

Empirical Models for Utility Disturbances in the Process Industry

Anna Lindholm * Krister Forsman **

* *Department of Automatic Control, Lund University, Box 118,
SE-221 00, Lund, Sweden (e-mail: anna.lindholm@control.lth.se).*

** *Perstorp AB, SE-284 80 Perstorp, Sweden*

Abstract: In this paper, a general method for identifying common disturbances in the supply of utilities to a process industrial site is presented, with focus on the chemical process industry. The method aims at finding typical utility disturbance trajectories and define them by simple measures, such as duration and time between failures. Statistical models are suggested for each of the measures. The estimated disturbance trajectories may be used as inputs to optimization problems for determining the optimal supply of utilities to each area of a site during a disturbance. Use of such optimization procedures can improve both proactive and reactive disturbance management for utility disturbances. Industrial data for some utilities, among others steam and cooling water, have been used to retrieve disturbance models for utilities at a specific site. These models may also give some information on the general characteristics of utility disturbances.

Keywords: Chemical industry, Process control, Models, Disturbance rejection, Statistical analysis, Utility functions

1. INTRODUCTION

Utilities such as steam and cooling water are commonly used at process industrial sites, and the utility costs often represent a large part of the total operating cost (Iyer and Grossmann (1998)). Disturbances in the supply of utilities may also cause the production to slow down or stop, which implies great loss of revenue. Since utilities are often shared between production areas, the problem becomes even more intricate. At the occurrence of a disturbance, critical information about the disturbance is often unknown (Lee and Weekman (1976)). The survey of utility disturbances that is presented in this paper aims to find common disturbance characteristics for each utility, which may give operators at a site some insight in what could be likely scenarios when a disturbance occurs. The objective is to present a general method for finding typical disturbances for each utility, given historical measurement data. The method is by no means the single approach to obtain these typical disturbance trajectories, and each of the steps of the method could be developed and augmented in order to ensure more accurate results for all cases. However, we do believe that there is a need for such a method, since little work seems to have been done on statistical modeling of utility disturbance characteristics within the chemical process industry. A related study for the pulp and paper industry has been performed in Khanbaghi et al. (1997), where statistical analysis of paper break data from a typical paper mill is used to form a Markov chain model for the paper-break process. Modeling the process as a continuous Markov chain implies that the failure rate is constant, due to the memoryless character of the exponential distribution. In Ogawa (2003), the nature of the input disturbances to a broke storage tank at a

pulp and paper production site are investigated, and it is found that break and normal durations can be modeled by exponential distributions. However, the exponential distribution with its constant failure rate does not seem to fit as well for utility disturbances. This is further discussed in section 2.3 of this paper.

The method proposed in this paper is applied to industrial data from Perstorp, a world leading company within several sectors of the specialty chemicals market (Perstorp (2013)). The disturbance characteristics for the utilities at Perstorp may be useful also for other companies, since the utilities used at the Perstorp site are very common also at other industrial sites. The estimated disturbance trajectories can be used as input to optimization problems for finding the optimal supply of a utility to different production areas at the occurrence of a disturbance. A method for formulating such optimization problems is presented in Lindholm and Giselsson (2013). The results from the optimization may be used by process operators to improve disturbance management for utility disturbances.

The objective of the method presented in this paper is to find statistical models for common disturbances in utilities. As a consequence, unlikely but possibly serious disturbances are disregarded. For that topic, we refer to literature on risk management, e.g. Greenberg (1991).

2. A GENERAL METHOD FOR FINDING EMPIRICAL UTILITY DISTURBANCE MODELS

A utility disturbance is according to Lindholm and Johnsson (2013) defined as when the measurement of a utility parameter, such as temperature or pressure of the utility,

goes outside the limits for normal operation¹. This definition enables to detect disturbances directly from measurement data, which gives the possibility to estimate typical utility disturbance trajectories for each utility. These trajectories may be used to evaluate how such disturbances should be handled, e.g. by the approach introduced in Lindholm and Giselsson (2013). A method to find these typical disturbance trajectories from measurement data is presented here. The method consists of three steps:

- (1) Prefilter the utility measurement data
- (2) Identify type(s) of disturbances for the utility
- (3) Find one or more typical disturbance(s) for the utility

Step 3 is divided into different cases based on which type(s) of disturbances that were identified in step 2.

