ANALYSIS AND FORECAST OF THE STORAGE NEEDS OF SOYBEANS IN BRAZIL

MARCO T. O. PATINO\textsuperscript{1}, MARINA F. MACHADO\textsuperscript{2}, GERALDO T. DO NASCIMENTO\textsuperscript{3}, MILLA R. DE ALCANTARA\textsuperscript{4}

ABSTRACT: The existence of a minimum storage capacity of grains as a condition for the maintenance of regulator physical stocks has been used as a strategic factor in the agribusiness expansion. However, in Brazil the storage infrastructure has not followed the growth of the agricultural sector. This fact is evident in the case of soybeans that currently represent 49\% of grain production in the country, whose volume production has been increasing significantly over the years. This study aimed to predict the future needs of static storage capacity of soybeans from historical data to estimate the investment needed to install storage units in Brazil for the next five years. A statistic analysis of collected data allowed a forecast and identification of the number of storage units that should be installed to meet the storage needs of soybeans in the next five years. It was concluded that by 2015 the soybean storage capacity should be 87 million tons, and to store 49\% of soybeans produced, 1,104 storage units should be installed at a cost of R$ 442 million.

KEYWORDS: agribusiness, Glycine max, temporal series, harvest seasons.

INTRODUCTION

Soybean is one of crops with greater economic significance in Brazil. According to data from the National Food Supply Company - CONAB (2012), it currently corresponds to 49\% of the cultivated areas of grains in the country and, for the 2011-2012 season is expected to produce 75.31 million tons with an average productivity of 3,000kg ha\textsuperscript{-1}. Projections indicate for the harvest of 2019/2020 a production of 81.95 million tons, an increase of 2.86\% per year. (MAPA, 2010)
Due to its importance for the Brazilian agriculture, soybean is the agricultural product that generates more export volume (in tons) to the country and makes it the second largest producer worldwide, requiring a planned and suitable logistics infrastructure for storing the product (PONTES et al., 2009).

The grain storage and maintenance of regulator physical stocks have often been used as a strategic factor in the expansion of the agricultural business. According to the Food and Agriculture Organization (FAO, 2011), static storage capacity of a country should be 1.2 times your annual agricultural production. According to the eighth grain survey of CONAB (2012), the production is expected to be 160,060,000 tons in 2011/2012 harvest, while the Brazilian static storage capacity should be 192 million tons. However, this static storage capacity registered in Brazil is currently 142.51 million tons, showing a deficit of 49.49 million tons, showing that the storage infrastructure for agricultural products in the country has not matched the growth of production in this sector (CONAB, 2012).

TORERO & VON BRAUN (2011) explain that even though the difference between the minimum and adequate stocks of grains is relatively small, the lack of sufficient stocks can lead to significant price increases and distortions in the functioning of markets. The difference in world grain stocks for 2004-05 and 2007-08 was about 60 million tons or 2.7% of global production, but with the sharp rise in prices in 2007-08, this difference in grain stocks combined with the increase in prices was enough to cause serious problems in the market.

The use of predictive models for agricultural policy analysis has allowed the examination of governmental mechanisms of price stabilization in the context of private producers and storage companies (SUMNER et al., 2010). Facing the increasing production of soybeans in the country, there is the need to provide proper storage of grains, enabling a strategic infrastructure for the expansion of agribusiness. Additionally, despite all the technology available to the Brazilian agriculture, losses during the process of post-harvest of the grains are not under control, and, during storage, the grain mass is constantly subjected to external factors that can change the quality or cause spoilage of the product (FARONI et al., 2009).

On the other hand, in Brazil, grains harvested in the crop are immediately taken to storage units located outside the farms, which creates additional problems due to the transportation. Cases like these do not happen in other soybean producing countries, like Canada and Argentina, where farmers have storage units and even storage about 80% and 25% of the volume harvested each season, respectively (JASPER et al., 2006) and in Brazil, about 13% have this infrastructure storage location (TRAMONTINA et al., 2008). It is noteworthy that the storage of grain on farms provides better conservation and marketing conditions and lower costs, with consequent effects on profitability of farmers (CONAB, 2006).

The storage units play an important role in agribusiness, especially regarding the flow of grain crops and supply policies (FREDERICO, 2010), but Brazil has serious deficiencies in this sector because grain production in the country is higher than the static storage capacity. In Brazil, the lack of storage units within farms is a major problem in the storage structure of the country (SOUSA JUNIOR et al., 2011); since it forces farmers to sell their products immediately after harvest with lower market prices (TEFERA et al., 2011).

