Forecasting weekly Canary tomato exports from annual surface data

Gloria Martín Rodríguez (gmartinr@ull.es)
José Juan Cáceres Hernández (jcaceres@ull.es)
University of La Laguna

Selected Poster prepared for presentation at the 28th International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguaçu, Brazil, 18-24 August, 2012

Forecasting weekly Canary tomato exports from annual surface data

Gloria Martín Rodríguez (gmartinr@ull.es)

José Juan Cáceres Hernández (jcaceres@ull.es)

Department of Economía de las Instituciones, Estadística Económica y Econometría

University of La Laguna

España

Abstract

Sea shipping is the main transport mode used by Canary farmers to export tomatoes to the European markets. Provincial associations of Canary growers negotiate charter fees with the shipping companies for the whole exporting period and, therefore, provide a unified sea transport service. When such a negotiation takes place each year, the individual growers' decisions about planting surface are usually known. However, the forecasting of the distribution of tomato exports over the whole harvesting period would help Canary associations make more timely and effective decisions. In this paper, a model is proposed to forecast weekly Canary tomato exports conditioned on a given total planting surface.

A seasonal model is formulated to deal with the weekly seasonal pattern of Canary tomato yields per hectare by means of evolving splines. Such a model is a useful tool to forecast weekly yields. From these forecasts, weekly tomato exports beyond the end of the sample are also forecast by taking the total planting surface into account. To illustrate the aptness of this framework, the proposed methodology is applied to a weekly series of tomatoes exported to the European markets from 1995/1996 to 2010/2011 harvests.

JEL classification: C22, Q17

Key words: tomato exports, surface, weekly data, seasonal splines.

1. Introduction

European markets are the main destination of Canary tomato exports. During the last harvests, some exporters have made the decision of sending tomatoes to mainland Spain by ship and, from there, they cover the distance to the final destination by lorry. However, the transport mode usually chosen by farmers is sea shipping from Canary Islands to the destination markets. In fact, the cost of sea transport to the European ports is a major part of the distribution costs up to the point of sale at wholesale markets. Currently, charter fees, insurance, and port operations —including loading and handling at the port of origin, and handling and unloading at the destination port — cost about 1.15 euros per box of six kg. These estimates reveal that insularity and distance to markets have a devastating effect on the competitiveness of Canary farms.

Each year, in an attempt to reduce such a cost, ACETO (Provincial association of Tenerife tomato growers) and FEDEX (Provincial federation of associations of vegetables and fruits exporters from Las Palmas) hire ships to transport tomatoes from Canary Islands to the ports of Rotterdam and Southampton during the harvesting period, from October to May. The charter fees are conditioned by both the number of ships hired and the frequency of shipments over the whole harvest, and the carrying capacity is negotiated to guarantee that production satisfying quality standard is exported. However, once a charter fee for a sea shipment is fixed, the final freight charge per kilogram rises when the weekly export volume is noticeably lower than the capacity of the ship warehouse. Therefore, before the beginning of a harvest, there is a need to forecast weekly exports over the whole harvesting period.

At the point in time in which the forecast is needed, farmers have usually made their decision about planting surface, because this information is required of them for insurance companies. Of course, the Canary tomato production is strongly related to the planting surface, which has fallen dramatically during the last years. Due to yields per hectare are not expected to change so suddenly, the distribution of weekly yields over a coming harvest could be more accurately forecast from its evolution observed in the past. In this way, once the planting surface is known for the coming harvest, the forecasts of weekly exports can be obtained from the forecasts of weekly yields.

Such a two-step procedure is proposed in this paper. In the next section, a seasonal model is formulated to forecast weekly time series. Third section deals with the application of such a model to forecast weekly yields of Canary tomato per hectare. Forecasts of weekly exports are also obtained. Finally, section 4 states the concluding remarks.

2. Seasonal model

Weekly data tend to show a pseudo-periodic seasonal behaviour, as the length of the seasonal period does not remain the same over time. Weekly prices of an agricultural commodity are only registered when the product is on the market, and the marketing period may change from year to year depending on climate or managerial strategies. Moreover, movements in the weekly series of supply for such a product over the marketing period are relevant from an economic point of view, whereas the figures corresponding to the non-marketing period do not provide any information other than that concerning the length of such a period.

To forecast this type of data, conventional models have serious difficulties in capturing seasonal patterns. In fact, the application of such models usually requires that some original values of the series be deleted (Martín-Rodríguez and Cáceres-Hernández, 2005) or that some artificial missing values be included (Harvey and Koopman, 1993; Cabrero et al., 2009) or eliminated by aggregation to a lower sampling frequency (Jumah and Kunst, 2008) to get a constant seasonal period. When, from time to time, no observations correspond to some of the seasons that belong to the seasonal period or, above all, in time series in which the length of the seasonal period does not remain the same over time, a method is needed to provide forecasts of seasonal effects whatever the season in coming seasonal periods. From this point of view, conventional models are too rigid because they may possibly not provide a forecast at each of these seasons. Furthermore, disturbance terms are assumed to drive seasonal movements in the series to capture changes in the magnitude of seasonal effects over the sample. However, such disturbance terms are expected to be null. Therefore, the forecast of the seasonal effect at a season is strongly conditioned by the adjusted value of the seasonal effect at the same season in the last observed seasonal period, but the trend movement in this value is left aside. This drawback is more noticeable when long-range predictions are needed to make economic decisions.