2.1 Step 1: Prefilter the Utility Measurement Data

The utility measurement data could be measurements of e.g. pressure, flow, or temperature of the utility. Since the definition of utility disturbances in Lindholm and Johnsson (2013) is used, the detection of a disturbance is entirely defined by the disturbance limit that is set for the utility. Special care should be taken to two cases:

Very short disturbances To avoid the detection of many short disturbances when a measurement oscillates around the disturbance limit, the data may have to be prefiltered before continuing to the second and third step of the method. The situation is illustrated in Fig. 1, where measurements of a utility parameter are given for approximately two hours of operation, during which one major utility disturbance occurred. As seen in the figure, many consecutive disturbance limit crossings give very short disturbance durations and time between failures. The disturbance limit is marked with a red dashed line in the figure. A simple filtering strategy is to simply merge disturbances where the time between failures is shorter than some limit, e.g., 10 minutes. For the example in Fig. 1, this gives the result shown in Fig. 2. Another option is to use a *spike-pass filter*, which is a nonlinear filter using change detection to identify abrupt changes. Between abrupt changes, it is a simple low-pass filter. In this way, high frequency noise is reduced, while capturing sudden changes in the signal. The principle is described in Gustafsson (2000). Spike-pass filtering gives a filtered signal that is similar to the resulting signal when the simple approach described above is used. Simple low-pass filtering can also be an alternative for the data prefiltering. The parameters of the filter should, independently of the choice of filter, be tuned in such a way that the result of the filtering agrees with the intuitive notion of what is high frequency noise, and what is actual disturbances. There is obviously no general choice of filtering parameters that works for any type of data, from any plant.

Very long disturbances Another filtering of the data that may have to be performed before continuing to steps 2 and 3 is to remove disturbances of very long durations, since these disturbances typically correspond to unlikely situations. One way to identify if these disturbances are very unlikely to occur is to look at the percentage of the

¹ Hereafter denoted *disturbance limits*.

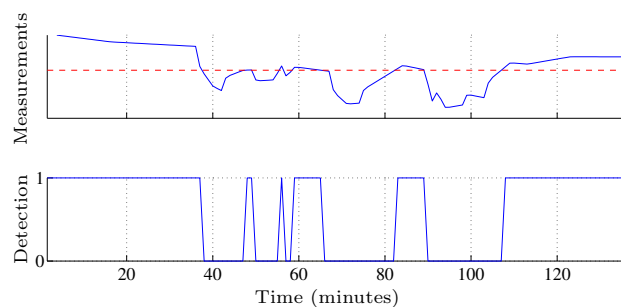


Fig. 1. Measurement data and disturbance detection.

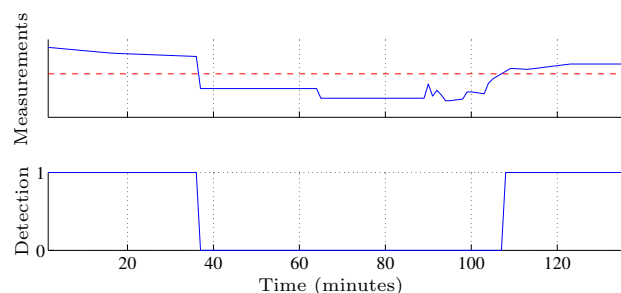


Fig. 2. Measurement data and disturbance detection after filtering.

number of disturbances that are of shorter duration than a threshold. This is visualized in Fig. 3. The suggestion is to identify a threshold for the duration, for which most disturbances are of shorter durations. This may be done by studying figures like Fig. 3 and see if a 'knee' can be identified. If no knee can be identified, include all disturbance data. For the example in Fig. 3, one possible threshold seems to be at approximately 10 hours, which is marked by a red dashed line in the figure. The data used in the figure are real industrial data for the middle-pressure steam utility from a site operated by Perstorp. This data will be used as an example for nonperiodic disturbances throughout remainder of the section.

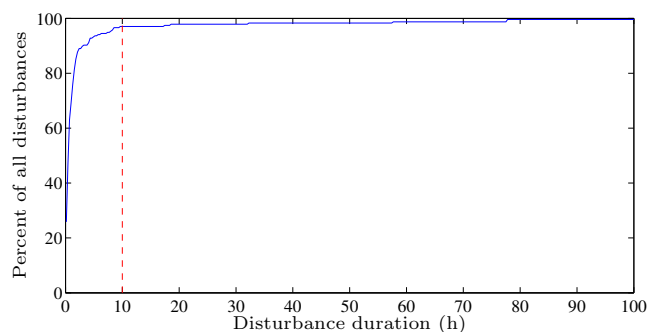


Fig. 3. Percentage of disturbances of duration shorter than a threshold.

