Taylor & Francis
Taylor & Francis Group

International Journal of Production Research

ISSN: 0020-7543 (Print) 1366-588X (Online) Journal homepage: http://www.tandfonline.com/loi/tprs20

Robust optimisation approach applied to the analysis of production / logistics and crop planning in the tomato processing industry

Cleber D. Rocco & Reinaldo Morabito

To cite this article: Cleber D. Rocco & Reinaldo Morabito (2016): Robust optimisation approach applied to the analysis of production / logistics and crop planning in the tomato processing industry, International Journal of Production Research, DOI: 10.1080/00207543.2016.1181284

To link to this article: http://dx.doi.org/10.1080/00207543.2016.1181284

Published online: 10 May 2016.

International Journal of Production Research, 2016
http://dx.doi.org/10.1080/00207543.2016.1181284

Robust optimisation approach applied to the analysis of production / logistics and crop planning in the tomato processing industry

Cleber D. Rocco[a] and Reinaldo Morabito[b]*

[a]*School of Applied Science, Universidade Estadual de Campinas, Limeira, Brazil;* [b]*Department of Production Engineering, Universidade Federal de São Carlos, Sao Carlos, Brazil*

(*Received 5 October 2015; accepted 10 April 2016*)

The soluble solids content in the tomato fruit, also known as 'brix', and the crop yield are the most relevant uncertain parameters to determine technical and economic performance in the tomato processing industry. This paper presents a linear programming model and three robust optimisation models to deal with data uncertainty in the analysis of crop, logistics and industrial tactical planning in this industry. We focused the analysis on the production and logistics costs due to the impacts of unfavourable disturbances on the amount of soluble solids and the quantity of tomatoes processed in the system. A typical industry in this sector collaborated with this study by providing real data of its production, logistics and crop plans and with in-depth discussions. From the results, some general conclusions were outlined and we discuss the benefits of adopting the robust optimisation approach instead of a deterministic one. The robust approach proved to be a powerful tool for elaborating scenarios for uncertainty analysis in medium-term decisions, as described in this study, and clearly has potential to be employed in real contexts.

Keywords: aggregate production and logistics planning; tomato processing industry; crop planning; linear programming; robust optimisation

1. Introduction

A significant amount of information is required for the tactical planning of agricultural and industrial activities in the tomato processing industry. Appropriate information management with decision support tools based on optimisation techniques can offer a competitive advantage in this highly competitive sector worldwide. Examples of decisions that are usually based on information management are: planning the planting and harvesting of tomatoes, transporting tomatoes from producing regions to processing plants, medium term production and inventory management of concentrated tomato pulp and final products (e.g. ketchup, sauces, etc.) for consumers, besides the transportation of these products between demand points in the supply chain. A profitable business in the tomato processing sector requires the combined coordination of agricultural and industrial activities, which is rarely done without effective analytical tools. In this study, a typical tomato processing company in Brazil contributed by supplying data and analysing the results. Brazil was the 7th largest producer of processing tomatoes in 2014, after the USA, China, Italy, Spain, Iran and Turkey (WPTC 2015). The entire Brazilian tomato sector (processing and fresh) is the 12th largest in production quantity, and is larger than the potatoes, beans, pork, cotton and coffee sectors, with an estimated production value of around US$ 2 billion (IBGE 2015).

In this paper, we introduce a deterministic linear optimisation model that can be used to analyse and support tactical planning decisions in the tomato processing industry. This model was first presented in Rocco (2014), and a related model with additional company features of a case study, such as energy consumption in the evaporation, was reported in Rocco and Morabito (2015). Here, the deterministic model is extended via robust optimisation techniques to approach uncertainty in the values of the two most relevant parameters of this agro-industrial system, viz., the soluble solids content in the tomato fruit and the crop yield. By means of the robust approach, we generated several useful results using real data for the decision-makers to corroborate its practical use.

Production and logistics planning models in agro-industry have been reported for several sectors, such as sugar and alcohol (Kawamura, Ronconi, and Yoshizaki 2006; Paiva and Morabito 2009; da Silva, Silva Marins, and Barra Montevechi 2013), citrus (Caixeta-Filho 2006; Munhoz and Morabito 2014), pulp and paper mill (Santos and Almada-Lobo 2012) and grain (Rosentrater and Kongar 2009; Junqueira and Morabito 2012). However, to the best of

*Corresponding author. Email: morabito@ufscar.br

our knowledge, we are not aware of studies reported in the literature that modelled the tomato processing industry and employed robust optimisation techniques to approach uncertainties in input data in this line of research, particularly to analyse the crop and production/logistics tactical planning decisions.

The main contribution of this study is to present three robust linear optimisation models for the tomato processing industry that are capable of analysing and support tactical planning decisions in agricultural and industrial activities. The robust approach showed advantage over the deterministic one in the elaboration and analysis of risk scenarios considering data uncertainty. For instance, some of the interesting insights obtained by the model's results are that managers should give preference to high brix levels rather than high crop yields with low soluble solids content in the tomato fruit. As the models presented here are based on a typical tomato processing industry, we believe that they can be directly applied or be easily adapted to deal with other companies and related agro-industrial sectors (e.g. sugar-alcohol, guava). This paper is organised in five additional sections, besides this introduction. In Section 2, agricultural and industrial tactical planning for the tomato processing industry is briefly presented. In Section 3, the deterministic linear model is described. In Section 4, the robust variants for the tomato brix and crop yield parameters, and both jointly in the same model, are presented. In Section 5, the results from computational experiments by the robust models using real data are discussed. Finally, Section 6 outlines the concluding remarks and perspectives for future studies.

2. Agricultural and industrial planning in the tomato processing industry

In this section, a brief description of agricultural and industrial planning in the tomato processing industry is presented. The description is focused on the tactical decisions, which must be made every year to design the tomato crop and industrial operations, and should be revisited each month to deal with possible data changes and uncertainties. These decisions could be organisationally grouped into two main stages according to the work teams: the agricultural and industrial stages. In the agricultural stage, some of the key decisions are the sizes of the areas to cultivate tomatoes for supplying the processing plants and the agricultural tomato variety in each cultivation period in the producing regions, since there are several tomato varieties with distinguished agronomic features, such as crop yield, resistance to pests, disease and drought, maturation curves (brix progress), etc. The logistics of transporting tomatoes from the fields to the plants is also planned by the agricultural team. Some companies have a single processing plant, whereas others have several plants in different regions that need to be supplied with tomatoes.

In the industrial stage, the decisions are related to the production planning of concentrated tomato pulp, also referred to as semi-finished product, which can be of two types under different concentrations of soluble solids. One type is simply called 'paste', which is the concentrated tomato pulp without seeds in a concentration typically ranging from 18 to 30%. The other type is called 'crush', which is the concentrated tomato pulp with a higher particle size and containing tomato seeds, usually with a concentration of soluble solids of 16% up to 22%. Concentrated tomato pastes with lower concentrations of soluble solids are consumed directly to produce final products to consumers, without being stored, whereas high brix concentrated pastes are sent to the stocks that will be consumed throughout the period between tomato seasons.

The plant configuration to produce semi-finished products determines whether the company will have enough paste and crush to meet the demand for manufacturing final products throughout the year. If not enough pulp was produced, the company would need to buy it from the market at prices far above the production costs. Companies owning more than one plant should plan where to produce pulp and crush, and the corresponding quantities in each plant. Processing plants may be installed close to the agricultural regions or near the consumers in urban areas. A well-known adopted strategy for this matter is to produce semi-finished products in plants close to the growing fields and transport them to plants that are close to the cities, where the semi-finished products are consumed for manufacturing final products.

Tomato harvesting in Brazil takes place over a period of four months a year, from approximately July to October. The entire quantity of harvested tomatoes should be processed immediately after arriving at the plant due to its perishability, since the losses in terms of quantity and quality increase as time goes by, as pointed out in Gameiro et al. (2008). Therefore, the production from the agricultural stage is figuratively pushed to the industrial stage until the production of concentrated pulp. On the other hand, the production of final products in the industrial stage is pulled by the demand from the markets, which varies mostly according to the season of the year. Thus, there is a kind of decoupling point in the tomato processing industry specifically positioned between the production of semi-finished products and final products for consumers. Thus, from the point of view of production planning and control, the industrial stage may be divided into two parts, separating the make-to-stock operation (1st industrial stage: production of concentrated tomato pulp) and the make-to-order operation (2nd industrial stage: production of final products for consumers). Figure 1 presents a conceptual scheme of the tomato agro-industrial system modelled mathematically in this study.

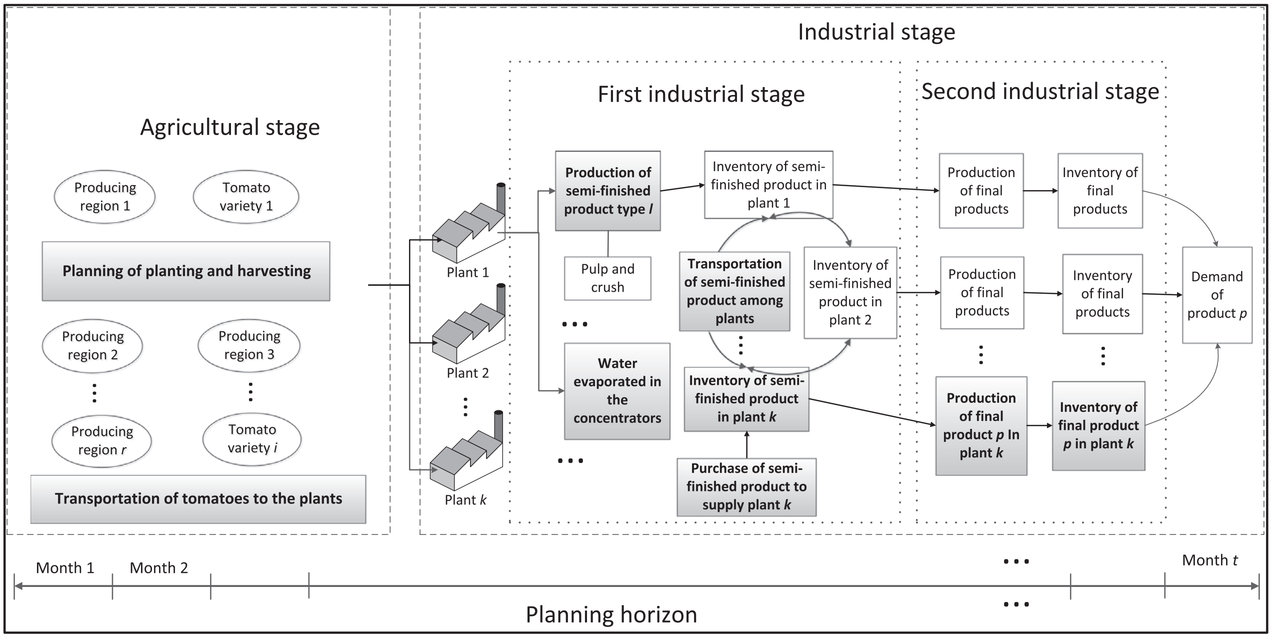

Figure 1. Conceptual scheme of the tomato agro-industrial system.

