

Two-Stage Stochastic Optimization of a hydrogen network

Gloria Gutierrez*, Anibal Galan*, Daniel Sarabia**, Cesar de Prada*

* *Department of Systems Engineering and Automatic Control, University of Valladolid
c/ Real de Burgos s/n. Sede Mergelina EII, 47011, Valladolid, Spain*
(e-mails: gloria@autom.uva.es, anibalsantiago.galan@uva.es, prada@autom.uva.es)

***Department of Electromecanic Engineering, Escuela Politécnica Superior, Universidad de Burgos,
Avda. Cantabria s/n, Burgos, 09006, (dsarabia@ubu.es)*

Abstract: This paper discusses how to deal explicitly with uncertainty in the optimal management of the hydrogen network of a petroleum refinery. The current system is based on a RTO/MPC system for supervision and on-line optimization that includes a robust data reconciliation to estimate consistent values of the process variables and update the model parameters. It has been extended with a two-stage stochastic optimization to take care of the effect of crude changes in operation of the network. The paper analyses how to formulate the problem in order to obtain implementable solutions and presents results that compare the deterministic and stochastic solutions using real plant data.

Keywords: Two-step stochastic optimization; Real-time optimization; Model predictive control; Oil refineries; Hydrogen networks.

1. INTRODUCTION

Hydrogen has become one the main utilities in oil refineries due to the combined needs of removing sulphur from petrol products and converting heavy hydrocarbons into lighter ones as a result of the new environmental legislation and the aim of increasing profitability of the refinery business.

Hydrogen is used in different reactors in order to accomplish these tasks, being supplied through a complex distribution network to the different plants that perform the hydro-desulphurization and hydro-treating operations. Due to technical reasons, hydrogen has to be used in excess in the reactors, with the unreacted hydrogen being recycled or sent to the fuel-gas network. As hydrogen is an expensive utility, a good management of the network implies minimizing the production of fresh hydrogen or conversely of losses to fuel-gas. Nevertheless, reducing the available hydrogen can limit the processing of the more valuable hydrocarbons, so that an optimum balance has to be reached between maximizing the flow of hydrocarbons and minimizing the required hydrogen, while satisfying the operation constraints, using the hydrogen generation and distribution in the network, membranes, and other process elements as degrees of freedom.

In the oil refinery of Petronor, in Northern Spain, a system driven by these aims is in operation, Sarabia et al. (2012). It is composed by a RTO working in supervisory mode and a LP/MPC controller that implements the optimal policies on-line. Additionally, a data reconciliation module provides consistent estimates of the process variables and updates the model parameters for the RTO. The system works well, nevertheless, as the refinery processes crudes of different origins and properties whose detailed compositions are rarely known beforehand, periodic disturbances take place during

feed changes. In the end, these unknowns affect negatively the operation and profitability of the plants and hydrogen network. In order to limit these consequences, the uncertainty in the properties of the new hydrocarbons, such as hydrogen demand or molecular weight of the light ends, should be incorporated explicitly in the optimization formulation.

The data reconciliation module above mentioned is able to deal with parametric uncertainty, but only provides indication of the changes once they have revealed in the process. Ideally, one should make decisions before one knows the value and possible effects of the uncertainty, and not only reacting to the consequences of a choice, which leads to the use of stochastic optimization methods instead of deterministic ones. Additionally, other methods such as Modifier Adaptation, are more focused on the process-model mismatch instead of the decision making in unknown and changing scenarios and are limited in the size of problems they can sensibly deal with. Considering these aspects, the two-stage stochastic optimization methods seem to be a good approach to the problem of incorporating the uncertainty of the hydrocarbon properties into the RTO system, as they offer the required flexibility in the description of uncertainty as different scenarios, and reflect the current practice of decision making and a-posteriori correction.

This paper discusses the formulation of the optimal management of a hydrogen network in an oil refinery as a two-stage stochastic optimization problem and presents results obtained with real data. The main contributions are linked to the proposal of an architecture and problem specification that allows using this advanced tool in a large scale system as the hydrogen network, which involves the joint operation of eighteen process plants.

