revista GEPROS

Editor Responsável: Prof. Dr. Hermes Moretti Ribeiro da Silva

Recebido:
26/01/2020
Aprovado:
18/08/2020

# Hybrid model for sales forecasts in a context of the consolidation of new products on the market: an application in the agro-industrial sector

Guilherme Arcoverde Wanderley[1], Federal University of Pernambuco, Pernambuco, Brazil

Heitor Oliveira Duarte[2], Federal University of Pernambuco, Pernambuco, Brazil

## RESUMO

**Objetivo -** Este artigo objetiva avaliar o desempenho do processo de consolidação de um produto recém-inserido comercialmente em portfólio através de um modelo de previsão de vendas que, focado em diversos segmentos da carteira de clientes, sirva como suporte à tomada de decisão comercial.
**Desenho / metodologia / abordagem –** A abordagem apresentada utiliza a metodologia ABC para delimitar a análise em segmento de relevância, e, em seguida, faz uma integração entre dois modelos: (i) modelos markovianos, em tempo discreto, de passo anual, para abordar o comportamento de incorporação do novo produto em substituição aos já consolidados em portfólio; (ii) suavização exponencial, em primeira e segunda ordens, para abordar a evolução temporal da demanda agregada do conjunto de produtos. O modelo foi aplicado em uma empresa distribuidora de sementes, tomando como base sua carteira de clientes e histórico de faturamento entre 2011 e 2019.
**Resultados** – O cenário *benchmark* de previsão de vendas, projetado para condições de estabilidade no comportamento de aderência ao novo produto e estratégia comercial, resultou em um suporte quantitativo para direcionamento e acompanhamento dos esforços comerciais no sentido de maximizar o desempenho macro deste processo.
**Originalidade / valor –** Além de apresentar uma nova aplicabilidade de Cadeias de Markov na área de gestão comercial, o modelo desenvolvido introduz na literatura uma ferramenta quantitativa para direcionamento de processos de gestão da carteira de clientes no contexto de substituição no portfólio de produtos.
**Palavras-chave** - Curva ABC. Cadeias de Markov. Previsão de Vendas. Suavização Exponencial. Gestão da Carteira de Clientes.

## *ABSTRACT*

***Purpose –*** *This paper aims to evaluate the performance of the consolidation process of a product recently included commercially in a portfolio through a sales forecasting model that, focused on several segments of the customer portfolio, supports the commercial decision-making.*
***Design/methodology/approach –*** *This approach uses the ABC Curve methodology to define the analysis in segments of relevance and then integrates two methods: (i) the Markovian models in discrete time and annual step to predict the transition behavior between the new replacement product and the consolidated ones; (ii) first-order and second-order exponential smoothing time series forecasting method to predict products in aggregate demand. This model was applied to a seed distributor company based on its customer portfolio and historical data for sales between 2011 and 2019.*
***Findings –*** *The sales forecasting benchmark scenario, designed for stable conditions in the behavioral adherence of new products and the commercial strategy, resulted in a quantitative support for targeting and monitoring commercial efforts to maximize the global performance for this process.*
***Originality/value*** *– Besides presenting a new Markov Chains commercial management approach, the model developed introduces a quantitative tool into the literature for targeting the customer portfolio management processes in the context of replacement in a product portfolio.*
***Keywords -*** *Curve ABC. Markov Chains. Sales Forecasting. Exponential Smoothing. Customer Portfolio Management.*

1.Rua Doutor Edgar Valois, 211 ,55612-550 Vitória de Santo Antão, PE, guilherme.a.wanderley@gmail.com, https://orcid.org/0000-0002-0032-7666; 2. heitorod@gmail.com; https://orcid.org/0000-0001-5870-5562
WANDERLEY, G.A.; DUARTE, H.O. Hybrid model for sales forecasts in a context of the consolidation of new products on the market: an application in the agro-industrial sector. **GEPROS. Gestão da Produção, Operações e Sistemas**, v.16, nº 2, p. 01 - 38, 2021.
**DOI**: http://dx.doi.org/10.15675/gepros.v16i2.2693

## 1. INTRODUCTION

In scenarios of high market competitiveness, a fast and efficient product development process has become a priority for companies. Treating this process according to a systematic methodology with an interdisciplinary approach influences over the speed, efficiency and quality, and for this reason, it is a main aspect for the company to obtain gains in competitive advantage and create new business opportunities (AMARAL; DE SOUZA; FILHO, 2020).

Within the business decision-making systematic scope, a good customer portfolio management, i. e., according to Albrecht *et al*., (2015), the routines in which the management of the positions and movements that clients take over certain period of time occurs, is fundamental for the success of the product development process.

However, it is not simple to predict the individual movement of the different customer portfolio segments, especially in the context of strong instability and changes in the pattern of acquisitions (RIBEIRO, 2016). Hence, usual demand forecast methods (e.g., moving average and exponential smoothing) are not able to predict the sales transition from one product to another and vice versa, although these methods are useful for aggregate (or global) demand forecasting. Consequently, the total sales amount forecasting may be constant, but the sales transition from one product to another over the next few years is not explored. Well planned transitions (from products with a lower profit margin to a higher one) may cause considerable gains.

In addition, historical projection methods require a reasonable amount of historical data, besides they are not suitable in contexts of changing behavior patterns (FAVERO, 2015). In qualitative approaches such as Delphi method, according to Marques e Freitas (2018), its result highly depends on the preparation of questionnaires and selection of specialists, than often generates a high amount of data for analysis, which could not allow its application in complex conditions as approached in this article.

Recent studies, such as LI, *et al.* (2014), Zhou *et al.* (2018) Gagliardi, Kapelan and Franchini (2017), have presented several satisfactory approaches based on the Markov Chains (MC) method. In these applications, the MC method is characterized by short-term applications, little computational effort and, mainly, to successfully represent the transitory behavior of demand.

Despite being a key element in the product portfolio management, i. e., the group of products that the company commercializes, the commercial decision-making choices assume a predominantly qualitative comportment in product development processes.

This article aims to develop a sales forecasting model that can support the customer portfolio management processes in the product development background. It also intend to apply and test this model in a supplier company in the agro-industrial sector that commercializes vegetable seeds to retailers in all regions of Brazil. Thereunto, this method maximizes the performance of the marketing product insertion process by quantitatively directing commercial efforts towards more relevant customer segments.

The current model assumes a hybrid quality and is the result of two single models approach: (i) markovian models to represent the transitory characteristics of sales distribution among portfolio items; (ii) and first or second order exponential smoothing, when it is necessary to represent the trend behavior of total sales demand per customer.

This approach provides information necessary to answer questions such as:

- Does the process of replacing items in the product portfolio occur in a behavior that will consolidate the new product in the customer portfolio in the short term?
- Which are the segments of the customer portfolio (e.g., geographic region and customer categorization by quantity purchased) most susceptible to the process of consolidating the new product?
- And which are deficient in this process?
- How will the new product sales amount be distributed for the following year in conditions of the current management scenario maintenance?
- Which are the customer portfolio segments that commercial actions can focus and maximize the new product sales results for the company?
- Which segments with high-expected aggregate demand may present a potential direction to the new product sales efforts?

This supplier company's vegetable seeds product portfolio, headquartered in the Pernambuco state in the Brazilian Northeast, contemplates a diverse number of species, i. e., reproductively isolated group of morphologically similar individuals that produces fertile descendant (e.g., lettuce, coriander, onion, tomato). In turn, it can be subdivided into the different cultivars, i. e., species subgroup related to the set of plants with the same characteristic according to agricultural production criteria, obtained from genetic improvement, and that differentiate by presenting unique traits of flavor, shape, productivity, resistance to diseases, among other criteria.

This article analyzes the coriander species, which is responsible for most of this company's historical sales amount. After the HTV 9299 coriander cultivar release, in 2016, the sales forecast model developed aims to estimate its results in terms of the new cultivar amount sold in comparison to this result for all species. In order to consolidate this new product in the portfolio, the proposed modeling seeks to follow the evolution in coriander seeds sales distribution among the three cultivars responsible for its sales (*Verdão*, HTV 0699 and HTV 9299) and that represent the same demand group for each customer.

Understanding the movements observed among the customer segments, whether in frequency, intensity or destination, is essential for sales management to occur effectively. After all, it is manager's role to interfere, through the actions taken in these movements, so that they are in concomitance with the company's objectives. Consequently, the importance of client portfolio management surpasses as a management support tool, and presents itself as a solid process to assist business development, in line with guidelines and strategies taken by the commercial team (ALBRECHT *et al*., 2015).

The decision-making process in situations of high degree of uncertainty regarding the future conditions of the environment is characterized by the limited rationality of decision makers, after all, these agents do not have the capacity to obtain and process all relevant information for this. In this context, simulation and future scenarios forecasting models are suitable tools for analyzing complex systems and assist the search for results that maximize the actions taken (MELO; FUCIDJI, 2016).

The remaining content of this article is organized as follows. First, a literature review is carried out regarding the main topics covered. Subsequently, a sales forecasting model that acts in accordance with the problem addressed is proposed. Finally, the results are presented, discussed and validated when compared with the one accomplished so far this year.

## 2. THEORETICAL FOUNDATION

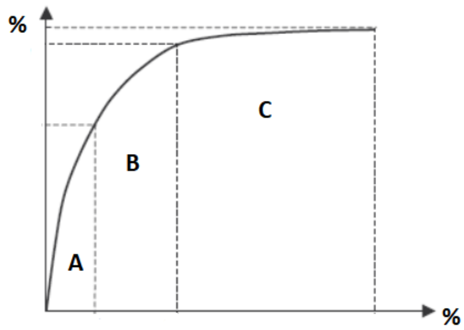
This section sets out the main concepts that will be included throughout this paper. In the first subsection, the concepts of the ABC analysis methodology, its applications and procedure are defined. In the second subsection, the demand forecasting approaches for exponential smoothing are detailed. In the third subsection, the concept of MC is presented, and, subsequently, a review of the literature is made on the forecasting models that apply MC.

*2.1 The ABC Curve*

Usually quoted in industrial applications, the ABC Curve is a methodology applied to support the decision-making processes based on the Pareto Principle, which requires a differentiation of the stock item's treatments based on some parameter of analysis (MOREIRA, 2012).

Each item is classified in terms of the total percentage of the analysis parameter in degree of importance, e. g., the cost of each item in stock. Subsequently, by drawing a graph, in which the X-axis represents the quantity of items, it’s observed, according to the Y-axis, that the curve rises quickly, followed by slow growth, indicating the less relevant items in the overall result (MOREIRA, 2012). This behavior can be observed according to Figure 1.

