Applied statistical quality control on field measurement data

Wibog, Erik
Statkraft / NTNU
erik.jacques.wiborg@statkraft.com

Hulaas, Harald
Norconsult
harald.hulaas@norconsult.com

Francke, Håkon
Flow Design Bureau
hf@fdb.no

Nielsen, Torbjrn Kristian
NTNU
torbjorn.nielsen@ntnu.no

August 5, 2014

Abstract

Field efficiency measurements on hydro power plants usually do not have a validation check during the ongoing field work. If the end results of the field test is analysed during the test itself, field engineers have a chance to check for obvious mistakes. Results clearly deviating from the expected trend (subjective to the field engineer) may be questioned and a repetition of the measurement is then performed. This does not usually apply for errors that are relatively small (but not necessarily insignificant). Operation of a secondary turbine during a performance test may alter the flow in the waterway, potentially leading to an error in the measurements. Disturbances like water surges, temporarily blocked sensors, temporary resonance; there are multiple reasons for a measurement to be slightly off during testing. Statistical Quality Control (SQC) of the data, running continuously within the acquisition system is a possible solution to this problem. Logged data can with this tool be checked for validity before advancing the measurement schedule. This would allow the field engineers an instant check, and the possibility to immediately re-measure any doubtful measurement. Such a system in being designed during the phd work of EJW and tested on data provided by Norconsult and by FDB on three different Statkraft power plant. Results are interesting, underlining the usefulness and the potential of the method. There are however still some unresolved issues, they will be presented in the hope of getting feedback towards possible solutions.

1 Background

Field measurements on hydro power plants are costly, not only because of the measurement itself, but also because of down time on the runner in question. This is why power plant owners wish the measurement to be as short as possible. At the same time, the measurement results may have large financial impacts and need to be done properly. During the measurements, many different types of data logging are performed, and one can simply not control manually every measurement manually. An automated check would be helpful for this task.

The paper is a continuation of the paper ”Statistical Quality Control and Field Measurements in Hydro Power Plants” [1] still unpublished. Most of the theory presented here can also be found in this first paper.

2 The method - general approach

To validate data during field measurements a strong statistical tool is employed. The method itself is inspired by process control, and the use of Shewhart diagrams [2].

In process control, a good way of monitoring the stability of a production line is to perform sample measurements displaying the results in a Shewhart diagrams, and check if the data follows some simple rules of normality. This enables the owner to control that the production maintains stable, accounting for natural variations in the production, which should by default be normally distributed around a target mean. A set of rules ensuring the normality are proposed [2]. A rule violation implies that the measurements show deviation from the expected distribution, and additional control should be performed.

One of the simplest and most commonly used rule is the "three-sigma control limit" [2]. Based on the 25[1] first measurement points one can compute a mean and standard deviation, the basic values needed to set up our control limits. This is what one could call "phase 1". During the continuation of the measurement ("phase 2"), If a measured value crosses the control limits, it is in violation, see Figure 1.

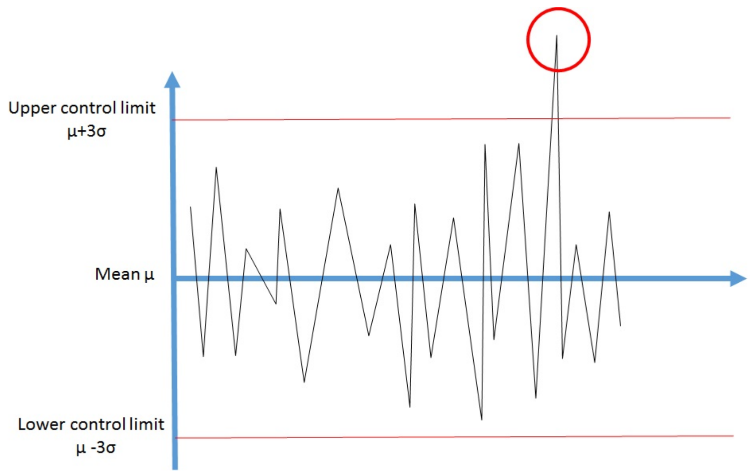

Figure 1: Shewhart Diagram Example - Three-sigma rule violation

In this study the three-sigma rule has been chosen, along with an other of the most commonly used rules: Two out of three consecutive measurements cross the same two-sigma control limit. This rule will be denoted as the double-two-sigma rule henceforth.