2.2 Step 2: Identify Type(s) of Disturbances for the Utility

To begin with, the characteristics of the disturbances have to be evaluated. Two main types of utility disturbances may be identified: *periodic* and *nonperiodic*.

Periodic disturbances Periodic disturbances can be characterized by a shape (e.g. a sinusoid or a square wave) and a period time/frequency. A sinusoidal disturbance may be

described as $A\sin(\omega t)$, where A is the amplitude and ω the angular frequency. A utility that could typically suffer from periodic disturbances is the cooling water utility, because of daily and yearly variations of the outdoor temperature.

Nonperiodic disturbances The idea is to characterize nonperiodic disturbances by three parameters: duration, severity, and time between failures. The severity is evaluated as the depth of the drop or the height of the peak in the measurements, and the time between failures is defined as the duration from the end of one disturbance until the beginning of the next disturbance. The disturbance measures are visualized in Fig. 4.

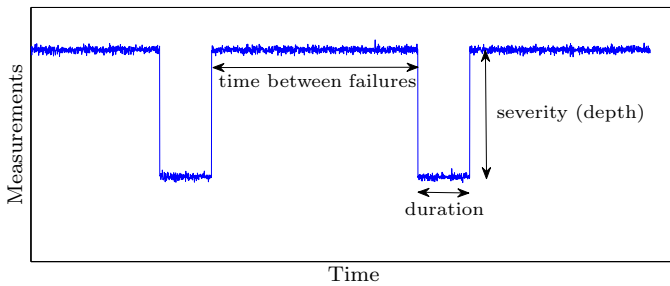


Fig. 4. Disturbance measures for nonperiodic disturbances.

2.3 Step 3: Find One or More Typical Disturbance(s) for the Utility

This step is divided into two parts: One for periodic disturbances and one for nonperiodic disturbances. There could be several typical both periodic and nonperiodic disturbances for a utility. If the results at one of the substeps seems to be inconsistent, two or more typical disturbances can be identified, with different amplitude and/or period time for periodic disturbances, or with different disturbance duration and/or severity for nonperiodic disturbances.

Periodic disturbances The existence of periodic disturbances may be seen by studying the measurement data, to see if a shape could be identified. There could also be physical indications that there should exist periodic disturbances, such as daily or yearly variations of the outdoor temperature. When the shape has been identified, identify the amplitude and period time of the variations.

An example is taken from a site operated by Perstorp. Fig. 5 shows the cooling water temperature together with the identified disturbance limit of 27°C , and the outdoor temperature during almost three years. The measurement data imply that there are both daily and yearly variations of the temperature, and the variations seem to be approximately sinusoidal-shaped. The temperature $T(t)$ may thus be described as

$$T(t) = T_{\text{norm}} + A_d \sin\left(\frac{2\pi}{24 \cdot 60} t\right) + A_y \sin\left(\frac{2\pi}{24 \cdot 60 \cdot 365} t\right) \quad (1)$$

where $T_{\text{norm}} = 24^\circ\text{C}$ is the normal cooling water temperature, A_d the daily temperature variations, and A_y the yearly temperature variations. The daily and yearly variations could be studied separately to find the amplitudes A_d

and A_y . Given the data in Fig. 5, amplitudes of $A_d = 1.5^\circ\text{C}$ and $A_y = 4^\circ\text{C}$ seem appropriate. The resulting disturbance trajectories for the daily and yearly temperature variations are shown together with the measurement data in Fig. 6.

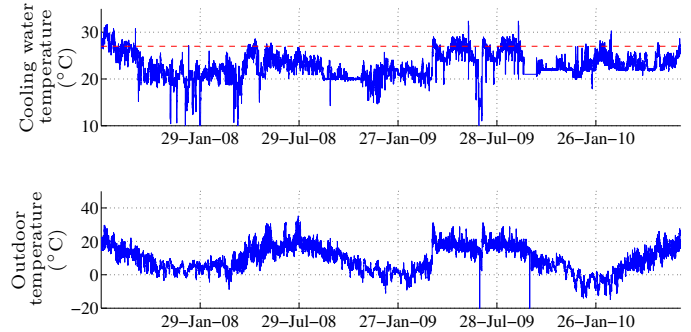


Fig. 5. Temperature of cooling water and outdoor temperature at an industrial site.

