LeonardRice ENGINEERS, INC.

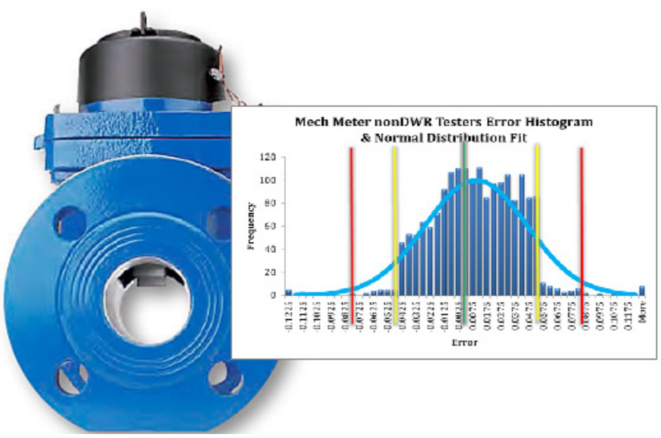
Statistical Analysis and Programmatic Evaluation of the Division 3 Well Measurement Program
FINAL REPORT

Prepared for:

Colorado Division of Water Resources
and
Rio Grande Decision Support System Technical Advisory Committee

May 30, 2012

Project # 1401CWR01



The following members of the Leonard Rice Engineers, Inc. staff contributed to the preparation of this report.

Project Engineer Jim Brannon

2000 Clay Street, Suite 300 | Denver, Colorado 80211
303-455-9589 | www.LREwater.com

Table of Contents

Table of Figures

1 INTRODUCTION

*"In order for the State Engineer and Division Engineer for Water Division No. 3 to obtain information needed for administration of the waters in Water Division No. 3 (the Rio Grande Basin within the State of Colorado) and to assist in compliance with the Rio Grande Compact, it is necessary to adopt rules governing the measurement of ground water diversions located in Water Division No. 3." [from RULES GOVERNING THE MEASUREMENT OF GROUND WATER DIVERSIONS LOCATED IN WATER DIVISION NO. 3, THE RIO GRANDE BASIN
in 2005]*

Colorado Division of Water Resources' (DWR) Division 3 has implemented the afore mentioned rules in the Rio Grande River basin. They require wells to be metered (with some exceptions described in the rules, available at:
http://water.state.co.us/DWRIPub/Documents/div3measurementrules.pdf)
and require well owners to submit total annual pumped volume to Division 3 each year. This metering and reporting requirement has generated a lot of new activity in the basin, both for water users and the State. The purpose of this feasibility-level study is to provide an objective analysis of the well pumping data collection program. This will be accomplished by evaluating the data collection and management systems that have been developed by and for Division 3, and by analyses of the data already collected as part of the program.

An overview and evaluation of the data collection and data management systems follows in sections 2 and 3, respectively. Section 4 contains data analyses of aspects of the collected data: an analysis of the complete data set relating to completeness, a comparison of aggregate meter totals to well pumping totals, and several analyses of the meter testing data. Section 5 summarizes the project's general conclusions and recommendations.

Data analysis, from classical statistics to free-form exploratory data analysis (EDA), is about observing the behavior of the systems in the world around us via the data we deem represents them, and thereby gaining understanding. The patterns, or the lack of patterns, tell us how they work, how they are structured internally, or what rules they follow. And, when the patterns do not match what we expected, it tells us something about ourselves, too, which also changes our understanding.

2 EVALUATE DIVISION 3 WELL PUMPING DATA COLLECTION PROGRAM

2.1 DESCRIPTION

This section describes and evaluates "what" comprises the DWR Division 3 well pumping measurement program.

2.1.1 Background

Division 3 was fortunate to have the experience and results of the Division 2 well data collection effort to learn from when designing its own program. Division 2 staff experience and documentation, like Burbidge-Thompson well inspection inventory protocol (1994), were invaluable resources when developing the Division 3 program.

The program began in Division 3 with a thorough field well inspection and inventory, section by section, starting with the information in existing inventories of wells (electronic and paper) located in files and databases located in Division 3 and in Denver and then augmenting it with the collected new information. The various inventories complement each other. For example, it was useful to start the field inventory with a well list from a DWR well database download. And, conversely, information found in the field both verified data in the DWR databases and also identified data to be corrected. It should be noted, however, that integrating multiple inventories of wells and resolving inconsistencies creates an on-going process as long as new data continues to be collected.

Next, the wells in the inventory were categorized so the appropriate rules and data collection processes can be applied. Exempt, active, and inactive wells all have special rules documented in the link above in Section 1, and in other documents on DWR's website. For example, active wells in Division 3 now require sufficient and approved measurement equipment to be installed so that annual pumping volumes can be calculated. Totalizing flow meters (TFM) of various kinds are frequently used, though some power conversion coefficient (PCC) equipment still exists. The specific rules for the installation, reporting, and testing of these meters, are also contained in these documents.

2.1.2 Data collection and processing

Measured flow volumes for each well pumping meter are required to be submitted to Division 3 each year. But the relationship between meters and wells is not always "one to one". Often a single meter is connected to multiple wells or multiple meters are connected to one or more wells. Therefore more information from the field is required before individual well pumping volumes can be calculated. At the end of each irrigation season the annual flow volumes for meters are converted to annual pumping volumes for wells using the system piping and connection information that the well owners also supply to Division 3.

An annual flow volume is required each year for each meter. Therefore a minimum of one meter reading is required each year, though often multiple readings are reported during a year. For example, meter testing that occurs when metering systems are inspected, updated, changed, or repaired generates intermittent TFM measurements that are also reported. These data are all entered into the GWDMS (groundwater data management system) database by Division 3. Denver DWR staff use algorithms coded into SQL scripts in the GWDMS SQL Server database to combine the data into a single annual volume per meter.

Meter readings are entered into the GWDMS throughout the year. At the end of each year, the raw meter readings data need to be reviewed for quality (QA/QC), evaluated, annotated and corrected. It is an iterative process that eventually yields a data set ready to be processed into structure (well) pumping values. Some of these activities take place inside the SQL Server database system, where DWR database staff run queries developed for this purpose.

Other of the necessary QA/QC and data processing activities require Division 3 and field staff expertise - for example, using information about how the wells and meters are connected to convert the meter readings into well pumping values. At the time of the writing of this report this information is not completely stored within the GWDMS. Therefore, data from the GWDMS is exported from the database into MS Excel workbook tables for combining with field data and further editing and analysis.

This latter exercise is tedious, vulnerable to errors, and difficult to document. It requires combining database meter reading volumes with the "on the ground realities" of how pipes and wells and meters and delivery systems are connected. Real world systems are often more complex than the database structure can accommodate, and they are constantly changing. As a result, judgments and special circumstances are often made in the calculations, record by record. These require more documentation in the form of notes added to the data processing workbooks.

Finally, once the annual well pumping values have been calculated at Division 3, they are sent back to DWR in Denver contained in a spreadsheet and are imported into HydroBase (not GWDMS). Then the process starts over again for the next year.

2.2 EVALUATION

The Division 3 office staff have developed and adopted thorough and rigorous business logic rules and processes to aid their data processing activities. To the extent possible these are designed to maximize the reliability, redundancy and accuracy of the data that is recorded and eventually stored into the GWDMS database. The data collection and reporting policies, forms, and processes have evolved and Division 3 continues to adapt them to new situations and challenges that arise.

The DWR database staff members in Denver are adept at designing and running sophisticated database and data processing systems for water resources related information. They designed and built the GWDMS database and forms that streamline the data collection process and make it more efficient. The sheer numbers of wells and meters, much less the increasing complexity of the systems and operations, make server-based database systems a requirement for this kind of effort.

The difficulty, though, of integrating "database centered" systems into real world operations cannot be overstated. Sometimes the necessary data structures are either more complex than many users want to deal with in day to day operations, or else do not handle every variation of data or circumstances that real world data present. Discussions with the DWR database staff have indicated that they certainly understand this and endeavor to build systems that are useful for the Division 3 users.

Most of us use database centered web systems like this in our daily lives on computers at home and work. It is a common experience that over time users' needs change, that the data evolves and gets more complex, and that the users' skills with the systems increase, and therefore they "outgrow" the original systems. Therefore it should be expected that updates to the GWDMS systems will be needed on a regular basis.

Such as the meter to well conversion described above, there are places in Division 3 data collection and processing activities where the DWR expertise and systems could be further used to make current Division 3 work easier and more productive. This has to be balanced against the limited DWR database staff resources and budgets. This is the same dilemma that many organizations face. However, as happens frequently these days, the rapid pace of technological advancement in data management provides new ways to bridge that gap in creative and less expensive ways if the organization stays aware and responsive to new technology.

2.3 CONCLUSIONS

Based on the size and complexity of the data collection program, and also the success of the systems already developed, DWR should continue to focus database system design and development to support the needs and workflow of the Division 3 office and field staff. Keeping the communication going between system developers and system users is an arduous process, but it is essential to the long term viability of the data management systems being built. Without that focus, it is easy for separate priorities to take hold, causing a divergence of objectives and eventually the creation of incompatible business processes that create conflict.

A brief review of the Division 3 systems (Excel workbooks) currently used to process meter volumes into well volumes indicates that they would benefit greatly if those data and processes were integrated into the GWDMS database. One example is the metadata describing the well – meter – piping connections that are needed to convert meter volumes into well pumping. This same issue is encountered at other organizations that collect and process large amounts of well data via meters (e.g. Central Colorado Water Conservancy District in Greeley), and they have benefitted from moving these data and processes to a central database server.

One benefit is the standardization it creates. These data get complex and with so many wells to process, each person who tackles the task will naturally do it differently, yielding perplexing data inconsistencies when data are combined again. The upfront cost of moving the data and processing systems onto a database and web server is much less than the perpetual cost of recreating the same logic and tools over and over again, by different people, in different ways. There is also the risk of poor business decisions being made down the road when the needed data are buried in these complex work products that are difficult to understand and utilize.

3 EVALUATE DIVISION 3 WELL PUMPING DATA MANAGEMENT SYSTEMS

3.1 DESCRIPTION

Evaluating data management systems is an auxiliary to the review of the data collection systems in the previous section. This section describes and evaluates more of the "how" and "conceptual technical approach" of the data management systems in place in Division 3 and in Denver at DWR, though some of the same topics overlap.

3.1.1 Physical hardware and software systems

Much of the well pumping data starts its journey on paper forms submitted to Division 3. These data are processed from paper to electronic records in the GWDMS database via custom online forms built for Division 3 by DWR Denver. Staff at Division 3 use desktop PC's connected securely to DWR to access the GWDMS forms for data entry. The GWDMS system in Denver is located on a virtual MS Windows Server hosts running MS SQL Server as the back-end database. Formerly DWR had its own hardware and software systems, but now they are hosted and managed by the Colorado Governor's Office of Information Technology (OIT).

The GWDMS, as well as HydroBase and other DWR data systems, are hosted Microsoft Windows server systems (often called MOSS systems, Microsoft Office, SQL-Server, SharePoint). This is the default and official platform for OIT hosted systems. It uses .NET (library) based languages like V.NET and C# for custom application development.

When GWDMS database data are sent back to Division 3, they are exported to Excel files. This is a familiar and powerful environment for the Division 3 staff doing the data processing. Once the conversion of meter volumes into well pumping volumes is complete, Excel workbooks containing the data are sent back to DWR Denver. The data is then imported into the HydroBase database on the DWR server. This makes the data available to all HydroBase users and DMI's.

3.1.2 Data Management Concepts

Through HydroBase and Colorado's Decision Support Systems' (CDSS) projects over the years, Colorado has been at the forefront of water resources technology for advanced data management designs and systems. "Data centered" designs have been a key component. This experience and these capabilities are a boon to the Division 3 well pumping data collection program.

Because the users of these systems are generally water resources experts, DWR systems often expose the underlying data structures of water resources data to the user. This design allows knowledgeable users the most flexibility and power in their use of the data.

The other side of that coin, though, is that less experienced users and the general public can find the information in the data perplexing. Therefore DWR should continue to refine and build tools like the StateCU Wizard that provide a simplifying, easy to use layer for the occasional user that insulates them from the complexity of the sophisticated data management systems behind the scenes. Continuing to fund the building of more "open standards"-based and web-based data services (restful XML web services, XML-RPC, etc.) will make it easier to create user-friendly interfaces to the DWR's complex databases for the general public.

Evaluation

The hardware and software systems in use at the State are sufficient and appropriate for the systems and purposes it needs. The OIT centralization and standardization of IT systems will yield cost efficiencies, but perhaps inhibit some flexibility and creativity when DWR is developing new custom systems.

The GWDMS data and data processing centralization has helped create a consistent approach for some of the data management processes involved in the Division 3 well data collection program. It also provides a platform for building more of the needed data processing tools in the future. As noted before, more of the Division 3 data processing activities can benefit from integrating with GWDMS.

A key weakness in the current Division 3 data management system is the "chain of custody" of the critical meter and well data. Once data are "normalized" into database structures, it can be dangerous to the integrity of the data to then export the official version of the data to Excel files, which essentially are flat unstructured files. Losing things like the uniqueness of primary keys and the enforcement of foreign key values in important database fields allow bad data to get into the official database, and to remain undetected.

A lesser concern but worth mentioning is that the structure of the current institutional and regulatory rules might be influencing some of the underlying data structure designs, and also the forms and processing tools. This is the fastest way to create systems, for sure. But the cost is it risks a shorter life and utility of the work products being developed for Division 3 and DWR. It can require more extensive overhauls of systems when rules and regulations and uses of the data change down the road.

3.2 CONCLUSIONS

A high priority for future Division 3 well pumping data collection efforts should be to integrate the final processing step (meter volume to well pumping) into the GWDMS systems, and to develop online forms and tools as necessary to keep this critical data inside the database, from cradle (GWDMS) to grave (HydroBase).

Well-designed database systems have been a key strength of DWR over the past years, and DWR management should continue to encourage and give resources toward building those kinds of far-sighted systems (like HydroBase) that outlast quick (current need only) solutions.

4 DIVISION 3 WELL PUMPING DATA ANALYSIS

This section describes the data collection and data analysis activities undertaken in this study.

4.1 DATA COLLECTION

4.1.1 GWDMS database

A copy of the GWDMS database (meter readings and meter test data) including the tables and some of the SQL queries (meter reading data processing) was obtained on January 8, 2012. The SQL Server database tables had been exported to a MS Access file, and the SQL into text files. After reviewing the schema and data, primary keys were re-established to several of the more important tables, and some foreign key relationships rebuilt, too. The database schema is well designed, and understanding the schema made generating queries possible for extracting data sets to use in data analysis activities.

It should be noted that the GWDMS is a continuously evolving database with multiple activities ongoing simultaneously (data entry, QA/QC, etc.). Therefore the January 8 copy was a snapshot, and the 2011 data was still in the QA/QC process. This is useful to understand when analyzing the data it contains, and provided some interesting insights.

4.1.2 Data processing workbooks

Leonard Rice Engineers (LRE) was provided copies of MS Excel workbooks used by DWR and Division 3 staff for well pumping data processing and data analysis. These workbooks provided insight into the data processing steps that Division 3 undertakes, and into the complexity of the data processing issues downstream of the GWDMS meter readings database.

4.2 DATA ANALYSIS

The purpose of this report is to provide an objective analysis of the Division 3 data collection and data management systems (sections 2 and 3 above), but also to analyze the collected data in various ways, the results of which will be presented in this section. The GWDMS database contains several tables comprising meter reading data and also meter test data. The analyses of these data are reported below.

4.2.1 Glossary of Statistical Terms

This section describes statistical concepts, terms and measures that are used in the following sections. [most definitions courtesy of http://www.wikipedia.com]

- *Asymptotic*
  An *asymptote* of a curve is a line such that the distance between the curve and the line approaches zero as they tend to infinity
- *Coherence*
  A property of self-consistency. In this particular case, used to describe data patterns that persist across subsets.
- *Random*
  A *random* process is a sequence of random variables describing a process whose outcomes do not follow a deterministic pattern, but follow an evolution described by *probability distributions*.
- *Probability Distribution*
  In probability theory, a *probability distribution* is a function that describes the probability of a random variable taking certain values.
- *Normal Distribution*
  In probability theory, the *normal* (or Gaussian) *distribution* is a continuous probability distribution that has a bell-shaped probability density function, known as the Gaussian function or informally the bell curve. For example, **the observational error in an experiment** is usually assumed to follow a normal distribution. Note that a normally-distributed variable has a symmetric distribution about its mean.
- *Mean*
  The *mean* is the expected value of a random variable, which is also called the population mean. For a data set, the mean is the sum of the values divided by the number of values.
- *Standard Error*
  The *standard error* is the standard deviation of the sampling distribution of a statistic -- in this case, it is the mean.
- *Median*
  The *median* is described as the numerical value separating the higher half of a sample, a population, or a probability distribution, from the lower half.
- *Standard Deviation*
  The *standard deviation* shows how much variation or "dispersion" exists from the average (mean, or expected value).
- *Kurtosis*
  Kurtosis is a descriptor of the shape of a probability distribution. In probability theory and statistics, *kurtosis* is any measure of the "peakedness" of the probability distribution of a real-valued random variable.
- *Skewness*
  In probability theory and statistics, *skewness* is a measure of the asymmetry of the probability distribution of a real-valued random variable.

- *Confidence Level (95.0%)*
  The meaning of the term "*confidence level*" is that, if confidence intervals are constructed across many separate data analyses of repeated (and possibly different) experiments, the proportion of such intervals that contain the true value of the parameter will approximately match the confidence level. In this case, it means that we are 95% sure that the actual mean is within the +/- the *Confidence Level* of the estimated mean.

4.2.2 Meter Readings Analyses

4.2.2.1 Description

The meter reading data are primarily time series data, but at a very coarse temporal resolution, generally annual, and for a very short period, essentially 5 years. However they exist for thousands of sites (meters) which have spatial and categorical dimensions that provide several areas of useful data exploration.

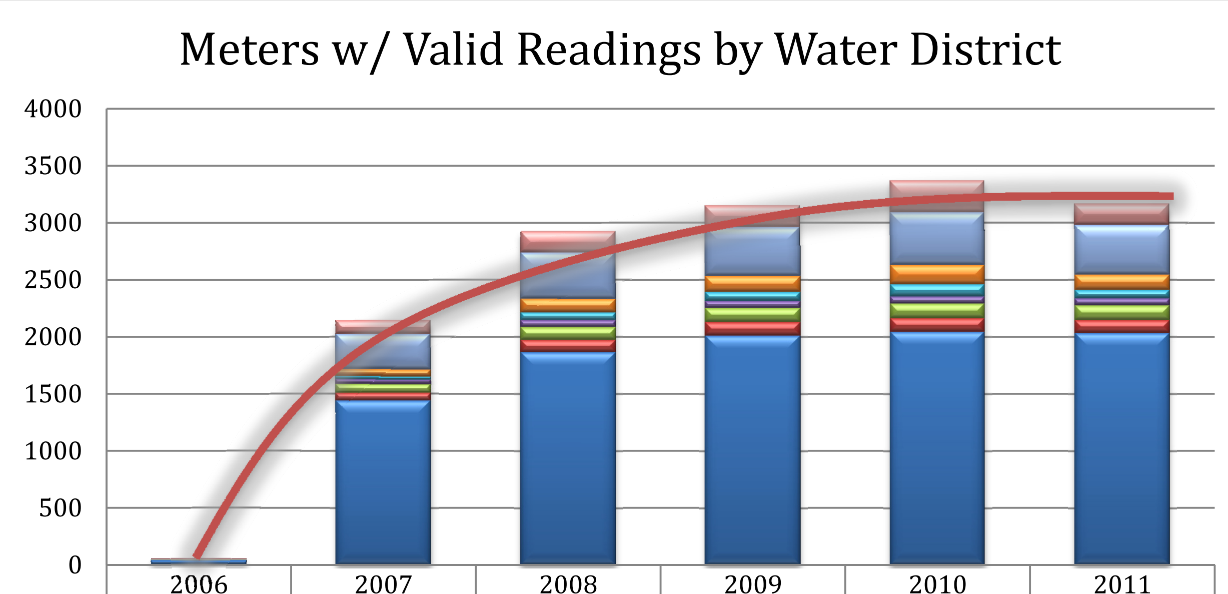
4.2.2.2 Temporal characteristics

The shape of the time series over a collection of sites at an annual time step for such a short period would not usually be expected to yield much information. However, at the beginning of a massive data collection effort such as this, it can yield important information about the program's progress toward "full implementation". In other words, the data can indicate an answer to the question "How close are we getting to collecting ALL the available pumping data?"

If the program has made significant progress and is being aggressively pursued, and the system being measured is naturally constrained, then it might be seen by an asymptotic shape of the aggregate data temporal graph, both in count and total pumping, but especially in count. If it exists in the total data set, then investigating the coherence (similar patterns) of the asymptotic pattern in selected subsets of data (by category or spatially defined) is useful for determining where data collection effort has reached full implementation and where more work may still be needed.