To overcome such limitations, there is a need to reflect on the concept of seasonal pattern. If the seasonal pattern is regular, such that the seasonal effect at a point in time is assumed to depend on the proportion of the whole seasonal period at this point in time, the seasonal component can be formulated as a function of such a proportion, which means that the periodicity of the seasonal pattern is measured in a different scale than the usual one. As Martín-Rodríguez and Cáceres-Hernández (2010) point out, this new approach provides the seasonal model with a noticeable gain in flexibility and parsimony.

In fact, when the seasonal effect at any season is assumed to be a function of a proportion that takes values into the continuous interval (0,1), a parametric formulation such as a spline function (Poirier, 1976; Eubank, 1988) is a useful tool to model and forecast the seasonal pattern. Note that such a formulation can capture regular seasonal patterns. The piecewise polynomial functional form and the periodicity and continuity restrictions that conventional splines are required to satisfy provide such a model with a degree of regularity. On the other hand, splines make feasible the definition of different seasonal models for different sample periods; therefore, enough flexibility is provided to capture seasonal patterns likely to evolve over time instead of fitting a fixed pattern.

By means of splines, the seasonal effect at any proportion of the seasonal period length can be expressed as a function of the seasonal effects at specific proportions. In this section, a parametric formulation of the changes in seasonal effects at these specific proportions is provided as a useful tool to forecast the seasonal pattern in weekly series. The aim is not to forecast the exact values of the seasonal effect at a season, but the whole shape of the seasonal pattern. The new formulation captures trends that drive the dynamic process of change in the shape of the seasonal pattern. Therefore, a function that can describe such a pattern in the near future is obtained. Once the seasonal pattern is described as a function of the proportion of the seasonal period length, the forecast of the seasonal effect at any season can be calculated once two assumptions are made: one, about the length of the seasonal period, and the other, about the season in which the seasonal variation begins.

In a weekly time series model formulated as

$$y_t = \mu_t + \gamma_t + \varepsilon_t \,,\ t = 1,\ldots,T \,, \tag{1}$$

where μ_t and γ_t are the trend or level component and the seasonal component, respectively, and ε_t is the irregular component, the seasonal pattern can be assumed to be completed in a period whose length does not remain the same over time.

Time series can be divided into as many sub-samples as complete seasonal periods are registered. In this way, a seasonal pattern is completed in the set of weeks that define one of these m sub-periods. Let s_c be the number of weeks in seasonal period c, $c=1,...,m$. Let γ_t be defined as $\gamma_t = \gamma_{c,w}$ if the observation at time t and sub-period c corresponds to week j_c in such a way that $w = \frac{j_c}{s_c}$. Note that w is the proportion of the seasonal period elapsed up to week j_c, $j_c = 1,..., s_c$, and, therefore, $0 < w \le 1$.

In the next sub-sections, the phases of building a model for this type of seasonal variations are developed and the way such a model is able to provide seasonal pattern forecasts is explained. Firstly, an evolving periodic cubic spline is formulated to capture changes in the shape of the seasonal pattern over time. In this way, the evolution of the seasonal effects at break points, which define the shape of the seasonal pattern, can be observed. Secondly, a non-periodic spline is formulated to describe the changes in each of these seasonal effects. Thirdly, the parametric formulation of such a dynamic process is introduced into the original evolving spline. Finally, the parametric formulation developed in previous sub-sections provides forecasts of seasonal effects at the break points, and, from these forecasts the whole seasonal pattern is forecast. Of course, forecasts of time series can also be obtained from both the seasonal and the level component forecasts.

2.1. Evolving spline without constraints between sub-periods

If seasonal variation is assumed to change in a smooth manner from one season to the following, a periodic cubic spline is a suitable model for these variations. That is,

$$\gamma_{c,w} = g_c(w) + \xi_{c,w}, \tag{2}$$

where $\xi_{c,w}$ is a residual term and $g_c(w)$ is a third-degree piecewise polynomial function,

$$g_c(w) = g_{c,i}(w) = g_{c,i,0} + g_{c,i,1}w + g_{c,i,2}w^2 + g_{c,i,3}w^3,\ w_{c,i-1} \le w \le w_{c,i},\ i=1,...,k, \tag{3}$$

where $w_{c,0} = 0$ and $w_{c,k} = 1$. Note that the number of segments is assumed to be fixed for each of the m sub-periods. The continuity of the spline function and of its first and second derivatives are enforced by the following conditions:

$$g_{c,i,0} + g_{c,i,1} w_{c,i} + g_{c,i,2} w_{c,i}^{\ 2} + g_{c,i,3} w_{c,i}^{\ 3} = g_{c,i+1,0} + g_{c,i+1,1} w_{c,i} + g_{c,i+1,2} w_{c,i}^{\ 2} + g_{c,i+1,3} w_{c,i}^{\ 3}, \tag{4.a}$$

$$g_{c,i,1} + 2g_{c,i,2} w_{c,i} + 3g_{c,i,3} w_{c,i}^{\ 2} = g_{c,i+1,1} + 2g_{c,i+1,2} w_{c,i} + 3g_{c,i+1,3} w_{c,i}^{\ 2}, \tag{4.b}$$

$$2g_{c,i,2} + 6g_{c,i,3} w_{c,i} = 2g_{c,i+1,2} + 6g_{c,i+1,3} w_{c,i}, \tag{4.c}$$

for $i = 1,...,k-1$. Furthermore, by assuming that the spline is natural, the following conditions are obtained:

$$2g_{c,1,2} + 6g_{c,1,3} w_{c,0} = 0, \tag{5.a}$$

$$2g_{c,k,2} + 6g_{c,k,3} w_{c,k} = 0. \tag{5.b}$$

Then, following a similar procedure to that developed in Martín-Rodríguez and Cáceres-Hernández (2010), the spline $g_c(w)$ can be expressed as a linear function