**Figure 1**: Standard ABC Curve Model.




Source: Adapted, Peinaldo e Graeml (2007).

The less relevant to the overall result items correspond to the places where the Y-axis value is close to 100%. The main regions of the graph, and, consequently, the categories (or classes), are defined, according to Moreira (2012) as:

I. Class A: corresponds to a small number of items, responsible for the highest accumulated percentage of investments. These are the most important items, which should receive special attention;

II. Class B: corresponds to an intermediate number of items, responsible for an accumulated percentage also intermediate of the investments. They should receive attention, but less than the items in region A;

III. Class C: corresponds to the largest number of items, responsible for small investments. They should be controlled with less relative rigor than the items of the previous classes or regions;

### *2.1.1 Other business applications of ABC Analysis*

Despite being mainly associated with the management of the degree of importance of the items in a stock based on their respective values, the ABC classification can be used in any situations that involve the establishment of priorities (VIANA, 2006).

There are countless possibilities for using this method in the business environment, which some of them are mentioned below. Peinado and Graeml (2007) use the method for a company's suppliers management. Vago (2013) applies the method to question the portfolio shape, through the supply of consumable material in a warehouse analysis. While Albrecht *et al*. (2015) make commercial analyzes as to the relevance that certain customers have in the results of business sales.

According to Peinaldo and Graeml (2007), the larger amount of revenue concentration in small groups of products can also mean a potential risk associated with high dependence on these products. In addition, the ABC classification, in this case, can be used to allocate greater efforts and resources in reducing costs and controls related to category A products.

### *2.1.2 ABC Curve application procedure in customer portfolio management*

Despite presenting different methods of application according to the context or purpose of the application, these procedures have some aspects in common. Albrecht et al. (2015), when applying the procedure in a context of client portfolio management, use the following procedure described in Table 1:

**Table 1** - Procedure for applying the ABC Curve in customer portfolio management.

| STEPS | DESCRIPTION |
|---|---|
| 1 | - Sort customers in descending order according to their absolute revenue amounts |
| 2 | - Build the percentage that the revenue of a given customer represents in the total revenue amount of the portfolio. Divide the customer's revenue by the total portfolio sales amount. |
| 3 | - Build the accumulated percentage of the values by adding the current value to the immediately previous value |
| 4 | - Assign the classes to the accumulated values. For this, conventional distribution values are adopted. |

Source: Adapted, Albrecht *et al*. (2015).

The accumulated percentage limits, as well as the number of groups with the same degree of relevance, are not fixed, and may vary according to the application. For Dias (2010) and Moreira (2012), this classification obeys only criteria of common sense and convenience. However, Dias (2010) indicates a maximum percentage of 20% of units in class A and 30% in class B.

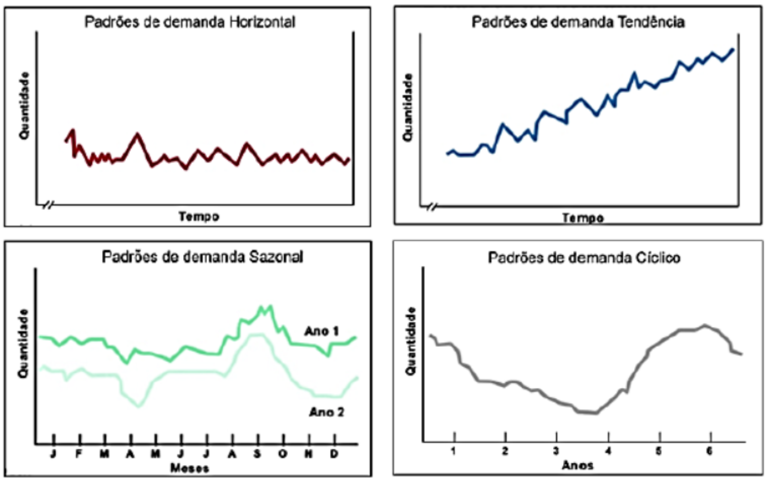
*2.2 Time series demand forecasting*

The definition of time series is associated to a sequence of observations of any variable (in this case, demand) over time. For this type of demand forecast, the variable is not associated with any other variable, and future values must be predicted exclusively based on the past values (MOREIRA, 2012).

Ribeiro (2016) highlights the existence of four basic behaviors (or patterns) in time series analysis for sales forecast, shown in Figure 2:

I. Horizontal: data fluctuation around a constant average;
II. Trend: systematic increase or decrease in the series average over time;
III. Seasonal: sales oscillation during the year;
IV. Cyclic: less predictable gradual increase or decrease in demand over longer periods of time;

**Figure 2** - Time series basic components.




Source: Ribeiro (2016).

According to Moreira (2012), the decomposition of time series models process a time series and their predictions based on these four components. There are several mathematical models that can be used, which may vary based on the data set obtained.

With the large availability of data for time series demand forecasting, the selection criteria of the applied model has presented as a prerequisite the possibility of automation that supports the large amount of required forecasts. The models addressed in the next subsections are frequently used in applications in which the robustness criterion in the forecast automation is fundamental and are characterized by simplicity, reliability and well-established historical record in its application (BARROW *et al*., 2020).

*2.2.1 Simple exponential smoothing demand forecasting*

Simple (or first order) exponential smoothing assumes the oscillation of demand around a constant base demand, which is adjusted according new data for the variable is incorporated into the time series. An adjustment between the predicted value and the previous real demand is done by a smoothing coefficient, which can be estimated based on some optimization model (RIBEIRO, 2016). Table 2 presents the parameters and an empirical formula that describes the subsequent step demand forecast.

**Table 2** - Parameters for exponential smoothing demand forecasting model.

| Parameter | Description | Restriction |
|---|---|---|
| $D_t$ | Forecast for the period t | $D_t = D_{t-1} + \alpha * (Y_{t-1} - D_{t-1})$ |
| $D_{t-1}$ | Forecast for the period t-1 | |
| $\alpha$ | 1st order smoothing constant | $0 \leq \alpha \leq 1$ |
| $Y_{t-1}$ | Real demand for the period (t-1) | |

Source: Moreira (2012).

As well as the moving average demand forecasting method, however, with greater sophistication, the simple exponential smoothing model is more suitable for time series that do not show trend and/or seasonality effects (RIBEIRO, 2016).

*2.2.2 Trend-adjusted exponential smoothing demand forecasting*

Trend-adjusted exponential smoothing follows the model principle used in first order application. However, instead of using the parameters of real demand, the method is applied over the first order smoothing results, and for this reason it is also known as second order smoothing, as highlighted in Table 3 (RIBEIRO, 2016).

**Table 3** - Parameters for trend-adjusted exponential smoothing demand forecasting model.

| Parameter | Description | Restriction |
|---|---|---|
| $D_t'$ | 2nd order forecast for the period *t* | $D_t' = D_{t-1}' + \beta * (D_{t-1} - D_{t-1}')$ |
| $D_{t-1}'$ | 2nd order forecast for the period *(t-1)* | |
| $\beta$ | 2nd order smoothing constant | $0 \leq \beta \leq 1$ |
| $D_{t-1}$ | 1st order forecast for the period *(t-1)* | |

Source: Moreira (2012).

The application of this second variable results in a greater smoothing in relation to the real time series values, and, therefore, more distant than the first order smoothing results. In this way, the second order forecast presents a kind of “correction of the trend effect”, being more appropriate to a time series that presents a trend effect, but not seasonality (RIBEIRO, 2016).

*2.3 Stochastic processes*

Stochastic processes represent the behavior of all systems whose state varies through a distribution of probabilities as a function of time. There are countless real phenomena that allow modeling through stochastic processes (HINOJOSA; MILANÉS, 2011).

For a function {X (t), t ∈ T}, a stochastic process is considered if, for each t ∈ T, X (t) is a random variable. A stochastic process is defined by a set of random variables that describe the temporal evolution of a process with T being defined as the parameter space. A stochastic process can have its parameter space in continuous or discrete time, as shown below (HINOJOSA; MILANÉS, 2011).

I. Discrete time: related to when the time variable is enumerable, i. e., $\{X_n, n = 0{,}1{,}2; 3 \ldots\}$

II. Continuous time: related to when the time variable is continuous, i. e., $\{X(t), t \geq 0\}$

The term state space refers to the set of values that represent the process variables. For stochastic processes, it is assumed that the state space is characterized according to a discrete or continuous time variable (HINOJOSA; MILANÉS, 2011). Figure 3 represents the types of stochastic processes with respect to the state space and parameters.

**Figure 3** - List of possibilities for stochastic processes in relation to the state space and parameters.

| **Stochastic Processes** | | |
|---|---|---|
| | **Enumerable E** | **Non-enumerable E** |
| **Enumerable t** | Process in discrete time and discrete state space | Process in discrete time and continuous state space |
| **Non-enumerable T** | Process in continuous time and discrete state space | Process in continuous time and continuous state space |

Source: Adapted, Hinojosa e Milanés (2011).

In the next topic, a specific case of stochastic process characterized as one of the main areas of modern probability and which has been highlighted in recent applications to quantify the behavior of these phenomena will be approached.

*2.3.1 Markov Chains*

Markov chains can be defined as a very important type of stochastic process whose purpose is related to the prediction of future states based exclusively on the initial state of the system (MAGELA, 2015).

A MC has the characteristic of being a memoryless stochastic process, i. e., the result of any state depends only on the result of the immediately previous state, and does not depend on other states (MAGELA, 2015).

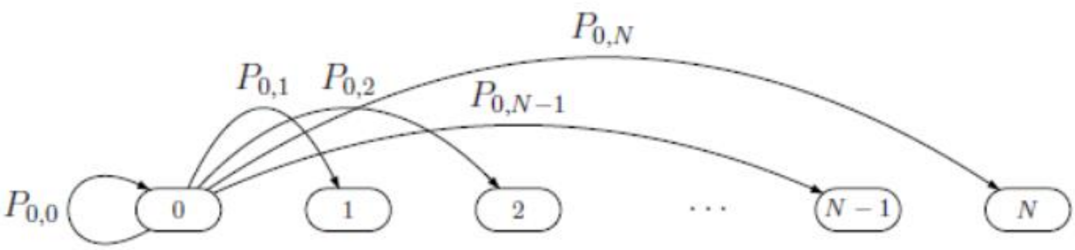
The transition (or conditional) probability of system from a state $s_i$ to a state $s_j$ in step n is defined by (MAGELA, 2015) as: $p_{ij}(n) = P(X_n = s_i | X_{n-1} = s_J)$.

