Parameter Studies for Energy Networks with Examples from Gas Transport

Tanja Clees

Abstract The focus of this chapter is on methods for the analysis of parameter variations of energy networks and, in particular, long-distance gas transport networks including compressor stations. Gas transport is modeled by unsteady Eulerian flow of compressible, natural gas in pipeline distribution networks together with a gas law and equations describing temperature effects. Such problems can lead to large systems of nonlinear equations with constraints that are computationally expensive to solve by themselves, more so if parameter studies are conducted and the system has to be solved repeatedly. Metamodels will thus play a decisive role in the general workflows and practical examples discussed here.

Keywords Metamodeling • Response surfaces • Design of experiments • Parameter analysis • Robustness • Gas networks

MSC codes: 93B51, 65D17, 62K20, 62K25

1 Introduction

Networks rule the world. The well-known social networks are just one example. Infrastructure for transport of gas, electricity, or water, but also electrical circuits inside technical devices are other important instances. Due to the ongoing transformation of our energy production and incorporation of increasingly larger amounts of renewable energy sources, energy networks of different types (electrical grid, gas, heat, etc.) have to form an integrated system allowing for balancing of supplies and demands in the future. Conversions between different energy media (power-to-gas, power-to-heat, etc.) and storages provided by, for instance, pipeline systems and caverns will play a decisive role. This increases the demand for enhanced cross-energy simulation, analysis and optimization tools.

T. Clees (✉)
Fraunhofer SCAI, Sankt Augustin, Germany
e-mail: tanja.clees@scai.fraunhofer.de

S. Koziel et al. (eds.), *Simulation-Driven Modeling and Optimization*, Springer Proceedings in Mathematics & Statistics 153,
DOI 10.1007/978-3-319-27517-8_2

Transport networks for energy or water as well as circuits can be mathematically modeled in a very similar fashion, based on systems of differential-algebraic equations. Their numerical simulation can be performed based on the same or at least similar numerical kernels.

In Section 2, the physical model for flow of compressible, natural gas in pipeline distribution networks with several technical elements is sketched. In a similar fashion, one can model electrical grids, pipeline systems and other energy transport networks as well.

Afterwards, in Section 3, we describe analysis tasks which are common for many energy networks and define some terms important there. Section 4 gives a short overview on a selection of methods frequently used. A flow chart asking some key questions and several workflows are outlined in 5. Section 6 summarizes several versatile visualization techniques. Based on these workflows, several examples from gas transport analysis are studied in some detail in Section 7. Finally, Section 8 concludes this chapter.

2 Gas Transport Network Modeling

In the following, the physical model considered here is sketched. More details can be found in [24]. Several ongoing research aspects such as model order reduction, ensemble analysis, coupled network and device simulation are discussed in [5, 13, 18, 20, 21, 28], for instance. Throughout this chapter, we use MYNTS (MultiphYsical NeTwork Simulation framework), see [1, 4]. In MYNTS, the physical model described in the following is implemented.

2.1 *Isothermal Euler Equations*

A gas transport network can be described as a directed graph $\mathscr{G} = (\mathscr{E}, \mathscr{N})$ where $\mathscr{N}$ is the set of nodes and $\mathscr{E}$ is the set of directed edges denoted by tuples of nodes. If we choose to not consider more involved components as a start, each edge constitutes a pipe and can thus be specified by length and width. Nodes, on the other hand, can be imagined as the points where pipes start, end or meet. The physics of a fluid moving through a single pipe can be modeled by the isothermal Euler equations. Discarding terms related to kinetic energy for simplification we get:

$$\partial_t \rho + \partial_x q = 0 \tag{1}$$

$$\partial_t q + \partial_x p + \partial_x \left(\rho v^2\right) + g\rho\partial_x h + F = 0 \tag{2}$$

$$p = R_s T z \rho. \tag{3}$$

The hydraulic resistance is given by the Darcy-Weisbach equation:

$$F = \frac{\lambda(q)}{2D}\rho v|v| \tag{4}$$

Here, $\rho = \rho(x,t)$ denotes the density, $p = p(x,t)$ the pressure, $q = \rho * v$ the flux, $v = v(x,t)$ the velocity, $T = T(x,t)$ the temperature, $h = h(x)$ the geodesic height, $D = D(x)$ the pipe diameter, z the compressibility, R_s denotes the specific gas constant (see also Section 2.3), and $\lambda = \lambda(q)$ being the friction coefficient.

Together with Kirchhoff's equations, the nonlinear system to be solved is defined. The meaning of the equations is as follows:

- the continuity (1) and Kirchhoff's equations modeling the mass (or molar) flux
- a pipe law (2) and Darcy-Weisbach modeling pressure-flux (see Section 2.2)
- a gas law (3) modeling pressure–density–temperature (see Section 2.3)

ρ, p, q, v, T form the set of unknown dynamical variables.

2.2 Pipe Laws

The friction coefficient in the Darcy-Weisbach equation can be modeled by, e.g.:

$$\lambda = \left(2\log\left(\frac{4.518}{\mathrm{Re}}\log\left(\frac{\mathrm{Re}}{7}\right) + \frac{k}{3.71D}\right)\right)^{-2} \quad \text{Hofer} \tag{5}$$

$$\lambda = \left(2\log\left(\frac{D}{k}\right) + 1.138\right)^{-2} \quad \text{Nikuradze} \tag{6}$$

where Re is the Reynolds number and κ a parameter describing the roughness of the pipe currently considered. For high Reynolds numbers, Hofer approaches the easier Nikuradze equation.

2.3 Gas Laws

For ideal gases, $pV = nRT$ holds where p denotes pressure, V volume, n amount (in moles), R ideal gas constant, and T temperature of the gas. Defining the specific gas constant R_s as the ratio R/m with m being the mass, we get $p = \rho R_s T$. For the non-ideal case, compressibility z is introduced, and Eq. 3 holds. Several gas laws are frequently used:

$$z = 1 + 0.257p_r - 0.533\frac{p_r}{T_r} \quad \mathrm{AGA(upto70bars)}$$

$$z = 1 - 3.52\exp(-2.260T_r)p_r + 0.247\exp(-1.878T_r)p_r^2 \quad \mathrm{Papay(upto150bars)}$$

$$z = 1 + B\tilde{\rho} - p_r\sum_n C_n + \sum_n C_n\left(b_n - c_n k_n \rho_r^{k_n}\right)\rho_r^{b_n}\exp(-c_n\rho_r^{k_n}) \quad \mathrm{AGA8-DC92}$$

Here, p_r, T_r, ρ_r denote the reduced pressure, temperature, and density, respectively. A reduced property is obtained by dividing the property by its (pseudo-)critical value. If AGA8-DC92 shall be applied, the fractions of all components in the gas mix have to be computed as well, cf. 2.5. AGA8-DC92 contains several further coefficients and constants which are not further explained here. Papay is quite popular. However, for an accurate simulation, AGA8-DC92 should be used though.

2.4 Network Elements

Several types of nodes should be distinguished:

- Standard supplies: Input pressure, temperature and gas composition are defined.
- Standard demands: Either output mass flow or volume flow or power is defined.
- Special supplies: Either input mass flow or volume flow or power is defined.
- Interior nodes: The remaining nodes, where nothing is defined (besides (h).

For all nodes, their geodesic height h has to be given.

Besides nodes and pipes, several elements are present in many gas transportation networks. The following elements are considered here, a typical selection:

- Compressors, described by a characteristic diagram (engine operating map; see, for instance, Figure 4).
- Regulators: Input pressure, output pressure, output volume flow are typical regulation conditions; regulators might be described by a characteristic curve.
- Coolers, heaters.
- Valves, resistors, flaptraps.
- Shortcuts: A special element which can be seen as an extremely short pipe.

2.5 Gas Mixing and Thermodynamics

Natural gas consists of 21 components, and the by far largest fraction is methane. For modeling the molar mix of gas properties, such as combustion value, heat capacity, fractional composition, etc., the system has to be enlarged and reformulated to take 21 gas components into account.

Also for modeling thermodynamical effects, the composition is incorporated. Several important effects have to be considered:

- Heat exchange (pipe-soil): A (local) heat transfer coefficient is used.
- Joule-Thomson effect (inside the pipe): Temperature change due to pressure loss during an isenthalpic relaxation process.

The system is then enlarged by an equation describing the molar mix of enthalpy (a form of internal energy). Arguments include, in particular, the heat capacity and the critical temperature and pressure, as modeled by the gas law chosen. For modeling the gas heating inside compressors, the isentropic coefficient has to be considered as well. Several models for approximating this effect exist, cf. [23], for instance.

