DATA: Use It; Don't Lose It

Sharon Jones, Demand Management Institute
Jennifer Atlee, Demand Management Institute

ABSTRACT

Program managers and analysts alike must understand that data is imperfect. Analysts must know how to work with incomplete and inconsistent data to perform accurate energy studies, whereas program managers need to understand data in order to critically review information and make well-reasoned decisions.

This paper discusses the following principles for good data and successful analysis: take more data than necessary, mine the data and learn from inconsistencies, check for sufficient accuracy and precision, and use consistent methodology.

The paper describes several cases of good and bad data sources, how to deal with bad data, and successful use of good data. It also briefly discusses the value of good data, stating that it is worth spending 10% of project cost to verify data quality, because the trustworthiness of quantitative analysis depends on the reliability of the data especially if cost effectiveness is questionable. Understanding data will aid the responsible analyst and encourage program managers to ask critical questions and get a second opinion.

Introduction

In our roles as energy efficiency analysts, both as proponents and evaluators, we are routinely looking for data. We have been doing energy studies in industrial plants for almost twenty years combined. Few plants have all the data we could want to do a thorough analysis. At DMI we do our work with a combination of historical data, hearsay, and spot-measurements. Data is rarely ideal, so those who use it must know how to tell good data from bad, how to milk data for unexpected variables as well as desired information, and how to determine whether there is sufficient data to perform an analysis with a reasonable degree of accuracy. Those who rely on others for information also need to understand data in order to judge the quality and validity of the information with which they are presented.

In this paper we will discuss general principles for good data and successful analysis and we will briefly describe several cases in which data was unavailable, unreliable, or inaccurate with suggestions of how to mitigate such situations. We will also relate a success story in voluminous, detailed and available data, when analyzed, showed the path to detail where significant energy savings with no capital cost.

General Principles

Data can be used for long-term process improvement as well as for real-time process control. It is frustrating to see effort and expense devoted to collecting data that for one reason or another cannot be used for analysis. Ideally, data should be accurate, be stored in a form that can be retrieved later for analysis, and be reviewed and analyzed periodically.

Data is rarely ideal. As consultants we frequently have to work with incomplete and inaccurate data. The following suggestions are based on our experiences working with real-world data and data collection.

1) Take more data than is required. It will not all agree with itself so choose which data to believe. Having excess data will facilitate double-checking of energy savings estimates, and may indicate possible problems and complicating variables. Facilities benefit from digitally recording process data, and digitally recorded data can sometimes be obtained from facility computers.

2) Learn to mine the data to get the information needed. If the required data is not readily available find it by triangulating using other data. "The devil is in the details" is a cliché because it is applicable. Become a detective. Inconsistencies in data can be a gold mine of information about process peculiarities and variables that will affect savings. Inconsistencies may also suggest alternate savings options.

3) All the data in the world is of no help if it is inaccurate, or of too poor resolution for use in calculations. Check the gauges being used, make sure they are calibrated, and keep an eye open for inconsistencies. Standard temperature gauges are only good to ± 2°F unless high accuracy gauges are purchased. Temperature gauges with ±1° or ±1/2° accuracy are available but are usually expensive. Pressure gauges go bad quickly. A test to determine whether a pressure gauge is operating correctly is to read two gauges and switch their locations to see if the readings are duplicated. Don't rely on the site's gauges; bring a pressure gauge to the site. Liquid-filled pressure gauges maintain their calibration longer but are not as accurate as non-liquid-filled.

4) Don't confuse precision with accuracy: "oh it's pretty accurate – it's digital". It is easier to read the precision on a digital gauge; it is not necessarily more accurate. Digital gauges can provide a very precise but wrong answer.

5) Pay attention to scale. Subtracting two reasonably accurate large numbers from one another can result in a highly questionable small number. For example, a 2% margin of error on the total energy use of a pump system may be small, but if the potential energy savings is 4% of the total energy use, the possibility of having zero savings is within that margin of error. If subtraction is used, it is desirable to use the same method for both large number calculations in order to minimize the effect of unaccounted variables. Calculating savings directly, rather than using subtraction, ensures a consistent method.