In addition to the strategic issue, storage also contributes to a more efficient marketing of products, more competitive sector against foreign competitors, control of quantities traded in the market, as well as being an essential element in the policy of storage of commodities for food security (MARTHA JUNIOR, 2010). In this sense it is necessary to plan future storage needs and define the infrastructure investments needed to allow a normal market behavior.

Therefore, this study aimed to estimate the future needs of static storage capacity of soybeans from historical data to estimate the investment needed to install storage units in Brazil for the next
five years and validate the prediction model used. These estimates can be used as important information on trading the futures market and also to plan applicable policies in the area.

**MATERIAL AND METHODS**

This study was divided into four stages. The first consisted of collecting data on specialized sites such as the National Food Supply Company (CONAB) and the Ministry of Agriculture, Livestock and Supply (MAPA).

In the second step, statistical analyzes were performed using data from spreadsheets and the Minitab software. The test of normality was used to observe the behavior of the data series studied and to determine if the data follow a normal distribution. Normalization is necessary to avoid the appearance of too large numbers, thus making them unintelligible. In this study the normal probability plot and the Anderson-Darling test belonging to the class of quadratic statistics based on the empirical distribution function (EDF) were used, since they use the squared differences between the empirical and the hypothetical distributions (STEPHENS, 1986).

The forecasting techniques tested to find the best adjustment were the ARIMA model, which represents the more general class of models for time series analysis, and Winter additive, multiplicative, exponential and simple exponential methods. Winter’s methods describe appropriately demand data where there is the linear trend, besides a seasonal component, characterized by the occurrence of cyclic variation patterns that are repeated at relatively constant intervals of time, typical of the study.

By analyzing one or more series, the graphical representation of data sequentially over time is critical and can reveal important patterns of behavior, such as growth or reduction trends, cyclical patterns, structural changes, aberrant observations, among others (SOUZA, 2008).

The models used to describe a time series use stochastic processes, or processes controlled by probabilistic laws. Among the goals of time series analysis, according to MORETTIN & TOLOI (2004), the investigation of the mechanism generating the series is highlighted; conducting forecasts of future values of short and long term, describing the behavior of the series, with graphical verification of trends, cycle and seasonal variations and searching for periodicity in the data. In this context, Winter’s method is an exponential smoothing method that takes into account the seasonal components of the series of observed data. To format the adjustments, Winter’s method employs, at each reporting period, three components: level (equation 1), trend (equation 2) and seasonality (equation 3). These components are updated in each period for the smoothing parameters, and the initial values for the level and trend components are also used, which are obtained from a time linear regression (equation 4) in accordance with FARIA (2009). The initial values for the seasonal component are obtained from the regression of the dummy variable using rectified data.

\[
L_t = \alpha (Y_t / S_{t-p}) + (1 - \alpha)[L_{t-1} + T_{t-1}] \tag{1}
\]

\[
T_t = \gamma [L_t - L_{t-1}] + (1 - \gamma) * T_{t-1} \tag{2}
\]

\[
S_t = \delta (Y_t / L_t) + (1 - \delta) * S_{t-p} \tag{3}
\]

\[
\hat{Y}_t = (L_{t-1} + T_{t-1}) * S_{t-p} \tag{4}
\]

where:

- \(L_t\) – is time t level, \(\alpha\) the weight for the level;
- \(T_t\) – is the trend in time t, \(\gamma\) is the weight for the trend;
- \(S_t\) – is the seasonal time in time t, \(\delta\) is the weight of the seasonal component;
- \(p\) - is the seasonal period;
- \(Y_t\) – is the value of data in time t;
- \(\hat{Y}_t\) - is the estimate value, or a period after a forecast, in time t.
To compare the predictive models tested and select the one with the best adjustment, the three precision measures provided by the Minitab software were used: MAPE (Mean Absolute Percentage Error), which measures the accuracy of the values of equivalent series of time and it is shown as a percentage, MAD (Mean Absolute Deviation), which accurately measures the time series of the values supplied, shown in the same units as the data, which helps to conceptualize the amount of error, and MSD (Mean Squared Deviations), which measures the accuracy of the serial values of the time given and it is calculated using the same denominator (the number of predictions), regardless of the model, allowing a comparison of the values of MSD in different models, and, therefore, comparing the accuracy between models.