The paper is organized as follows: After the introduction, section 2 describes the hydrogen network and its operation. In section 3, the current RTO and MPC control system is presented, followed by the core section 4 which deals with the two-stage stochastic optimization and the way it is adapted and formulated for optimization of the network operation. Then, section 5 gives results of the stochastic optimization and compares them with the deterministic case. The paper ends with some conclusions and references.

2. PROCESS DESCRIPTION

2.1 Hydrogen network

In the refinery of reference, high purity hydrogen is produced in steam-reforming furnaces in two plants, named H3 and H4. Additionally, two platformer plants (P1 and P2) generate lower purity hydrogen as a by-product of the catalytic reforming process so that their flows can be considered as non-controllable disturbances to the network. From these four plants, hydrogen is distributed to the consumer ones using several interconnected networks at different purities and pressures, as can be seen in the schematic of Fig.1. The network interconnects a total of eighteen plants, four producers and fourteen consumers, mainly hydrodesulphurization (HDS) plants.

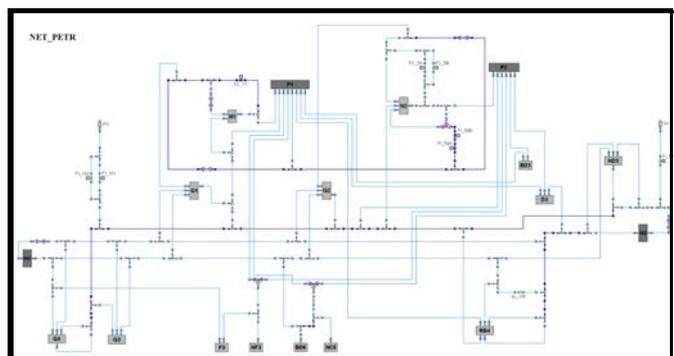


Fig 1. Schematic of the hydrogen network of the Petronor refinery. Dark grey boxes represent producer plants, while light grey ones refer to hydrogen consumer units

A typical HDS receives hydrogen from different sources, and after mixing it with the hydrocarbon load, the mixture is processed in bed reactors where the hydrogen must be in excess to prevent shortening of the life-cycle of the expensive catalyst. The excess hydrogen is partly recycled internally (in some cases using membranes to increase its purity), partly purged to the fuel-gas network or recycled to a low purity header (LPH) to be used in other plants. The global operation of the network can be explained using Fig. 2, which is a simplified representation where only a small number of producer and consumer plants are represented. The generated hydrogen is distributed to the consumer plants through the corresponding high purity headers. The hydrogen demand of each plant depends on the quantity and composition of the hydrocarbons being treated, which may experience strong changes every two or three days according to the crude that is

being processed. Excess hydrogen from these plants is partially collected in the low purity header (LPH) and recycled back to the consumer plants, while the rest goes to the fuel gas network, where it is mainly burnt in furnaces.

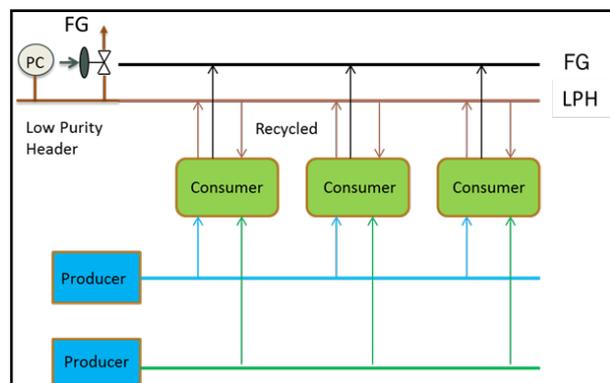


Fig 2. Schematic of producer and consumer plants with the main hydrogen distribution headers and fuel gas network.

2.1 Network operation

Both, plants and networks, are operated from control rooms equipped with Distributed Control Systems (DCS) implementing basic controls (flow, pressure, ...) and several MPCs (DMC) in charge of more complex multivariable tasks, such as sulphur removal in the plants.