3. The deterministic model

A linear mathematical programming model was built to represent and optimise the key tactical decisions in the agricultural and industrial stages in the tomato processing industry. This model is deterministic and its building concept is based on the aggregated production planning performed in coupled multiple capacitated stages, with multiple products and periods. The objective function (OF) is designed to reflect the trade-off between the production and logistics costs and the benefit of making stocks of semi-finished products. The OF minimises the costs of transporting tomatoes from the growing fields to the plants, the cost of semi-finished products procurements due to eventual shortages to meet the demand, the cost of keeping stocks of semi-finished products and final products and the cost of transporting semi-finished products among plants. The economic benefit of making stocks of semi-finished products is performed by accounting for the quantities stored in the last period of the planning horizon. The aim of the model is to calculate the allocation of the tomato areas in the available regions, and consequently the production of tomatoes to supply the processing plants, aligned with the production and inventory policy of semi-finished products, for the purpose of meeting the demand for manufacturing final products.

The model was designed for a time horizon of 12 months, from January to December, comprising one tomato season and the intercrop periods. The final products of the demand were grouped in a few families according to the similarity of their compositions in terms of soluble solids content. The capacities of the concentrators/evaporators in a given plant were aggregated, without distinguishing multiple pieces of equipment. Using the model, an entire year's scenario for the key agricultural and industrial activities in the tomato processing industry can be created, with single or multiple processing plants, related to the production of tomatoes in agricultural regions, industrial operations to produce semi-finished products and final products and the logistics involved, i.e. transport and inventory management.

Besides all these model capabilities, the scope of this paper is to use the robust models derived from the deterministic model as a tool to perform sensitivity analysis for investigating the economic effects of disturbances to unfavourable cases in the tomato brix and crop yield data, which are considered the most relevant parameters in the tomato processing industry.

Model sets

T Time periods of the planning horizon, $t \in T$.
I Tomato varieties, $i \in I$.
R Producing regions, $r \in R$.
K Processing plants, $k, \hat{k} \in K$.

L Semi-finished products, $l \in L$.
G Group of semi-finished products, $g, \hat{g} \in G$.
P Family of final products for consumers, $p \in P$.
L_g Set of semi-finished products in group g.
L_p Set of semi-finished products which may be used in final product family p.
P_k Set of final products which may be produced in plant k.

Model parameters

β_{kt} Reception capacity of tomatoes in plant k in period t (ton/month).
γ_r Available area to cultivate tomatoes in growing region r (hectare).
δ_{irt} Yield of tomato variety i cultivated in region r in period t (ton/ha).
ε_{irt} Brix of tomato variety i cultivated in region r in period t (%).
η_{rk} Cost of transporting tomatoes from region r to plant k (\$/ton).
τ_l Storage cost for semi-finished product l (\$/ton).
θ_l Brix of semi-finished product l (%).
μ_l Price to purchase semi-finished product l on the market (\$/ton).
$\rho_{k\hat{k}}$ Cost of shipping semi-finished products from plant k to plant $\hat{k}$ (\$/ton).
h Storage cost of final products (\$/ton).
υ_l Economic benefit of storing semi-finished product l (\$/ton).
λ_{gp} Soluble solids content of semi-finished products from group g in the composition of products in family p (% of mass).
d_{pt} Aggregate demand for final products in family p in period t (ton/month).
π_{kt} Aggregate capacity of water evaporation in plant k in period t (ton/month).

Model variables

Z Variable to account for the OF value of the model (\$).
Q^{tom}_{irkt} Quantity of tomatoes of variety i harvested in region r to supply plant k in period t (ton).
P^{raw}_{lkt} Production of semi-finished product l in plant k in period t (ton).
M^{raw}_{lkt} Procurement of semi-finished product l to supply plant k in period t (ton).
$T^{\text{raw}}_{lk\hat{k}t}$ Transport of semi-finished product l from plant k to plant $\hat{k}$ in period t (ton).
I^{raw}_{lkt} Inventory of semi-finished product l in plant k in period t (ton).
C^{raw}_{lpkt} Consumption of semi-finished product l for manufacturing final product p in plant k in period t (ton).
I^{prod}_{pkt} Inventory of final product p in plant k in period t (ton).
P^{prod}_{pkt} Production of final product p in plant k in period t (ton).
C^{prod}_{pkt} Consumption of final product p from plant k in period t to meet the demand (ton).

Mathematical formulation

$$\text{Minimise}: \ Z = \sum_{\substack{irkt \\ \delta_{irt} \neq 0}} \eta_{rk} Q^{\text{tom}}_{irkt} + \sum_{lkt} \tau_l I^{\text{raw}}_{lkt} + \sum_{lkt} \mu_l M^{\text{raw}}_{lkt} + \sum_{\substack{lk\hat{k}t \\ k \neq \hat{k}}} \rho_{k\hat{k}} T^{\text{raw}}_{lk\hat{k}t} + \sum_{pkt} h I^{\text{prod}}_{pkt} - \sum_{kl} \upsilon_l I^{\text{raw}}_{lk|T|} \tag{1}$$

$$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} Q^{\text{tom}}_{irkt} / \delta_{irt} \leq \gamma_r \quad \forall\, r \tag{2}$$

$$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} Q^{\text{tom}}_{irkt} \leq \beta_{kt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0 \tag{3}$$

$$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} \geq \sum_{l} \theta_l P^{\text{raw}}_{klt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0 \tag{4}$$

$$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} Q^{\text{tom}}_{irkt} - \sum_{l} P^{\text{raw}}_{klt} \leq \pi_{kt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0 \tag{5}$$

$$I^{\text{raw}}_{lkt} = I^{\text{raw}}_{lk,t-1} - \sum_{p \in P_k} C^{\text{raw}}_{lpkt} + P^{\text{raw}}_{lkt} + M^{\text{raw}}_{lkt} + \sum_{\hat{k}} T^{\text{raw}}_{l\hat{k}kt} - \sum_{\hat{k}} T^{\text{raw}}_{lk\hat{k}t} \quad \forall\, l,\, k,\, t \tag{6}$$

$$\sum_{l \in \left(L_g \cup L_p\right)} \theta_l C^{\text{raw}}_{lpkt} \geq \lambda_{gp} P^{\text{prod}}_{pkt} \quad \forall\, g,\, t,\, k,\, p \in P_k \tag{7}$$

$$I^{\text{prod}}_{pkt} = I^{\text{prod}}_{pk,t-1} - C^{\text{prod}}_{pkt} + P^{\text{prod}}_{pkt} \quad \forall\, p,\, k,\, t \tag{8}$$

$$\sum_{k} C^{\text{prod}}_{pkt} \geq d_{pt} \quad \forall\, p,\, t,\, d_{pt} > 0 \tag{9}$$

$$\begin{aligned} &Q^{\text{tom}}_{irkt} \geq 0 \quad \forall i,\, r,\, k,\, t;\; P^{\text{raw}}_{lkt} \geq 0 \quad \forall l,\, k,\, t;\; M^{\text{raw}}_{lkt} \geq 0 \quad \forall l,\, k,\, t;\; T^{\text{raw}}_{lk\hat{k}t} \geq 0 \quad \forall l,\, k,\, \hat{k},\, t;\; I^{\text{raw}}_{lkt} \geq 0 \quad \forall l,\, k,\, t;\; C^{\text{raw}}_{lpkt} \geq 0 \\ &\forall l,\, p,\, k,\, t;\; I^{\text{prod}}_{pkt} \geq 0 \quad \forall p,\, k,\, t;\; P^{\text{prod}}_{pkt} \geq 0 \quad \forall p,\, k,\, t;\; C^{\text{prod}}_{pkt} \geq 0 \quad \forall\, p,\, k,\, t. \end{aligned} \tag{10}$$

The OF (1) minimises the main production and logistics costs in the tomato processing industry. The first cost component accounts for the transportation of tomatoes from the growing fields to the processing plants. The second is the inventory cost of semi-finished products. The third is the cost of semi-finished product procurements from the market. The fourth cost component refers to the transportation of semi-finished products among plants. The fifth is the inventory cost of final products, and the last component, with a negative sign, is the economic benefit of making stocks of semi-finished products in the last period of the planning horizon.

Constraints (2) define the limit of area available to cultivate tomatoes in the growing regions. Constraints (3) establish the reception capacity of tomatoes for each plant in the system. Constraints (4) are responsible for making semi-finished products from the tomato soluble solids. Constraints (5) set the capacity of water evaporation in each plant to make semi-finished products. Equations (6) are the inventory balance of semi-finished products. Constraints (7) are responsible for blending the semi-finished products for making the final products. Equations (8) are the inventory balance of final products. Constraints (9) refer to the demand of final products, which the company must meet without delay. Constraints (10) are the non-negativity of the decision variables in the model. In the worst case, the total number of model constraints is given by $|R| + |T|[|K|(3 + |L| + |G||P| + |P|) + |P|]$ and the model variables by $|K||T|[|I||R| + |L|(|K|/2 + |P| + 3) + 3|P|]$.

The mathematical model (1)–(10) appropriately represents the main tactical planning decisions in the tomato processing industry. Some model parameters naturally carry uncertainties in their values, for example, the tomato brix and the crop yield at harvest in the growing regions. From the historical data, the average values for the brix and crop yield of each tomato variety, region and harvesting time could be obtained, which were represented in the model by parameters ε_{irt} and δ_{irt}, respectively. However, there are no guarantees that the estimated average values will be observed in practice. Agricultural teams know that adverse weather may cause significant changes in the values of the tomato brix and crop yield. For instance, if there are heavy rains throughout the tomato cultivation cycle, the crop yield and brix may drop significantly, affecting industrial yield and business profitability. These two parameters from the agricultural stage are considered the most relevant by managers in this industry and are the target for applying the robust optimisation approach to provide solutions outlined in scenarios that managers could use to guide their decisions. Other parameters are also subject to uncertainties, such as the demands of final products, the costs of the components in the OF, the availability of areas for tomatoes, etc. However, these are not considered in this study and are interesting topics for future research.

