BEFORE THE WASHINGTON UTILITIES AND
TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

DIRECT TESTIMONY OF THOMAS D. DUKICH
REPRESENTING AVISTA CORPORATION

Exhibit T___(TDD-T)

I. INTRODUCTION

Q. Please state your name, position, and business address.

A. My name is Thomas D. Dukich and my business address is S. 7222 Brookshire Court, Spokane, Washington, 99223. For 20 years I held various management positions in the Rates and Regulatory Affairs Department at Avista including the position of Director. Prior to that, I was the Supervisor of Research and Forecasting for the Company. In that position, I conducted and directed numerous studies involving statistical sampling and statistical analysis as well as producing the Company's short-term load forecast. My last day of full time employment for Avista was November 1, 2001.

I joined Avista in 1978 after having been previously employed for seven years as an Assistant and then an Associate Professor at Gonzaga University where I taught courses on research methodology. I have also taught Market Research at Eastern Washington University. I am now an independent consultant testifying on behalf of Avista.

Q. Would you describe your educational background?

A. I graduated from the University of Minnesota in 1967 with a B.A. in Psychology and Business and from the University of Montana in 1972 with M.A. and Ph.D. degrees in Experimental Psychology and Statistics. Both my undergraduate and graduate training included numerous courses in statistics and I have subsequently attended courses and seminars on strategic planning, Box-Jenkins forecasting, finance, accounting, rate design and pricing.

Q. Have you previously testified before this Commission?

A. Yes. I have appeared before this Commission on numerous occasions.

Q. Would you briefly describe your assignment for Avista?

A. I was asked to examine what historical period of streamflows would likely best reflect pro forma power costs for Avista.

SUMMARY OF TESTIMONY

Q. Would you please summarize your testimony?

A. I first examined the kinds of questions we try to answer when we look to the historical water year record during the rate setting process. After formulating what I believe to be the most relevant questions, I then follow with four different approaches that I feel are responsive to those questions. All four approaches lead me to the same conclusion: the most appropriate course of action is to use all the available data applicable to Avista's resources. In this, case 60 years of streamflow record.

I also reviewed the 1985 testimony of Commission Staff witness Mr. Winterfeld in Cause No. U-85-36 ("Winterfeld testimony"). I did so because the Commission relied extensively on his testimony when it ordered the use of the rolling 40-year average in that fully litigated Avista rate case. Much subsequent discussion has centered on Mr. Winterfeld's conclusions. After a detailed examination of Mr. Winterfeld testimony, I concluded that his results are largely driven by the mechanics of the way he set up his analysis. I don't believe they have much merit when it comes to actual streamflow data.

INTRODUCTION

Q. What is needed by the rate making process with regard to hydro generation.

A. Pro forma, normalized power supply cost should reflect the expected cost of serving customer loads. If they do not, the Company will chronically under recover or chronically over recover costs. Put another way, customers will either chronically underpay or overpay.

Q. How is this cost determined?

A. As Mr. Norwood and others Company witnesses cover in detail in their testimonies, a period of historical water years, <u>each</u> <u>one</u> separately run through the power supply model, are used to determine what Avista expects costs to be in the future. For ratemaking purposes we should also ask: 1) what period of future cost are we trying to reflect in rates, and 2) how many historical water years should we run through the model to best reflect a given future time period?

Regardless of ones view on how much of the historical record should be used, we ought to determine as best we can the answer to the question regarding what period of future costs we are trying to reflect in rates. The answer to this question is important because it determines the "dependent variable" in any analysis. Our success should be measured against this variable. If we are interested in only next year's costs, then that is the benchmark. If we are interested in the average cost over the next five years, then that's the benchmark, and so on. If we decide we are estimating cost to some indeterminate length of time, that may demand a different analytical approach.

Q. What is your opinion as to what period of future costs should be reflected in rates?

A. Based on my analysis here and on my experience with the rate making process, I believe an average of a five-year or longer period is reasonable.

Q. What about the question regarding what historical record should be used to estimate these future costs?

A. No matter what period of future costs we choose to use as the criterion, I conclude that using 60 years of record is the best overall approach.

Q. Why did you draw that conclusion?

A. My reasons fall into four different categories and represent four somewhat different approaches to the problem: one approach might be called theoretical, one empirical, one traditional, and one anecdotal. I'd like to begin with the theoretical.

THEORETICAL

I realize that there is a risk of eyes glazing over when a theoretical discussion is presented in testimony. However, I believe in the current instance a theoretical discussion is extremely relevant and should not be ignored. Over the years, it has been interesting to me to observe how during the rate making process we often struggle to make sure that decisions are consistent with the rule of law and the principles of accounting. I believe that we should do no less when it comes to the principles of statistical analysis as they apply to decisions regarding the selection of the number of water years.

There is a mathematical theorem that has been gaining more exposure in the statistical literature over the past several years even though this theorem was proven true in 1933. It goes by the name of the Glivenko-Cantelli Theorem, after the two mathematicians who first published the proof of the theorem. How various experts have described the Glivenko-Cantelli Theorem reflects its importance. It has been called the existence theorem for statistics as a branch of applied mathematics. It has also been called the fundamental theorem of statistics. It has been said that, in effect, without the Glivenko-Cantelli Theorem modern statistics would not exist, as we know it today. It is <u>the</u> central truth of modern statistics. It is even more fundamental than the Central Limit Theorem, which is the foundation of much inferential or sampling statistics.

Q. What is the essence of the Glivenko-Cantelli Theorem?

A. What Glivenko-Cantelli says is that as you build your observed or empirical distribution function by collecting more and more data points, this empirical distribution will converge to the true distribution. In other words, more data is better if you wish to reflect the true distribution function, or as statisticians frequently say, if you wish to reflect the true state of nature. And, as you collect more and more information, you get closer and closer to the true distribution, how close you get is a direct function of how much information you have collected.[1] This seems intuitively obvious, but until Glivenko-Cantelli, convergence to the true distribution was not mathematically proven.

In my opinion, the analysis of how many historical water years to use really needs to go no further than Glivenko-Cantelli. It is clear to me that the Glivenko-Cantelli Theorem indicates we should use as much reliable data as we have. Sixty (60) years is preferable to 40, if we are interested in most accurately reflecting how water years are truly distributed across different values. And as Mr. Norwood and others testify, it is 60 individual years, not the average of the 60 that are modeled to reflect pro forma power supply costs.

In sum, more data is better if you wish to reflect the true state of a distribution. This fact is the foundation of modern statistics. I hope that it can also become one of the foundations of utility regulation, along with the rule of law and the principles of accounting.

EMPIRICAL APPROACH

Q. Would you now describe what you have called the empirical approach to the water year issue?

A. Yes. The empirical approach follows from the theoretical discussion of Glivenko-Cantelli. It is the practical application side. In this approach, I have used what has come to be called "computer intensive" or "resampling" methods to answer questions regarding water years.

Q. Would you explain the term "resampling".

---

[1] A little more formally, the Glivenko-Cantelli Theorem can be stated as follows. The empirical distribution function f_n estimates F to any desired degree of precision uniformly in x for sufficiently large sample size n. The true distribution function F can be rediscovered from the data; or the empirical distribution function f_n can be said to look like the true distribution function F for large sample size n.

A. Yes. The resampling method repeatedly draws random samples from the existing data in order to gain insights into the "true" distribution. The sample is said to bootstrap itself into the true distribution following the logic of the Glivenko-Cantelli Theorem. As a result, resampling is also called "bootstrapping" in the statistical literature.

Q. Is resampling the same as what is called a Monte Carlo method?

A. Not really. In Monte Carlo, samples are usually drawn according to some assumed probability distribution to reflect the range of possible outcomes. In resampling, no assumption is made about the underlying distribution and no a priori distribution is used to determine the sampling plan. In fact, the goal of resampling is to rediscover the true distribution from the empirical data on hand through the resample methodology.

Q. Is resampling a proven technique in common use?

A. Yes. One author, David Salsburg, had this to say about resampling: "The implications of this method have been so extensive that almost every issue of the mathematical statistical journals since 1982 has contained one or more articles involving bootstrap." (The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century, Freeman, 2001.)

Q. How did you apply the resampling methodology to the water year question?

A. I started with 60 years of natural inflow data measured at the Dalles in cubic feet per second. In my testimony I refer to this as simply streamflow. As Mr. Norwood testifies, this is all the streamflow data we have that is directly applicable to Avista's unique configuration of hydro generation.

However, just because we start with a 60-year database does not mean that fewer years might not in some way be proven to be a better reflection of the average streamflows or power supply expenses. But of course, it follows from Glivenko-Cantelli that 60 years are best if we are interested in working toward approximating the true state of nature.