The main network operation aims are:

- Distribute the available fresh hydrogen and the recycled hydrogen (including internal plant recycles) so that the requirements of hydrogen at the reactors' inputs in all plants are satisfied.
- Maximize the hydrocarbon loads to the plants, approaching the production targets established by the refinery planning system.
- Balance the hydrogen that is produced and the hydrogen that is consumed so that the hydrogen losses to fuel gas are minimized.

The main decision variables are the fresh hydrogen production of H3 and H4 plants, the hydrocarbon feed to the consumer plants and the hydrogen distribution and reuse in the network, including the use of membranes where available. The overall operation is framed by the specific production targets given by the planning system of the refinery that change according to the market conditions and crudes available, and it is constrained by the physical and operational limitations imposed by the equipment.

3. RTO AND MPC

3.1 Data reconciliation

The operation of the system is difficult not only due to its complexity and the presence of significant disturbances that affect the process, but because the information available

about many key variables is limited and unreliable. In particular, molecular weights of the impurities are unknown, which stops the computation of sensible mass balances. To avoid this problem, a data reconciliation (DR) system was developed with the aim of estimating consistent values of all plant variables from available on-line measurements based on a process model.

A first principles model of the network and associated plants was developed to provide support in process optimization Gomez (2016). It is based on mass balances of hydrogen and light ends (considered as a single pseudo-component) in the pipes and units. In addition, it incorporates other equations for compressors, membranes, separation units (including a solubility model), etc., some of which are reduced order models fitted to experimental data or with some adjustable parameters. Taking into account the much faster dynamics of the hydrogen flows compared to the dynamics of the reactors, the hydrogen distribution model is static, having flows, purities, molecular weights of hydrogen and light ends of all streams and hydrogen consumption in the reactors as its main variables.

Data reconciliation requires redundancy in measurements, and takes advantage of the fact that the core of the model, being based on mass balances, does not present structural errors. The DR problem is solved as a large NLP one in GAMS using IPOPT and incorporating robust estimators as the Fair function to compensate gross errors, Nicholson et al. (2014). The implementation involves more than 4400 variables and 4700 equality and inequality constraints. It provides consistent estimations of the measured and unmeasured variables while, at the same time, enables the update of certain unknown model parameters.

3.2 Deterministic Real Time Optimization

Once reliable information of the network and a model are available, it is possible to formulate the following optimization problem, that translates the aims described in section 2.1, and is executed at regular time intervals:

$$\min \sum_{i=1}^2 p_{H_2i} F_{H_2i} - \sum_{i=1}^{14} p_{HCi} HC_i + \sum_{i=1}^{14} p_{Ri} R_i \quad (1)$$

s.t.

Process model

Process constraints

Refinery planning specifications

where HC_i refers to the hydrocarbon loads to consumer plants, F_{H_2i} denotes the fresh hydrogen generated in the steam reforming plants and R_i are recycles of hydrogen in the consumer plants, which are linked to the operation of the recycle compressors. Here, p_{HC} , p_{H_2} and p_R stand for prices associated with hydrocarbons, fresh hydrogen and compressors in order to provide an economic meaning to the cost function. The problem has to be solved under the constraints imposed by the model and operation of the units, taking also into account the specifications coming from the refinery planning. Constraints apply mainly to pipes'

capacity, recycle purity in the consumer plants, ratio hydrogen/hydrocarbon at the reactors' input, operating range of membranes, producer plants' capacity, reciprocating and centrifugal compressors' capacity, etc. Important parameters, such as the specific hydrogen consumption or equilibrium constants are fixed in the model according to the DR estimation. Again, the problem is a NLP one and is solved in the GAMS environment, involving nearly 2000 variables and more than 1800 equality and inequality constraints, with the IPOPT algorithm in less than one minute. CPU time, running every two hours, and its results are available in an Excel HMI and through the refinery Osisoft PI system.