2.6 Outputs, Parameters and Criteria

The following terms are used here:

- Input: Settings/values which have to be defined before starting a simulation.
- Output: Results stemming from a simulation.
- Parameter: An input which shall be varied for a specific analysis task.
- Criterion: A single value or distribution computed from one or more outputs; a criterion might be a global value or one with only a local meaning for the network given.

Several general questions arise for each parameter:

- Of which type is the parameter: discrete (e.g. state of a valve) or continuous (e.g. input pressure)?
- Which (range of) values shall be considered for the parameter for the specific analysis task? Examples are:
  - "on" and "off" for the state of a valve
  - the interval [52.0; 60.0] for an input pressure
- Can all parameters be varied independently? If not, is the dependency be known in advance or result of another process?
- Which type of distribution shall be considered for the parameter for the specific analysis task?
- How is this distribution be defined?
  - Function: Analytic (physical model) or fitted to data stemming from measurements or simulations (attention: the method for and assumptions behind the fitting might have a large impact).
  - Histogram (raw data) resulting from another process.

See Table 1 for a selection of input parameters, varied in the examples (see Section 7), as well as output functions, analysed in more detail.

Table 1 Input and output functions (selection).

shortcut	function	node/edge	remark
pset	input: pressure at supply	node	values in *bar* here
qset	input: volume flow at demand	node	values in $1000 Nm^3/h$ here
m	mass flow	edge	values in kg/s here
p	pressure	node	values in *bar* here
pslope	pressure difference	edge	$p_2 - p_1$ on the edge
t	temperature	node	values in Kelvin here
tslope	temperature difference	edge	$t_2 - t_1$ on the edge
had	head of compressor	edge	change of isentropic enthalpy $[kJ/kg]$
qvol	volume flow through compressor	edge	volume per second $[m^3/s]$

A General Remark. Depending on the concrete physical scenario to be solved and, in particular, conditions for compressors and regulators, either a set of equations or a set of equations and constraints (and possibly an objective function) has to be solved. We call both cases "simulations" in the following, though.

3 Analysis Tasks and Ensembles

Parameters of the model can vary, depending on their meaning and physical and/or numerical scenarios considered. A parameter variation might cover a tiny up to a huge range of values, in one or more intervals, and different types of distributions might be considered. Typical ones are

- uniform
- (skewed) Gaussian (see, for example, Figure 9)
- based on a histogram harvested from measurements

Different tasks can be solved. Important ones include

- comparison of scenarios and visualization of differences on the net (Section 6.3)
- stability analysis (Section 3.1)
- parameter sensitivity and correlation analysis (Section 3.2)
- robustness analysis and analysis of critical situations (Section 3.3)
- robust design-parameter optimization (RDO, Section 3.4)
- calibration of simulation models (history matching, Section 3.5)
- analysis of the security of energy supplies: this has to be carried out, in particular, for electrical grids; the so-called $(N-1)$-study is usually performed. If N is the number of (at least) all (decisive) components, N simulations are performed in each of which another one of these components is assumed to fall out.

If the data basis for such an analysis task is a collection of results from several simulation runs or measurements, we call this data basis *ensemble* here.

In the following, we describe the tasks listed above in more detail. Afterwards, we will present and discuss methods for creating ensembles and analysing them.

3.1 Stability Analysis

We call (the design of) a scenario *stable* if tiny changes of initial conditions and physical properties have a tiny impact on the results only. We call it *instable*, if tiny changes in the initial conditions lead to substantially different results with a large portion of purely random scatter. We call it *chaotic*, if tiny changes in the initial conditions lead to substantially different and even unpredictable results. Instability might stem from physical and/or numerical issues, and it is difficult to separate these effects in many cases.

A pragmatic way to perform a stability analysis is given by workflow STAT, see Section 5.1, where many parameters of the model as well as numerical settings are randomly varied in a tiny range each.

3.2 Parameter Sensitivity and Correlation Analysis

We call (the design of) a scenario *sensitive*, if small changes of the initial conditions lead to substantially different, still predictable results.

There are several ways to measure sensitivity and to perform sensitivity analysis. A simple method and quick check for nonlinear behaviour is based on a star-shaped experimental design (design-of-experiment (DoE), see Section 4.1). Per parameter, 3 values (left, center, right) for checking dependencies are available then. This simple analysis might be the first part of workflow ADAPTIVE (Section 5.3).

If a deeper analysis shall be performed directly, or if only one set of simulation runs is possible, either workflow STAT (Section 5.1) or workflow UNIFORM (Section 5.2) can be performed. Workflow UNIFORM has the advantage that a metamodel is constructed as well. Global impacts are reflected by correlation measures (see Section 4.2). Local sensitivities, 2D histograms and (approximations of) cumulative density functions give a deeper insight, see also Sections 6.1 and 6.1.

3.3 Robustness Analysis and Analysis of Critical Situations

We call (the design of) a scenario *robust* if small changes of the initial conditions will only have small and w.r.t. to the "quality" affordable impacts on the results. In particular, robustness analysis is a typical way to examine critical situations based on simulations.

One has to carefully distinguish between robustness and reliability. Roughly speaking, robustness aims at the behaviour for the majority of the cases (between the 5- and 95-percent-quantile or the 1- and 99-percent-quantile, say), whereas reliability asks for the seldom cases (outside the area considered for robustness analysis). In practice, reliability might be very difficult to compute accurately, whereas one can at least characterize robustness. Some robustness measures are (cf. [25, 27]):

- If a certain objective should be optimal in average, an expected value can be minimized (Attention: The distribution of the target function itself is allowed to have big outliers).
- If a certain objective should vary as less as possible, dispersion can be minimized. It is mandatory to combine this measure with one for controlling quality itself.

- If a certain objective must not fall below a given threshold, worst-case analysis and a respective measure can be used.
- If a given percentage of values of a target isn't allowed to fall below a threshold, a quantile measure can be applied.

Each measure is reasonable, there are more alternatives, and several measures can even be used simultaneously. Note that the decision for one or more measures depends on the intention of the designer.

3.4 Robust Design-Parameter Optimization (RDO)

RDO means parameter optimization with robustness aspects. One or more robustness criteria can be added to the optimization process. However, minimization of the value of a target function can lead to a higher dispersion and vice versa: usually, you cannot achieve both. Compromises might be found by a weighted objective function or the computation of Pareto fronts or a more substantial change of design.

3.5 History Matching

The adjustment of parameters of a model with respect to (historical) data stemming from physical measurements is quite often called *calibration* or *history matching*.

The goal is to ensure that predictions of future performance are consistent with historical measurements. History matching typically requires solving an ill-posed inverse problem, and thus, it is inherently non-unique. One can obtain a set of matching candidates by means of solving a multi-objective parameter-optimization problem.

Besides the parameters and their ranges, one or more optimization criteria have to be set up measuring the quality of the match. Often, differences of decisive properties such as pressures, fluxes, temperatures, etc., measured in, e.g., the L1- or L2-norm, are used.

4 Methods

In order to solve one of the analysis tasks discussed above, methods have to be selected and a workflow set up. Here, methods are outlined. In Section 5, several workflows as well as a flow chart supporting the selection of a workflow are discussed.

4.1 Experimental Designs

Experimental designs (design-of-experiment, DoE) considered here are based on some standard sampling schemes. Among them can be Monte Carlo (MC), Quasi Monte Carlo (QMC), Latin hypercube sampling (LHS), stratified sampling (SS), Centered stratified sampling (CSS). A detailed description of sampling schemes can be found in [19, 25], for instance.

A special DoE useful for a rough sensitivity analysis is the star-shaped DoE. It consists of $2n_p + 1$ experiments where n_p is the number of parameters. The central design plus, per parameter, a variation to a smaller as well as a larger value (typically with the same distance to the central point) is performed.

Note that the choice of the DoE is depending on the analysis task and the concrete step performed.

4.2 Correlation Measures

The Pearson correlation measure is an often-used, yet easily misleading one, because only monotonous correlations are captured (a typical example is depicted in Figure 1 (on the left)), and particularly for the case depicted in Figure 1 (on the right), it completely fails.