I applied the resampling methodology to the water year question, by breaking the analysis down into a series of questions that relate back to my earlier discussion regarding what the pro forma should reflect. First I asked, does using more historical years produce a more accurate estimate of the average or the distribution of streamflows over the last 60 years? Then I asked what might we expect if we try to predict streamflow for a single year versus what we might expect if we try to predict it for an average of a 5-year period or for an average of a 10-year period of streamflows?

Q. Starting with the first question regarding whether larger samples produce more reliable estimates of the average streamflow, what conclusions did you draw?

A. Using more water years clearly produces more reliable estimates of the average streamflow. My conclusion is based upon the resampling experiment described below. Often times a particular resampling methodology is called an "experiment" because there is more emphasis on setting up the proper questions to ask rather than on the statistical formulae involved.

Starting with the 60 historical water years, I drew 1,000 different random samples, with replacement, of various sizes, i.e., 1,000 individual random samples of 30, 1,000 random samples of 40, 1,000 random samples of 50, and 1,000 individual random samples of 60.

For each of the 1,000 individual samples of a given size, I calculated the mean value. Then I calculated the average of all the mean values for each sample of that given size (the grand mean or the mean of the means). I also calculated how far each of the 1,000 individual means was from the grand mean (the average deviation from the grand mean). Then I plotted those values as shown in Exhibit ___(TDD-1).

Several things are worth noting in Exhibit ___(TDD-1). First, the grand mean across different sample sizes is very stable. This is what we would expect since each of the samples is estimating the same "true" mean. It's also quite clear that as sample size increases, the average deviation around the grand mean decreases significantly. Exhibit ___(TDD-2) shows this same average deviation information but with a chart that is purposely scaled to focus attention on just the changes in the average deviation as a function of sample size.

One way to interpret these results is to conclude that as sample size increases, the reliability or accuracy of the estimate increases, i.e., the error of the estimate decreases. We may be tempted to compute traditional "confidence intervals" around our estimates in order to quantify this decrease in estimating error. Recall however that this is a resampling experiment. We make no assumption about the parameters of the underlying distribution. So we probably should not, for example, use plus or minus 1.68 times the standard deviation of the means to calculate the 90% confidence limit. But I have saved that calculation for the discussion of the traditional approach.

We can, however, express the improvement in our estimate in a similar way. It has became traditional in resampling experiments to use percentiles instead of standard deviation or standard error units. So for the result in Exhibit ___(TDD-3), I show the 5th and the 95th percentile ranks. One way to verbalize these results is to say that 90 times out of a 100, the mean will be included in this range. It has also been common to say that we are 90% confident that the <u>true population mean</u> is between these values. I will return to Exhibit ___(TDD-3) again later on in my testimony when I address the question of how well the current 40-year rolling average represents the average streamflows.

In sum, it is very clear from these results that using more water years clearly produces a more reliable estimate of the average streamflow, i.e., 60 years is better than 50 which is better than 40 which is better than 30.

Q. Would you now turn to the next questions. What might we expect if we try to predict streamflow for a single year versus what we might expect if we try to predict it for an average of a 5-year period or for an average of a 10-year period of streamflows?

A. It takes some effort to describe how this analysis was accomplished. Let me start by explaining what I did for the question regarding what best predicts the streamflow for a single year? Starting with the 60 historical water years, I randomly selected, with replacement, 100 different individual single years.[2] Then for each one of those single years I examined what kind of estimating errors would occur using 30, 40, 50, or 60 years of history to forecast that fixed single year. I did this 100 times. The end result is that for <u>each</u> randomly selected single year, there are 400 attempts to forecast that year, 100 using 30 years of history, 100 using 40 years, 100 using 50, and 100 using 60 years of history. The single year question involves 1,800,100 data points.

The same process was repeated for the question of what best predicts a five-year period. I again begin with the 60 historical years of water but this time randomly select, again with replacement, 100, five-year periods and took the mean of each of those five-year periods. For the mean of each five-year period, I again examined what kind of estimating errors occurred using 30, 40, 50, or 60 years of history to forecast that particular five-year forecast period. I did this 100 times, once for each five-year period. Again, the end result is that for each randomly selected five-year period, there are 400 attempts to forecast that

---

[2] It is possible to select 100 years out of 60 because we are sampling <u>with replacement</u>. This means that a single year may be selected more than once. We could sample <u>without replacement</u> which would, of course, limit the total number of single samples to 60. But sampling without replacement leads to dependencies that are difficult to evaluate even with a resampling methodology. It is reasonable to sample with replacement because there are no dependencies in the streamflow data and the same level of streamflow does in fact occur more than once in the actual 60-year history. For example, in 1950 and 1983 and again in 1942 and 1963 streamflows were within two, one-one hundredths of a percent (0.02%) of each other. If we round the data to the nearest 100 cubic feet per second, there are 21 "repeats" in the actual 60-year record.

period, 100 using 30 years of history, 100 using 40 years, 100 using 50, and 100 using 60 years of history. The five-year question involves 1,800,500 data points.

An analogous procedure was followed for the question of what historical period best predicts the average of a 10-year period of streamflows. There were 1,801,000 data points are involved in evaluating the 10-year average.

For the three periods, a total of 5,401,600 data points were analyzed. Of course this helps explain why resampling is classified as a "computer intensive" endeavor.

Q. What did you conclude as a result of this resampling experiment?

A. Attempting to forecast a single year is very difficult no matter how many prior years are used to do it. Forecasting five-year and ten-year averages is much easier in the sense that the possible variation in a five and in a ten-year average is much smaller than the variation in a single year. In fact, the single year average deviation is over two times the deviation of the 5-year and the 10-year averages.

The results are further summarized in Exhibit ___(TDD-4) and Exhibit ___(TDD-5). Exhibit ___(TDD-4) shows the 90% confidence interval for the average deviation and Exhibit ___(TDD-5) shows the shape of the distributions for the single year, for the 5-year averages, and for the 10-year averages.

Q. What about trends and cycles in the historical data? Wouldn't that justify using a continuous record of just the most recent data rather than examining various non-continuous periods?

A. There is no evidence that I am aware of that indicates that there are trends or cycles in the adjusted streamflow data. Mr. Winterfeld also concluded that there are no such trends in streamflows. It is reasonable to conclude that streamflow is a "continuous random variable" in statistical terms, meaning that knowing what happened in any one year does not significantly improve the chance of forecasting what will happen in any other year.

TRADITIONAL

Q. The third area you mentioned is the traditional approach. How is that different from the empirical approach you just discussed?

A. The empirical approach assumes nothing about the shape of the underlying distribution or the true state of nature. The traditional approach usually assumes that the underlying distribution is normally distributed or at least that the sampling statistic that is used is normally distributed. For example, because of the Central Limit Theorem, it can be safely assumed that the distribution of sample means drawn from the historical water data is normally distributed even though the distribution of individual data points themselves may not be normally distributed.

The traditional approach is dependent upon picking the correct formula for calculating the sample statistic, the standard error of the means, etc. This may not be a problem when we ask questions such as: what is the 90% confidence interval around the mean of a sample of 40 water years? However, when we ask more complicated and unique questions, like just discussed in the prior section, finding the correct formula may involve some very clever mathematics.

However, even without the advanced mathematics, I believe it is still valuable to examine the more traditional approach to see what insights it may offer. It is also an approach that many of us are more comfortable with. I suspect that will change, as the resampling approach becomes better known.

Q. What analysis did you conduct?

A. To begin, it should be pointed out that we can view this problem two different ways. One way is to treat the flows that have occurred so far as the total population. In that case, there is no need to use inferential statistics since we already know the true state of nature. The other and more traditional way is to assume that the flows that have occurred so far are a sample of all flows that have or will ever occur. In this latter case, we don't know the true state of nature. This distinction can matter.

For example, since we can't assume that the historical mean to date is the true state of nature, we really must be careful about using that mean as a criterion to make definitive judgements about one method or the other. We don't really know the true state of nature. However, the historical mean to date is the only estimate we have so we can, with some caution, use it as a guideline for making necessary judgements.

I treated the historical record as a sample of the unknown true state of nature. I looked at what we might expect to conclude about the reliability of estimating the mean streamflow given sample sizes of 30, 40, 50, and 60 years. The results are shown in Exhibit ___(TDD-6) and in Exhibit ___(TDD-6).