3.3 On-line implementation with DMC

One of the main problems related to the implementation of the RTO solutions is the fact that, being a static optimization executed at low frequency, it is not able to cope with disturbances and changes that must be taken into account and corrected at a higher frequency, Darby et al. (2011). At the same time, when operating the RTO, is possible to identify a set of patterns in the optimal solutions that can be implemented as partial targets and that define the global network optimization: For instance, maintaining the purge from the LPH at minimum, or keeping the purity of the recycles at a certain low value. These partial targets have been implemented in the LP layer of a commercial DMC controller, providing set points to the MPC controller below similar to the ones of the RTO, but computed on-line at the frequency of the controller, De Prada et al. (2017). The global architecture is summarized in Fig.3. The LP/DMC only acts on the six more important plants of the network. The RTO operates in supervisory mode, with its solutions being computed for the whole network and providing a reference framework for the on-line operation of the DMC.

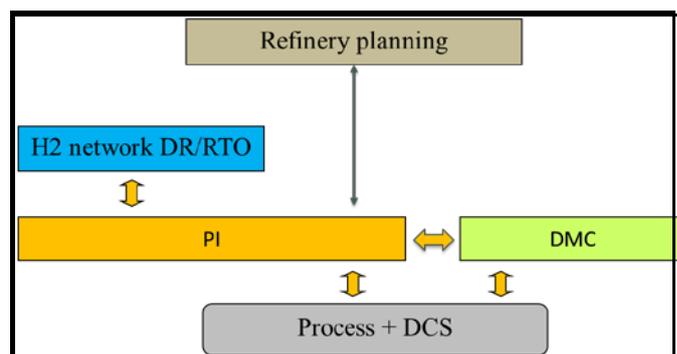


Fig. 3. Block diagram of the main elements involved in the hydrogen network management

4. TWO-STAGE STOCHASTIC REAL TIME OPTIMIZATION

4.1 Load changes

As mentioned above, an oil refinery normally receives petroleum supplies from different sources every two or three days that are processed continuously to generate a wide range

of refinery products. The quality and composition of these supplies may vary quite a lot, depending of the country of origin and the type of oil. This means that, after being processing in the crude distillation tower, the different streams and products that have to be treated in the downstream hydro-treatment processes may present significant changes of the hydrogen demand over time. The operation of the plants involved in the hydrogen network, alternates in this way periods of relatively stable hydrogen consumption with transients where the estimation or prediction of the specific hydrogen demands in the reactors is difficult to perform. Of course, there are many others sources of uncertainty in the operation of the network, including the state of the equipment, the effect of disturbances, changes in the molecular weight of the hydrocarbons and light ends generated in the reactors, etc. but this is the one that has a mayor impact in production and appears with a frequency low enough to be treated in the optimization layer, while other changes of higher frequency require more frequent corrections in the range of the control actions.

As a consequence of the oil supply changes, the optimization of the hydrogen production and distribution in the network and the computation of the maximum admissible hydrocarbon load to the hydro-treating plants is more difficult to perform in the transient periods and would benefit of integrating the associated incertitude explicitly in the formulation of the optimization problem.

4.1 Two-Stage Stochastic Optimization

A common way of considering uncertainty in the decision making process is incorporating the concept of recourse variables: If one have to make a decision now and some parameters of the problem are unknown, one should take into account the different values that these parameters may get (different scenarios), but also should consider when making the decision the possibility of correcting the initial one (changing the so called recourse variables) when, later on in the future, the uncertain parameters may become known.

In this approach, the time horizon is divided in two stages: in the first one we decide on the value of the “here and now” variables, while in the second stage, we decide on the values of the recourse variables, which have different values for each of the scenarios considered and depend also on the first stage decisions. The idea is represented in Fig. 4 (left), where we can see that the first stage variables are unique for all scenarios, while the second stage ones are particular of every scenarios. Fig.4 (right), displays the evolution of the process states according to the realization of the uncertain. In the first stage, once the first stage decision variables are applied to the process, the state may evolve to different values depending on the value of the uncertain parameters, while in the second stage the evolution will depend on the specific value of the recourse variables for each scenario.