4. The robust optimisation approach

Robust optimisation techniques to build robust models have been widely discussed by some authors in the specialised literature, such as Gabrel, Murat, and Thiele (2014), Beyer and Sendhoff (2007), Bertsimas and Thiele (2006), Bertsimas and Sim (2003, 2004), Ben-Tal and Nemirovski (1998, 1999, 2000), El Ghaoui, Oustry, and Lebret (1998), El Ghaoui and Lebret (1997), Mulvey, Vanderbei, and Zenios (1995) and Soyster (1973). Some recent examples of practical applications in production planning can be found in Bohle, Maturana, and Vera (2010), Leiras, Hamacher, and Elkamel (2010), Alvarez and Vera (2011), Alem and Morabito (2012), Najafi, Eshghi, and Dullaert (2013) and Munhoz and Morabito (2014).

The motivation for choosing the robust optimisation approach over stochastic programming (Birge and Louveaux 1997) to deal with uncertainty in model parameters is because the robust approach does not require the probabilistic distribution of the data, which is rarely available in the real context of the tomato processing industry, as well as to avoid dealing with a large number of scenarios for the solution analysis. Here, we briefly describe the essence of the robust optimisation technique; for a comprehensive and detailed description of the technique used here, the readers are referred to Bertsimas and Sim (2003, 2004) as a suggestion.

Given the linear optimisation model $\text{Min}_{x \geq 0}\{c^T x | Ax \leq b\}$, a set J_i is defined as the coefficients subject to uncertainty in row i of the technological matrix A. Each parameter subject to uncertainty, defined as $\tilde{a}_{ij}$, $j \in J_i$, is an independently symmetric distributed random variable with a mean equal to the nominal value a_{ij} and varying into the interval $\left[a_{ij} - \hat{a}_{ij},\, a_{ij} + \hat{a}_{ij}\right]$. The scale deviation $z_{ij} \in [0, 1]$ is defined as $z_{ij} = \left(\tilde{a}_{ij} - a_{ij}\right)/\hat{a}_{ij}$, where $\hat{a}_{ij}$ can be understood as the maximum variation which the uncertain parameter may take.

Bertsimas and Sim (2004) claimed that it is unlikely that all parameters $\tilde{a}_{ij}$ take the worst-case value at the same time in the optimal solution. Therefore, the goal in building a robust solution is to be protected against a certain number of parameters in their worst-case values, i.e. either $a_{ij} - \hat{a}_{ij}$ or $a_{ij} + \hat{a}_{ij}$, depending on constraint i and parameter j. The approach introduces in constraint i (where $\tilde{a}_{ij}$ is present) the parameter Γ_i, which takes the value in the interval $[0, |J_i|]$ to adjust the robustness of the solution. Parameter Γ_i can control the degree of conservatism for every constraint i, protecting it deterministically against its violation. The robust model is equivalent to the deterministic case when $\Gamma_i = 0$, in which all parameters in constraint i assume their nominal values, i.e. a_{ij}. At the other extreme, the model is equivalent to Soyster's (Soyster 1973) formulation when $\Gamma_i = |J_i|$, where all parameters subject to uncertainty assume their worst-case values and the optimal solution value of the model is usually heavily penalised. The trade-off between the deterministic less conservative solution and Soyster's highly conservative solution is defined by the level of conservatism established by the robust parameter Γ, also called the *price of robustness*. The robust optimisation approach presented by Bertsimas and Sim (2004) can be written as the following model:

$$\text{Minimise} \sum_j c_j x_j$$

$$\sum_j a_{ij} x_j + \lambda_i \Gamma_i + \sum_{j \in J_i} \mu_{ij} \leq b_i \quad \forall\, i$$

$$\lambda_i + \mu_{ij} \geq \hat{a}_{ij} x_{ij} \quad \forall\, i, j \in J_i$$

$$x_j \geq 0 \quad \forall\, j$$

$$\mu_{ij} \geq 0 \quad \forall\, i, j \in J_i$$

$$\lambda_i \geq 0 \quad \forall\, i.$$

Variables λ_i and μ_{ij} in the robust model above are the dual variables that appear in the method and that avoid non-linear formulations. Note that the robust model above, derived from the deterministic linear model, is also a linear optimisation model. The probability of violation of the ith constraint is calculated in the present study using the De Moivre-Laplace approximation of the binomial distribution $B(|J_i|, \Gamma_i) \approx 1 - \Phi\left(\frac{\Gamma_i - 1}{\sqrt{|J_i|}}\right)$, where: $\Phi(\theta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta} exp\left(-\frac{y^2}{2}\right) \mathrm{d}y$ is the cumulative distribution function of a standard normal for all i. A study on the probability bounds of constraint violation in robust optimisation can be found in Bertsimas and Sim (2004).

In this section, the deterministic model (1)–(10) presented previously is modified to build three robust variants, which incorporate uncertainty in the parameters of the technological matrix. The first variant incorporates uncertainty in the tomato brix parameter (ε_{irt}) only, the second variant incorporates uncertainty in the crop yield (δ_{irt}) only and the third variant incorporates uncertainty in these two parameters into the same robust model. For the sake of departing from the simplest start, we first present the robust model for the brix parameter, since the crop yield parameter is in the denominator of the constraints and requires a few algebraic adjustments before performing the robust optimisation technique, which is not necessary in the robust model for the tomato brix parameter.

4.1 *Uncertainty in the tomato brix*

The parameter in the model which indicates the tomato brix (ε_{irt}) is relevant because it determines the concentration of soluble solids in the tomato juice. If the tomato brix is reduced for whatever reason (e.g. excessive rain at harvest), the less concentrated the tomato juice will be and more water will be necessary for evaporation to produce the concentrated tomato pulp. As the production capacity of semi-finished products is derived from the capacity of the concentrators to evaporate water, indicated by parameter π_{kt}, the tomato brix directly influences the industrial capacity. From model (1)–(10), constraints (4) are picked for modification to build the 'Robust Model Brix', shortened simply here as RMB.

Considering a given constraint (k, t) of the family of constraints (4), let J_t be the set of coefficients ε_{irt} in row t of the technological matrix, which is subject to uncertainty $(i, r \in J_t)$. We define $\hat{\varepsilon}_{irt}$ as the maximum range in which the deterministic parameter ε_{irt} may be disturbed. Let p be the disturbance in percentage relative to the deterministic value of ε_{irt}. The maximum uncertainty variation for the parameter tomato brix is given by $\hat{\varepsilon}_{irt} = \varepsilon_{irt}p$. Now, let $\tilde{\varepsilon}_{irt}$ be the uncertain parameter of ε_{irt}. Parameter $\tilde{\varepsilon}_{irt} \in J_t$ takes a value according to a symmetric distribution with a mean equal to the nominal value ε_{irt} in the interval: $[\varepsilon_{irt} - \hat{\varepsilon}_{irt}, \varepsilon_{irt} + \hat{\varepsilon}_{irt}]$. In the Appendix, we describe the key steps to building the robust counterpart of model (1)–(10) that consider the uncertainty in the tomato brix (Appendix 1). For every (k, t) constraint, we introduce the parameter Γ_{kt}, not necessarily an integer, that can take values in the interval $[0, |J_t|]$. The role of parameter Γ_{kt} is to adjust the robustness of the solution; in other words, it controls the trade-off between the probability of violating the constraint with the uncertain parameter and the conservative effect of the solution and the OF in the deterministic problem. As mentioned before, Bertsimas and Sim (2004) called this trade-off the *price of robustness*. The RMB derived from the deterministic model (1)–(10) is shown below. It has the same number of constraints as the deterministic model, in addition to the new constraints from the dual sub-problems of the protection function, which is calculated by the expression $|I||R||K||T|$. With the number of variables listed, the RMB adds $|K||T| + |I||R||K||T|$ compared to the deterministic version.

$$\text{Maximise: } (1)$$
$$\text{subject to: } (2), (3), (5)\text{–}(10),$$

$$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} - \Gamma_{kt}\,\lambda_{kt} - \sum_{(i,r)\in J_t} \mu_{irkt} \geq \sum_{l} \theta_l P^{\text{raw}}_{lkt} \quad \forall\ k,\ t,\ \sum_{ir} \delta_{irt} \neq 0 \tag{11}$$

$$\lambda_{kt} + \mu_{irkt} \geq \hat{\varepsilon}_{irt}\ Q^{\text{tom}}_{irkt} \quad \forall\ k,\ t,\ (i, r) \in J_t,\ \delta_{irt} \neq 0 \tag{12}$$

$$\lambda_{kt} \geq 0 \quad \forall\ k,\ t \tag{13}$$

$$\mu_{irkt} \geq 0 \quad \forall\ k,\ t,\ (i, r) \in J_t. \tag{14}$$

4.2 *Uncertainty in the crop yield*

The crop yield parameter in agricultural planning models determines the production level of the outputs. Overvalue or undervalue means obtaining non-realistic solutions for the entire model. Crop yield is of course subject to uncertainties from nature due to events that can undermine productivity in the field, such as rainfall, drought, pests, diseases and poor agricultural management practices. Therefore, crop yield parameters are the best examples for investigating the effects of uncertainties in agricultural planning models. Model (1)–(10) can be modified to cope with uncertainty in this parameter. We called the modified model 'Robust Model Yield', or simply RMY. Constraints (2) are picked for modification according to robust techniques. Considering a given constraint r of the family of constraints (2); let J_r be the set of coefficients δ_{irt} in row r of the technological matrix that is subject to uncertainty $(i, t \in J_r)$. For each constraint r, we introduce parameter Γ_r, which defines the budget of uncertainty in the model, and it can take values in the interval $[0, |J_r|]$.