This is, of course, a simplified view of a complex system. Complicating the total pumping evaluation is the reality that pumping demand is actually not constant, and is affected by many factors including climate, commodity prices, groundwater levels, etc. But the resource is not infinite. It is apparent from many studies, like the Rio Grande Decision Support System, that the active and mature agricultural economy in Division 3 may have already reached the limits that the basin's water supplies can sustain over the long term. So it would be expected to find these long term constraints apparent in the temporal patterns.

The following graph demonstrates this pattern exists in the Division 3 Well Metering program data annual time series.



	2006	2007	2008	2009	2010	2011
35	4	121	183	189	275	183
27	4	308	405	427	457	436
26	2	64	121	136	170	131
25		24	66	83	106	77
24	2	43	59	60	61	57
22	3	77	111	123	131	131
21	1	65	112	121	120	114
20	37	1441	1863	2009	2044	2033

Figure 1. Meters w/ Valid Readings by Water District

We will look for this behavior in both the sample (meter) count AND the aggregate meter volumes, as well as in selected subsets. If it exists then the next question to consider is "what is this limit"? Is it the total well count and pumping volume total within the data collection area?

4.2.2.3 Data Filtering

Before beginning the analysis, the meter readings data had to be filtered into two sets, valid and invalid, to remove spurious data that was masking patterns in the legitimate data. As was mentioned before, the database being analyzed is a snapshot from January 8th, and QA/QC of 2011 data is ongoing, so it is not surprising to find meter reading values outside expected bounds in 2011.

A record in the database represents a single TFM meter reading. When compared to the immediately previous reading (if one exists), a total pumping "delta" and total number of days in the period can be calculated. These two values are also used to calculate an "average pumping per day" rate for the period. These three values are used to separate valid from invalid meter readings using the following criteria.

The total pumping value must be >=0 and not null. Meter roll-over events initially create a negative "delta" value until one of the GWDMS data processing scripts finds and corrects them. These records in the 1/8/2012 GWDMS snapshot are ignored.

"Pumping per day" and "days in period" values can also be outside of reasonable bounds. The pumping per day histogram is below. After discussing the maximum possible well pumping rates with Division 3 staff, it was decided that this value should be less than 20 AF day which would correspond roughly to a 5000 gpm well pumping continuously over the whole period.

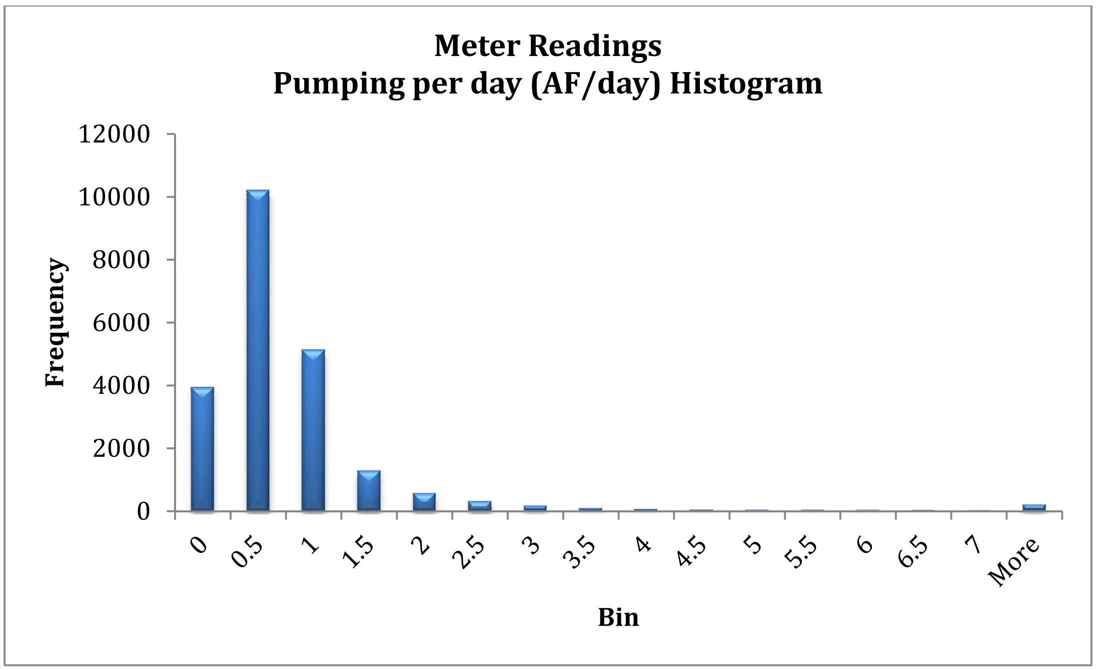


Figure 2. Meter Readings - Pumping per day (AF/day) Histogram

The "days in period" value should theoretically have a maximum value close to 365 since readings are required once per year. However, readings are not always taken in the same month, so a value up to 650 days was considered valid. The histogram for this value is below.

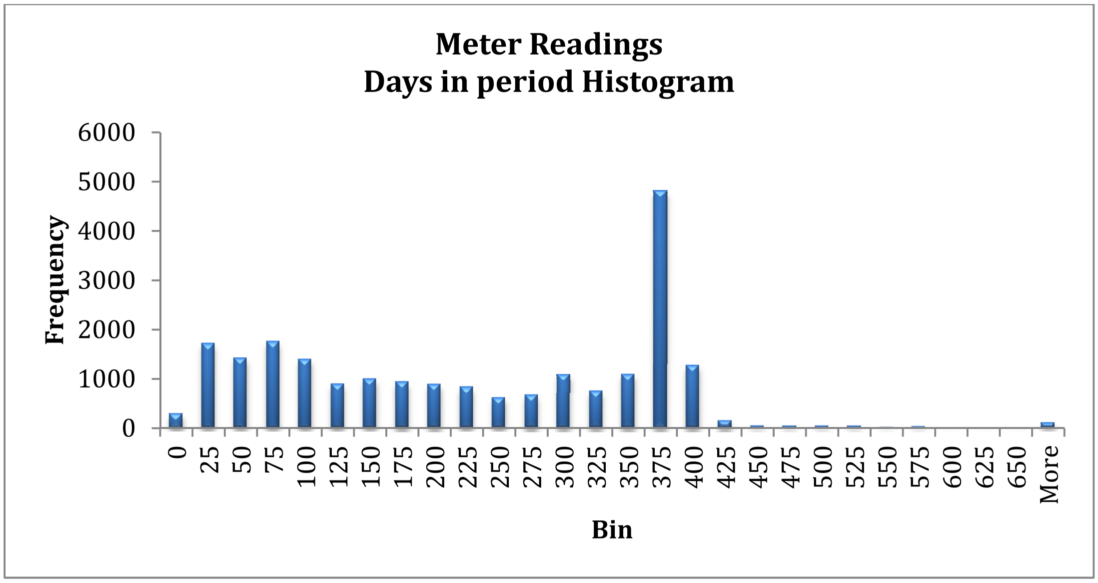


Figure 3. Meter Readings - Days in period Histogram

Using these criteria, the meter readings data was separated into two data sets for analysis. A time series for the invalid meter record count is below:

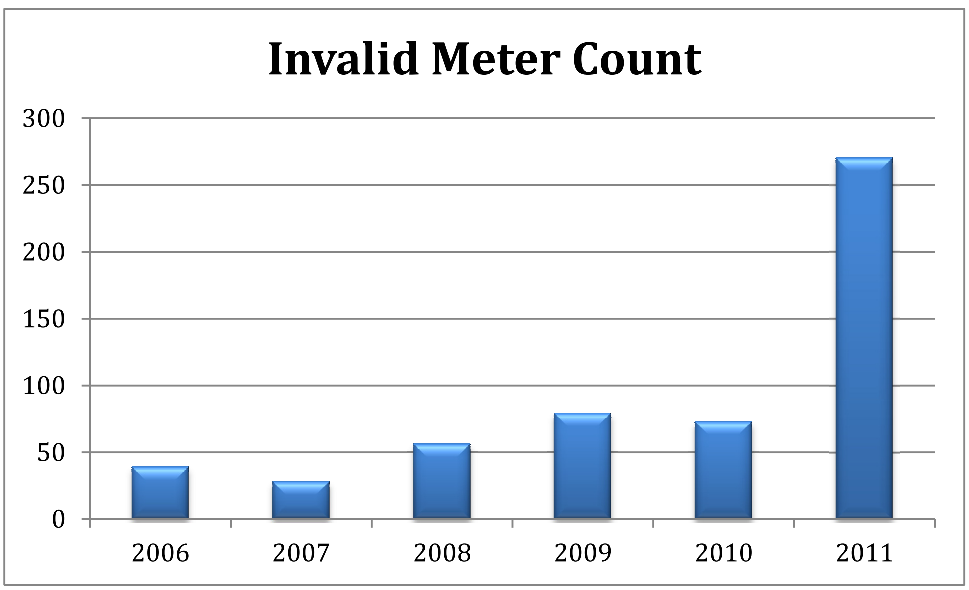


Figure 4. Invalid Meter Readings Count

It can be seen that the 2011 invalid count is much higher than previous years. Based on discussions with the Division 3 and DWR staff, this is because processing of the 2011 data was not fully complete on the date (January 8th) that the GWDMS database snapshot was made for this analysis.

4.2.2.4 Temporal graphs of total data set

Summary time series graphs for total volume pumped (AF) and the number of active meters (with a non-null "delta" value) are below.

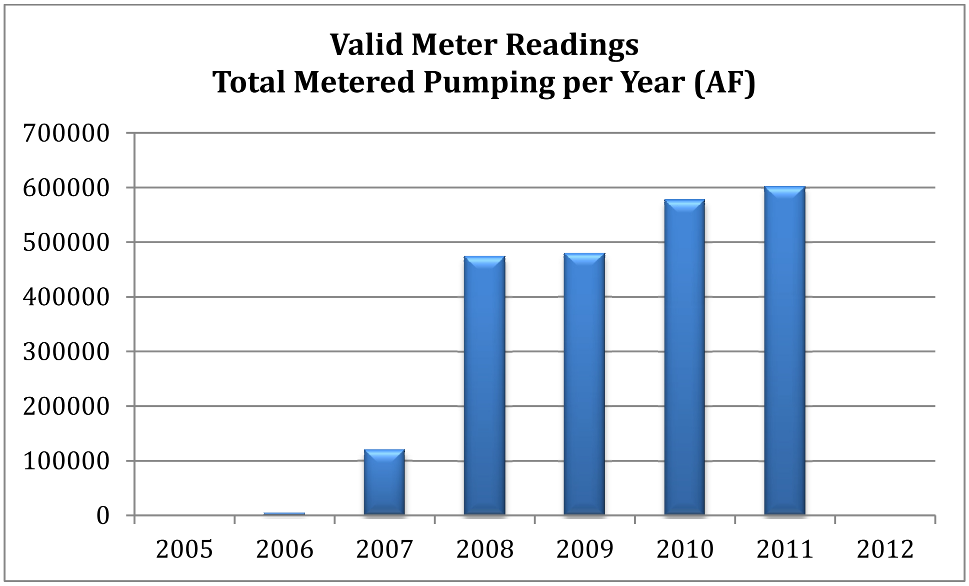


Figure 5. Valid Meter Readings - Total Metered Pumping per Year (AF)

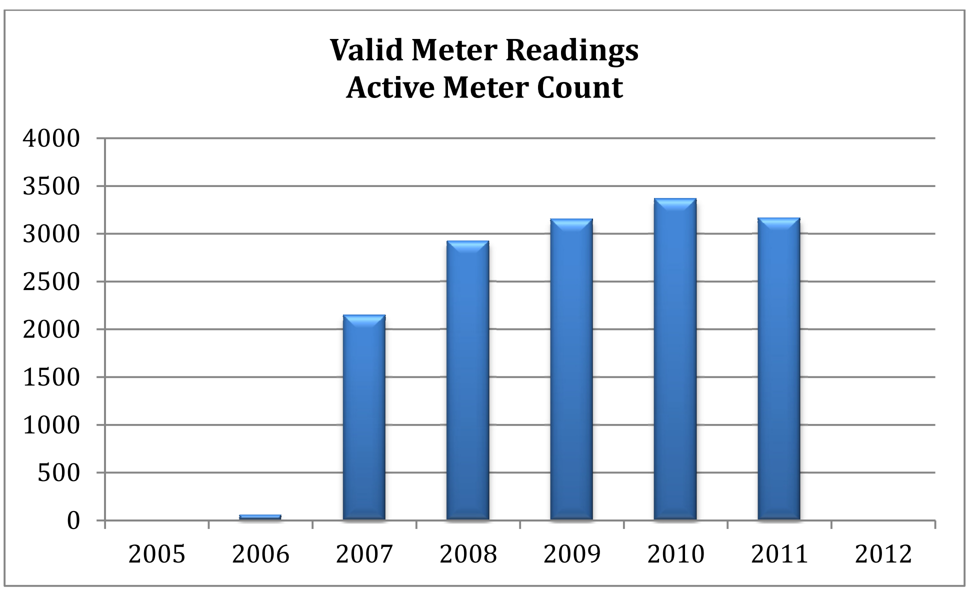


Figure 6. Valid Meter Readings - Active Meter Count

Two things are apparent in these graphs, the quick ramp up of the data collection program from 2006 to 2008, and the asymptotic shape of both of the time series graphs. An exception to the asymptotic shape was the slight drop in valid meter count in 2011. However, as mentioned before, there are still more QA/QC processing steps (meter roll-over corrections, etc.) to be done for the 2011 data. The plot of the "invalid" meter counts above confirms this explanation and reinforces that the asymptotic shape will be maintained when those meters switch back to the valid category.

4.2.2.5 What is the limiting value being approached?

The asymptotic shape discussed before is evident in the aggregate dataset. Since it is indicated in BOTH the meter count AND the pumping volumes over a very large sample set, a hypothesis is that the well count and meter volume total are approaching a threshold constraint or limit which causes the asymptotic behavior. What is this limiting value being approached?

Ideally it would be the total well count and pumping in the data collection area. However, it is possible that the limit being approached is not the total pumping but some fraction of that pumping.

In the case of this metered pumping database, the limiting value is the *total observed measurable pumping*. This should be defined carefully:

1) *total* means simply the sum of all the values (meter volumes) in the set
2) *observed* means these are values observed on the meters. They may not be *actual* pumping if the meters are biased or in error. This will be investigated in a later section of this report using meter test data.
3) *measurable* means that it is a well that can be discovered and metered. We have to allow for the chance that it is possible that there exist pumping wells that can never be discovered or measured.

Consider the following reasonable assumptions:

1) Uncertainty about water use activity in the basin supplied by pumping has shrunk significantly in the last two decades (and especially the last decade) because:
    a) Remote sensing (aerial, satellite) technology capabilities.
    b) Easier access to remote sensing data.
    c) Prevalent mobile communication and photography devices with GPS technology.
    d) Public awareness and concern of water rights and environmental issues make people more observant.

2) Division 3 is aggressively implementing and enforcing the promulgated rules to meter all known (non exempt) pumping activities. Voluntary cooperation as well as advertisement of enforcement penalties applied to violators has resulted in water user compliance with the rules.

3) Total aggregate pumping (in terms of basin-wide well count and demand) has not changing significantly between 2008-2011.

If these assumptions are valid (and if the meter test analysis shows that the meters are accurate in aggregate), then the measurable pumping is very close to the actual pumping. Therefore the actual Division 3 total pumping is the limiting value being approached by the collected meter data.

The following section investigates how "coherent" (persistent) the asymptotic pattern is in various subsets of the data.

4.2.2.6 Additional Temporal Graph Analyses

The larger set can be broken into significant and independent subsets using spatial (water district and sub-district meter collections) and categorical (meter size, etc.) attributes of the meters. If the asymptotic behavior still is evident, especially in subsets with very little invalid meter data to create uncertainty, this gives the theory additional credence, and can also indicate where data collection activities are more or less successful in terms of completeness.

Database queries were written to create these data subsets for aggregate pumping and meter count time series. Below are important examples. Additional data will be in Appendix A.

4.2.2.6.1 Sub-district 1 Wells

First is a spatially defined subset, the "sub-district 1" wells. The valid meters were associated with their connected structures (wells) and then joined to the structure table where this attribute is stored. Below are time series graphs for meters inside and outside this group.

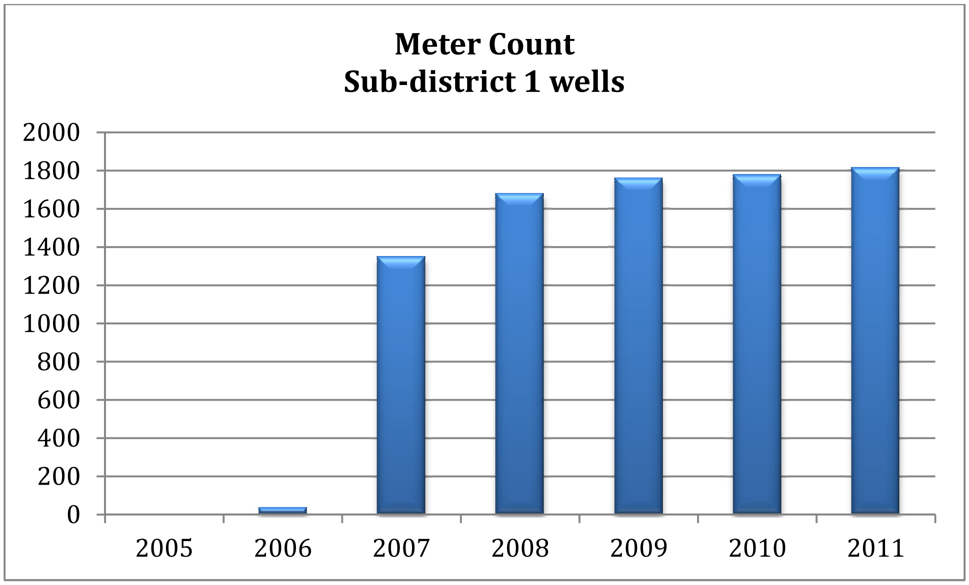


Figure 7. Meter Count - Sub-district 1 wells

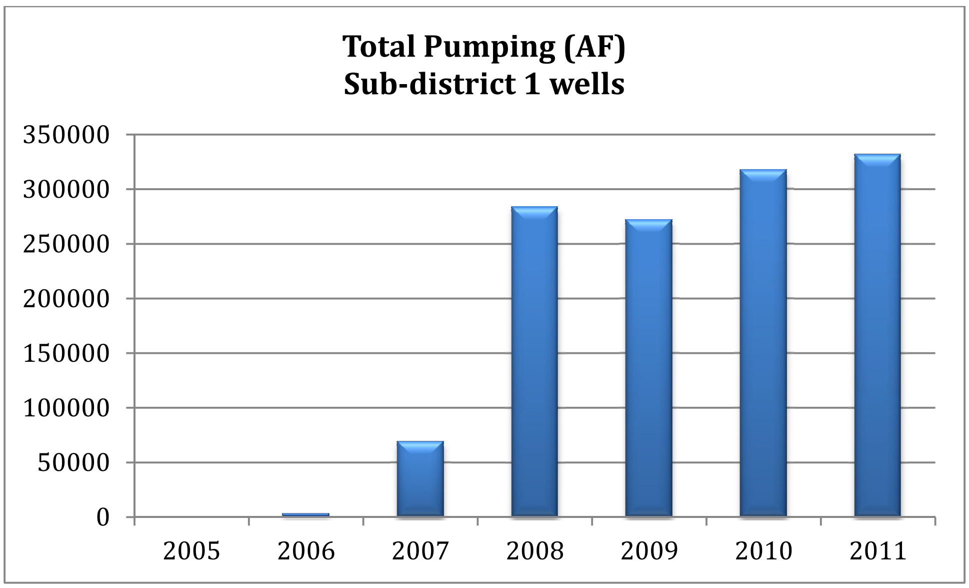


Figure 8. Total Pumping (AF) - Sub-district 1 wells

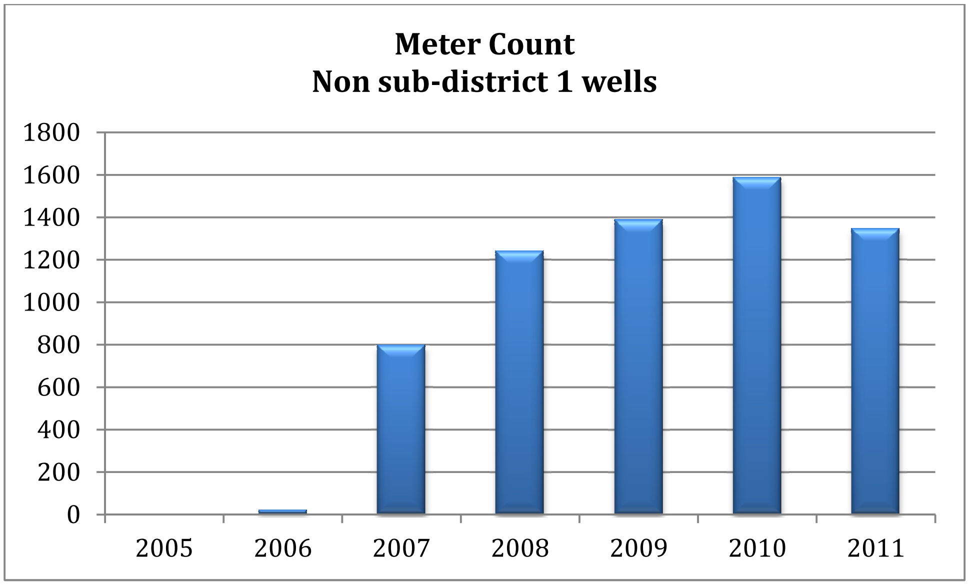


Figure 9. Meter Count - Non sub-district 1 wells

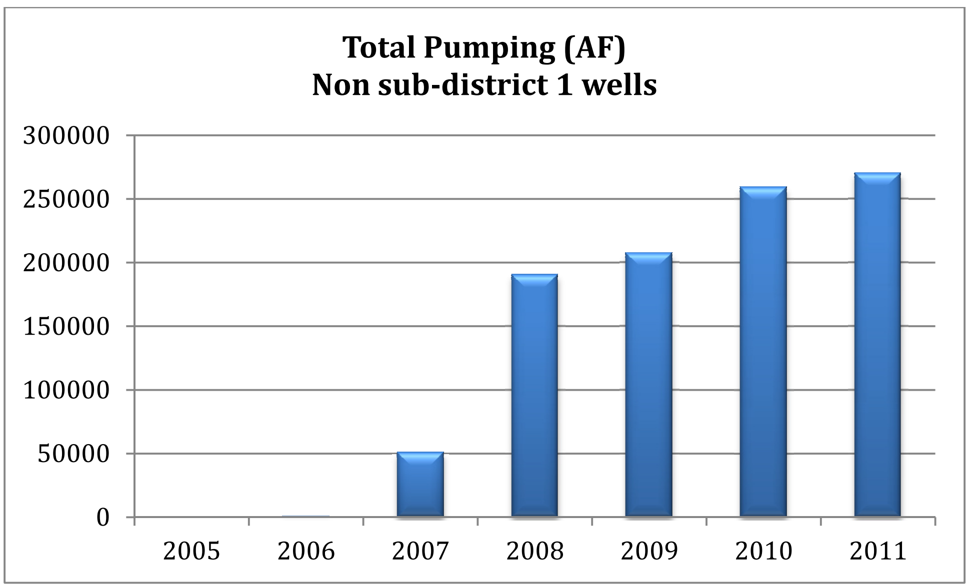


Figure 10. Total Pumping (AF) - Non sub-district 1 wells

The graphs indicate that the sub-district 1 meter data are close to an apparent asymptotic limit. This makes intuitive sense - these wells have been identified as part of a groundwater sub-district legal battle and are the subject of much scrutiny. Therefore it is likely that almost all of them are identified, metered and measured in the database already.