For homogeneous (or stationary) Markov Chains, i. e., the conditional probabilities are independent of the system step, $p_{ij}(n)$ is denoted just as $p_{ij}$ since it is not independent of n (MAGELA, 2015).

If a state space j is reachable from a parameter n, for some n> 0, the transition probability has the restrictions $p_{ij} > 0$ and $\sum_{j=1}^{n} p_{ij} = 1$ (ROSS, 2010).

An alternative way to display the transition probabilities is to use a representation called the State-Transition Diagram. In this representation, the transition probabilities from some state are identified by arrows. Figure 4 represents all the transition probabilities from any state (MAGELA, 2015).

**Figure 4** - Transition Probabilities from State 0.




Source: Hinojosa e Milanés (2011).

*2.3.2 Recent Markov Chains forecasting applications*

Recent applications have shown a wide scope in the application of the CM method in several areas with satisfactory results in their applications. Some of these applications even combine MC characteristics with other demand forecasting methods. Hajirahimi and Khashei (2019), when analyzing the use of several hybrid models in recent publications, conclude that these hybrid models can enhance the performance compared to the single model application, combining the positive aspects of each method.

Gagliardi, Kapelan and Franchini (2017) compare the demand forecasting efficiency of a MC application with two more computationally robust methods, based on Artificial Neural Networks (ANN) and Naïve's method. As a result, there is a similar performance, with greater simplicity of implementation, in addition to good adaptation of the model to situations with low availability of historical data.

LI *et al.* (2014) apply the MC method in a marketing context with the objective to forecast the market share of the main automotive brands in a city. It is noteworthy that this demand forecast is based on the transitions pattern between brands in a market stability context. Despite highlighting that the MC method is mostly used in short-term forecasts, the authors present an approach of the method for long-term forecasts under stable conditions.

Zhou *et al.* (2018), through the application of a hybrid demand forecasting model, efficiently merging the MC method characteristics with regression models, predict both the individual behavior of the various stations, as well as the aggregate traffic demand of a shared bicycle system in a busy urban area.

In the agribusiness context, Ashoka (2013) analyzes the changes in the market share pattern in the export of Indian coffee. The transition probability matrix, for this case, is obtained by applying linear programming in order to minimize the average absolute deviation of annual coffee exports for each importing country. Consequently, the author aims to obtain a parameter of historical loyalty, i. e., the importer's ability to retain its market share over the years.

Santos and Braga (2012) apply MC in a convergence analysis of the plant productivity in the cities that belong to the Legal Amazon region. In this article, the authors determine productivity classes and investigate the cities transitory behavior between classes, while recognizing the limitation of the analysis according to the MC method to measure the impact of exogenous agents, since the transition probabilities remain constant over time.

As highlighted in Hajirahimi and Khashei (2019), although there are several publications that classify the time series forecasting models according to a variety of criteria, none of them properly considers the hybridization structures. Despite the aforementioned models are a reference for the objective of this paper, mainly because the hybrid characteristic they present, none of them is able to predict the transition demand pattern for this case.

In the context of the behavior transition changes that the insertion of a new product causes in the sales distribution of a similar group of products, the transitions between states are not accounted explicitly according to historical data, but are presented implicitly in the progression of the annual sales distribution. Therefore, in this work, due to the gap in the literature of applications in this context, an innovative model will be developed in order to measure this transient characteristic.

## 3. METHODOLOGICAL PROCEDURES

The main objective of this paper is to develop a demand-forecasting model for the following year that supports the strategic commercial decision making systematic during the process of consolidating a new item in its product portfolio. The model will be applied in a vegetable seeds supplier company that serves several retailers in different regions of the Brazilian territory. The model’s validation occurs through a comparison between the forecast and the partial results obtained for 2019 and brainstorming, i. e., according to Alves, Campos and Neves (2006), a technique widespread in several areas for generating ideas regarding a problematic) with a sector’s specialist, in order to qualitatively analyze the impact of commercial actions on the results obtained.

To achieve the proposed objective, the following steps are performed:

I. To delimit the analysis of the customer portfolio;
II. To model transitory demand;
III. To model aggregate demand;
IV. To integrate the models and forecast demand by group of customers;
V. To compare the actual results with results obtained by the model;

The data collection took place covering the interval between January 2011 to November 2019, according to the database provided by the company. The HTV 9299 coriander company's product portfolio incorporation occurred in January 2016. The data consists in the annual amount commercialized for the cultivars *Verdão*, HTV 0699 and HTV 9299, in which consists the coriander seeds demand for the entire customer portfolio. Associated with this customer portfolio, data that defines the geographic location of each of these is also. All of this information was obtained according the company's database through the ERP software.

The first phase of the data analysis procedure consists to delimit the customer portfolio for the most relevant group in terms of the historical revenues data, through the application of the ABC Curve methodology.

With the results obtained from the segmentation in degree of relevance, the transitional behavior per client of the coriander sales distribution between their cultivars, and the time variation for all species demand. Combining the effect of these two behaviors, it is possible to trace a projection of sales, in volume, for the coriander HTV 9299 per customer in 2019 and 2020.

By grouping customers into segments that have some similarity on the marketing decision-making criteria, e. g., customer size or geographic area, the demand forecast for this group is forecasted and, therefore, can serve as a quantitative direction of the commercial strategy and benchmark scenario for monitoring the commercial performance in that customer portfolio segment.

There is no well-defined methodology for the proposed objectives and analyzed context. For this reason, the methodology developed in this article by the authors is the result of a series of individual adaptations based on other proposed methods, in which the procedure of application is described in the topics below.

### *3.1 Customer Portfolio Segmentation*

Based on the approaches described in Peinado and Graemil (2007) and Viana (2006), in addition to the adaptations promoted by Albrecht et al. (2015), c.f., Table 1, for a customer portfolio management process, the ABC Curve method is applied to classify customers who purchased any product from this demand group during the interval between January 2016 and December 2018.

The accumulated percentage limits of 80%, 90% and 100%, respectively, are used for groups A, B and C, and as long as they satisfy the maximum limit proposed by Dias (2010) of 20% of units in class A and 30% in class B.

The set of clients that will have their behavior analyzed individually corresponds entirety to those classified as group A or B. Therefore, all those that are classified as less important, i. e., class C, will not have their demand patterns analyzed individually.

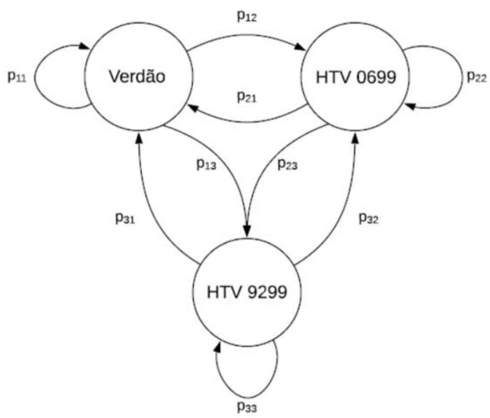
At the end of this topic, it is assigned an ordination for the coriander seeds results by ABC analysis for the customer portfolio. In this way, it is possible without compromising the final results to restrict the number of transactional demand analyzes between products per customer, which will be addressed in the next section.

*3.2 Transitional Demand between Products*

Given that the new product was included in portfolio in January 2016, the transition pattern from year t to year t + 1 is modeled based on this product's sales distribution between January 2016 and December 2018.

The transactional demand for this customer demand niche is modeled by a MC, in discrete time, with an annual step. As highlighted by Gagliardi, Kapelan and Franchini (2017), the MC method in suitable in situations with low historical data, in addition to the simplicity of implementation, which, due to this context of low time horizon since the launch of the new product and high number of iterations, this method adapts to the needs of this modeling. Its State Transition Diagram, in terms of the coriander aggregate demand distribution, is presented in Figure 5, while Table 4 describes its variables and parameter constants.

**Figure 5 -** State Transition Diagram per customer.




Source: Authorial.

**Table 4 -** Parameterization constants, restrictions, considerations and variables for applying the Markov Chain.

| Variable | | | Symbol |
|---|---|---|---|
| Transition probability between *Verdão* Coriander and *Verdão* Coriander | | | $p_{11}$ |
| Transition probability between *Verdão* Coriander and HTV 0699 Coriander | | | $p_{12}$ |
| Transition probability between *Verdão* Coriander and HTV 9299 Coriander | | | $p_{13}$ |
| Transition probability between HTV 0699 Coriander e *Verdão* Coriander | | | $p_{21}$ |
| Transition probability between HTV 0699 Coriander e HTV 0699 Coriander | | | $p_{22}$ |
| Transition probability between HTV 0699 Coriander e HTV 9299 Coriander | | | $p_{23}$ |
| Transition probability between HTV 9299 Coriander e *Verdão* Coriander | | | $p_{31}$ |
| Transition probability between HTV 9299 Coriander e HTV 0699 Coriander | | | $p_{32}$ |
| Transition probability between HTV 9299 Coriander e HTV 9299 Coriander | | | $p_{33}$ |
| **Parameterization Constants** | **Symbol** | **Description** | **Value** |
| *Verdão* Coriander Annual Demand | $Y_1(t)$ | *Verdão* Coriander historical sales volume (kg) | According database |
| HTV 0699 Coriander Annual Demand | $Y_2(t)$ | HTV 0699 Coriander historical sales volume (kg) | According database |
| HTV 9299 Coriander Annual Demand | $Y_3(t)$ | HTV 9299 Coriander historical sales volume (kg) | According database |
| Coriander Annual Demand (kg) | $Y(t)$ | Coriander historical sales volume (kg) | $Y(t) = Y_1(t) + Y_2(t) + Y_3(t)$ |
| *Verdão* Coriander Real Representativeness | $R_1(t)$ | Portion of *Verdão* Coriander in the total sale of cilantro in historical data in year t | $\frac{Y_1(t)}{Y(t)}$ |
| HTV 0699 Coriander Real Representativeness | $R_2(t)$ | Portion of HTV 0699 Coriander in the total sale of cilantro in historical data in year t | $\frac{Y_2(t)}{Y(t)}$ |
| HTV 9299 Coriander Real Representativeness | $R_3(t)$ | Portion of HTV 9299 Coriander in the total sale of cilantro in historical data in year t | $\frac{Y_3(t)}{Y(t)}$ |

Source: Authorial.

The state vector for this model application corresponds to the representativeness function projected for each type of product as a function of time, which is represented by the following equation:

$$S(t) = [s_1(t), s_2(t), s_3(t)] \quad (1)$$

The parameterization functions, as well as the restrictions to which they are subject, are characterized in Table 5.