Examples of Unavailable/Inaccurate Data

The following examples illustrate ways of recognizing and, where possible, mitigating situations where data is unavailable, unreliable, or inaccurate.

Poor Quality Data

DMI was contracted to review the savings of a chilled water system retrofit at a plastics manufacturing plant. To perform this analysis we needed to determine the chiller load. One way to calculate load is to use the difference between the entering and leaving chilled water temperatures. At this site the operator manually recorded the entering and leaving chilled water temperatures from a stick thermometer having an accuracy of ± 1°F and low resolution. Assuming a visual operator error of 1°F and a meter accuracy of ± 1°F for both entering and leaving temperature results in an error of ± 4°F. The entering and leaving chilled water temperature differential was 10°F when the unit was fully loaded. An error of ± 4°F could result in a misestimate equaling 40% of the total possible load. The poor resolution of the data rendered it useless for analysis. In this case we used chiller amperes to estimate the load.

Inaccessible Data of Poor Quality

DMI was contracted to evaluate the energy savings resulting from meeting a facility's additional cooling load with a chilled water plant expansion instead of DX units. To determine savings we needed to estimate the cooling load based on flow and temperature of the chilled water through a condenser coil. Although this data was available, it was in graphical form only and there was no storage and retrieval mechanism. While this may have been effective for real time process control, substantial work was needed to extract data for analysis. Acquiring the necessary information required printing 22 pages of graphs daily for six months. After all that, the temperature accuracy was too poor for load calculations (Figure 1). We printed 4,000 pages, only to throw them away as useless. Think about the level of accuracy required for analysis before putting extensive effort into data collection.