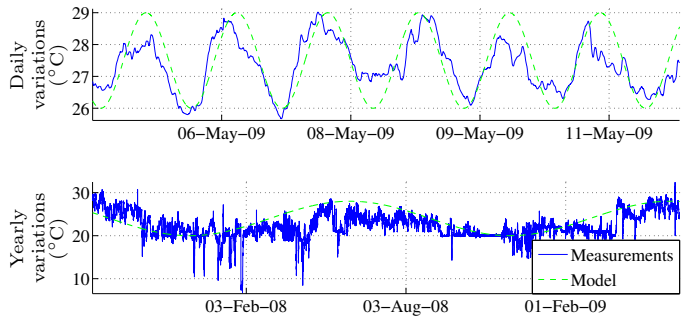


Fig. 6. Model for daily and yearly variations of cooling water temperature.

Nonperiodic disturbances For nonperiodic disturbances, the procedure is to find the duration, severity, and time between failures for one or more typical utility disturbances. How to find these measures by forming statistical models is discussed below.

I. *Estimate most probable disturbance duration:* Some different probability distributions were evaluated as candidates for the distribution of the disturbance durations, among others the normal distribution, lognormal distribution, exponential distribution, and the Weibull distribution. Probability plots (see e.g. Rice (2007)) were used to assess the fit of data to the distributions. The lognormal distribution showed to give the best fit for the utility data from Perstorp, why this is our suggestion for the statistical model of the disturbance durations for utilities. The duration of a disturbance is related to the concept of *repair time* for applications with repairable units. Time to repair is according to e.g. Hamada et al. (2008) and Schroeder and Gibson (2010) also well characterized by a lognormal distribution, which further motivates the choice of this distribution for the disturbance durations.

The suggestion is thus to fit a lognormal distribution to the disturbance duration data. The lognormal distribution

has two parameters, μ and σ and the probability density function (pdf) is given by

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} x^{-1} e^{-(\ln x - \mu)^2 / (2\sigma^2)} \quad (2)$$

where $0 < x < \infty$. Further information about the lognormal distribution and its applications is given e.g. in Crow and Shimizu (1988).

In the upper subfigure of Fig. 7, a histogram of the disturbance durations for the middle-pressure steam utility at the Perstorp site is given together with the lognormal probability density function that gives the best fit (red line). The median disturbance duration is marked with a green dashed line in the figure, and the average disturbance duration with a blue dash-dotted line. As seen in the figure, the average duration is significantly longer than the median, because of some long and uncommon disturbances. The parameters of the distribution that gives the best fit are $\mu = 3.22$ and $\sigma = 1.44$. The median disturbance duration is $e^\mu = 25$ minutes and the average duration $e^{\mu + \sigma^2/2} = 70$ minutes. The cumulative probability density function (cdf, red line) is plotted in the bottom subfigure of Fig. 7 together with the measured disturbance durations (blue dots). According to this plot, the lognormal distribution seems to be a good fit for the disturbance durations.

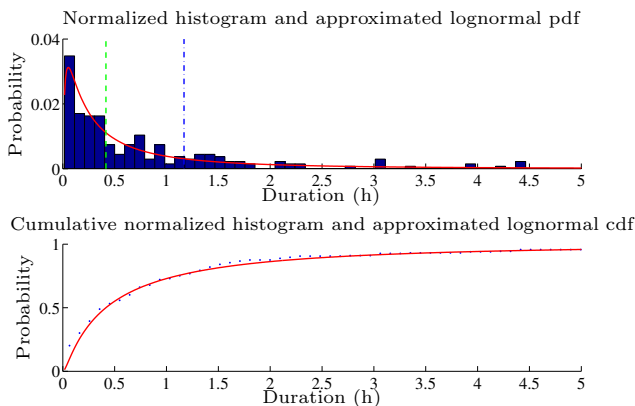


Fig. 7. Statistical model for disturbance durations for middle-pressure steam.

