March 2007

InTech®

ISA™

Crude gets smart

By Pavel Kataev, Wulf Slobodkin, Alexei Slavnov, and Lou Heavner

Neural network technology proves best choice for virtual sensing in crude refining units

FAST FORWARD

- Russian crude refinery needed more than lab operator skills to keep products within spec.
- Incorporating neural networks into new automation an ideal solution.
- Collecting data remains a major task; results show decision was best for overall production.

In crude refining units, keeping products within specifications is a challenge, especially when relying only on feedback from the lab and operator skills. At Perm, Russia's, LUKOIL refinery, the lab had been analyzing samples once or twice a day and reporting results to the operator to help him keep the process under control and products within specification. The operator tried to control key temperatures and other variables by manipulating flows including reflux, furnace fuel, pump-arounds, and product draws. Without online analyzers, product qualities had to be controlled open-loop. The company wanted to reduce conventional intermittent sampling and lab analyses. Online analyzers were costly and prone to needing maintenance, but incorporating neural networks into the new automation seemed an ideal option.

In an effort to optimize operation of a crude unit in its refinery, the company upgraded from pneumatic control to advanced digital process automation. The project included upgraded advanced regulatory control and model predictive control (MPC) technology that is fully integrated into the new automation platform. Upgrades occurred in the facility's pre-flash, atmospheric, and vacuum columns.

The three towers at the refinery operate in series with the preheated crude feeding the pre-flash tower. Here a heavy naphtha (gasoline) product recovers overhead along with liquefied petroleum gas. The reduced crude heats in a fired heater and feeds to the atmospheric tower, where a light naphtha product produces overhead. Sidedraws recover kerosene and atmospheric diesel. The bottom product, atmospheric gas oil, heats in a second fired heater and feeds the vacuum tower. The overhead vapors absorb with atmospheric diesel and remove as vacuum diesel. In addition, there are side draws for low, medium, and high viscosity vacuum gas oils and a slop stream with vacuum residue removed as the bottom product.

Various choices

The company needed to decide whether to deploy online analyzers or to use inferential models for many of the dependent variables configured as inputs to the advanced controls. They chose to use artificial neural networks for online, real-time inferential property estimation in what could be one of the world's largest installations (in terms of the number of neural models) on a single crude unit.

The neural networks run as function blocks within the automation system's controllers. Several blocks can execute in the same controller simultaneously; they need no separate computer platform; and the property models execute and update as fast as once per second. Since the neural models and MPC functionality reside in the system controllers, the advanced controls enjoy the same level of redundancy and security as traditional control loops.

Neural network pre-processing, design, training, and verification activities occurred automatically using the automation system's standard step-by-step graphical mode. An engineering workstation also called on an expert graphical mode to work off-line to change the sensitivity of some inputs by modifying the correlation between input and output. The expert mode also allowed the adjustment of detailed parameters—outlier limits, maximum/minimum number of hidden neurons, maximum number of training epochs, and the like.

The neural models are empirical and require process data from a variety of points in sufficient quantities and of high quality to accurately represent the process. We needed nearly 100 lab samples of each variable to train the models. Data collection and model refinement continues today at the refinery.

Measuring fuels

The company initially purchased licenses for 10 neural networks and applied them to develop the virtual properties of seven extraction cuts spread over the three columns:

The second gasoline end boiling point is essentially the same from a control perspective as the kerosene flash point, representing the cut point between light naphtha and kerosene. The lab test for kerosene crystallization point never ran to completion, but only until the sample was within specification. Modeling the property would have required modifying lab procedures. The other two vacuum tower properties were less important as control or constraint variables. Still, it will be possible in the future to consider building neural models for any or all of these properties.

Going digital

For the crude unit column control upgrade project, the company replaced all panel-based single-loop controls with digital process automation. While its high reliability and diagnostic capabilities increased control system availability and avoided shutdowns and slowdowns, the most significant benefits came from improvements in control design using strategies formerly impractical with single-loop controllers. When the process is stable, operation occurs closer to its economic limits, whether these limits occur at a constraint limit or at an optimum point within the acceptable operating range.