This method is the theoretical basis for a validation software (created in LabView) that can be implemented to the data acquisition system used during field testing.
The main goal being, of course, to check that the measurand (pressure, vibration, flow, level, temperature, and so forth) is stable during the data acquisition phase. The assumption of stability being, for process control, a normal distribution around a mean. Dealing with water power one expects there to be residual pressure waves in the system that may disturb the measurements. Some waves may take too long to dissipate and the software will have to accept the presence of a wave, which could lead to a discrepancy in terms of distribution. This topic is addressed later in the paper.

3 The method, adapted to hydro power field measurements

It is important to underline that some rule violations are bound to occur, even in stable conditions. By definition there ia a 0.3 % chance for a measurement point to violate the three-sigma rule [3], for example. We wish to avoid unnecessary hold-ups during field testing, our software therefore needs to accept a certain amount of rule violations. For the purpose of assessing a correct amount of authorized rule violations, a

[1]This number is suggested by Montgomery [2]. Based on the assumption of normal distribution, one should have enough statistical information with these measurements to lay down "rules of normality" that are well balanced for this use.

study was done earlier, and presented in the first paper dealing with this issue [1].
The main task of the study was to find the relationship between the number of sample points and the expected amount of natural violation. Using repeated randomly generated values with a normal distribution (therefore, by definition without any causal disturbances) a linear relationship was found for both of the rules and their combination.

Another issue that became apparent while developing the validation software. The sensitivity to the first 25 measurement points ("phase 1", see chapter 2) was troublesome. This will be explained in details further along this paper, but mainly one would have problems with measurands that stabilized themselves during testing, leading to false (too wide) control limits. The need for a second check became apparent. A back-to-back approach was therefore undertaken.
Two checks are therefore performed by the software. Once in the normal way described i section 2, and then another one, this time with phase 1 based on the 25 last values. The check with the highest amounts of violations is the one selected for the analysis. This way, if the measurand stabilized itself during the acquisition phase, the second measurement would detect higher variation in the beginning of the acquisition period and visa-versa.

Finally, it became quickly obvious that some of the measurements, branded as invalid by the tool, where in fact usable. The software can not be entrusted to automatically discard data. As a results, a report will now be displayed for invalid measurements, leaving the final judgement in the hand of the chief of test. Figure 2 is an example of such an report. Examples of such usable data will be described in the next chapter.

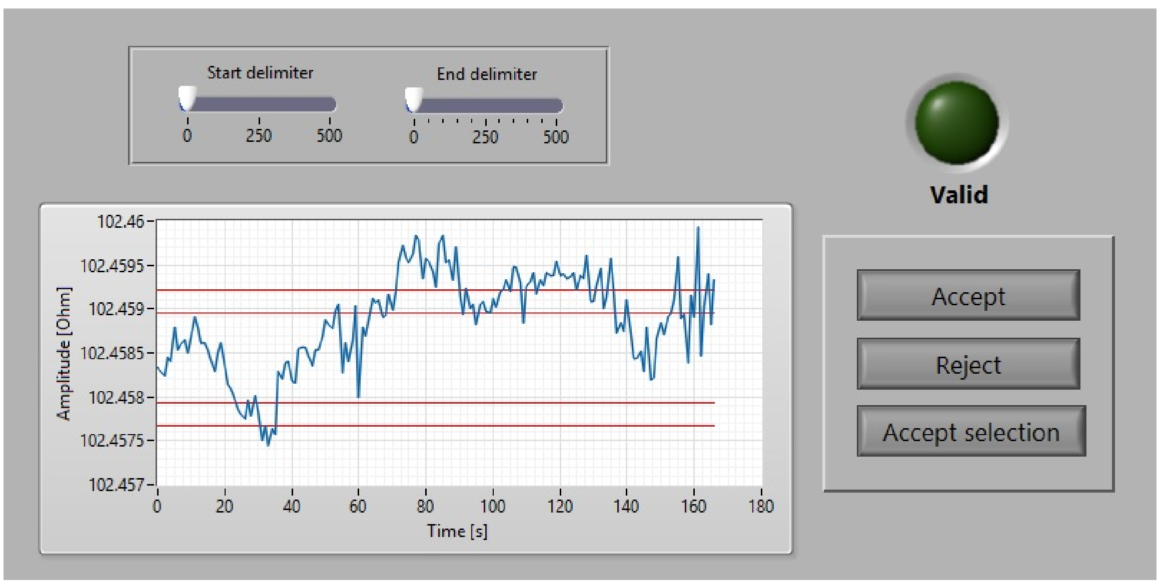

Figure 2: Validity report

4 Testing on real data

Upon testing on real data the use of the software proved its worth. Most of the data from the three site tests where accepted by the software, with right. Some measurements where however picked up as invalid by the system. Bellow are some selected results, examples of different cases of issues.
This is more or less a repetition of the different situation uncovered in the first analysis [1].