II. *Estimate most probable severity of disturbance:* The maximum deviation in the measurements from the normal operating point is correlated to the duration, but the correlation is not very strong. In fact, the probability distribution of the deviation is not at all lognormal in the cases studied here. It does not fit any of the "standard" distributions (normal, lognormal, exponential, Beta, Weibull, or Gamma) to a statistically significant level. The suggestion is therefore to use a normalized histogram with an appropriate number of bins to produce an empirical probability density function for the severity of the disturbance. A normalized histogram of the maximum deviation from the nominal pressure at disturbances in the middle-pressure steam utility at the Perstorp site is shown in Fig. 8. In the histogram, the median and average deviation are marked with a green dashed line and a blue dash-dotted line, respectively. For the example, the median is approximately 5 bar, whereas the average is around 6 bar.

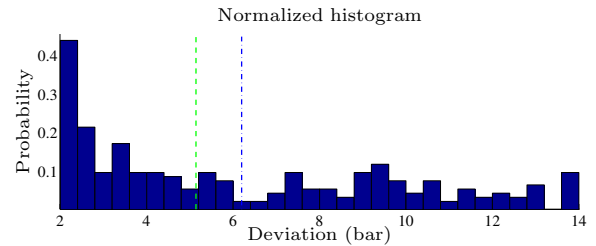


Fig. 8. Normalized histogram of maximum deviation from nominal pressure for middle-pressure steam.

III. *Estimate most probable time between failures:* The Weibull distribution is often used to model the time to failure or life span of a component (Pal et al. (2005)). Some application areas include computing systems, for which among others Schroeder and Gibson (2010) and Zhang et al. (2005) model the time between failures using the Weibull distribution. By evaluating the probability plots for some common distributions (as in step I), it was found that the Weibull distribution also seems to fit the time between utility failures well, based on the data from Perstorp. The distribution has two parameters, one scale parameter, a , and one shape parameter, b . The probability density function is given by

$$f(x|a, b) = \frac{b}{a^b} x^{b-1} e^{-(x/a)^b} \quad (3)$$

where $0 < x < \infty$, and $a, b > 0$. More information about the Weibull distribution can be found in, e.g., Pal et al. (2005).

In the upper subfigure of Fig. 9, the Weibull probability density function (red line) that gives the best fit for the middle-pressure steam utility at the Perstorp site is plotted together with a normalized histogram of the time between failures for the utility. The median and average time between failures are marked by a green dashed line and a blue dash-dotted line, respectively. The cumulative distribution function (red line) is plotted together with the measurements (blue dots) in the lower subfigure of Fig. 9. These plots show that the Weibull distribution seems to be a good fit for the example data. The parameters of the Weibull distribution that gives the best fit are $a = 2500$ and $b = 0.46$, for disturbance data in minutes. The median time between failures for the distribution is

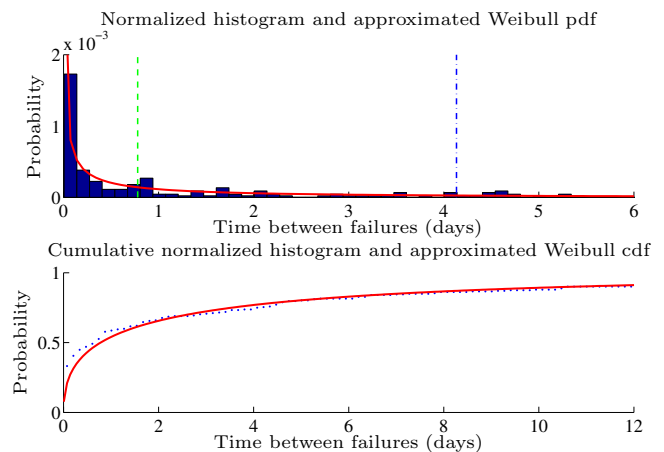


Fig. 9. Statistical model for time between failures for middle-pressure steam.

$a(\ln 2)^{1/b} = 19$ hours, whereas the mean time between failures is $a\Gamma(1 + 1/b) = 4$ days for this data set, where Γ is the gamma function. Mean time between failures (MTBF) is a measure that is often used for different kinds of repairable systems.

2.4 Results

The typical disturbance trajectories are obtained after performing the three steps described in this section. For nonperiodic disturbances, distributions for the duration, severity, and time between failures of the disturbance were identified in substeps I-III of step 3 of the method. A typical disturbance trajectory is obtained as a realization using the statistical models of these measures. An alternative is to use the median or average for the distributions to generate a typical disturbance trajectory.