**Table 5** - Parameterization functions and variable restrictions for Markov Chains application.

| Parameterization Functions | Symbol | Description | Function |
|---|---|---|---|
| *Verdão* Coriander projected representativeness as time function | $s_1(t)$ | Expected *Verdão* Coriander portion of total coriander sales for the year t | $s_1(t+1) = p_{11} * s_1(t) + p_{21} * s_2(t) + p_{31} * s_3(t)$ |
| HTV 0699 Coriander projected representativeness as time function | $s_2(t)$ | Expected HTV 0699 Coriander portion of total coriander sales for the year t | $s_2(t+1) = p_{12} * s_1(t) + p_{22} * s_2(t) + p_{32} * s_3(t)$ |
| HTV 9299 Coriander projected representativeness as time function | $s_3(t)$ | Expected HTV 9299 Coriander portion of total coriander sales for the year t | $s_3(t+1) = p_{13} * s_1(t) + p_{23} * s_2(t) + p_{33} * s_3(t)$ |

| Restriction | Description |
|---|---|
| $p_{11} + p_{12} + p_{13} = 1$ | The sum of the transition probabilities from state 1 is unitary |
| $p_{21} + p_{22} + p_{23} = 1$ | The sum of the transition probabilities from state 2 is unitary |
| $p_{31} + p_{32} + p_{33} = 1$ | The sum of the transition probabilities from state 3 is unitary |
| $s_1(t) + s_2(t) + s_3(t) = 1$ | Coriander sales are distributed exclusively among the 3 cultivars |
| $s_3(t) = 0 \; \forall \in [2011{,}2015]$ | The HTV 9299 cultivar has no sales history between 2011 to 2015 |

Source: Authorial.

Based on the use of linear programming as applied in Ashoka (2013), the transition matrix per client is obtained according 2 nonlinear optimization models, described below, which have the transition probabilities as variables. The first of these models aims to minimize the sum of the squared errors for the projected representativeness function between 2012 and 2015, in order to measure the transition characteristics between products prior to the portfolio modification. As $s_1(t) + s_2(t) = 1$ in this period, it is irrelevant whether the optimization process occurs by $s_1(t)$ or $s_2(t)$. At the end of this step, the variables values of the variables $p_{11}, p_{12}, p_{21}$ $e$ $p_{22}$ are obtained. The objective function for this model is:

Minimize:

$$SQ_e = \sum_{t=2012}^{2015} (s_1(t) - R_1(t))^2 \qquad (2)$$

Subject to the boundary condition:

$$s_1(t = 2011) = R_1(t = 2011) \qquad (3)$$

For the second optimization model, it aims to estimate the transition behaviors of all products group prior to the change in the portfolio due to the new product. At the end of this step, through the optimization process and adopted considerations, see Table 6, the variables $p_{13}, p_{23}, p_{31}, p_{32}$ $e$ $p_{33}$ are obtained. The objective function of this optimization model is described by:

Minimize:

$$SQ_e{}' = \sum_{t=2017}^{2018} (s_3(t) - R_3(t))^2 \tag{4}$$

Subject to boundary condition:

$$s_3(t = 2016) = R_3(t = 2016) \tag{5}$$

**Table 6** - Markov Chains application considerations.

| Considerations | Value | Description |
|---|---|---|
| The transition rate for the new product as the same regardless of the starting state | $p_{13} = p_{23}$ | The adopted transition probability is the same for $p_{13}$ and $p_{23}$ |
| The transition rate from coriander HTV 9299 acts to preserve the historical sales distribution among consolidated products | $p_{31} * \sum_{t=2016}^{2018} \frac{Y_1(t)}{Y(t)} = p_{32} * \sum_{t=2016}^{2018} \frac{Y_2(t)}{Y(t)}$ | The transition probability $p_{31}$ times the portion of *Verdão* Coriander between 2016 to 2018 is equal to the transition probability $p_{31}$ times the portion of HTV 0699 Coriander between 2016 to 2018 |

Source: Authorial.

The choice of applying the optimization in the period prior to the modification in the portfolio, as well as the considerations adopted, are resulted, in this context, from the low historical sales data availability for HTV 9299 Coriander. The model must determine the values for all 9-transition probabilities; however, the post-incorporation period of the new product includes only two steps. The considerations adopted in Table 6, although not ideal, are necessary for the optimization model.

This model is implemented for each customer and then replicated for the entire customer portfolio, according to section 3.1, by the Solver tool in the MS-Excel® software, applying the Visual Basic for Applications (VBA) programming language. This algorithm, whose function is to return the transition probabilities for the whole set analyzed, is detailed as shown in Figure 6. This procedure, in this case, it is characterized by simplicity in its application, which enables a large set of iterations.

**Figure 6** - Solver tool application

```
Sub Gerar_Matrizes_de_Transição()

ultimalinha = Sheets("Rep. Rel.").Range("A100000").End(xlUp).Row

For linha = 3 To ultimalinha
'Preenche os dados do novo cliente
    Sheets("Solver").Range("B1").Value = Sheets("Rep. Rel.").Range("A" & linha).Value

'Utiliza o Solver para Produto 1 vs Produto 2
    Plan8.Activate
    SolverReset
    SolverOptions precision:=0.0001
    SolverOk SetCell:="$K$20", MaxMinVal:=2, ValueOf:=0, ByChange:="$B$16:$B$19", _
        Engine:=1, EngineDesc:="GRG Nonlinear"
    SolverAdd CellRef:="$N$16", Relation:=2, FormulaText:="1"
    SolverAdd CellRef:="$N$17", Relation:=2, FormulaText:="1"
    SolverSolve userfinish:=True

'Utiliza o Solver para Produtos (1 + 2) vs Produto 3
    SolverReset
    SolverOptions precision:=0.0001
    SolverOk SetCell:="$K$27", MaxMinVal:=2, ValueOf:=0, ByChange:="$B$25:$B$28", _
        Engine:=1, EngineDesc:="GRG Nonlinear"
    SolverAdd CellRef:="$N$25", Relation:=2, FormulaText:="1"
    SolverAdd CellRef:="$N$26", Relation:=2, FormulaText:="1"
    SolverSolve userfinish:=True

'Preenche a Matriz de transição por cliente
    Sheets("PREVISÕES").Range("B" & linha).Value = Sheets("Solver").Range("B16").Value * Sheets("Solver").Range("B25").Value
    Sheets("PREVISÕES").Range("C" & linha).Value = Sheets("Solver").Range("B17").Value * Sheets("Solver").Range("B25").Value
    Sheets("PREVISÕES").Range("D" & linha).Value = Sheets("Solver").Range("B26").Value
    Sheets("PREVISÕES").Range("E" & linha).Value = Sheets("Solver").Range("B18").Value * Sheets("Solver").Range("B25").Value
    Sheets("PREVISÕES").Range("F" & linha).Value = Sheets("Solver").Range("B19").Value * Sheets("Solver").Range("B25").Value
    Sheets("PREVISÕES").Range("G" & linha).Value = Sheets("Solver").Range("B26").Value
    Sheets("PREVISÕES").Range("H" & linha).Value = Sheets("Solver").Range("B28").Value * _
    (Sheets("Solver").Range("B12").Value / (Sheets("Solver").Range("B12").Value + Sheets("Solver").Range("C12").Value))
    Sheets("PREVISÕES").Range("I" & linha).Value = Sheets("Solver").Range("B28").Value * _
    (Sheets("Solver").Range("C12").Value / (Sheets("Solver").Range("B12").Value + Sheets("Solver").Range("C12").Value))
    Sheets("PREVISÕES").Range("J" & linha).Value = Sheets("Solver").Range("B27").Value

Next linha

End Sub
```

Source: Authorial.

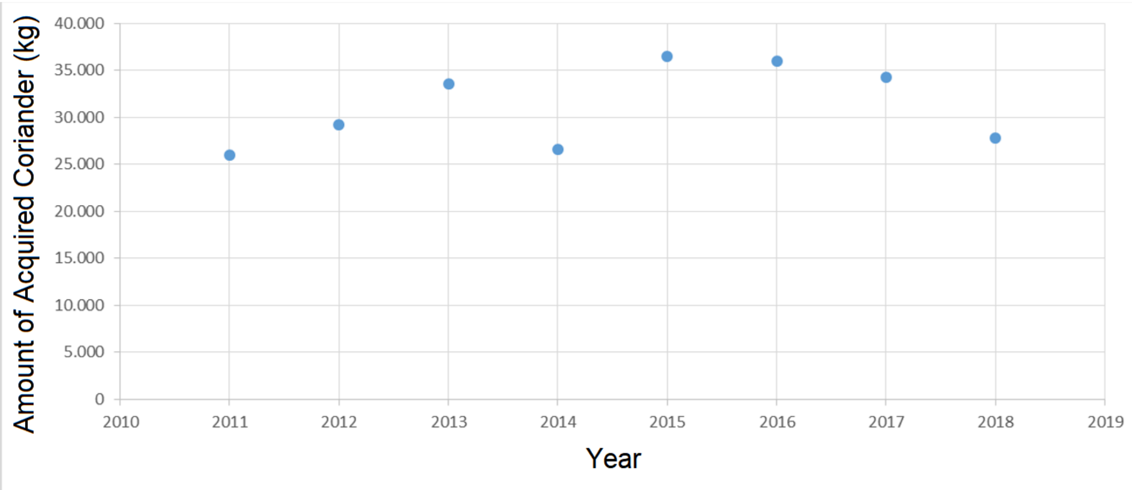
At the end of this stage, the Markovian process is characterized by the transition matrices obtained for each customer. Consequently, it's possible to estimate the annual distribution of coriander demand per customer, $S(t)$, for the following years.

*3.3 Coriander Aggregate Demand*

While there is a transitional relationship between the sales distribution for the cultivars *Verdão*, HTV 0699 and HTV 9299, it is also likely that there will be changes in the expected aggregate demand for all coriander cultivars for subsequent years. Therefore, it is necessary, for forecasting the demand for the new product, that the proposed model assume a hybrid character, as described in Zhou *et al.* (2018), when merging the transactional characteristics, addressed in the previous section, with the influence caused in the time series projection of all coriander demand per client.

Each time series, which represents the annual coriander seeds sales history per customer, is characterized based on the data in the Dispersion Diagram, shown in Figure 7.

**Figure 7** - Dispersion Diagram for a Generic Customer Portfolio Segment.




Source: Authorial.

The demand forecast, in this case, is carried out individually based on the different customer groupings by segments of relevance to the customer portfolio, c. f., next subsection. This procedure is applied in such a way that each iteration of the demand forecast presents a considerable sampling, in addition to reducing the size of the data set.