$$g_c(w) = g_{c,1,0} X^{g}_{c,1,0,w} + g_{c,1,1} X^{g}_{c,1,1,w} + g_{c,2,0} X^{g}_{c,2,0,w} + ... + g_{c,k,0} X^{g}_{c,k,0,w}, \tag{6}$$

where $X^{g}_{c,1,0,w}, X^{g}_{c,1,1,w}, X^{g}_{c,2,0,w},..., X^{g}_{c,k,0,w}$ are appropriate functions of proportion w and break points $w_{c,i} = w_i$, $i = 0,...,k$, and $g_{c,1,0}, g_{c,1,1}, g_{c,2,0} ..., g_{c,k,0}$ are free parameters to be estimated. Note that the locations of the break points are assumed to remain the same for each of the m sub-periods. The previous specification is flexible enough to capture a seasonal pattern in which γ_t evolves over time. The seasonal pattern in the m sub-periods into which the series is divided can be jointly modelled as $\gamma_t = g(t) + \xi_t$, where $g(t)$ is the evolving spline

$$g(t)=\sum_{c=1}^{m}\left[g_{c,1,0}X_{1,0,t}^{g}+g_{c,1,1}X_{1,1,t}^{g}+g_{c,2,0}X_{2,0,t}^{g}+\ldots+g_{c,k,0}X_{k,0,t}^{g}\right]D_{c,t}^{sp}\ , \qquad (7)$$

where $D_{c,t}^{sp}=\begin{cases}1, & t\in sub-period\ c\\ 0, & in\ other\ case\end{cases}$, $c=1,\ldots,m$, and $X_{1,1,t}^{g}=X_{c,1,1,w}^{g}$ and $X_{i,0,t}^{g}=X_{c,i,0,w}^{g}$, $i=0,\ldots,k$, if the observation at time t and sub-period c corresponds to week j_c in such a way that $w=\dfrac{j_c}{s_c}$.

The seasonal pattern in sub-period c, $c=1,\ldots,m$, is assumed to satisfy the condition

$$\sum_{i=1}^{k}\int_{w_{i-1}}^{w_i}\left(g_{c,i,o}+g_{c,i,1}w+g_{c,i,2}w^2+g_{c,i,3}w^3\right)dw=0\ . \qquad (8)$$

That is to say, the area under the spline function is equal to zero over the whole seasonal period. Once estimates of parameters $g_{c,i,j}$, $c=1,\ldots,m$, $i=1,\ldots,k$, $j=0,1,2,3$, are obtained, if such a condition is not satisfied in such a way that

$$\sum_{i=1}^{k}\int_{w_{i-1}}^{w_i}\left(\hat{g}_{c,i,o}+\hat{g}_{c,i,1}w+\hat{g}_{c,i,2}w^2+\hat{g}_{c,i,3}w^3\right)dw=A(c), \qquad (9)$$

estimates of seasonal effects $\gamma_{c,w}$ in sub-period c should be corrected by subtracting the constant term $A(c)$.

To identify the changes in the shape of the curve that describes the seasonal variation over the sample, an alternative formulation in a similar sense to Koopman (1992) and Harvey et al. (1997) would be useful. However, given that γ_t is defined as $\gamma_t=\gamma_{c,w}$, the seasonal variation at any proportion w can be expressed as a function of the values of seasonal effects at specific proportions $w_{c,i}=w_i$, $i=1,\ldots,k$. In this sense, the spline $g_c(w)$ is expressed as

$$g_c(w) = \gamma_{c,w_0} X^{\gamma}_{c,0,w} + \ldots + \gamma_{c,w_k} X^{\gamma}_{c,k,w}, \tag{10}$$

where $X^{\gamma}_{c,0,w}, \ldots, X^{\gamma}_{c,k,w}$ are appropriate functions of proportion w and break points w_i, $i = 0,\ldots,k$, and $\gamma_{c,w_0}, \ldots, \gamma_{c,w_k}$ are free parameters to be estimated.

Again, the previous specification is flexible enough to capture a seasonal pattern in which γ_t evolves over time. The seasonal pattern in the m sub-periods into which the series is divided can be jointly modelled as $\gamma_t = g(t) + \xi_t$, where $g(t)$ is the evolving spline

$$g(t) = \sum_{c=1}^{m} \left[\gamma_{c,w_0} X^{\gamma}_{0,t} + \ldots + \gamma_{c,w_k} X^{\gamma}_{k,t} \right] D^{sp}_{c,t}, \tag{11}$$

where $D^{sp}_{c,t} = \begin{cases} 1, & t \in sub-period\ c \\ 0, & in\ other\ case \end{cases}$, $c = 1,\ldots,m$, and $X^{\gamma}_{i,t} = X^{\gamma}_{c,i,w}$, $i = 0,\ldots,k$, if the observation at time t and sub-period c corresponds to week j_c in such a way that $w = \frac{j_c}{s_c}$.

If $\gamma_{c,w_i} = \gamma_{w_i}$, $i = 0,\ldots,k$, $c = 1,\ldots,m$, occurs, the model in Equation (11) provides a valuable gain in parsimony with regard to traditional models in which the seasonal variation is defined as a function of the season. However, these parameters could also be assumed to evolve over time, even according to parametric formulations. This being the case, a noticeable gain in parsimony is also feasible.

