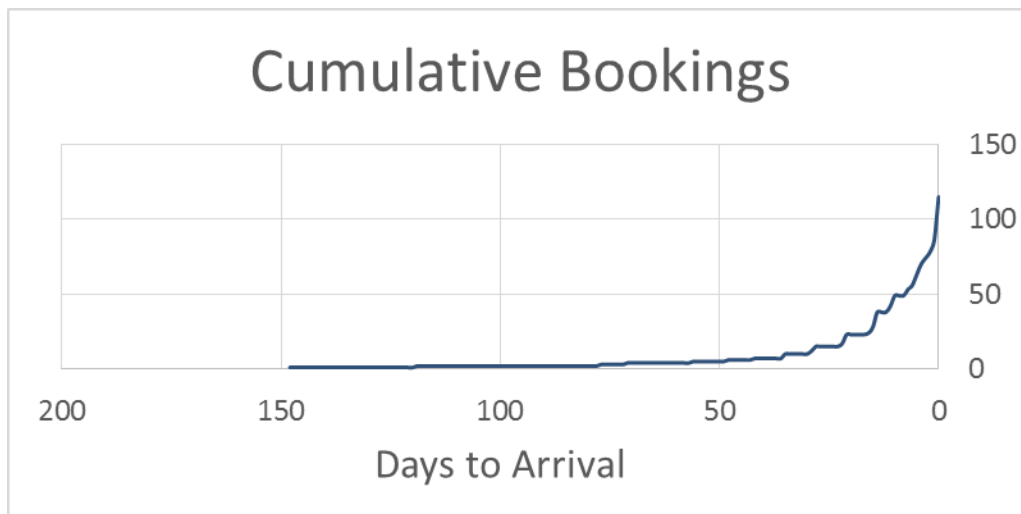


Figure 1. Example Booking Curve

Days to arrival appears on the x-axis, while the cumulative rooms booked appears on the y-axis. Note that bookings typically pick up quickly within approximately 30 days of the stay date. The shape of the booking curve can vary based on a variety of factors. Commonly, different customer segments, days of week, and seasons can have different booking curve shapes. Special events and rate changes might also impact the booking curve. Figure 2 shows two different booking curves. Note how a minor deviation at 60 days to arrival becomes a major difference in the final number of rooms booked by the arrival date—nearly 120 rooms booked in one case but just 35 booked in the other.