Virtual data points

The columns separate the crude oil into specific cuts by relative volatility, although the cuts are complex mixtures rather than pure components. At times, it is good to vary the cutpoints of a product such as kerosene to increase production or to meet requirements for a lighter or heavier grade. The LUKOIL laboratory determines product quality and composition using analytical tests such as ASTM distillation (usually initial and end boiling points), flash point, crystallization point, color, and viscosity.

Historically, operators de-veloped the ability to regulate quality and composition—and to maintain these values between lab analyses by monitoring various other measured properties. Every operator had a slightly different way of controlling the process, and some operators did a better job than others. However, even the best operators had other tasks that precluded them from continuously adjusting column controls.

In locations where measurements or laboratory analyses are impossible or too costly, virtual property estimation technologies give LUKOIL the opportunity to supplement measurements and lab work. Property estimators are not new, but the use of neural networks for such calculations are not common. The goal of

Pre-flash tower	
1st Gasoline (Heavy naphtha)	End boiling point
Atmospheric tower	
Kerosene	Initial boiling point
Kerosene	End boiling point
Atmospheric Diesel	50% boiling point
Atmospheric Diesel	End boiling point
Vacuum tower	
Medium Viscosity VGO	Flash point
Medium Viscosity VGO	Viscosity
High Viscosity VGO	Viscosity
Vacuum Diesel	End boiling point
Residue	Viscosity

The refinery had considered four other possible models including:

Atmospheric tower	
2nd Gasoline (Light naphtha)	End boiling point
Kerosene	Crystallization point
Vacuum tower	
High Viscosity VGO	Color
Low Viscosity VGO	End boiling point

a property estimator is to provide an accurate gauge of product quality, especially after lab results have become stale, which is most of the time. Property estimators are not intended to eliminate lab analyses, although the frequency of analyses may lessen once estimators are proven. Even though estimators may not be as accurate as lab analyses, they can be worthwhile calculated variables to help engineering and operations personnel monitor, troubleshoot, or understand and control the process.

Property estimator models can derive from first principles or develop empirically. First principles models are more robust but also more difficult to develop and may require measurements not physically obtainable. Empirical models, such as neural networks, take advantage of existing measurements to calculate new ones, and they can incorporate process knowledge to some extent. These models are not limited to variables that have a causal relationship and can discover variables that are highly correlated simply as a reflection of the process. Empirical models are not as robust as first principle designs, however, and they do not extrapolate reliably outside the range of collected data to model and train them. The latter caution is particularly true of non-linear neural models. For this reason, take care when using these property estimators as inputs to a control strategy.

While it is important to monitor the accuracy of neural models with laboratory sample data, it is also important for operators who collect samples and lab technicians who analyze them to follow procedures to assure data integrity. When an unexpectedly large discrepancy occurs between values from the lab and the model, the company determines the reason as quickly as possible. However, having continuous input from the property estimator gives the operator a new and improved gauge for monitoring the process between lab samples. Abnormal conditions produce an off-spec value and receive immediate flags, which alert the operator to take appropriate control action, even if the property estimator is not part of a closed-loop control algorithm.

Neural model nuances

Some steps to follow when building any empirical model, including neural models are data collection, pre-processing, design, training and testing, and verification. One of the challenges of building neural models is the inherent pseudo-correlations evident in process data. Auto-correlation, cross-correlation, and other statistical concerns will often fool a modeler into believing a true correlation exists when it is actually a temporary or coincidental one.

The best way to assure identifying true

correlations is to train a neural model with as much laboratory data and process data from multiple sources as possible. Ideally, no step-testing or manual disturbance of the process is necessary during network development and training. A highly stable process can make it easier to correlate process data with lab analyses, but that is a double-edged sword. You need some process variability to observe a correlation. Empirical models do not reliably extrapolate. If the process does not vary much, the model will not be reliable if the process wanders into a range with no collected training data. Slow changes or steps held for long periods are the best inputs for training the neural model.