3. A CASE STUDY

In this section, the general method described in section 2 is used to identify typical disturbances for selected utilities at a site operated by Perstorp. The utilities that have been studied are

- Middle-pressure (MP) steam
- High-pressure (HP) steam
- Cooling water
- Nitrogen
- Combustion device

Nitrogen is used at the site to maintain pressure in vessels and the combustion device for combustion of tail gas. Other utilities are also used at the site, but for these utilities, the disturbances have been so few during the considered time period that the data set is not sufficiently large to produce reliable disturbance models. The data that are available span approximately three years of operation.

3.1 Step 1: Prefilter the Utility Measurement Data

The approach that was selected for prefiltering of the data was to merge all disturbances where the time between failures were shorter than 10 minutes. To remove very unlikely disturbances of long durations from the data sets, Fig. 10 was studied. The figure shows the percentage of all disturbances for the utilities that are shorter than a threshold. MP steam was omitted in the figure, since this utility has already been treated in Fig. 7. The red dashed lines in Fig. 10 mark the thresholds that were set for the considered utilities: approximately 10 hours, 5 hours, 100 hours, 2 hours, and 21 hours for MP steam, HP steam, cooling water, nitrogen, and the combustion device, respectively.

3.2 Step 2: Identify Type(s) of Disturbances for the Utilities

In this step, it was concluded that the cooling water utility was the only utility that indicated a periodical behavior.

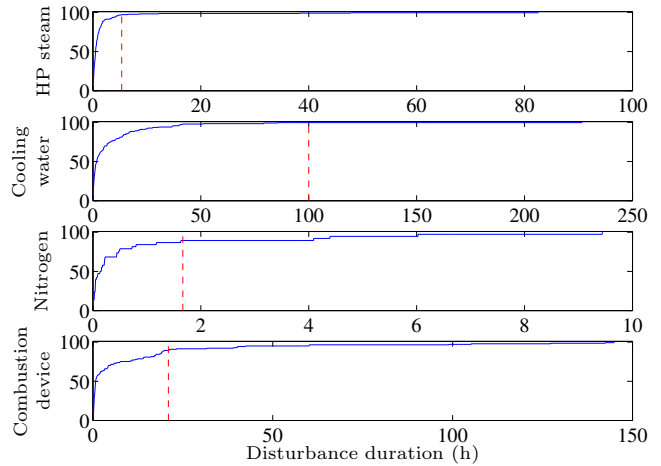


Fig. 10. Percentage of disturbances of duration shorter than a threshold.

3.3 Step 3: Find One or More Typical Disturbance(s) for the Utilities

One typical disturbance for each utility was identified; a periodic disturbance for cooling water, and nonperiodic disturbances for the MP steam, HP steam, nitrogen, and combustion device utilities. The procedure is described in more detail below. Here, MP steam and cooling water have been left out, since these utilities were used as examples in section 2, and the details can be found there.

Periodic disturbances The identification of the characteristics of the periodic cooling water disturbance was handled as an example in section 2, with the resulting disturbance trajectory structure (1) for the temperature.

Nonperiodic disturbances The three steps to identify typical nonperiodic utility disturbances are here carried out for high-pressure steam, nitrogen, and the combustion device simultaneously.

I. *Estimate most probable disturbance duration:* The estimated lognormal probability distributions for the disturbance durations are shown together with the normalized histograms in Fig. 11 for the three utilities. The median disturbance durations are marked with green dashed lines and the average durations with blue dash-dotted lines. The parameters μ and σ of the best fit lognormal distributions are given in table 1, for disturbance data in minutes.

Table 1. Parameters of lognormal distribution for disturbance duration.

Utility	μ	σ
HP steam	3.34	1.16
Nitrogen	1.87	1.34
Combustion device	4.25	1.51

II. *Estimate most probable severity of disturbance:* Normalized histograms of the maximum deviation from the normal operating point at disturbances in the three utilities are shown in Fig. 12. For HP steam and nitrogen, the measurement unit is pressure in bar, and for the combustion device it is flow in kg/h. For all three utilities,

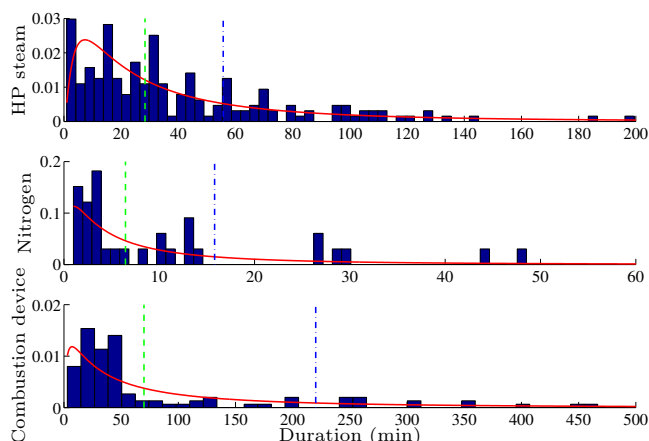


Fig. 11. Statistical models for disturbance durations.