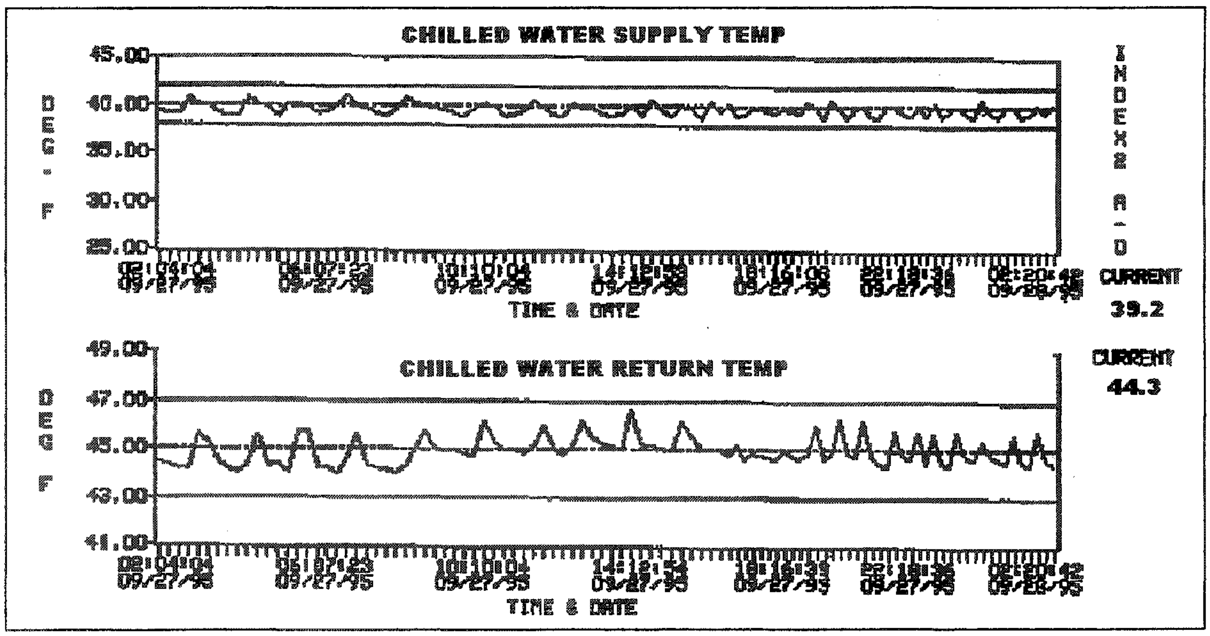

Figure 1. Graphical Temperature Data

Assumptions and Inconsistencies

A flexible circuit board manufacturer uses 156,000 gal/day of city water for a temperature-sensitive plating process. DMI was investigating energy savings that could be gained by using city water destined for the warm water system (100°F) to precool 60°F chilled water (Figure 2). Savings depended on entering city water temperature of 57°F or less. While this assumption seemed reasonable to us and to the plant manager, an operator argued that he had seen the entering city water temperature rise above 80°F. The plant manager vigorously disagreed, explaining that the water was supplied from a river where he had successfully fished, and fish couldn't live in 80°F water. There was no record of temperature, and we were performing the study in the middle of winter. To resolve the issue we called the city water supply company, which recorded water temperature once a week. It turned out that city water from the river at times exceeded 80°F, and was higher than 57°F for most of the summer. If we had not investigated the inconsistent responses of plant personnel by tracking down the data from an alternate source we could have drastically overestimated the savings potential.

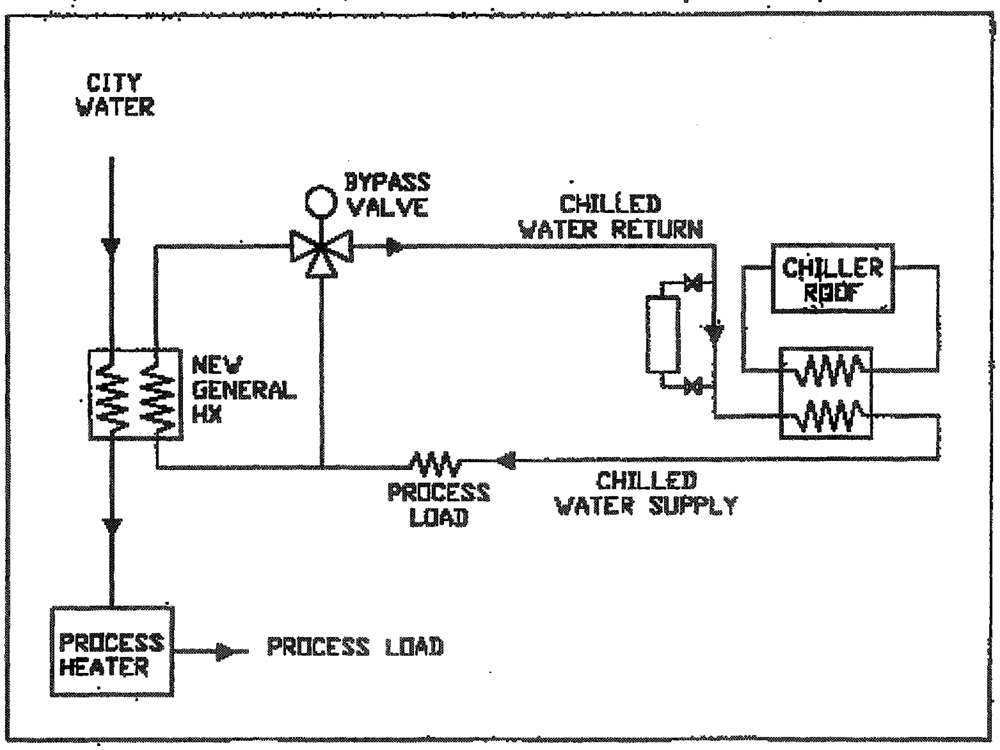

Figure 2. Schematic of Heat Recovery System

Minimum Data Requirements

A fabric manufacturing plant had a process in which dye was forced at high pressure through rolls of fabric. System pressure was generated by constant speed pumps and manually controlled by throttling valves. We were to determine the potential energy savings for installing variable frequency drives (VFDs) on the pumps thus saving energy by generating only as much pressure as required by the process. To estimate savings we needed to know the flow through the pump and the pressure drop across the valve.

There were no operational gauges to measure pressure drop across the valve directly, however the pressure drop across the valve could also be calculated from flow and valve position using the Cv ratio (a characteristic of the valve). There was a pressure gauge that indicated the pressure input to the pneumatic throttling valve (corresponding to position), however its value was not observed or recorded.