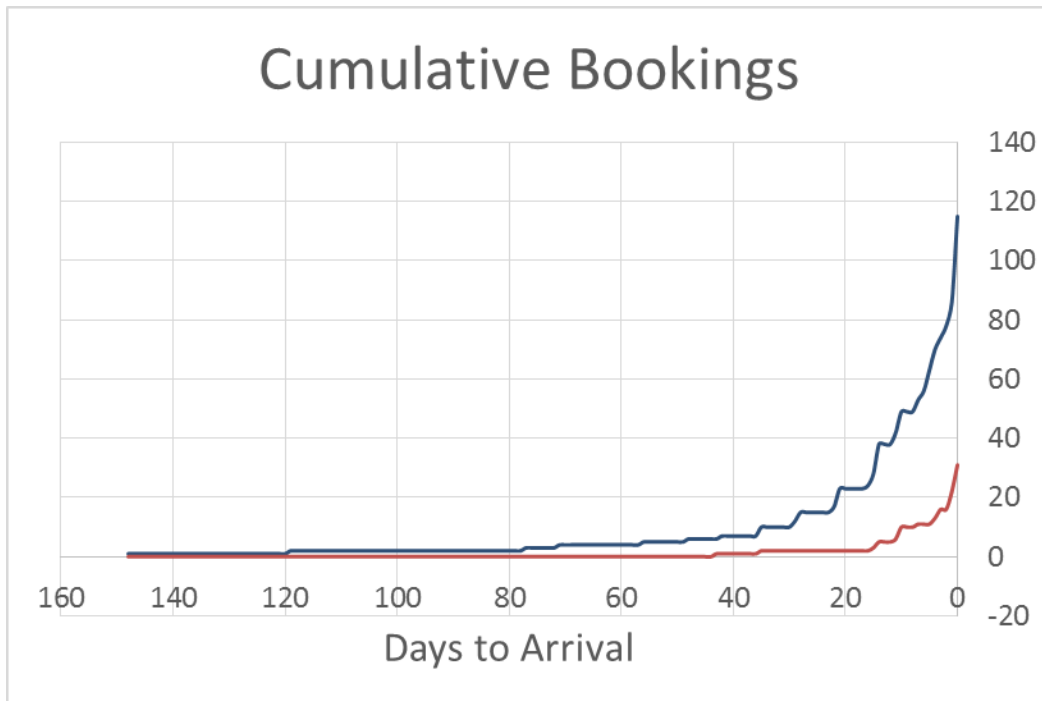


Figure 2. Example of Different Booking Curves

Identifying such variation can have a major impact on the ultimate forecast accuracy. PROC SIMILARITY can be used to get a sense of the relationships between different series. Booking curves for the same customer segment, but different days, can show significant variation. The output of this procedure includes a number of similarity measures, including distance metrics, distance matrices, and path statistics (Leonard 2016).

REGRESSION MODEL

To experiment with PROC SSM, estimate a simple model of cumulative bookings. The model takes the following form:

$$C = \alpha + \beta \exp(-TTS) + \varepsilon$$

The variable C represents cumulative bookings. The variable TTS represents time to service. The parameter α represents the intercept, the parameter β represents the time to service effect, and ε is white noise. In the SAS procedure, the data is grouped by month and customer segment.

Estimates

You can estimate the parameters of this model using both PROC SSM and PROC REG. Since this model does not contain any state variables, the two procedures will yield the same result. Figure 3 and Figure 4 show the parameter estimates from PROC SSM and PROC REG, respectively.

Regression Parameter Estimates					
Response Variable	Regression Variable	Estimate	Standard Error	t Value	Pr > t
In_cum_bk	intercept	2.4493	0.019696	124.36	<.0001
In_cum_bk	neg_tts	0.0234	0.000316	74.19	<.0001

Figure 3. Regression Parameter Estimates from PROC SSM

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	2.44928	0.01970	124.36	<.0001
neg_tts	1	0.02343	0.00031578	74.19	<.0001

Figure 4. Regression Parameter Estimates from PROC REG

The regression parameter estimates are the same for both procedures, although the PROC REG output shows one additional decimal point.

Model Results

Figure 5 shows the resulting model, compared to three actual booking curves.

