## PSIG 2417



## **Automatic Tuning of Temperature and Pressure Coefficients on Pipeline Simulation**

John Hooker<sup>1</sup>, Daniel Theis<sup>1</sup>, Vadim Shapiro<sup>1</sup>

<sup>1</sup> Statistics & Control, Inc.

© Copyright 2024, PSIG, Inc.

This paper was prepared for presentation at the PSIG Annual Meeting held in Charleston, South Carolina, May 7 – 10 May 2024.

This paper was selected for presentation by the PSIG Board of Directors following review of information contained in an abstract submitted by the author(s). The material, as presented, does not necessarily reflect any position of the Pipeline Simulation Interest Group, its officers, or members. Papers presented at PSIG meetings are subject to publication review by Editorial Committees of the Pipeline Simulation Interest Group. Electronic reproduction, distribution, or storage of any part of this paper for commercial purposes without the written consent of PSIG is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of where and by whom the paper was presented. Write Librarian, Pipeline Simulation Interest Group, 945 McKinney, Suite #106, Houston, TX 77002, USA – info@psig.org.

## **ABSTRACT**

In the face of escalating climate change and the increasingly erratic daily temperature fluctuations, accurately simulating pipeline systems has become a daunting task. The complexity arises when precision in thermal modeling is crucial, especially for event detection. An illustrative case occurred in a Montana town, where the temperature swung by an astonishing 103 degrees Fahrenheit within a mere 24 hours. The primary challenge lies in developing a dependable thermal model for pipelines while contending with the substantial disturbances caused by fluctuating ambient temperatures. Even a minor shift in ambient temperature can trigger ripple effects throughout the entire system, resulting in inaccurate simulations. In response to this issue, this paper presents an innovative approach: the automatic tuning of model coefficients in real time. This method ensures the continual accuracy of the model, allowing it to adapt to varying conditions. Furthermore, it moves away from the necessity of employing multiple models, each finetuned with distinct coefficients, to accommodate diverse environmental circumstances effectively.

### **INTRODUCTION AND BACKGROUND**

Correction coefficients, also known as calibration coefficients or adjustment factors, have been used in simulation modeling for several decades. The exact origins of their use are challenging to pinpoint, but they have been a fundamental aspect of simulation modeling since the early days of computerbased simulation.

One of the earliest documented instances of using correction coefficients in simulation modeling can be traced back to the field of operations research and industrial engineering in the mid-20th century. During this time, researchers and practitioners began employing computer simulation techniques to model complex systems such as manufacturing processes, supply chains, and queuing systems. As these complex systems became modelled, researchers were able to apply these modelling techniques to the oil and gas process and industry.

In these early simulation models, it became apparent that there were often discrepancies between the simulated behavior of the system and real-world observations. The discrepancies could arise due to various factors such as inaccuracies in input data, simplifications in the model structure, or uncertainties in the underlying processes being modeled.

To address these discrepancies and improve the accuracy of simulation results, practitioners began incorporating correction coefficients into their models. These coefficients were used to adjust model parameters, input data, or simulation outputs to better align with observed real-world behavior. By iteratively adjusting these coefficients based on empirical data or expert judgment, modelers could fine-tune their simulation models to better reflect reality.

Over time, the use of correction coefficients in simulation modeling has become increasingly sophisticated, with advancements in statistical methods, optimization techniques, and simulation software enabling more systematic approaches to calibration and validation. Today, correction coefficients remain a common tool in the simulation modeling toolkit, employed across various disciplines. This paper will discuss automatic changing of correction coefficients and large disturbances in these operating conditions.

# **THEORETICAL BACKGROUND**

In order to develop a reliable robust auto-tuning procedure it is necessary to start with an accurate and versatile pipeline model.

A description of the model that is used to determine the fluid flow and pressure drop can be found in our previous paper on Long Haul Natural Gas Pipeline Compressor Station Optimization which shows the gas flow model.[1] That paper discusses how equations of state and other empirical formulas that are used to compute gas and liquid properties, provides an overview of the conservation equations (of mass, momentum, and energy) that form the foundation of the pipeline simulation, and describes how those formulas are used to compute the changes in pressure and flow rate across the pipeline.