Q. What did you conclude?

A. As we might expect, the larger the sample size, the narrower the confidence interval around the mean, and the more reliable the estimate of the mean streamflow. Mr. Winterfeld drew the same conclusions in his testimony. This comes about because the formula for the standard error of the mean, the variance of the sample means from the estimated grand population mean, has sample size in the denominator.[3]

Each of the samples makes an estimate of the very same true population mean, a major difference being that ones with smaller sample size produce less reliable or accurate estimates. In other words, the variance of the estimates is greater for smaller sample sizes. Since it is no more expensive to use 60 years than it is 40 years for our sample, the data having already been collected, it is most reasonable to use as large a sample as we can,

[3] The formula for the standard error of the mean is the standard deviation of the sample observations divided by the square root of the sample size. So as sample size increases, we would expect the standard error of the mean to decrease and, at the same time, confidence limits to narrow, all things being equal.

namely 60 years. And we should continue to add to the sample size, as the data become available. Again, 60 years is preferable to 50, to 40, to 30, and so on.

Q. Are these results consistent with the empirical approach you discussed earlier?

A. Yes. If we compare at Exhibit ___(TDD-) and Exhibit ___(TDD-) we see the same pattern. For both the resampling methodology and for the traditional approach, estimates of the mean are more reliable as sample size, in this case the number of historical water years, increases.

ANECDOTAL APPROACH

Q. Would you turn now to the fourth area you mentioned, or the anecdotal approach.

A. The term anecdotal is sometimes used to reflect an unscientific report. However, in this instance what I mean to reflect by this term is that with very little analysis or basically just simple observation, certain conclusions could be drawn regarding the historical water year record.

I believe that we are all aware that Avista and the region are in the midst of record low streamflows. For Avista in particular, this is one of the very lowest recorded streamflow in the last 75 years. Mr. Norwood and Mr. Kalich both discuss the streamflow record in their respective testimonies. Prior to the 2001 water year, there had been speculation that the record lowest water year was a very rare event that we may never experience in our lifetimes. There was even the suggestion that perhaps it was reasonable to ignore very low water years since these were unlikely to ever occur again. As Mr. Norwood discusses, that is just what the use of the adjusted 40 water-year record accomplishes. It ignores the very lowest water years.

Q. What is your opinion of omitting certain years from the historical record because they appear to be unlikely to occur again?

A. It is interesting to note history behind this approach. It developed because it was assumed that extreme highs and lows were in fact "errors" indicating that something had gone wrong with the observation or experimental technique. Modern approaches to analysis which focus more on distributions have come to recognize that extremes likely reflect the underlying true distribution.

We may be tempted in the current situation to exclude the two highs and the two lows, for instance, because they may influence the mean of the streamflows more than we would like or because they produce rate impacts that are not pleasant to contemplate. As a rationale, we may even note that the mean streamflow is not significantly influenced by excluding these data points.

There are several problems with this approach. First, there is no indication that the extreme values in the streamflow data are the result of a failure or errors in the methodology used to measure the flows. Consequently, excluding these legitimate values biases the estimate of the true distribution. Secondly, the fact that the historical mean values is not significantly influenced by excluding some values is somewhat misleading since it is only an estimate of the true mean. In addition, it is quite possible to select any number of combinations of less than 60 values from the historical record that produce the same mean as the 60-year historical record.

Some of these combinations may well imply very unusual distributions such as one with no values in the "middle". For example, taking the average of eight relatively low water years along with two high water years may produce nearly the same mean as for all 60 years of the water record. But this is hardly a rationale for excluding the 50 other years that are more toward the center of the distribution of historical water years.

Lastly, pro forma power supply costs that are determined by modeling what would happen if each and every water year of the record we choose were run through the hydro resources that now exist. It is <u>not</u> the "average water year" that is run through the hydro plants that are now on the rivers. So, using the average as the criterion doesn't reflect how the cost is developed.

Nevertheless, excluding "outliers" as they are sometimes called, has become such a common practice that Microsoft's Excel spreadsheet program actually has a built-in function specifically designed to accomplish this purpose. It's called TRIMMEAN. I used TRIMMEAN to evaluate the impact of excluding various data points from the 60-year water record and compared those results to the 60, 50, and 40-year water record means. The results are shown in Exhibit ___(TDD-8).

There is not much impact on the mean until we exclude 12 water years from the analysis—the 6 highest and the 6 lowest. In other words, the mean of the series with excluded years is very close to the 60 water year record and remains significantly lower than either the last 50 or the last 40 years of continuous record.

Q. Would you summarize what you have called the anecdotal approach?

A. We now have had a new record low water year established for Avista, 2001. In my opinion, this event by itself is a validation for using the entire 60 years of historical streamflows if our interest is to reflect, as best we can, the true distribution and the true state of nature. It is also worth noting that, if history is any guide, it will likely be 20 years or more before the 2001 record low is actually reflected in the water record Avista uses to establish power supply costs. The current record goes through 1988 and the next 10-year update, taking us to 1998, will almost certainly not capture the 2001-water year.

WINTERFELD TESTIMONY

Q. The Commission relied extensively on Mr. Winterfeld's testimony when it adopted the 40-year rolling average standard in its order in the 1985 Avista case. What is your impression of his testimony?

A. First of all, Mr. Winterfeld examined several different kinds of data, <u>none of them</u>, as near as I could tell, were really streamflow data.[4] For example, he compares "good" and "poor" streamflows to a series of coin flips or tosses of a die. He then represents these outcomes with a series of ones (1) and zeros (0). A one (1) meaning that year was good and a zero (0) meaning the year was poor. Nowhere that I could find does he define good and poor in terms of the streamflow data, comparing it to above or below the mean for example. He then does a Monte Carlo computer simulation on this 1 and 0 data even though he could have calculated the exact probabilities and variances for the binomial

---

[4] Mr. Winterfeld repeatedly refers to his 1 and 0 data as "water conditions." In fact, he does so three times in a single paragraph on page five of his testimony stating for example, "Column (b) contains the water condition for years 1 through 40" and "adding intervening water conditions to the average." This creates the impression that streamflow data is being used when it is not. He even refers to the "40-year water record" on page six when he is referring instead to a series of forty 1s and 0s.

data series he created, as he himself admits. In addition, it is difficult for me to determine exactly how he did the Monte Carlo.

In his analysis Mr. Winterfeld eliminates the sign of the errors in order to guard against large positive and negative errors canceling each other out. Then he calculates an average deviation score. Normally, this same end is accomplished by calculating a traditional variance and standard deviation. Because of the non-standard way that Mr. Winterfeld calculated his deviations, it is difficult to interpret his results and put them the context of traditional analysis.

Toward the end of his testimony, just when I thought he would test the efficacy of his "model" on the actual 50 years of streamflow record that then existed, he switches to a 112 year interpolated streamflow series. This series was apparently created by converting the 50-year record to a 112-year record through a weighting scheme. He then ran additional simulations on this 112 years series but <u>not</u> on streamflows. Rather, he ran his 1-year rolling average simulations on the <u>revenues</u> produced by these data after they had been run through Avista's 1984 power supply model.

Q. Besides being confusing is there anything wrong with what Mr. Winterfeld did?

A. In my opinion, there is. Mr. Winterfeld "loses the data" as he progresses through his analysis. If fact, he <u>never</u> actually does any analysis on streamflows as far as I can tell. He did his analysis on a coin flip experiment, on the output of the power supply model, and on the prices as they existed or were forecasted in 1984. He also created a 112-year water record rather than using the 50-year record more relevant to Avista's resource mix and dam locations.

Q. But don't all these problems cancel each other out in the Monte Carlo simulation?

A. I don't believe so, partly because the various comparisons he does are so confounded with different levels of data. In addition, nothing is added by treating streamflow as a binary variable, i.e., as either good or poor. In fact, much is lost and misleading conclusions likely have been drawn.

Q. Would you please give some detailed examples of the problems you found in Mr. Winterfeld's methodology?

A. Yes. First of all, recall that Mr. Winterfeld concludes that the most recent continuous 40 "water years" of history are preferable to using 50 water years because the "cumulative error" is greater for 50 "water years" than for 40 water years.[5] Recall again that his analysis did not include real water years but rather 0s and 1s with no defined reference to any actual streamflow parameter.

In addition, I believe that Mr. Winterfeld's conclusion favoring 40 years of historical water is largely driven by an artifact in his methodology. In other words, the arithmetic necessarily follows merely from the mechanics of the way he set up his comparisons even though Mr. Winterfeld himself may not have been aware of it. And because so much has been made of Mr. Winterfeld conclusions, it is worth while to examine his methodology in some detail.