From the analysis of data normality, in the third step, it was possible to forecast for the next five years the static storage capacity of soybeans, comparing methods and adopting the best adjustment.

In the last stage of the study, an estimate was made of the required amount of storage soybean units, assuming that 100% of the soybeans produced in Brazil will be stored in the next five years. In this calculation, it was considered that each unit had static storage capacity of 30,000 tons and consisted of dryers, pre-cleaning and cleaning machines and an elevator, having an installed cost of approximately R$400,000. This value was supplied by D’Andrea Agrimport, manufacturer of storage units (DANDREA, 2011).

RESULTS AND DISCUSSION

From the statistical data collected, it was possible to estimate the percentage of participation of the static capacity of soybeans in relation to the total static capacity of grains in Brazil in the period of 2000-2010 (Table 1) in order to identify the storage requirements if 100% of the soybeans were stored.

TABLE 1. Values of annual production grains and soybeans and respective static storage capacity.

<table>
<thead>
<tr>
<th>Year</th>
<th>PRODUCTION OF GRAINS</th>
<th>PRODUCTION OF SOYBEANS</th>
<th>STATIC GRAIN CAPACITY</th>
<th>%*</th>
<th>STATIC SOYBEAN CAPACITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t)</td>
<td>(t)</td>
<td>(t)</td>
<td></td>
<td></td>
<td>(t)**</td>
</tr>
<tr>
<td>2000</td>
<td>100,266,877</td>
<td>38,431,800</td>
<td>87,833,000</td>
<td>43.8%</td>
<td>38,646,520</td>
</tr>
<tr>
<td>2001</td>
<td>96,799,000</td>
<td>42,230,000</td>
<td>89,277,000</td>
<td>47.3%</td>
<td>41,960,190</td>
</tr>
<tr>
<td>2002</td>
<td>123,168,000</td>
<td>52,017,500</td>
<td>89,734,200</td>
<td>58.0%</td>
<td>52,045,836</td>
</tr>
<tr>
<td>2003</td>
<td>119,114,200</td>
<td>49,792,700</td>
<td>93,358,600</td>
<td>53.3%</td>
<td>49,480,058</td>
</tr>
<tr>
<td>2004</td>
<td>114,695,000</td>
<td>52,304,600</td>
<td>100,056,000</td>
<td>52.3%</td>
<td>52,029,120</td>
</tr>
<tr>
<td>2005</td>
<td>122,530,783</td>
<td>55,027,100</td>
<td>106,056,000</td>
<td>51.9%</td>
<td>55,149,120</td>
</tr>
<tr>
<td>2006</td>
<td>131,750,600</td>
<td>58,391,800</td>
<td>121,987,700</td>
<td>47.9%</td>
<td>59,554,096</td>
</tr>
<tr>
<td>2007</td>
<td>144,137,300</td>
<td>60,017,700</td>
<td>123,300,000</td>
<td>48.7%</td>
<td>60,417,000</td>
</tr>
<tr>
<td>2008</td>
<td>135,134,500</td>
<td>57,165,500</td>
<td>125,700,000</td>
<td>45.5%</td>
<td>56,565,000</td>
</tr>
<tr>
<td>2009</td>
<td>149,254,900</td>
<td>68,688,200</td>
<td>135,700,000</td>
<td>50.6%</td>
<td>69,207,000</td>
</tr>
<tr>
<td>2010</td>
<td>161,535,471</td>
<td>72,227,800</td>
<td>137,412,920</td>
<td>52.6%</td>
<td>71,828,848</td>
</tr>
</tbody>
</table>

*Percentage of soybeans in the grain total static capacity; ** Estimating that 100% of the soybeans are stored.

In Table 1 it can be seen that in 2000 the production of grains was 100.26 million tons, and soybeans represented 38% of this value. Thus, to store all soybean production of this year, an available static capacity corresponding to 44% of total capacity was necessary. From these historical data it is possible to see that 2002 was the most critical year in terms of storage requirements, because soy represented 58% of the total capacity of storage, i.e., from the 89.73 million tons, 52.04 million were intended only to soybean storage. In 2009 the total static capacity...
was of 135.7 million tons, and, if 100% was stored, 69.20 million tons of total static capacity would be used.