In this instance we will focus on the temperature model especially with heat transfer to the environment which hasn't been explained in quite as much detail. The conservation of energy determines the amount of energy that is transferred from a fluid into the environment as it travels along a pipeline. It also influences how the temperature of the fluid will change. Mathematically, the conservation of the specific energy of the fluid  $E_m$  (at a given point of along the pipeline), can be expressed by the following equation:

$$
\frac{\partial E_m}{\partial t} = -v \frac{\partial E_m}{\partial x} - \frac{v}{\rho} \cdot \frac{\partial P}{\partial x} - \frac{4K_{Total}(T - T_A)}{\rho D}
$$

where:

- *P*, *T*, *v*, and  $\rho$  specify the pressure, temperature, superficial velocity, and density of the fluid,
- *D* specifies the inner diameter of the pipe,
- *T<sup>A</sup>* specifies the ambient temperature of the environment that surrounding the pipe,
- and *KTotal* specifies the total heat transfer coefficient that is associated with both the fluid and the environment that surrounds the pipe.

The three terms that appear in the equation above represent the changes to  $E_m$  (for a given point along the pipeline) that result from: the flow of fluid into and out of that point of the pipeline; the work that is performed on the fluid due to the changes in the pressure across the pipeline; and the heat that dissipates from the fluid, through the pipe wall, to the surrounding environment. Now that the change in energy of the fluid is known we can move to a change in temperature.

Starting from the equation above, the following expression can be derived and used to determine the changes that will occur to the temperature of the fluid as it flows through the pipe:

$$
\frac{\partial T}{\partial t} = -v \frac{\partial T}{\partial x} + \frac{1}{c_P} \cdot \left( \frac{fv^3}{2D} - \frac{4K_{Total}(T - T_A)}{\rho D} \right)
$$

where  $c<sub>P</sub>$  denotes the specific heat capacity of the fluid and *f* denotes the friction factor that is used to compute the frictional force that opposes the movement of the fluid.

The first term identifies the change to temperature that results

from the flow of fluid into and out of a given point of the pipeline. The remaining terms determine the increase in temperature due to friction and the decrease in temperature due to the heat that dissipates out of the pipe. To predict the changes that occur to temperature and energy in the fluid accurate values must be identified for *f* and *KTotal*. The remainder of this section will focus on the evaluation of *KTotal*. This will help determine the overall heat transfer coefficient that we're looking to automatically update.

As heat is transferred from the fluid to the environment it passes through the following substances:

- a thin film of the fluid which coats the inner surface of the pipe,
- the wall of the pipe,
- and any insulation that surrounds the pipe.

The physical properties related to each of those substances influence the rate that heat is transferred from the fluid to the environment (*i.e.*, they influence *KTotal*). The physical properties of the environment also influence *KTotal* (its properties control the dispersion of the heat throughout the environment).

Based on those observations *KTotal* can be separated into contributions from each of the substances listed above:

$$
\frac{1}{K_{Total}} = \frac{1}{h_{Film}} + \frac{1}{h_{pipe}} + \frac{1}{h_{Ins}} + \frac{1}{h_{Env}}
$$

where:

- *hFilm* denotes the heat transfer coefficient of the fluid film,
- *hPipe* denotes the heat transfer coefficient of the pipe wall,
- $h_{Ins}$  denotes the heat transfer coefficient of the insulation material,
- and *hEnv* denotes the heat transfer coefficient of the environment.

When heat moves through a solid material it is transferred via thermal conduction. The rigid structure of solid materials and the well-known characteristics of thermal conduction allowed scientists to determine simple yet highly accurate formulas for the heat transfer coefficients of solid substances (even for complex substances like the ground).

For pipes and pipe insulation the following formulas can be used:

$$
h_{pipe} = \frac{2k_{pipe}}{D \ln \frac{D_{out}}{D}}
$$
 and 
$$
h_{Ins} = \frac{2k_{pipe}}{D \ln \frac{D_{Ins}}{D_{out}}}
$$

where:

• *kPipe* and *kIns* specify the thermal conductivities of the pipe and the pipe insulation,

- *D* specifies the inner diameter of the pipe,
- $D_{Out}$  specifies the outer diameter of the pipe (which is also the inner diameter of the pipe insulation),
- and  $D_{Ins}$  specifies the outer diameter of the pipe insulation.