2.2. Splines for modelling changes in seasonal effects at the break points over the sub-periods

Let $\{\gamma_{c,w_i}\}_{c=1,\ldots,m}$, $i = 0,\ldots,k$, be the sets of estimates of the seasonal variation at a season in which the proportion of the seasonal period elapsed up to such a season is w_i, $i = 0,\ldots,k$, for each of the m sub-periods into which the series is divided. If seasonal variations at proportion w_i are assumed to change in a smooth manner from one sub-period to the following, a non-periodic cubic spline is a suitable model for each one of these variations. That is,

$$\gamma_{c,w_i} = g_i(c) + \xi_{i,c}\,, \tag{12}$$

where $\xi_{i,c}$ is a residual term and $g_i(c)$ is a third-degree piecewise polynomial function,

$$g_i(c) = g_{i,j}(c) = g_{i,j,0} + g_{i,j,1}c + g_{i,j,2}c^2 + g_{i,j,3}c^3\,,\ c_{i,j-1} \le c \le c_{i,j}\,,\ j = 1,...,r\,, \tag{13}$$

where $c_{i,0} = 1$ and $c_{i,r} = m$. Following a similar procedure to that developed in subsection 2.1, the estimates $\{\gamma_{c,w_i}\}_{c=1,...,m}$, $i = 0,...,k$, can be expressed as

$$g_i(c) = \gamma_{i,0} Y^{\gamma}_{i,0,c} + ... + \gamma_{i,r} Y^{\gamma}_{i,r,c}\,, \tag{14}$$

where $Y^{\gamma}_{i,0,c},...,Y^{\gamma}_{i,r,c}$ are appropriate functions of sub-period c and break points $c_{i,j} = c_j$, $j = 0,1,...,r$, and $\gamma_{i,0},...,\gamma_{i,r}$ are free parameters to be estimated, which represent the seasonal effects at the proportion w_i of the seasonal period in sub-periods $c_0, c_1,...,c_r$. Note that the locations of break points are assumed to remain the same for each one of the splines $g_i(c)$.

To forecast the values of the seasonal effects at break points, an alternative formulation can also be obtained in such a way that the estimates $\{\gamma_{c,w_i}\}_{c=1,...,m}$, $i = 0,...,k$, are expressed as a linear function

$$g_i(c) = g_{i,1,0} Y^{g}_{i,1,0} + g_{i,1,1} Y^{g}_{i,1,1} + g_{i,2,0} Y^{g}_{i,2,0} + ... + g_{i,r,0} Y^{g}_{i,r,0}\,, \tag{15}$$

where $Y^{g}_{i,1,0}, Y^{g}_{i,1,1}, Y^{g}_{i,2,0}, Y^{g}_{i,r,0}$ are appropriate functions of sub-period c and break points $c_0, c_1,...,c_r$, and $g_{i,1,0}, g_{i,1,1}, g_{i,2,0} ..., g_{i,r,0}$ are free parameters to be estimated.

2.3. Evolving spline with constraints between sub-periods

From the formulation in sub-section 2.2, the parameters $\{\gamma_{c,w_i}\}_{c=1,\dots,m}$, $i=0,\dots,k$, could also be assumed to evolve over time according to the parametric model corresponding to the functions in Equation (14). That is, if parameter γ_{c,w_i} is expressed as $\gamma_{c,w_i}=g_i(c)=\gamma_{i,0}Y^{\gamma}_{i,0,c}+\dots+\gamma_{i,r}Y^{\gamma}_{i,r,c}$, then the evolving spline in Equation (11) can be written as a function of parameters $\{\gamma_{i,0},\dots,\gamma_{i,r}\}_{i=0,\dots,k-1}$, as follows:

$$\begin{aligned} g(t)=\ &\gamma_{0,0}U_{0,0,t}+\gamma_{0,1}U_{0,1,t}+\dots.+\gamma_{0,r}U_{0,r,t}+ \\ &\gamma_{1,0}U_{1,0,t}+\gamma_{1,1}U_{1,1,t}+\dots.+\gamma_{1,r}U_{1,r,t}+\dots+ \\ &\gamma_{k,0}U_{k,0,t}+\gamma_{k,1}U_{k,1,t}+\dots.+\gamma_{k,r}U_{k,r,t} \end{aligned} \tag{16}$$

where $U_{i,j,t}=\left[\sum_{c=1}^{m}Y^{\gamma}_{i,j,c}D^{sp}_{c,t}\right]X^{\gamma}_{i,t}$, $i=0,\dots,k$, $j=0,\dots,r$. Therefore,

$$g(t)=\sum_{i=0}^{k}\sum_{j=0}^{r}\gamma_{i,j}U_{i,j,t}\ . \tag{17}$$

Note that $\gamma_{i,j}$ is the seasonal variation at proportion w_i of the seasonal period corresponding to the break point located at sub-period c_j. Therefore, the number of parameters to be estimated is equal to $(k+1)*(r+1)$.

It can be demonstrated that $\sum_{i=0}^{k}\sum_{j=0}^{r}U_{i,j,t}=1,\forall t$. Therefore, once one of these regressors is deleted, the formulation in Equation (17) can be introduced into a time series model to estimate conjointly the seasonal component and the remainder of components in the original series. These estimates could be introduced as a dependent variable in Equation (7). Then, from the results of estimating such a model, seasonal effects are properly corrected to satisfy Equation (8).

