POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Executive Summary of the Thesis

Optimal management of a biogas production plant: optimization models with uncertainty constraints

Laurea Magistrale in Automation and control Engineering - Ingegneria dell'automazione e del controllo

Author: Elvis Sheyler Jimenez Portillo

Advisor: Prof. Pietro Luigi Belotti

Co-advisor:

Academic year: 2021-2022

---

1. Introduction

In the face of climate change and the search for new forms of renewable energy, biogas presents itself as a new opportunity and a solution to both problems. The objective of this work is the study and implementation of a model of anaerobic digestion plant for the production of biogas, with the aim of optimizing the costs related to the operation of the plant. Robust optimization models have been implemented to deal with the possible presence of uncertain parameters present in the plant. Multiple uncertainty models have been used with the aim of obtaining solutions that can deal with uncertainty in the least conservative way.

2. The structure of the plant

A biogas plant transforms biomass into biogas through a series of anaerobic transformations. Generally the structure of a biogas plant is quite standardized, and its main components are: storage areas, digester, gasometers, biogas treatment systems, cogeneration and digestate storage areas. The process that leads to the formation of biogas is called anaerobic digestion, which is, in extreme synthesis, a complex biological process in which, in the absence of oxygen, the organic substance is transformed into biogas. From anaerobic digestion, in addition to biogas, another product is formed, called digestate, which is made up of the non-degraded organic fraction rich above all in nitrogen. Digestate is often used as a natural fertilizer. In this work a 1 MW biogas plant is considered. The plant is small in size and works on a local scale. In this case, the biomasses treated for the production of biogas are corn harvesting waste and pig manure. The latter have been added as biomass entering the plant for various reasons. The first is that they bring useful substances to the biogas production process, while the second is that they are used to bring the percentage of dry matter below a certain threshold (we will see in more detail what this second point means later). The biomasses are bought and transported from the respective farms/pig farms, to then be prepared for the process inside the anaerobic digesters. Inside the digesters, through the process of anaerobic digestion, biogas and digestate, which will be sold as fertilizer, will be formed.

3. The model

First of all, an “ideal” biogas plant model was implemented, without taking into consideration any uncertain parameters. In particular, an optimization model was built, because as previously described, the objective of this work is to obtain optimal management of the plant, and therefore to maximize its profit. An optimization model is a system composed of specific characteristics which are the objectives, the variables, the parameters and the constraints. The goal is usually defined by an objective function. The variables are controlled by us and depend on the relationships defined within the model, while the parameters, which describe the data of the problem, are predetermined and cannot be modified. In our model, the variables considered are the quantity of corn and slurry purchased, essential for the operation of the plant, the quantity of biogas and fertilizer sold, and the number of trips necessary for the transport of biomass and fertilizer. Speaking of constraints instead, these define the necessary conditions that the model must respect in order to find feasible solutions. Usually the constraints are represented by equations or inequalities. In our model there are several constraints that can be divided according to the functions or characteristics they represent. For example, there are production constraints, which limit the maximum amount of corn/slurry produced for a farm/pig farms; there are constraints that relate biomass and biogas/fertilizer, or the relationship constraints for transport, which define the link between the quantities of biomass purchased and the number of trips necessary for their transport. As a demonstration, we represent in mathematical form the constraint that defines the limit of dry substance present in the digester. In fact it should be emphasized that biomass cannot enter the digester as it is, but must be processed and treated beforehand in special storage areas to prepare them for the anaerobic digestion process. It is important that the percentage of dry matter, which is the part of the material sample that remains after dewatering, does not exceed a certain threshold. If the dry part were to exceed this threshold, the anaerobic digestion process would not guarantee the correct production of biogas and digestate. The total dry matter threshold admissible in the digester in this case cannot be more than 10%, while for corn and slurry it is respectively thirty and five per cent (data found in the literature). The constraint that describes it is the following:

$$0.3 \sum_{j=1}^{J} X_j + 0.05 \sum_{k=1}^{K} Y_k \leq 0.10 \left(\sum_{j=1}^{J} X_j + \sum_{k=1}^{K} Y_k\right)$$

where J and K represent respectively the number of farms and pig farms considered, while X and Y are the variables representing the quantity of corn and slurry purchased.
The objective function, as previously described, has the purpose of optimizing the profit of the biogas plant, therefore to maximize the difference between revenues and total costs. All the profits we get from the sale of biogas and digestate that are produced by the plant are considered revenues. The costs, on the other hand, refer to the purchase of biomass and the transport costs (of biomass and digestate).