I have recreated a portion of Mr. Winterfeld's Exhibit T-76 in my Exhibit ___(TDD-9). In Exhibit T-76, Mr. Winterfeld examined only a single <u>future</u> water year. Later he studies only <u>continuous</u> running averages even though he acknowledges that there are no

[5] We now have 60 years of adjusted water data available that further complicates the re-examination of comparisons done in Mr. Winterfeld's original simulation. In a few years we will have 70 years. For the purposes of my discussion, I usually compare the rolling 40-year rolling method to using the full 60 years worth of data. Unfortunately, some 20 years ago, these data were not available to Mr. Winterfeld. He also did not have available the computing power nor the software that we have today. In 1984, the newest Macintosh had an 8-Megahertz clock speed. I ran these resampling experiments on a desktop computer with a 1.7-Gigahertz clock speed. Over 200 times faster.

dependencies in streamflows. If there are no dependencies, any 40-year record cold have been examined, including a non-continuous 40-year record. This should have been done in order to avoid, as I discuss later, the real possibility that his results were merely a coincidence resulting from the particular pattern of streamflows that has historically occurred already. His Monte Carlo method does not correct for this possibility because, as far as I could tell, it sampled only from the immediately prior 40-year period so in a sense it preserved the uniqueness of the immediately prior 40 years.

In addition, when Mr. Winterfeld created the scenario to test the one-year rolling average, he built in a dependency that constrains or limits the kind of errors that can be made in the one year rolling data. He did this by "backcasting"or creating an after the fact scenario. The estimate for the year forecasted is always what happened in the immediately prior year. This creates a dependency that does not exist in actual streamflow data. If we use the immediately prior year to forecast the next year, in the real world that next year could be any year at random because actual streamflow data have no dependencies, unlike Mr. Winterfeld simulation.

Mr. Winterfeld's method produces an error that is always a type of inverse of the original estimate. Note in Exhibit ___(TDD-9) the places where the state changes from 0 to 1 or 1 to 0. If the data goes from 1 to 0, you can predict the error by simply putting a 1 in the immediately preceding line of the error column. These occurrences are marked with a symbol the looks like a check mark with an arrow on one end of it. If it goes from 0 to a 1, a minus 1 will appear. If there is no change, 1 to 1 or 0 to 0, a 0 will be in the immediately preceding line of the error column.

Q. What about how Mr. Winterfeld sets up the continuous records method.

A. The way this is calculated appears to produced a bias against the continuous records method. I believe I can explain why this occurs by discussing how the variance of a series of 1s and 0s like Mr. Winterfeld used (a discrete random variable) is influenced by the length of the period of is 1s and 0s and by discussing some "memory" characteristics of statistical models.

We need to begin by reviewing how the continuous records approach and the rolling average method created their respective estimates. If we refer again to Exhibit ___(TDD-9), the very last estimate in the continuous records column (0.564), we see that this estimate is made up of an average of the prior 39-conditions in the "streamflow" condition column. Now in the same continuous records column, go to line 7 to the estimate of 0.667. This estimate is an average of the six prior "streamflows" in the streamflow condition column.

When we examine the estimates in the 1-year rolling average column, we note that they are not an average of prior conditions at all, but as I just discussed, simply the immediately prior condition. In other words, the estimate is just what happened last time. Examples of this are indicated on Exhibit ___(TDD-9) by a long line with an offset at the end.

Q. You referred to reasons related to a discussion of statistical modeling. How are those related?

A. A model that changes as a result of the feedback it gets from its environment can be said to have a "memory". Though they do so in different ways, both the continuous records model and the 1-year rolling average model change as a result of feedback from their environments. However, the continuous records model takes into account the entire past history, whereas the 1-year rolling average model considers only the just prior single year. The continuous records model can be said to have an infinite memory and the 1-year rolling average model only a single-year memory.

Exhibit ___(TDD-10) through Exhibit ___(TDD-23) are a related series of exhibits that show step-by-step what happens to the estimates as years are added to the 1 and 0 data series. It is important to understand this "estimate--error--new estimate" cycle so that we can fully realize the consequences of the simulation Mr. Winterfeld set up.

Exhibit ___(TDD-10) shows what happens when a long series of sixteen 1s abruptly changes to a long series of eighteen 0s. The rolling average model, shown with filled squares, had been going along simply estimating that the next year will be what happened last year, that is, a one (1). And because early on the 1s continue, it did not make any errors until things changed abruptly. This is shown on Exhibit ___(TDD-11).

What are the consequences of this error? The rolling average model then immediately changes its estimate to a 0, or what happen the last time. The single error causes the model to changes its estimate from a 1 to a 0 and it suffers an error of only 1 unit and continues on estimating 0s.

Exhibit ___(TDD-10) also shows how the continuous records model operates under the identical pattern of 1s and 0s. First, note that the continuous records model (open circles) has been estimating a value between 0.5 and 1.0 even though only 1s have been occurring. It does so because it has a memory longer that 1 year and because Mr. Winterfeld started the example in his Exhibit T-76 at year seven. Why, I'm not exactly sure. In any event, because the continuous records models has an infinite memory, it uses the average of the first six years, rather than just year six alone, to estimate year seven. So the estimate for year seven is the average of what has happened so far: 1+0+1+0+1+1 divided by 6 equals 0.667. This is shown as the first open circle in Exhibit ___(TDD-10).

The next estimate is the average of the first seven years (six plus the 1 that just occurred) or 1+0+1+0+1+1+1 divided by 7 or 0.714. Estimates proceed in this fashion, each taking into account all the information to date. In a sense this is a changing rolling average of "N", where N is the number of years that have already occurred.

When the series abruptly changes from a long run of 1s, the continuous records model does not drastically change its estimate. Instead, a 0 just gets added to the calculation of the average. The pattern of estimates as the 0s continue is shown in Exhibit ___(TDD-10).

The error pattern suffered by the continuous records model is shown in Exhibit ___(TDD-12). It is considerably more complicated than the error pattern produced by the 1-yr rolling average model. Exhibit ___(TDD-13) shows the two error patterns on the same chart.

Without going into further detail, one can begin to get a sense of what happens when these two models are compared. When the world flip-flops, goes from all 1s to all 0s, the 1-year rolling average model adjusts after making just one error. However, the continuous records model is penalized because it is forced to have an infinite memory. It continues to factor in all the past 1s even though the world has changed to all 0s and as a result it accumulates a large error total. This is shown graphically in Exhibit ___(TDD-14), which compares the cumulative absolute error made by the two models.

Different patterns of 1s and 0s produce different estimating and error patterns. Exhibit ___(TDD-15) shows the <u>estimating</u> pattern for the two models when there are eight 1s in a row, followed by nine 0s in a row, followed again by eight 1s in a row, followed once more by nine 0s in a row. Exhibit ___(TDD-16) shows the <u>error</u> pattern for the two models and Exhibit ___(TDD-17) shows the <u>cumulative</u> absolute error. Once again, the 1-year rolling average model accumulates a lower error total than the continuous records model.

Things get more complicated as the 1 and 0 pattern changes and the value of a memory in a more complicated world begins to show. Exhibit ___(TDD-18) shows the resulting estimating pattern when two 0s in a row alternate with two 1s in a row, i.e., 00110011, and so on. Exhibit ___(TDD-19) shows the error pattern. Exhibit ___(TDD-

20) shows the cumulative error for the two methods. The cumulative error is the <u>same</u> in this case.

Following along the same logic, Exhibit ___(TDD-21) shows the estimating pattern for the two models when 1s and 0s alternate, i.e., 0101010101. Exhibit ___(TDD-22) shows the resulting error pattern. Exhibit ___(TDD-23) compares the cumulative error for the two models. The 1-year rolling average model now has a much <u>greater</u> error total than the continuous records model. In a sense, we might say memory is most valuable when the world is constantly changing.

Depending on the pattern of 1s and 0s, sometimes the continuous records approach leads to a lower cumulative error and sometimes it is the 1-year rolling average. I will discuss this further a little later in my testimony.

Q. How do all these factors you have discussed so far relate to your statement that there is an artifact in Mr. Winterfeld method that drives his conclusions?

A. Putting together all that I've discussed so far, it is possible to demonstrate that his results appear to simply follow from the mathematics of the way the comparisons are set up. Rather than running a Monte Carlo simulation as Mr. Winterfeld did, I calculate the expected result of the comparison between the two models.

To calculate the variance of a binomial distribution such as a series of 1s and 0s, we can take the expected probability of 1 (usually called "p"), times the expected probability of 0 (usually called "q"), time the series length (usually designated as "N"). So the formula for the variance is: $p \cdot q \cdot N$.