4.1 Hasty start

Depending on the power plant, the time needed to reach a stable situation after load change may vary. The necessary waiting time will of course vary accordingly. A hasty start may be the cause of faulty measurements, figure 3 is an example of just that.

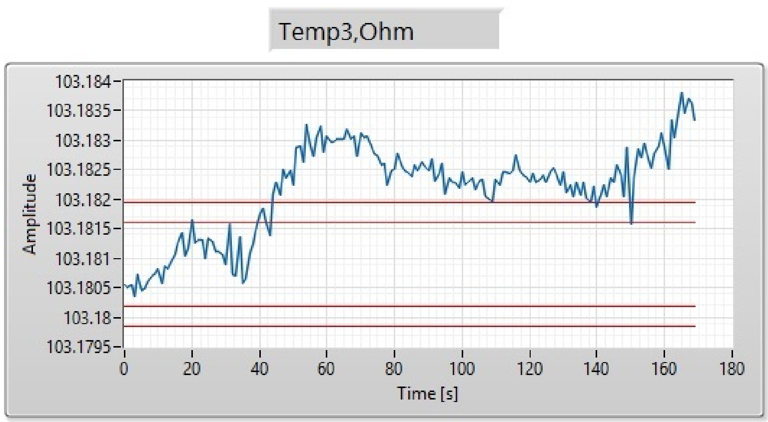

Figure 3: PP1 101rv - Temperature measurement (Ohm)

Hasty starts are not uncommon, but it may be counter productive to automatically discard the whole measurement just because of a small amount of offset measurement points.

This is a typical case where one could use the "accept selection" choice of the softare report (see figure 2).

4.2 Causal variation

This was the original case that the software was supposed to detect. Picking up variations due to causal effects (in contrast to random errors, caused by static).

Figure 4: PP2 107 Temperature Diff (Ohm) - Causal variation

4.3 False alarm

Here are examples of why there is a need for human judgement (see 3). Statistics are never 100% accurate, and some rejections may not be justified.

In the first example (figure 5) the number of violations is indeed above the suggested limits, but the mean seems stable and centred. The cause of the invalidity of this measurement is simply an above average erratic static noise. It would be wrong to discard this measurement.

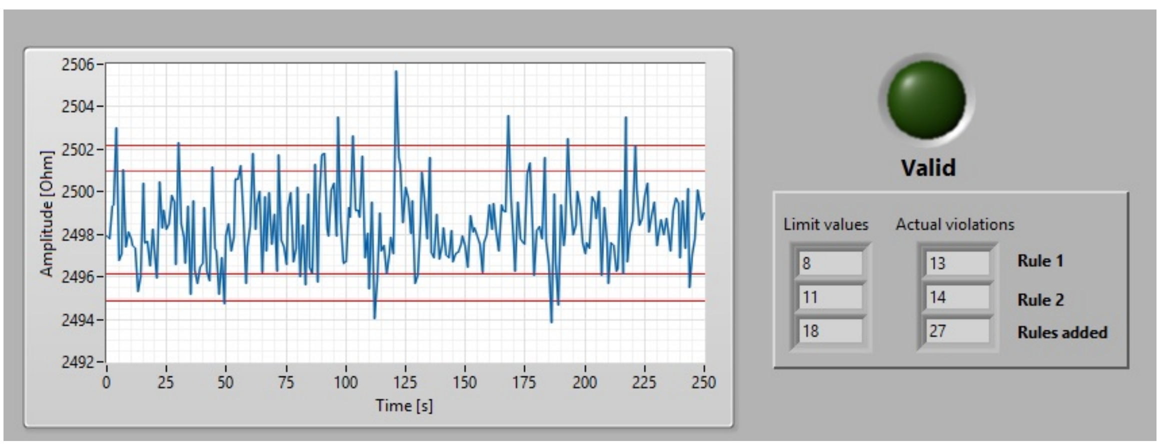

Figure 5: PP3 05 Pressure (kPa) - Possible false alarm

In the second example (figure 6) there are still surges that need to dampen, this is actually a case of hasty start mentioned above. However the wave amplitude and phase would suggest that the mean does not deviate much from the true mean. By using the validity report (see figure 2) one could verify the mean for the whole set, and compare it with the mean of a selection after the wave has dampened. Leaving the option of excluding the first wave with the largest amplitude.
This is a good example of why the back-to-back approach was needed. The control limits are the ones computed by the second check, with phase 1 based on the last 25 measurement points.