A measure reflecting nonlinearities is necessary as, for instance, the DesParO correlation measure, developed for RBF metamodels (see next section). In order

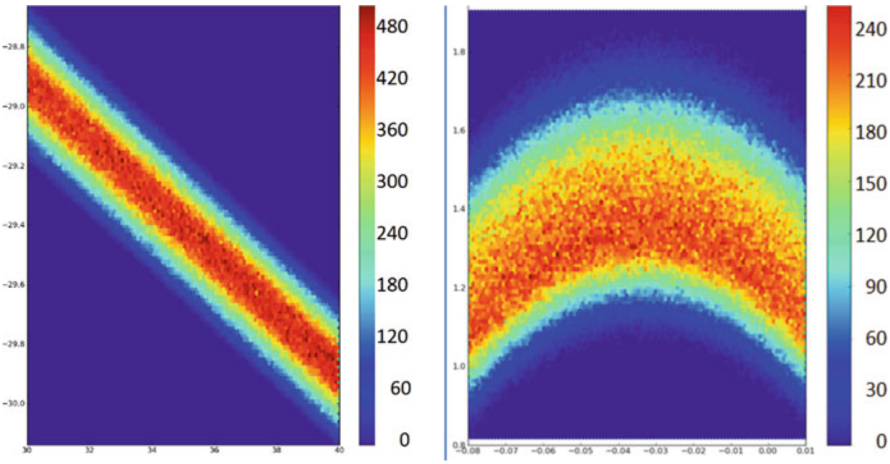

Fig. 1 Exemplary correlation plots. For the situation on the right, the Pearson measure will be 0 which is quite misleading. The color represents the number of samples in the respective hexagonal bin. Note that the actual range of values on the x- and y-axis are not relevant here

to roughly check which parameter-criteria dependencies are still linear or already nonlinear, both measures can be compared. This has been done exemplarily in Figs. 5, 8 and 15.

4.3 Metamodeling (Response Surfaces) and Adaptive Refinement

Classically, ensemble evaluations and Pareto optimizations rely on many experiments (simulation runs) - usually a costly procedure, even if the number of parameters involved is reduced beforehand.

For drastically reducing the number of simulation runs, one can set up fast-to-evaluate metamodels (response surfaces). This way, dependencies of objective functions on parameters are interpolated or approximated. Metamodels are quite often a good compromise for balancing the number of simulation runs (or physical measurements) to set up the model and a sufficient accuracy of approximation.

In the DesParO software [6], we use radial basis functions (RBF; e.g. multiquadrics, ANOVA), see [3], with polynomial detrending and an optional adjustment of smoothing and width.

We developed a measure for the *local tolerance* of the model w.r.t. leave-one-out cross-validation. It has some similarities to what can be done when working with Kriging. By means of this measure, interactive robustness analysis can be supported, in addition to quantile estimators. Analogously, we developed a nonlinear global correlation measure as well. Figs. 5, 8, 15 show examples of DesParO's metamodel explorer. Current tolerances are visualized by red bars below the current value of each criterion. Visualization of correlations is explained in Section 6.1.

As an orientation for the number of experiments n_{exp} which shall be used for constructing a basic RBF metamodel, one can use the following formula, assuming that the order of polynomial detrending is 2 and n_p denotes the number of parameters:

$$n_{exp} \geq C \left(2 + n_p + \frac{n_p(n_p + 1)}{2} \right) \tag{7}$$

C is an integer which can be set to 3, 4, or 5, say, to obtain a rough, small-sized, or medium-sized metamodel, respectively.

A standard measure for the quality of a metamodel is PRESS (predicted residual sums of squares). Originally, it stems from the statistical analysis of regression models, but can be used for other metamodels as well. If a local tolerance estimator is available, as for DesParO, quality can also be assessed locally.

More details and applications are discussed in [2, 8–12, 31], for instance.

A metamodel can be adaptively refined, in particular, if a local tolerance measure is available, as is the case in DesParO. We developed an extension and modification of the expected-improvement method (cf. [30] and references given therein to

Keane's method) for determining points (i.e. sets of parameters), the addition of which to the DoE is expected to improve interpolation. De facto, a hierarchical model results. See [7, 15], for instance.

4.4 *Quantiles and Robust Multi-Objective Optimization*

Several methods for computing quantiles and their applications are discussed in, e.g., [14, 17, 25–27, 29]. Methods for robust multi-objective optimization are presented and their practical applicability discussed in, e.g., [7, 15, 16, 22, 30].

5 Workflows

In the following, several workflows for tackling the analysis tasks summarized in Section 3. The specific task to be considered is denoted with **Analysis Task**.

The flow chart sketched in Figure 2 can be used as an orientation. In order to balance computational effort and accuracy while working with simulations, one should know, in addition, how fast a single simulation run is.

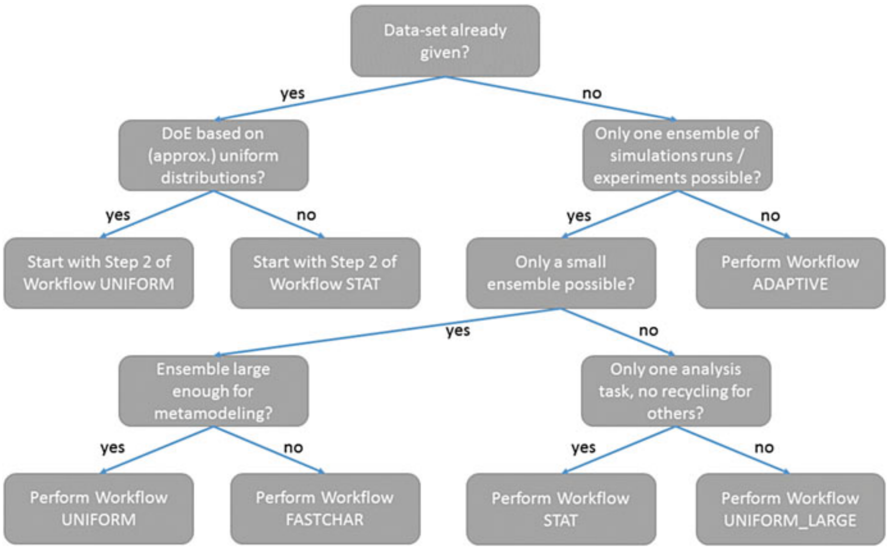

Fig. 2 Flow chart

5.1 *Workflow STAT*

A standard workflow for performing the **Analysis Task** directly based on simulation runs or experimental data is sketched below:

1. Ensemble set-up
   a. Determine the set of parameters
   b. For each parameter, determine its range of values and distribution according to the **Analysis Task**
   c. Set up a DoE according to the parameter ranges and distributions determined above; for determining the size of the DoE, find a balance between effort and quality
   d. Perform corresponding simulation runs / experiments
2. Perform the **Analysis Task** based on the ensemble

5.2 *Workflows UNIFORM and UNIFORM-LARGE*

A standard workflow for performing the **Analysis Task** employing a metamodel is sketched below:

1. Ensemble setup
   a. Determine the set of parameters
   b. For each parameter, determine its range of values
   c. Set up uniform DoE; for determining the size of the DoE, find a balance between effort and quality; in case of UNIFORM, find orientation in Eq. 7; in case of UNIFORM-LARGE, estimate the number of experiments necessary for a classical QMC method, say
   d. Perform corresponding simulation runs / experiments
2. Metamodel and quality assessment
   a. Set up a metamodel using the ensemble created above
   b. Check model tolerances (*global PRESS value*, local tolerances)
   c. Check *correlation measures*
   d. Reduce parameter space for analysis task, as far as possible
3. If metamodel ok: Perform the **Analysis Task** employing the metamodel

5.3 *Workflow ADAPTIVE*

An iterative workflow for adaptive hierarchical metamodeling and optimization of decisive metamodeling parameters is sketched now:

1. (Optional:) Set up a star-shaped DoE for performing a basic sensitivity analysis
2. (Optional:) Based on its results, reduce the parameter space
3. Perform workflow UNIFORM
4. This includes the first run of the **Analysis Task**. In case of RDO, a rough Pareto optimization for finding candidate regions should be performed
5. If necessary, perform *model refinement* (cf. Section 4.3), then go to step 4
6. Perform the final run of the **Analysis Task** employing the metamodel

6 Visualizations

Besides the methods for approximating dependencies and statistical measures, visualization techniques play a decisive role. Without appropriate representation of results, the sometimes immense output cannot be digested and efficiently interpreted. Visualization can efficiently support, for instance, pointing to interesting features of a problem and interactive exploration of parameter-criteria dependencies. Some important techniques are summarized in the following.