We could estimate flow through the pump at the time of the visit from our metered power draw and the pump curve. However this estimate depends on an assumed motor efficiency, and the assumption that the pump is performing to specification. In addition, due to the lack of records we could not determine whether that flow was representative.

The savings estimate was very sensitive to both valve position and flow. Assuming a flow of 1,000 gpm, a valve position of 75% open would result in negative savings, due to the energy penalty of the VFD, whereas a valve position of 25% open would result in savings of 25% of the pre-retrofit energy use. Assuming a valve position of 25% open, a flow of 400 gpm vs. 1,000 gpm would also result in savings ranging from negative to 25%. The site could either install considerable monitoring equipment or make the VFD decision without any assurance that there would be savings. The vendor of the dye machines claimed the new system could provide an average savings of 30%, but did not explain the basis of the claim. The site decided to collect more data.

The Value of a Consistent Method & a Sleuth's Eye

We were contracted to review the actual savings of a previously installed VFD retrofit at a wastewater pumping station. The original analysis, conducted pre-retrofit, estimated energy usage by the post-retrofit system using a bin analysis based on assumptions of usage, nominal head of the pump, static head in the system, and pump efficiency. The analyst calculated the energy use of the pre-retrofit system from six months of billing data, assuming that the pump was responsible for all of the energy use. The methods used for both the pre- and post- retrofit calculations were reasonable but inconsistent: the former used data while the latter relied entirely on assumptions. We wanted to use a consistent methodology in our review to minimize the possible effects of unaccounted variables. Our options were to review what the theoretical analysis showed for the base case, or collect electrical billing data for the installed case. Working with the rule that more data is better, we did both and metered the power input to the pump as well.

Our metering data for the pump did not correspond to the billing data for the same time period. Through interviews of site personnel we discovered there were electric heaters in the pump house.

Once we determined that electric heat was an issue, we used the billing data in conjunction with historical weather data from NOAA and flow data from the site to do a 2-variable regression on both heating degree-days and flow. We ran one pre-retrofit and one post-retrofit regression and calculated the annual energy use in each case as the sum of products of the flow for each billing period and the flow coefficient of the regression. Our analysis showed that actual savings were only 11% of the original estimate.

Neglecting the heaters was an easy mistake to make. The real underlying error was using inconsistent methodology so that the heating energy was essentially added to the savings resulting in a drastic overestimate. We discovered this error by insisting on consistent methodology and investigating the discrepancy we found in the data. The original analyst did not have the benefit of post-retrofit metering data, however using a theoretical analysis for the pre-retrofit case as well as the post-retrofit case would have been methodologically consistent and would have been less likely to severely overestimate savings. In addition, the analyst could have discovered the heaters by comparing the pre-retrofit theoretical analysis with the billing data.

A Warehouse of Data in Need of Analysis – The Best of All Worlds

DMI was contracted to determine the potential energy savings at a large wastewater treatment plant that could be gained by installing VFDs on the blowers and by maintaining the dissolved oxygen (DO) setpoint in the aeration tanks at 1.0 mg/l. This unusual plant had a plethora of data: air flow meters for every tank, amperage measurement and inlet valve position on every blower, DO in each tank, total plant flow, BOD in and out of the aeration tanks, and influent temperature. For once we had all the data we could possibly want.

There were automatic controls available for the blowers, but operators were uncomfortable with the control scheme and turned blowers on and off manually. Manual control rarely leads to optimal efficiency since operators adjust controls less frequently than an automated system. Because several hours could elapse between the time when the DO level dropped below the setpoint and the time when the operators started another blower, the operators bought time by programming in a high DO setpoint (3.0 mg/l). However the average DO concentration for the aeration tanks was well below that setpoint (1.9 mg/l), and the DO ranged between 0.1 mg/l during peak loading during the day up to 9 mg/l at night (Figure 3). The relationship between DO and energy use is not a linear function; a slightly higher DO level requires a substantially greater amount of energy.

