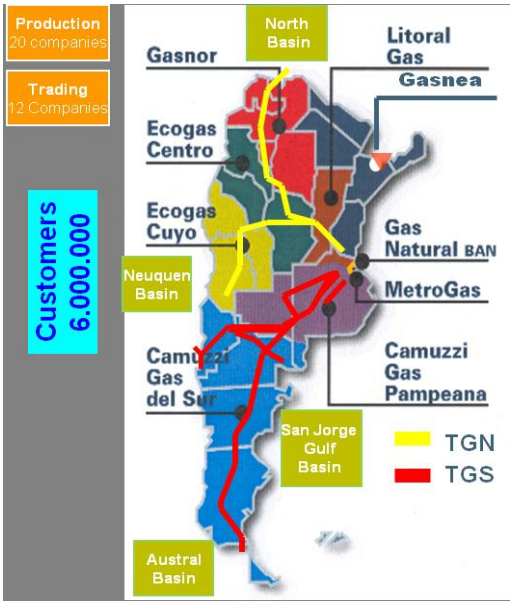
Unaccounted for gas diminishing applying Artificial Neural Networks, a tool to fight against frauds & thefts

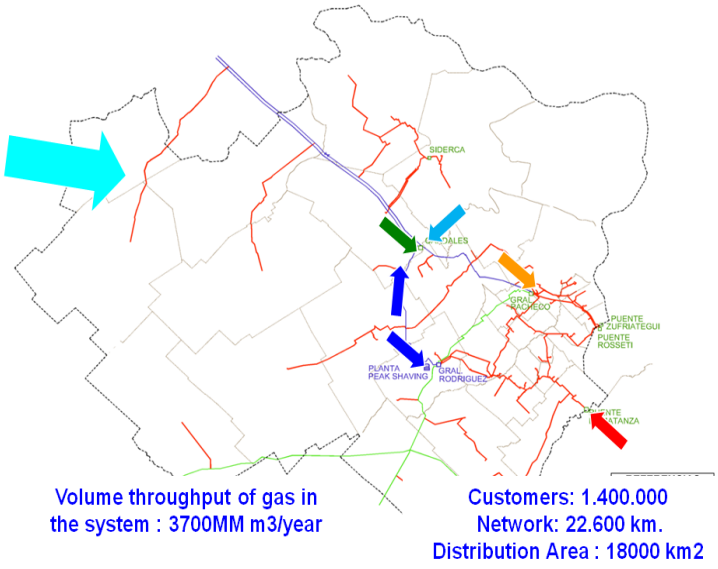
Jorge Doumanian
Guillermo Lanzillotti

Gas Natural BAN S.A.

Business Scenario

The study case was developed in Argentina, a mature natural gas market.





The business scenario is a distribution company located in the North West Buenos Aires area.

These are the main characteristics data of the distribution situation:

Volume throughput of gas in the system: 3700MM m3/year
Customers: 1.400.000
Network: 22.600 km.
Distribution Area: 15.000 km2 of the North West Buenos Aires area in Argentina

What was really happening?

During the performance of a study to determine the status of the universe of residential customers' meters, the existence of a significant number of violated meters was determined.

Those meters were intentionally manipulated in order to alter the measurement and reduce the amounts to pay for the service.

This situation contributes strongly to the unaccountable for gas (UFG), and there were no "technical reasons" that justified it.

We can see the main characteristics of these frauds & thefts types & situations in the pictures shown below.

What was the purpose of the project?

The purpose of the project was developing a methodology to find the situations of frauds & thefts inside a universe of 1.4MM customers in order to reduce the associate UFG.