Mathematically, a typical two-stage stochastic optimization problem is formulated as the minimization of a cost function under constraints involving a set of stochastic variables ξ as in (2):

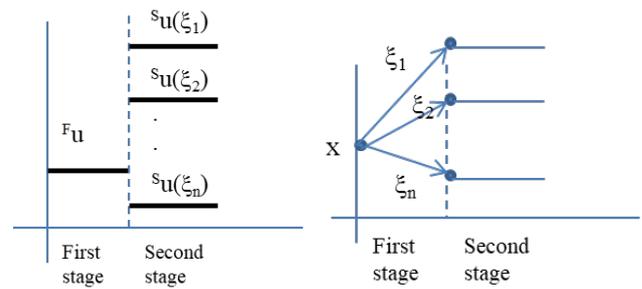


Fig 4. (Left) First stage variables are unique for al scenarios, while the recourse (second stage) variables are particular of every scenario considered for the uncertainty. (Right) The state of the process, when applied the first stage decision variables, may evolve to different values according to the realization of the uncertain variables ξ_i , but the subsequent evolution is associated to a specific action $S_u(\xi_i)$ for every scenario.

$$\begin{aligned} \min_{F_u, S_u(\xi)} \quad & {}^F J(F_u) + E\{{}^S J(F_u, S_u(\xi), x(\xi))\} \quad \forall \xi \in \Xi \\ & {}^F h({}^F x(\xi), {}^F u) = 0, \quad {}^F g({}^F x(\xi), {}^F u) \leq 0 \\ & {}^F u(\xi) = {}^F u \\ & {}^S h({}^F u, S_u(\xi), S_x(\xi)) = 0, \quad {}^S g({}^F u, S_u(\xi), S_x(\xi)) \leq 0 \end{aligned} \quad (2)$$

The notation requires some explanation: F(.) refers to variables or functions in the first stage and S(.) denotes the ones in the second stage. The decision variables are denoted as u and the remaining ones as x. The uncertainty is represented by the parameters ξ that can take values within a set Ξ according to a certain probability distribution. Normally this set is sampled and only a finite number ξ_i , $i = 1, 2, 3, \dots, n$ of elements is considered which constitute the scenarios that will represent the uncertainty. $E\{.\}$ stands for the expected value.

The cost function is composed of two terms: The first one, ${}^F J$, is the cost in the first stage which depends on the first stage decisions ${}^F u$. These are decisions made and applied at current time, without knowing the particular realization of the uncertainty ξ , that will be maintained over the time horizon covered by the optimization problem. Consequently, they are the same for all values of ξ , what is represented by the constraint ${}^F u = {}^F u(\xi)$ known as non-anticipativity one. Nevertheless, we can correct the effects of the ${}^F u$ decisions in the second stage once the value of the ξ parameters materializes, using the recourse variables $S_u(\xi)$ that take a particular value for each realization of the uncertainty. The second term of the cost function, $E\{{}^S J(.)\}$, represents the effect of these second stage corrections on the total value of the cost function, which also depends on the ${}^F u$ decisions and the uncertainty ξ .

The variables of the problem have to satisfy the constraints imposed by the model h and additional inequality constraints g in every stage for all possible scenarios considered. In (2), the corresponding equations, that depend on the stochastic

parameter ξ , should be interpreted as being fulfilled with probability one.

Notice that the two-stage stochastic optimization approach to dealing with decision problems in uncertain environments provides more degrees of freedom, represented by the use of the recourse variables, than classical robust ones. Robust optimization formulates the problems as min max ones, providing decisions that fit the worst case of all scenarios without considering the second stage corrections, leading consequently to more conservative solutions than the two-stage approach. At the same time, two-stage is not as computationally demanding as a considering the full stochastic problem.

4.3 Stochastic optimization formulation

The task of formulating an optimization problem that considers explicitly the uncertainty in the refinery hydrogen network is not easy and has to balance different aspects, taking as starting point the configuration of the existing RTO system. Critical elements of the formulation are the selection of the uncertain variables and scenarios, the choice of the first and second stage decision variables, the coherence with the global operation and the feasibility of computing the solutions in a short time so that they can be useful for the on-line operation of the network.