6.1 Correlations

Global correlation measures can be visualized by means of tables with boxes. The magnitude of the box represents the magnitude of the absolute correlation value, its color the direction of correlation, e.g., blue for monotonously decreasing, red for monotonously increasing, black for nonmonotonous behaviour.

Examples can be found in Figure 5 (on the right), for instance.

Correlations can directly be visualized by means of two-dimensional scatter plots of all pairs of values involved. However, especially if larger areas of the two-dimensional space are filled this way, a good alternative are two-dimensional histograms (see the next section).

6.2 Histograms and Alternatives

Classical techniques to visualize distributions are

- (one-dimensional) histograms: an example can be found in Figure 14 (on the left)
- approximate CDF curves (CDF: cumulative density function) an example can be found in Figure 10 (on the bottom)
- boxplots

In addition to a histogram, a *plot of sorted values* is often of help. To create one, all values of interest have to be sorted first, decreasing or increasing by (absolute) value. All values (or selected ranges only) of the resulting vector v are plotted then, i.e., all data points $(i,)v(i)$ are plotted. An example can be found in Figure 14 (on the right).

2D histograms (hexbins) are a good alternative to scatter plots if enough data points are available. Examples can be found in Figs. 1, 7 and 12, for instance.

6.3 2D Network Representations

Colors and thicknesses of nodes and lines can be chosen differently representing values of different inputs or outputs. A classical version is shown in Figure 13. Pressure values are used for coloring the nodes, averaged pressure values for coloring the edges, and pipe widths for determining the thickness of the edges. A 2D network representation is also a good choice for showing differences of values for two scenarios of an ensemble. In order to find areas with large differences, node and/or edge sizes should dependent on the local difference of the output function considered. Typical applications are the comparison of different physical laws, parameter variations (extreme cases), or local differences between the maximal and minimal value in the ensemble for the output function considered.

Manipulating the coordinates is another possibility. Quite often, the coordinates do not reflect the geometrical situation but is a compromise of real locations and a schematic view. Important areas can be given more space this way. Alternatively, algorithms might be used which perform mappings of coordinates in order to highlight areas with, for instance, large mass flows.

6.4 3D Network Representations

A classical setting is to use (x, y, h) for 3D plots. This setting allows for debugging of height data - drastic drop downs to zero, for instance, might indicate missing or wrongly specified height values.

For analysing function profiles, one might use, for instance, pressure or temperature as z-coordinate. Figure 3 provides an example for a realistic pressure profile. Analogously, means, medians, quantiles or differences between minimum or maximum values or two selected quantiles can be visualized.

7 Examples from Gas Transport

Applications of the methods discussed above are illustrated by means of two examples from long-distance gas transport, namely a compressor station and a mid-sized pipeline network with several supplies and regulators, for instance.

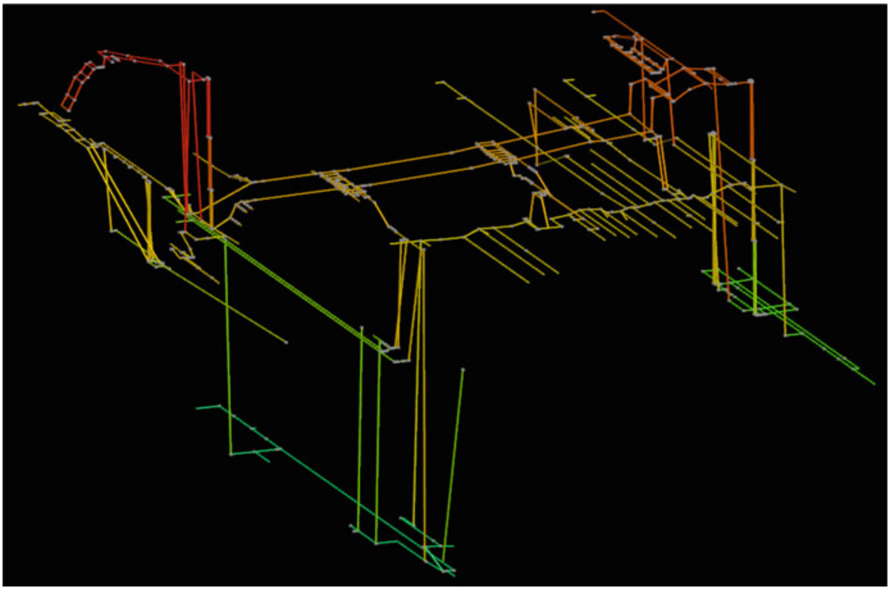

Fig. 3 3D plot (MYNTS' interactive OpenGL-based 3D viewer) for a decisive part of a large realistic network - pressure profile: color and z-coordinate are based on pressure values, clearly showing the different pressure levels and transition areas (due to compressor stations, for example)

7.1 Example 1 - Compressor Station

The first example, see Figure 4, is a simple compressor station consisting of two machine units with a compressor, a drive, a cooler, a regulator, two resistors and a master each, as well as a master and several valves and pipes for controlling the station and switching tracks. One supply and one demand node are attached.

The parameters and criteria investigated overall are listed in Table 2. Two scenarios are analysed. Their settings and results are described and discussed in the following sections.

7.1.1 Scenario 1

In Scenario 1, only the first compressor is up and running, and only QSET is varied.

Workflow UNIFORM is used. The DesParO metamodel resulting from a standard uniform DoE with 50 experiments (after automatically excluding 9 parameter values which are very close to others) is shown in Figure 5. 50 experiments are not really necessary here, given that only one parameter is varied. The metamodel would react more or less identically if only 10 experiments are used, say. However, the same ensemble can be used to create a model, evaluate it randomly and plot 1D and 2D histograms, see Figure 6, as well as to plot finely resolved curves for parameter-criteria dependencies directly, see Figure 7.

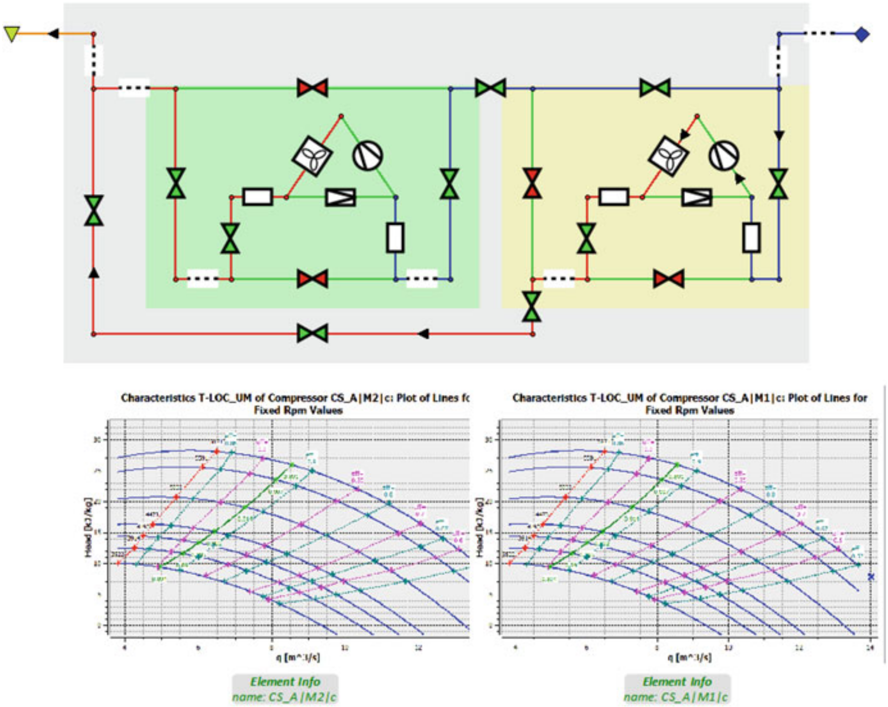

Fig. 4 Example 1: (*top*) schematics of the network (*diamond* on the right marks the supply, *triangle* on the left the demand node, other elements as explained in the text); (*bottom*) characteristics maps of the two compressors

Table 2 Example 1: parameters and criteria

property	element	type	description
PSET	junc0	parameter	input pressure at supply
QSET	junc1	parameter	output volume flow at demand
Had	CS_A|M1|c	criterion	head of compressor 1
QVOL	CS_A|M1|c	criterion	volume flow through compressor 1
tslope	CS_A|M1|c	criterion	tslope of compressor 1
pslope	CS_A|M1|c	criterion	pslope of compressor 1
m	CS_A|M1|c	criterion	mass flow through compressor 1
pslope	junc0^locE	criterion	pslope at pipe beginning at the supply
tslope	junc0^locE	criterion	tslope at pipe beginning at the supply
m	junc0^locE	criterion	mass flow at pipe beginning at the supply
T	junc1	criterion	temperature at demand
P	junc1	criterion	pressure at demand
m	locA^junc1	criterion	mass flow at pipe ending at the demand
tslope	locA^junc1	criterion	tslope at pipe ending at the demand
PSLOPE	locA^junc1	criterion	pslope at pipe ending at the demand