2.4. Forecasting seasonal patterns

From estimating the time series model, estimates of seasonal variations are obtained. The results of estimating a spline for $\{\gamma_{c,w_i}\}_{c=1,...,m}$ in terms of Equation (15) are useful to obtain a forecast of $\{\gamma_{m+h,w_i}\}_{h=1,...,}$, $i = 0,...,k$. Then, the forecasts of the seasonal pattern h sub-periods ahead are obtained, according to Equation (10), as

$$g_{m+h}(w) = \gamma_{m+h,w_0} X^{\gamma}_{m+h,0,w} + ... + \gamma_{m+h,w_k} X^{\gamma}_{m+h,k,w}, \tag{18}$$

where $X^{\gamma}_{m+h,0,w},..., X^{\gamma}_{m+h,k,w}$ are appropriate regressors defined as indicated in first phase. Note that a possibly different seasonal pattern is forecast for each sub-period *m+h*, as a function of the dynamic process of change in parameters $\gamma_{c,w_0},...,\gamma_{c,w_k}$. That is, features of the seasonal pattern over the sample help in predicting changing seasonal patterns beyond the end of the sample. To correct the seasonal pattern obtained from the estimates of $\{\gamma_{m+h,w_i}\}_{h=1,...,}$, the following alternative formulation, in terms of Equation (6), is also useful:

$$g_{m+h}(w) = g_{m+h,1,0} X^{g}_{m+h,1,0,w} + g_{m+h,1,1} X^{g}_{m+h,1,1,w} + g_{m+h,2,0} X^{g}_{m+h,2,0,w} + ... + g_{m+h,k,0} X^{g}_{m+h,k,0,w}, \tag{19}$$

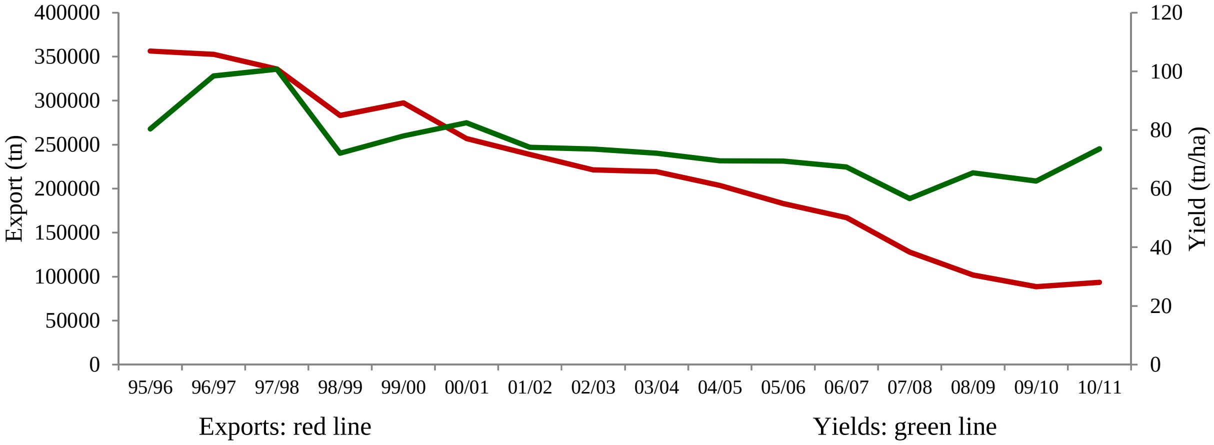
where $X^{g}_{m+h,1,0,w}, X^{g}_{m+h,1,1,w}, X^{g}_{m+h,2,0,w},..., X^{g}_{m+h,k,0,w}$ are appropriate regressors defined as indicated in first phase. Given that the seasonal variation is defined as a function of the proportion of the whole seasonal period elapsed up to the season, the seasonal period length must be selected to obtain forecasts of seasonal effects at specific seasons.

3. An application to weekly series of tomato exports

The usefulness of the methodological proposal is illustrated by using weekly series of Canary tomato exports to the European markets between the 1995/1996 and 2010/2011 harvests. Export data are provided by ACETO and FEDEX. The search for profitability has led growers to concentrate exports in the winter period (Cáceres-Hernández, 2000, 2001; Martín-Rodríguez and Cáceres-Hernández, 2005). Therefore, tomatoes are not usually exported for some weeks, especially during the summer period. However, the length of the harvest has evolved over time not only as a consequence of the agronomic characteristics of the tomato plants but also as a result of changes in the trade rules regulating the accession of agricultural production to the European markets.

The relevant seasonal variation is defined during the marketing period. Therefore, missing values located in the summer period have been deleted. In fact, missing data in the summer period do not provide relevant economic information beyond that derived from the length of the non-marketing period, which is provided implicitly by the modelled seasonal variation. Thus, a new series is obtained and is referred to as $\{x_t\}_{t=1,\ldots,625}$. From the export data, weekly exports per hectare are obtained. The new yield series is referred to as $\{y_t\}_{t=1,\ldots,625}$.

Figure 1. Evolution of Canary tomato exports (tonnes) and yields (tonnes per hectare)



Source: ACETO, FEDEX, and Consejería de Agricultura, Ganadería y Pesca del Gobierno de Canarias.

As shown in Figure 1, the Canary export volumes have dropped sharply last decade as a result of the impact on its competitiveness of other supplies and the consequent diminishing of the planting surface. However, tomato exports per hectare are likely to move around a slower varying level. Therefore, the forecasting of the weekly yields from its evolution observed in the past is expected to be more accurate than the forecasting of the unstable long term movements in the export series. So, the seasonal model proposed in the statistical section is applied to the weekly yield series, in which, as shown in Table 1, the length of the seasonal period oscillates noticeably.