Next, we proceed with a minor algebraic adjustment in constraints (2) in order to change the parameter $\delta_{irt}$from the denominator. Let Δ_{irt} be the inverse of δ_{irt}, i.e. $\Delta_{irt} = 1/\delta_{irt}, \delta_{irt} > 0$, which indicates the area required for producing a given quantity of tomatoes. Let p, already defined, be the disturbance, in percentage, of the parameter δ_{irt} The perturbed parameter for the unfavourable cases of the problem is Δ^p_{ikt}, calculated by expression (15). The boundary value in which the uncertain parameter $\tilde{\Delta}_{irt}$ may vary according to the random variable Z^Y_{irt} is given by $\hat{\Delta}_{irt}$ (expression (16)), i.e. $\tilde{\Delta}_{irt} \in \left[\Delta_{irt} - \hat{\Delta}_{irt}, \Delta_{irt} + \hat{\Delta}_{irt}\right]$.

$$\Delta^p_{ikt} = \left[\frac{1}{\delta_{irt}} + \frac{1}{\delta_{irt}\, p}\right] \cdot \left[\frac{1}{(1-p) + \frac{\delta_{irt}(1-p)}{\delta_{irt}\, p}}\right] \tag{15}$$

$$\hat{\Delta}_{irt} = -\left(\Delta_{irt} - \Delta^p_{ikt}\right) \tag{16}$$

In the Appendix, the key steps for building the robust model that considers uncertainty in the crop yield are presented (Appendix 2). The procedure is similar to the above described for the parameter tomato brix. The robust model for the crop yield parameter (RMY) derived from the deterministic model (1)–(10) is written below. In the worst-case scenario, it has the same number of constraints and variables of the deterministic model in addition to new constraints calculated by the expression $|I||R||T|$ and new variables calculated by $|R| + |I||R||T|$.

$$\text{Maximise: } (1)$$
$$\text{subject to: } (3)\text{–}(10),$$

$$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} \Delta_{irt}\, Q^{\text{tom}}_{irkt} + \Gamma_r \lambda_r + \sum_{(i,t) \in J_r} \mu_{irt} \leq \gamma_r \quad \forall\, r \tag{17}$$

$$\lambda_r + \mu_{irt} \geq \hat{\Delta}_{irt} \sum_k Q^{\text{tom}}_{irkt} \quad \forall\, r,\ (i,t) \in J_r,\ \delta_{irt} \neq 0 \tag{18}$$

$$\lambda_r \geq 0 \quad \forall\, r \tag{19}$$

$$\mu_{irt} \geq 0 \quad \forall\, r,\ (i,t) \in J_r. \tag{20}$$

4.3 *Uncertainty in the tomato brix and crop yield parameters*

The robust model considering uncertainty in the tomato brix and crop yield parameters simultaneously is referred to here as RMBY. This model is a combination of the modified constraints from RMB and RMY. The procedure for building the robust model is similar to that described in the Appendix and was omitted here for the sake of brevity. Compared to the deterministic model, the RMBY has $|I||R||T|(1 + |K|)$ new constraints and $|T|[|I||T|(1 + |K|) + |K|] + |R|$ new variables.

$$\text{Maximise: } (1)$$
$$\text{subject to: } (3), (5)\text{–}(10),$$

$$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} - \Gamma_{kt}\lambda_{kt} - \sum_{(i,r)\in J_t} \mu_{irkt} \geq \sum_{l} \theta_l P^{\text{raw}}_{lkt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0 \tag{11}$$

$$\lambda_{kt} + \mu_{irkt} \geq \hat{\varepsilon}_{irt} Q^{\text{tom}}_{irkt} \quad \forall\, k, t, (i, r) \in J_t,\ \delta_{irt} \neq 0 \tag{12}$$

$$\lambda_{kt} \geq 0 \quad \forall\, k,\, t \tag{13}$$

$$\mu_{irkt} \geq 0 \quad \forall\, k,\, t,\, (i, r) \in J_t \tag{14}$$

$$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} \Delta_{irt} Q^{\text{tom}}_{irkt} + \Gamma_r \lambda_r + \sum_{(i,t)\in J_r} \mu_{irt} \leq \gamma_r \quad \forall\, r \tag{17}$$

$$\lambda_r + \mu_{irt} \geq \hat{\Delta}_{irt} \sum_{k} Q^{\text{tom}}_{irkt} \quad \forall\, r, (i, t) \in J_r,\ \delta_{irt} \neq 0 \tag{18}$$

$$\lambda_r \geq 0 \quad \forall\, r \tag{19}$$

$$\mu_{irt} \geq 0 \quad \forall\, r,\, (i, t) \in J_r. \tag{20}$$

5. Model results

This section presents some results obtained from the RMB, RMY and RMBY robust models using real data supplied by a Brazilian company that collaborated with this research. The models were implemented in the General Algebraic Modelling System and solved by CPLEX 12.5 in its default settings. A simple personal computer with an Intel I3 processor was employed. In the data-set, there are four agricultural regions, three tomato varieties and three processing plants. In the first industrial stage, there are three types of concentrated tomato paste and two types of crush that can be produced. The final products for consumers are grouped into three families and they have monthly demands. More information about the experiments presented here and the data-set used can be found in Rocco (2014).

In the results presented below, the deterministic value of the parameters subject to uncertainty were perturbed unfavourably on five levels defined arbitrarily to outline the scenarios for analysis, in which the disturbance p assumed values of 0.1, 0.2, 0.3, 0.4 and 0.5. In other words, the deterministic value of the parameter subject to uncertainty had a reduction of p in the worst-case solution. For example, consider a given region r and a given tomato variety i that has a monthly agricultural yield of 60, 70, 85 and 95 ton/hectare and a tomato brix value of 3.78, 4.02, 4.27 and 4.3% over the crop season, from July to October. Figure 2 shows the nominal and disturbed tomato brix values and Figure 3 presents the curves of the values for the nominal and disturbed area requirement, which is the inverse of the crop yield parameter calculated by expression (15).

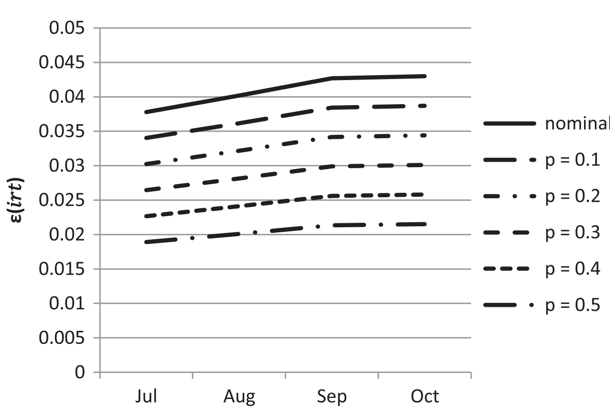

Figure 2. Brix parameter.

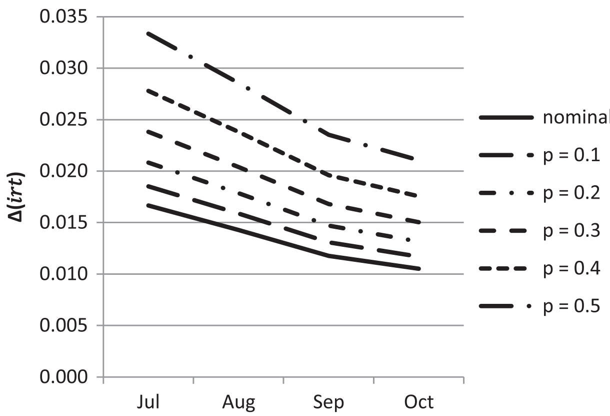

Figure 3. Area requirement.

5.1 *Results of the RMB*

The number of coefficients ε_{irt} subject to uncertainty in constraints (4) defined for each plant k and period t is 12, due to the three different types of tomato varieties ($|I| = 3$) multiplied by the four agricultural regions used to cultivate them ($|R| = 4$). The parameter to adjust the robustness of the model is defined by parameter Γ_{kt}, such that $\Gamma_{kt} = 0$ provides the deterministic solution and $\Gamma_{kt} = 12$, the equivalent to Soyster's. The RMB is solved successively with a unitary increment of Γ_{kt}, from $\Gamma_{kt} = 0$ to $\Gamma_{kt} = 12$, and the optimal OF value with its cost and benefit components and other relevant results is depicted in Figure 4: Graphs: A–J.

It is known that the optimal worst-case solution can be achieved with a number of uncertain parameters at the worst-value lower than the cardinality of the set of all possible parameters taking uncertainties. This can be observed in the results presented in Table 1. Note that for all p values, the optimal OF achieves its worst value with the robustness parameter Γ_{kt} lower than its cardinality, which is 12. This means that a few parameters (i.e. regions and varieties) have a greater positive contribution to the system. Nonetheless, all parameters have their contribution, which is demonstrated when we analyse the increase in the disturbance level, where more parameters should assume the worst-case value to reach the Soyster solution. An interesting result shown in Table 1 and Figure 4 is that the total quantity of tomatoes processed in the system varies near the maximum processed quantity indicated by the deterministic solution (243,837 ton), and the amount of cultivated area (2850 ha) does not change considering the variation of Γ_{kt}. The disturbance on the brix parameter causes a reduction in soluble solids content in the tomatoes and consequently in the system, which is confirmed in the results shown in Table 1 and Figure 4: Graph G. The model reallocates the cultivated tomato areas as compared to the deterministic solution, seeking to minimise the damage caused by the brix reduction, and does not rule out any available area.

Insofar as the brix parameter is disturbed unfavourably, the cost of transporting tomatoes from the agricultural regions to the processing plants increases significantly, as can be easily observed in Figure 4: Graph B, particularly for intermediate values of Γ_{kt}. The reason is that the model follows the logic of allocating the region-plant pair with the highest brix values (or the less damaged brix values, in this case) rather than the lowest transportation cost, despite the fact that the criterion is taken into consideration in building the solution. Reducing the total quantity of soluble solids in the system (Figure 4: Graph G) shows the need for additional procurements of semi-finished products to meet the demand for manufacturing final products (Figure 4: Graph C), as well as decreasing the formation of stocks of semi-finished products that provide economic benefits for the company in the last period of the planning horizon of the model (Figure 4: Graph F). The approximate probability of violating a given constraint (k, t) is shown in Figure 4: Graph I. From this graph, we observe that the highest probability of constraint violation, and model infeasibility, is in the deterministic solution ($\Gamma_{kt} = 0$). Figure 4: Graph J shows the optimal OF value in light of the probability that the constraint is feasible. Note that the OF value changes monotonically as Γ_{kt} increases (Figure 4: Graph A), as well as the probability of constraint feasibility (Figure 4: Graph J).

5.2 *Results of the RMY*

For each constraint r that determines the limit of area available to cultivate tomatoes (constraints (2)), the number of coefficients subject to uncertainty would be 36 due to the multiplication of the cardinalities of sets $|I| = 3$ and $|T| = 12$.