The two methods, continuous records versus 1-year rolling average, are structured such that for the continuous records approach, on average, the error (variance) is calculated each time a new year is added. So in year one, the variance is 0.25 (.5x.5x1); in year five it is

1.25 (.5x.5x5); in year ten it is 2.5 (.5x.5x10); in year twenty-five it is 6.25 (.5x.5x25), and so on.[6]

If we do the same for the 1-year rolling average approach, the flaw in Mr. Winterfeld method begins to shows up. In year one the variance is 0.25 (.5x.5x1). The same as in the continuous records approach. However, in year two, the variance is again 0.25 (.5x.5x1) because the rolling one-year approach does not take into account anything more than a single immediately prior year. For each added year the variance is the same: 0.25. To get the "cumulative error" we would again add up all these individual, identical variances each based on an N of only one. So for the 1-year rolling average model we have constant variance (error) and for the continuous records model we have a variance that increases as N increases.

The end result is that the rolling 1-year average produces a lower cumulative error simply because of the mechanics of the way the comparison was set up. **Necessarily, the sum of p·q·N (the continuous records method) will always be greater than sum of p·q·1(the 1-year rolling method), where N is greater than 1**.

This result follows from probability theory. So I conclude that Mr. Winterfeld's results appear to be based on the mechanics of the way he set up his comparisons. They say little about the merits of one method compared to the other with regard to streamflows.

---

[6] Mr. Winterfeld transformed all the "errors" at this stage into absolute values and summed them up to get what he called his "cumulative error" metric. He did not calculate the variance or standard deviation at this point. Mr. Winterfeld created absolute deviations and then took the standard deviation of those values, the standard deviation of average deviations. The interpretation of the average deviation is much different than the standard deviation. The standard deviation from the mean has very convenient mathematical properties; one of the most important is that it is one of the two parameters of a normal distribution. This cannot be said of the average deviation. Hence, it is more difficult to interpret. I discuss these issues with regard to resampling and note that traditional inferential statistics are not always appropriate for resampling experiments.

Q. Does this flaw happen for cases where the expected probability of 1s and 0s is different than 0.5.

A. Yes. For any p and q, this will always happen.

Q. Why does the continuous records approach ever turn out better than the 1-year rolling average method?

A. Because at any point in the series and for any particular unique series, the p and q values don't remain constant. Remember that the continuous records method has a memory. Wherever it is, it looks back to the start of the series and calculates an average to predict the next outcome. So, if six 1s in a row have just occurred in the middle of a series, it will base its estimate for the next outcome on the average it saw since the beginning of the series rather than the long term expected value of p or q. When the errors are summed up, they are summed over different p and q combinations. However, for the 1-year rolling average method, the p and q values are always very close to the expected long run values because it has a very short memory—one year in fact. And because it always and only predicts the same as just last year, it stays very close to the developing probabilities of p and q whereas the continuous records methods can get further away because it is averaging more and more history.[7]

At 30 years for example, the continuous records model has built up a reserve of knowledge of the past. If all prior 30 years were 1s, it would take 30 years of 0s to equal the number of 1s and the prediction for the next occurrence would still not be a 0. But rather, it would be 0.5, the long-term or expected value. The continuous records approach

---

[7] It could be argued that what Mr. Winterfeld has done is to set up a process where we are allowed to witness the unfolding of an estimate of the variance for a binomial distribution (continuous records) vs. just the estimate of the probability of a 1 (the 1-year rolling average model) as N increases. In this case it is not clear what comparing the cumulative errors between these two estimating processes really means.

minimizes its maximum error by hedging its bets, so to speak. The 1-year rolling average will predict 0s after just one mistake. The continuous records method never clears its memory. It assumes all the past is equally valid so it does not discard it. On the other hand, the 1-year rolling average method remembers only yesterday. All else is irrelevant. Remind anyone of some teenagers?

Of course, as the rolling average length increases to 5, 10, or a 40-year rolling average, memory gets longer. The length of the memory is equivalent to a judgement regarding the relevance of historical data. A 40-year rolling average says that no data over 40 years old is worthwhile or relevant. If new data is added, older data is discarded as a result of the, "over 40 years is irrelevant rule", set up by this approach. In contrast, the continuous records approach keeps all data and gives it equal relevance.

All these factors combine to produce somewhat unpredictable results as several of my prior exhibits show. An example of how these factors combine can be demonstrated by turning to Exhibit ___(TDD-24).

This exhibit shows that depending upon the unique series of 1s and 0s that have occurred, and depending upon the series length we focus on, either method may produce the lower cumulative error. The series of 1s and 0s that are shown in my Exhibit ___(TDD-24) are the <u>same</u> as the series Mr. Winterfeld chose for his Exhibit T-76. At the 35th year, the continuous records approach is lower; at the 30th year the rolling average approach is lower; at the 25th year the methods are essentially identical with regard to cumulative error.

In the same vein, as Exhibit ___(TDD-23) shows what happens to cumulative error when two 0s alternate with two 1s: 00110011001100110011. The result is that in this instance cumulative error is essentially identical for the two methods.

Q. Mr. Winterfeld looks at more than a 1-year rolling average. Doesn't that correct the deficiencies you discussed?

A. No, it just incrementally waters them down until at exactly 40 years the two methods are essentially the same, each with a 40-year memory. Then as the continuous records methods adds years beyond 40, the differences again reappear.

Q. In light of all this, are you implying that a continuous record approach is better in all situations? Should we always use all the data that is available?

A. What method is better depends on a detailed analysis of the data series in question.[8] If there are trends or cycles in the data, it may indeed be better to use a rolling average approach, the exact length being uncertain. Of course, even a simple rolling average approach gives equal weight to all the data within its length. For a 40-year rolling average, the second year in the 40-year period is given the same weight as the thirty-ninth year. There are methods that differentially weight the age of the data in the period so that older data may have less weight or so that each tenth year is given more weight if there is a 10-year cycle in the series, etc.

There is, however, no indication that such a model is appropriate for the streamflow data we are discussing here. In fact, it has been mentioned already that there is widespread agreement that there are no dependencies or cycles in the streamflow history. Absent a reason to say that older data is irrelevant or should be given less weight, there is no statistical reason to prefer a model, such as a 40-year rolling average, that uses less than all the applicable data. In fact, as I have already testified, I believe that to discard any data reduces the accuracy of our estimates and is contrary to theory in mathematically statistics,

---

[8] The one-year rolling average method may be effective in a 1 and 0 world where simply the raw number of exact, correct "hits" is what is important, as opposed to minimizing its maximum error. For example, betting maroon (a combination of black and red that may well perfectly represent the relative frequency of red and black) on a roulette wheel will always produce more losses than betting red all the time, or black for that matter. However, if I am interested in forecasting blended or continuous values, and want to be closest most often, a 1 and 0 model probably will not be my choice. In fact it may always be wrong.

especially the Glivenko-Cantelli Theorem. And of course, the 2001 record low streamflow demonstrates quit convincingly that the low streamflows that occurred some 60 years ago during the 1930s are not irrelevant, contrary to the way the 40-year rolling average model treats the same data.

Q. Do you have an opinion on Mr. Winterfeld use of only a single forecasted year as the criterion for deciding on how many historical water years to use?

A. Mr. Winterfeld does base his conclusions on the premise that we are most interested in only yearly estimates of streamflows and power supply costs (pages 3 and 5 of his testimony). My experience has been that when we construct a pro forma test period for rate making, we attempt to make that pro forma reflect a period longer than one year. We average out injuries and damages for example. When we evaluate surplus sales, we don't just use the next year's prices but also examine a longer period.

If fact, the first full year that rates are in effect (the "rate year") is typically over a year past the pro forma period. This occurs simply because of the time it takes to litigate a case and because it is very difficult to file a case with the newest possible pro forma test year. It is also prudent to plan on not coming in to adjust rates each year so, as long a period as is practically possible, should be reflected by the pro forma.

As I discussed earlier, we ought to ask what period we are trying to forecast for the pro forma, since that may become the criterion we use to evaluate various approaches. I believe it is more reasonable to think about a five-year period, not only because it fits better with what has been done in the past but also because it is easier to forecast an average period than a single year, as Exhibit ___(TDD-4) Exhibit ___(TDD-5) and reflect. Mr. Winterfeld never evaluated anything but a single forecasted year in any of his comparisons.

Q. Did you do any analysis with regard to the type of bias a continuous 40-year method may introduce?

A. Yes I did. I compared the distribution of individual streamflow years for all 60 years of the record as well as for the last 40 years. I created histograms or frequency polygons, as they are sometimes called, for both the 60-year series and the 40-year series. I standardized these series to adjust for the fact that the 60-year series obviously has more data points than the 40-year series. The results are shown in Exhibit ___(TDD-25).