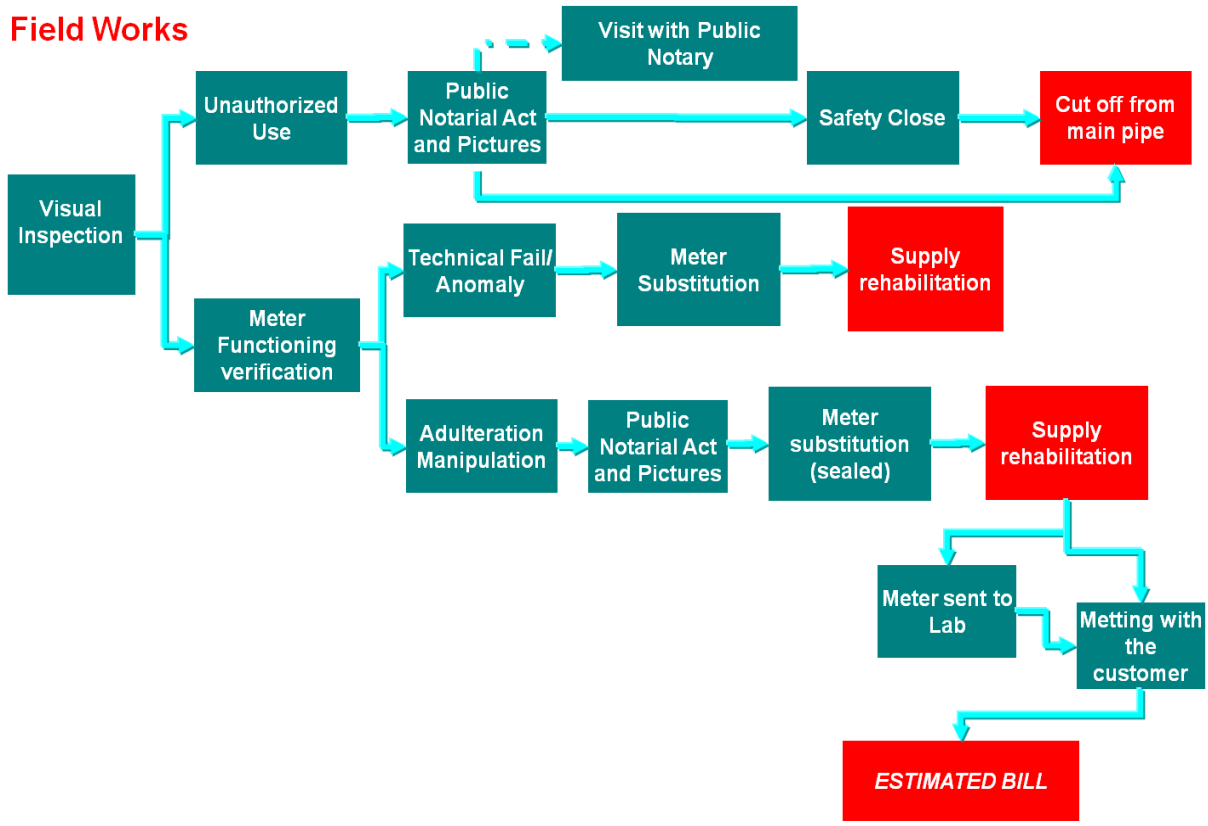
This paper describes the methodology to develop the frauds situations searching tool, and does not pretend to generalize the results in different scenarios. It shows the successfully applied of the methodologies and tools used in this particular case.

Key Findings

How was the job organized?

Two main subjects of activities were developed: Detection & Field Works.

On the Field Works, the organization is shown on the graph below



For the Detection, considering the number of installed meters (aprox.1.400.000), and taking also in consideration an annual inspection plan, we treated to determine some initial search vectors as meters reading´s warnings & consumption variations warnings, in order to avoid a complete overhaul of the meters universe.

From the studies conducted in order to find the meters cluster responsible for the UFG, and in general, no associations were found with the age of the meters.

A possibility of recording or no low flows (up to 180 / 100 l / h) by the meters was determined.
In the following table, each row is overturned with the results of the calibration values for each meter.

Then the cells were stained according to the range of error according to the following criteria;

e = -100 %
-100 % < e < -6 %
-6 % < e < 0 %
0 < e

The table shows that when a flow calibration error is -100%, for all low flows in that meter will also be the error of -100%.

The sample shows that 15% of the meters have a flow point with error equal to -100%. So, apparently earlier, 15% of the meters of the sample had an error equal to -100% in the flow of 10 l / h (about pilot flow).

This graph is the origin of the search method by neural networks, as it sets the attribute search.

This resulted from the question, does a meter registered consumption when it is driving about 100 l / h (consumption of a small burner)?

For this analysis it was not necessary to know the measurement error, but only determine whether the error was equal to -100%.

Upon removal of the base all meters that have not registered this consumption and run the simulation program, it was verified that the UFG categorically declined.

This allowed the development of a methodology from a field test, looking for those meters that did not record consumption of 100 l / h.

Meters that did not register consumption at that flow rate should be removed. This situation changed the need of calibration for a field test that consists in generating a low consumption and to check visually if the meter recorded any consumption, regardless of the error.

