The paper was published in the proceedings of the 11th International Symposium on Field Monitoring in Geomechanics and was edited by Dr. Andrew M. Ridley. The symposium was held in London, United Kingdom, 4-7 September 2022.

11th International Symposium Field Monitoring in Geomechanics

The Analysis of Cold Geotechnical Instrumentation Data

Garrett BAYRD

Barr Engineering, Minneapolis, United States of America
Corresponding author: Garrett Bayrd (GBayrd@barr.com)

Abstract

One of the more challenging tasks for a scientist or engineer who specializes in instrumentation is the cold data review. In these cases, a client or colleague asks for assistance in interpreting the validity of instrumentation data that is unfamiliar. This paper details a set of methods to use to understand this data. The method proposed involves determining if the instrument is mechanically functioning, and then interpreting what that data from that instrument could be saying.

In the first section, mechanical analysis of instrumentation, the goal is to gain a deeper understanding of the instrument and setting of the site. The intent is to assess the functionality of the instrument – to determine if the transducer is accurately recording the physical properties of the world that this instrument is designed to measure. If you determine that the instrument is not functioning during the first section, you don't proceed to the second section – the data is erroneous, and there is no need for further analysis.

In the second section, the interpretation of instrumentation data, the goal is to figure out what story the data from the instrument is telling. This step is more imaginative, and entails developing a full, single, coherent set of facts that describe what is happening with that instrument. Fundamentally, the data from the instrument needs to be representative of a true site condition. During this step, we begin by trusting the data, examining all the potential counterfactuals, and determining the most likely honest story about what is occurring with the instrument and data in question.

Keywords: Data Analysis, Instrumentation, Data Review, Data Processing

1. Introduction

Throughout history, humans have shared stories that explain the world around them and the data that they've gathered about that world. To an extent, they may not care if the story is true or not. They want an appealing story to explain whatever situation they find themselves in. They want to feel like they know something, even if it is not true.

In my professional practice specializing in instrumentation and data analysis, I'm often asked to determine what potentially erroneous instrumentation data means. This means that clients and colleagues want me to tell them a story. The story could be that of an incorrect instrument installation, poor data management, a faulty sensor, or an unexpected site condition. Even in conditions where the data in question is not dangerous, and does not exceed a threshold value, stakeholders want to feel that they understand the site. The existence of this story is more important, in their mind, than the truth of the story. Therefore, the burden of truth rests on me to ensure that the story I offer them is truthful. It is my job as a scientist to convey truthful information to other interested parties, while trying to keep the story as clear as possible. On my best days, I can convey the most truth with the best story, and help others clearly understand the situations that we are monitoring.

To discover what is happening at an instrumented site that is giving potentially erroneous data, I analyse data to determine the most accurate story with two main steps. The first is mechanical, and easier to document. The second is imaginative, and harder to convey. The intent with the first section is to attempt to determine the functionality of the instrument – to determine if the transducer is accurately recording the physical attributes of the world that this instrument is designed to measure. The intent with the second section is to determine what story the data from the instrument is telling.

2. Mechanical data analysis of instrumentation data

The first step in instrument data analysis is determining if the instrument is functioning as designed. Functioning as designed means two things:

- The transducer on the instrument is accurately recording and converting the physical property we seek to measure into an electrical signal.
- The instrument is installed in such a manner that the transducer is experiencing the physical property we intend for it to measure.

To determine if the instrument is functioning, I need to develop a deep understanding of the instrument, determine the installation methods, understand the transducer type, and gain full access to all the raw data from that instrument. There are several tasks that need to be done to perform this mechanical data analysis. Some of these tasks will be more important than others depending on the instrument. For example, a shape array has all the calibration factors programmed into the data analysis software, whereas an in-place inclinometer relies on entering the correct calibration data.

2.1 Develop your understanding of the transducers and the dataloggers.

Determine how the transducers work in the instruments in question. Determine the fundamental physical property of the world that is being measured, as well as the technology that the transducer uses to measure physical phenomena. Is your inclinometer MEMS or an older style? Is your piezometer using vibrating wire technology or something different? Different types of transducers tend to fail in different ways and have different accuracy levels.