The darker line in the exhibit is the 60 year series, the lighter line the 40. The shaded areas show the significant bias in the 40-year series compared to the 60 years of streamflow record. The rolling 40-year average very clearly produces a distribution that significantly under represents low water years and over represents the high water years.

Avista witnesses have discussed this bias in prior cases and they discuss it again in this case. The continuous last 40-year of record not only produces a change in the estimate of the mean of streamflows but it also produces a different shaped distribution of yearly flows, one that is skewed away from values below the mean value.

Q. Did you do any other analysis on the differences between the 40-year rolling average method and the 60 year water record?

A: Yes. My results are shown in Exhibit ___(TDD-26) and Exhibit ___(TDD-27). These two exhibits are a recast of my Exhibit ___(TDD-2) and Exhibit ___(TDD-3) with data added to enable a comparison of how the series of unique water years represented by the existing continuous 40-year rolling average compares with many other outcomes that may well have occurred. In fact, resampling theory implies that these possible outcomes are a better estimate of the true state of nature than the single unique 40-year series recommended by Mr. Winterfeld.

Exhibit ___(TDD-27) shows that the 40-year continuous rolling average is just barely within the 90% confidence interval for a sample size of 40 years and is clearly outside the 90% confidence interval for sample sizes of 50 or 60 water years. What this indicates is that the 40-year rolling average is one of the very least representative series that should be used to reflect the true state of streamflows.

Q. You demonstrated that the continuous record approach sometimes produces a lower cumulative error than the rolling average approach. Did you conduct any resampling experiments comparing the two methods?

A. Yes. This resampling experiment is an element of my critique of Mr. Winterfeld's method. It involved using the resampling technology to directly compare Mr. Winterfeld's rolling average method to the continuous records approach for actual streamflow data. I believe the outcome of the resampling experiment better reflects the efficacy of the two models because it avoids the use of 1s and 0s and because it does not limit the data range to simply the immediately prior 40 year period, although that period was allowed as a possible selection.

Q. Would you outline the experiment you did?

A. I randomly selected, with replacement, 60 years of streamflows. I then compared the continuous records approach and the 1-year rolling average approach with regard to cumulative, absolute error, calculated the way Mr. Winterfeld did in his testimony.

I ran 1,000 iterations of the experiment with the streamflow data, taking a new random sample each time. I found that the continuous records approach produced a lower cumulative, absolute error 993 times out of the 1,000, or 99.3% of the time. In other words, in only 7 times out of 1,000 did Mr. Winterfeld's rolling average approach do better than the continuous records approach when actual streamflow data were used. The other 993 times the continuous records approach was better. I would be happy to run additional iterations during my cross-examination since the model can be demonstrated on a laptop computer with a projection screen.

So even though I do not believe that the cumulative error metric is very relevant to picking the best historical period, it is nonetheless consistently and overwhelmingly lower for 60 years of record compared to 40 years.[9]

SUMMARY OF WATER YEAR FINDINGS

Q. Would you please summarize your testimony with regard to the number of water years that should be used to best reflect power supply costs for rate making purposes?

A. All four approaches I review lead me to the same conclusion. The entire 60 years of streamflow record should be used for rate setting purposes. In addition, my analysis of forecasting errors leads me to believe that we can much more accurately forecast a 5-year average for streamflows than we can a single year. However, even though streamflow data may be best viewed as an average, the financial impacts may not occur "on average" but year-to-year. Mr. Norwood discusses the implications of this conclusion and how it relates to Avista's request for a power cost adjustment mechanism.

With regard to Mr. Winterfeld testimony, it is first of all dated. It is now nearly 20 years since his analysis. There is another 10 years of water record available and there may soon be another 10 available. A new record low streamflow has been established for 2001. Avista has updated its power supply model. The computer intensive techniques that are now routinely available were not available to Mr. Winterfeld. And, most importantly, there

---

[9] When we evaluate whether we should us 40, 50, or 60 years of water record, we don't have to build that record from year one and accumulate errors as we add years. The data are already available. N goes directly to 40, 50, or 60 without having to be forced through sizes of 2, 3, 4, or 25 as Mr. Winterfeld's comparison demands. The only error term that's relevant in such a case is the last one, not the cumulative error. As Mr. Winterfeld himself says on page 4 of his testimony, "If this is essentially correct, a continuous record provides an ever increasing sample size, which should provide a better estimate of average water conditions." In other words, the final N is the one that matters most.

are significant and fatal flaws in his methodology. Finally, duplicating his method and using actual streamflow data led me to a conclusion just the opposite of the one he drew.

Compared to the entire 60-year streamflow record, the continuous 40-year rolling average under represents poor water years and over represents good water years. This not only biases the estimate of the mean streamflows, but also distorts the distribution of streamflows that are used to determine power supply costs. In fact, the 40-year rolling average method produces one of the very least representative streamflow periods.

All applicable data should be used; 60 years of streamflow data is preferable to 40, rolling or otherwise. This conclusion is consistent with statistical theory. More data is better.

Q. Does this conclude your testimony?

A. Yes it does.

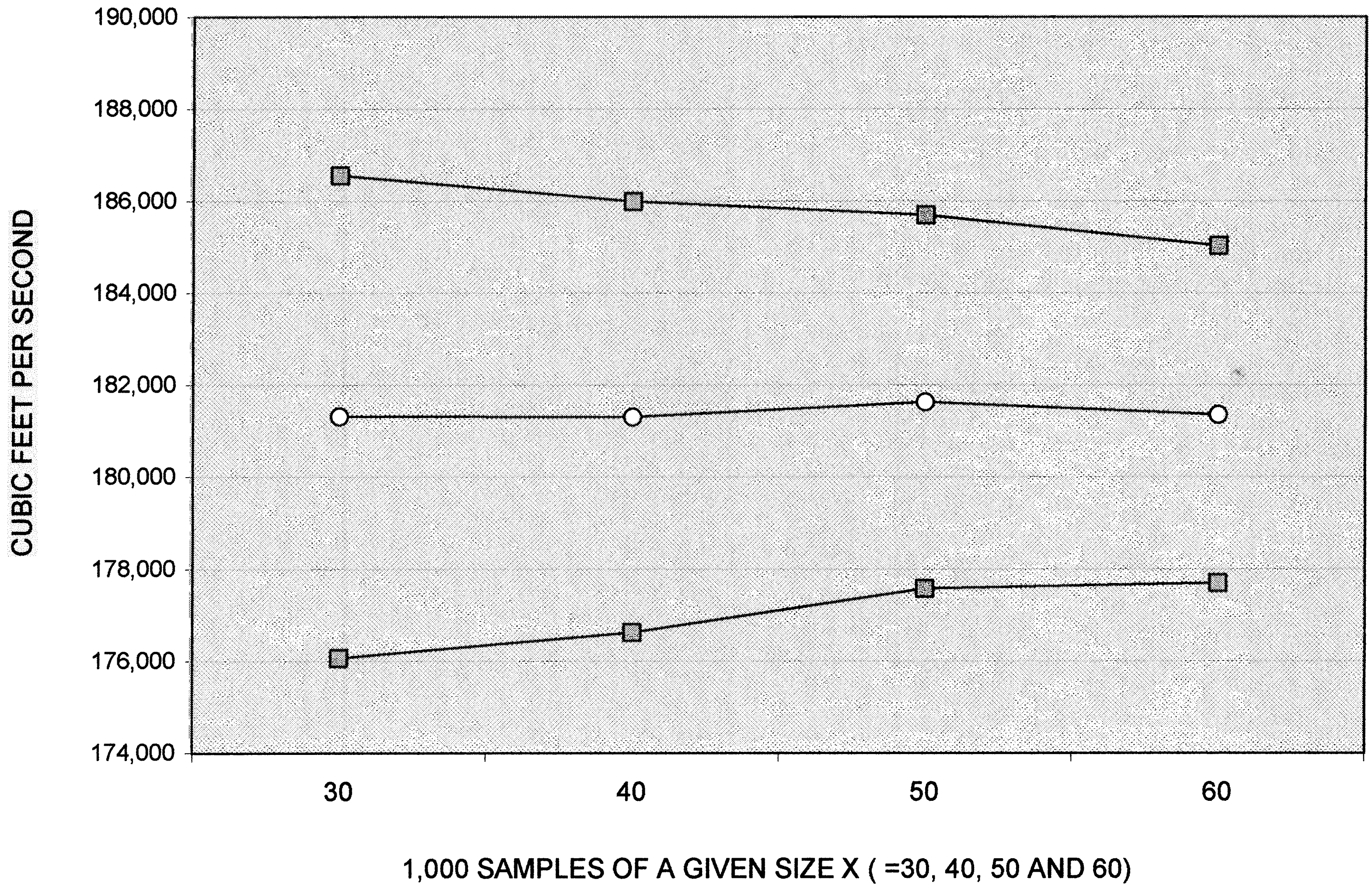
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-1)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