3.1. The model uncertainty

After implementing the model in the absence of uncertainty, we move on to consider the uncertainty that may be present in the plant. This serves to make the model more realistic due to the random nature of the yields, but requires the use of another type of optimization, called robust, which will be described later. Now the quantities of corn produced by the farms (S_j^M) are no longer considered constant but subject to uncertainty. In fact, it is very probable to think that during the year, due to various factors, the production of corn is not constant, but subject to uncertainty. Climatic and environmental factors can occur throughout the year and significantly influence corn production. Obviously the model will be subject to modifications, the most important of which is the introduction of a new variable, called r, which represents the total quantity of corn imported into the plant, which is now uncertain, and depends on the X_j, a variable that is redefined to represent the percentage of production capacity of a farm j. The X_j variable was changed in meaning because if it remained as the amount of corn bought, as it was in the initial model, it would become an uncertain variable beacuse the amount of corn produced is uncertain in this model. Therefore it has become necessary to introduce a new variable, namely r, which by defining the to-

tal amount of corn imported into the plant, is not directly dependent on S_j^M. In this way it will be possible to build a robust model that is not too conservative, as we would have only one constraint in the system in which the uncertain parameter will be present, and which also determines the value of r (this constraint will be shown later). The importance of the variable r arises from the fact that it relates the uncertain parameter S_j^M to our optimization model.
The implemented constraints are mostly those defined in the model without uncertainty. Of particular interest, however, is the constraint that defines the relationship between the total quantity of corn produced on each farm and the total quantity of corn imported into the plant, described below:

$$\sum_{j=1}^{J} S_j^M \; X_j \geq r, \tag{1}$$

This constraint is fundamental because it tells us that r is limited from above by the total quantity of corn produced, which is the sum of the products between the uncertain capacities S_j^M and the proportions X_j. It could be said that this constraint defines the level of uncertainty of the model, because it contains all the variables and parameters affected by uncertainty. It is precisely from this constraint that the implementation of the robust optimization model will start.
As regards the objective function, however, it is always the same, with the aim of maximizing the profit of the plant.
Let us start talking about the optimization model that we will use to deal with uncertainty.

4. Robust optimization

Robust optimization is a paradigm for modeling optimization problems under uncertainty. In a problem of this type, in addition to known parameters and variables, there are also uncertain parameters, not known, of which we have no control. These parameters can vary within a well-defined set of uncertainty. This set depends on the type of uncertainty being treated and can be defined by both linear and non-linear inequalities and equations. For any uncertain parameter value, the constraints defining the uncertainty set must be satisfied. It is also important to underline the fact that we must try to define a set/range of uncertainty that is not too large but narrow enough, to prevent a too conservative model.
The goal of robust optimization is to find a solution that is feasible for all possible values of the uncertain parameters within the uncertainty set, regardless of the realization of those parameters. To explain it in the simplest way, one can think that the uncertain parameters are decided and controlled by an "opponent" [1]. After we solve our optimization problem, the opponent sees our solution and chooses the values of the uncertain parameters that are most likely to harm us. In practice the opponent solves an optimization problem in which the uncertain parameters are its variables, and with the aim of finding a solution that most damages ours (for example by violating our constraints). Consequently we will have to try to anticipate the moves of our opponent. So we need to build a "robust counterpart" to the optimization problem, in other words we have to incorporate our opponent's optimization problem into the model. In practice we will have to solve another optimization problem, called the "robust counterpart", which will have our uncertain parameter as variables and will have the objective of finding the solution which is more unfavorable in our case. Here is an example of what was said:

$$\max \; 2x + 3y$$

$$1 \leq x \leq 7$$

$$2 \leq y \leq 9$$

$$x + y \geq 6$$

$$\max(ux + y - v : u \geq 0, v \geq 0, u + v \leq 1) \leq 5$$

where x and y are the variables, while u and v are the uncertain parameters. The last constraint, the one in which the uncertain parameters are present, represents the robust counterpart, i.e., the optimization problem solved by our opponent, who in this case, through the maximization operator, tries to violate the constraint. However, by solving the optimization model in this way, it may happen that a non-linear constraint is obtained in the model which significantly increases the complexity of the problem, as happens in the example above. A solution to this eventuality can be found by

applying the concept of duality.
The dual of a problem is a fundamental concept in linear programming. In fact, each linear programming problem can be associated with another problem, called dual, which provides important information relating to the solution. By resorting to duality we could transform the nonlinear problem into a linear one, without resorting to overly complex computational efforts. In this case one could consider the model implemented by the opponent as the primal problem, on which the dual will then be built by applying the primal/dual correspondence rules that can be found in any linear programming manual. Finally, by replacing the dual problem with the nonlinear constraint in which the uncertain parameters are present, the robust optimization model will be obtained.
In the following section we will show how the robust optimization model was implemented using the approach just described.