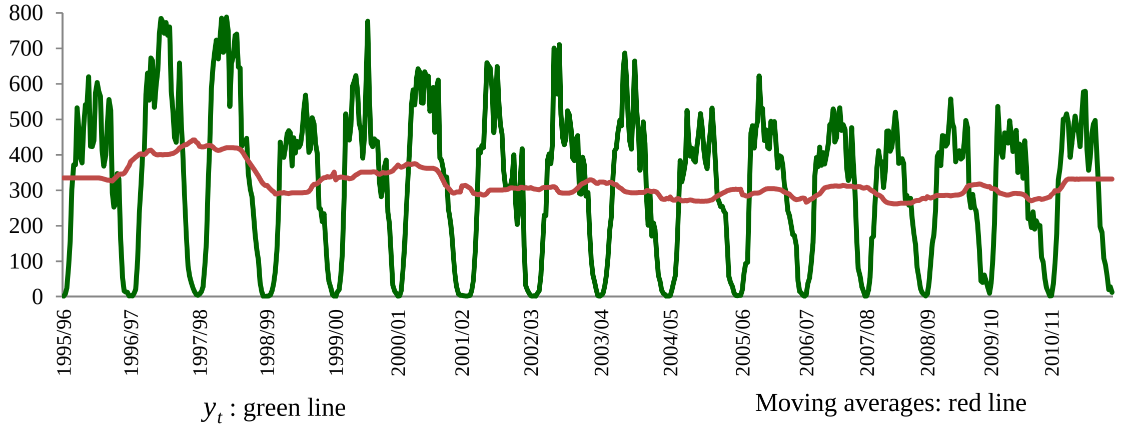
Table 1. Exporting period by harvest

Harvest	Period	Harvest	Period	Harvest	Period	Harvest	Period
1995/96	39-26 (40)	1999/00	39-23 (37)	2003/04	39-27 (41)	2007/08	41-24 (36)
1996/97	39-27 (41)	2000/01	39-24 (38)	2004/05	39-28 (43)	2008/09	41-26 (38)
1997/98	39-26 (40)	2001/02	34-22 (41)	2005/06	40-25 (38)	2009/10	42-24 (36)
1998/99	38-25 (41)	2002/03	35-24 (42)	2006/07	40-23 (36)	2010/11	41-25 (37)

Note: The exporting period is indicated by the beginning and ending weeks. The length of the seasonal period is enclosed in brackets. Due to the presence of leap years in the sample, there are observations corresponding to week 53 in years 1998, 2004, and 2009.

Some stylized facts about yields per hectare are worth noting. In Figure 2 an increasing phase is always observed that begins in October and finishes in January or February, followed by a decreasing phase that continues until May or June. That is, the seasonal pattern is more or less regular, with maximum and minimum levels almost always located around the same proportion of the length of the harvesting period. To model these movements around the trend level, seasonal variations could be formulated by means of spline functions.

Figure 2. Weekly yield series (boxes of 6 kg per hectare)



To this end, the first step consists of estimating an approximation of seasonal effects. In this sense, an approximation of the long-term movement can be extracted by calculating moving averages with period s_c, corresponding to observations at a seasonal period whose length is s_c. Then, the deviations from the original series to the moving average series are previous estimates of the seasonal variations, $\{\hat{\gamma}_t^1\}_{t=1,\ldots,625}$, from which the number and locations of the break points are selected. To obtain a more parsimonious formulation, the break points are assumed to be the same for all harvests. These points are also assumed to belong to the set $\left\{\frac{l}{1000}\right\}_{l=1,\ldots,1000}$, which divides the continuous interval $(0,1)$ into short enough subintervals. From the results of estimating a regression model in terms of Equation (7), a six-segment spline has been specified. The number of break points is selected so that the spline captures the main changes in the shape of the observed seasonal pattern, whereas the set of locations of these points is the one that provides the best fit[1].

[1] The locations of the break points are: $w_1 = 0.148$, $w_2 = 0.177$, $w_3 = 0.267$, $w_4 = 0.794$, and $w_5 = 0.902$.

According to the results of estimating the parametric models in Equation (7) or (11), when parameters are assumed to evolve over the sample, new approximations of seasonal variations are obtained. Once these estimates are corrected in such a way that the area under the spline function is equal to zero over each harvest, new estimates $\{\hat{\gamma}_t^2\}_{t=1,\ldots,625}$ are obtained (Figure 3). The previous estimates of the seasonal variations do not show a common pattern over the sample. According to the methodological section, the evolution of the seasonal effect at a specific proportion of the seasonal period can be modelled by means of a non-periodic cubic spline. A two-segment spline has been selected to model such an evolution of the seasonal effects at each break point. When these constraints are enforced, a new formulation of the original evolving splines, in terms of Equation (17), is obtained:

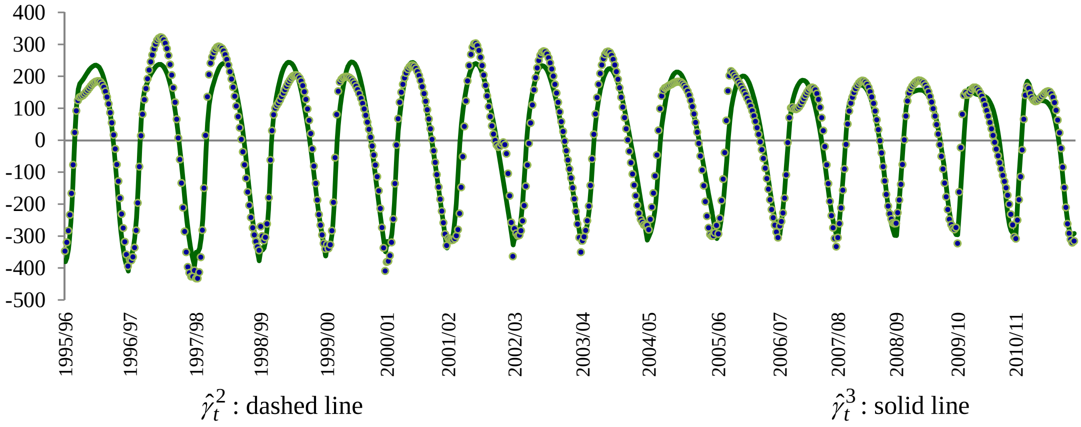
$$g(t) = \sum_{i=0}^{6} \sum_{j=0}^{2} \gamma_{i,j} U_{i,j,t} , \tag{20}$$