GRAND MEAN AND THE AVERAGE DEVIATION (PLUS OR MINUS) OF SAMPLE MEANS FROM THE GRAND MEAN FOR VARIOUS SAMPLE SIZES BASED ON 1,000 MEANS FOR EACH SAMPLE SIZE



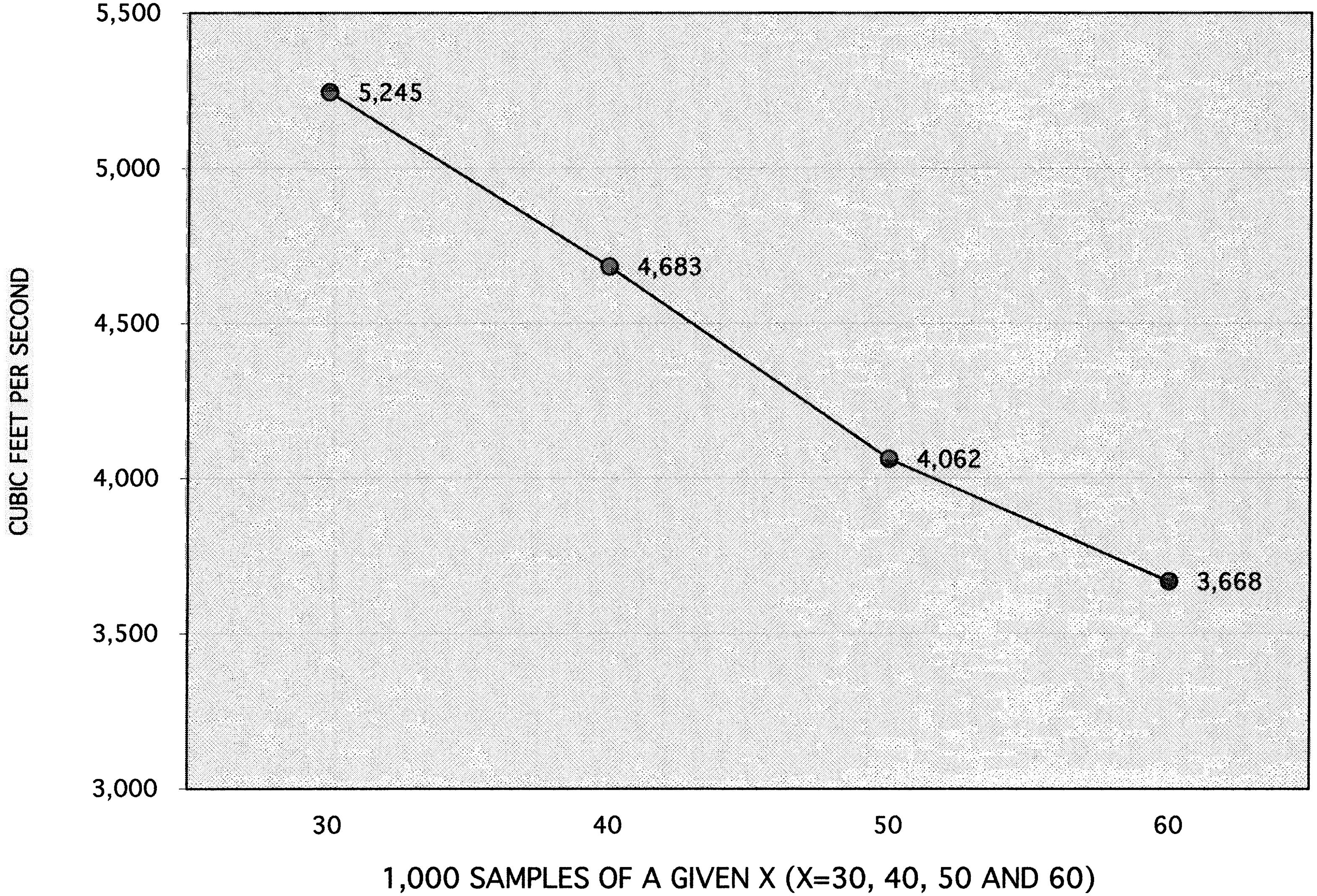
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-2)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

AVERAGE DEVIATION FOR THE MEAN FROM THE GRAND MEAN FOR VARIOUS SAMPLE SIZES BASED ON 1,000 MEANS FOR EACH SAMPLE



BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-3)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

RESAMPLING 90% CONFIDENCE LIMITS FOR THE MEAN
AROUND THE GRAND MEAN FOR VARIOUS SAMPLE SIZES
BASED ON 1,000 MEANS FOR EACH SAMPLE SIZE