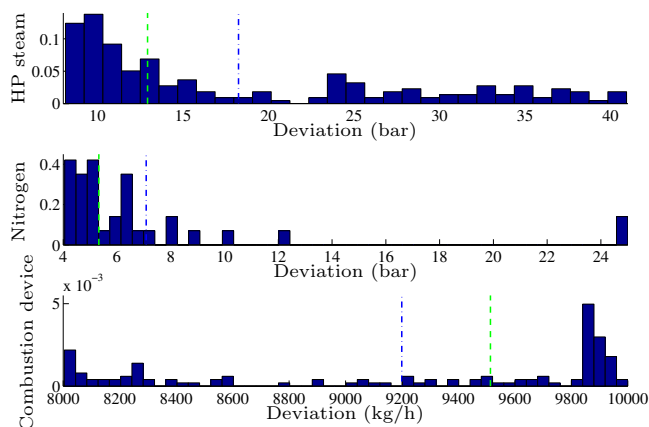


Fig. 12. Normalized histograms of maximum deviation from normal operating point.

a disturbance is a drop in the pressure or flow. The median depths of the drops are marked with dashed green lines and the average with dotted blue lines.

III. *Estimate most probable time between failures:* Fig. 13 shows histograms of the time between failures for the three utilities together with the suggested Weibull distributions. The median time between failures is marked with dashed green lines and the average with blue dotted lines. The scale and shape parameters a and b of the best fit Weibull distributions are given in table 2, for disturbance data in minutes.

Table 2. Parameters of Weibull distribution for time between failures.

Utility	a	b
HP steam	2676	0.45
Nitrogen	3875	0.32
Combustion device	2630	0.40

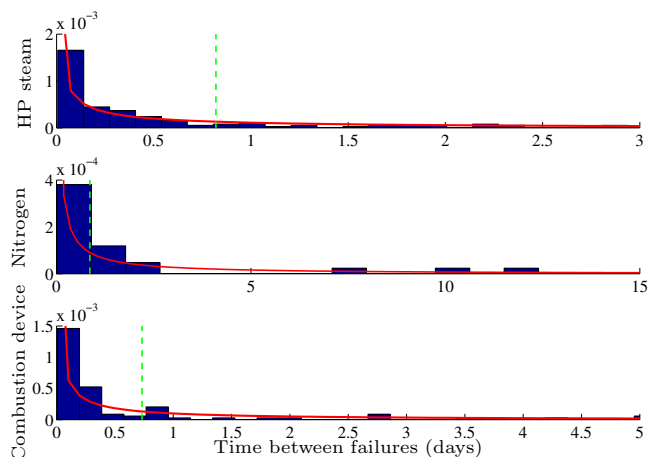


Fig. 13. Statistical models for time between failures.

3.4 Results

The typical periodic disturbance for the cooling water utility is given by the temperature trajectory in °C:

$$T(t) = 24 + 1.5 \sin\left(\frac{2\pi}{24 \cdot 60} t\right) + 4 \sin\left(\frac{2\pi}{24 \cdot 60 \cdot 365} t\right) \quad (4)$$

The other utilities suffer from nonperiodic disturbances. The distributions of the duration, severity, and time between failures for the nonperiodic disturbances were estimated in step 2 of the method. To get a feeling for what the typical disturbance looks like, the median of the disturbance duration (l), the median of the severity parameter (d), and the mean time between failures (tbf) are presented in table 3.

Table 3. Typical utility disturbances at the Perstorp site.

Utility	l	d	tbf
MP steam	25 min	5 bar	4 days
HP steam	28 min	13 bar	5 days
Nitrogen	6 min	5 bar	19 days
Combustion device	70 min	9513 kg/h	6 days

3.5 Discussion

For the periodic cooling water disturbance, the suggested sinusoidal disturbance seems to capture the disturbance characteristics quite well, even though it can be seen in Fig. 5 that the periodicity of the measurements of the outdoor temperature is more distinct than those of the cooling water temperature. This might be due to the feedback effect: if the plant is (for other reasons than cooling water problems) run at reduced speed, less cooling effect is needed and the cooling water temperature will be lower. This might explain the drops in the measurements of the cooling water temperature that can be seen in the figure.