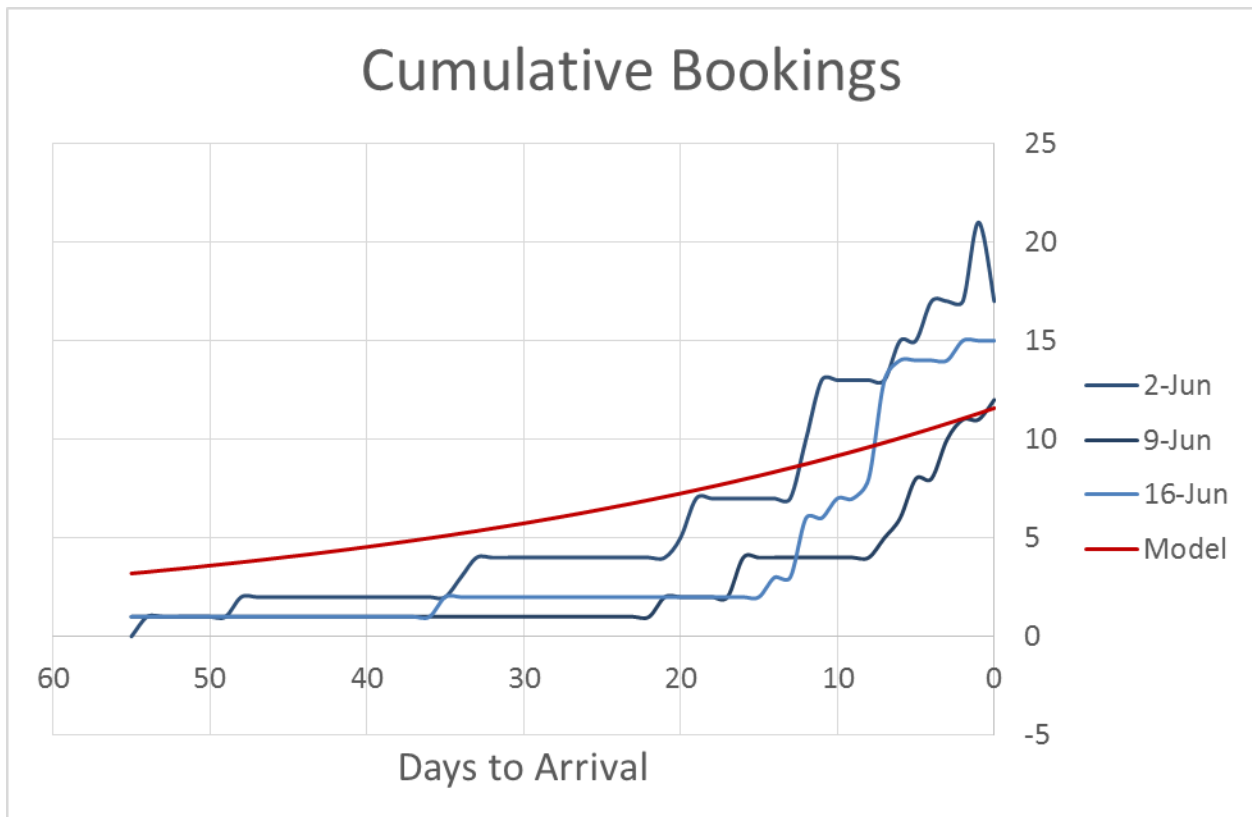


Figure 5. Three Booking Curves with Regression Model

The slope of the booking curve increases as the days to arrival decreases. This feature is not adequately captured by the simple regression model shown. The forecast error on the date of stay is shown in Table 1.

	June 2	June 9	June 16	MPE	MAPE
Regression Model	-45%	-3%	-23%	-24%	24%

Table 1. Booking Curve Error % on Day of Arrival

Performance

PROC SSM shows a significant performance improvement. For the same set of data, PROC REG completed in 1 minute and 19 seconds. The model ran in only 36 seconds through PROC SSM.

STATE SPACE TREND MODEL

The state space procedure affords significant flexibility in modeling. PROC SSM allows you to quickly add a pre-defined trend to the model. One available trend type, the Ornstein-Uhlenbeck decay pattern, closely resembles the shape of the booking curve. You can find additional detail in *SAS/ETS 14.1: User's Guide* (SAS 2013). The data is segmented by customer segment, month, and day of week. One model included only the trend and white noise. A second model, for comparison, also included a price effect.

Model Results

The second model, with the price effect, yielded slightly better results. The model also showed a much better fit relative to the regression model. Figure 6 shows the booking curve actualization for June 2, 2013, compared to the model with trend.