The non sub-district 1 wells have a somewhat asymptotic shape, though clearly not as near the limit as the sub-district 1 wells. A review of the data indicates that about 100 meters will be added back to the count in 2011, they appear to be invalid meters pending data processing that will make them valid.

4.2.2.6.2 Meter size = 8”

Another statistically large subset is the set of meters whose size is specified as 8”. This common meter size should be spatially distributed pretty evenly across the basin and therefore independent of the previous subset.

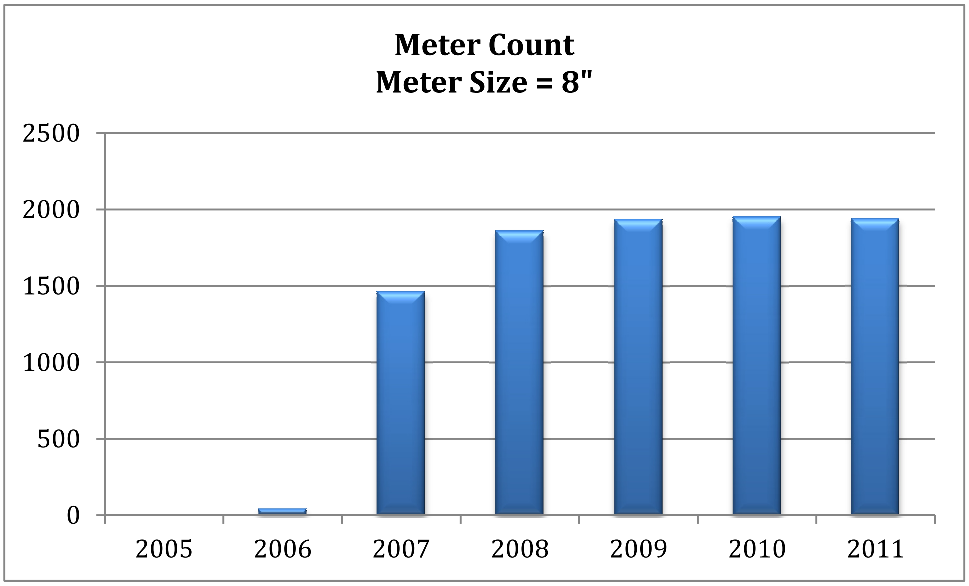


Figure 11. Meter Count - Meter Size = 8"

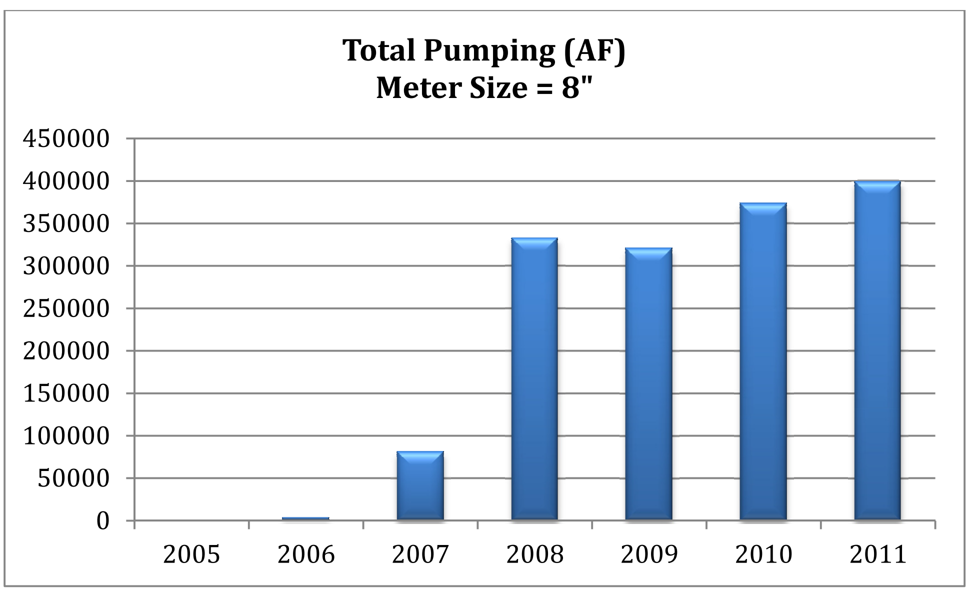
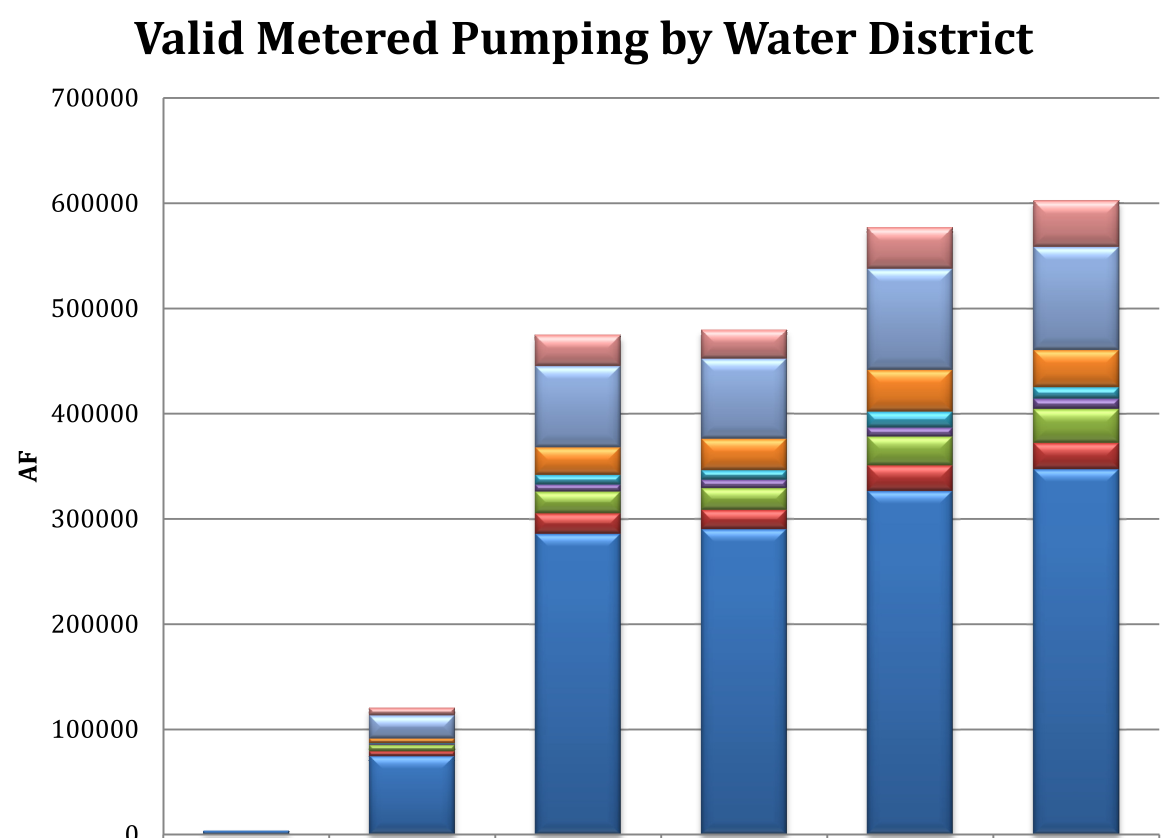


Figure 12. Total Pumping (AF) - Meter Size = 8"

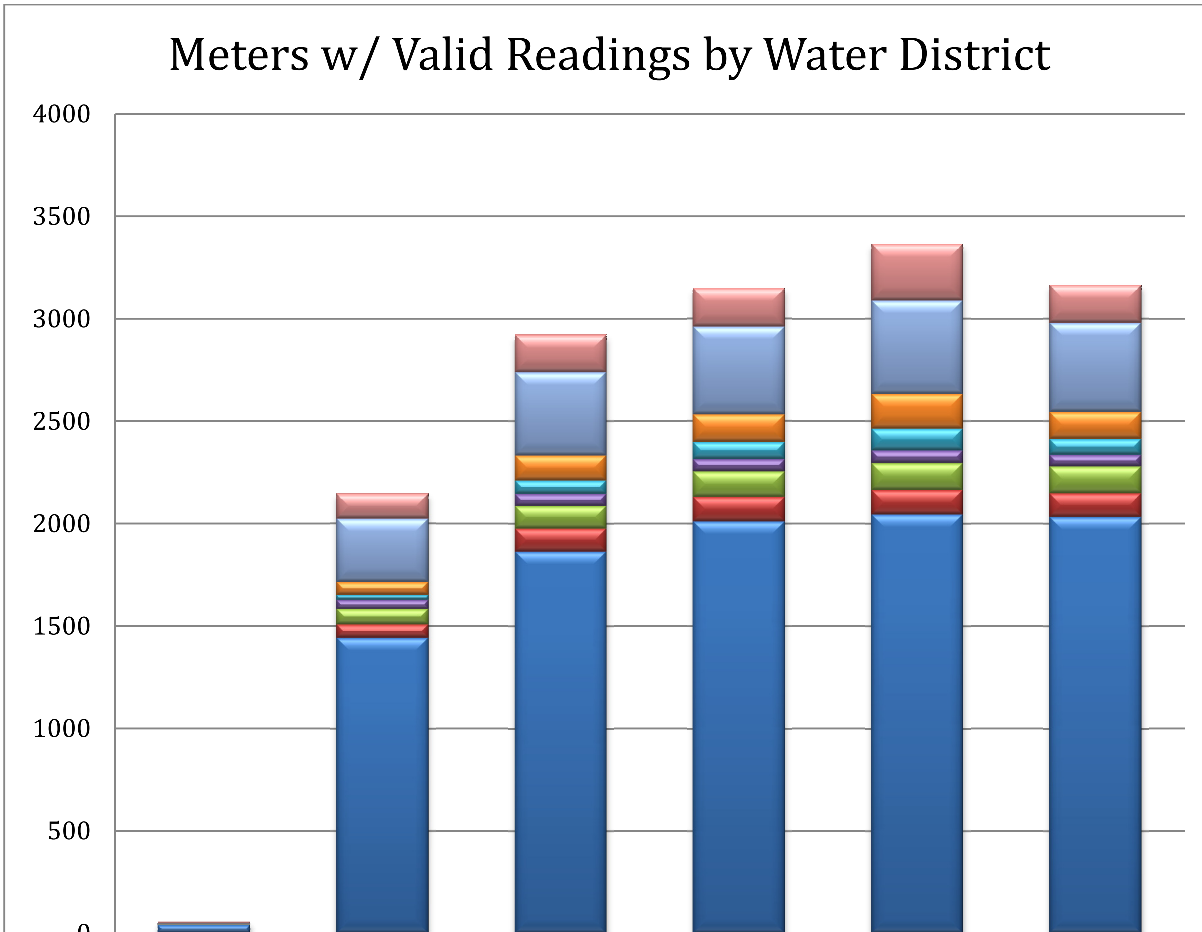
4.2.2.6.3 Water Districts

The following graphs illustrate the total pumping and total meter counts broken down by water district.



	2006	2007	2008	2009	2010	2011
35	39	7203	29210	27645	39345	43252
27	160	21465	76692	75649	95661	97896
26	0	4726	26432	30031	39924	35804
25		636	9380	9096	15067	10741
24	0	921	6683	7771	8223	9895
22	10	6181	20279	20686	27708	31828
21	0	4405	19639	18455	24306	24896
20	3136	74513	285804	290074	326421	347489

Figure 13. Valid Metered Pumping by Water District

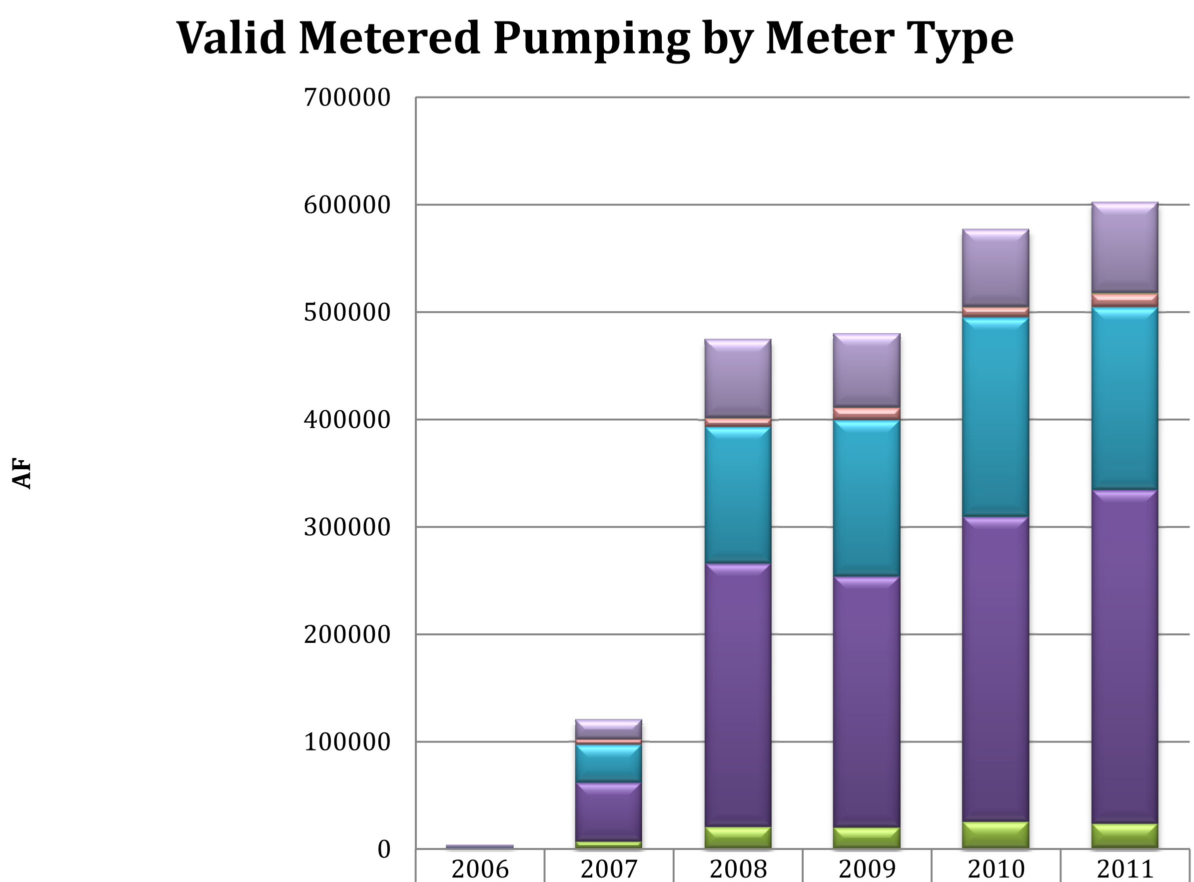


	2006	2007	2008	2009	2010	2011
35	4	121	183	189	275	183
27	4	308	405	427	457	436
26	2	64	121	136	170	131
25		24	66	83	106	77
24	2	43	59	60	61	57
22	3	77	111	123	131	131
21	1	65	112	121	120	114
20	37	1441	1863	2009	2044	2033

Figure 14. Meters w/ Valid Readings by Water District

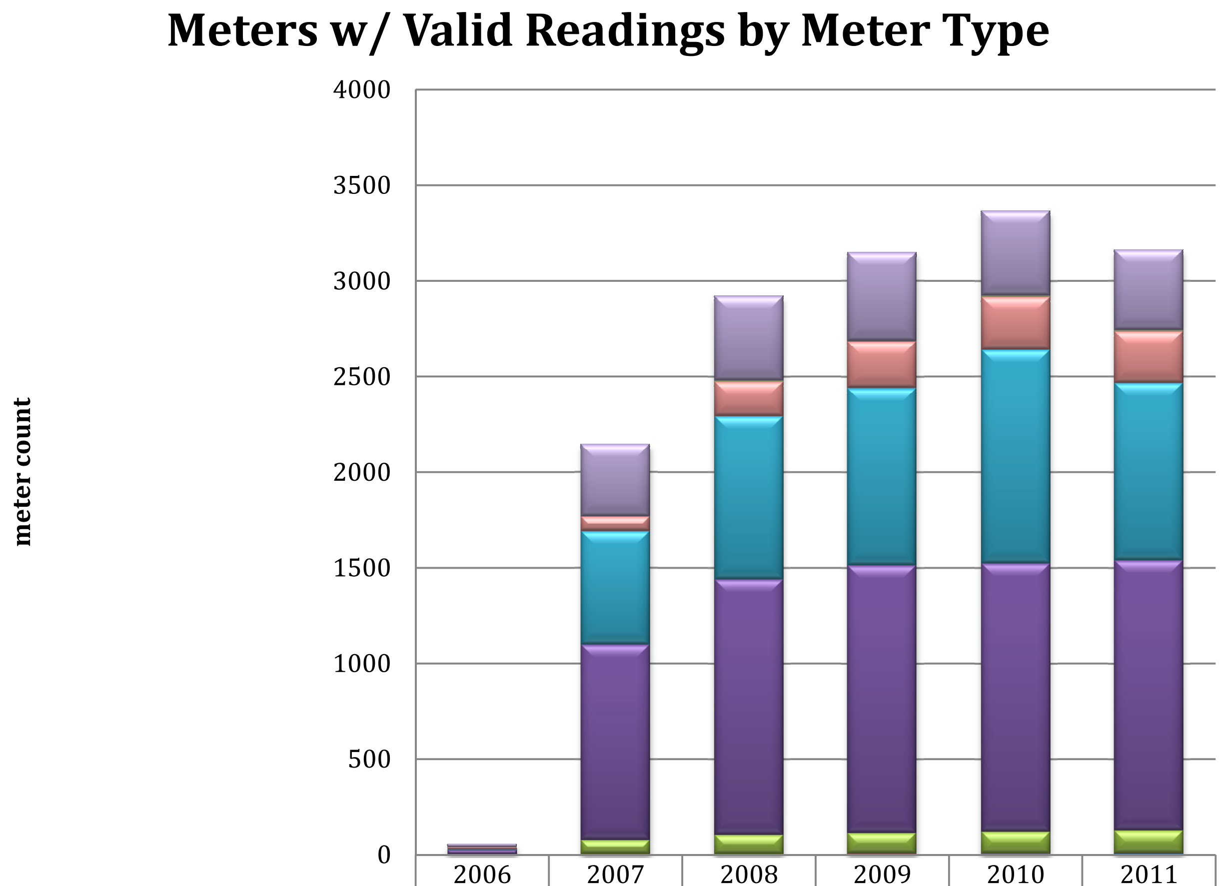
4.2.2.6.4 Meter Type

The following graphs illustrate the total pumping and total meter counts broken down by meter type.