### *3.3.1 Customer Clustering*

The customers are grouped into certain similarity groups, in this case, some customer portfolio segments that may exhibit similar sales behavior according to marketing strategy criteria. For this paper, pre-defined classes of behavior similarity are considered: (i) the ABC classification of each client for each year between 2011 to 2018; (ii) the geographic influence per customer, characterized by the federative unit to which each customer registration is related. These factors, described according to Table 7, act as aggregate demand multipliers based on their influence on coriander sales result.

**Table 7** - Demand adjustment factors by class based on client portfolio segmentation criteria.

| Parameter | Description |
|---|---|
| $F_{ABC}$ | Demand adjustment factor according ABC Classification criteria |
| $F_{UF}$ | Demand adjustment factor according location criteria |

Source: Authorial.

The combination of these effects over the aggregate demand results in a final adjustment factor, $F_R$, defined according to the equation below, based on its Euclidean distance between the parameters described in Table 7. The forecast for coriander aggregate demand per customer, therefore, for a time t + 1 is defined as the product between the real demand observed in t with the resulting adjustment factor, $F_R$.

$$F_R = \sqrt{F_{UF}^2 + F_{ABC}^2} \quad (6)$$

The procedure for obtaining all parameters $F_{ABC}$ and $F_{UF}$ individually, is covered in the following subsection.

*3.3.2 Adjustment factors for relevant segments*

The method described in this topic aims to have its demand value forecast for 2019 based on the historical volume (kg) commercialized from 2011 to 2018 by each customer segment. This time series forecasting occurs by the first and second orders exponential smoothing methods.

The choice between the two methods occurs in so that the projection is more suitable to segments either they have a trend line behavior or not. For this, the selection occurs according to the smallest sum of the squared errors, after applying both methods. In search of a better fit, the first and second order smoothing constants are determined according to a nonlinear optimization model that will be described below.

The constants, variables and restrictions for both optimization models, as well as method selection and demand characterization in 2019 are described in Table 8.

**Table 8** - Parameterization constants, restrictions and variables for applying the 1st and 2nd order exponential smoothing methods.

| Variables | Symbol | Value |
|---|---|---|
| 1st order forecast for the period t | $D(t)$ | $D(t) = D(t-1) + \alpha * (Y(t-1) - D(t-1))$ |
| Trend adjustment function | $b(t)$ | $b(t) = \beta * \left(D'(t) - D'(t-1)\right) + (1-\beta) * b(t-1)$ |
| 2nd order forecast for the period t | $D'(t)$ | $D'(t) = \alpha * Y(t) + (1-\alpha) * (D'(t-1) + b(t-1))$ |
| 1st order smoothing constant | $\alpha$ | According optimization process |
| 2nd order smoothing constant | $\beta$ | According optimization process |
| Squared errors sum for adjustment by 1st order exponential smoothing | $SQ_e$ | $\sum_{t=2012}^{2018} (D(t) - Y(t))^2$ |
| Squared errors sum for adjustment by 2nd order exponential smoothing | $SQ_e{}'$ | $\sum_{t=2012}^{2018} (D'(t) - Y(t))^2$ |

| Parameterization Constants | Symbol | Value |
|---|---|---|
| Annual demand for the entire product group (kg) | $Y(t)$ | According database |

| Restrictions | Description |
|---|---|
| $0 \leq \alpha \leq 1$ | The 1st order smoothing coefficient assumes values between 0 and 1 |
| $0 \leq \beta \leq 1$ | The 2nd order smoothing coefficient assumes values between 0 and 1 |
| $D_t \geq 0$ | Expected demand cannot be negative |
| $D_t' \geq 0$ | Expected demand cannot be negative |

Source: Authorial.

The first application of the nonlinear optimization model has the following equations as an objective and boundary condition.

Minimize the sum of the squared errors:

$$SQ_e = \sum_{t=2012}^{2018} (D(t) - Y(t))^2 \tag{7}$$

Subject to the boundary condition:

$$D(t = 2011) = Y(t = 2011) \tag{8}$$

The second application of the nonlinear optimization model has the following equations as an objective and boundary condition.

Minimize the sum of the squared errors:

$$SQ_e{}' = \sum_{t=2012}^{2018} (D'(t) - Y(t))^2 \tag{9}$$

Subject to the boundary condition:

$$b(t = 2011) = \frac{\frac{Y(t = 2012)}{Y(t = 2011)} + \frac{Y(t = 2013)}{Y(t = 2012)} + \frac{Y(t = 2014)}{Y(t = 2013)}}{3} \tag{10}$$

The parameter selection that defines the influence factor that this type of segmentation imposes on the demand variable is expressed according to the criteria below.

$$SQ_e' > SQ_e \rightarrow F = \frac{D'(t = 2019)}{Y(t = 2018)} \tag{11}$$

$$SQ_e' \leq SQ_e \rightarrow F = \frac{D(t = 2019)}{Y(t = 2018)} \tag{12}$$

The demand time series forecasting method selection is based on the minimum sum of the squared errors, as well as its parameterization and subsequent iteration for all groups of customers, also occurs with assistance of the MS-Excel® software, by Solver tool, using the Visual Basic for Applications (VBA) programming language.

At the end of this step, it is possible to trace the demand forecast for coriander seeds per customer for the following two years as result of the product between its demand for the previous year with its respective resulting factor.

*3.4 Demand forecasting for relevant customer portfolio segments*

The groups of clients considered relevant to the customer portfolio management that will have their results presented to support the commercial decision-making processes are:

1. All customer portfolio: in order to analyze the whole performance of the new product's post-development process;
2. Customers by ABC category: in order to analyze the process performance for customers with the same size in terms of revenue;
3. Customers by region: in order to analyze the process performance according geographic criteria;

*3.5 Monitoring the results obtained*

In this section, the results presented in section 3.4 represent a reference scenario (benchmark), which is, the forecast for the maintenance of sales historical behavior observed in the customer portfolio.

The brainstorming carried out with the specialist of the area in the company aims to identify the commercial actions carried out in 2019 and their effects that may cause a divergence between the benchmark and real scenario. In addition to validating the proposed model, this step aims to measure the effect of these commercial actions and present other divergences not identified by the company's management.

The applied method is based on the one proposed by Baxter (2000) as classic brainstorming, and its application can be described in the following phases:

1. Orientation: Initial brainstorming phase in which the nature of the problem is determined;
2. Preparation: In this phase, all data related to the problem are gathered;
3. Analysis: Examines previous orientation and preparation phases
4. Ideation: It is the creative phase, in which the various ideas that represent possible solutions to the problem are generated;
5. Incubation: With the decrease in the fluency of ideas, the session is suspended for one day for a deliberate departure from the problem. After the period of absence, a new fluency of ideas may emerge in the resumption of the session;
6. Synthesis and evaluation: Ideas are judged and a selection is made according to the nature of the problem defined in the orientation phase;

The actual results collected from the historic sales data that serves as the basis for the brainstorming session are those obtained partially for the year 2019, between the months of January and November, compared with the forecast obtained and the nature of the problem corresponds to the reasons why data may have presented diverged compared to the benchmark scenario.

## 4. RESULTS AND DISCUSSION

The methodological procedures are applied in the real case of a vegetable seed distribution company based in the state of Pernambuco, Brazil. The context involves the consolidation process of the HTV 9299 coriander cultivar, joined in the product portfolio in 2016. Its results are based on the historical behavior of the other coriander cultivars (*Verdão* and HTV 0699), already consolidated in the company's historical sales.

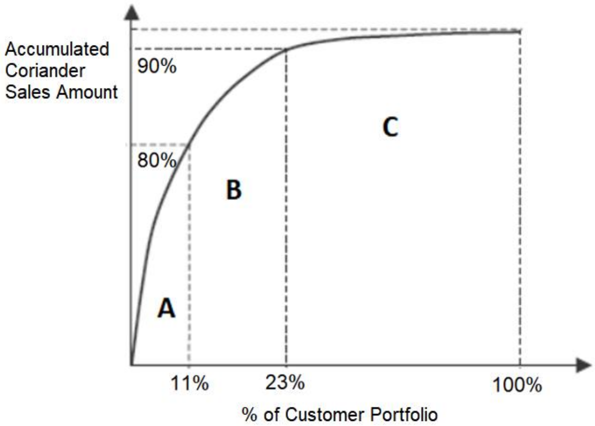
The specific results of each stage are presented in each subsection of this topic, in which, also, their results are discussed and analyzed. Finally, the results presented must show a quantitative direction to the commercial decision-making processes by bringing a benchmark scenario for the subsequent two years from the analyzed period.

### *4.1 Analysis segmentation by the ABC Curve method*

There are a total of 751 clients registrations in the customer portfolio according to the database, in which, 342 of these had some coriander seeds buying record after January 2016 after the HTV 9299 coriander portfolio inclusion.

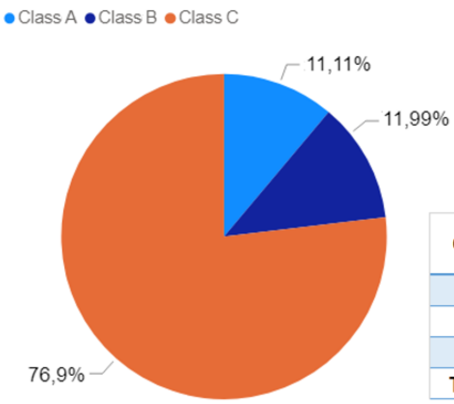
The the ABC Curve method results for the customer portfolio segmentation, in a manner similar to Albrech *et al.* (2015), according to the accumulated percentage parameters defined above, are summarized in Figure 8 and Figure 9.

**Figure 8 - ABC** Curve for customer portfolio.




Source: Authorial.

**Figure 9** - Customer distribution by ABC category (2016 a 2018).




| Class | Number of Clients |
|---|---|
| A | 38 |
| B | 41 |
| C | 263 |
| **TOTAL:** | **342** |

Source: Authorial.

The ABC classification results showed a maximum percentage of units in the categories A and B, well below the limits of 20% and 30% of the customer portfolio in these classes, respectively, as suggested by Dias (2010) and adopted in the development of the methodology, c. f. section 3.1, validating the accumulated percentages adopted. In addition, a high quantity of customers represents low relevance for this model. Therefore, all class C customers do not have their transitory demands analyzed individually, but the MC application, unlike class A and B customers, occurs only once for this entire ABC category. Thus, it is possible to affirm that, because of this methodology, the units for individual analysis was reduced by 76.9%, which allowed a better operationalization of this demand forecasting approach with low loss in its degree of reliability.