The table below displays typical values for the thermal conductivities of several substances:





Several accurate empirical formulas for the heat transfer coefficient of the ground (for buried pipes) have also been developed by adding empirical corrections to the following core equation:

$$
h_{Env} = \frac{2k_{Ground}}{D \ln \frac{2z_{Depth}}{D_{External}}}
$$

where:

- *kGround* specifies the thermal conductivities of the ground,
- *zDepth* specifies distance between the surface of the ground and the center of the pipe,
- and  $D_{External}$  equals  $D_{ins}$  for insulated pipes and  $D_{Out}$  for non-insulated pipes.

Building on the thermal model discussed above a tunable model can be generated by introducing tunable coefficients. To adjust the accuracy of values that are computed for the pressures and flow rates, tunable coefficients ( $\alpha_{frict}$ ) are introduced into the expression that computes the pressure drop due to friction

$$
\Delta P = \alpha_{frict} \cdot \frac{fv^3}{2D}
$$

To adjust the accuracy of the values that are computed for the temperatures, tunable coefficients ( $\alpha_{heat}$ ) are introduced into the expression that is used to compute the rate that heat dissipates into the environment

$$
\frac{\partial Heat}{\partial t} = -\alpha_{heat} \cdot \frac{4K_{Total}(T - T_A)}{\rho D}
$$

Now that the thermal model is set up and adjustment coefficients have been created, an auto-tuning algorithm can be created to adjust the coefficients to simulate an accurate model.

#### Auto-tuning and Accuracy

Maintaining the accuracy of the simulated data with respect to the real process is the primary purpose in creating a digital twin. Improvements to the model's accuracy are achieved using signal processing, numerical optimization, and statistical process control algorithms. Using those algorithms optimal values are computed for the tuning coefficients. Those values are determined by minimizing the variance that exists between the simulated and field data for several key process variables.

The amount of time it takes for the digital twin to reach its target accuracy depends on several factors, including the pipeline's complexity, the number of process variables that are tracked, the available amount of field data, and the presence/absence of transient events. For most pipelines, the desired accuracy is typically reached after 1-2 weeks of operation. By using accuracies from multiple components (process variables) to compute (and minimize) the cumulative accuracy, the results of the adjusted simulation become dependable and robust.

Before optimizations are performed to determine the tuning coefficients, a data filtering algorithm is used to remove outliers from the field measurements that were recorded for the process variables. The filtering algorithm uses the following formula to identify which data points are outliers:

$$
Test\left(\hat{F}_k(t)\right) = \begin{cases} Use & (\left|\hat{F}_k(t) - \mu_k\right| \ge \delta_k) \\ Reject & (\left|\hat{F}_k(t) - \mu_k\right| < \delta_k) \end{cases}
$$

where:

- $\hat{F}_k(t)$  denotes the value that was measured for process variable *k* at time *t*,
- $\mu_k$  denotes the mean value that was determined for process variable *k* using the full set of data points,
- and  $\delta_k$  denotes the threshold criterium for the outliers of process variable *k*.

The value of  $\delta_k$  is either a function of measurement distribution of process variable *k* (*e.g.,* a multiple of its standard deviation) or it is manually specified by the process engineer (*e.g.,* an upper/lower control limit). The output of the data filtering procedure is the set of all the acceptable field measurements for the process variables.

Once the data filtering procedure is complete simulated date points are determined for each of the acceptable field measurements. In the expressions that follow,  $\hat{S}_k(t)$  denotes the simulated data point that corresponds to the field measurement  $\hat{F}_k(t)$ . The accuracy of the model is then optimized by using least squares fitting techniques to minimize the deviation between the simulated and field measurements  $\hat{F}_k(t) - \hat{S}_k(t)$ . Additional rounds of data collection, data filtering, coefficient optimization will continue until the overall accuracy of the model reaches a preset accuracy criterion.