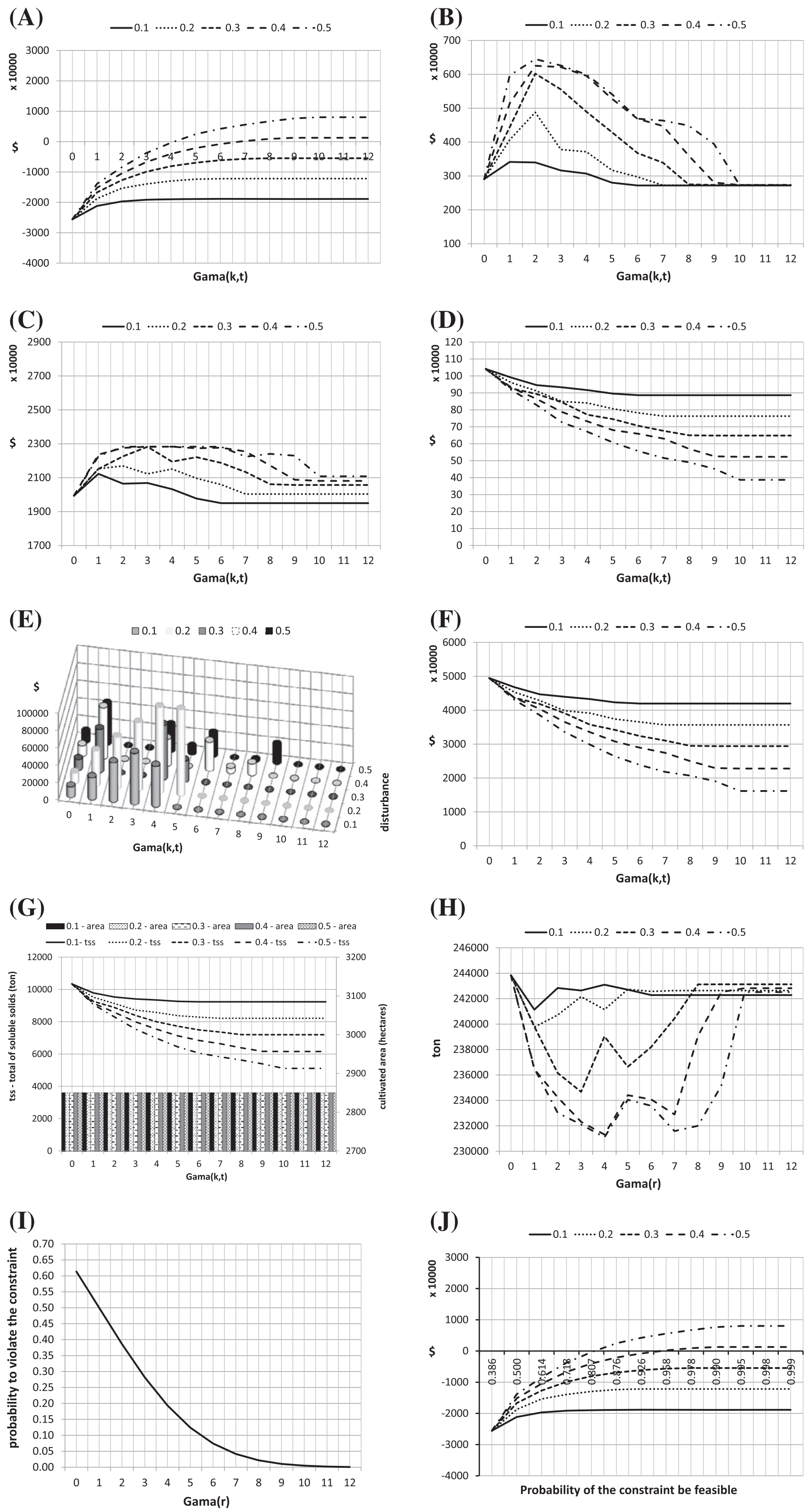

Figure 4. (A) Graph A: OF value – RMB; (B) Graph B: Transportation cost of tomatoes – RMB; (C) Graph C: Cost of procurements of semi-finished products – RMB; (D) Graph D: Cost of inventories of semi-finished products – RMB; (E) Graph E: Cost of transporting semi-finished products among plants – RMB; (F) Graph F: Economic benefit of stocking semi-finished products – RMB; (G) Graph G: Cultivated area and total amount of soluble solids in the plants – RMB; (H) Graph H: Quantity of tomatoes processed in the plants – RMB; (I) Graph I: Probability of violating the constraint with the uncertain brix parameter $(\tilde{\varepsilon}_{irt})$ – RMB; (J) Graph J: OF value and the probability of the constraint being feasible – RMB.

Table 1. The worst-case solution of RMB for each p over a budget of Γ_{kt}.

p	Γ_{kt}	OF value ($)	Total soluble solids in the system (ton)	Amount of processed tomatoes (ton)
0.1	6	−18,853,892	9238	242,282
0.2	7	−12,160,132	8217	242,637
0.3	9	−5,470,344	7198	243,134
0.4	10	1,282,282	6155	242,825
0.5	10	8,020,505	5113	242,518

However, the cardinality of set T is not 12 elements (12 months), because the parameter δ_{irt} assumes non-zero values only in four periods along the planning horizon, i.e. only during the four months of the tomato harvest. The parameter to adjust the robustness of the model is defined as Γ_r and it sets up the number of coefficients δ_{irt} in each constraint r that will take the worst value according to the disturbance level of p. Figure 5: Graphs: A, B, C, D, E, F, G, H, I and J shows the main economic and logistical results in the system, as well as the probability of observing the optimal OF value according to the budget of uncertainty considered by Γ_r in the model.

The disturbance on the crop yield causes a reduction in the quantity of tomatoes produced per area and consequently a worsening in the OF value (Figure 5: Graph A), mainly due to increases in semi-finished product procurements (Figure 5: Graph C) and lower stocks of semi-finished products that provide an economic benefit to the company (Figure 5: Graph F). The cost of transporting tomatoes from the fields to the plants is reduced because a lower quantity of tomatoes is transported (Figure 5: Graph B), unlike what was observed in the RMB solution (Figure 4: Graph B), in which there is a cost increase due to the reallocation of the region-plant pairs, particularly when the intermediate values of Γ_{kt} were observed. The inventory cost of semi-finished products decreases as p and Γ_r increase because fewer semi-finished products are produced and stored (Figure 5: Graph F).

The robust solution patterns of transporting semi-finished products among plants in the RMY were slightly affected over the disturbance of p and Γ_r (Figure 5: Graph E) compared to the deterministic $(\Gamma_r = 0)$ and RMB (Figure 4: Graph E) solutions. This observation can support the argument that the transportation planning of these products can be done disregarding the uncertainties in the crop yield, observing the total demand, the stocks and the production planning for each plant. However, this inference is only valid when the demands are fully aggregated for all plants (d_{pt}) and the model allocates them according to their production settings. If the demand is defined per plant (d_{ptk}), the proper model requires the transportation of final products among plants to prevent it from becoming infeasible, and it is likely to see different patterns for the transportation of semi-finished products as the disturbance and budget of uncertainties increase.

Due to the reduced crop yield, a lower quantity of tomatoes is produced in the available area (Figure 5: Graph H). Therefore, the equivalent area loss relative to the nominal crop yield was calculated and displayed in Figure 5: Graph G. In the managers' opinion, this is an important result to estimate the economic losses of cultivated areas that suffered crop loss. The worst-case solution in the RMY was reached with the number of uncertain parameters lower than 12, as can be observed in Table 2. It means that a few parameters δ_{irt} have the highest contribution to the system. The approximate probability of violating the constraint is shown in Figure 5: Graph I. Note that the slope of the curves for RMY and RMY are equal because both have 12 parameters coincidentally subject to uncertainty in the models. The optimal OF value of the RMY against the probability of the constraint being feasible is shown in Figure 5: Graph J.

5.3 *Results of the RMBY*

Similarly to the experiments for RMB and RMY, 60 solutions were obtained with RMBY to generate the set of scenarios displayed below. The first solution with $\Gamma_r = \Gamma_{kt} = 0$ corresponds to the deterministic scenario. For every set of scenarios, there was an increase in one unit in the value of the robustness tuning parameters (Γ_r, Γ_{kt}) such that in the last set it was for $\Gamma_r = \Gamma_{kt} = 12$. The number of all possible uncertain parameters in the RMBY is 24, of which 12 are controlled by Γ_r in the constraints where δ_{ikt} is present, and another 12 are controlled by Γ_{kt} in the constraints where ε_{ikt} appears. The Sosyter results for the RMBY are shown in Table 3 and, as expected, they were reached with a set of uncertain parameters lower than 24. Not surprisingly, the effects observed in the solution of the RMB and RMY models, insofar as the p and Γ values increase, were combined in the RMBY model, and its optimal solution deteriorates more intensively compared to RMB and RMY. The RMBY optimal OF value became positive, i.e. higher costs than profit,