Once you know about the transducers, it's important to consider the dataloggers that record these measurements. First, determine if the data is manually read or read by a datalogger. If it's manually read, are there any trends depending on the staff performing the readings? If it has automated readings, determine if the datalogger is performing any data processing or correction internally, prior to submitting the data. If the datalogger is programmable, pull and review the datalogger's code. Have any changes happened with the datalogging system? Review the system's power levels, especially solar powered systems. In some cases, the system may have enough power to power the datalogger but not the instrumentation, which leads to erroneous readings.

The goal with this section is to determine that you believe the raw data as provided to you is correct. The instrument is appropriate for measuring this physical property, and the datalogger is appropriate for measuring the instrument.

2.2 Examining the raw data

The first step is reviewing the raw data. Access to the raw data is crucial to any analysis of instrumentation. If you don't have the raw data, you won't be able to fully analyse the potential root causes of any instrumentation issues. If possible, observe the data collection at the moment of measurement (Tufte, 2020). If the raw data is manual, talk to those who took the measurements to determine the methods of measuring and the site conditions during the measurements.

Start the analysis by getting the raw data open in a viewable manner. Excel or Notepad++ work well for analysis of the most common data formats. Do an order of magnitude sanity check for the raw data – see if it looks reasonable compared to other raw data you've encountered for similar functional instruments in similar site conditions. Compare the raw data to other raw data for similar instruments, or the instruments in question prior to the shift in data. If the data is open in Excel, conditional formatting on specific columns can help look for jumps or steps in the data.

Typically, cold data reviews are prompted by a change in the data at a site, or unexpected data at the site. A review of the raw data can help inform if the change in data is due to a change in the outputs from the transducer, or if the change in data is due to a new error in the data processing (Figure 1). Additionally, review of the raw data can determine if the transducer or datalogger has failed.

TIMESTAMP	RECORD	SERIAL_NUMS	SAA2_ACC_VALUES(1,1)	SAA2_ACC_VALUES(1,2)	SAA2_ACC_VALUES(1,3)	SAA2_ACC_VALUE
4/7/2022 17:30	2243	389175	32061.56	32553.9	17028	32
4/7/2022 18:00	2244	389175	32061.8	32554.1	17028.05	32
4/7/2022 18:30	2245	389175	32061.69	32553.91	17028.12	32
4/7/2022 19:00	2246	389175	32061.85	32553.91	17027.46	32
4/7/2022 19:30	2247	389175	32061.72	32554.08	17027.59	32
4/7/2022 20:00	2248	389175	32061.36	32554.03	17028.17	32
4/7/2022 20:30	2249	389175	32061.78	32554.12	17028.09	32
4/7/2022 21:00	2250	389175	32061.73	32553.97	17028.08	32
4/7/2022 21:30	2251	389175	32061.83	32553.87	17028.04	32
4/7/2022 22:00	2252	389175	32061.9	32554.13	17027.93	32
4/7/2022 22:30	2253	389175	32061.65	32554.22	17027.97	32
4/7/2022 23:00	2254	389175	32061.38	32554.12	17027.86	32
4/7/2022 23:30	2255	389175	32061.49	32554.07	17027.73	32
4/8/2022 0:00	2256	389175	32061.4	32554.01	17027.95	3
4/8/2022 0:30	2257	389175	32061.31	32553.98	17027.38	32
4/8/2022 1:00	2258	389175	32061.84	32554.56	17028.04	32
4/8/2022 1:30	2259	389175	32061.59	32554.11	17027.96	32

TIMESTAMP	RECORD	Sensor_cumDispX_001	Sensor_cumDispX_002	Sensor_cumDispX_003	Sensor_cumDispX_004	Sensor_cumDisp
4/7/2022 17:30	2628	0	-0.00023	-0.00054	-0.00023	
4/7/2022 18:00	2629	0	-0.00014	-0.0004	-0.00019	
4/7/2022 18:30	2630	0	-0.00011	-0.00031	0.00016	
4/7/2022 19:00	2631	0	-0.00013	-0.00037	-0.00014	
4/7/2022 19:30	2632	0	-0.00018	-0.00039	-0.00005	
4/7/2022 20:00	2633	0	-0.00019	-0.00046	-0.00026	
4/7/2022 20:30	2634	0	0.00027	0.00014	0.0468	
4/7/2022 21:00	2635	0	0.00031	0.00016	0.0466	
4/7/2022 21:30	2636	0	0.00016	0.00008	0.04687	
4/7/2022 22:00	2637	0	0.00007	0.00004	0.04711	
4/7/2022 22:30	2638	0	0.00023	0.00012	0.04693	
4/7/2022 23:00	2639	0	0.00019	0.0001	0.04656	
4/7/2022 23:30	2640	0	0.00028	0.00014	0.04664	
4/8/2022 0:00	2641	0	0.00021	0.00011	0.04694	
4/8/2022 0:30	2642	0	0.00009	0.00005	0.04735	
4/8/2022 1:00	2643	0	0.0002	0.0001	0.04692	

Figure 1: Raw and processed data for a shape array. In this case the lack of distinct pattern in the raw data (left) and the pattern in the processed data (right) indicates a processing error.