The hydrogen network involves eighteen plants so that if the uncertain variables are not chosen carefully, the number of scenarios, generated from combinations of the values of the uncertainties in all of them, can blow up easily. Fortunately, if we assume that the main source of incertitude is the change of quality of the oil supplies as discussed above, and considering as its main effect the variations of specific hydrogen consumption in the reactors, we can notice that the changes in quality affects in parallel to all plants. This means that if the refinery receives crude with e.g. a higher sulphur content, the hydrogen demand will increase in all hydrodesulphurization plants, avoiding the need of covering all possible combinations of increments and decrements in every plant, that can be substituted by a small number of scenarios all of them in the same direction of increment or decrement of the specific hydrogen consumption specified as a set of per cent changes over the current estimation of the specific hydrogen consumption given by the data reconciliation module according to the analysis performed in the crude being processed.

The reformulation of the RTO as a two-step stochastic optimization problem has to consider that RTO is basically a static optimization one where the aim is to compute targets for the different variables involved, while the problem of dealing with the supply changes has a certain dynamic character linked to the load transients. One important aspect of the problem is the fact that the hydrogen producer plants have slow dynamics, needing around two hours to reach a new production target. Because of that, once a hydrogen production aim is given to a producer plant, if there are sudden changes in the hydrogen demand, in order to avoid wasting hydrogen to the fuel-gas network or facing hydrogen

defaults in the reactors, changes in that production aim are not effective in the short term. Instead, one has to act on other faster variables such as hydrocarbon loads to the plants, fuel-gas purges or recirculation purity. Taking this into account, we propose to use the hydrogen generated in the producer plants as first stage variables and the other ones as recourse variables.

The two-stage stochastic optimization can then be formulated as (3):

$$\begin{aligned} \min_{F_{H2i}, s_u(\xi)} \sum_{i=1}^2 p_{H2i} F_{H2i} - E \left\{ \sum_{i=1}^{14} p_{HCi} HC_i(\xi_j) - p_{Ri} R_i(\xi_j) \right\} \\ \forall \xi_j \in \Xi \end{aligned} \quad (3)$$

s.t.

Process model(ξ_j)
 Process constraints(ξ_j)
 Refinery planning specifications

Here, the process model and constraints are the same as in the deterministic case, but particularized for every scenario, which largely increases the number of variables and equations. Notice that the first stage decision variables F_{H2i} are the same for all scenarios, according to the non-anticipativity constraints. The first stage cost corresponds to the production cost of fresh hydrogen, while the second stage includes the expected value of the hydrocarbons processed and the cost of the hydrogen recycles. The aim is maximizing the hydrocarbon load (HC) to consumer plants, minimizing the use of fresh hydrogen generated in the steam reforming plants (F_H) and minimizing the internal recycles of hydrogen (R) in the consumer plants, considering all possible values of the uncertainty. s_u refers to the remaining variables of the model.

Within a stochastic environment, particularly relevant in periods of petroleum supply changes, the formulation tries to find the best choice of the fresh hydrogen production targets such that, when the actual hydrogen demands in the reactors are revealed, it is possible to recourse to the correction action of other variables (load changes, membranes, purges, etc.) such that the operation constraints are satisfied for all scenarios considered and the expected value of the cost function in (3) is optimized.

5. RESULTS

2.1 Implementation

The problem (2) has been implemented in the GAMS environment and solved using the EMP feature. Due to the large amount of variables involved, it is not possible to present all of them in a paper. In addition, due to confidentiality reasons, we cannot offer information of the actual value of many key variables. Considering these constraints, we will present values of the main variables in percent, taking as 100% the corresponding value obtained in the data reconciliation step before optimization. The following tables compare the solution obtained with the

deterministic approach (1) and the stochastic one for three scenarios S1, S2, S3, where the specific hydrogen demand in the reactors has been assumed to be the same, a 5% higher and a 10% higher than the one computed in the DR step, with probabilities 0.6, 0.3 and 0.1 respectively. The first stage solution is also given in the tables. The first column displays the acronym of the plants involved. Two plants, D3 and RB4, were not in operation at the time when the data were collected.