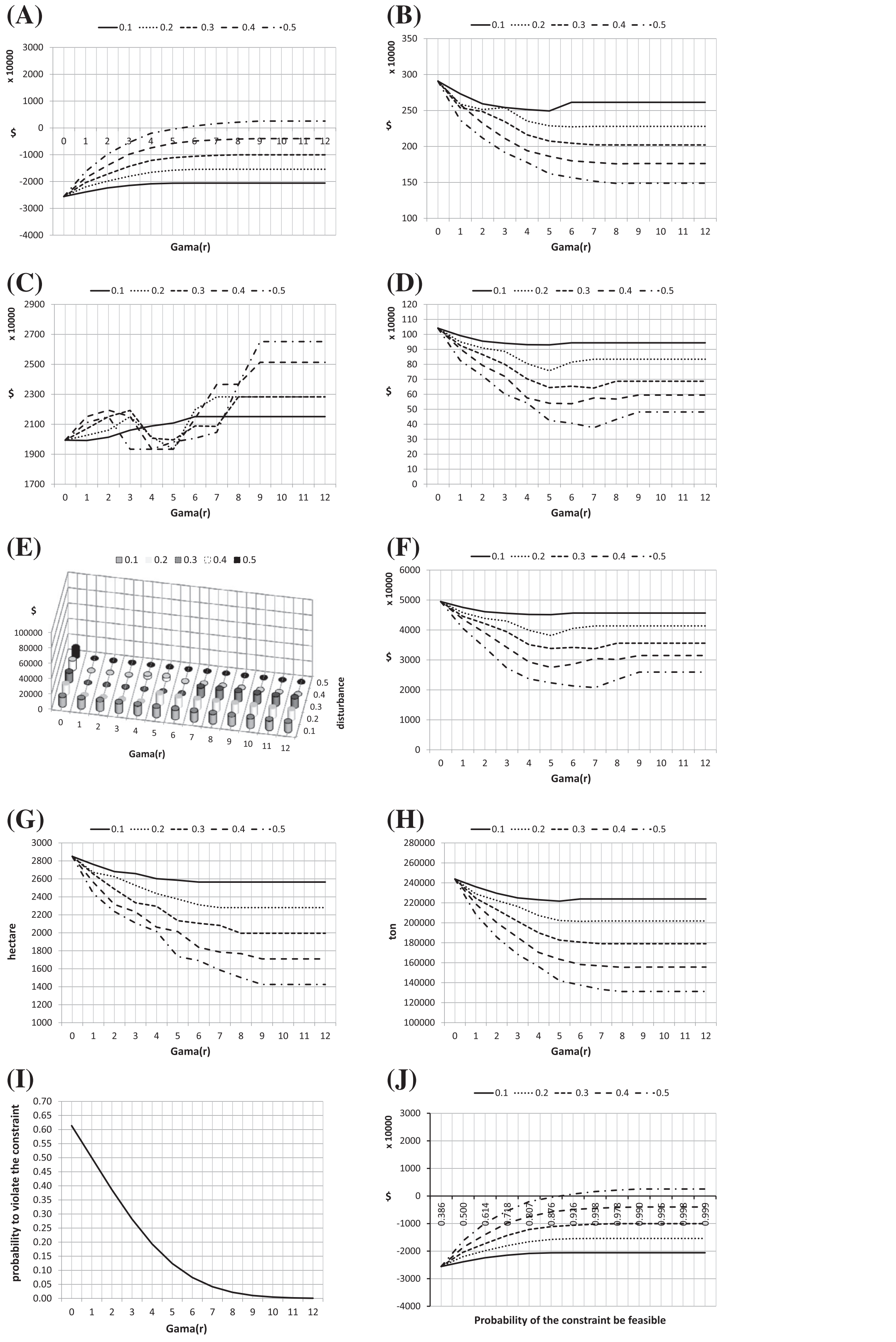

Figure 5. (A) Graph A: OF value – RMY; (B) Graph B: Cost of transporting tomatoes from the field to the plants – RMY; (C) Graph C: Cost of procurements of semi-finished products – RMY; (D) Graph D: Cost of inventories of semi-finished products – RMY; (E) Graph E: Cost of transporting semi-finished products among plants – RMY; (F) Graph F: Economic benefit of stocking semi-finished products – RMY; (G) Graph G: Equivalent area cultivated with tomatoes – RMY; (H) Graph H: Equivalent quantity of tomatoes processed in the plants – RMY; (I) Graph I: Probability of violating the constraint with the uncertain yield parameter $\left(\delta_{irt}\right)$ – RMY; (J) Graph J: OF value and the probability of the constraint being feasible – RMY.

Table 2. The worst-case solution of RMY for each p over budget of Γ_r.

p	Γ_r	OF value ($)	Equivalent cultivated tomato area (hectare)	Equivalent quantity of tomatoes processed (ton)
0.1	6	−20,570,674	2565	223,931
0.2	7	−15,407,290	2280	201,866
0.3	8	−10,033,332	1995	179,066
0.4	9	−3,985,714	1710	155,653
0.5	9	2,549,303	1425	131,220

Table 3. The worst-case solution of RMBY for each p over a budget of $\Gamma_r = \Gamma_{kt}$.

p	$\Gamma_r = \Gamma_{kt}$	OF value ($)	Equivalent cultivated tomato area (hectare)	Total soluble solids (ton)	Equivalent quantity of tomatoes processed (ton)
0.1	6	−14,375,145	2565	8584	223,232
0.2	8	−4,047,081	2280	6956	201,349
0.3	9	5,305,689	1995	5452	178,227
0.4	10	14,123,205	1710	4156	156,069
0.5	11	21,824,162	1425	2910	130,023

for pairs of (p, Γ) equal to $(0.3, 4), (0.4, 3)$ and $(0.5, 2)$. In the RMB, this happened when $p = 0.4, \Gamma_{kt} = 7$ and $p = 0.5, \Gamma_{kt} = 5$, and in RMY only for $p = 0.5, \Gamma_r = 6$.

The RMBY result for the cost of transporting tomatoes from the fields (Figure 6: Graph B) and the benefits of stocking semi-finished products (Figure 6: Graph F) is clearly a combination of the results observed in RMB and RMY. Regardless of the disturbance level, the RMBY solutions for Γ lower than 7 concerning the region-plant pairs and purchases of semi-finished products roughly resembles what was observed in the RMB solutions, whereas for Γ greater than 7, the RMBY solutions become more similar to those observed in RMY. The shipment pattern of semi-finished products among plants was distinct in the RMBY solution compared to RMB and RMY (Figure 6: Graph E). The economic benefit of making stocks of semi-finished products was greatly damaged when yield and brix were jointly disturbed (Figure 6: Graph F), and this led to lower total inventory costs (Figure 6: Graph D). The approximate probabilities of violating the constraints are shown in Figure 6: Graph I; for this calculation, complete independence and no correlation between parameters δ_{irt} and ε_{irt} in the model were assumed.

5.4 *Logic of the results*

Considering the experiments conducted on the robust models, some conclusions about the impact of unfavourable changes in the tomato brix and crop yield data in the agro system of the tomato processing industry can be drawn. Both parameters are at the core of agricultural and industrial performance, and unfavourable disturbances may cause significant damage in the company's business profitability.

We observed from the results that, in general, the disturbances to the tomato brix had a greater negative impact on the OF value of the models compared to the disturbances on the crop yield at the same budget of uncertainty imposed by Γ. Considering the number of parameters at the worst-case value less than or equal to four $(\Gamma \leq 4)$ in the constraints of the RMB and RMY, the brix parameter had a greater negative impact on the optimal OF value when disturbances p were small $(p \leq 0.3)$. This means that small unfavourable changes in the brix value are more relevant than small changes in the crop yield. For larger disturbances $(p \geq 0.4)$, the difference between the negative impact in the optimal OF value caused by the brix and by the yield at the same level of Γ until $\Gamma \leq 4$, is minor, and is almost the same impact for the disturbance for $p = 0.5$ with $\Gamma \leq 4$. In practice, this means that large unfavourable changes in the brix and yield values have almost similar impacts on the optimal OF value of the models.

When the disturbance is applied on the tomato brix, the yield of the manufacturing process in the industrial stages is affected and the logistical costs of the agricultural stage increase as a reflection of adjustments in some transportation decisions. Besides this fact, transported tomatoes have a lower soluble solids content, and the shipping cost is calculated on the quantity of tomatoes and not the amount of soluble solids. On the other hand, when a disturbance is on the crop

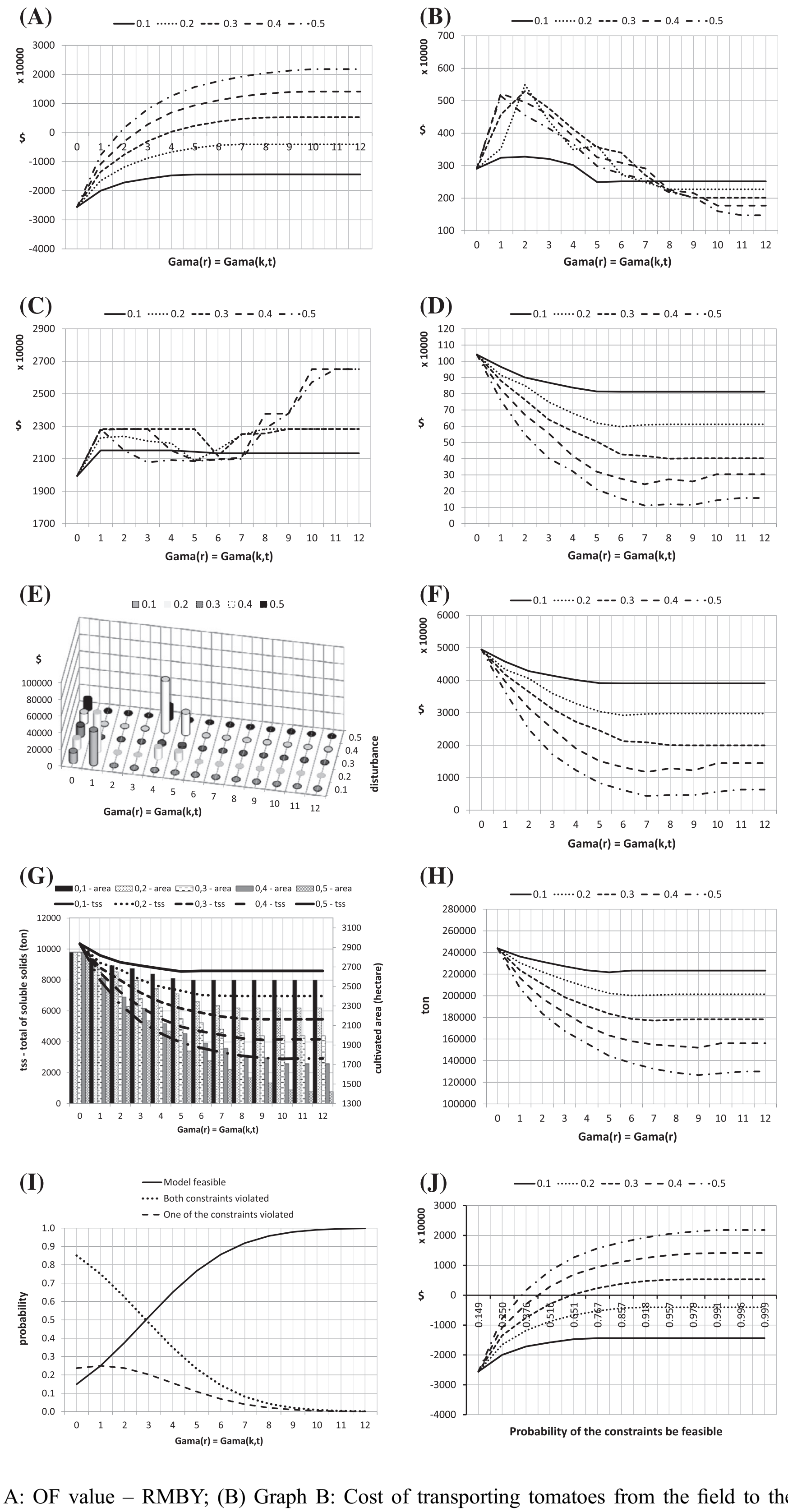

Figure 6. (A) Graph A: OF value – RMBY; (B) Graph B: Cost of transporting tomatoes from the field to the plants – RMBY; (C) Graph C: Cost of procurements of semi-finished products – RMBY; (D) Graph D: Cost of inventories of semi-finished products – RMBY; (E) Graph E: Cost of transporting semi-finished products among plants – RMBY; (F) Graph F: Economic benefit of stocking semi-finished products – RMBY; (G) Graph G: Cultivated area and total amount of soluble solids in the plants – RMBY; (H) Graph H: Equivalent quantity of tomatoes processed in the plants – RMBY; (I) Graph I: Joint probability of constraint violation and feasibility – RMBY; (J) Graph J: OF value and the probability of the constraints being feasible – RMBY.

yield, the manufacturing process is affected by a lower quantity of tomatoes to process. In this situation, the logistical cost of the agricultural stage is reduced because a lower quantity of tomatoes is produced in the fields and transported to the plants. Both situations cause a reduction in the tomato raw material to process (one is in the amount of tomato fruit and the other in the soluble solids content), but the disturbances over the tomato brix are shown to be more striking considering the logistical costs and the economic benefit brought by stocking semi-finished products compared to the disturbances over the crop yield.