*4.2 Markov Chains transitional demand*

Through the multiple iterative processes of non-linear optimization, 80 homogeneous transition matrices are generated (38 for class A customers, 41 for class B customers and 1 representing the aggregate behavior for all class C), which have their results stored in a spreadsheet. Among the group analyzed, only 20 exhibited some transitory behavior for the new product, i. e., $p_{13} > 0$ or $p_{23} > 0$, so that the state space related to the new product acquisition is not reachable for most of this group of customers. Table 9 lists all the probabilities, which characterize the transition matrices, for the set of customers with expected adherence to the new product.

**Table 9 -** Transition probabilities from year t to year t+1 for customers that incorporated the new product, highlighted those most susceptible to its incorporation.

| CUSTOMER | STATE-TRANSITION MATRIX | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | P11 | P12 | P13 | P21 | P22 | P23 | P31 | P32 | P33 |
| **1** | **88,184%** | **0,000%** | **11,816%** | **11,457%** | **76,726%** | **11,816%** | **37,958%** | **39,244%** | **22,798%** |
| 2 | 80,276% | 10,512% | 9,211% | 90,789% | 0,000% | 9,211% | 19,367% | 2,977% | 77,656% |
| 3 | 94,854% | 1,259% | 3,887% | 96,113% | 0,000% | 3,887% | 0,000% | 0,000% | 100,000% |
| 4 | 16,813% | 76,202% | 6,985% | 0,000% | 93,015% | 6,985% | 0,000% | 100,000% | 0,000% |
| **9** | **87,115%** | **0,000%** | **12,885%** | **7,547%** | **79,569%** | **12,885%** | **21,390%** | **78,610%** | **0,000%** |
| 11 | 33,350% | 64,381% | 2,269% | 3,213% | 94,518% | 2,269% | 7,782% | 92,218% | 0,000% |
| 16 | 88,780% | 1,767% | 9,453% | 1,023% | 89,524% | 9,453% | 10,020% | 89,980% | 0,000% |
| 23 | 76,054% | 15,422% | 8,523% | 21,091% | 70,386% | 8,523% | 39,144% | 60,856% | 0,000% |
| 32 | 94,738% | 0,972% | 4,290% | 0,000% | 95,710% | 4,290% | 0,000% | 0,000% | 100,000% |
| 36 | 97,990% | 0,000% | 2,010% | 0,000% | 97,990% | 2,010% | 0,000% | 0,000% | 100,000% |
| 37 | 98,399% | 0,000% | 1,601% | 0,000% | 98,399% | 1,601% | 19,680% | 80,320% | 0,000% |
| 39 | 93,962% | 0,000% | 6,038% | 0,009% | 93,953% | 6,038% | 0,000% | 0,000% | 100,000% |
| 51 | 92,090% | 0,000% | 7,910% | 0,000% | 92,090% | 7,910% | 1,622% | 98,378% | 0,000% |
| **55** | **22,273%** | **64,342%** | **13,385%** | **59,085%** | **27,530%** | **13,385%** | **57,181%** | **42,819%** | **0,000%** |
| 56 | 3,026% | 87,358% | 9,616% | 0,000% | 90,384% | 9,616% | 100,000% | 0,000% | 0,000% |
| 69 | 95,735% | 0,000% | 4,265% | 21,497% | 74,238% | 4,265% | 52,759% | 47,241% | 0,000% |
| 70 | 21,357% | 76,990% | 1,653% | 26,932% | 71,415% | 1,653% | 24,145% | 75,855% | 0,000% |
| 73 | 97,400% | 0,000% | 2,600% | 0,000% | 97,400% | 2,600% | 0,827% | 99,173% | 0,000% |
| 78 | 7,024% | 85,434% | 7,542% | 26,887% | 65,572% | 7,542% | 24,021% | 75,979% | 0,000% |
| OTHERS | 92,572% | 0,000% | 7,428% | 7,642% | 84,931% | 7,428% | 77,459% | 22,541% | 0,000% |

Source: Authorial.

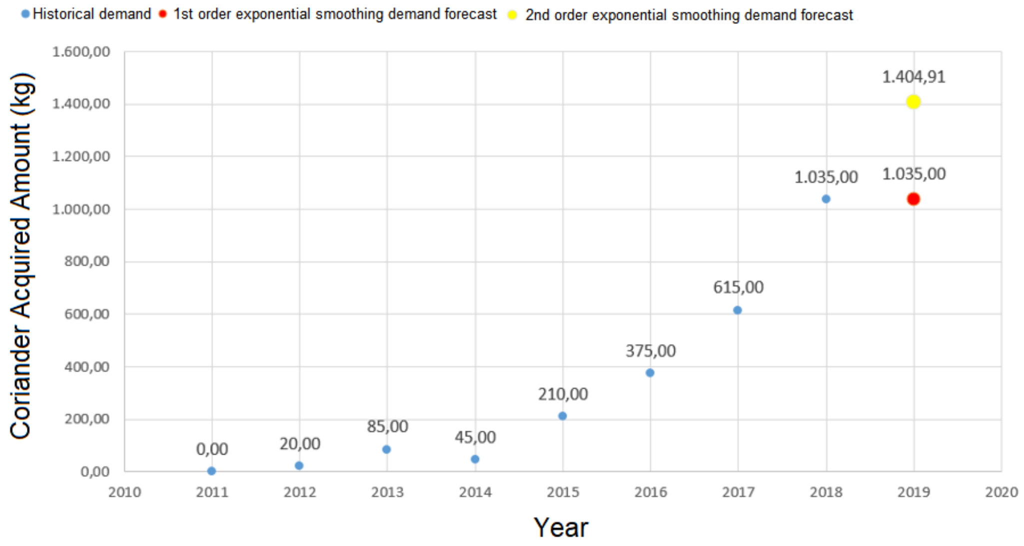
It is possible to highlight customers who are more susceptible to the incorporation process according their current purchasing patterns. Among those mentioned, Customers 1, 9 and 55, in bold, stand out with the transition probabilities $p_{13}$ and $p_{23}$ above 10%, which can serve as a benchmark for directing commercial activities to other clients. For example, if a customer has similar conditions, e. g., geographic or revenue amount, to one of the three highlighted clients, but still presents an inefficiency in the incorporation process, its commercial actions can look for the actions taken among these benchmark clients.

Combining the results presented in this section with the ABC classification, it is possible to notice that 28.95% of the customers classified as highly important (category A) presented some adherence behavior to the new product, whereas, for the category B, this value corresponds to 19.51%. Among the 10 customers that had the highest sales volume of this type between 2016 and 2018, for example, only 5 of them have any sales record forecast for the new product. It is suggested that the commercial actions for the new product's market expansion should be concentrated on category A customers. Since the customers of this category are more relevant to the overall results for the new cultivar incorporation, the focus on product penetrability on these types of customers would result in more significant aggregate results for this process.

*4.3 Exponential smoothing aggregate demand forecasting*

Among the 23 groups of customers analyzed individually (3 by ABC classification and 23 by Federative Units (FU)), 6 of these groups were better suited to the exponential smoothing method, indicating that these time series are more susceptible to the trend effect. As described in 2.2, the seasonality effect is not considered to be relevant in this case. Due to the great variability in the data distribution obtained, the option to choose between the exponential smoothing methods allowed a greater adjustment to the presence (or absence) of trend effect for that group. Figure 10 exemplifies the divergence in the response presented by the methods and the trend correction presented by the second order smoothing method.

**Figure 10** - Comparative between the first (in red) and second (in yellow) exponential smoothing methods for the State of




ES.

Source: Authorial.

As for the customer grouping by ABC analysis criteria, the factors obtained have very similar values and close to the unit value. Therefore, it is not observable that this type of segmentation presents different patterns of considerable similarity for the global demand for coriander seeds, according to Table 10.

**Table 10** - Adjustment factor obtained by customer grouping by ABC analysis.

| *Class* | $F_{ABC}$ |
|---|---|
| *A* | 1,000 |
| *B* | 0,998 |
| *C* | 0,999 |

Source: Authorial.

In contrast, the grouping by geographic regions indicated a reduced influence on the demand for 2019. As exemplified in Table 11, Pernambuco (PE), which represents the largest sales region, should maintain a constant demand for coriander seeds in 2019, while relevant regions such as Ceará (CE) and Paraíba (PB) are expected to have an increase in their percentage in the coriander sales distribution by states. On the other hand, regions such as Pará (PA) and Sergipe (SE) showed a predicted drop in their indexes. It suggests that some regions will lose sales force, and, therefore, the company's management must further analyze their causes.

**Table 11** - Aggregate sales distribution forecast comparison for FU in 2019 and 2020, based on ajust factor, *versus* the results in 2018.

| *FU* | $F_{UF}$ | *Sales % in 2020* | *Sales % in 2020* | *Sales % in 2020* | *Variation (2019/2018)* | *Variation (2020/2018)* |
|---|---|---|---|---|---|---|
| *PE* | 1,000 | 22,370% | 22,545% | 21,72% | +0,79% | -2,90% |
| *PA* | 0,910 | 17,344% | 15,902% | 14,06% | -8,32% | -18,93% |
| *CE* | 1,047 | 13,055% | 13,776% | 13,93% | 5,53% | +6,73% |
| *PB* | 1,051 | 12,781% | 13,533% | 13,75% | 5,89% | +7,56% |
| *SE* | 0,779 | 12,763% | 10,021% | 7,99% | -21,48% | -37,38% |
| *AL* | 1,000 | 6,987% | 7,042% | 6,78% | 0,79% | -2,90% |
| *RN* | 1,000 | 6,744% | 6,797% | 6,55% | 0,79% | -2,84% |
| *GO* | 1,474 | 4,286% | 6,365% | 10,48% | 48,51% | +144,44% |
| *MT* | 0,967 | 1,236% | 1,205% | 1,12% | -2,53% | -9,02% |
| *PI* | 0,856 | 0,754% | 0,650% | 0,55% | -13,69% | -26,97% |
| *BA* | 1,442 | 0,618% | 0,898% | 1,43% | 45,34% | +130,79% |
| *ES* | 1,357 | 0,464% | 0,635% | 0,91% | 36,81% | +96,82% |
| *MA* | 1,000 | 0,323% | 0,325% | 0,31% | 0,79% | -2,68% |
| *RO* | 0,953 | 0,229% | 0,220% | 0,20% | -3,91% | -11,23% |
| *DF* | 1,747 | 0,040% | 0,071% | 0,16% | 76,05% | +299,82% |
| *AC* | 2,002 | 0,007% | 0,014% | 0,04% | 101,76% | +510,79% |
| *RR* | 0,333 | 0,000% | 0,000% | 0,00% | -66,47% | -69,95% |

Source: Authorial.