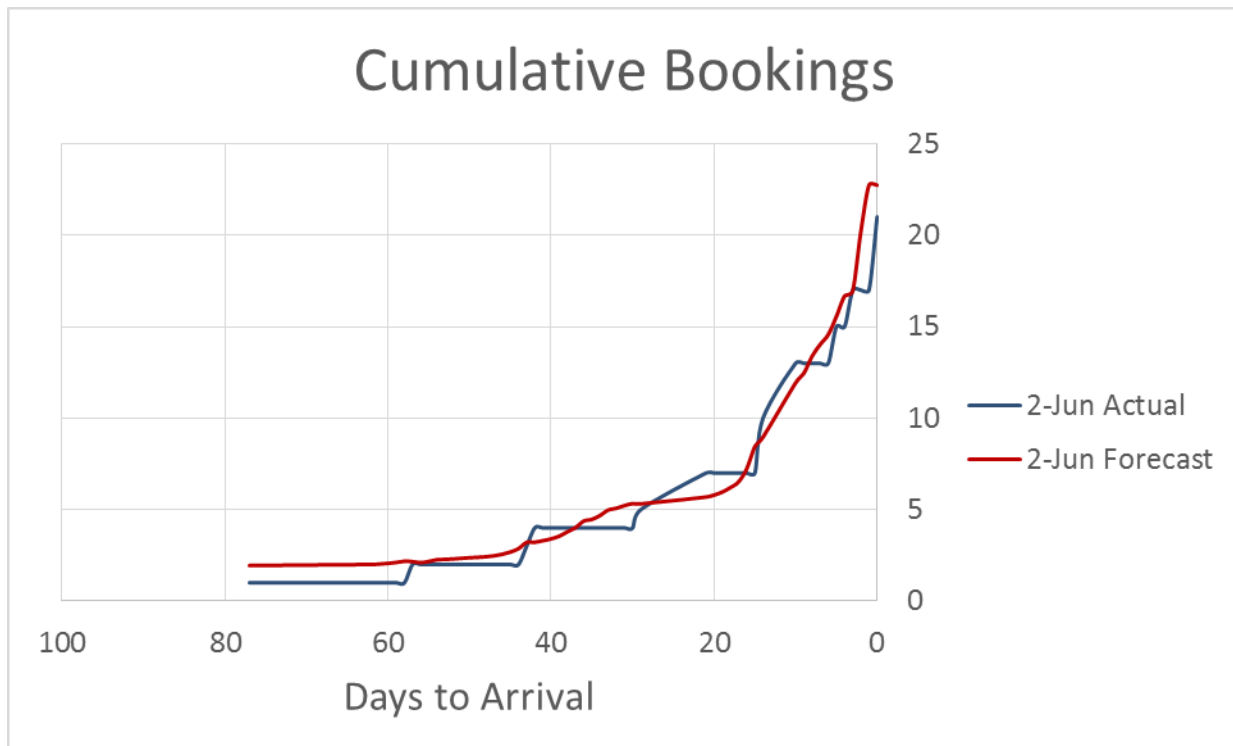


Figure 6. Booking Curve with Trend Model

This model clearly does a better job fitting the booking curves, as you can see in Table 2.

	June 2	June 9	June 16	MPE	MAPE
Regression Model	-45%	-3%	-23%	-24%	24%
State Space Trend	8%	37%	1%	16%	16%

Table 2. Booking Curve Error % on Day of Arrival

The overall error metrics—MAPE and MPE—are better with this model. In particular, the June 2 booking curve has an improved fit. The predicted cumulative bookings for June 2 has an error of 8%.

Model Run Time

This version of the model ran in 1 minute and 51 seconds.

STATE SPACE TREND AND RANDOM EFFECTS MODEL

For the trend model in the previous section, the data is segmented by customer segment, month, and day of week. A steady state model formulation easily allows the inclusion of customer segment, month, or day of week effects in the model. This could replace the need to segment the data. To test the approach, start by segmenting the data only by month and day of week. To account for different behavior between customer segments, you can include a customer-specific random effect in the steady state model.

Model Results

The fit is slightly better for the June 2 booking curve, as seen in Figure 7.