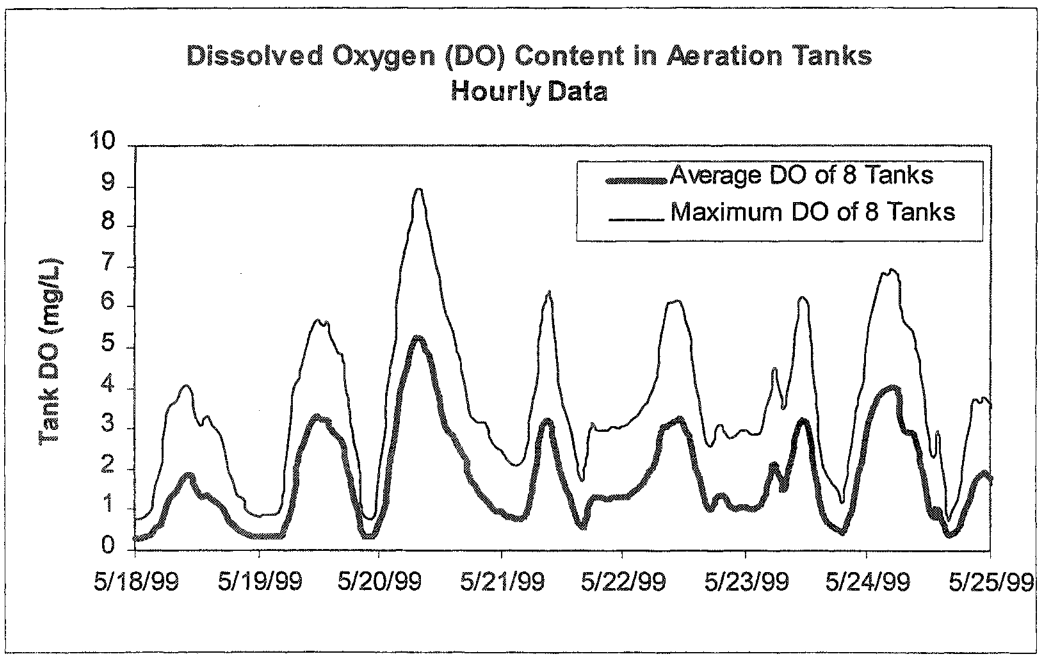

Figure 3. Hourly DO Level in Aeration Tanks

We analyzed six months' data (over 200,000 data points) and found that, in addition to the savings for VFDs ($21,000/yr) and better DO control ($28,000/yr), the plant could save $54,000/yr simply by running fewer blowers. Running fewer blowers at a greater load would also reduce the risk of blower surge, something the plant operators were keenly trying to avoid and their primary reasoning for manual control.

Data was invaluable in the analyses of these measures. More remarkable however, the data itself led us to discover an additional measure that could save $30,000/yr or more for almost no cost.

When a tank is taken off-line a small amount of clean water is put in the off-line tank to protect the dome diffusers from the weather. A small amount of airflow is required to prevent the diffusers from plugging up. Initially we had ignored the data showing air flow to the off-line tanks, assuming that it was noise or trouble with the metering equipment. When the total energy use calculated for the pre-retrofit condition differed substantially from pre-retrofit billing data we took a second look and discovered that one third of the total airflow went to off-line tanks.

Because the off-line tanks had a lower water level, the resulting backpressure was lower than in on-line tanks. Even with the valves pinched down, the air flowed toward that lower pressure. While some air is deemed necessary in off-line tanks to prevent the diffusers from clogging, a flow equal to or greater than that going to on-line tanks (Figure 4) is clearly excessive. Our conservative estimate was that an average of 2,000 cfm of air, costing $30,000/yr or more, was wasted supplying excess air to off-line tanks.

Without the large amount of useful data in digital form, we could not have quantified this additional load or justified efforts to reduce it. We only discovered the added load by investigating inconsistencies within the data.