	2006	2007	2008	2009	2010	2011
GENERIC FLOW	1275	18416	72379	68057	71668	83998
FLOW - ULTRASONIC (AF)			820	914	932	1188
FLOW - MECHANICAL (GAL)	252	4817	8322	11134	9098	12178
FLOW - MECHANICAL (CF)					0	
FLOW - MECHANICAL (AI)		0	130	137	149	186
FLOW - MECHANICAL (AF)	275	34915	126886	145213	185319	170152
FLOW - MAGNETIC (REGISTER)	1362	54634	245066	233769	283737	310517
FLOW - MAGNETIC (NO REGISTER)	182	7097	19852	19646	25075	22760
COLLINS TUBE (REGISTER)		159	475	238	346	415
COLLINS TUBE (NO REGISTER)		11	188	299	332	407

Figure 15. Valid Metered Pumping by Meter Type



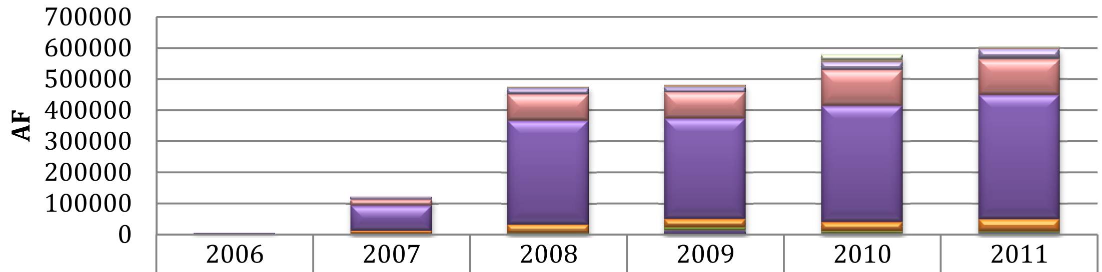
	2006	2007	2008	2009	2010	2011
GENERIC FLOW	14	378	443	461	441	418
FLOW - ULTRASONIC (AF)			6	6	8	8
FLOW - MECHANICAL (GAL)	10	76	180	240	275	270
FLOW - MECHANICAL (CF)					1	
FLOW - MECHANICAL (AI)		1	1	1	1	1
FLOW - MECHANICAL (AF)	6	592	853	927	1116	926
FLOW - MAGNETIC (REGISTER)	22	1019	1331	1399	1402	1412
FLOW - MAGNETIC (NO REGISTER)	1	72	101	107	114	118
COLLINS TUBE (REGISTER)		4	4	4	2	3
COLLINS TUBE (NO REGISTER)		1	1	3	4	6

Figure 16. Meters w/ Valid Readings by Meter Type

4.2.2.6.5 Meter Size

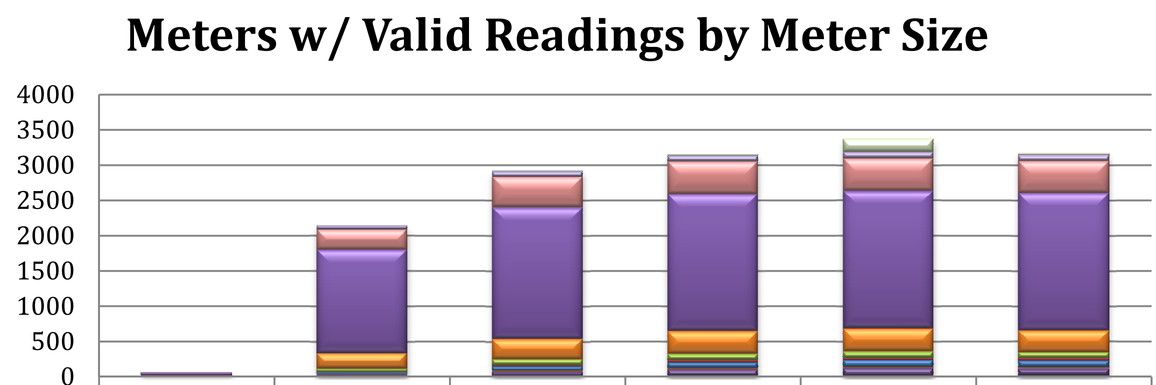
The following illustrate the total pumping and meter counts broken down by meter size.

Valid Metered Pumping by Meter Size



	2006	2007	2008	2009	2010	2011
(blank)		279	1033	1329	17716	1943
80		0	188	241	260	216
16		5	165	267	247	477
15		731	3006	3596	3620	3481
14		203	897	921	1319	1118
12		7862	18635	17682	25949	30836
11			386	55	431	638
10	330	18086	83971	84344	113588	114653
9		6	287	62	51	160
8.25		73	226	42	155	155
8.12		52	233	193	174	207
8	2963	81252	332715	320271	372976	398783
7.81		113	134	93	108	57
7		99	681	795	979	870
6.25		99	197	176	215	273
6	42	9633	25132	27190	29771	38393
5		1	22	39	40	22
4.5				162	204	212
4		866	3426	8416	5623	4311
3	9	26	680	1442	1043	1175
2	0	407	897	1740	1294	2867
1.5	0	20	370	73	87	99
1.25						11
1	2	76	265	10000	442	545
0.75		0	1	3	2	2
0.625			0	0	0	0
0		160	572	275	364	298

Figure 17 Valid Metered Pumping by Meter Size



	2006	2007	2008	2009	2010	2011
(blank)		4	8	16	178	20
80		1	1	1	1	1
16		1	1	1	1	1
15		2	2	2	2	2
14		2	3	3	3	2
12		52	73	80	85	83
11			1	1	1	1
10	3	282	430	457	457	448
9		2	3	3	2	2
8.25		1	1	1	1	1
8.12		1	1	1	1	1
8	39	1459	1856	1934	1948	1939
7.81		1	1	1	1	1
7		2	5	6	6	5
6.25		2	2	2	2	2
6	6	218	286	316	316	311
5		2	2	2	2	2
4.5				1	1	1
4		42	82	90	92	81
3	1	8	23	31	34	36
2	1	30	61	86	98	92
1.5	1	4	16	16	18	18
1.25						1
1	2	20	52	86	104	102
0.75		1	3	4	5	5
0.625			1	1	1	1
0		6	6	6	4	3

Figure 18 Meters w/ Valid Readings by Meter Size

4.2.2.7 Meter (GWDMS) vs. Well (HydroBase) Pumping Comparison

Between the meter readings data collection residing in GWDMS and the official DWR annual well pumping diversion values in the DWR HydroBase, is a complex and lengthy set of data processes, some even occurring outside the database completely. As described in earlier sections of this report, the latter require exporting of data to spreadsheets and then re-importing of calculation results. The data during these calculations is preprocessed with automated queries, exported from GWDMS to an Excel workbook, QA/QC'ed manually by Division 3 staff, integrated with "meter to well connection" information by Division 3 staff, and then is re-imported back to the DWR HydroBase for final processing into official annual diversion values.

Initially, pumping volume errors caused by meter roll-overs, unit errors, reporting and transcribing errors in dates, decimal places, meter serial numbers, and etc. need to be removed. Next these pumping volumes for meters need to be converted to annual values and then assigned to wells. In database parlance, the relationship between meters and well is "many to many". Often there is one meter per well. However, frequently there are multiple wells sharing a single meter, or multiple meters on a single well, and sometimes even multiple meters connected to multiple wells. Splitting and/or combining the values correctly require the experience and judgment of the Division 3 staff and cooperation of the water users. Complicating this further is the mixing of surface water supplies through some of the meters, requiring additional measurements, metering and calculations to isolate the groundwater pumping portion of irrigation deliveries. The table below is an excerpt from the 2011 workbook that helps illustrate the complexity of this annual task.

WDID	STRUCTURE NAME		METER SERIAL #	TOTAL PUMPED		Formula	Diversion	Units
2008220	82CW161 WELL NO 01 82CW161 WELL NO 02 82CW161 WELL NO 03 PERMIT 47983-F-	a	10062787/10062806	142.08	C	a/4	**35.5**	AF
2008221							**35.5**	AF
2008222							**35.5**	AF
2014050							**35.5**	AF
2008229	82CW164 WELL NO 16 82CW164 WELL NO 17 82CW164 WELL NO 18 PERMIT 47996-F-	a	07-8-1397	191.88	C	a/4	**48.0**	AF
2008230							**48.0**	AF
2008231							**48.0**	AF
2014271							**48.0**	AF
2005454	W3686 WELL NO 05	a	10063599/10063660-100490	215.6	D	(a-b)	**123.76**	AF
2005496	81CW061 WELL NO 08S	b	10063529/10063707-100446 - 3A	91.84		b	**91.84**	AF
2008934	W0372 WELL NO 03	a	01061414/01061441	192.9	D	a/2	**96.45**	AF
2008932	W0372 WELL NO 01	b	04-8-1588	0.43		a/2+b	**96.88**	AF
2008414	W0091 WELL NO 08	a	01062047/01062063	203.5	D	a/2	**101.75**	AF
2008410	W0091 WELL NO 04	b	07-8-1151	2.71		a/2+b	**104.46**	AF
2008449	W0102 WELL NO 07	a	10063617/10063636-100534	320.34	D	(a-b)	**144.78**	AF
2008450	W0102 WELL NO 08	b	10063616/10063658-100535	175.56		b	**175.56**	AF
2008448	W0102 WELL NO 06	a	10063612/10063634-100529	240.49	D	(a-b)	**147.83**	AF

WDID	STRUCTURE NAME		METER SERIAL #	TOTAL PUMPED		Formula	Diversion	Units
2008451	W0102 WELL NO 09	b	10063607/10063633-100531	92.66		b	**92.66**	AF
2008940	W0376 WELL NO 02	a	10060696/10061097-100164	252.95	D	a/2	**126.48**	AF
2008939	W0376 WELL NO 01	b	10060697/10061089	12.17		a/2+b	**138.65**	AF
2013996	PERMIT 20380-F-	a	10063574/10063696-100515	246.06	D	a-b	**149.79**	AF
2011425	W1830 WELL NO 02	b	10063624/10063635-100527	96.27		b	**96.27**	AF
2006266	79CW013 WELL NO 01A	a	03063675/03063703	242.11	D	a/2	**121.06**	AF
2008427	W0098 WELL NO 02	b	GP10-2560- OPEN DISCHARGE METE	37.45		(a/2)+b	**158.51**	AF

Figure 19 Excerpt from 2011 Meter to Well Workbook

This process for 2009, 2010, and 2011 corrected, improved and refined the groundwater pumping data. The figures below demonstrate the magnitude of the changes, comparing the GWDMS meter database pumping totals to the HydroBase well database pumping totals, both basin-wide and in selected water districts.

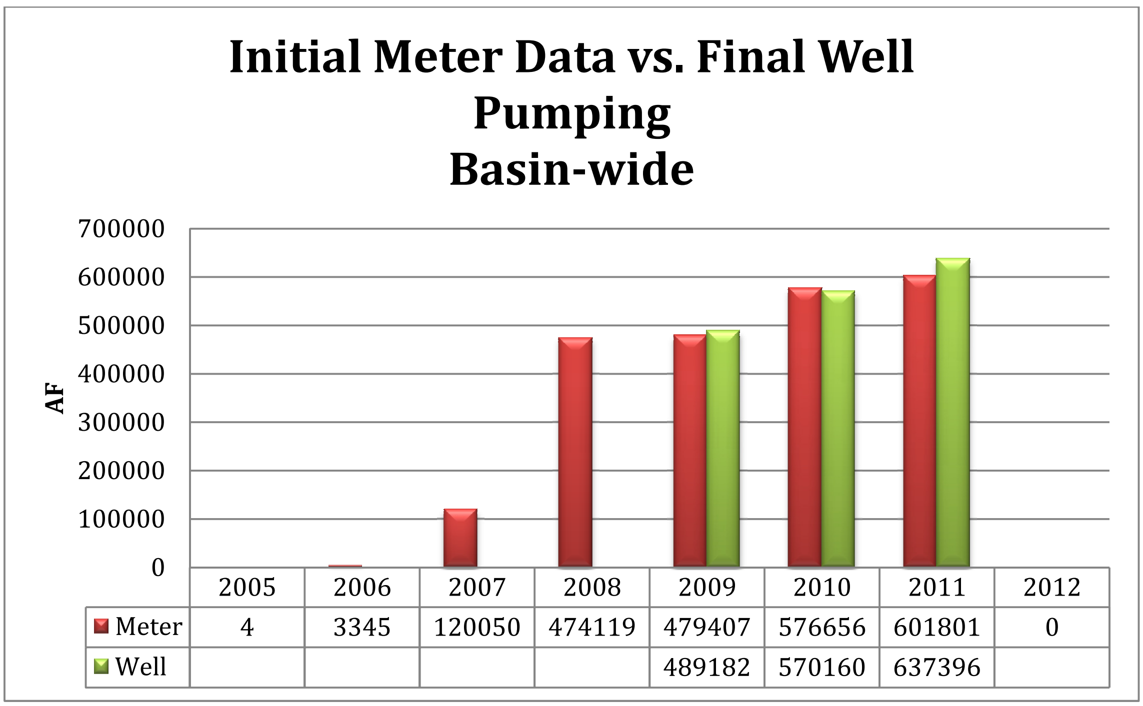
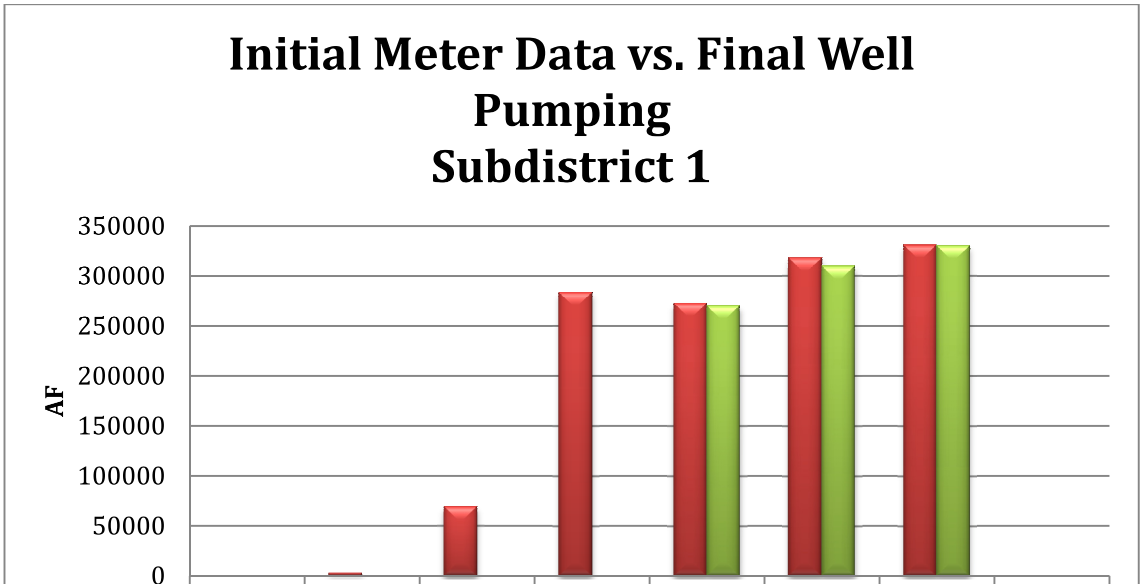
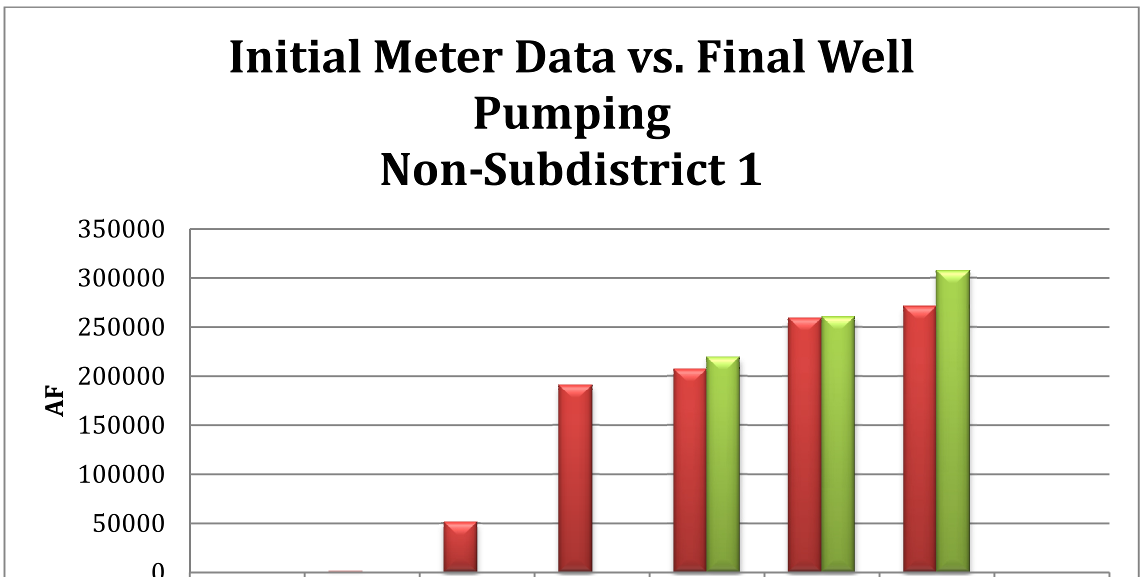


Figure 20 Initial Meter Data vs. Final Well Pumping - Basin-wide



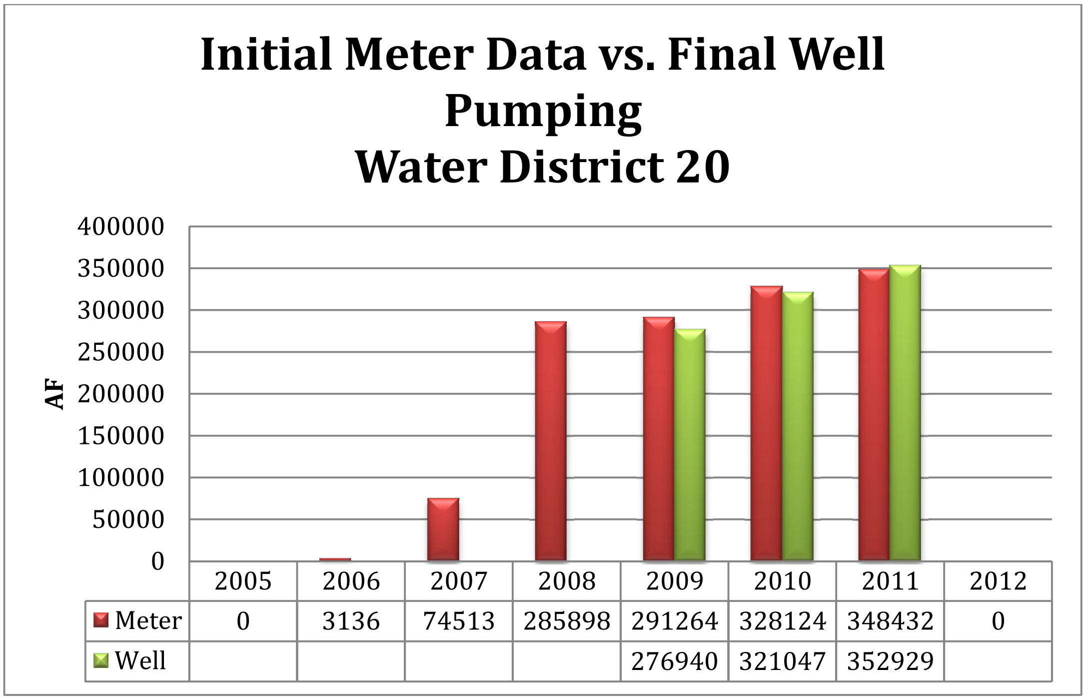
	2005	2006	2007	2008	2009	2010	2011	2012
Meter	0	2778	69231	283525	272364	317847	331037	0
Well					269778	309881	330208	

Figure 21 Initial Meter Data vs. Final Well Pumping - Subdistrict 1



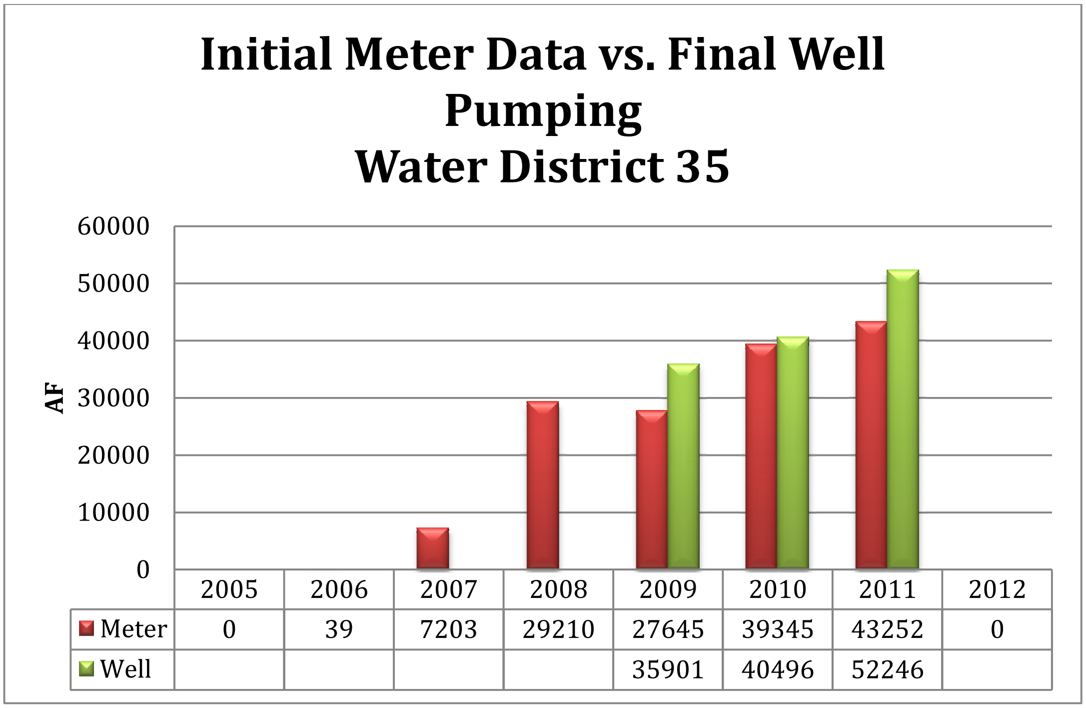
	2005	2006	2007	2008	2009	2010	2011	2012
Meter	4	568	50819	190593	207042	258809	270763	0
Well					219404	260280	307188	

Figure 22 Initial Meter Data vs. Final Well Pumping - Non-Subdistrict 1



	2005	2006	2007	2008	2009	2010	2011	2012
Meter	0	3136	74513	285898	291264	328124	348432	0
Well					276940	321047	352929	

Figure 23 Initial Meter Data vs. Final Well Pumping - Water District 20



	2005	2006	2007	2008	2009	2010	2011	2012
Meter	0	39	7203	29210	27645	39345	43252	0
Well					35901	40496	52246	

Figure 24 Initial Meter Data vs. Final Well Pumping - Water District 35

The combination of a well-designed system of rules and regulations, modern GPS and metering technology, experienced and dedicated Division 3 staff, efficient DWR data management tools and the cooperation of water users in the basin has yielded high quality groundwater pumping diversion data for 2009 through 2011 in HydroBase.