where $\gamma_{i,j}$ is the seasonal effect at proportion w_i in the seasonal period corresponding to harvest c_j. These hypotheses about the seasonal pattern can be introduced into a structural model (Harvey, 1989) formulated as

$$y_t = \mu_t + \sum_{i=0}^{5} \sum_{j=0}^{2} \gamma_{i,j} U_{i,j,t} + \sum_{j=0,1} \gamma_{6,j} U_{6,j,t} + \varepsilon_t , \tag{21}$$

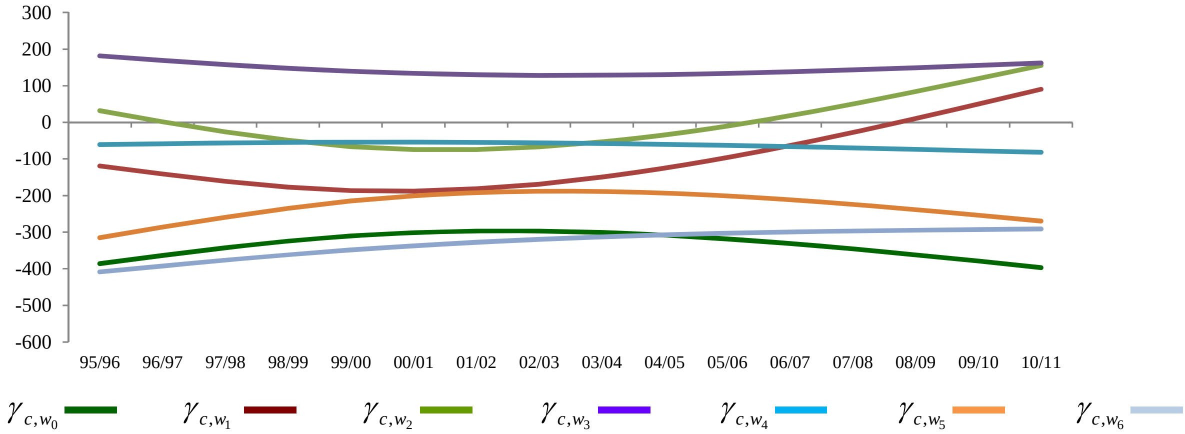
where the level component is modelled as a stochastic level. Note that one of the regressors $U_{i,j,t}$ is deleted to avoid multicollinearity problems. The estimates of seasonal variations from the structural time series model, $\{\hat{\gamma}_t^3\}_{t=1,\ldots,625}$ (Figure 3), have been corrected in such a way that the area under the spline function over each harvest is equal to zero. The correction applied to the estimates of seasonal variation has been taken into account to correct the estimates of the long-term component, which are not shown to save space.

Figure 3. Estimates of seasonal effects over the sample



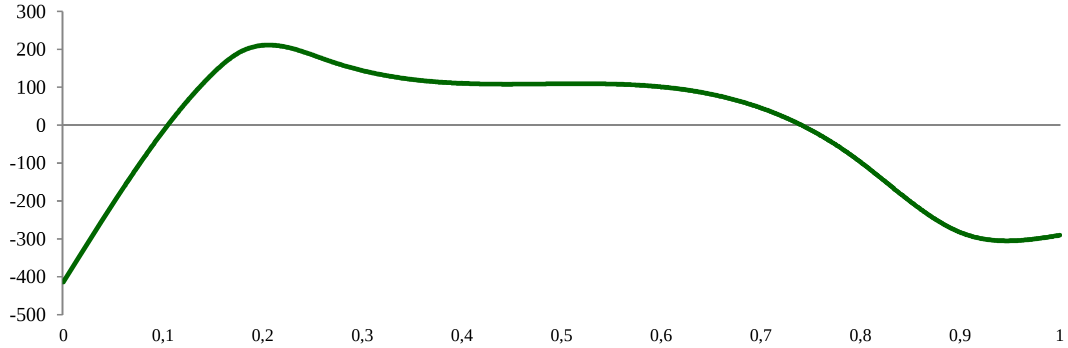
The estimates $\hat{\gamma}_t^3$ do not match exactly the seasonal pattern of each harvest defined in terms of estimates $\hat{\gamma}_t^2$ over the sample. Besides of that, the most important remark is that estimates $\hat{\gamma}_t^3$ show the direction in which the shape of the whole seasonal pattern is changing. This behaviour can also be explained from the estimates of seasonal effects at the break points into the seasonal period shown in Figure 4.