*4.4 Demand forecast in relevant segments of the customer portfolio*

Subsequently to the hybrid demand forecast model application, it is forecasted a commercialized coriander seeds volume slightly similar from that presented in 2018. In 2020, the forecasted amount is only 2.63% higher than that observed in 2018. The company's database indicates a total of 12,526 kg (5.62% of the volume of the species) marketed in 2018 for the HTV 9299 cultivar. The model projects, for 2019, a HTV 9299 commercialized volume of 15,195 kg (6.81% of the volume of the species), that is, an increase of 21.31% in relation to 2018, and, for 2020, a volume of 16,589 kg, which represents an increase of 32,51% in relation to 2018. This shows that, in fact, coriander HTV 9299 is being incorporated into the market.

Despite this, this incorporation process does not occur in a behavior that indicates this complete (or major) portfolio substitution possible over time. Thus, if the company has as a key objective this process, it is necessary that there is a change in the post-development commercial processes for this new product. However, as it is characteristic of the demand forecasting applications applying MC methods, e.g., LI, et al. (2014), Zhou et al. (2018) and Ashoka (2013), this methodology allows an individual analysis of the customer portfolio main segments.

Table 12 represents the new product's expected sales volume by customer size in terms of volume purchased. In this table, there is a high variation in percentage terms for class B customers, but, even so, it is considerably lower than the other classes, as in 2018. In parallel, the less relevant group obtained in terms of volume commercialized should be less susceptible for consolidating the new product in their portfolio.

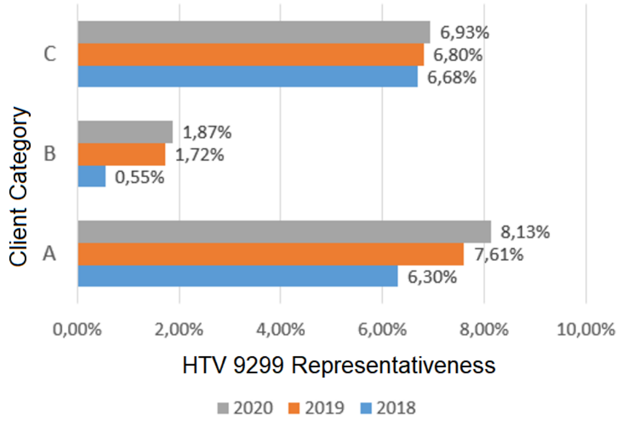
**Table 12** - Forecasted amount for the HTV 9299 coriander for customer grouping by ABC analysis criteria (kg).

| *Class* | *Amount in 2018 (kg)* | *Forecasted amount in 2019 (kg)* | *Forecasted amount in 2020 (kg)* | *Variation (2019/2018)* | *Variation (2020/2018)* |
|---|---|---|---|---|---|
| *A* | 11.205,00 | 13.458,4 | 14.680,85 | +20,11% | +31,02% |
| *B* | 150,00 | 474,04 | 523,90 | +216,02% | +249,27% |
| *C* | 1.171,00 | 1.262,78 | 1.393,51 | +7,84% | +19,00% |

Source: Authorial.

Figure 11 highlights the low representativeness that the new cultivar still has in its expected sales volume. On the other hand, the main customers (Class A) present a forecast of greater representativeness for this product, even though only 28.95% of this category's customers show transitory behavior foreseen for this item, according section 4.1.

**Figure 11** - 2019 and 2020 forecasted new product's representativeness in sales volume based on customer grouping by ABC analysis.




Source: Authorial.

Figure 12 and Figure 13 show the new product's total volume distribution forecasted for 2019 and 2020, respectively, between the federative units. It is possible to infer that the portfolio consolidation process still presents a geographical restriction, since most of the new product's forecasted sales volume (58.22% in 2019 and 60.83% in 2020) concentrate in the same state (PE), where it is located the company's headquarters. Therefore, it indicates that the process of new product incorporation has not yet reached sufficient geographical coverage to replace the consolidated products in the portfolio.

Combining these results with the one obtained in the previous section, it can be notice that two of the three main states in terms of the expected coriander seeds sales volume still show high resistance to the portfolio substitution process. The States of Pará (PA) and Ceará (CE), highlighted in bold in Figure 12 and Figure 13, are an example of this geographical resistance. The author suggests that, in the case of geographic expansion of the new product's customer incorporation, commercial efforts should be prioritized in the client portfolio in these regions, given its significance in the aggregate demand for coriander seeds.

**Figure 12** - HTV 9299 expected sales geographical distribution (2019).


| UF | % HTV 9299 CORIANDER SALES (2020) | % CORIANDER SALES (2020) |
|---|---|---|
| PE | 58,22% | 22,55% |
| SE | 19,50% | 10,02% |
| PB | 9,65% | 13,53% |
| GO | 6,73% | 6,37% |
| **PA** | **1,59%** | **15,90%** |
| RN | 1,01% | 6,80% |
| BA | 0,96% | 0,90% |
| **CE** | **0,82%** | **13,78%** |
| AL | 0,37% | 7,04% |
| MT | 0,35% | 1,21% |
| RO | 0,24% | 0,22% |
| MA | 0,23% | 0,33% |
| PI | 0,13% | 0,65% |
| ES | 0,09% | 0,64% |
| DF | 0,09% | 0,07% |
| AC | 0,02% | 0,01% |
| RR | 0,00% | 0,00% |
| RJ | 0,00% | 0,00% |
| SP | 0,00% | 0,00% |
| RS | 0,00% | 0,00% |
| MS | 0,00% | 0,00% |
| SC | 0,00% | 0,00% |
| PR | 0,00% | 0,00% |
| MG | 0,00% | 0,00% |
| TO | 0,00% | 0,00% |
| AM | 0,00% | 0,00% |
| AP | 0,00% | 0,00% |

Source: Authorial.

**Figure 13** - HTV 9299 expected sales geographical distribution (2020).


| UF | % HTV 9299 CORIANDER SALES (2020) | % CORIANDER SALES (2020) |
|---|---|---|
| PE | 60,83% | 21,72% |
| SE | 14,32% | 7,99% |
| GO | 9,41% | 10,48% |
| PB | 8,78% | 13,75% |
| **PA** | **1,55%** | **14,06%** |
| BA | 1,27% | 1,43% |
| ES | 0,93% | 0,91% |
| RN | 0,90% | 6,55% |
| **CE** | **0,73%** | **13,93%** |
| AL | 0,31% | 6,78% |
| MT | 0,29% | 1,12% |
| MA | 0,20% | 0,31% |
| RO | 0,19% | 0,20% |
| DF | 0,15% | 0,16% |
| PI | 0,09% | 0,55% |
| AC | 0,04% | 0,04% |
| RR | 0,00% | 0,00% |
| AM | 0,00% | 0,00% |
| AP | 0,00% | 0,00% |
| SP | 0,00% | 0,00% |
| MG | 0,00% | 0,00% |
| RJ | 0,00% | 0,00% |
| RS | 0,00% | 0,00% |
| PR | 0,00% | 0,00% |
| TO | 0,00% | 0,00% |
| SC | 0,00% | 0,00% |
| MS | 0,00% | 0,00% |

Source: Authorial.

*4.5 Brainstorming for 2019 results analysis*

The expected results for the year 2019 were compared with the actual results obtained between 01/01/2019 and 25/11/2019. Subsequently, the presented divergences had their possible causes elucidated by brainstorming with the company's manager.

In 2019, while the benchmark scenario indicates 28.95% class A customers that showed some adherence behavior to the new product, for the real values obtained, 42.11% of the class A customers showed some record of the new product acquisition. Although not yet consolidated in these customers, this demonstrates a greater coverage of HTV 9299 coriander in this category. As a result, the sales performance on this product significantly exceeded the expected benchmark scenario, according Table 13.

**Table 13** - HTV 9299 coriander representativeness comparison between real and benchmark scenarios from customer grouping by ABC analysis

| *Class* | *Projected Representativeness (2019* | *Real Representativeness (2019)* | *Variation (Real/Benchmark)* |
|---|---|---|---|
| *A* | 7,61% | 13,21% | +73,58% |
| *B* | 1,72% | 4,82% | +180,19% |
| *C* | 6,80% | 22,66% | +233,24% |

Source: Authorial.

Table 14 and Table 15 show the comparison between the actual and projected results by customers' geographic region.

**Table 14** - Coriander real and benchmark scenarios representativeness for coriander by geographic criteria.

| **UF** | **Coriander Expected Representativeness (2019)** | **Coriander Real Representativeness (2019)** | **Variation (Real/Benchmark)** |
|---|---|---|---|
| **PE** | 22,55% | 21,85% | -3,10% |
| **SE** | 10,02% | 15,28% | 52,44% |
| **PA** | 15,90% | 14,55% | -8,48% |
| **CE** | 13,78% | 14,43% | 4,74% |
| **PB** | 13,53% | 13,00% | -3,90% |
| **RN** | 6,80% | 6,00% | -11,72% |
| **GO** | 6,37% | 5,75% | -9,64% |
| **AL** | 7,04% | 4,97% | -29,36% |
| **MT** | 1,21% | 1,68% | 39,44% |
| **ES** | 0,64% | 1,16% | 83,19% |
| **BA** | 0,90% | 0,70% | -21,98% |
| **MA** | 0,33% | 0,26% | -20,43% |
| **PI** | 0,65% | 0,19% | -70,17% |
| **AP** | 0,00% | 0,12% | - |
| **RO** | 0,22% | 0,04% | -81,40% |

| | | | |
|---|---|---|---|
| **AC** | 0,01% | 0,01% | -42,30% |
| **DF** | 0,07% | 0,00% | - |
| **MG** | 0,00% | 0,00% | - |
| **RR** | 0,00% | 0,00% | - |
| **SP** | 0,00% | 0,00% | - |

Source: Authorial.