Arriving at this point, it was necessary to find the meters that had this characteristic.

What tools, technologies or methodologies were used & which are the resources involved in the project?

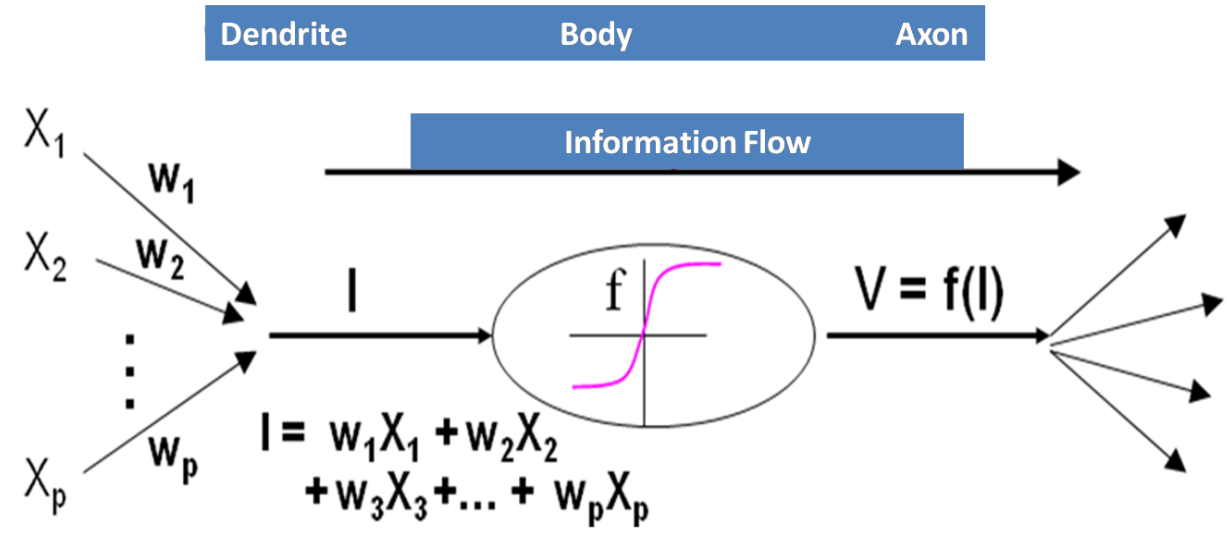
Taking in consideration the varied casuistry to determine the universe of meters to be found, it was decided to explore the possibility of using the technique of neural networks.

A thousand customers were randomly selected and performed a test was done. It consisted in closed all consumption at the house except for a small cooking stove with minimal valve opening. If the meter recorded consumption was given as good. Otherwise it was retreated to its analysis.

A data matrix for each customer was completed adding the following information:

Lector's warnings (if exists); House type; Geographical data; Consumption habits; Meter data; Test results.

Some Basic´s on Artificial Neural Networks:



Receives inputs x1, x2 ... ... xp from other neurons or from the environment.
Weighted inputs
Make the entry as a weighted sum of all total entries.
The transfer function (Activation Function) converts input into output
The output goes to other neurons or to the environment

Training Model:

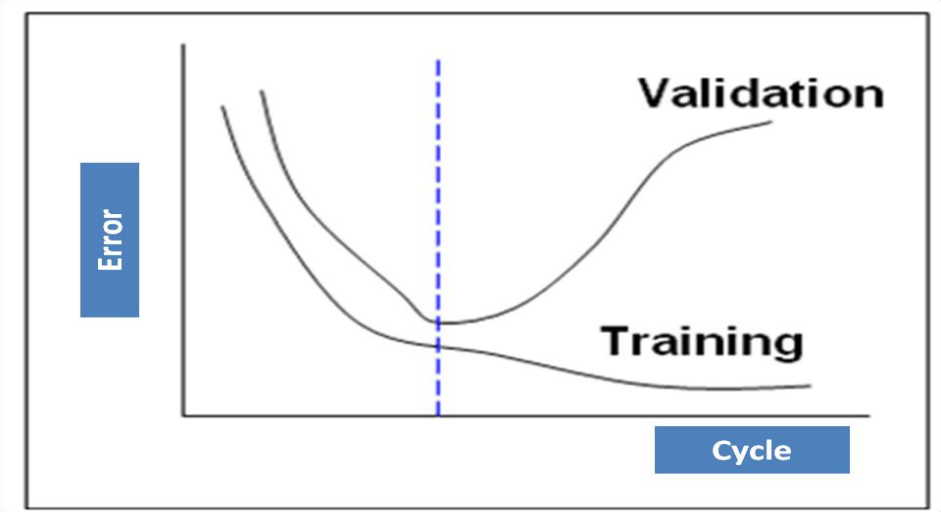
The training performed by the neural network methodology, is the detection of the patterns that govern the search criteria.