Fig. 5 Example 1, Scenario 1: metamodel and correlations for the parameters and criteria listed in Table 2. Criteria from top to bottom: tslope, Had, QVOL, pslope, m of the compressor, pslope, tslope and m of junc0^locE, T and P of junc1, m, tslope and pslope of locA^junc1

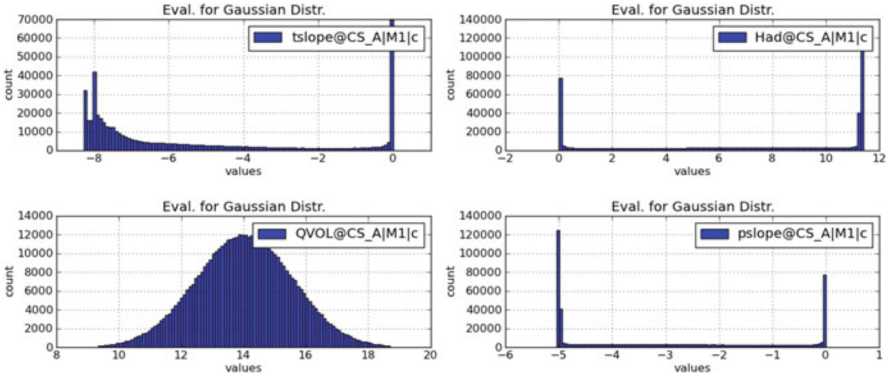

Fig. 6 Example 1, Scenario 1: histograms (Gaussian distribution for QSET, one million evaluations) of the first 4 criteria. x-axis: values (clockwise: approx. [-8.5;0.5], [-2;12], [-6;1], [8;20]), y-axis: counts (clockwise: up to 70,000, 140,000, 140,000, 14,000)

Figure 5 shows both Pearson and DesParO correlation results. The magnitude of the correlation values is very similar among the parameters. However, several correlations are nonlinear (black box in the DesParO correlation plot), and Pearson indicates a large monotonous correlation instead. Especially, Had@CS_A|M1|c reacts in a strongly nonlinear fashion to changes of QSET, see also Figs. 6 and Figure 7. This criterion is decisive for describing the behaviour of the compressor. Hence, the nonlinearity cannot be neglected, and a linear model and Pearson correlation are not sufficient in this small and quite simple test case involving one compressor only.

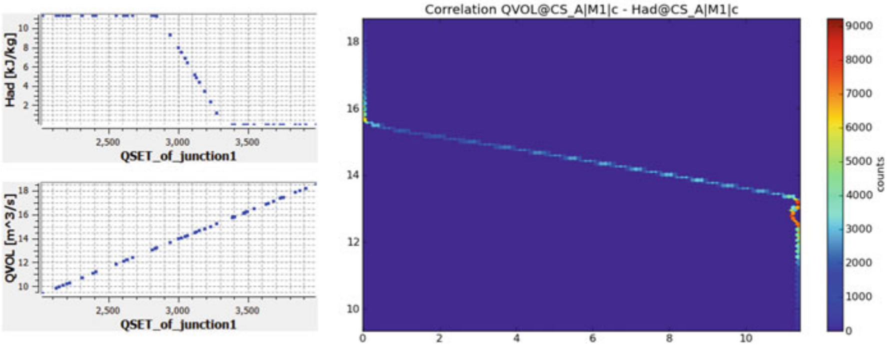

Fig. 7 Example 1, Scenario 1: exemplary ensemble curves for HAD of the compressor, and 2D histogram (Gaussian distribution for QSET) for QVOL vs. HAD. x-axis: [2,000;4,000], y-axis: HAD [0;12], QVOL [10.5;18.5]. Counts (color) on the right: [0;9,000]

7.1.2 Scenario 2

In Scenario 2 both compressors are up and running in parallel, and both parameters (PSET and QSET) are varied: $[50; 60]$ for PSET, and $[2,000; 10,000]$ for QSET. We already learned from analysing Scenario 1 that several parameter-criteria dependencies are expected to be strongly nonlinear, depending on the concrete range of variations.

Again, workflow UNIFORM is used, and, since we have only two parameters here, a standard full-factorial DoE with $7^2 = 49$ experiments is chosen as a basis for direct analysis as well as creation of a DesParO metamodel. Ensemble curves, i.e., raw-data plots of parameter-criteria dependencies, are shown in Figure 11.

Indeed, Scenario 2 shows several interesting effects. The ranges are de facto chosen here so that the compressors cannot completely fulfill their task of creating an output pressure of 60 bars. Fig. 11 (top-right plot) clearly shows that 60 bars cannot be reached for most combinations. Analogously, Figure 11 (bottom-left plot) shows that the compressor goes to de facto bypass mode (zero head) for QSETs above approximately 7,000.

The metamodel for the ensemble is shown in Figure 8. The model has a reasonable quality (see PRESS values), however, DesParO's tolerance measure cannot be neglected here. Based on the parameter distributions depicted in Figure 9, the metamodel is evaluated and 1D and 2D histograms for several exemplary interesting dependencies shown in Figs. 10 and 12. Note that skewed Gaussian distributions are used, in case of QSET only for a part, namely $[2,000; 7,000]$, of the range covered by the metamodel (Figure 8).

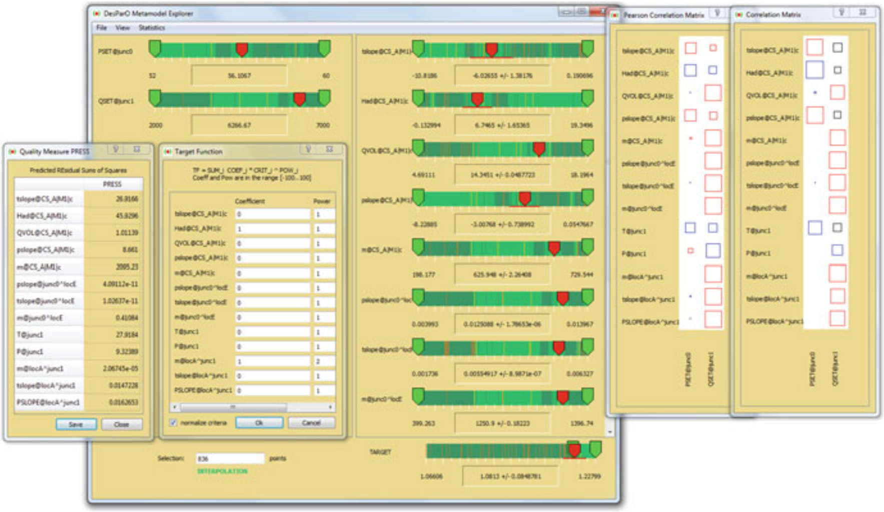

Fig. 8 Example 1, Scenario 2: metamodel, correlations, PRESS values, target function

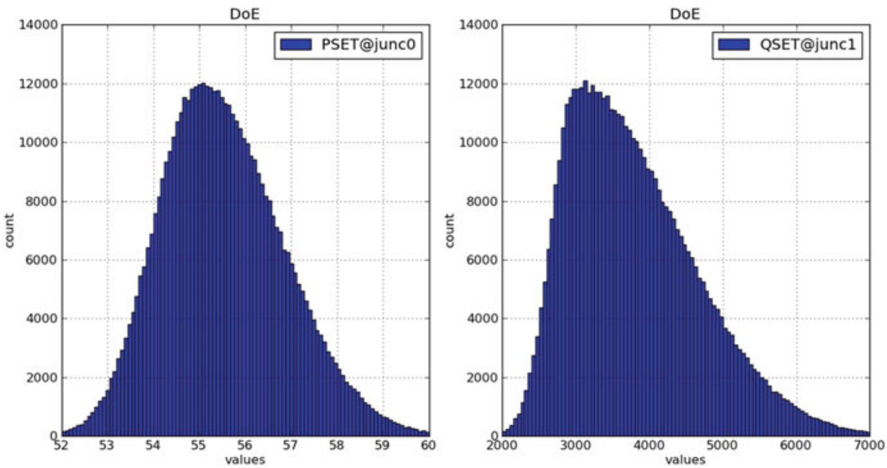

Fig. 9 Example 1, Scenario 2: histograms of the skewed Gaussian distributions of parameters PSET [52;60] and QSET [2,000;7,000]. Counts (y-axis): [0;14,000]

Looking at Had@CS_A|M1|c and QVOL@CS_A|M1|c, one can study effects of nonlinear, linear or very weak dependencies on variations of PSET and QSET here. QVOL has to react linearly on variations of QSET, as the 1D histogram in Figure 10 as well as the 2D histogram in Figure 12 show. As long as the compressors are able to completely fulfill their common task (PSET large enough), QVOL does weakly react on PSET. For smaller PSETs and large QSETs, the compressors go to their limit, and the distribution of QVOL is wide.