Figure 4. Estimates of seasonal effects at break points



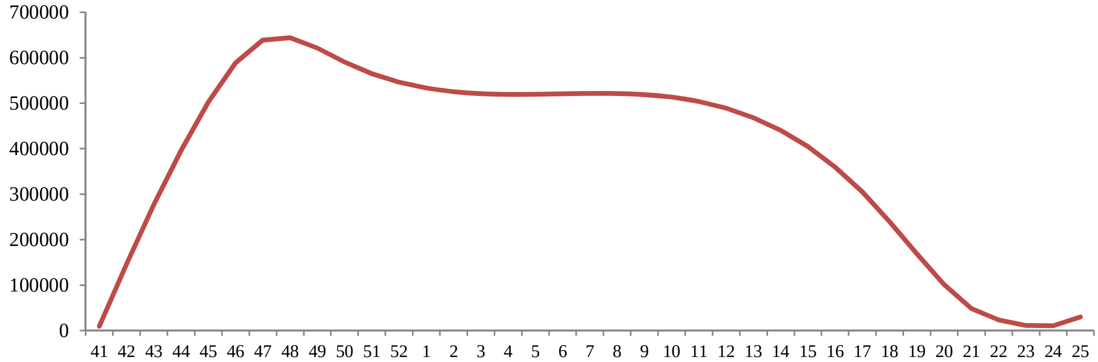
In this way, the final seasonal model, noticeably parsimonious, provides a formulation useful to forecast seasonal effects at the proportions of the seasonal period corresponding to the break points in coming harvests. Then, such a model makes feasible the forecasting of the seasonal effect at each proportion of the seasonal period. Figure 5 shows the forecasts of these seasonal effects corresponding to the harvest 2011/2012.

Figure 5. Forecasts of the seasonal effects at any proportion in weekly yield series



The forecasting of seasonal effects at specific seasons is conditioned by the assumptions about both the length of the seasonal period and the beginning week. In the coming harvest, the beginning week and the length of the seasonal period are assumed to be the same as those observed in the harvest 2010/2011. On the other hand, predicted values of the trend component are obtained according to the estimates of parameters of a cubic spline fitted to the stochastic level. Then, the weekly yields and, finally, the export forecasts shown in Figure 6 are also obtained.

Figure 6. Forecasts of export series at any week



4. Concluding remarks

Canary tomato growers make some decisions before the beginning of a harvest to match highest sea transport capacity periods with highest production periods. To this aim, there is a need to forecast weekly exports of Canary tomatoes for a coming harvest. In this paper, a method is proposed to provide this type of forecasts from data about the total planting surface in such a harvest. At a first step, the method provides forecasts of pseudo-periodic seasonal patterns in weekly yields. Then, forecasts of weekly export levels are also obtained by taking the total planting surface into account.

The proposed method provides noticeable advantages as regard to conventional models, which are not suitable enough for modelling or forecasting this type of seasonal patterns, in which the length of the seasonal period does not remain the same over time. The proposed formulation provides the model with enough flexibility to capture the shape of the seasonal pattern whatever the number of weeks of the seasonal period. Furthermore, the model proposed in this paper can capture changes in seasonal effects over the sample by means of a parametric formulation of the dynamic process of change in the shape of the seasonal pattern. Therefore, one of the virtues of the approach taken in this article is parsimony, but, above all, the method can provide forecasts for the seasonal effect at whatever proportion of the continuous interval in which the seasonal period is completed. In this way, series forecasts can be obtained whatever the specific seasons in coming seasonal periods. From this point of view, the method proposed in this paper is shown to be a useful tool to forecast weekly agronomic production series, or even daily or hourly agricultural price series.

Acknowledgements

The authors wish to express our gratitude to Domingo Mendoza of ACETO for his efforts to provide us with useful statistical information. This research would not have been possible without the financial support by the Canary Islands Government in the AGROEXPCAN research project.

References

Cabrero, A., Camba-Méndez, G., Hirsch, A. and Nieto, F., 2009. Modelling the daily bank notes in circulation in the context of the liquidity management of the European Central Bank. *Journal of Forecasting* **28**, 194-217.

Cáceres-Hernández, J.J., 2000. *La Exportación de Tomate en Canarias. Elementos para una Estrategia Competitiva*. Ediciones Canarias, La Laguna.

Cáceres-Hernández, J.J., 2001. Optimalidad del patrón estacional de las exportaciones canarias de tomate. *Estudios de Economía Aplicada* **18**, 41-66.

Eubank, R.L., 1988. *Spline Smoothing and Nonparametric Regression*. M. Dekker, New York.

Harvey, A.C., 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Harvey, A.C., Koopman, S.J., 1993. Forecasting hourly electricity demand using time varying splines. *Journal of the American Statistical Association* **88**, 1228-1236.

Harvey, A.C., Koopman, S.J., Riani, M., 1997. The modelling and seasonal adjustment of weekly observations. *Journal of Business and Economic Statistics* **15**, 354-368.

Jumah, A., Kunst, R.M., 2008. Seasonal prediction of European cereal prices: good forecasts using bad models? *Journal of Forecasting* **27**, 391-406.

Koopman, S.J., 1992. *Diagnostic Checking and Intra-Daily Effects in Time Series Models*. Thesis Publishers Tinbergen Institute Research Series **27**, Amsterdam.

Martín-Rodríguez, G., Cáceres-Hernández, J.J., 2005. Modelling weekly Canary tomato exports. *Agricultural Economics* **33**, 255-267.

Martín-Rodríguez, G., Cáceres-Hernández, J.J., 2010. Spline and the proportion of the seasonal period as a season index. *Economic Modelling* **27**, 83-88.

Poirier, D.J., 1976. *The Econometric of Structural Change with Special Emphasis on Spline Functions*. North Holland Publishing Company, Amsterdam.