2.3 Review the data processing steps

Check that the data has been processed correctly. Typically, the manufacturer calibrates the instrumentation and provides an equation with variables on a calibration sheet with the instrumentation. Make sure to acquire these sheets – manufactures usually provide them if you have the instrument serial numbers. Check the equations provided by the manufacturer against the equations used in the data processing steps. Make sure the constants used in these equations match the constants you find in the calibration sheets for the instruments. Determine if there are there any added constants in the equations used to convert the raw data. Check to see if any large values are added in the equation (like depths or elevations).

Ensure that the data you are reviewing has been processed using the best practices for instrument data management, which means that the data is stored in one single location and processed once. If the data is processed in steps, check each step individually, and the conversation between steps. If that doesn't produce good data, re-process the raw data into calculated data using a separate method. I like to use Python as a check for online or automated data processing tools.

2.4 Review the installation documentation and information on the current conditions onsite

Check the installation documentation to determine how the installation was performed. It's important to know the standard practices for installing instrumentation of this type. Determine if the installations in question are consistent with the standard practices, or do they differ in any important manner. It's important to know the issues that can arise during the installation. If your instrument was installed in a borehole, how was the boring drilled and backfilled? If your instrument was installed on a structure or at ground surface, how was it attached to the area in question? It's also important to review installation materials to determine the ways in which the installation of the instrument can affect the data you gather. If you are measuring a piezometer, is it measuring a zone of elevation in a boring, or a specific point in the boring?

There are many reasons to modify and improve standard installation methods with new projects. We should not feel tied to historical installation methods just because of tradition. Different installation methods will lead to different results, however, and we should be aware of the effects these installation methods will have on our projects.

2.5 Summary of the mechanical analysis

Once you've gone through steps 2.1 through 2.4, you should be able to determine whether the instrument is functioning or not. A functioning instrument:

- Has raw data that is reasonable, given the type of instrument and installation method.
- Is using the correct data processing steps for the instrument type and installation method.
- Is an appropriate type of instrument for the desired measurements.
- Is using the provided calibration factors and equations for data reduction.
- Was installed in a manner that will allow it to experience the physical properties it is intending to measure.

The intent of the mechanical step is to determine if you should trust your data. The goal, at this point, is to know if your transducer is functioning as intended and the data is being processed in the correct manner. If you find errors with the data processing, correct them, and then convey the accurate story of why the errors occurred, what has been done to correct them, and what has been done to prevent the errors from occurring in the future.

Do not use the discovery of errors as an opportunity to assign blame, except private discussions to teach and prevent further mistakes.

Once the mechanical analysis of the data is complete, you should know whether to trust your data. If you do trust your data you can move onto the next section, in which we try and determine what the data means.

3. Imaginative analysis of instrumentation data

The second phase is more imaginative than the first. There are no equations for comfort, no absolute certainty. After we've completed step one, we are now at a stage where we believe our data. We believe that the instruments are functioning well and accurately recording the physical parameter of the earth that we intend to monitor. The goal with this second phase is to determine what the instruments are telling us, what story about the site in question is most likely to be true.

3.1 Understand what is being measured

The map is not the territory (Korzybski, 1933). The sensors we use to monitor the earth, much like our own eyes and ears, convey a representation of the actual physical properties to us (Figure 2). They don't convey the actual physical situation. Vibrating wire piezometers measure pressure, not groundwater elevation. Inclinometers measure the displacement of a casing from its prior location, not ground motion. We need to understand the manner in which the transducer functions, what it is measuring, and how it converts that measurement into an electrical signal that we monitor.

Figure 2: The Treachery of Images by Rene Magritte (1929)

Understanding what is being measured allows us to determine potential ways in which the measurements can be mis-interpreted. For example, a poor inclinometer casing installation can lead to displacements in the casing that are not representative of displacements of the ground. Different installation methods for piezometers can lead to a "sump" of water that is retained after the water level in the unit dips below the instrument.