The accuracy of each individual reading is computed as a scaled error:

$$
Scaled Error = \frac{\hat{F}_k(t) - \hat{S}_k(t)}{S_{k,max} - S_{k,min}} \times 100
$$

where:

- $\hat{S}_k(t)$  and  $\hat{F}_k(t)$  correspond to the simulated and field values that were determined for process variable *k*,
- And  $S_{k,max}$  and  $S_{k,min}$  correspond to maximum and minimum limits that are imposed on process variable *k*.

Normally auto-tuning optimizations need to determine the set of modeling coefficients that reduce deviations between the field and simulated measurements for multiple process variables. The overall accuracy of the whole model is determined by calculating individual scaled errors for each process variable and using those values in following formula:

$$
Accuracy = \frac{\sum_{k=1}^{P.V.} \{w_k \cdot \varepsilon_k\}}{\sum_{k=1}^{P.V.} \{w_k\}}
$$

where  $\varepsilon_k$  specifies the scaled error of process variable  $k$ ,  $w_k$  specifies the weight that process variable  $k$  makes to the overall accuracy, and *P.V.* denotes the fact that the summation occurs overall every process variable.

Over time the accuracies should form a distribution, *dP,* that is similar to the distributions shown in figure 1. Successfully optimized tuning coefficients should produce an accuracy distribution that has a mean which is close to zero and has its majority of accuracies below a predefined Upper Specification Limit, *USL*. The concept is illustrated by the curves displayed in figure 1.



**Figure 1 – Accuracy Submodule Concept**

The algorithm calculates a process capability index which is illustrated in the equation below. For the accuracy calculation we only care about the Upper Specification Limit.

$$
C_{p,u} = \frac{USL - \mu}{3\sigma}
$$

where  $\mu$  is the mean of *dP* distribution and  $\sigma$  is the standard deviation within a parametrically defined time period. The accuracy algorithm allows the user to change *USL* online. In this case, the accuracy calculations will be adjusted in real time. Since accuracy has one-sided specifications, the demands on *Cp,u* values are not as stringent as with two-sided specifications. Thus, following generally accepted statistical process control principles, the process is considered "capable" or "in control" whenever  $C_{p,u} \geq 1.33$ . Typically, the system learning and optimization processes will continue until  $C_{p,u} \geq 1.67$ . In this case, the desired accuracy level would be considered reached.

The final output of the algorithm is a set of adjusted coefficients that are provided to the initial model. The overall algorithm repeats whenever a process change occurs, or the overall accuracy will fall out of range. Now that the algorithm is established it needs to be scalable to handle multiple coefficients with many loads, real time data, and pipeline information.

Software system scalability describes maintaining efficiency and effectiveness when many components are added. The accuracy algorithm has been designed to do just that by utilizing parallel processing available in multi-core hardware systems as well as by working with very large data sets, also known as big data. The system is prepared to carry out advanced control actions for massive multi-unit operations. For example, a gas pipeline may contain hundreds of miles of pipes with several compressor stations, which results in thousands of parameters that need to be simulated and modeled. These parameters may need to be tracked at the second level, which produces massive

quantities of data (thousands of columns, multiple millions of records) that need to be analyzed, smoothed, and used in the model process. The accuracy algorithm handles this by saving the model coefficients and learning when it's acceptable to use different coefficients based on the input data, operating mode of the process, and main variables that are fed into the process. With the algorithm and scalability, we can see how well it can be used in the field.

## **CASE STUDIES**

#### Test Study of Steam Distribution System

The accuracy algorithm was run with historical data that was taken from a geothermal steam project. The geothermal project had 17 steam wells that were connected to 2 steam turbines that generated about 100 MW of electricity each. The system was looking for an accuracy of 5% and was able to reach that value within a week where it achieved a value of 2.14 *Cp,u* and an overall model accuracy of 99%









Table 2 shows accuracy values for every modeled process characteristic as well as the overall plant model. The **Accuracy %** column is the mean of the *dP* distribution during the oneweek time period the system has been in training. The table also shows current field measurements (shown in green) and the simulated values (shown in blue) for each process variable during the corresponding time scan. The numbers in red indicate process variables that have scaled errors above 5% (the upper specification limit). The Weight column indicates the

importance of each subcomponent in calculating the overall accuracy, which is the weighted average of its subcomponents. The user is allowed to change weights online. The accuracy calculations will instantaneously reflect those changes.