Because of the careful metering, QA/QC and data processing, most of the individual well diversion values in HydroBase are representative of the actual pumping. Given the scale and complexity of the effort, however, it is inevitable that errors will still persist on a few individual well diversion values. In aggregate, though, the findings of this report demonstrate that the values are very likely to be accurate and reliable.

Naturally, room for improvement still exists in the system. The 2009-2011 QA/QC process improved the data, but also leaves it vulnerable to the introduction of new data errors inside the QA/QC workbooks.

Also, the 2009-2011 processes were "one-directional". In other words, many data corrections and refinements made it *forward* into the final annual well pumping values, but not *backward* into the raw meter collection data. This means that within the database systems (HydroBase and GWDMS), the meter pumping to well pumping conversion is not a reliably repeatable or reversible process. As suggested in earlier sections of this report, this is an important component of data integrity. It should be addressed. In the future it will be very important WITHIN THE DATABASE to be able to easily recalculate well pumping when meter data are updated, and also to be able to display how any particular well pumping value was derived from the meter data.

[February/March 2012 UPDATE: Based on discussions with DWR, the database features described in the previous paragraph are already being designed and implemented in a new version of the DWR data management system. It should be ready for the 2012 data year.]

4.2.3 Meter Test Data

4.2.3.1 Description

The meter test data set in the GWDMS database can be used to analyze the meter readings data. As explained earlier all meters are required to be tested when installed, every 4 years subsequently, and whenever anything about the well or meter changes. The resulting detailed test data is recorded on forms and then entered by Division 3 into the GWDMS database.

The test data is simple to evaluate and relates directly to the meter readings data, because it compares test equipment volume readings (assumed to be correct) with volume readings from the installed meter on the well. The error observed is incorporated into a "correction factor" that is calculated for every test.

Current rules specify that meters with correction factors between .95 and 1.05 are considered accurate and the meter readings are used as is. Meters with correction factors between .92-.95 and 1.05-1.08 are allowed to be used with a “variance” status, and the correction factor is used to “correct” the metered volume. Meters with correction factors <0.92 or >1.08 are considered to have failed the test and cannot be used until fixed.

At least one meter test should exist for every well in the database, and sometimes multiples exist. This creates a statistically viable sample set that is also coherent with the meter readings data, since it covers all the same meters. The calculated correction factor is normalized, therefore so that large wells have an identical weight as small wells. This means that the statistical relationship being evaluated is to the meter population not the meter pumping. This should be sufficient for the test being undertaken. Additionally, the same test will be run on selected data subsets (based on meter size, meter type, etc.).

For each test an error percentage was calculated as (QI – QT) / QT where QI is the well meter measured volume and QT is the test meter measured volume. If the error distribution is symmetrically distributed around 0 (i.e. zero error) and center-weighted, or even better if it is similar to a classical “normal” distribution, then it indicates the *meter reading* volumes in aggregate will likely accurately represent the *actual pumping* volumes in aggregate. The collection of meter test error percentages in aggregate over all meters was extracted from the database and the frequency histogram is shown below.

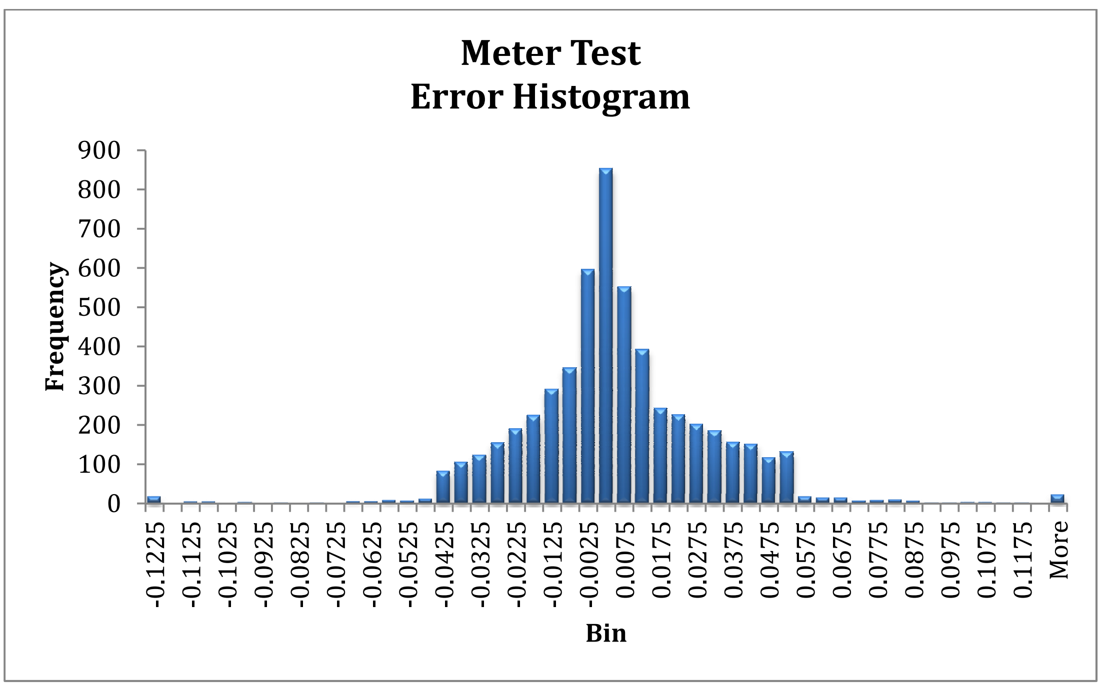


Figure 25 Meter Test - Error Histogram

However, it was discovered that the meter test data in the database for "magnetic" type meters is most often entered AFTER a pretest is used to correct the meter's "K factor." This means that these particular meter test records will NOT represent raw meter error, but "corrected meter error." In addition, sometimes a mechanical meter that fails a field test will be removed for repair or replacement and the failed test data will not be submitted. Both of these issues cause bias in the meter test errors in the existing GWDMS database prior to 2012. Division 3 will be implementing new procedures to encourage failed test data to be submitted. These and other data issues will be addressed in the following analyses.

4.2.3.2 Analyses

Although the aggregate meter test error histogram above generally indicates the desired symmetrical error distribution roughly centered on 0, the data need to be analyzed more carefully. Some of the meter test data needed to be excluded before a rigorous analysis could proceed. "Legitimate errors in actual tests" are the only records that should be used in the analysis. Therefore those that have correction factors of 0 or abnormally large values were excluded.

It should be noted that a symmetrical error distribution centered on 0 (or correction factor centered on 1) will cause the pumping volumes for large groups of meters in the GWDMS to be accurate in aggregate. . If this distribution pattern generally exists across subsets of data, then it causes individual meter errors to effectively cancel out instead of accumulate.

Because of the potential (and expected) difference in meter test error distributions for meter types and meter test personnel, database queries were constructed to isolate meter test data based on these and other criteria.

4.2.3.2.1 Meter Types

Three general types of meters are prevalent in the Division 3 database: mechanical meters, magnetic meters and generic flow meters. Other types exist, too, but only rarely. Meter error distributions will be investigated for these three prevalent meter types.

4.2.3.2.1.1 Mechanical Meters

Four mechanical meter types exist in the database for this category, but only the first two are significantly used:

- Mechanical (AF)
- Mechanical (Gal)
- Mechanical (CF)
- Mechanical (AI)

Slightly less than half (by count and pumping volume) of the meters in the basin are in this category.

Error terms were calculated for every mechanical meter test record in the database. These were used to create the error histogram below. A normal distribution (representing random behavior) was fit to this data and is shown on the histogram. Also, the key reference points in the error data are designated using vertical lines, the 0 mark (no error), the +/- 5% error range, and the +/- 8% error range. These are critical components of the Division 3 Well Pumping Data Collection Rules and Regs. Descriptive statistics about the data are shown in Figure XX below.

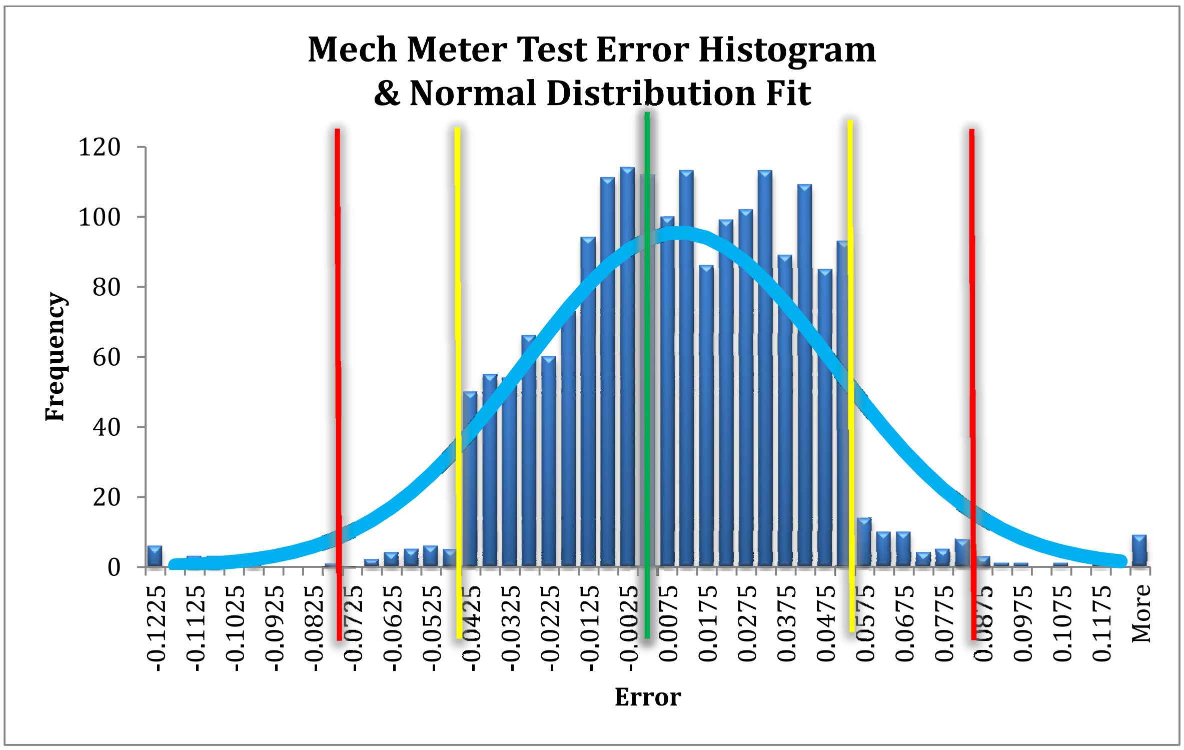


Figure 26 Mech Meter Test Error Histogram & Normal Distribution Fit

It is apparent that the error distribution is not a perfect "normal distribution" showing pure random error. The mean is higher than 0 and it is non-symmetrical, skewed to the right. This indicates that the actual pumping through mechanical meters is slightly lower than the reported values. This is true ONLY in aggregate over all the mechanical meters and over all the years.

Another concern in the mechanical meter test data is the appearance of abrupt shoulders at the +/- 5% error locations. This is not unexpected given that at least two causes of bias could have been introduced into the data:

- When a mechanical meter fails a field test (performed by non-Division 3 staff), sometimes the meter is repaired or replaced and a new test performed before the data is submitted to

Division 3. This practice introduces a significant source of bias into the meter test error data that will create the shoulders shown in the above figure. Division 3 is introducing new data collection procedures to encourage more of the failed test data to be submitted.

- To a lesser extent some understandable human bias could be introduced into the data. Given that the calibrated meter test equipment is not itself perfect and considering the costs incurred when a meter fails a test, it is likely that some bias will be introduced by testers on borderline test cases when the errors are very close to the limits. It is difficult to detect this bias in the test data, though, given the known significant bias caused by the previous issue. However, the error distribution data for individual testers can be investigated and compared to other testers to find anomalous distributions. A web-based data exploration tool for Division 3 to use was developed for this purpose. In the future when more of the failed meter test data is submitted and stored in the database, detecting this bias will become easier.

4.2.3.2.1.2 Magnetic Meters

Two magnetic meter types exist in the database for this category:

- Magnetic (Register)
- Magnetic (No register)
- Slightly less than half (by count and pumping volume) of the meters in the basin are in this category.

As noted previously, many of the meter test error data in the current database for magnetic meters (performed by non-Division 3 staff) is subsequent to adjusting the meter's "K- factor" which essentially eliminates the meter error that had developed in the meter since the last test. Therefore, much like the repair/replacement of inaccurate mechanical meters, this introduces a bias into the collected magnetic meter test data. In these cases the meter error results for magnetic meters will measure how much error is left after K-factor adjustments. It is expected to be very similar to theoretical random meter error distributions – very tightly and symmetrically grouped around 0 but still with some indication of shoulders at the 5% error limits. Division 3 is introducing new data collection procedures to encourage meter error or K-Factors to be reported BEFORE the K-factor is adjusted, in order to better understand how magnetic meter errors develop over time.

Error terms were calculated for every magnetic meter test record in the database. These were used to create the error histogram below. A normal distribution (representing random behavior) was fit to this data and is shown on the histogram. Also, the key reference points in the error data are designated using vertical lines, the 0 mark (no error), the +/- 5% error range, and the +/- 8% error range. These are critical components of the Division 3 Well Pumping Data Collection Rules and Regs. Descriptive statistics about the data are shown in a table below.

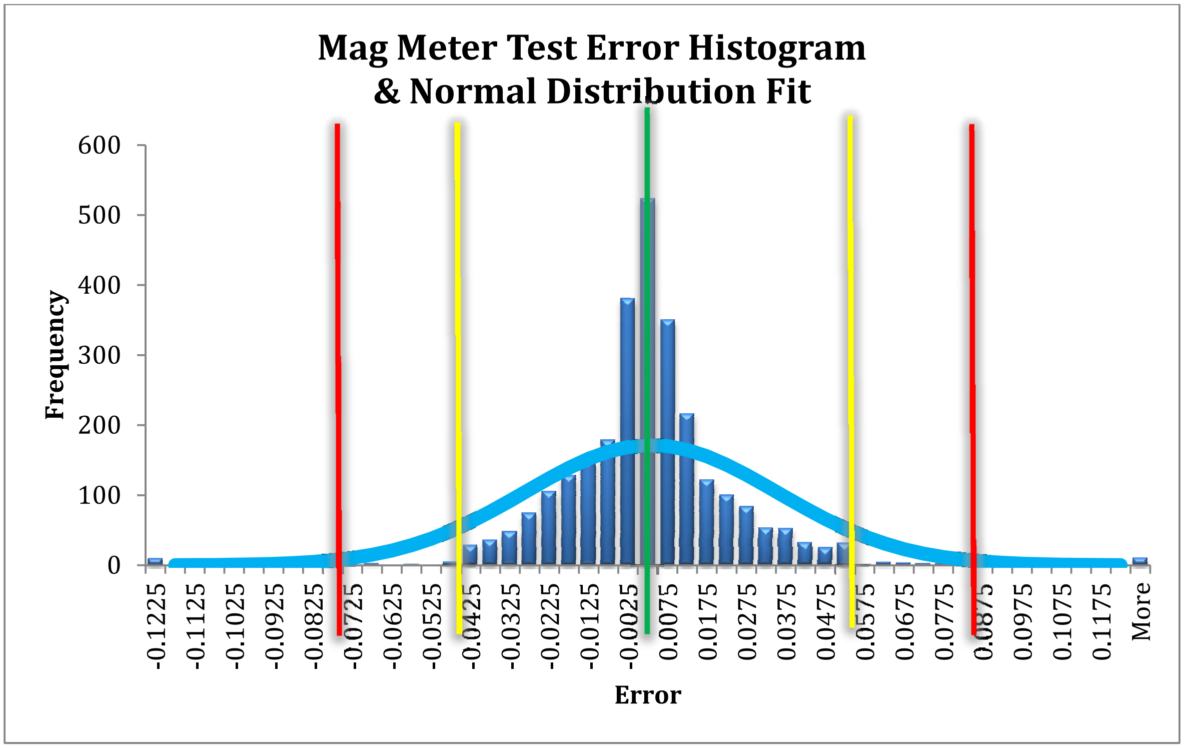


Figure 27 Mag Meter Test Error Histogram & Normal Distribution Fit

When comparing Figure 26 to Figure 27, it is apparent there is much less meter error in the magnetic meter data. The error is even more center weighted than a normal distribution, is symmetrical, is centered on 0, and as expected has slight shoulders at the 5% limits. However, it is difficult to determine how much of this behavior is due to the historical practice of resetting K-Factors before testing and submitting test data, or how much is due to the inherent accuracy of magnetic meters. In the future as more pre-adjustment K factors are added and/or pre-adjustment test data are submitted, it will become clearer.

4.2.3.2.1.3 Generic Flow Meters

About 15 percent (by count and pumping volume) of the meters in the basin are in this category.

Error terms were calculated for every generic flow meter test record in the database. These were used to create the error histogram below. A normal distribution (representing random behavior) was fit to this data and is shown on the histogram. Also, the key reference points in the error data are designated using vertical lines, the 0 mark (no error), the +/- 5% error range, and the +/- 8% error range. These are critical components of the Division 3 Well Pumping Data Collection Rules and Regs. Descriptive statistics about the data are shown in a table below.