The new configuration software can automatically perform data collection and model training needed to build and prove the neural networks and to even stop over-training when detected. Configuring data collection, model building, and model training involves familiar Windows graphical conventions.

The digital automation system's historian collects up to 20 process variables automatically for use as inputs in configuring each neural network. The configuration allows viewing of input sensitivities, exclusion of abnormal operating conditions, and eliminates variables with little or no effect on model outputs.

Collecting data a major task

As in similar projects, lab data collection was the biggest concern and challenge. The lab data must be accurate, and any error introduced during sample collection, analysis, or reporting will affect the quality of the neural model. One well known concern is accurate time-stamping of the lab data. The time stamp should reflect the time of data extraction from the process—not when it was scheduled for sampling, or when the lab technician performed the analysis, or when they reported the lab results. The sample time records in the system historian as a discrete event with a time-stamp coordinated with other system measurements.

Since no previous system history existed for modeling, we collected the data from scratch. The original goal of collecting samples during process step-testing for 10 models soon proved overwhelming. So we eventually collected data for model development for several months and used it to build preliminary models. While these models were promising, they were still not good enough. So we continued normal lab sampling, and over time the models showed great improvement.

Another consideration of data collection is to avoid filtering or manipulating the process data. Raw snapshot data usually makes for the best models. However, the long time to steady-state and the large amounts of data needed meant considering some kind of data compression. This approach can hide important information, especially transients, but the risk was accepted since the columns normally operate fairly calmly. Whenever the online model sees input data (or intermediate values) that are outside of the training range, the model status will be changed to "Uncertain" and the operator can be alerted. This is a limiting characteristic inherent in any empirical model. When we developed the online models built with the compressed data, we filtered the inputs to provide a similar averaging affect. We could perform further testing to better understand the true affect of input filtering in this application, but for now, the refinery is happy with the results.

The tool is able to identify the variables having the greatest significance. It uses statistical techniques like partial least squares and principle component analysis to identify not only the sensitivity of each input, but sensitivity over the time to steady-state. This enables the identification of time delays between when you observe process measurement and when you observe the corresponding change in lab property. Consider each input with its own time delay. And the preprocessing analysis helps identify and eliminate inputs that do not contribute to the model.

Sometimes, reasons for ignoring inputs that seem to be important are the input may not have had any variability, or another input or combination of inputs (causal or reflective) may have provided the same information. Including the additional input can be useful by increasing robustness in the event that a process measurement fails, but the overall model is degraded because the new information is redundant and noise inherent in any measurement is additive. One of the problems experienced early on was a lack of lab data. We overcame this by using part of the data twice. While it is not typically a good practice and may not be mathematically sound, it overcame the constraints when the tool did not have enough data to build a model and allowed for building and testing initial models. Later, when we had collected enough data we discontinued this practice. However, by the time we had collected and prepared the process data and lab data and selected the best inputs and time delays, most of the work was done.

The tool splits the data into two sets, one for training and one for testing. The training and testing occurs so as to identify the ideal number of hidden layer neurons and best weighting factors. This occurs on the system processor in a matter of seconds, or minutes at the most. You can verify the resulting model against the training and testing data or another set of data. Initially, when there was limited data, you could only verify against the training and testing data, but later, you could verify it against data it had not seen during training and testing. This ensured the model was truly able to perform as a general model. Continued monitoring of the predicted value with lab data identifies an outlier condition or when the model performance begins to deteriorate.

The neural estimating tool has the ability to apply an online bias that can update information on a new lab result. The ability to bias online should improve short-term performance of the estimator, but LUKOIL has not deployed it yet.

ABOUT THE AUTHORS

Pavel Kataev is an engineer of automation at LUKOIL-Inform in Perm, Russia; Wulf Slobodkin is a process analysis and modeling consultant to LUKOIL in Moscow. Alexei Slavnov is a project manager at Emerson Process Management in Moscow; and Lou Heavner is an advanced applied technologies consultant for Emerson Process Management in Austin, Tex.