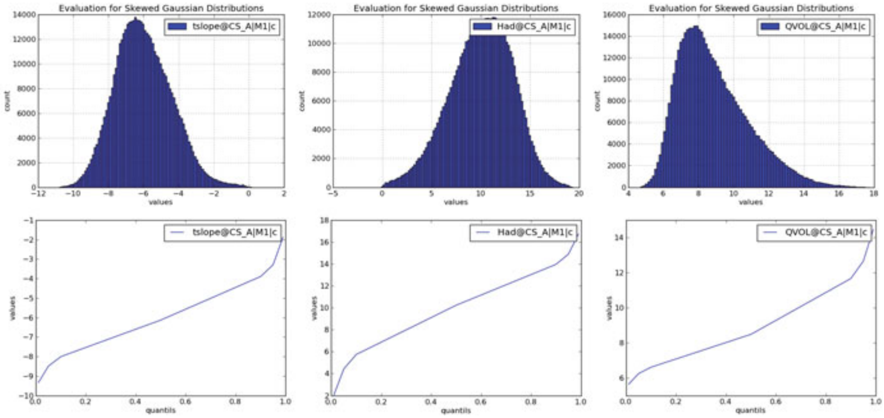

Fig. 10 Example 1, Scenario 2: histograms and CDFs (for distributions shown in Figure 9) of the first 3 criteria. Figures on top: count [0;14,000] vs. TSLOPE [-12;2], count [0;12,000] vs. HAD [-5;20], count [0;16,000] vs. QVOL [-4;18]. Quantile plots at the bottom: x-axis [0;1], y-axis: TSLOPE [-10;1], HAD [2;18], QVOL [5.5;14.5]

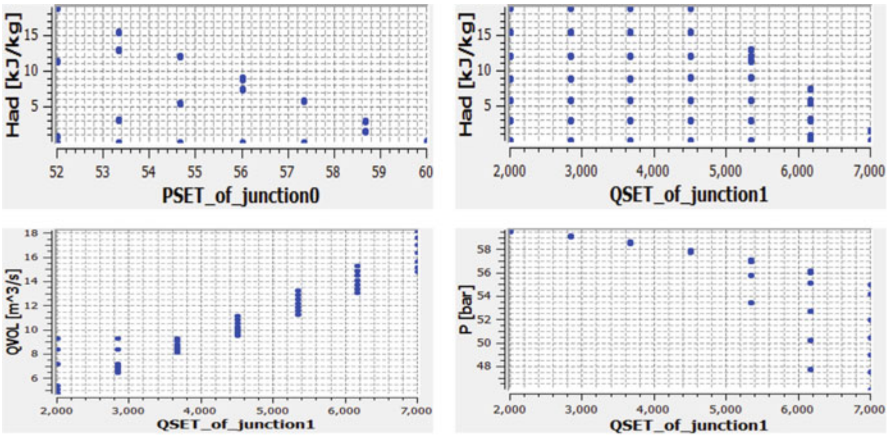

Fig. 11 Example 1, Scenario 2: exemplary ensemble curves for HAD [0;20] vs. PSET [52;60] and QSET [2,000;7,000] as well as QVOL [5;18] and P@junction1 [46;59.5] vs. QSET [2,000;7,000]

Based on the metamodel, RDO tasks can be set up and solved. Figure 8 shows such a task and its visual exploration. Here, the following target is set:

$$\max \text{Had@CS\_A|M1|c} + (\text{m@locA\^{}junc1})^2 \tag{8}$$

One could also use, for instance, QVOL@CS_A|M1|c instead of m@locA^junc1. By visual inspection, one can see that the parameter space separates into two parts: one is with mid-sized PSET and QSET, one with small PSET and large QSET. DesParO's tolerance measure gives a first indication of robustness of results.

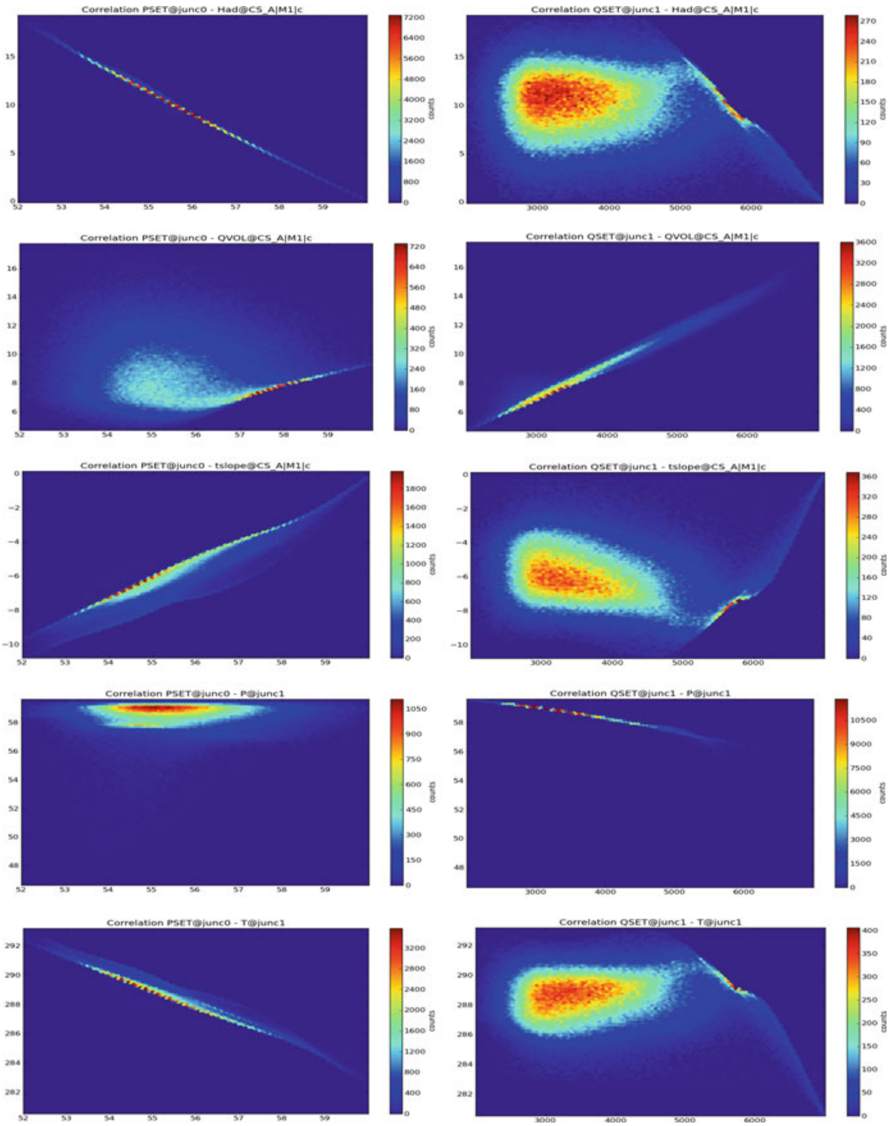

Fig. 12 Example 1, Scenario 2: exemplary 2D correlation plots for exemplary criteria vs. PSET (on the left) or QSET (on the right). Ranges for values analogously as before, counts up to (left) 7200, 720, 2000, 1100, 3600, (right) 270, 3600, 360, 12000, 400