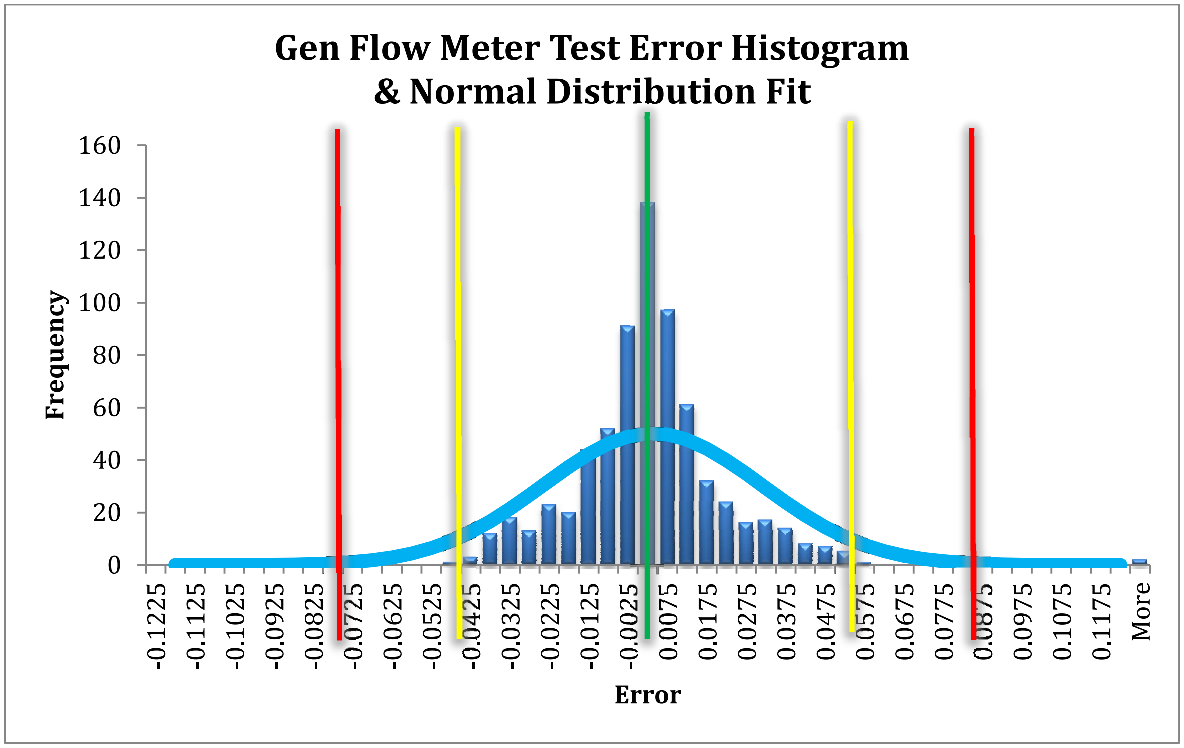


Figure 28 Gen Flow Meter Test Error Histogram & Normal Distribution Fit

The above error distribution is similar to the magnetic meter data distribution. The error observed is more center weighted than a normal distribution, is symmetrical, and is centered on 0.

4.2.3.2.1.4 Statistical Comparison

In Figure 29 on the follow page are some typical quantitative statistics for the meter error distributions for the three major types of meters in the data. The numbers quantify the general patterns described already in the previous sections. Some interesting numbers to note are the high positive values of "kurtosis" (more "peaked" around 0) and the positive skewness (distribution "fatter" on the left or negative) for magnetic and generic flow meter test errors.

	Mechanical	Magnetic	Generic Flow
Mean	0.007909	0.000772	0.001196
Standard Error	0.000905	0.00061	0.00105
Median	0.008213	-0.00041	0.000289
Standard Deviation	0.039276	0.032056	0.027763
Sample Variance	0.001543	0.001028	0.000771
Kurtosis	51.55201	239.6219	201.9813
Skewness	-1.11788	10.62984	10.87926
Range	1.045537	0.984536	0.584624
Minimum	-0.62537	-0.19975	-0.04784
Maximum	0.420168	0.784788	0.536786
Sum	14.885	2.130323	0.836133
Count	1882	2759	699
Confidence Level(95.0%)	0.001776	0.001197	0.002062

Figure 29 Comparison of Descriptive Statistics for Meter Types

4.2.3.2.2 Years

Field experience suggests that the meter error, especially mechanical,is progressive over time. To test this hypothesis the mechanical meter error data were further broken down by year to see if the pattern is progressive.

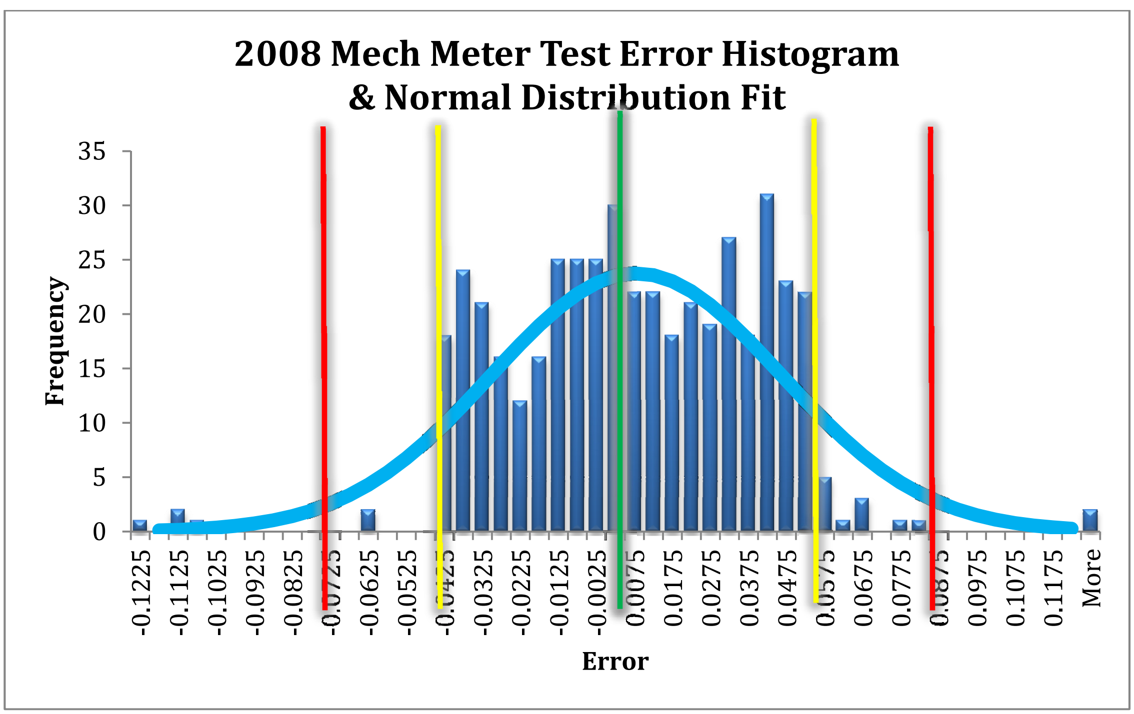


Figure 30 2008 Mech Meter Test Error Histogram & Normal Distribution Fit

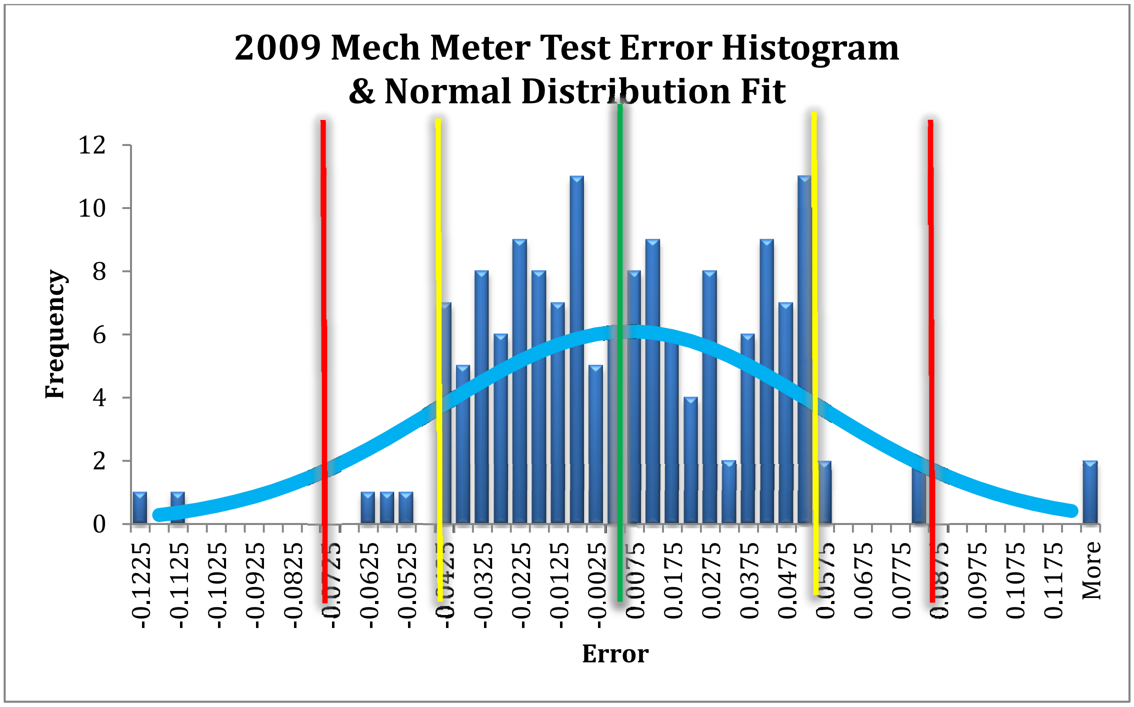


Figure 31 2008 Mech Meter Test Error Histogram & Normal Distribution Fit

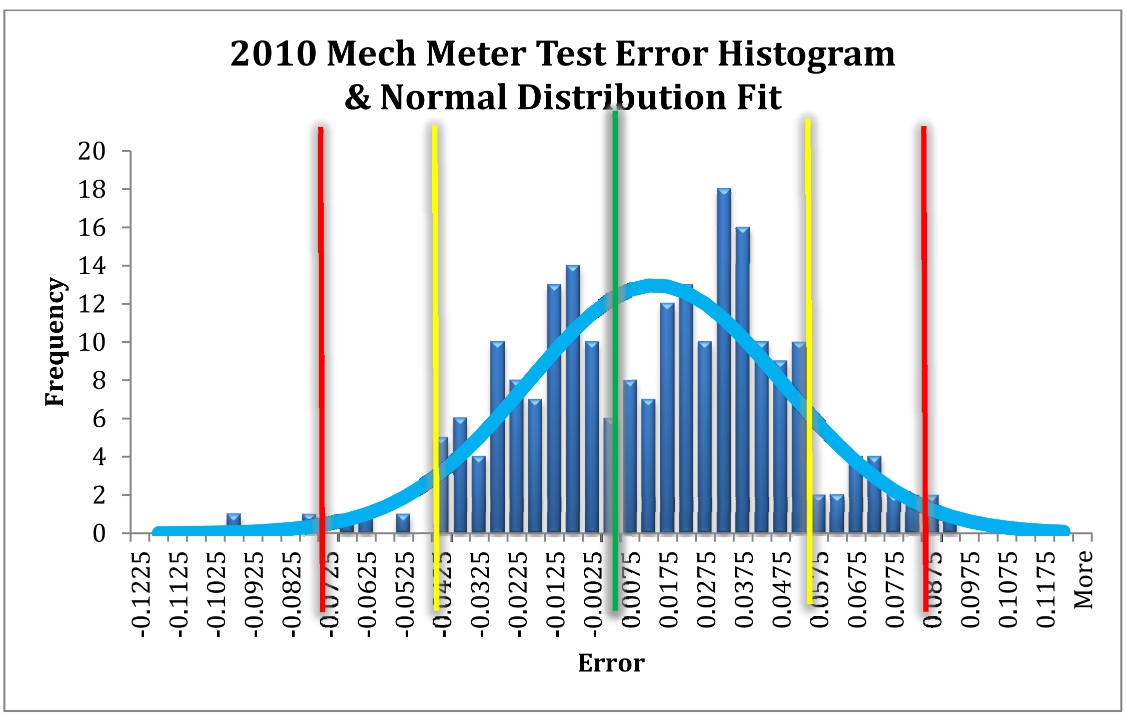


Figure 32   2010 Mech Meter Test Error Histogram & Normal Distribution Fit

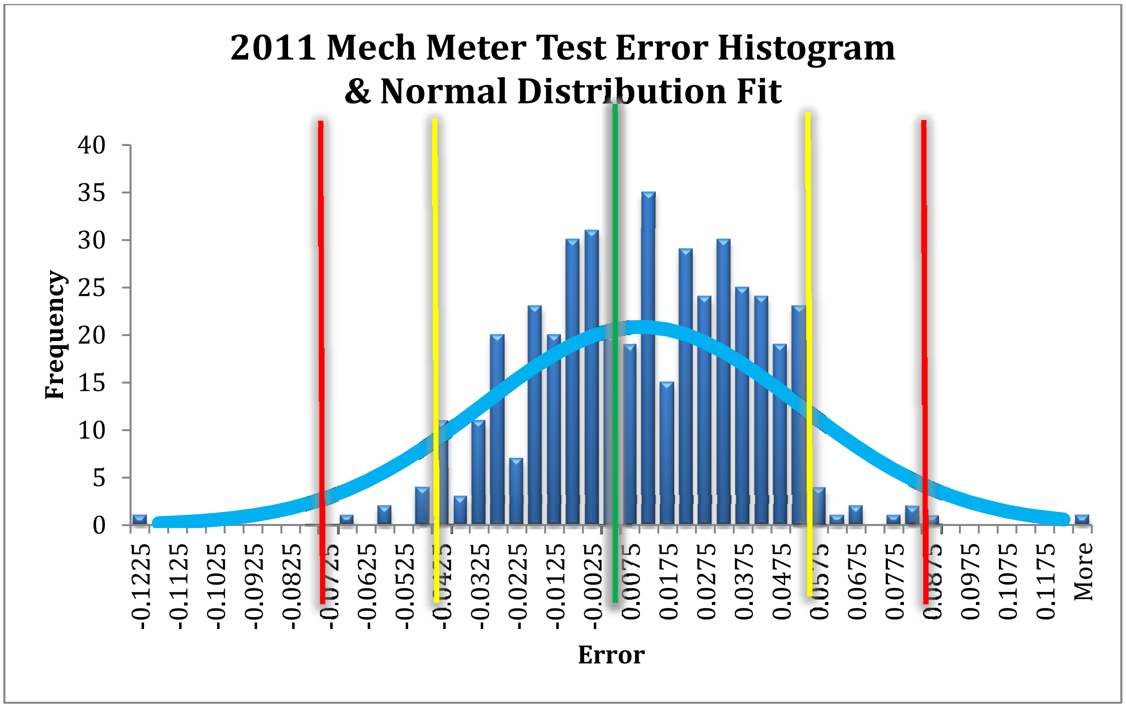


Figure 33 2011 Mech Meter Test Error Histogram & Normal Distribution Fit

	2008	2009	2010	2011
Mean	0.005515	0.003781	0.011612	0.007954
Standard Error	0.00179	0.004045	0.00228	0.002002
Median	0.003702	0.002304	0.016176	0.009656
Standard Deviation	0.038135	0.05003	0.033816	0.041957
Sample Variance	0.001454	0.002503	0.001144	0.00176
Kurtosis	22.2722	18.65379	-0.16259	118.0372
Skewness	1.608148	0.660061	-0.1952	-7.78323
Range	0.595508	0.619649	0.187616	0.785187
Minimum	-0.21692	-0.28916	-0.09821	-0.62537
Maximum	0.378592	0.330492	0.089407	0.159818
Sum	2.503611	0.578494	2.554556	3.491972
Count	454	153	220	439
Confidence Level(95.0%)	0.003517	0.007991	0.004493	0.003936

Figure 34 Comparison of Descriptive Statistics for Years of Mech Meter Tests

Considering both the mean and skewness statistics, it appears that that the mechanical meter errors do progress over time to become skewed (negative skewness) toward positive errors – in other words, the meters tend to record higher volumes than actual. And it is reasonable to infer this correlates with higher meter age in the first 4 years of the data collection program. An analysis with more years of data and calculating the actual individual meter age would be necessary to define more rigorously the error vs. age patterns for mechanical meters.

4.2.3.2.3 Meter Testers

After discussion with Division 3 staff, it was decided to review the error distributions in individual meter tester data. There are about fifty meter tester id's in the database. A summary table is presented first. Selected individual detailed examples are presented next. Finally, an online data viewer for the GWDMS meter test errors was developed for Division 3 staff to use to investigate the data for particular testers in more detail.

meter test results by tester						
tester_num	fail <92	variance 92-95	pass 95-100	pass 100-105	variance 105-108	fail >108
0	0	0	0	1	0	0
1	5	3	78	112	1	1
2	2	0	163	100	1	2
3	0	0	276	215	0	0
4	0	0	70	82	0	0
5	0	0	89	132	1	0
6	0	0	35	18	0	0
7	0	0	80	121	0	0
10	0	2	36	25	0	0
11	0	3	120	120	1	0
12	1	1	43	121	2	2
15	0	1	42	77	4	0
17	0	0	2	1	0	0
19	6	9	211	174	1	3
21	0	0	27	19	0	1
30	5	0	70	112	0	2
32	1	4	165	226	5	1
35	0	0	1	2	0	0
36	1	0	15	32	0	0
37	0	1	14	21	0	0
38	2	3	174	98	0	0
41	0	0	0	1	0	0

meter test results by tester						
tester_num	fail <92	variance 92-95	pass 95-100	pass 100-105	variance 105-108	fail >108
42	0	0	33	28	0	0
45	2	8	74	48	1	1
46	0	0	15	20	0	0
47	0	1	82	100	0	1
48	0	1	41	42	0	0
51	0	0	34	33	0	0
91	1	0	20	28	0	0
92	6	15	37	21	1	7
94	0	0	0	1	0	0
95	0	0	17	7	0	0
96	0	0	3	0	0	0
102	1	1	158	124	1	0
104	0	1	1	2	0	0
131	0	0	19	16	0	0
142	6	8	36	29	3	6
143	0	0	87	90	0	0
144	0	0	3	1	0	0
145	2	1	48	25	0	0
148	0	0	54	85	1	0
158	0	0	2	0	0	0
192	0	0	2	0	0	0
196	0	0	0	1	0	0
197	3	0	25	71	1	1
199	0	0	12	2	0	0
201	0	0	0	2	0	0
202	1	0	44	54	1	0
203	0	0	55	57	0	0
204	0	0	10	1	0	0

Figure 35 Meter Test Results By Tester Number

It can be seen that some testers observe meter error outside the allowed bounds, while other do not. But because of the meter types, it is hard to determine whether this is because of the tester or the meters (and meter types) they happen to work with.

Figures 36 and 37 show mechanical meter test errors for DWR testers vs. nonDWR testers. Unfortunately, out of 1882 mechanical meter tests, only 82 were performed by DWR staff, so the sample is too small to make conclusive observations. However, the same negative skewness or positive error (measurements higher than actual) is evident. And there is one marked difference - the DWR test errors are much more likely to cross the 5% variance boundary, as seen in the graph, even in the small sample set. This is expected because of the practice by non-Division 3 testers of adjusting magnetic meter K-Factors and repairing/replacing mechanical meters before errors are submitted.