The analysis of the robust outputs show that for small disturbances ($p \leq 0.2$), regardless of increasing Γ, the behaviour of robust solutions for the variables is quite close to the deterministic solution. As the disturbance increases, other solution patterns appear, particularly related to the allocation of region-plant pairs and procurements of semi-finished products. In the opinion of the managers who analysed the results, the Soyster solutions for the variables are not very conservative, despite the significant losses accounted for by the OF of the models. However, these experts also claimed that it would be unlikely to have unfavourable disturbances above 20% of the nominal values, even less so in all possible data of brix and crop yield, as performed in the analysis.

Considering the opposite of the unfavourable disturbance, i.e. a favourable disturbance in which the brix and crop yield are increased, it is not surprising that the brix parameter has a greater positive impact on the OF value compared to the crop yield. Increasing the brix value means increasing the amount of soluble solids content in the same quantity of tomatoes. The system ends up transporting the same amount of tomato mass determined by the crop yield, but with a higher soluble solids content. More soluble solids in the system can produce more quantity of semi-finished products, partially reducing the procurements from the market and increasing the stocks that bring economic benefit to the company. When the crop yield is favourably disturbed, the cost of transporting tomatoes from the agricultural fields to the plants increases because more tomatoes are produced. Still in the context of this research, a sensitivity analysis of disturbances in the tomato brix and crop yield parameters can be found in Rocco (2014).

Experiments using larger data-sets, for example, by doubling the planning horizon (24 periods) and including several processing plans (at least 4), regions (at least 10), tomato varieties (at least 5), semi-finished (at least 10) and final products (at least 10), etc., did not cause problems in building the solution or in the computational processing times. The results were quickly obtained by CPLEX (less than 5 s) and were consistent to the system. However, the model does not perform two harvests in the same available area, once there are no equations to represent this and the period index for year.

6. Concluding remarks

This paper presented a deterministic linear optimisation model and three robust optimisation models applied to the analysis of production/logistics costs, crop and industrial planning in the tomato processing industry. The deterministic model itself is able to represent the key agricultural and industrial decisions to plan the annual tomato crop, despite its simplifying assumptions, particularly regarding the aggregated capacities of concentrators and demand for final products, as discussed in Rocco (2014). The robust optimisation approach applied in this study allowed us to investigate the effect of data uncertainty on unfavourable cases in two relevant parameters in this agro-industrial system, viz., the tomato brix and crop yield at the moment of harvesting, which are crucial for determining the economic outcome of the business. An assumption that should be mentioned is the equal probability of data uncertainty assumed by the robust approach, which may not be true in agricultural data.

We believe that a key contribution of this paper is its practical application to solve real production / logistics tactical planning problems, taking uncertainty into account in the most relevant data to the tomato processing industry. These models can be embedded in a friendly interface for final users, allowing them to generate a number of risk scenarios to support decisions. According to the industry experts consulted during the study, they are not aware of a similar analytical approach for planning agricultural and industrial activities, as the optimisation models presented here allow. They agree that companies should invest in tools to support and optimise decisions, especially in the accuracy of monitoring and forecasting the tomato brix and crop yield in order to reduce uncertainty in the data used in the models.

By means of several experiments using real data, the effects of five levels of unfavourable disturbances on tomato brix and crop yield parameters could be analysed, while varying the budget of uncertainty considered for building the solutions. As the entire approach is based on linear programming, the computing times are only a few seconds (less than 3 s) to solve the models using a standard personal computer, which is very convenient for practical use. From the results, it can be observed that for the same budget of uncertainty Γ, the uncertainties in the tomato brix have a greater economic impact on the system as compared to uncertainties in the crop yield. This conclusion is quite important for managers in the sector, who may give preference to high brix levels by means of agricultural management practices, for instance, through irrigation, rather than obtaining high crop yields with low soluble solids content in the tomato fruit.

Further relevant insights from the robust scenarios are: creating a policy for stock coverage; transportation and purchase plans for semi-finished products to compensate for the data uncertainty in the brix and crop yield parameters.

Future studies considering the purpose of addressing management issues are on the authors' agenda. One is to incorporate the uncertainty in the resource area available to cultivate tomatoes and in the demands of final products for consumers. These parameters could also bring relevant insights concerning the integrated system capacity, particularly when the disturbances are meant for favourable cases. Energy consumption is a very relevant topic in the tomato processing industry due to the high consumption of steam in the concentration of tomato juice. Several authors have addressed this subject (Dale, Okos, and Nelson (1982), Rumsey et al. (1984), Rocco and Morabito (2014)) and we think that the integrated approach of optimisation modelling, as presented here, may help managers to search for ways to save energy and build a sustainable business. Finally, despite the fact that this study was performed in a particular tomato processing industry, we believe that the analytical approach used here could be applied or easily adapted for use by other tomato companies and other agro-industrial systems, such as the citrus and guava sectors.

Acknowledgements

The authors would like to thank the two anonymous reviewers for their useful comments and suggestions, and the Brazilian company for its collaboration with this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was partially supported by CAPES; CNPq; FAPESP, Brazil.

ORCID

Cleber D. Rocco http://orcid.org/0000-0002-7988-6136
Reinaldo Morabito http://orcid.org/0000-0002-3948-305X

References

Alem, D. J., and R. Morabito. 2012. "Production Planning in Furniture Settings via Robust Optimization." *Computers & Operations Research* 39: 139–150. doi:10.1016/j.cor.2011.02.022.

Alvarez, P. P., and J. R. Vera. 2011. "Application of Robust Optimization to the Sawmill Planning Problem." *Annals of Operations Research* 219: 457–475. doi:10.1007/s10479-011-1002-4.

Ben-Tal, A., and A. Nemirovski. 1998. "Robust Convex Optimization." *Mathematics of Operations Research* 23: 769–805. doi:10.1287/moor.23.4.769.

Ben-Tal, A., and A. Nemirovski. 1999. "Robust Solutions of Uncertain Linear Programs." *Operations Research Letters* 25: 1–13. doi:10.1016/S0167-6377(99)00016-4.

Ben-Tal, A., and A. Nemirovski. 2000. "Robust Solutions of Linear Programming Problems Contaminated with Uncertain Data." *Mathematical Programming* 88: 411–424. doi:10.1007/PL00011380.

Bertsimas, D., and M. Sim. 2003. "Robust Discrete Optimization and Network Flows." *Mathematical Programming* 98: 49–71. doi:10.1007/s10107-003-0396-4.

Bertsimas, D., and M. Sim. 2004. "The Price of Robustness." *Operations Research* 52: 35–53. doi:10.1287/opre.1030.0065.

Bertsimas, D., and A. Thiele. 2006. "A Robust Optimization Approach to Inventory Theory." *Operations Research* 54: 150–168. doi:10.1287/opre.1050.0238.

Beyer, H.-G., and B. Sendhoff. 2007. "Robust Optimization – A Comprehensive Survey." *Computer Methods in Applied Mechanics and Engineering* 196: 3190–3218. doi:10.1016/j.cma.2007.03.003.

Birge, J. R., and F. Louveaux. 1997. *Introduction to Stochastic Programming*. Book Springer Series in Operations Research and Financial Engineering.. New York: Springer. ISBN: 978-1-4614-0236-7 (Print), 978-1-4614-0237-4 (Online). doi:10.1007/978-1-4614-0237-4.

Bohle, C., S. Maturana, and J. Vera. 2010. "A Robust Optimization Approach to Wine Grape Harvesting Scheduling." *European Journal of Operational Research* 200: 245–252.

Caixeta-Filho, J. V. 2006. "Orange Harvesting Scheduling Management: A Case Study." *Journal of the Operational Research Society* 57: 637–642. doi:10.1057/palgrave.jors.2602041.

da Silva, A. F., F. A. Silva Marins, and J. A. Barra Montevechi. 2013. "Multi-choice Mixed Integer Goal Programming Optimization for Real Problems in a Sugar and Ethanol Milling Company." *Applied Mathematical Modelling* 37: 6146–6162. doi:10.1016/j.apm.2012.12.022.

Dale, M. C., M. R. Okos, and P. Nelson. 1982. "Concentration of Tomato Products: Analysis of Energy Savings Process Alternatives." *Journal of Food Science* 47: 1853–1858. doi:10.1111/j.1365-2621.1982.tb12898.x.

El Ghaoui, L., and H. Lebret. 1997. "Robust Solutions to Least-squares Problems with Uncertain Data." *SIAM Journal on Matrix Analysis and Applications* 18: 1035–1064. doi:10.1137/S0895479896298130.

El Ghaoui, L., F. Oustry, and H. Lebret. 1998. "Robust Solutions to Uncertain Semidefinite Programs." *SIAM Journal on Optimization* 9: 33–52. doi:10.1137/S1052623496305717.

Gabrel, V., C. Murat, and A. Thiele. 2014. "Recent Advances in Robust Optimization: An Overview." *European Journal of Operational Research* 235: 471–483. doi:10.1016/j.ejor.2013.09.036.

Gameiro, A. H., J. V. Caixeta Filho, C. D. Rocco, and R. Rangel. 2008. "Modelagem e gestão das perdas no suprimento de tomates para processamento industrial [Modeling and Management of Losses in the Supply of Tomatoes for Manufacturing Process]." *Gestão & Produção* 15: 101–115. doi:10.1590/S0104-530X2008000100010.