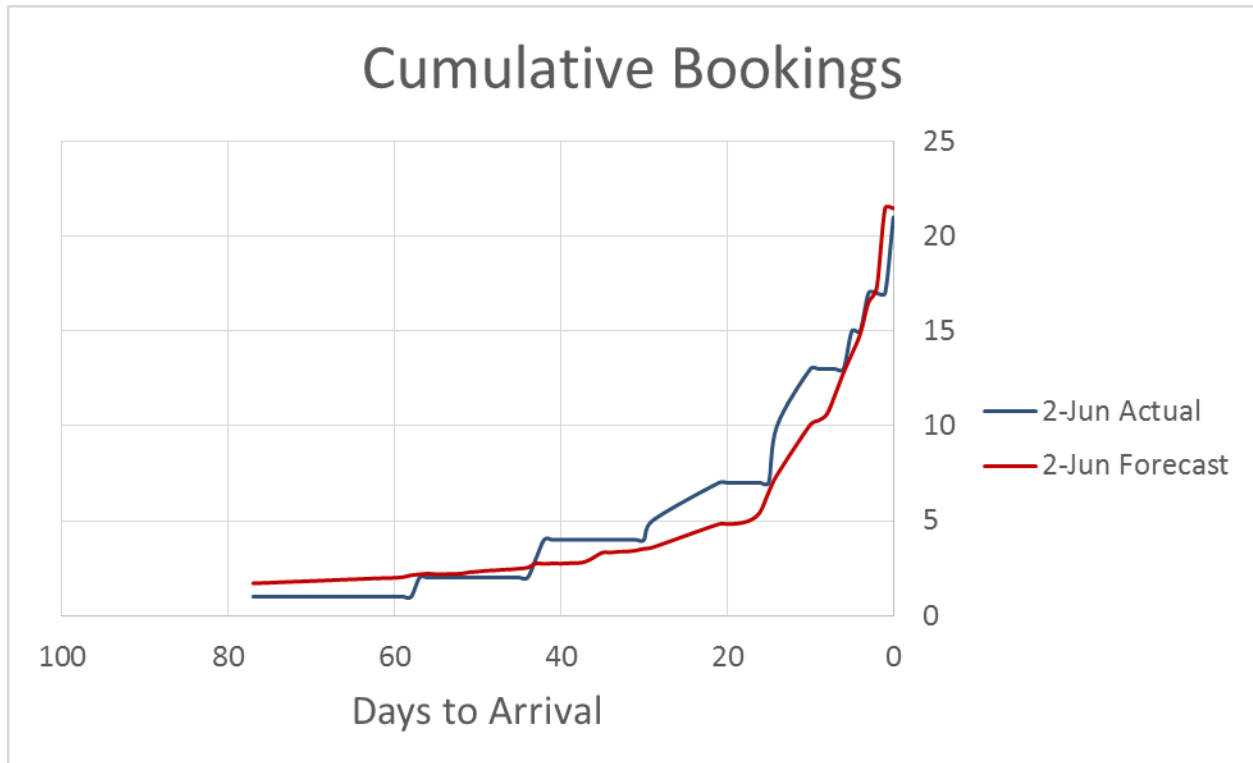


Figure 7. Booking Curve with Customer Effect

However, overall error on the day of arrival, when considering all three stay dates, is slightly worse, as shown in Table 3.

	June 2	June 9	June 16	MPE	MAPE
Regression	-45%	-3%	-23%	-24%	24%
Steady State Trend	8%	37%	1%	16%	16%
Steady State Trend + Customer Effect	-2%	68%	12%	26%	27%

Table 3. Booking Curve Error on Day of Arrival

Performance

The model with trend and the customer-specific effects took 2 minutes and 33 seconds to run.

STATE SPACE MODELS OF RETAIL INVENTORY

Inventory position is critical to retail stores. A store could use predicted inventory levels to decide when to mark products down or to plan display changes. The inventory path of one product could vary between stores. The same product can have very different inventory positions and paths in different stores, as shown in Figure 8 [Figure 8. Inventory Units by Day](#).