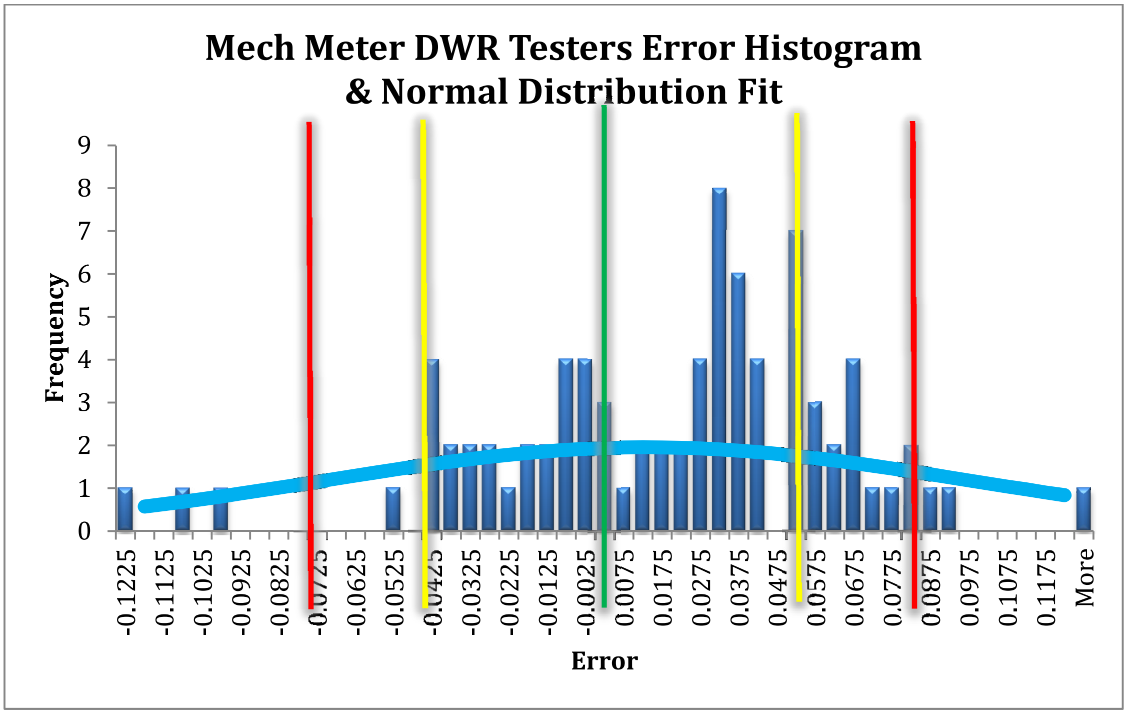


Figure 36 Mech Meter DWR Testers Error Histogram & Normal Distribution Fit

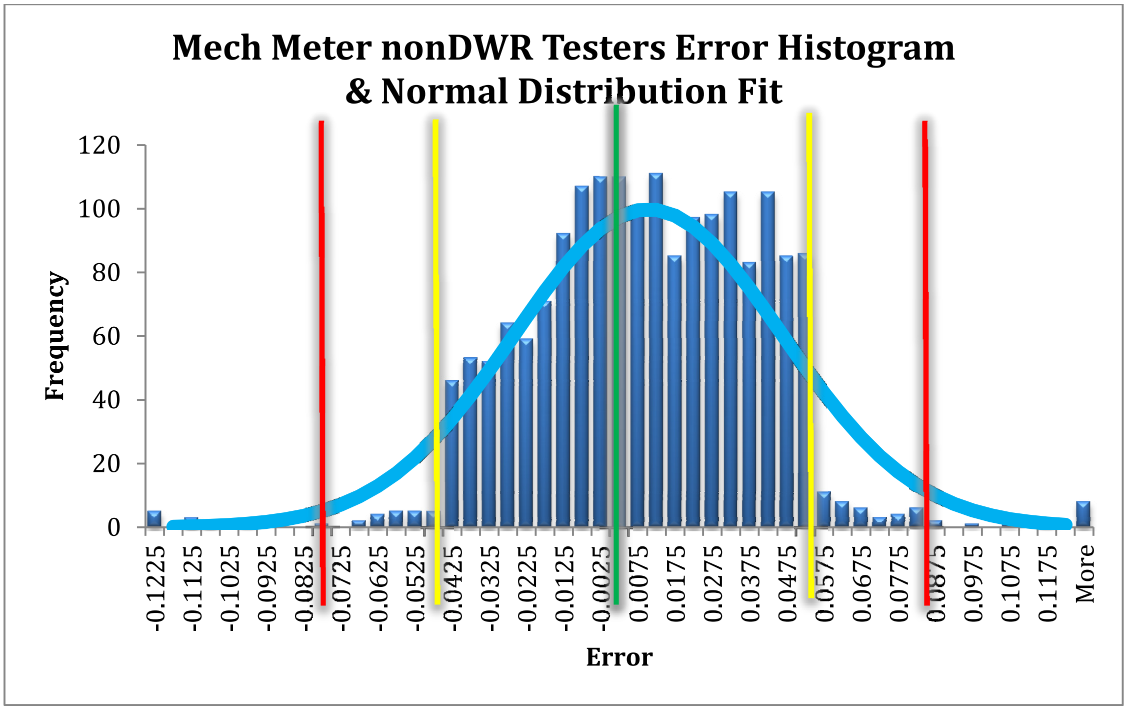


Figure 37 Mech Meter nonDWR Testers Error Histogram & Normal Distribution Fit

	DWR	Non-DWR
Mean	0.011062	0.007765
Standard Error	0.009239	0.000848
Median	0.02705	0.007749
Standard Deviation	0.08366	0.035999
Sample Variance	0.006999	0.001296
Kurtosis	41.77583	23.95198
Skewness	-5.50347	1.466024
Range	0.797942	0.709325
Minimum	-0.62537	-0.28916
Maximum	0.172574	0.420168
Sum	0.907062	13.99341
Count	82	1802
Confidence Level(95.0%)	0.018382	0.001663

Figure 38 Comparison of Descriptive Statistics for DWR vs. Non-DWR Testers

Below are magnetic and mechanical meter test error distribution graphs for selected individual meter testers out of the approximately fifty identified in the GWDMS database. Graphs for other testers can be generated using the online GWDMS data viewer. The graphs should show whether individual tester error distributions for magnetic and mechanical meters are similar to the basinwide data, and also whether the magnetic meter error patterns persist for Division 3 testers.