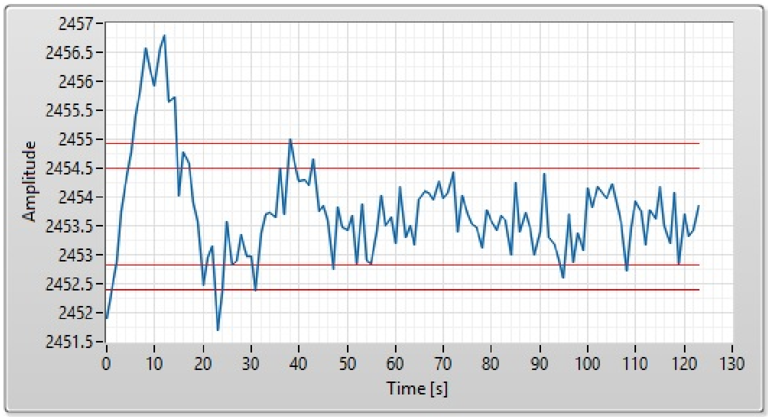

Figure 6: PP3 07Pressure (kPa) - Surges - Valid

5 Issues and future improvements

5.1 Phase 1

During phase 1, where the control limits and mean are computed, we must assume perfect stability and Gaussian distribution of the measurement values. This is of course rarely the case, surges give sinusoidal disturbances to pressure, flow and temperature. Especially for high frequency measurements, we are likely to encounter problems related to autocorrelation error if standing waves are present.
Highly correlated measurements will provide a poor initialization of the mean and control limits.

To study the effects of autocorrelation, an autocorrelation analysis was performed on some sample measurement series. This was done with the statistical tool R [4]. Bellow are example of the results provided by R.

For measurements with no apparent standing wave, like the one shown in figure 5 no autocorrelation was detectable allready after the second point, see figure 7. In this case the 25 points rule is deemed applicable.

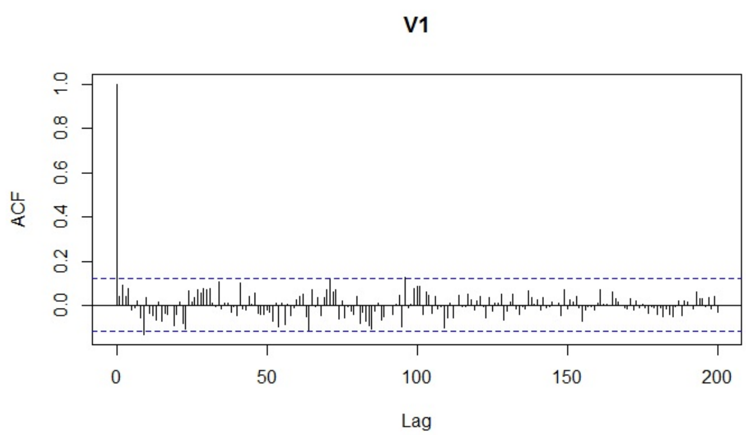

Figure 7: PP3 05 Pressure (kPa) - Possible false alarm - Autocorrelation

The measurement in figure 8 is the "PP1 101rv" measurement, displayed in figure 3.