**Table 15 -** HTV 9299 coriander real and benchmark scenarios representativeness for coriander by geographic criteria.

| UF | HTV 9299 Coriander Expected Representativeness (2019) | HTV 9299 Coriander Real Representativeness (2019) | Variation (Real/Benchmark) |
|---|---|---|---|
| **PE** | 58,22% | 36,55% | -37,21% |
| **SE** | 19,50% | 22,48% | 15,31% |
| **PB** | 9,65% | 16,85% | 74,64% |
| **ES** | 0,09% | 8,00% | 8429,15% |
| **AL** | 0,37% | 7,88% | 2027,94% |
| **GO** | 6,73% | 3,94% | -41,45% |
| **CE** | 0,82% | 1,64% | 98,80% |
| **BA** | 0,96% | 0,77% | -20,04% |
| **RN** | 1,01% | 0,73% | -27,82% |
| **PA** | 1,59% | 0,68% | -57,02% |
| **MT** | 0,35% | 0,48% | 36,66% |
| **AC** | 0,02% | 0,00% | - |
| **AP** | 0,00% | 0,00% | - |
| **DF** | 0,09% | 0,00% | - |
| **MA** | 0,23% | 0,00% | - |
| **MG** | 0,00% | 0,00% | - |
| **PI** | 0,13% | 0,00% | - |
| **RO** | 0,24% | 0,00% | - |
| **RR** | 0,00% | 0,00% | - |
| **SP** | 0,00% | 0,00% | - |

Source: Authorial.

The high divergence between the real and expected values do not invalidate the proposed model; neither do they mean that the forecasts are inaccurate, as this forecasting model does not aim at a forecast model that accurately reflects the results for a production planning, for example. It is worth remembering that the model proposes a forecast for the benchmark scenario, i. e., the maintenance of the customer acquisition behavior without the influence of strategic actions or other external factors such as stock instability, competition performance, among others. Thus, the proposed model supports the commercial management by: (i) directing annual strategic actions, according section 4.4; (ii) monitoring the results of these strategic actions and identifying any other factors that have an influence these scenarios, as presented in this section. In 2019, it can be notice that a number of factors showed a strong divergence between the scenarios.

In brainstorming with the company's manager, the key commercial actions taken in 2019 that may have influenced on the divergence between the real and forecast scenarios are described qualitatively in Table 16.

**Table 16 -** Commercial actions taken by the company in 2019 that have influenced the sales forecast

| Effect | Cause |
|---|---|
| **Increased HTV 9299 coriander sales in Sergipe (SE), Alagoas (AL) and Paraíba (PB) States.** | High period of HTV 0699 coriander stock unavailability retracted the sale indicators in these regions. |
| **Significant HTV 9299 coriander sales increase in the Espírito Santo (ES) State.** | A successful commercial action was taken to insert the HTV 9299 coriander in this region. |
| **Significant coriander sales increase in Sergipe (SE) State.** | In 2019, it was incorporated a main client in customer portfolio, but who is not prone to HTV 9299 coriander. |
| **HTV 9299 coriander sales below forecast in Goiás (GO) State.** | Cause not identified. |
| **Coriander sales below forecast in Alagoas (AL) State.** | Cause not identified. |
| **Greater adherence to HTV 9299 coriander for smaller customers (class C) compared to larger ones (class A).** | Cause not identified. |

Source: Authorial.

The benchmark scenario proposed in this model, although divergent from the real scenario, is consistent with the commercial actions outlined. In fact, the comparison between the quantitative sales forecast and commercial actions taken in 2019 may highlight the real effect that these had on sales results. In addition, this comparison makes visible divergences not easily observable by the company's management.

Although a more favorable behavior can be noticed for the HTV 0699 coriander incorporation in the customer portfolio, the commercial actions carried out differ from what was suggested in this analysis. For 2020, it is suggested that commercial actions focus on the regions of Pará (PA) and Ceará (CE) Sates and the further consolidation of the new product in class A customers, according section 4.4.

## 6. CONCLUSION

The article presents an analysis focused on the segments of the customer portfolio, in which it highlights that a probabilistic modeling in specific segments, added to a qualitative analysis, can have a high impact on the overall sales results of the new cultivar, as well as the entire coriander seeds group. In addition, the results presented may act as an indicator of greater or lesser adherence to the product consolidation process in the customer segment over the forecast year, since these are in concordance to the historical adherence behavior.

The proposed model presents a scenario of historical customer behavior maintenance, without any strategic action being taken by the sales team. Managerially, this base scenario (benchmark) provides quantitative support for the direction of commercial efforts to act in order to maximize the cultivar sales indicator, accelerating the portfolio replacement process and aligning with business strategies. In parallel, the comparison between the future and benchmark scenarios also shows itself as a possible managerial tool, by facilitating the identification of the causes that may have an effect on the sales behavior in the various segments of the customer portfolio.

The low computational effort required, due to the customer portfolio segmentation by ABC classification, in addition to the application of tools such as Solver and programming in VBA language trough MS-Excel® software, ensured the operationalization of demand forecast in several segments of the company's customer portfolio, presenting a characteristic of replicability for other similar forecasts, including with a more robust customer portfolio.

This article's proposed aims to create a benchmark scenario in opposition to the one influenced by changes in the company's commercial behavior and other exogenous factors. Therefore, it filled literature gap and introduces a quantitative tool for directing the customer portfolio decision-making processes in the context of product portfolio update. As result of a hybrid model based on usual demand forecasting models developing, this method combines the main individual advantages that the single application would result in, which is the transient demand pattern trough MC and the evolution of aggregate demand for exponential smoothing. In addition, it allows the characterization of this behavior in various segments of the forecast dataset.

In situations of product portfolio changes, characterized by a high instability in sales distribution, however, it is suggested for the method to be applied to the forecast only to a time step (year). Since the transient behavior between products is very sensitive to the strategic actions that will be taken each year, the subsequent years will suffer considerable changes in the sales distribution between products. In more stable scenarios analyzes, which involve less modification of the historical sales behavior, as already consolidated product portfolio, the demand forecast for Markov Chains has the potential to predict 2 or more years ahead, although it has not been tested in this work.

Some proposals for future work are: (i) customer grouping with a higher number of criteria, as sales channels, profitability, visibility offered, among others, in order to identify other possible relevant marketing action plans; (ii) data clustering techniques for similar behavioral grouping to compose the aggregate demand forecast per customer; (iii) updating the results obtained, incorporating the subjective aspect, after eliciting an expert opinion (for example, sales manager), according to modeling by Bayesian Networks; (iv) forecasting by this hybrid method under conditions of greater transactional stability and database robustness.

## References

ALBRECH, A.; FAUSTO, D.; GASPAR, V.; MIGLIAVACCA, R.; MILANI, A. Processo de gestão da carteira de clientes. **Revista iPecege**, v. 1, n. 2, p. 169-186, 2015.

AMARAL, G.; DE SOUZA, L.; FILHO, L. A study about the process of product-service development in a company of sharpening and recovery of industrial cutting tools. **GEPROS. Revista Gestão da Produção, Operações e Sistemas**, v. 15, n. 2, p. 175-203, 2020.

ASHOKA, N. **Dynamics of coffee marketing and export:** an agribusiness approach. Dharwad: University of Agricultural Sciences, 2013.

ALVES, H.; CAMPOS, F.; NEVES, A. Aplicação da técnica criativa "Brainstorming Clássico" na geração de alternativas na criação de games. In: SIMPÓSIO BRASILEIRO DE JOGOS PARA COMPUTADOR E ENTRETENIMENTO DIGITAL, n.6, v. 1., 2007, Porto Alegre. **Anais**… Recife: Universidade Federal de Pernambuco, PE, 2007.

BAXTER, M. **Projeto de produto:** guia prático para o design de novos produtos. São Paulo: Editora Blucher, 2000.

BARROW, D.; KOURENTZES, N.; SANDBERG, R.; NIKLEWSKI, J. Automatic robust estimation for exponential smoothing: perspectives from statistics and machine learning. **Expert Systems with Applications**, v. 160, n. 113637, 2020.

DIAS, M. **Administração de materiais:** uma abordagem logística. 5. ed. São Paulo: Atlas, 2010.

FAVERO, E. P. **Métodos de previsão de vendas:** um estudo de caso numa rede varejista. São Carlos: Escola de Engenharia da Universidade de São Paulo, 2015.

GAGLIARDI, F.; KAPELAN, Z.; FRANCHINI, M. A Probabilistic Short-Term Water Demand Forecasting Model Based on the Markov Chain. **Water**, v. 9, n. 7, p. 354–368, 2017.

HAJIRAHIMI, Z.; KHASHEI, M. Hybrid structures in time series modeling and forecasting: a review. **Engineering Applications of Artificial Intelligence**, v. 86, p. 83-106, 2019.

HINOJOSA, A.; MILANÉS, A. **Uma introdução aos processos estocásticos com aplicações.** Departamento de Estatística ICEx UFMG: Belo Horizonte, 2011.

LI, L.; LI, Y.; SUN, J.; XUAN, H. Mathematical model based on the product sales market forecast of markov forecasting and application. **Journal of Chemical and Pharmaceutical Research**, v. 6, n. 6, p. 1359–1365, 2014.

MAGELA, M. **Teoria básica das cadeias de Markov.** Vitória: Universidade Federal do Esprito Santo, 2015.

MARQUES, J.; FREITAS, D. Método DELPHI: caracterização e potencialidades na pesquisa em Educação. **Pro-Posições**, v. 29, n. 2, p. 389–415, 2018.

MELO, T.; FUCIDJI, J. Racionalidade limitada e a tomada de decisão em sistemas complexos. **Revista de Economia Política**, v. 36, n. 3, p. 622-645, 2016.

MOREIRA, D. **Administração da produção e operações.** 2. ed. rev. São Paulo: Cengage Learning BR, 2012.

PEINADO, J.; GRAEML, A. **Administração da produção:** operações industriais e de serviços. Curitiba: UnicenP, 2007.

RIBEIRO, G. **Proposta de classificação de métodos de previsão de demanda para novos produtos:** estudo no sistema brasileiro de franquias. 2016. 171 f. Dissertação (Mestrado em Engenharia de Produção) - Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2016.

ROSS, S. **Introduction to Probabity Models**. 10. ed. Los Angeles: Elsevier, 2010.

SANTOS, R. B. N.; BRAGA, M. J. Evolução das disparidades da extração vegetal e da silvicultura na Amazônia Legal: uma aplicação da cadeia de Markov. **Ciência Rural**, v. 42, n. 9, p. 1701-1706, 2012.

VAGO, F; FAGUNDES, A.; LARA, J.; MELO, J., SAMPAIO, D.; SOUSA, C. A Importância Do Gerenciamento De Estoque Por Meio Da Ferramenta Curva Abc. **Revista Sociais e Humanas**, v. 26, n. 3, p. 638–655, 2013.

VIANA, J. **Administração de materiais:** um enfoque prático. 1. ed. São Paulo: Atlas, 2006.

ZHOU, Y. TAN, Y. WANG, L. ZHONG, R. A Markov Chain Based Demand Prediction Model for Stations in Bike Sharing Systems. **Mathematical problems in engineering**, v. 2018, p. 1–8, 2018.