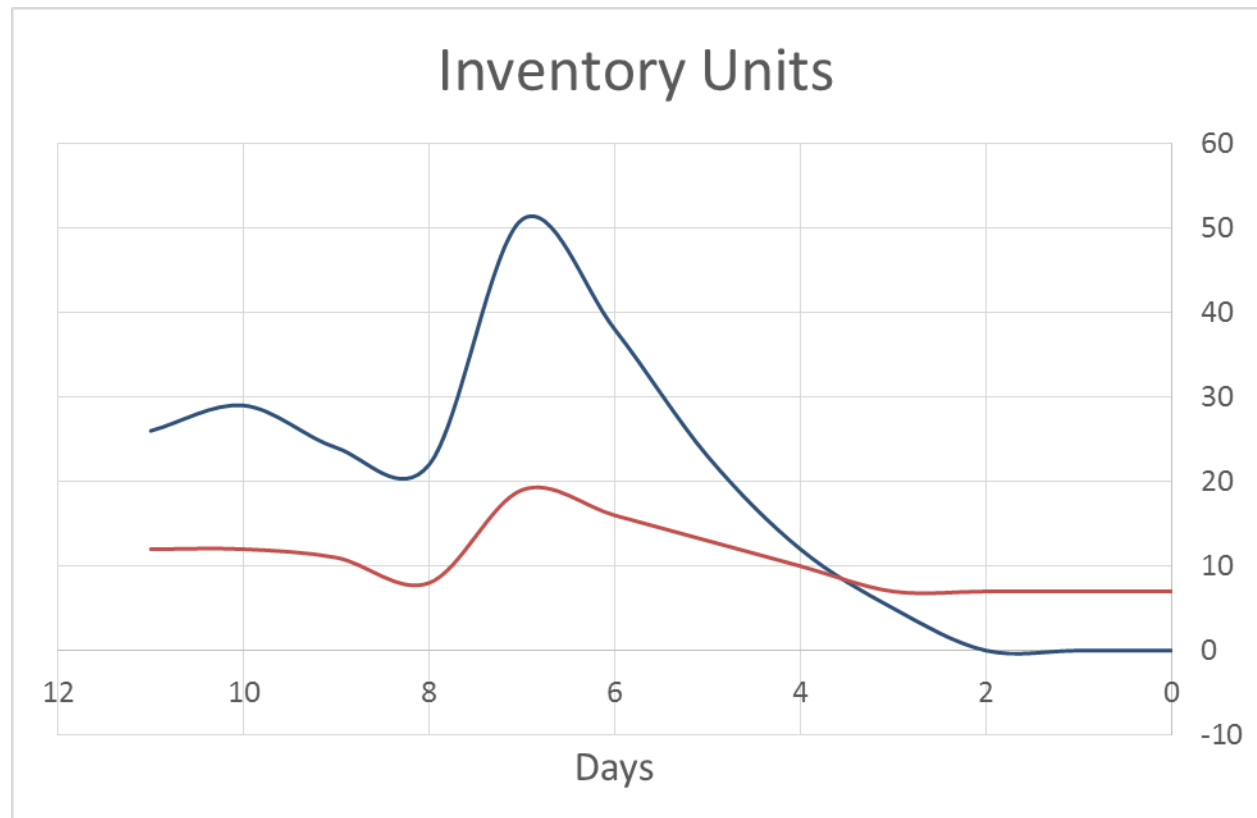


Figure 8. Inventory Units by Day

Note that both of the stores received additional inventory on Day 8, accounting for the increase in inventory. After this infusion of inventory, the inventory position declines again as expected. However, the shape for the store shown in red is much flatter than the curve of the store shown in blue. It is important to capture differences like this when planning.

REGRESSION MODEL

First, you can segment the data by product and run a regression model in PROC SSM. This model does not have a state space component. The structure of the model is:

$$I = \alpha + \beta \exp(\text{day}) + \delta \text{receipt} + \varepsilon$$

Model Results

The parameter estimates from PROC SSM for Item 1 are shown in Figure 9:

Regression Parameter Estimates					
Response Variable	Regression Variable	Estimate	Standard Error	t Value	Pr > t
CLOSE_INV_UNITS	intercept	-0.730	0.2985	-2.44	0.0145
CLOSE_INV_UNITS	RECEIPT_UNITS	5.379	1.9525	2.76	0.0059
CLOSE_INV_UNITS	price_fcn	0.881	0.2061	4.28	<.0001
CLOSE_INV_UNITS	time_to_end	0.063	0.0182	3.47	0.0005

Figure 9. PROC SSM Parameter Estimates

The same model can be run in PROC REG to yield the same parameter estimates for Item 1, as seen in Figure 10:

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-0.72952	0.29851	-2.44	0.0146
RECEIPT_UNITS	1	5.37924	1.95252	2.76	0.0059
price_fcn	1	0.88149	0.20615	4.28	<.0001
time_to_end	1	0.06301	0.01816	3.47	0.0005

Figure 10. PROC REG Parameter Estimates

The regression model captures some characteristics of the inventory curve for a sample item. It fits the shape of the increase in inventory as units are received, and then both the predicted and actual inventory decline until the end of life. This behavior can be seen in Figure 11.