5. Robust optimization model

5.1. Polyhedral uncertainty

In this case the uncertainty set in which the uncertain parameter S_J^M is defined is the following:

$I = \{S_j^M, j \in J : S_j^M \in [SmLB_j, SmUB_j],$
$\sum_{j=1}^{J} S_j^M \geq \sigma_{LB},$
$\sum_{j=1}^{J} S_j^M \leq \sigma_{UB},$
$S_{j_1}^M \geq 0.9\ S_{j_2}^M,\ S_{j_2}^M \geq 0.9\ S_{j_1}^M\ \forall (j_1, j_2) \in E\}$

The uncertainty set establishes that the corn production for each farm is within a certain range, defined by $SmLB_j$ and $SmUB_j$, which are the lower and upper limits of corn production for each farm, respectively. σ_{LB} and σ_{UB} are two parameters representing the lower and upper limit of total corn production. Finally, there are also proximity relations between neighboring farms. In other words in a pair of farms (j_1,j_2) included in the set E, which defines the set of pairs of neighboring farms, the farm j_1 cannot produce more than a certain amount of farm j_2 and vice versa. The set E will be formed by the pairs of farms that will be no more than 3 km apart. The equations/inequalities representing this interval define a polyhedron of possible values of S_j^M.
Now to build the robust model we apply the following reasoning. The uncertain parameter S_j^M is present only in the constraint represented by equation (1). In order to violate this constraint, the opponent could decide to build an optimization model where S_j^M is the controlling variable, and the objective function aims to minimize the left side of the constraint (1), in this way:

$$\min(\sum_{j=1}^{J} S_j^M\ X_j) \geq\ r \qquad (2)$$

Obviously the constraints of the opponent's optimization model will be those defined in the uncertainty set. The robust counterpart will be given by substituting the model just described in constraint (1). In this way, however, the linear model becomes non-linear with all the problems that derive from it, including for example a greater difficulty in its resolution. Remembering what has been said previously, using duality, we will be able to implement a linear optimization model and overcome this problem. Therefore from the primal problem, which would be the model implemented by the opponent, the dual problem is constructed through the rules of primal/dual correspondence. Finally, by substituting this problem for constraint (1) of the original model, the robust optimization model will be obtained.

5.2. Budget uncertainty

Now we will test another class of uncertainty sets to implement the robust model, which takes the name of "Budget uncertainty" or "Gamma uncertainty". It was Bertsimas and Sim who proposed this new approach in 2004 [3], with the aim of reducing the level of conservatism of the robust models implemented up to that point. In short, budget uncertainty is a constant that controls how many uncertain parameters can deviate from their nominal values. For example let us consider a set of parameters a_j belonging to a set J, and some of these are subject to uncertainty. Uncertain parameters deviate from their nominal value in a certain way. Statistically it is improbable that all the parameters $a_j, j \in J$ are subject to uncertainty, therefore a parameter θ can be introduced which represents the number of parameters which can deviate from the nominal value. θ

regulates the number of uncertain parameters that can be considered, and consequently θ controls the "trade-off" between the level of uncertainty and its effects on the objective function of the nominal problem. This trade-off is referred to as the "price of robustness". In practice, by adjusting θ, the robustness of the model can be controlled with respect to the level of conservatism of the solution.
Returning to our problem, let us construct a new uncertainty set that defines the uncertain parameter S_j^M, recalling the budget uncertainty approach defined previously. The uncertainty set is as follows:

$I = \{S_j^M, j \in J : S_j^M \in [SmLB_j, SmUB_j],$
$S_j^M \geq SmNOM_j\ (1 - \lambda b_j)\ \forall j \in J,$
$\sum_{j=1}^{J} b_j \leq \theta,$
$b_j \geq 0\ \forall j \in J,$
$b_j \leq 1\ \forall j \in J\}.$