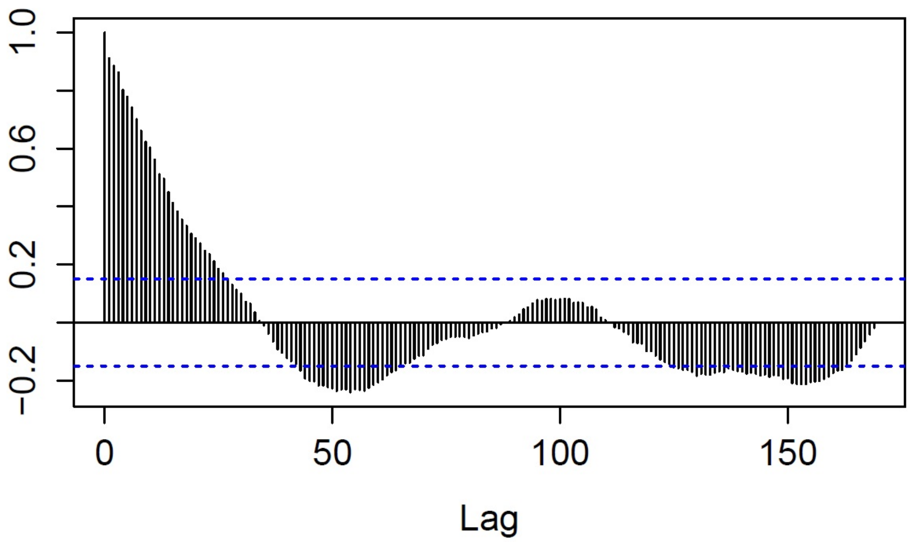

Figure 8: PP1 101rv - Autocorrelation plot

And as we can see from figure 8 we still have autocorrelation issues after 25 measurements. In this case we should have used at lest 30 measurements for phase 1.
Such an autocorrelation check should be implemented to the software, ensuring that we have a sound initialization of the control limits.

Note that for standing sinusoidal waves the autocorrelation may not stabilize itself. But once we have reached a satisfactory level (neglectable autocorrelation) we should have recorded at least half a period of the largest standing wave, which should be enough for a satisfactory phase 1 initiation.

This being said, the 25 measurement points offer a more strict limitation. One could argue that it would be better to check the invalidity report more frequently than necessary, only to accept falsely rejected measurements, rather than the alternative.

5.2 Instability consequences

As one may have noticed from the different figures presented in the paper, the amplitude of variation is usually very small. The validation method is purely statistical, and does not take into account the consequences for the end result.

As mentioned before, human judgement should always be included. A possible improvement to the software could be to have it quantify the effects of a (small) change of the measurand value on the end results. This way the chief of test can decide to accept unstable measurements if they are of little consequence to the end results.

5.3 Statistics criteria

As it has been mentioned a number of times, the software presented is purely a statistical tool. Statistical predictions need a certain amount of data to be valid, this is no exception.

The law of Small Numbers [5] is applicable here: Too few measurement points may not give a proper representation of the true state of the measurand, and extreme ratios are more likely to occur for smaller populations [5]. This implies that the tool will be more correct in its assessment of a valid/invalid measurement if it contains a large amount of measurement points. Experience from the studies performed with the tool would suggest a minimum requirement of about 100 points.

5.4 The rules of validity

Two rules have been used for the validity check up untill now (the ”three-sigma” and the ”double two-sigma” rules, see 2). These are the most commonly used rules for process control, and seem appropriate for field measurement use.
There is however one other rule that may be interesting to implement: one should check for a faulty sensor/transducer, should three or more consecutive measurements be repeated exactly (to the last decimals given by the instrument). One other situation that may give these results is if the measurand saturates the measurement sensor, in which case we also have an error i measurement.

This rule should be easy to implement and should provide extra safety during measurements, as long as the instrument is of high precision and static noise is unavoidable as long as it is functioning properly.

6 Conclusion

The statistical control tool could be beneficial for a quick check of measurement stability during field measurements. It will however probably never be good enough to be entrusted to automatically discard faulty measurements, as it may result in unnecessary delays. The only exception would be on-line measurements where time is abundant.
It is noteworthy that all the accepted measurements where in fact truly stable. The control function of the tool may be harsh and may reject acceptable measurements, but it is certain: accepted measurements are indeed stable and valid.

Some improvements to the software are still necessary, the repetition rule mentioned in chapter 5.4 should be implemented but probably also optional since it would not function properly for low resolution measurements. Also a quick example computation of the end result consequences of unstable measurements would help in choosing whether to discard a measurement or not.

This study has proven the importance of human judgement, there is no doubt that an efficient and easy-to-use human interface is vital for a proper use of the tool.

The tool has yet to prove its worth on a live field measurement. How helpful it actually will be remains to be seen.

7 References

References

[1] E. J. Wiborg, T. Nielsen, H. Hulaas, and H. Francke, "Statistical quality control and field measurements in hydro power plants," *Journal of Dynamic Systems, Measurement, and Control [UNPUBLISHED].*

[2] D. C. Montgomery, "Introduction to statistical quality control," 1991.

[3] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability and statistics for engineers and scientists.* Macmillan New York, 1993, vol. 5.

[4] "Autocorrelation function r." [Online]. Available: http://stat.ethz.ch/R-manual/R-patched/library/stats/html/plot.acf.html

[5] K. Daniel, "Thinking fast and slow," *Allen Lane, New York, USA*, 2011.