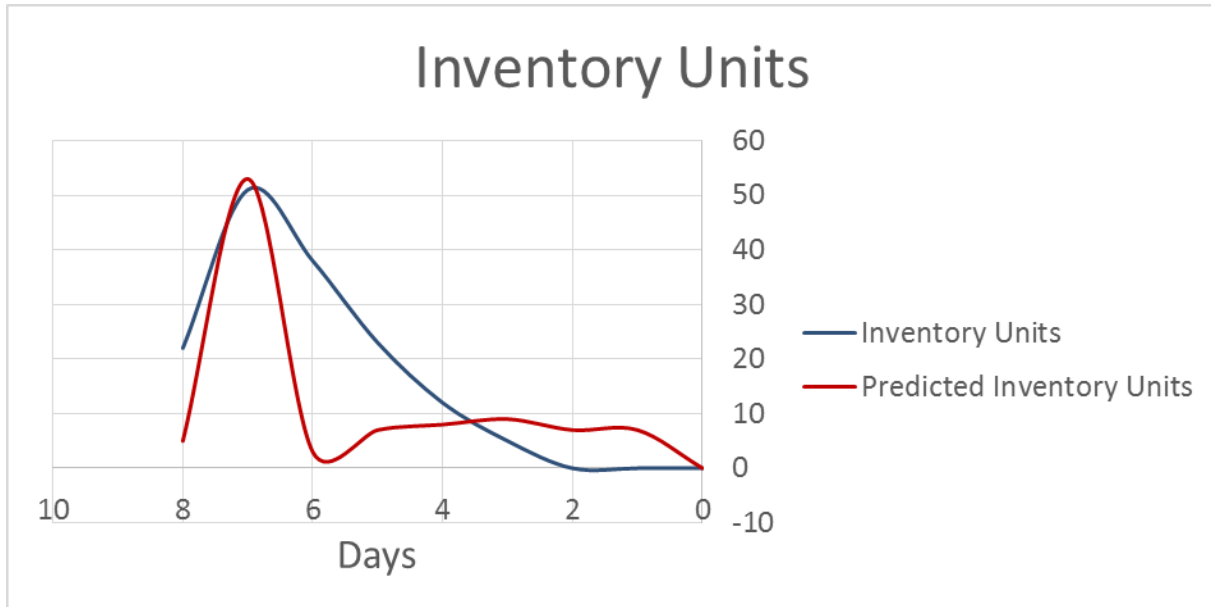


Figure 11. Inventory Curve Regression Model

Performance

The PROC SSM model ran in 25 seconds, while the PROC REG procedure took 2 minutes and 2 seconds. As demonstrated in the booking curve example, the SSM procedure has a clear performance improvement while yielding identical results to the regression procedure.

STEADY STATE MODEL WITH TREND

The PROC SSM model formulation makes it very easy to add trend to a model. After running the regression model via PROC SSM, you can define a state space model with the Ornstein-Uhlenbeck decay pattern. The model fit is slightly improved, but still does not capture the slow decline in inventory units that occurs after units are received. This can be seen clearly in Figure 12:

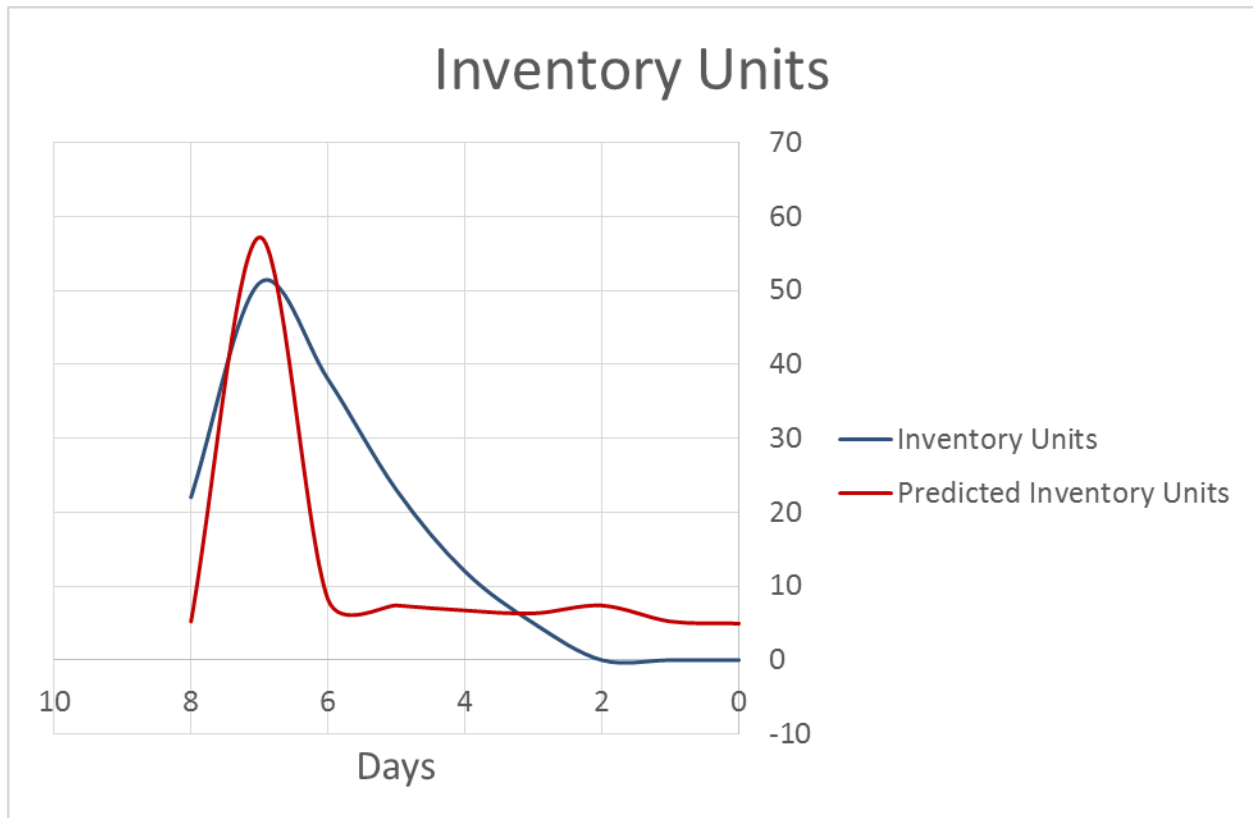


Figure 12. Inventory Curve with Steady State Trend Model

Performance

This model runs in 32 seconds. It does not take significantly longer than the regression model, when the regression model runs through PROC SSM.

CONCLUSION

The booking curve and the inventory curve examples showcase some of the capabilities of PROC SSM. The procedure's straightforward and flexible syntax makes it easy to make model variations. The common output across multiple model types allows easy comparison between models. Finally, the increased code efficiency can be a significant advantage in enterprise-level analyses.

REFERENCES

Leonard, M., J. Sloan, T. Lee, and B. Elsheimer. "An Introduction to Similarity Analysis Using SAS." Accessed March 2016. <https://support.sas.com/rnd/app/ets/papers/similarityanalysis.pdf>

SAS Institute Inc. 2013. SAS/ETS 14.1: User's Guide. Cary, NC: SAS Institute Inc. Available: <http://support.sas.com/documentation/cdl/en/etsug/68148/HTML/default/viewer.htm#titlepage.htm>.

Selukar, Rajesh. 2011. "State Space Modeling Using SAS." *Journal of Statistical Software*, 41.

ACKNOWLEDGMENTS

My thanks to Sanjeewa Naranpanawe for providing guidance on hospitality topics, Garrett Frere for helping with the coding, and Blanche Shelton for allowing me to take the time to conduct this research.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Beth Cabbage, PhD
Analytical Consulting Manager
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
919-531-9491
Email: beth.cabbage@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.