The next figure shows distribution parameters of the selected model. This distribution provides an "at-a-glace" view of the process and allows engineers and plant operators to monitor the inner workings of the system. This feature allows the user to enter the Upper Specification Limit as well as tolerance for accepting outlier detection errors. The date and time stamp on the left hand side of the screen shows the currently analyzed time scan. The counters of good and bad points indicate how many outliers have been rejected and how many observations have made it through the data justification calculation. The distribution parameters are then visually displayed via the box plot in the middle of the screen as well as the distribution function on the right-hand side of the screen. The *Cp,u* value at the bottom of the data table is the critical component that the user should monitor to identify whether desired accuracy has been reached. A larger version is available in the Appendix for easier viewing.



**Figure 2 – Accuracy Graphics**

The bottom part of the screen in figure above shows the learning process comparing measured process variables (blue line) to the simulated data (green line). All information is updated either by reading signal logs or by incorporating real-time process variable values.

Application to a Commercial Pipeline. The proposed autotuning and simulation techniques were also applied to a commercial gas pipeline. The pipeline is a 302 mile long natural gas transmission pipeline and is comprised of a 30-inch diameter. The line runs on average of 430 MMBTU/day of gas. The maximum allowable operating pressure is around 1440 psig and runs around 1200 psig. The pipeline has 16 mainline valve stations with a pressure and temperature probe at each valve station.

<b>Valve</b> <b>Station</b>	<b>Milepost</b>	<b>Miles</b> <b>of</b> <b>Pipe</b>
MLV1	17.2	17.2
MLV2	36.1	18.9
MLV3	53	16.9
MLV4	71.2	18.2
MLV5	87.7	16.5
MLV <sub>6</sub>	104.8	17.1
MLV7	123.3	18.5
MLV8	141.6	18.3
MLV9	160.9	19.3
MLV10	179.9	19
MLV11	198.5	18.6
MLV12	214.7	16.2
MLV13	233.3	18.6
MLV14	252.8	19.5
MLV15	276.6	23.8
MLV16	295.2	18.6
End	303	7.8

**Table 3 – Main Line Valve Stations**

Here the pipeline operator has temperature and pressure probes about every 20 miles which allows for consistent reading across the pipeline. The pipeline is buried for the most part and doesn't have any large elevation changes. The ground temperature is taken from two readings. One temperature at the beginning of the pipeline and one right next to MLV#9 station. For the gas pipeline it's a simple operation where the inputs are steady and one of the biggest changes in ambient temperature. The inputs for Scenario 1 are in the table below:

#### **Table 4 – Input for Scenario 1**



The accuracy algorithm will compare this scenario with Scenario 2 where the ambient temperature drops quite significantly. The inputs for Scenario 2 are similar except it's 0 Fahrenheit instead of 65.





## **RESULTS**

Looking at the Accuracy for the 65 degree Fahrenheit scenario the accuracy is achieved of less than 1% which is great. The pipeline didn't have as many fluctuations as the steam field, so the tuning was able to reach a CPU of less than 1% within 24 hours. Below is a table of accuracy calculations with the Value Field (green value) being the data from the sensors and the Value Sim (blue value) being the simulated pipeline numbers.

**Table 6 – Accuracy Calculation Gas Pipeline Scenario 1**

Model/ Object	<b>Accuracy</b> (%)	<b>Value</b> <b>Field</b>	<b>Value</b> Sim	Weight
<b>Steam</b> <b>Model</b>	0.98028	USL:1	Cpu: 0.98319 <b>Reached</b>	
$P$ (Psig): P101	0.03357	1201.81	1198.453	$\mathbf{1}$
$T (°F)$ : T <sub>101</sub>	0.00569	68.3	67.731	1
$P$ (Psig): MLV1	0.01558	1173.48	1171.922	1
$T (^{\circ}F)$ : MLV1	0.00443	67.5	67.057	$\mathbf{1}$
$P$ (Psig): MLV2	0.04073	1142.98	1138.907	$\mathbf{1}$
$T (^{\circ}F)$ : MLV <sub>2</sub>	0.00283	66.9	66.617	$\mathbf{1}$
$P$ (Psig): MLV3	$-0.00842$	1115.32	1116.162	1
$T (^{\circ}F)$ : MLV3	0.00129	67.1	66.971	1
P(Psig): MLV4	0.02605	1084.28	1081.675	1
$T (°F)$ : MLV4	0.00579	68	67.421	1
$P$ (Psig): MLV <sub>5</sub>	0.02737	1052	1049.263	1