For the nonperiodic disturbances, most figures seem to show similar results for the four utilities, with some exceptions. In step I, the probability distributions in Fig. 11 seem to fit better for utilities where the data contain a larger number of failures, in this case for the steam utilities. This is also seen in step III of the distributions for

the time between failures, but not as clearly. The weakest result is the estimation of the severity of the disturbances in step II. The histograms do not clearly show a peak or a tendency of the data to fit any of the standard distributions. For the nitrogen and combustion device utilities, it seems like there might be two types of disturbances; one smaller disturbance, and one with a greater deviation from the normal operating point. An attempt to improve the modeling could be to divide these into two separate cases and redo substeps I-III for each of the cases.

4. CONCLUSIONS

A general method for finding typical utility disturbance trajectories were presented and used to analyze disturbances in the supply of utilities at a site at Perstorp. The disturbances were categorized into periodic and nonperiodic disturbances, and characterized by simple measures: period time and amplitude for periodic disturbances, and duration, severity and time between failures for nonperiodic disturbances. The duration of a disturbance was modeled using the lognormal distribution, and the time between failures using the Weibull distribution. The resulting disturbance trajectories may be useful for both proactive and reactive disturbance management.

An interesting future work direction would be to combine the identified statistical disturbance models with an optimization framework for reactive utility disturbance management. It would also be valuable to further investigate the nature of periodic utility disturbances, to obtain a general method for identifying the period time and amplitude of such disturbances.

ACKNOWLEDGEMENTS

The research is performed within the Process Industry Centre (PIC) supported by the Swedish Foundation for Strategic Research (SSF). The first author is a member of the LCCC Linnaeus Center and the eLLIIT Excellence Center at Lund University.

REFERENCES

- Crow, E. L., Shimizu, K., 1988. Lognormal Distributions: Theory and Applications. Marcel Dekker, USA.
- Greenberg, H. R., 1991. Risk Assessment and Risk Management for the Chemical Process Industry. John Wiley & Sons, Inc., USA.
- Gustafsson, F., 2000. Adaptive filtering and Change Detection. Wiley, UK.
- Hamada, M. S., Wilson, A., Reese, C. S., Martz, H., 2008. Bayesian reliability. Springer, USA.
- Iyer, R. R., Grossmann, I. E., 1998. Synthesis and operational planning of utility systems for multiperiod operation. Computers & Chemical Engineering 22 (7), 979–993.
- Khanbaghi, M., Malhame, R., Perrier, M., Roche, A., 1997. A statistical model of paper breaks in an integrated tmp-newsprint mill. Journal of pulp and paper science 23 (6), J282–J288.
- Lee, W., Weekman, V. W., 1976. Advanced control practice in the chemical process industry: A view from industry. AIChE Journal 22 (1), 27–38.
- Lindholm, A., Giselsson, P., 2013. Minimization of economical losses due to utility disturbances in the process industry. Journal of Process Control 23 (5), 767–777.
- Lindholm, A., Johnsson, C., 2013. Plant-wide utility disturbance management in the process industry. Computers & Chemical Engineering 49, 146–157.
- Ogawa, S., 2003. Stochastic averaging level control and its application to broke management in paper machines. Ph.D. thesis, The University of British Columbia, Canada.
- Pal, N., Jin, C., Lim, W. K., 2005. Handbook of Exponential and Related Distributions for Engineers and Scientists. Chapman and Hall/CRC, UK.
- Perstorp, 2013. [online] <http://www.perstorp.com/>, June 2013.
- Rice, J. A., 2007. Mathematical statistics and data analysis. Thomson, USA.
- Schroeder, B., Gibson, G. A., 2010. A large-scale study of failures in high-performance computing systems. IEEE Transactions on Dependable and Secure Computing 7 (4), 337–350.
- Zhang, Y., Squillante, M., Sivasubramaniam, A., Sahoo, R., 2005. Performance implications of failures in large-scale cluster scheduling. In: Feitelson, D., Rudolph, L., Schwiegelshohn, U. (Eds.), Job Scheduling Strategies for Parallel Processing. Vol. 3277 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 233–252.