IBGE (The Brazilian Institute of Geography and Statistics). 2015. *Banco de dados agregados SIDRA* [Aggregated Database System].

Junqueira, R. A. R., and R. Morabito. 2012. "Production and Logistics Planning Considering Circulation Taxes in a Multi-plant Seed Corn Company." *Computers and Electronics in Agriculture* 84: 100–110. doi:10.1016/j.compag.2012.02.019.

Kawamura, M. S., D. P. Ronconi, and H. Yoshizaki. 2006. "Optimizing Transportation and Storage of Final Products in the Sugar and Ethanol Industry: A Case Study." *International Transactions in Operational Research* 13: 425–439. doi:10.1111/j.1475-3995.2006.00556.x.

Leiras, A., S. Hamacher, and A. Elkamel. 2010. "Petroleum Refinery Operational Planning Using Robust Optimization." *Engineering Optimization* 42: 1119–1131. doi:10.1080/03052151003686724.

Mulvey, J. M., R. J. Vanderbei, and S. A. Zenios. 1995. Robust Optimization of Large-scale Systems. *Operations Research* 43: 264–281. doi:10.1287/opre.43.2.264.

Munhoz, J. R., and R. Morabito. 2014. "Optimization Approaches to Support Decision Making in the Production Planning of a Citrus Company: A Brazilian Case Study." *Computers and Electronics in Agriculture* 107: 45–57. doi:10.1016/j.compag.2014.05.016.

Najafi, M., K. Eshghi, and W. Dullaert. 2013. "A Multi-objective Robust Optimization Model for Logistics Planning in the Earthquake Response Phase." *Transportation Research Part E: Logistics and Transportation Review* 49: 217–249. doi:10.1016/j.tre.2012.09.001.

Paiva, R. P. O., and R. Morabito. 2009. "An Optimization Model for the Aggregate Production Planning of a Brazilian Sugar and Ethanol Milling Company." *Annals of Operations Research* 169: 117–130. doi:10.1007/s10479-008-0428-9.

Rocco, C. D. 2014. "Modelagem do planejamento agrícola e industrial no setor de tomate para processamento utilizando programação matemática e otimização robusta [Modeling of Agricultural and Industrial Planning in the Tomato Processing Industry by Using Mathematical Programming and Robust Optimization]." Phd diss., Universidade Federal de São Carlos, São Carlos, Brazil.

Rocco, C. D., and R. Morabito. 2014. "Scheduling of Production and Logistics Operations of Steam Production Systems in Food Industries: A Case Study of the Tomato Processing Industry." *Journal of the Operational Research Society* 65: 1896–1904. doi:10.1057/jors.2013.84.

Rocco, C. D., and R. Morabito. 2015. "Production and Logistics Planning in the Tomato Processing Industry: A Conceptual Scheme and Mathematical Model." Technical Report. Departament of Production Engineering, Federal University of Sao Carlos, Sao Carlos, Brazil.

Rosentrater, K. A., and E. Kongar. 2009. "Modeling the Effects of Pelleting on the Logistics of Distillers Grains Shipping." *Bioresource Technology* 100: 6550–6558. doi:10.1016/j.biortech.2009.07.051.

Rumsey, T. R., T. T. Conant, T. Fortis, E. P. Scott, L. D. Pedersen, and W. W. Rose. 1984. "Energy Use in Tomato Paste Evaporation." *Journal of Food Process Engineering* 7: 111–121.

Santos, M. O., and B. Almada-Lobo. 2012. "Integrated Pulp and Paper Mill Planning and Scheduling." *Computers & Industrial Engineering* 63: 1–12. doi:10.1016/j.cie.2012.01.008.

Soyster, A. L. 1973. "Technical Note – Convex Programming with Set-inclusive Constraints and Applications to Inexact Linear Programming." *Operations Research* 21: 1154–1157.

WPTC. 2015. Releases of the World Production Estimatives. http://www.wptc.to.

Appendix 1. Procedure to build the Robust Model Brix (RMB)

Constraints (4) in the original form	$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} \geq \sum_{l} \theta_l P^{\text{raw}}_{klt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0$
Parameter subject to uncertainties $(\tilde{\varepsilon}_{irt})$ written based on the random variable Z^B_{irt}	$\tilde{\varepsilon}_{irt} = \varepsilon_{irt} + \hat{\varepsilon}_{irt} Z^B_{irt}$ $-1 \leq Z^B_{irt} \leq 1 \quad \forall\, i,\, r,\, t$
Constraints with the protection function maximising the deviation of the parameter for the worst-case value	$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} - \max \left\{ \sum_{\substack{(i,r) \in J_t \\ \delta_{irt} \neq 0}} \hat{\varepsilon}_{irt} Q^{\text{tom}}_{irkt} Z^B_{irt} \right\} \geq \sum_{l} \theta_l P^{\text{raw}}_{lkt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0$
Rewritten constraints with the protection function according to the level of Γ_{kt}	$- \sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} + \beta_{kt}\left(Q^{tom^*}_{irkt}, \Gamma_{kt}\right) \leq - \sum_{l} \theta_l P^{\text{raw}}_{lkt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0$
Primal sub-problem of the protection function β_{kt}	$\beta_{kt}\left(Q^{tom^*}_{irkt}, \Gamma_{kt}\right) = \max \left\{ \sum_{\substack{(i,r) \in J_t \\ \delta_{irt} \neq 0}} \hat{\varepsilon}_{irt} Q^{\text{tom}}_{irkt} Z^B_{irt} \right\}$ subject to: $\sum_{(i,r) \in J_t} Z^B_{irt} \leq \Gamma_{kt}$ $0 \leq Z^B_{irt} \leq 1 \quad \forall\, (i,r) \in J_t$
Dual model of the primal sub-problem of the protection function for each constraint (k,t)	$\text{Min} : \Gamma_{kt} \lambda_{kt} + \sum_{(i,r) \in J_t} \mu_{irkt}$ subject to: $\lambda_{kt} + \mu_{irkt} \geq \hat{\varepsilon}_{irt} Q^{\text{tom}}_{irkt} \quad \forall\, (i,r) \in J_t,\, \delta_{irt} \neq 0$ $\lambda_{kt} \geq 0$ $\mu_{irkt} \geq 0 \quad \forall\, (i,r) \in J_t$
Inserting the OF of the dual sub-problem into the original constraints of the deterministic model, and also adding the new constraints from the dual sub-problem	$\sum_{\substack{ir \\ \delta_{irt} \neq 0}} \varepsilon_{irt} Q^{\text{tom}}_{irkt} - \Gamma_{kt} \lambda_{kt} - \sum_{(i,r) \in J_t} \mu_{irkt} \geq \sum_{l} \theta_l P^{\text{raw}}_{lkt} \quad \forall\, k,\, t,\, \sum_{ir} \delta_{irt} \neq 0$ $\lambda_{kt} + \mu_{irkt} \geq \hat{\varepsilon}_{irt} Q^{\text{tom}}_{irkt} \quad \forall\, k,\, t,\, (i,r) \in J_t,\, \delta_{irt} \neq 0$ $\lambda_{kt} \geq 0 \quad \forall\, k,\, t$ $\mu_{irkt} \geq 0 \quad \forall\, k,\, t,\, (i,r) \in J_t$

Appendix 2. Procedure to build the Robust Model Yield (RMY)

Constraints (2) in the original form.	$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} Q_{irkt}^{\text{tom}} / \delta_{irt} \leq \gamma_r \quad \forall\, r$
Modified constraint to eliminating the parameter of the denominator.	$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} \Delta_{irt} Q_{irkt}^{\text{tom}} \leq \gamma_r \quad \forall\, r$
Uncertain written parameter based on the random variable Z_{irt}^{Y}.	$\tilde{\Delta}_{irt} = \Delta_{irt} + \hat{\Delta}_{irt} Z_{irt}^{Y}$ $-1 \leq Z_{irt}^{Y} \leq 1 \quad \forall\, i, r, t$
Constraints with the protection function maximising the parameter deviation for the worst-case value.	$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} \Delta_{irt} Q_{irkt}^{\text{tom}} + \max \left\{ \sum_{\substack{k,(i,t) \in J_r \\ \delta_{irt} \neq 0}} \hat{\Delta}_{irt} Q_{irkt}^{\text{tom}} Z_{irt}^{Y} \right\} \leq \gamma_r \quad \forall\, r$
Constraints rewritten with the protection function according to the level of Γ_r.	$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} \Delta_{irt} Q_{irkt}^{\text{tom}} + \beta_r \left(Q_{irkt}^{tom^*}, \Gamma_r \right) \leq \gamma_r \quad \forall\, r$
Primal sub-problem of the protection function β.	$\beta_r \left(Q_{irkt}^{tom^*}, \Gamma_r \right) = \max \left\{ \sum_{k,(i,t) \in J_r} \hat{\Delta}_{irt} Q_{irkt}^{\text{tom}} Z_{irt}^{Y} \right\}$ subject to: $\sum_{(i,t) \in J_r} Z_{irt}^{Y} \leq \Gamma_r$ $0 \leq Z_{irt}^{Y} \leq 1 \quad \forall\, (i,t) \in J_r$
Dual model of the primal sub-problem of the protection function.	Min : $\Gamma_r \lambda_r + \sum_{(i,t) \in J_r} \mu_{irt}$ subject to: $\lambda_r + \mu_{irt} \geq \hat{\Delta}_{irt} \sum_k Q_{irkt}^{\text{tom}} \quad \forall\, (i,t) \in J_r,\ \delta_{irt} \neq 0$ $\lambda_r \geq 0$ $\mu_{irt} \geq 0 \quad \forall\, (i,t) \in J_r$
Inserting the OF from the dual model in the constraints of the deterministic model and adding new constraints from the dual model into the deterministic model.	$\sum_{\substack{ikt \\ \delta_{irt} \neq 0}} \Delta_{irt} Q_{irkt}^{\text{tom}} + \Gamma_r \lambda_r + \sum_{(i,t) \in J_r} \mu_{irt} \leq \gamma_r \quad \forall\, r$ $\lambda_r + \mu_{irt} \geq \hat{\Delta}_{irt} \sum_k Q_{irkt}^{\text{tom}} \quad \forall\, r, (i,t) \in J_r,\ \delta_{irt} \neq 0$ $\lambda_r \geq 0 \quad \forall\, r$ $\mu_{irt} \geq 0 \quad \forall\, r, (i,t) \in J_r$