To allow the learning process, a number of resolved cases are an input to the system, which selects determined patterns depending on the particular error for its calculus. Then it checks with new cases if the prediction is correct.

The convergence criteria separate data into a "Training set" and a "Validation set."

It uses the training set to build the model, and the validation set to test the performance on unseen data.

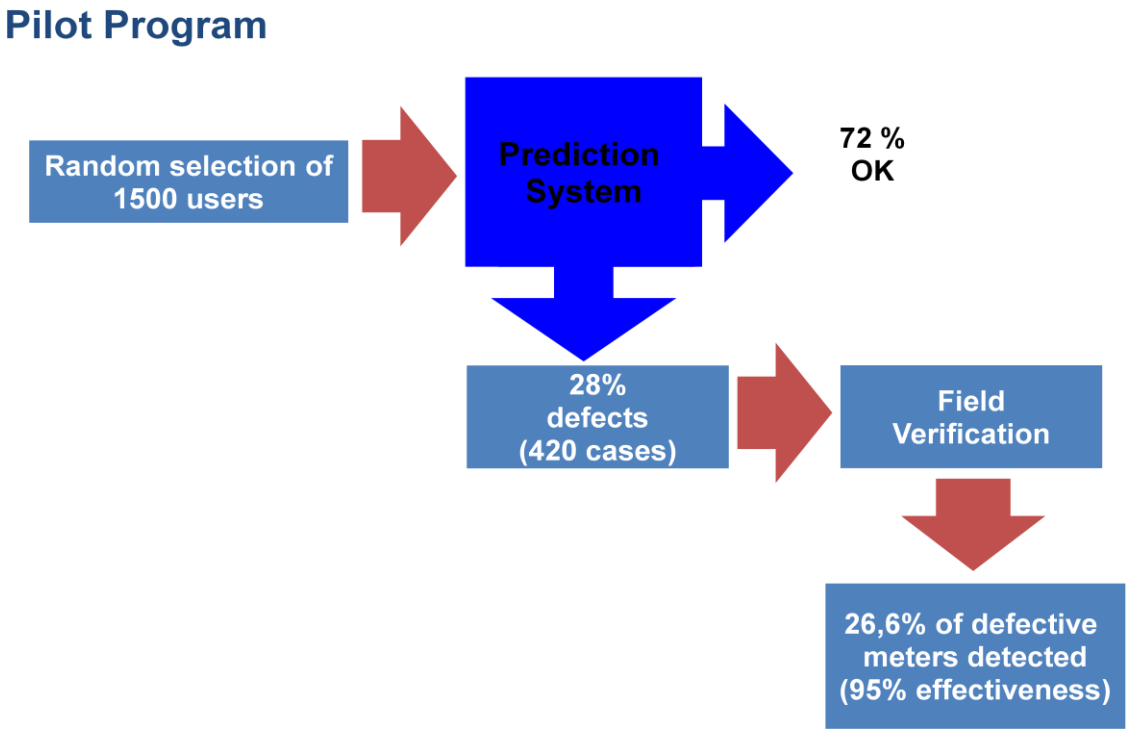
The training set error always decreases as the system adjusts the variables. The error of the validation set, by contrast first decreases and then begins to increase. This situation marks the best performance of the system and the training stops here, as it is shown in the graph below.



Pilot Program:

The Pilot Program works as follows:

First of all, based on a sample of 1500 customers, the system "learns" the pattern that characterized those meters that are violated or present large defects.