Table 1. Scenario probabilities and hydrogen demands.
*No change.

	S1	S2	S3
Probability	0.6	0.3	0.1
Hydrogen demand	NC*	+5%	+10%

Table 2. H₂ feed to the consumer units in %

Plants	Det.	1 st stage	S1	S2	S3
BD3	98.65	118.00	120.34	185.98	120.11
BD6	34.49	711.72	149.63	50.17	144.16
F3	94.53	121.38	109.23	126.89	110.17
G1	101.68	95.55	98.50	97.61	98.25
G2	122.42	108.80	113.98	110.45	112.10
G3	91.37	99.63	102.07	101.98	102.68
G4	100.	108.59	107.55	107.42	115.42
HD3	113.45	109.65	109.72	110.06	108.17
NC6	99.99	102.01	102.7	95.76	95.73

Table 3. HC loads to the consumer units in %

Plants	Det.	1 st stage	S1	S2	S3
BD3	100.	115.96	120.0	120.0	80.0
BD6	99.99	115.99	119.99	119.99	79.99
F3	100.0	116.0	120.0	120.0	80.0
G1	100.0	100.0	100.0	100.0	100.0
G2	100.0	80	80.0	80.0	80.0
G3	100.0	100	100.0	100.0	100.0
G4	99.99	100	99.99	99.99	99.99
HD3	100.21	100.18	100.211	100.21	99.949
N1	100.	116.0	120.0	120.0	80.0
N2	100.0	116.0	120.0	120.0	80.0
NC6	99.99	104.50	107.2	107.23	79.99
NF3	99.99	115.99	119.99	119.99	79.99

Table 4. Fresh Hydrogen production in %

Plants	Det.	1 st stage	S1	S2	S3
H3	70.02	100.9	100.9	100.9	100.9
H4	104.55	106.31	106.31	106.31	106.31

As can be seen, significant corrections can be made in the hydrogen to the plants via recirculation from the low purity header, while changes in the hydrocarbon loads are also affected in some scenarios.

6. CONCLUSIONS

The paper shows the applicability of two-stage stochastic approach in this type of problems, but further work is required before implementation, in particular in solving more efficiently the associated optimization problem, using decomposition methods as in Martí et al. (2015).

AKNOWLEDGMENT

The authors wish to express their gratitude for the financial support received from the Spanish Government with project INOPTCON (MINECO/FEDER DPI2015-70975-P) and the EU ITN PRONTO, Grant agreement No 675215. The authors are also grateful to Petronor and its management for supporting this study.

REFERENCES

- Darby, M.L.; Nikolaou, M.; Jones, J.; Nicholson, (2011) D. RTO: An overview and assessment of current practice. *J. Process Control*, 21, 874–884.
- De Prada C., Sarabia D., Gutierrez G., Gomez E., Marmol S., Sola M., Pascual C., Gonzalez R. (2017) Integration of RTO and MPC in the hydrogen network of a petrol refinery, *Processes*, 5(1), 3; doi:10.3390/pr5010003. 7
- Gomez, E. (2016) A Study on Modelling, Data Reconciliation and Optimal Operation of Hydrogen Networks in Oil Refineries. *Ph.D. Thesis, University of Valladolid, Spain*,
- Martí R., Lucia S., Sarabia D., Paulen R., Engell S., de Prada C. (2015) Improving Scenario Decomposition Algorithms for Robust Nonlinear Model Predictive Control, *Computers and Chemical Engineering*, 79 30–45 Elsevier
- Nicholson, B.; Lopez-Negrete, R.; Biegler, L.T. (2014) On-line state estimation of nonlinear dynamic systems with gross errors. *Comput. Chem. Eng.*, 70, 149–159.
- Sarabia, D.; de Prada, C.; Gómez, E.; Gutiérrez, G.; Cristea, S.; Mendez, C.A.; Sola, J.M.; González, R. (2012) Data reconciliation and optimal management of hydrogen networks in a petro refinery. *Control Eng. Pract.* 20, 343–354.