that are above the maximum error in accuracy calculations. In this example 5% is the upper limit and 1% is the lower limit for accuracy, so if a signal were 5% or greater it would show up on the chart in red. Now that the accuracy is at an achievable level the scenario was run through where the outside temperature went down from 65 degrees Fahrenheit to 0 degrees Fahrenheit.

#### **Table 7 – Accuracy Calculation Gas Pipeline Scenario 2**



The accuracy calculations would also show any rejected values



The model was trying to automatically adjust to the ambient temperature, but the red values show pressures that are out of range. The system accuracy is 3% which is still pretty good for the model coefficients for a 65 degree change in ambient temperature. The difficult part is how quickly the temperature can drop and the system can adjust to the ambient temperature change. The temperature of the system wasn't affected as much, but the pressure dropped due to the temperature drop.

## **CONCLUSIONS**

The system could keep up with relatively good accuracy. The heat transfer in the pipeline has room for improvement in being able to handle large atmospheric changes in temperature. Another potential area of improvement is to be able to have outside temperature at each main line valve station or ground temperature at each valve station to have a more accurate live version of the temperature where it's happening. The accuracy

calculation has proven to be able to handle multiple variables and being able to scale up to large gas pipeline fields as well as large geothermal fields. Another area of interest in the field is being able to track the accuracy and the model coefficients to be able to see the changes that happen over time and see which areas and processes are subject to the largest fluctuations in the system.

### **REFERENCES**

- 1. V. Shapiro, J. Hooker L. Youngblood, "Long-Haul Natural Gas Pipeline Compressor Station Optimization" *Pipeline Simulation Interest Group*, PSIG 1816, 2018.
- 2. M. V. Lurie *Modeling of Oil Product and Gas Pipeline Transportation*, WILEY-VCH Verlag GmbH & Co., Weinheim, 2008.
- 3. C. Long and N. Sayma, *Heat Transfer*, Ventus Publishing ApS, E-Book, 2009.
- 4. S. J. Amir, *Calculating Heat Transfer from a Buried Pipeline*, Process Heat Exchange, McGraw-Hill, 1979.
- 5. F. W. Dittus and L. M. K. Boelter, *Heat Transfer in Automobile Radiators of the Tubular Type*, University of California Publications in Engineering, 1930.
- 6. S. Churchill and H. Chu, *Correlating Equations for laminar and turbulent free convection from a horizontal cylinder*, Int. J. Heat Mass Transf. 18, 1049 (1975).

### **AUTHOR BIOGRAPHY**

**John Hooker** is a controls engineer and the Director of Engineering at Statistics & Control, Inc. He has a Bachelor of Science in Aerospace Engineering from Iowa State University. Over the past 14 years, John has participated in several global projects in the oil & gas and power generation sectors and has implemented numerous applications for control solutions, equipment monitoring, design testing, data analysis, and brownfield optimization.

**Dr. Daniel P. Theis** is a software engineer working at Statistics & Control, Inc. He has a Bachelor of Science in Chemistry from Bethel College, a Master of Science in Physical Chemistry from the University of Minnesota, and a Doctorate in Computational Chemistry from the University of North Dakota. Daniel has over 15 years of experience in the development and use of computer software designed to predict the properties and behavior of complex chemical systems.

**Vadim Shapiro** is co-founder and President of Statistics & Control, Inc. He manages all aspects of engineering, including development of the company's *OptiRamp*® Advanced Process Control software. Vadim has over 25 years' leadership experience in systems and software engineering, project management, and product development in the field of turbo

machinery control, advanced process control, and power management systems. He holds six patents in the areas of turbomachinery control and advanced process control, with several applications pending.

# **FIGURES**



**Figure 1 – Accuracy Calculation Concept**



**Figure 2 – Accuracy Graphics**