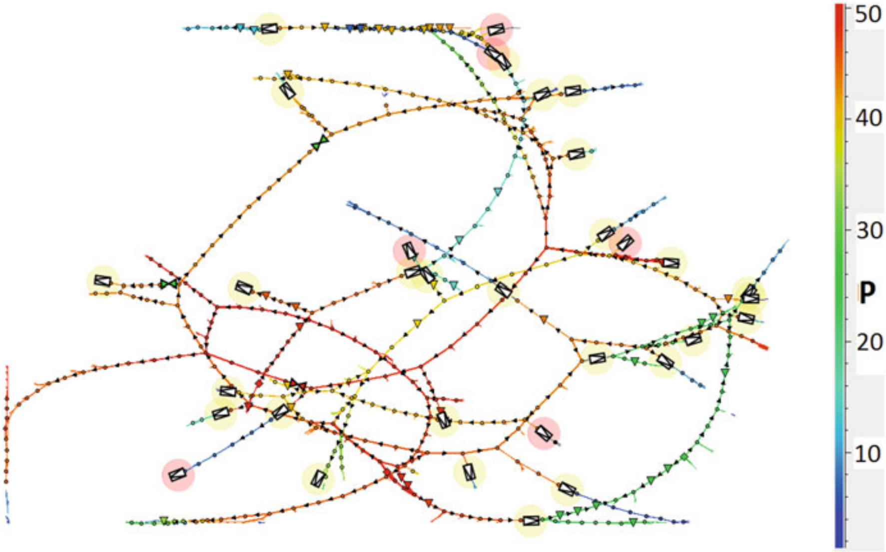

Fig. 13 Example 2: net with pressure results of the basic scenario considered

Table 3 Example 2: elements (table on the left), parameters and criteria (table on the right)

element type	number
pipes	907
regulators	45
heaters	45
valves	52
resistors	89
important supplies	5
important demands	5
remaining nodes	1069

element	parameter	min	max	distributions DoE; analysis
(all)	tsoil	267.15	287.15	uniform; Gaussian
out1	QSET	20	30	uniform; Gaussian
out2	QSET	50	60	uniform; Gaussian
out3	QSET	30	40	uniform; Gaussian
out4	QSET	20	30	uniform; Gaussian
out5	QSET	50	60	uniform; Gaussian
in1...in5	m	–	–	(criteria)

7.2 Example 2

The second example, see Figure 13, is a mid-sized network consisting of the elements mentioned in Table 3 (on the left). A typical distribution of pressures resulting from the simulation of a typical scenario is shown in Figure 14. In contrast to Example 1, this network does not contain compressors. The task is here to determine the influence of variations of the 5 largest demands as well as the soil temperature on the network, see Table 3 (on the right).

Again, workflow UNIFORM is used, and a standard uniform DoE with 50 experiments is chosen as a basis for direct analysis as well as creation of a DesParO metamodel. A thin full-factorial DoE for checking nonlinearities would already have $6^3 = 216$ experiments. As can be seen from Figure 15, the criteria depend

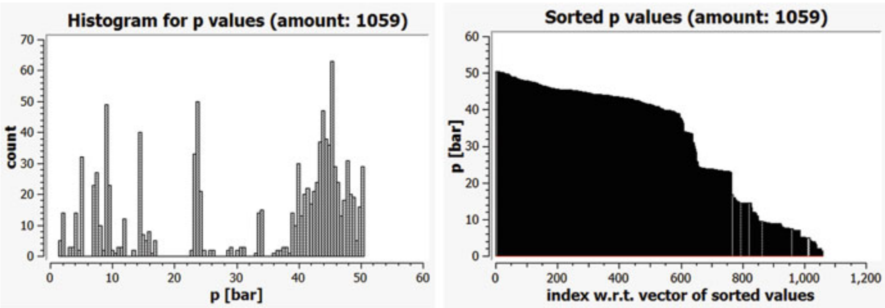

Fig. 14 Example 2: exemplary histogram (left: count [0;70] vs. pressure [0;60]) and plot of sorted values (right) for pressure (P) (y-axis: pressure [0;60], x-axis: index [0;1,200])

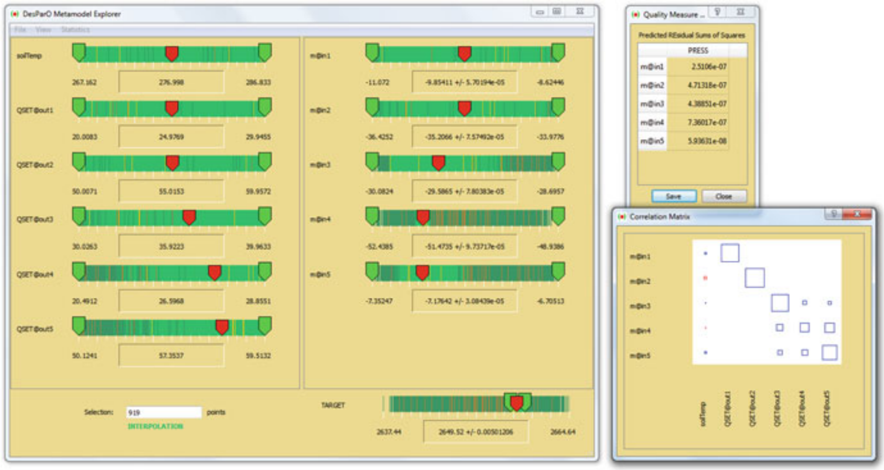

Fig. 15 Example 2: metamodel exploration, PRESS quality measure and correlation plot for the parameters and criteria listed in Table 3

monotonously from the parameters. They are de facto quite linear: the Pearson and DesParO correlation measures are more or less identical, the PRESS values for the quality of the metamodel tiny, of course. As can be seen from the correlation plots, the soil temperature does not play a decisive role. The interplay of the different supply demand combinations reveals a partition into three parts: two one-to-one constellations (out1-in1 and out2-in2), one three-to-three constellation. Figure 16 shows exemplary combinations of strongly dependent parameter-criteria combinations.

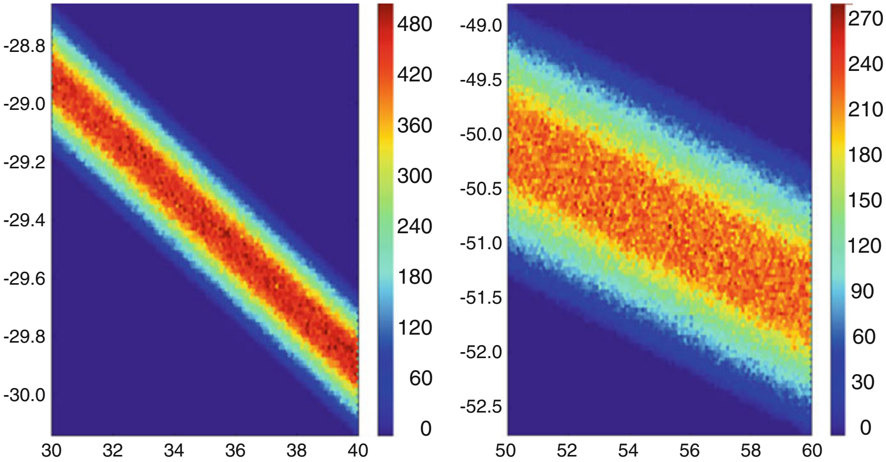

Fig. 16 Example 2: 2D histograms: (left) m@in3 [-30.2;-28.6] vs. QSET@out3 [30;40] which has a decisive impact on m@in3, (right) m@in4 [-52.7;-48.8] vs. QSET@out5 [50;60] which has an impact but shares this with two other parameters

In such a situation, maybe a cheaper way to proceed would be to use workflow ADAPTIVE. We could start with a simple star-style DoE ($2 * 6 + 1 = 13$ experiments) in order to check the linearity of the dependencies. The set of parameters could be reduced by removing soil temperature. In addition, the three separate constellations mentioned above could be analysed separately.

Based on a metamodel, one or more optimization tasks can be set up and solved, for instance, in order to fulfill certain contractual conditions. For illustration, a very simple optimization target is set up (here: m@in4 shall meet a certain value). Figure 15 shows how a visual exploration already reveals influences on the parameter combinations.

8 Conclusions and Outlook

Several methods for approximating parameter-criteria dependencies and determining statistical quantities, corresponding workflows and versatile visualization methods for analysing parameter variations in energy networks have been described and discussed. A common physical model for flow in gas transport networks has been described. Two exemplary gas networks with some typical effects have been studied in some detail using the methods and workflows discussed.

From examining the examples, one can see that nonlinearities might play a decisive role, especially in networks with compressor stations. Numerical methods for measuring correlations and approximating parameter-criteria dependencies have to be chosen which are able to reflect nonlinearities. Regression-based linear methods can sometimes support the analysis though - a comparison of the values provided by Pearson's correlation measure with the ones provided by DesParO's measure indicates nonlinearities in an intuitive fashion.