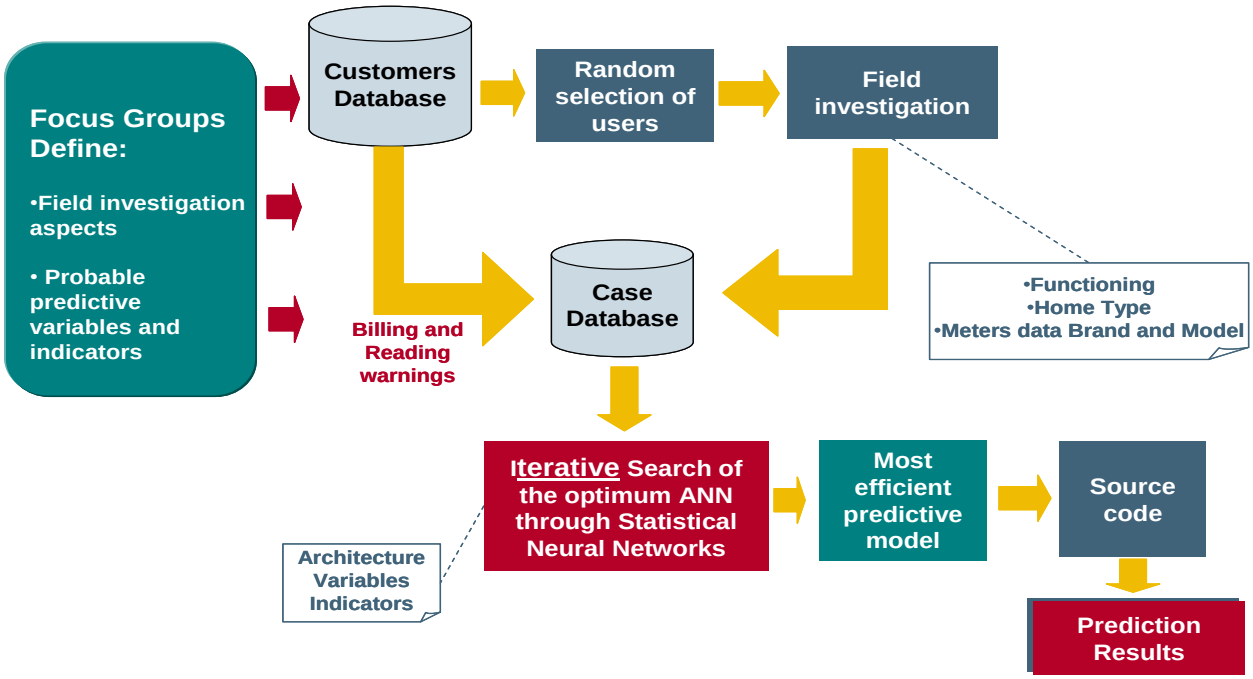
Once the model is trained it is fed with all the significant variables taking in consideration the whole residential customer´s database.

Then the program predicts which customers repeat the pattern found during its learning.

These customers are subject of a field visit and a consumption test is performed as it was already explained.

Those meters that do not register the flow are removed and taken to the calibration laboratory for analysis.

The Prediction Model works as it is shown on the following diagram:



What is the time frame?

12 months for the first results

Results

What results is the project helping to produce?

		Methodology	
Indicator	Description	Sampling	ANN
1	% of meters to verify	100%	26,5%
2	% of success on extracted meters	7- 8%	27,27%
3	% of success on total cases	100%	88,24%

- **18.659 field inspections performed**
- **8.882 cases of frauds & thefts detected & regularized**

As it is shown in the previous table, applying the Artificial Networks for the detection increased the percentage of success from 7-8% of the cases under a 100% verification to a 27% (that means a 300%), but only going to the 26.5% of the cases

Putting that in some numbers for the very beginning of the process: 18.659 field inspections performed & 8.882 cases of frauds & thefts detected & regularized

What is the value of the results?

19 MMm3/year potential UFG under control.

Conclusions

What lessons were learned from the project?

- ✓ It is possible to find a sand grain in the ocean if you open your mind
- ✓ Use IT tools that are common in other industrial applications for solving a problem that at the very beginning is very far for the final solution

What good practices were identified?

- ✓ Coordination between IT people & Field operations crew
- ✓ IT tools practical use
- ✓ Information analysis for frauds search

What are the main benefits?

- ✓ You can minimize costs and operation efforts finding with an intelligent tool the objectives you are looking for, with a high level of success
- ✓ The project let us drive the frauds search, improving the efficiency in 300% from 8% (random) to 27%
- ✓ You give a strong signal of success to the potential “thieves”

What are the critical issues?

- ✓ The field operations logistic, and the continuous maintenance and filling of the databases

How could the initiative be improved or maintained longer?

- ✓ Improved the efficiency with early detection of meters anomalies.
- ✓ Systemic reports emission showing potential measurement anomalies