To identify the most appropriate way to conduct the process of predicting the static capacity of soybeans, a graphic analysis was performed, through the normality test, since the number of observations is below 50. Figure 1 shows a normal probability curve which is formed by the set of points by a straight line and the estimated confidence interval, which has the function to represent the data storage capacity, and determining the reliability of the study. As can be seen in this figure, it is found that the distribution fits the data and thus the smallest statistic has the best fit with the descriptive level AD = 0.889 (calculated from the Anderson-Darling statistic), indicating that the data have a fitted distribution.

The null hypothesis is only rejected if the test provide a value lesser than the significance level, i.e., if \( p < \alpha \). In this study, it was adopted alpha = 0.05 and, therefore, the data follow a normal distribution, since the descriptive level (P-value) is equal to 0.020.

Confirmed the hypothesis of normality of the data sample, a comparison of forecasting models was carried out. Initially the ARIMA model was tested, but the results did not show the 25 interactions required for the analysis of optimal prediction.

Winter’s methods are recommended to smooth the data in a time series prediction and provide short or medium range. These methods were chosen because the series has presented the data with trend and seasonal pattern and provide a meaningful solution, considering the magnitude of the seasonal pattern is proportional to the data, consistent with the model when the seasonal features a double exponential relationship with the trend (FARIA, 2009).

Thus, a comparison was made with the models, Winter Additive, Winter’s Multiplicative, Winter’s smoothed Simple Exponential and Winter’s smoothed Double Exponential methods. The smoothed Double Exponential model was the one that showed better results (Table 2).
TABLE 2. Winter’s methods used.

<table>
<thead>
<tr>
<th>METHOD</th>
<th>MAPE</th>
<th>MAD</th>
<th>MSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive</td>
<td>6.619</td>
<td>3.6437</td>
<td>2.331</td>
</tr>
<tr>
<td>Multiplicative</td>
<td>6.659</td>
<td>3.6437</td>
<td>2.331</td>
</tr>
<tr>
<td>Simple Exponential</td>
<td>7.332</td>
<td>4.063</td>
<td>3.063</td>
</tr>
<tr>
<td>Double Exponential</td>
<td>5.387</td>
<td>3.062</td>
<td>1.491</td>
</tr>
</tbody>
</table>

The Winter’s Double Exponential method was used because it had the best explanation of the amplitude of seasonal variation, which can increase or decrease depending on the time.

The use of Winter’s method allowed the prediction of static storage capacity of soybeans and, with these data, Table 3 was elaborated. It shows the forecast values of static capacity of storage required over the next five years, to store 100% of soybeans produced in the country. These values were calculated using production and the static storage capacity data.

TABLE 3. Forecast of static storage capacity of soybeans for the next five years (tons).

<table>
<thead>
<tr>
<th>Period</th>
<th>Prevision</th>
<th>Percentage/year</th>
<th>Accumulated percentage</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>72,483,688</td>
<td>0%</td>
<td>0%</td>
<td>64,981,305</td>
<td>79,986,072</td>
</tr>
<tr>
<td>2013</td>
<td>75,560,342</td>
<td>4.2%</td>
<td>4.2%</td>
<td>67,464,871</td>
<td>83,655,813</td>
</tr>
<tr>
<td>2014</td>
<td>78,636,996</td>
<td>4.0%</td>
<td>8.2%</td>
<td>69,886,561</td>
<td>87,387,431</td>
</tr>
<tr>
<td>2015</td>
<td>81,713,650</td>
<td>3.9%</td>
<td>12.1%</td>
<td>72,259,226</td>
<td>91,168,074</td>
</tr>
<tr>
<td>2016</td>
<td>84,790,304</td>
<td>3.7%</td>
<td>15.8%</td>
<td>74,593,013</td>
<td>94,987,595</td>
</tr>
</tbody>
</table>

According to the estimates shown in Table 3, the static capacity growth over the five years under study is 15.8%, with an expectation for 2012 of 72.48 million tons of static capacity, with a range of 64.98 and 79.98 million tons. Similarly, the period of 2013, has a range of 67.46 and 83.65 million tons, while for the year of 2014, according to this study, 78.63 million tons are predicted, which limits are between 69.88 and 87.38 million tons. For the year of 2015, a total of 81.71 million tons was predicted, with 95% of confidence that the estimates will be between the limit with, a minimum of 77.580 million and maximum of 94.980 million and, finally, for the year of 2016 it was predicted 84.79 million tons, with limits of 74.59 and 94.98 million tons.

These results show the trend of increasing storage needs and complement the results of NOGUEIRA JUNIOR & TSUNECHIRO (2010), who found an increase of 77% in grain production between the years 2000 to 2010, while the static storage capacity had an increase of 52.6% during the same period, a deficit of 24.4 million tons.