The challenge of transforming our energy production by means of incorporating increasingly larger amounts of renewable energy sources can only succeed if also our energy networks are transformed to build up an integrated system allowing for balancing of supplies and demands with the use of energy conversion and storages provided by pipeline systems and caverns, to give just some examples. This motivates the author and her colleagues a lot. They will continue their work on physical modeling of energy networks, their efficient simulation, statistical analysis and optimization.

References

1. Baumanns, S., Cassirer, K., Clees, T., Klaassen, B., Nikitin, I., Nikitina, L., Tischendorf, C.: MYNTS User's Manual, Release 1.3. Fraunhofer SCAI, Sankt Augustin, Germany (2012). www.scai.fraunhofer.de/mynts
2. Borsotto, D., Clees, T., Nikitin, I., Nikitina, L., Steffes-lai, D., Thole, C.A.: Sensitivity and robustness aspects in focused ultrasonic therapy simulation. In: EngOpt 2012 – 3rd International Conference on Engineering Optimization. Rio de Janeiro, Brazil (2012)
3. Buhmann, M.: Radial Basis Functions: Theory and Implementations. Cambridge University Press, Cambridge (2003)
4. Cassirer, K., Clees, T., Klaassen, B., Nikitin, I., Nikitina, L.: MYNTS User's Manual, Release 2.9. Fraunhofer SCAI, Sankt Augustin (2015). www.scai.fraunhofer.de/mynts
5. Clees, T.: MYNTS – Ein neuer multiphysikalischer Simulator für Gas, Wasser und elektrische Netze. Energie — Wasser-Praxis **09**, 174–175 (2012)
6. Clees, T., Hornung, N., Nikitin, I., Nikitina, L., Pott, S., Steffes-lai, D.: DesParO User's Manual, Release 2.2. Fraunhofer SCAI, Sankt Augustin, Germany (2012). www.scai.fraunhofer.de/desparo
7. Clees, T., Hornung, N., Oyerinde, A., Stern, D.: An adaptive hierarchical metamodeling approach for history matching of reservoir simulation models. In: SPE/SIAM Conference on Mathematical Methods in Fluid Dynamics and Simulation of Giant Oil and Gas Reservoirs (LSRS). Istanbul, Turkey (2012). Invited presentation (T. Clees)
8. Clees, T., Nikitin, I., Nikitina, L.: Nonlinear metamodeling of bulky data and applications in automotive design. In: Günther, M., et al. (eds.) Progress in industrial mathematics at ECMI 2010. Mathematics in Industry, vol. 17, pp. 295–301. Springer, Berlin (2012)
9. Clees, T., Nikitin, I., Nikitina, L., Kopmann, R.: Reliability analysis of river bed simulation models. In: Herskovits, J. (ed.) CDROM Proceedings of the EngOpt 2012, 3rd International Conference on Engineering Optimization, no. 267. Rio de Janeiro, Brazil (2012)
10. Clees, T., Nikitin, I., Nikitina, L., Thole, C.A.: Nonlinear metamodeling and robust optimization in automotive design. In: Proceedings of the 1st International Conference on Simulation and Modeling Methodologies, Technologies and Applications SIMULTECH 2011, pp. 483–491. SciTePress, Noordwijkerhout, The Netherlands (2011)
11. Clees, T., Nikitin, I., Nikitina, L., Thole, C.A.: Analysis of bulky crash simulation results: deterministic and stochastic aspects. In: Pina, N., et al. (eds.) Simulation and Modeling Methodologies, Technologies and Applications, AISC 197. Lecture Notes in Advances in Intelligent and Soft Computing, pp. 225–237. Springer, Berlin, Heidelberg (2012)

12. Clees, T., Steffes-lai, D., Helbig, M., Sun, D.Z.: Statistical analysis and robust optimization of forming processes and forming-to-crash process chains. Int. J. Mater. Form. **3**, 45–48 (2010). Supplement 1; 13th ESAFORM Conference on Material Forming. Brescia, Italy (2010)
13. Grundel, S., Hornung, N., Klaassen, B., Benner, P., Clees, T.: Computing surrogates for gas network simulation using model order reduction. In: Koziel, S., Leifsson, L. (eds.) Surrogate-Based Modeling and Optimization, pp. 189–212. Springer, New York (2013)
14. Harrell, F.E., Davis, C.E.: A new distribution-free quantile estimator. Biometrika **69**, 635–640 (1982)
15. Hornung, N., Nikitina, L., Clees, T.: Multi-objective optimization using surrogate functions. In: Proceedings of the 2nd International Conference on Engineering Optimization (EngOpt). Lisbon, Portugal (2010)
16. Jones, D., Schonlau, M., Welch, W.: Efficient global optimization of expensive black-box functions. J. Glob. Optim. **13**(4), 455–492 (1998)
17. Jones, M.C.: The performance of kernel density functions in kernel distribution function estimation. Stat. Probab. Lett. **9**(2), 129–132 (1990)
18. Klaassen, B., Clees, T., Tischendorf, C., Soto, M.S., Baumanns, S.: Fully coupled circuit and device simulation with exploitation of algebraic multigrid linear solvers. In: Proceedings of the Equipment Data Acquisition Workshop. Dresden (2011)
19. Kleijnen, J.: Design and Analysis of Simulation Experiments. Springer, New York (2008)
20. Lorenz, J., Bär, E., Clees, T., Evanschitzky, P., Jancke, R., Kampen, C., Paschen, U., Salzig, C., Selberherr, S.: Hierarchical simulation of process variations and their impact on circuits and systems: results. IEEE Trans. Electron Devices **58**(8), 2227–2234 (2011)
21. Lorenz, J., Clees, T., Jancke, R., Paschen, U., Salzig, C., Selberherr, S.: Hierarchical simulation of process variations and their impact on circuits and systems: methodology. IEEE Trans. Electron Devices **58**(8), 2218–2226 (2011)
22. Maass, A., Clees, T., Nikitina, L., Kirschner, K., Reith, D.: Multi-objective optimization on basis of random models for ethylene oxide. Mol. Simul. Special Issue: FOMMS 2009 Conference Proceedings, vol. **36**(15), pp. 1208–1218(11) (December 2010)
23. Maric, I., Ivek, I.: Natural gas properties and flow computation. In: Potocnik, P. (ed.) Natural Gas. InTech (2010). ISBN: 978-953-307-112-1. doi: 10.5772/9871. Available from: http://www.intechopen.com/books/natural-gas/natural-gas-properties-and-flow-computation
24. Mischner, J., Fasold, H.G., Kadner, K.: gas2energy.net - Systemplanerische Grundlagen der Gasversorgung. Div Deutscher Industrieverlag München (2011). ISBN 978-3835632059
25. Rhein, B., Clees, T., Ruschitzka, M.: Robustness measures and numerical approximation of the cumulative density function of response surfaces. Commun. Stat. Simul. Comput. **43**(1), 1–17 (2014)
26. Rhein, B., Clees, T., Ruschitzka, M.: Uncertainty quantification using nonparametric quantile estimation and metamodeling. In: Eberhardsteiner, J., et.al. (eds.) European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2012). Vienna, Austria (2012)
27. Rhein, B., Ruschitzka, M., Clees, T.: A simulation framework for robust optimization based on metamodels. In: Proceedings of NAFEMS World Congress 2013, International Conference on Simulation Process and Data Management, Salzburg, 9–12 June 2013
28. Schöps, S., Bartel, A., Günther, M., ter Maten, E.J.W., Müller, P.C. (eds.): Progress in Differential-Algebraic Equations, Differential-Algebraic Equations Forum. Proceedings of Descriptor 2013, pp. 183–205. Springer, Berlin, Heidelberg (2014)
29. Sfakianakis, M.E., Verginis, D.G.: A new family of nonparametric quantile estimators. Commun. Stat. Simul. Comput. **37**, 337–345 (2008)
30. Sobester, A., Leary, S., Keane, A.: On the design of optimization strategies based on global response surface approximation models. J. Glob. Optim. **33**(1), 31–59 (2005)
31. Steffes-lai, D., Clees, T.: Statistical analysis of forming processes as a first step in a process-chain analysis: novel PRO-CHAIN components. Key Engineering Materials (KEM) **504–506**, 631–636 (2012). Special Issue Proceedings of the 15th ESAFORM Conference on Material Forming. Erlangen, Germany (2012)