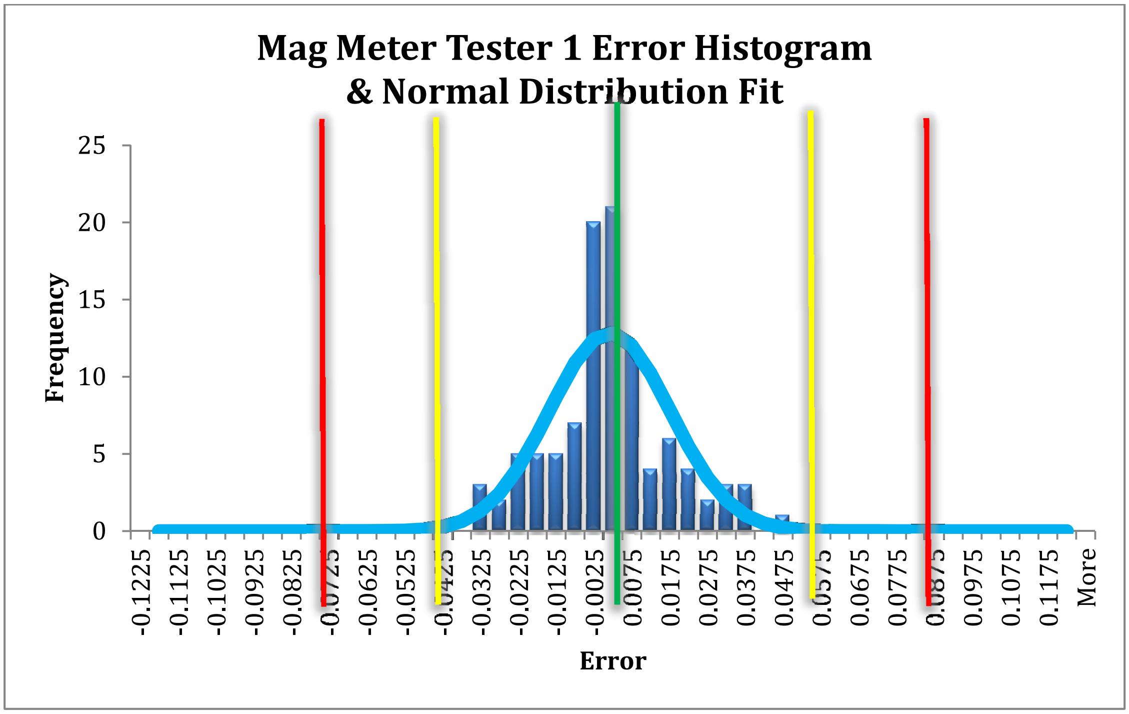


Figure 39 Mag Meter Tester 1 Error Histogram & Normal Distribution Fit

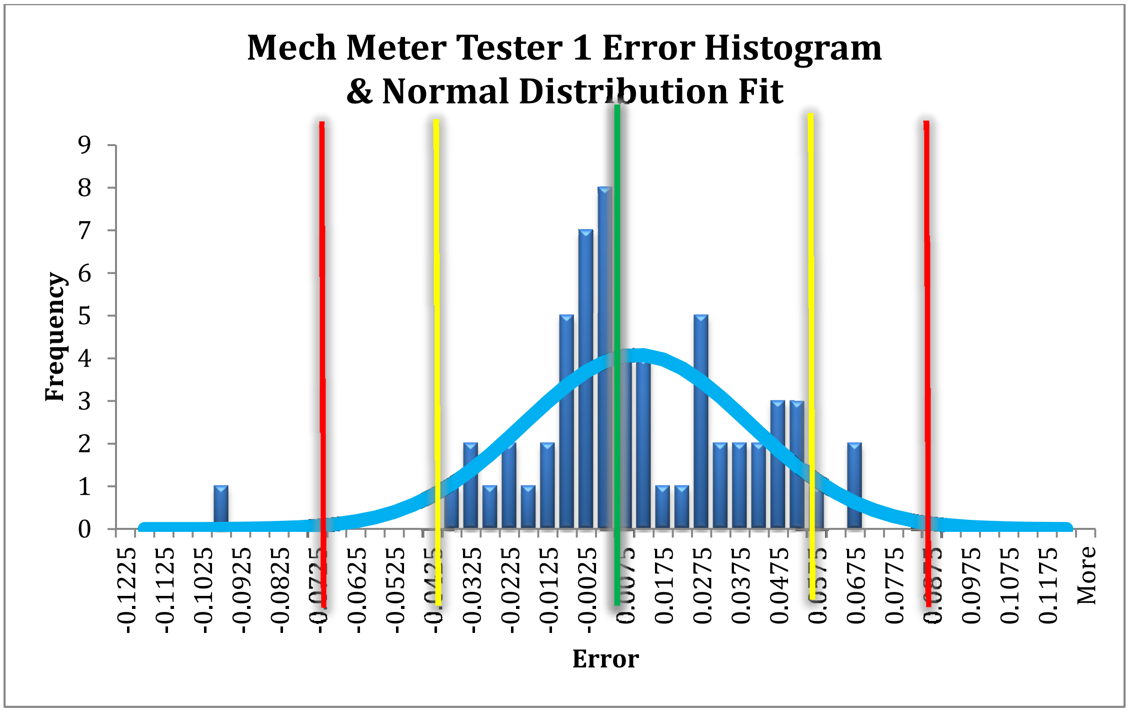


Figure 40 Mech Meter Tester 1 Error Histogram & Normal Distribution Fit

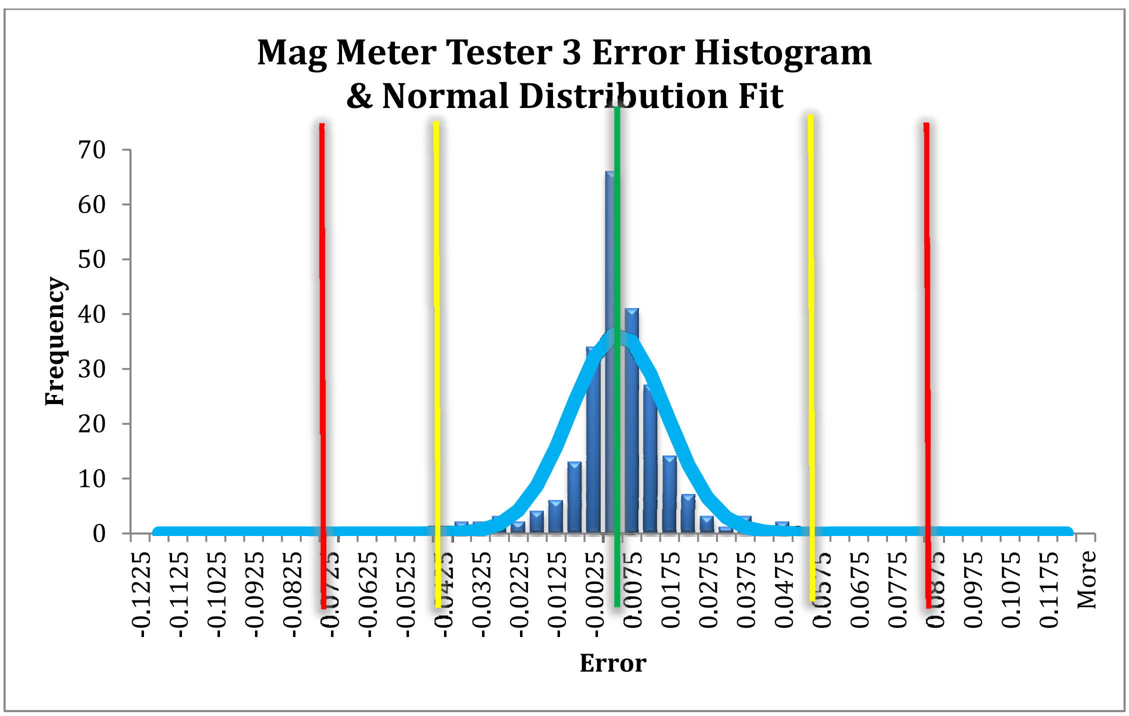


Figure 41 Mag Meter Tester 3 Error Histogram & Normal Distribution Fit

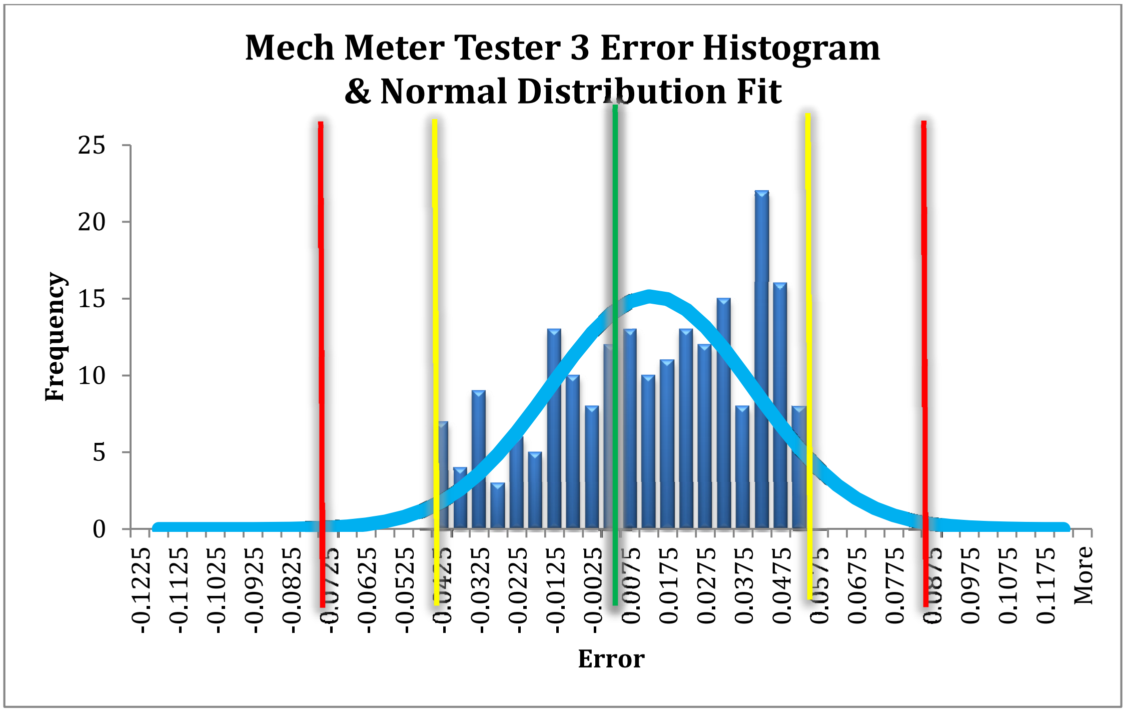


Figure 42 Mech Meter Tester 3 Error Histogram & Normal Distribution Fit

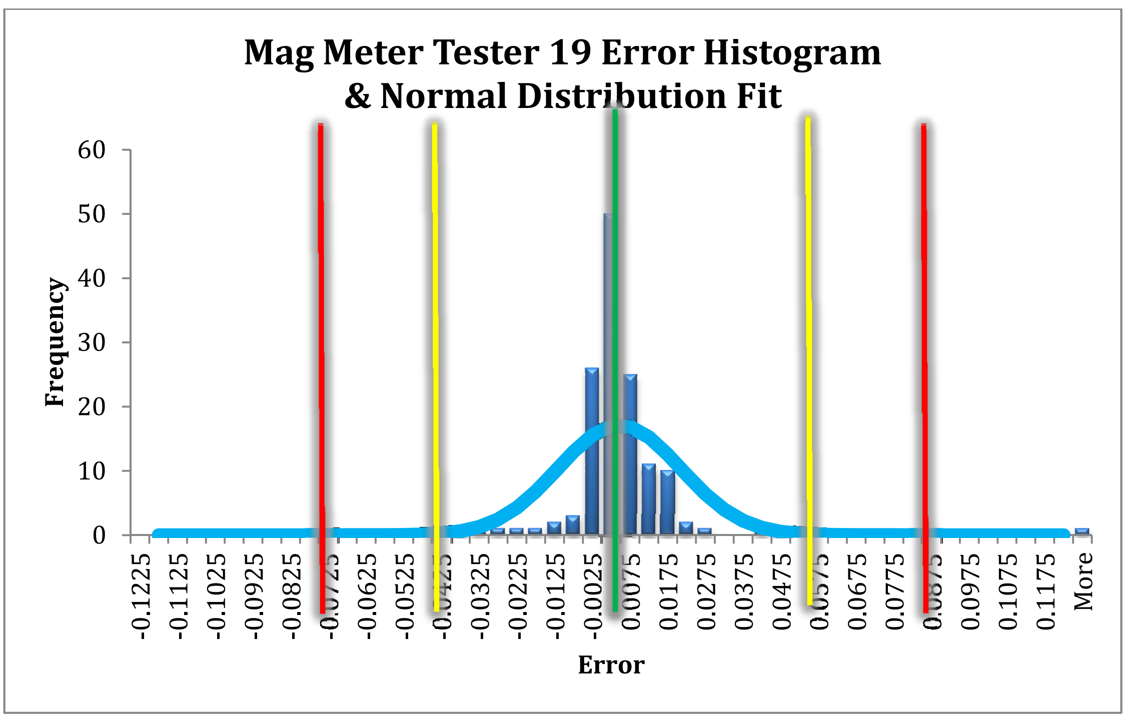


Figure 43 Mag Meter Tester 19 Error Histogram & Normal Distribution Fit

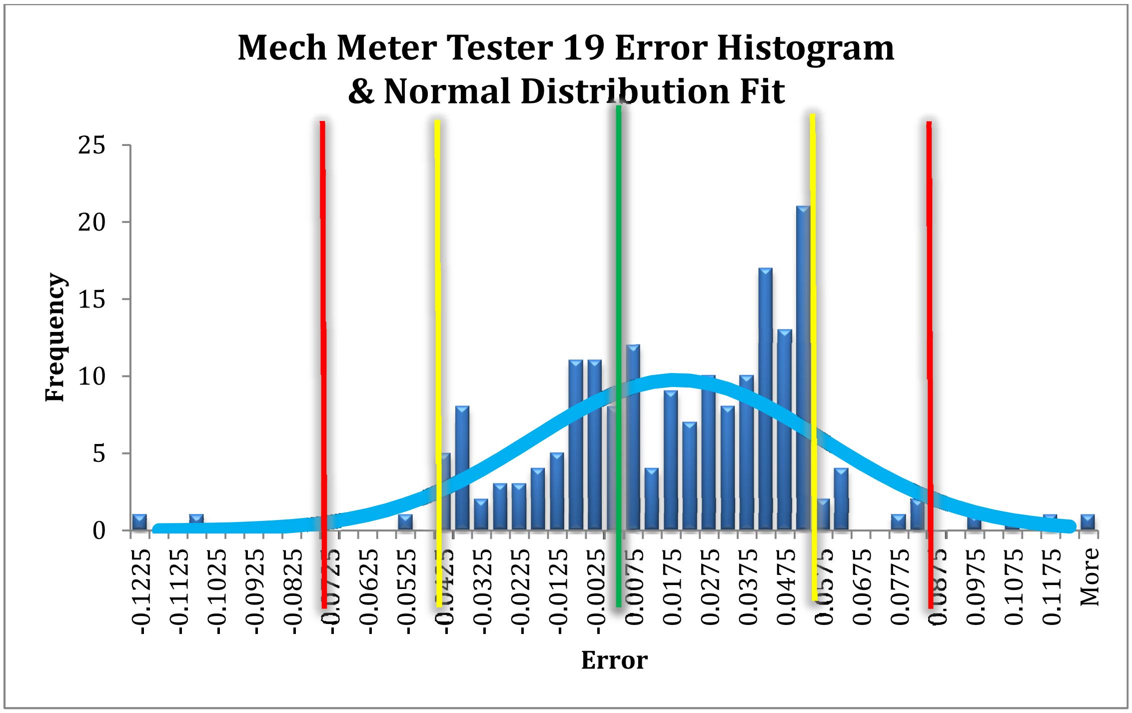


Figure 44 Mech Meter Tester 19 Error Histogram & Normal Distribution Fit

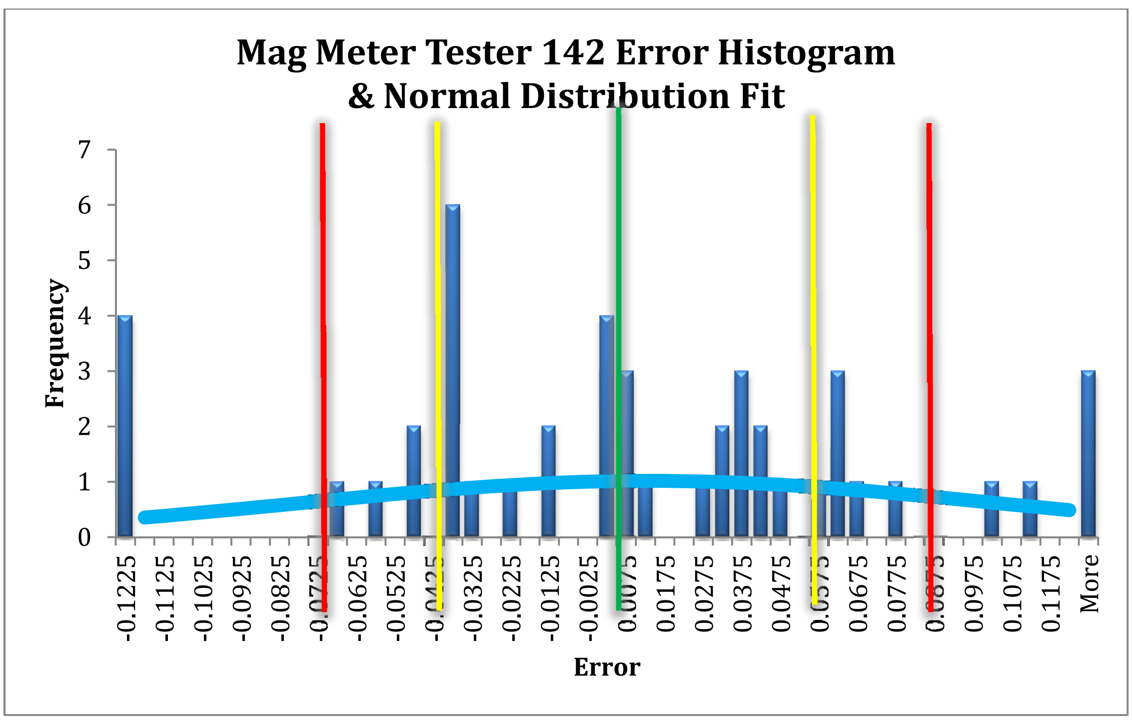


Figure 45 Mag Meter Tester 142 Error Histogram & Normal Distribution Fit

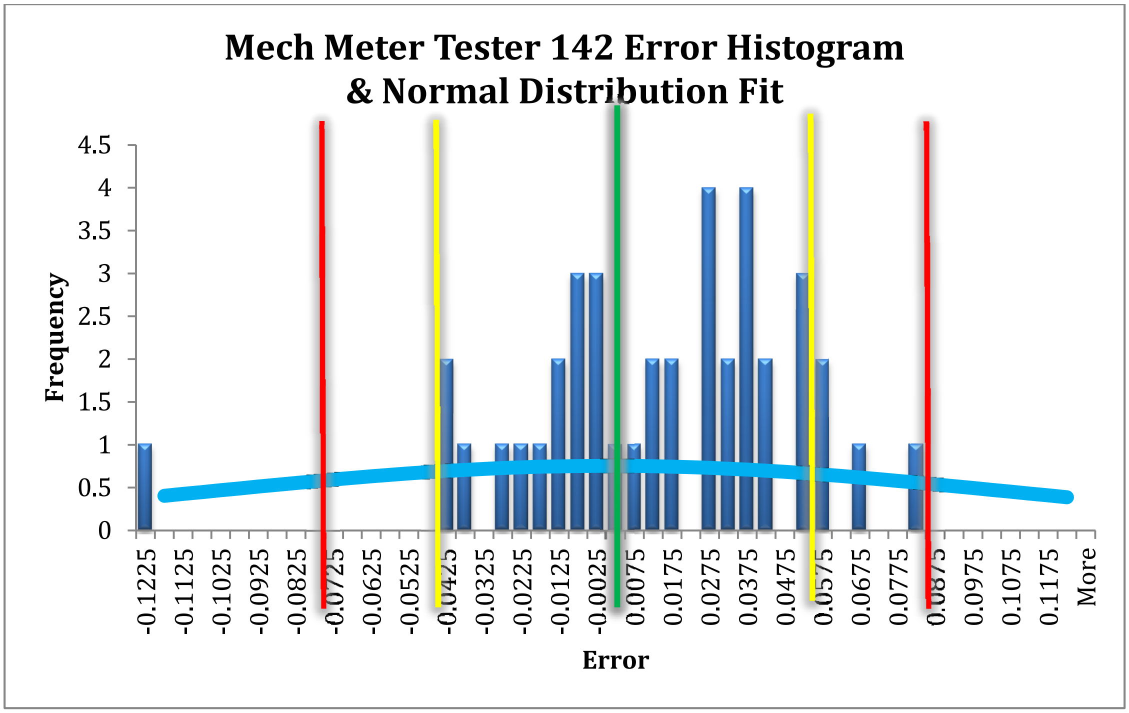


Figure 46 Mech Meter Tester 142 Error Histogram & Normal Distribution Fit

The previous graphs demonstrate that the error distribution patterns in the aggregate data also surface in individual tester data. As discussed before, magnetic meter data for non-Division 3 testersshows less error because of the K-factor adjustments being made before the submitted test data occurs. Mechanical meter tests by non-Division 3 testers show a sharp shoulder at the regulatory boundary (+/- 5%) because failed tests result in meter repair/replacement and non-submittal of the failed test data.

It is interesting to note the wide variation of DWR tester 142 errors for mag meters. Assuming this tester did not pre-adjust the K-Factors before the tests, it provides some evidence that the tight error distributions in magnetic meter test data are more likely the result of the K-Factor adjustments and not that magnetic meters are inherently more accurate.

4.2.4 Impacts of Meter Errors on Aggregate Pumping

When error distributions are non-random (skewed, not symmetrical, not centered on 0) it can potentially affect the accuracy of aggregate pumping totals.

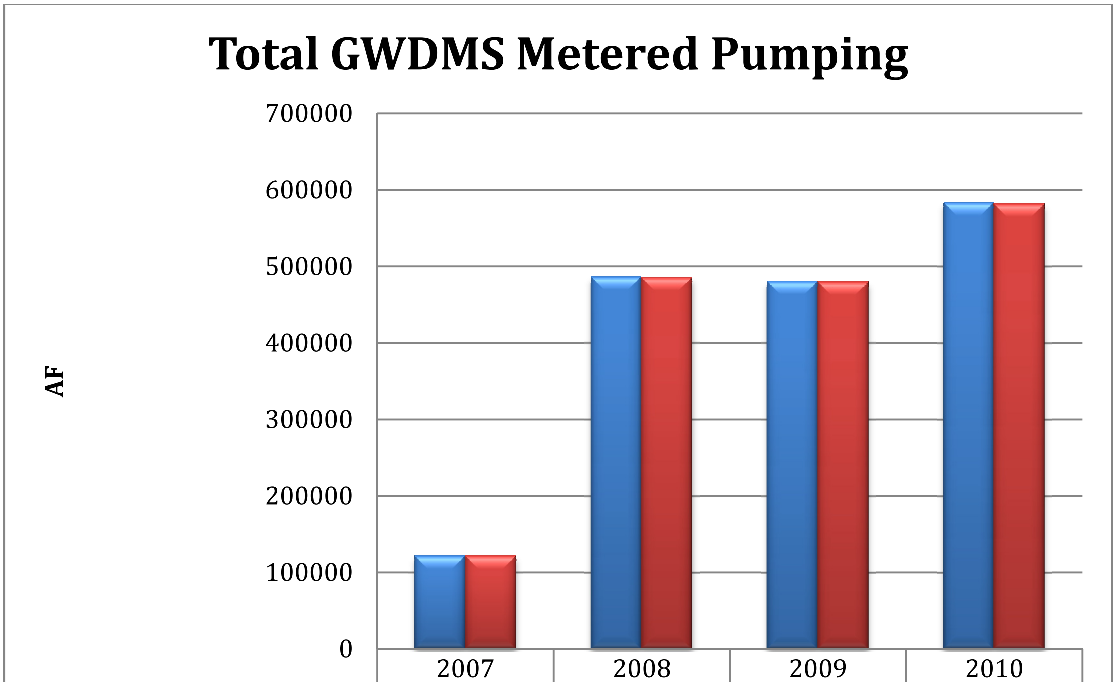
The current Rules and Regs specify that the metered pumping values are used **unadjusted** if the effective meter test at the time of the measurement has a correction factor between 0.95 and 1.05. If the error distributions are random, then over any sufficiently large set of these measurements, the meter errors effectively cancel and then the aggregate pumping value should be an accurate estimate. But if the error distributions are skewed and/or not centered on 0, then some error will be introduced when calculating aggregate totals.

For meters with correction factors between .92-.95 or 1.05-1.08, the meter values are adjusted using the correction factors and therefore the errors are removed from the data.Meters with correction factors less than .92 or greater than 1.08 are not used at all. It is desired to know whether aggregate pumping totals are significantly affected by not adjusting the reading using the calculated correction factor for meters with correction factors between 0.95 and 1.05.

The tables below show aggregate pumping calculated two ways, the DWR method ("DWR Calculated Pumping") described above and the actual metered pumping ("Metered Pumping") using each meter's measured error correction factor. Any difference will be caused by the error distribution anomalies.

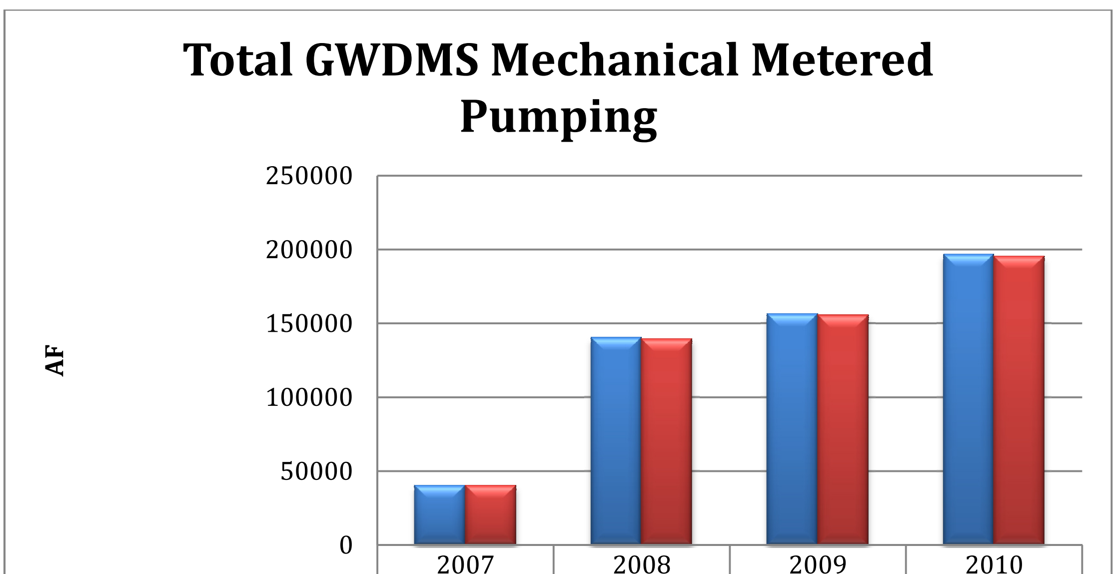
Note that 2011 data is not displayed because the 2011 meter reading adjustment process had not yet completed when the GWDMS database was provided to LRE. Using this data would lead to spurious differences that contain no useful information.

Fortunately, all the following graphs demonstrate that the aggregate pumping totals are uniformly accurate across the various aggregation types, and that therefore any error distribution imperfections in the meter test errors do not significantly affect metered pumping totals.



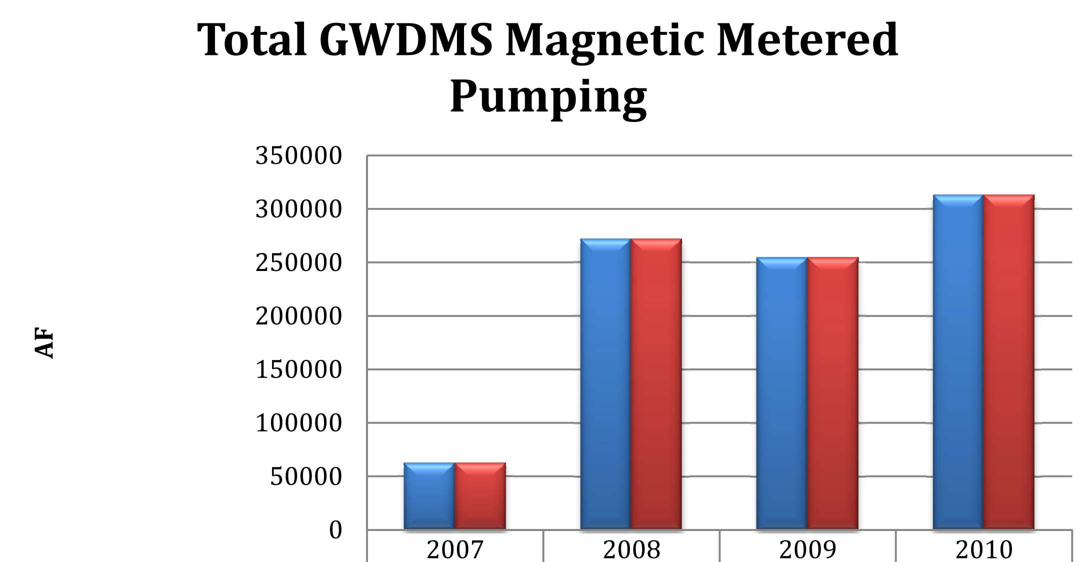
	2007	2008	2009	2010
DWR Calculated Pumping	121086	486148	480210	582935
Metered Pumping	121129	485244	479131	581536

Figure 47 Total GWDMS Metered Pumping



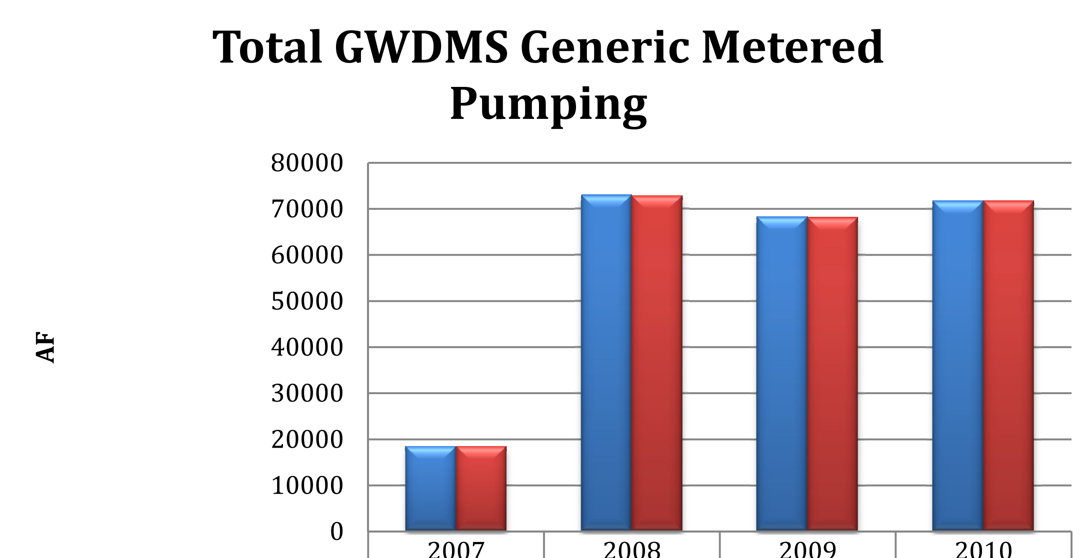
	2007	2008	2009	2010
DWR Calculated Pumping	40372	140295	156504	196867
Metered Pumping	40337	139510	155409	195465

Figure 48 Total GWDMS Mechanical Metered Pumping



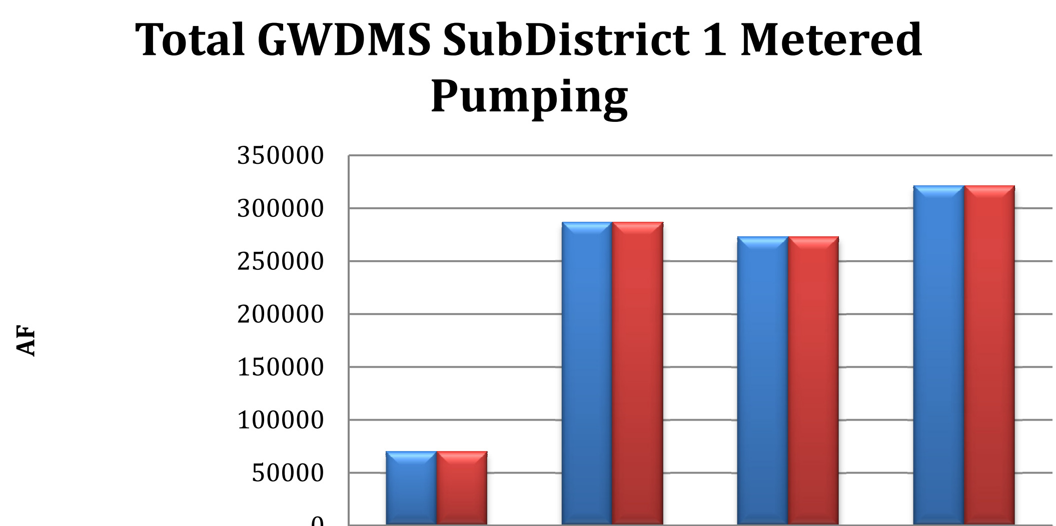
	2007	2008	2009	2010
DWR Calculated Pumping	61966	271427	254098	312790
Metered Pumping	62114	271420	254190	312826

Figure 49 Total GWDMS Magnetic Metered Pumping



	2007	2008	2009	2010
DWR Calculated Pumping	18418	72943	68157	71668
Metered Pumping	18348	72825	68075	71634

Figure 50 Total GWDMS Generic Metered Pumping



	2007	2008	2009	2010
DWR Calculated Pumping	69454	286734	272794	321183
Metered Pumping	69462	286464	272550	320870

Figure 51 Total GWDMS SubDistrict 1 Metered Pumping

5 CONCLUSIONS AND RECOMMENDATIONS

1. The design of the Division 3 data collection systems, rules and regulations are highly informed by previous experience (Division 2). It has evolved to incorporate better data management designs, modern technology (GPS, databases and web-based tools) and also local understanding (Division 3 staff and other experts).
2. The DWR database development team created effective data management systems for data collection and processing. Just as important, they continue to work with Division 3 and local entities to expand and evolve the systems. This is critical since data understanding is improving, data needs and uses are increasing, technology gets more powerful, etc. In order to be successful, systems like this must adapt because the objectives and expectations change so quickly.
3. The combination of a well-designed system of rules and regulations, modern GPS and metering technology, experienced and dedicated Division 3 staff, efficient DWR data management tools and the cooperation of water users in the basin has yielded high quality groundwater pumping diversion data for 2009 through 2011 in HydroBase.
4. An important data integrity improvement to the system should be to maintain the data and processes within the DWR database systems. This will help prevent additional error from being introduced during the QA/QC and meter to well calculations. It also will make it much easier to incorporate updated information (updated meter readings, updated meter to well connection data, etc.) and then recalculate well diversion data. And it will allow the database to directly show how any well diversion value was derived.
    a. According to recent discussions with DWR, new database and data management systems (HydroBase Data Management Console – HBDMC) are currently being designed and developed and will be coming online soon, in time for the 2012 data.
5. The analyses of meter *test* data demonstrate the importance of the meter testing program:
    a. Increases confidence in the collected meter readings
        i. Quantitative and objective
        ii. Even though some meter error data collection practices (K-factor adjustments on mag meters, replacing failed mech meters before submitting test results) have affected the data in the GWDMS, the error distributions generally are still centered on 0 and symmetrical in aggregate. Therefore aggregate pumping totals will be accurate because of resulting error cancelling behavior.
    b. Increases understanding of measurement technology
        i. Can compare error distributions and temporal changes of mechanical vs. magnetic meters.
    c. Provides objective information about measurement procedures and activities
        i. Mechanical meter error distributions indicate non-random (i.e. human bias) behavior at the critical +/- 5% and +/- 8% regulatory allowable meter error thresholds.

            1. This is likely caused by quick repair/replacement of a meter and then a submittal of successful test data rather than the submittal of failed test data.
            2. In the future, mechanical meter failed test results should be submitted to Division 3, even when the meter is going to immediately repaired/replaced .
        ii. Currently magnetic meter error data is biased because of K-factor adjustments prior to recording test results.
            1. Therefore the data represents post-adjustment error, not raw meter error.
            2. Post adjustment error distributions appear random, except more highly center weighted ("peaked") than a normal distribution.
            3. The test data collection program should be expanded to collect pre-adjustment error data. [Division 3 has indicated this will happen]

6. Meter readings
    a. In aggregate the collected meter readings time series data quickly became asymptotic, reaching a constraint that is likely the measurable Division 3 well pumping total.
    b. The non-random patterns observed in the mechanical meter error distributions did not significantly impact aggregate pumping totals.

The data and analyses in this report indicate that the Division 3 Well Pumping Data Collection Program has been successful in collecting accurate measurements of meter pumping data, in calculating annual well pumping estimates, in developing efficient and effective data management systems, and in continuing to improve the processes and systems.

LeonardRice
ENGINEERS, INC.

Leonard Rice Engineers, Inc. provides consulting services related to planning, management, and development of water at the highest level of technical excellence that is professional, ethical, and profitable.

Our company is devoted to serving our clients, developing staff, effectively using technology, communicating effectively, continuously improving, and contributing to our community.

We foster a challenging, stimulating, and fulfilling work experience.