Figure 2 was created with data from the predictions, and shows the values of the exponentially weighted averages using the standard 0.2 smoothing constant (level, trend and seasonal). This figure shows the current curve, the adjustment curve and prediction curve with its confidence interval. The accuracy of measurements of the MAPE, MAD and MSD were 5.387, 3.062 and 1.491 respectively. According to the software, it indicates a proper adjustment to the model.
Similarly, projections made by the Ministry of Agriculture, Livestock and Supply - MAPA (2010) showed a growth of 2.86% in soybean production until 2019/2020 that, if implemented, will determine a production of 75.06 million tons of soybeans in 2015.

FAO (2011) recommends that the static storage capacity of a country should be 1.2 times your annual agricultural production. Using this parameter, the optimal capacity for storing all soybeans produced in 2015 is 90.07 million, but according to the prediction carried out, the static storage capacity of soybeans for this year is 87.16 million tons, therefore a deficit of 3.23% in storage capacity.

The lack of investment in storage infrastructure can jeopardize the growth of agribusiness, since a lower storage capacity forces the producer to flow the production immediately after its harvest, and according to TRAMONTINA et al., (2008), only 13% of static capacity of grain storage lies within the farms.

Given this deficient panorama of grain storage capacity in Brazil, a projection of the static storage capacity of soybeans in Brazil was carried out, estimating that, for the next five years (2012-2016), there is going to be an increase of 13% to 49% (nine percent per year) in storage capacity (Table 4). In this projection it was considered that 100% of soybeans produced in the country are stored.

<table>
<thead>
<tr>
<th>Year</th>
<th>Projection of Capacity (100%)</th>
<th>% of Soybeans Stored</th>
<th>Capacity</th>
<th>Additional demand</th>
<th>Unit Numbers*</th>
<th>Cost (thousand)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>72,483,688</td>
<td>13</td>
<td>9,422,879</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2013</td>
<td>75,560,342</td>
<td>22</td>
<td>16,623,275</td>
<td>7,200,396</td>
<td>240</td>
<td>96,000</td>
</tr>
<tr>
<td>2014</td>
<td>78,636,996</td>
<td>31</td>
<td>24,377,468</td>
<td>7,754,193</td>
<td>258</td>
<td>103,200</td>
</tr>
<tr>
<td>2015</td>
<td>81,713,650</td>
<td>40</td>
<td>32,685,460</td>
<td>8,307,992</td>
<td>277</td>
<td>110,800</td>
</tr>
<tr>
<td>2016</td>
<td>84,790,304</td>
<td>49</td>
<td>41,547,248</td>
<td>8,861,788</td>
<td>295</td>
<td>118,000</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>428,000</td>
</tr>
</tbody>
</table>

* One storage unit is equivalent to 30 thousands of tons of grains.
From Table 4 it can be seen that for the year of 2013 are needed 240 new storage units to store 22% of soybeans produced, representing an investment of R$96 million, considering R$400 thousand as the price of each unit. For the next five years, a total investment of R$428 million is needed to increase the static storage capacity of 13% to 49%. This is not a significant investment, taking into account the value generated by soybean agribusiness, estimated by CEPEA (2012) of around R$36 billion of reais per year (BARROS et al., 2011).

These results can be achieved with proper planning of grain production storage on the farm, which for D’ARCE (2004), when done well, has significant advantages, such as minimizing production losses, savings on transportation, higher yield at harvest, better product quality and obtaining financing through credit lines specific for pre-marketing. Similarly, a study by DAMBROSIO et al. (2009) indicated that it is more economically advantageous for producers to store their product in their own warehouse, within the farm, than storing in an outsourced warehouse.

A further analysis indicates that with an average yield of 3.0 t ha\(^{-1}\) (CONAB, 2011), a storage unit of 30 thousand tons and properties of 125 hectares, each soybean storage unit will serve 10,000 hectares or 80 properties of this size.

**CONCLUSION**

The use of the forecasting model by the Winter’s method was validated and the results confirm the lack of investment in grain storage infrastructure in Brazil, which over the years has not matched the rapid growth in grain production. Projections indicate that in the period of 2012-2016, in the case of soybeans is necessary to build 1,070 units with capacity of 30 thousand tons each, to increase the static storage capacity of 13% to 49%.

**REFERENCES**