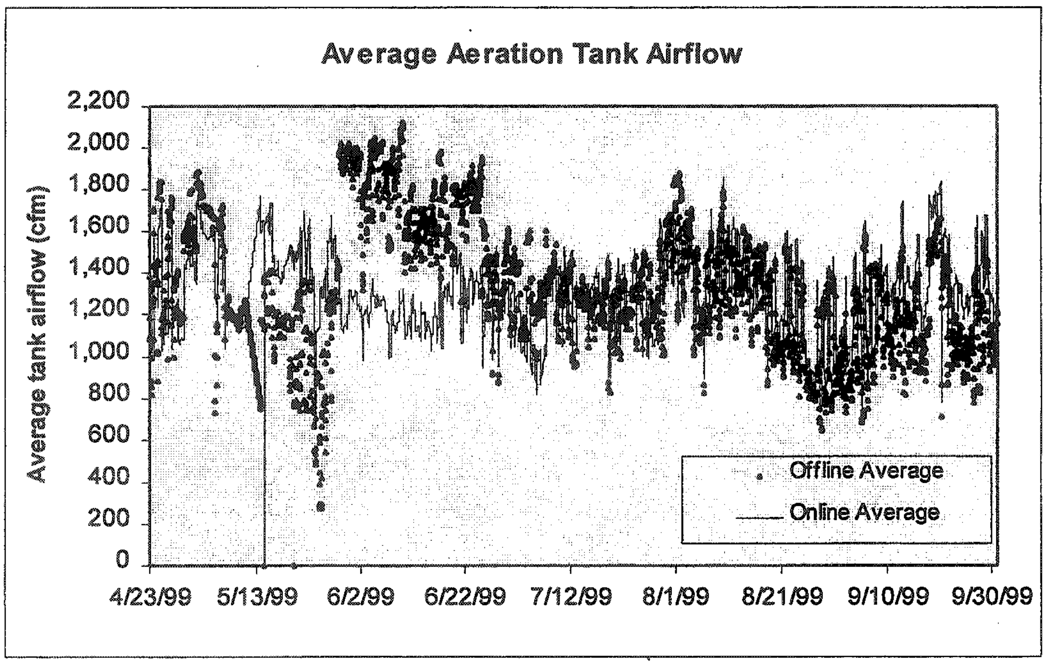

Figure 4. Average Airflow to an Aeration Tank

The Danger of Generalizations

We have described an example of how data can show the way to additional savings for a particular site. We also see numerous examples of unrealized savings that were estimated using a general rule and not the particular situation. Rules of thumb are only as good as the data behind them.

We have an easy-to-use chart from a VFD vendor that indicates pump energy savings from using a VFD based on the percent flow. While the basic principles are correct, and are the same as we use in our analysis, for individual applications the chart is correct as often as a stopped clock. If the application is an open pumping system, or if there is significant static head there might not be any savings. Depending on the specifics of the site, savings are often much less than the simple rule would predict.

Vendors may claim savings for an advanced model over the standard design without specifying on what conditions those savings are based. It is not uncommon for advertised savings to be based on continuous operation and 100% load. Vendor estimates may be unrealistic in other ways. One client commissioned us to create a baseline for air compressor costs and efficiency in order to keep vendors from artificially increasing savings by using an exceptionally inefficient model for comparison.

What is Data Worth?

A person who does not deal with data directly will still deal with the results of another's analysis. Knowing the quality of data on which that analysis is based is part of understanding how trustworthy it is.

Clients buy DMI's services because site-specific details, elucidated through data, determine whether a measure will save energy. It is certainly worthwhile to spend up to 10% of the project cost to make sure the data is good, because the trustworthiness of quantitative analysis depends on the reliability of the data.

The need for good data and accurate analysis is most obvious if the decision of whether the efficiency measure is cost effective is in the balance. There are times when the data will show that a measure, which looked great at first, is not cost effective or is even less efficient than the existing situation. In instances like the waste water treatment facility discussed above, the availability of good data and resulting good analysis led to substantial savings in an unexpected area.

Conclusion

Knowing how to work with data and to critically review results is vital for energy analysts and program managers alike. We have described techniques for working with imperfect data. These methods will aid the responsible analyst in the assessment and quantification of savings associated with energy conservation measures. An awareness of data-related issues will encourage program managers to review a study's underlying assumptions before its measures are adopted. Program managers may not have the resources to carefully check data sources; however they can ask critical questions about the analysis and get a skilled analyst to do a back of the envelope calculation to provide a reasoned second opinion. Sound decisions are based on trustworthy information, and trust must be earned.