The uncertainty set defines that the corn production for each farm is within the range defined by $SmLB_j$ and $SmUB_j$. Furthermore, a new variable called b_j is introduced, which determines the number of uncertain parameters present. The b_j variable is defined in the range of zero to one. In fact, if b_j is different from zero it means that the production of S_j^M corn is different from the nominal value. The parameter θ regulates the "price of robustness", i.e., it controls the trade-off between the level of uncertainty of the model and its effects on the objective function. If θ is equal to zero the sum of the variables b_j must be equal to zero and therefore there are no farms with uncertain corn production. Instead increasing the values of θ increases the upper limit on the sum of the values of b_j, and therefore increases the level of uncertainty. New parameters have been introduced, $SmNOM_j$ and λ, which respectively define the nominal production of corn for each farm and the decrease in the nominal production value of corn, when this production is uncertain. To build the robust optimization model we have to define the robust counterpart of our optimization problem. The reasoning is always the same: we must think that we have an opponent who wants to harm us in some way. The adversary controls the uncertain parameters, so he will try to damage us by violating the only constraint in which the uncertain parameter S_j^M is present, which is the inequality (1), trying to minimize the quantity to the left of the sign. As previously done, the optimization problem performed by our opponent is defined as a primal problem, with the difference that the constraints imposed in this case are those defined in the new uncertainty set. Then, again with the same procedure, the dual problem was constructed, and finally by substituting it for equation (1) of the original model, the new robust optimization model was obtained.

5.3. Scenario-based uncertainty

Finally we deal with a scenario-based type of uncertainty. Scenarios are an efficient way of constructing robust optimization problems using historical data of uncertainties [2]. Suppose we have a vector of uncertain parameters, but we do not know an uncertainty set that fits our model. However, if we had a database containing the value of each uncertain parameter for a given period of time, we could build a model for which one or more constraints are satisfied for each register of the uncertainty vector. So even if we do not have a well-defined set of uncertainty, we could build a robust model considering the set of scenarios as a sort of range of uncertainty. In this case, the historical values of corn production in a certain time range will be considered (the data do not refer to a real plant but will be obtained randomly using Python libraries). In practice now the uncertain parameter S_j^M is not a row vector, but a matrix of values. In fact, S_j^M must represent the corn production values for all farms considered over the years. Consequently now the equation (1) is modified in this way:

$$\sum_{j=1}^{J} S_{n,j}^M\ X_j \geq r,\ \ \forall n \in 1..N$$

with N defining the number of past years taken into account. It should be noted that by increasing the number of scenarios, and therefore of past years taken into consideration, the level of uncertainty of the problem increases and therefore a more conservative model is obtained.

5.4. Solution

The following table shows the values of biogas produced, in cubic meters, for the different optimization models implemented. We show these values because biogas is the main source of income, and therefore the profit of our plant directly depends on these results. The amount of biogas produced, without uncertainty, is obtained by solving the non-robust model for the given $SmNOM_j$. In this case the value corresponds to 6220k cubic meter of biogas. The following table shows the values obtained for the implemented robust optimization models.

Δ	Poliedr.	θ	Budget	N	Scenarios
0.99	6155k	1	6088k	5	6161k
0.95	5907k	2	5955k	10	5953k
0.90	5597k	3	5822k	20	5892k
0.75	5443k	5	5573k	40	5852k
0.5	5443k	10	5443k	80	5757k

Table 1: Quantity of biogas produced in cubic meters for the implemented models

The values in the table were obtained always starting from a given $SmNOM_j$ vector and considering a single set of farms/pig farms with their coordinates. Δ is a parameter that regulates the values of σ_{LB} and σ_{UB} in polyhedral uncertainty. The lower this value is, the more the interval defined by σ_{LB} and σ_{UB} widens and consequently the level of uncertainty in the model increases. In the same way, increasing the value of θ and N for the budget uncertainty and for the scenario uncertainty, we increase the level of uncertainty in the respective models. Obviously, an increase in uncertainty corresponds to more conservative models and consequently to more stringent solutions in terms of values, as can also be seen from the table.

6. Conclusion

From the table we can draw some conclusions about the utility of the type of uncertainty sets chosen. The table shows us that by increasing the level of uncertainty in each model we see a decrease in the amount of biogas produced. In fact, with the increase in uncertainty, the models become more conservative and consequently the solutions report lower values. However, this decrease in values is more or less rapid depending on the type of uncertainty considered.

In general, if the robust counterpart is implemented for two different sets of uncertainty, different values are obtained. However if one uncertainty set is the subset of another, then its robust counterpart will obtain larger values of the objective function regardless of the type of uncertainty set classes (in our case polyhedral, budget or scenario-based uncertainty). Ultimately it can be said that depending on the cases that may occur in reality, a set of uncertainty may be more or less useful. For example, it may be happen that not all plants have historical data available, and therefore in that case the scenario uncertainty model could not be used. On the other hand, if one would like to have tighter control over the trade-off between the level of uncertainty of the model and the value of the objective function, one could use budget uncertainty, and so on.

References

[1] Pietro Belotti, Zsolt Csizmadia, Susanne Heipcke, and Sebastien Lannez. Robust Optimization with FICO™Xpress. 2014.

[2] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*, volume 28. Princeton university press, 2009.

[3] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.