3.2 Avoid common pitfalls

Don't suggest the installation of more instruments if you don't understand what your current instruments are telling you. If we can't parse and understand the data we are presented with, more data is not helpful. Additional instruments are helpful if you do understand your instruments but seek to learn more about the site. For example, additional piezometers nested in a borehole may reveal gradients that explain the incongruence of a given piezometer with the phreatic surface. Extra prisms may reveal that lateral displacement was in fact twist.

Check any story with a sanity check. Avoid extremely implausible suggestions. It's unlikely a manual inclinometer probe stopped worked in one casing, stopped working for the second casing, and then worked again in the third casing.

Check the scale of the plots. Understand the accuracy and precision of the instrument. I have personally spent weeks of my time analysing displacement that turned out to be nothing more than the inaccuracies of the instrumentation. The accuracy levels reported for most instruments by the manufacturer are laboratory accuracy levels. They are not representative of the level of accuracy to expect in the field. New instruments, especially, suffer from a lack of published data on field accuracies. The field of Geotechnical Instrumentation and Monitoring would benefit if there were more publications on actual field accuracies achieved during monitoring and fewer breathless case studies on how amazingly a site was monitored by the authors employer.

3.3 Plot, plot, and re-plot your data

To analyse data properly, we need to be able to create a large variety of plots in a minimal amount of time. The best data analysis comes when the data can be plotted and re-plotted in different ways, against different variables. Good data analysis follows fast plotting capabilities.

There are two main ways to produce a wide variety of plots quickly, cleanly, and accurately. The first is to invest in personal knowledge and learn a data science programming language like Python or R. These languages allow individuals to query, process, and plot large data sets. These programs give the users complete control of the data visualizations. The second is to invest in a powerful data visualization software package like Tableau or Power BI. These programs take less long to learn, but limit the users control of the data visualization.

Begin by comparing the data with respect to other variables at the site. If this does not explain the data, investigate alternative publicly available sources of data that record variables at the site. Tide gauges, river levels, barometric pressure, and rainfall are usually available on publicly run servers. It can be appropriate to create plots for internal review, zoomed in or showing small sections of the data.

When you do create plots for submittal to the stakeholders, take the time to create beautiful plots of the data you have. To quote Edward Tufte "Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space" (Tufte, 1985).

3.4 Posit your counterfactuals. Create zones of truth.

Something must be correct. The data is coming from somewhere. There is one story that is true about the data coming from this instrument, and it's our job to get as close to that story as we can. If we are arguing that an initial story is wrong, we need to simultaneously argue that another story is correct. We can't just criticize an incorrect interpretation. Analyse the probabilities of the potential counterfactual situations that could be true in this case. Fundamentally, some story about the data is correct.

Don't be tempted to discount data you think is correct because it's convenient for the interested parties. Once you've figured out the most likely true story of the instrument, take your time to document how you came to this conclusion, and convey this solution to the interested parties in an accurate story.

4. Conclusions

This paper outlined a method for analysing cold geotechnical engineering data. These steps will not all be applicable for every instrument, site, or job. However, having a framework for the analysis of this data can aid in the review and understanding of new data from various sites. It is my hope that this paper assists other practitioners effectively review data.

In my experience, it is often easier to sell fifty thousand dollars of instrumentation purchases and installation fees than twenty thousand dollars of an expert's time. The fifty thousand dollars of instrumentation only increase the confusion if there is a not a firm understanding of the existing instrumentation onsite.

Acknowledgements

I would like to acknowledge Matthew Walker, of Barr Engineering, who provided much needed review and insight of this paper. I would also like to acknowledge Rob Clark, of Geokon, who taught me most of what I know about instrumentation. The writings of Edward Tufte and Jake VanderPlas were formative in my understanding of data presentation and data analysis. I would like to thank my wife, Megan, for an initial read through of this paper and just generally being an excellent wife.

References

Korzybski, A. (1933), A Non-Aristotelian System and its Necessity for Rigour in Mathematics and Physics 'Science and sanity: an introduction to non-Aristotelian systems and general semantics' , Lancaster, .

Tufte, E.R. (1985). The Visual display of quantitative information. Cheshire, Conn.: Graphics Press.

Tufte, E.R. (2020). Seeing with fresh eyes meaning, space, data, truth, Cheshire, Conn: Graphics Press.

Magritte, R. (1929). The Treachery of Images.