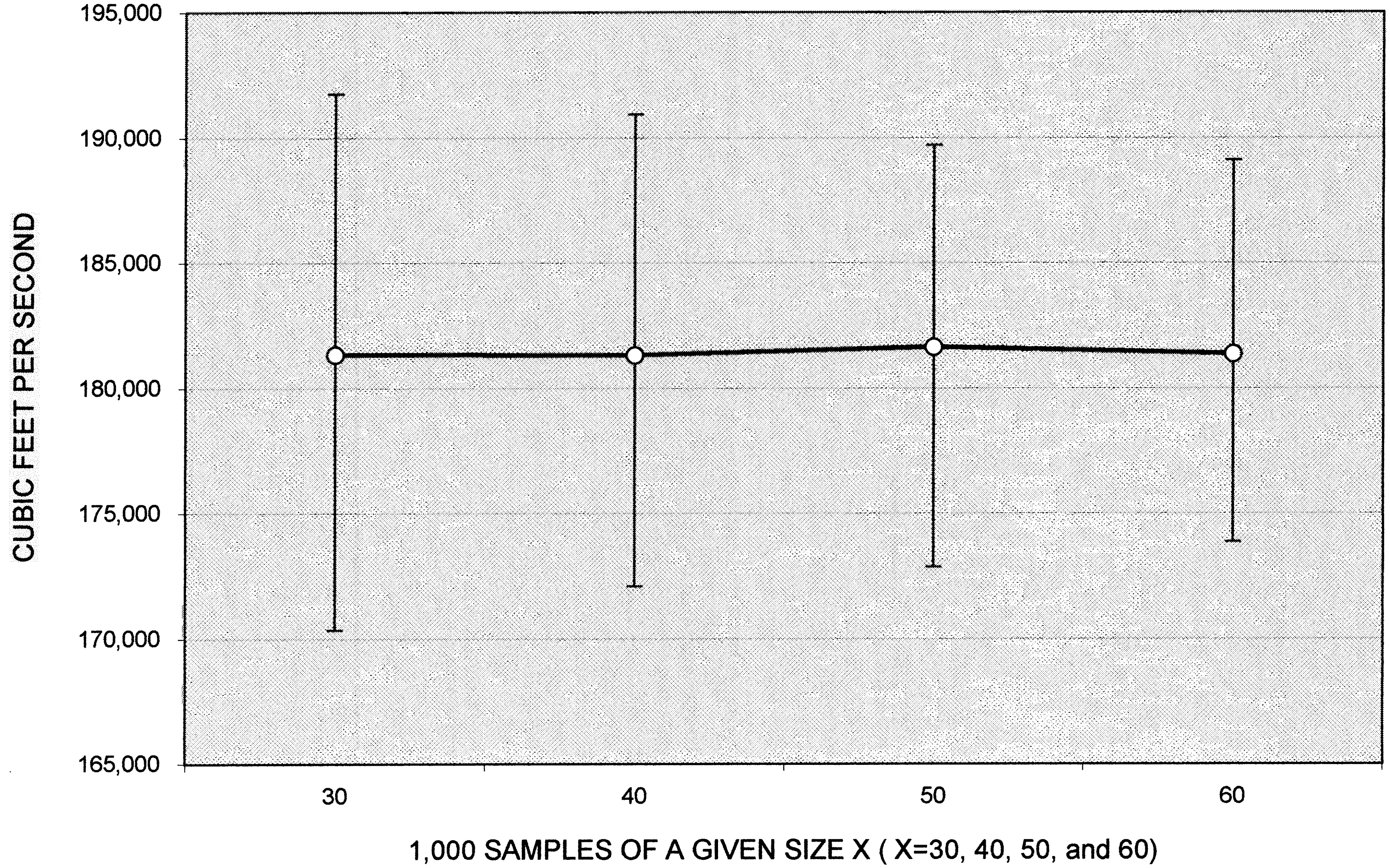


Exhibit No. ___ (TDD-3)
Docket No. UE-01 ___
Page 1 of 1

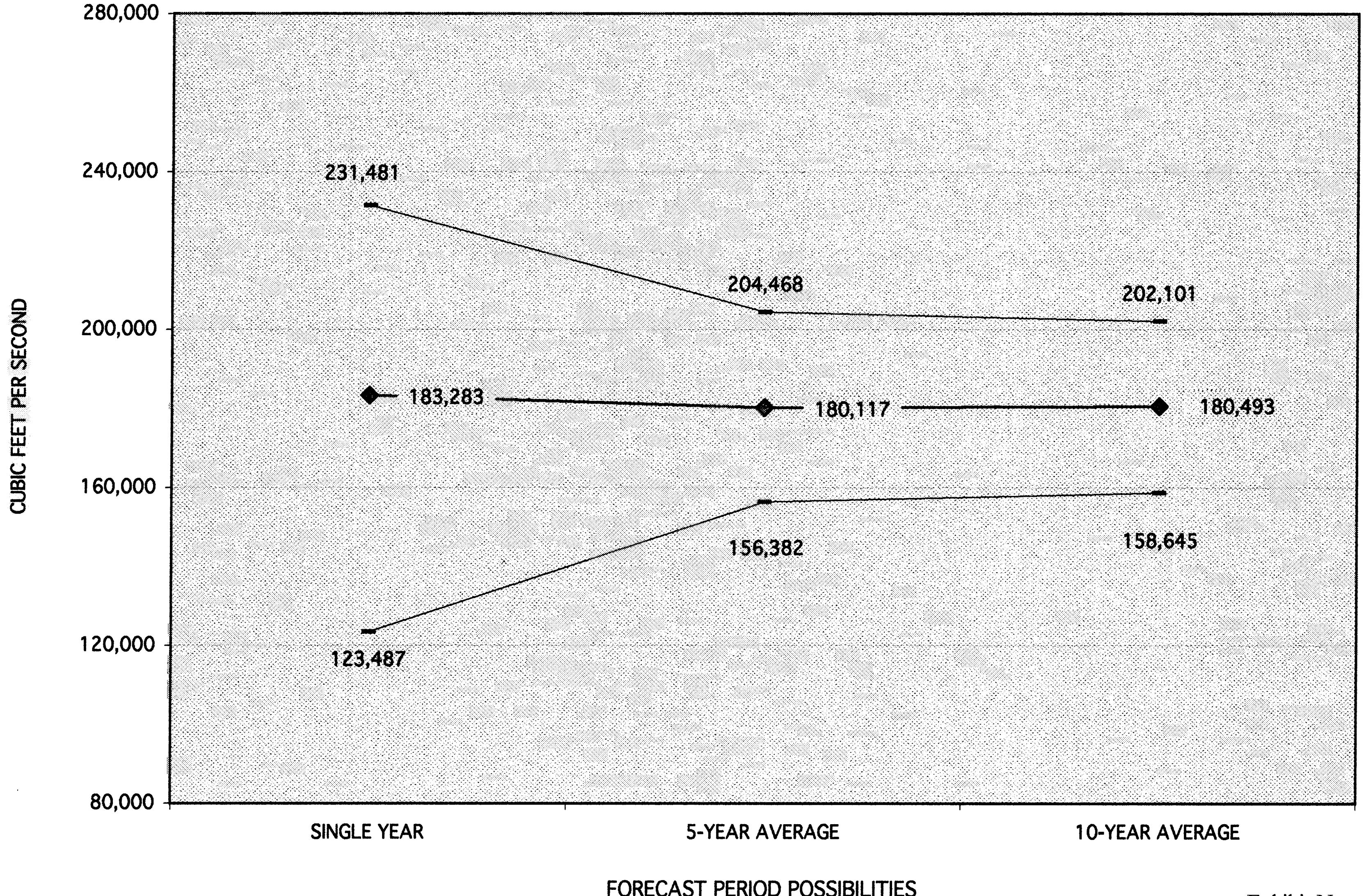
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-4)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

90% CONFIDENCE AROUND THE AVERAGE DEVIATION FOR 100 RANDOMLY SELECTED SINGLE YEARS AND FOR 100 RANDOMLY SELECTED 5-YEAR AVERAGES AND 10-YEAR AVERAGES



BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-5)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

DISTRIBUTION OF 100 RANDOMLY SELECTED SINGLE WATER YEARS, 100 RANDOMLY SELECTED 5-YEAR AVERAGES, AND 100 RANDOMLY SELECTED 10-YEAR AVERAGES

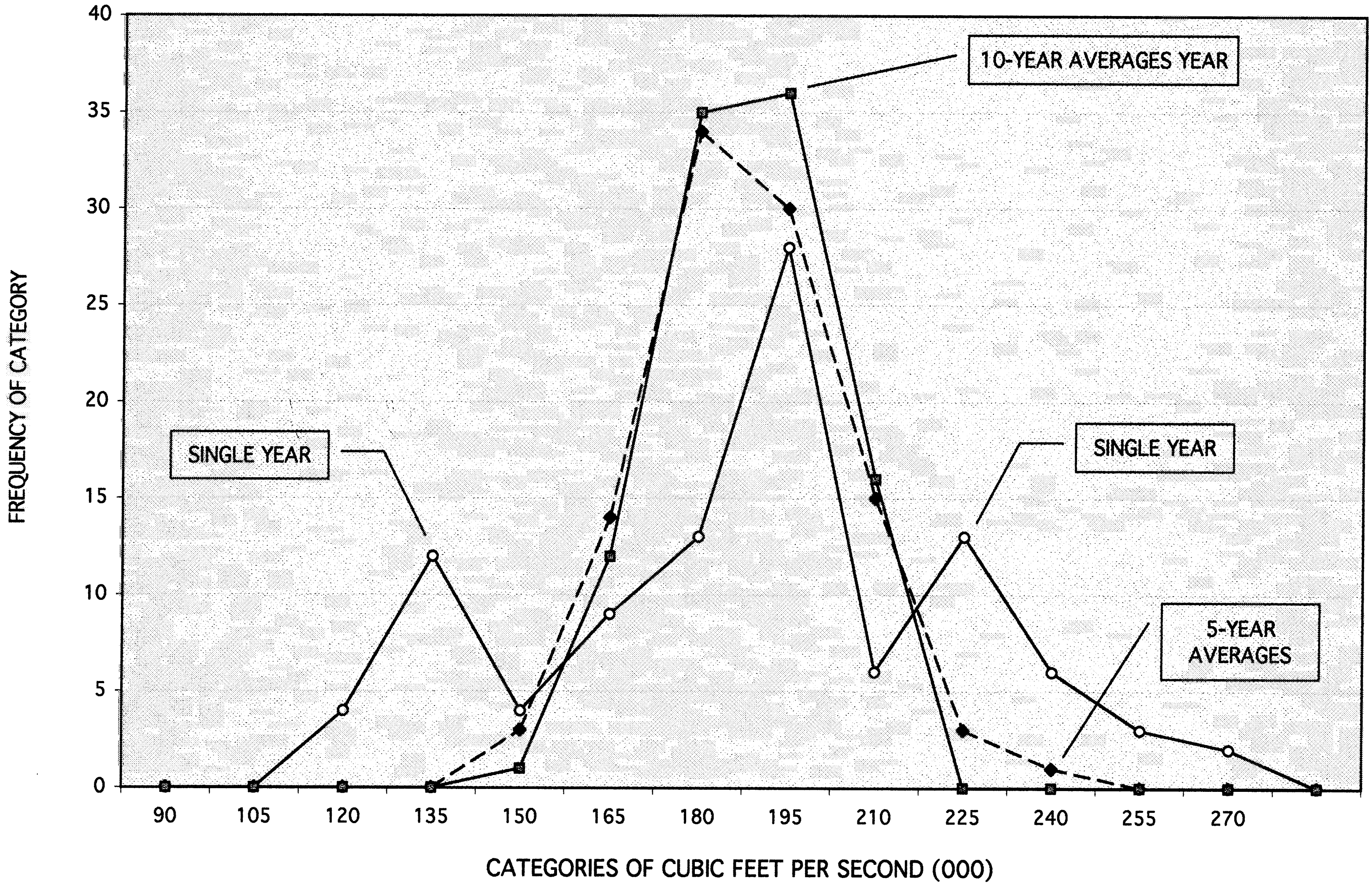


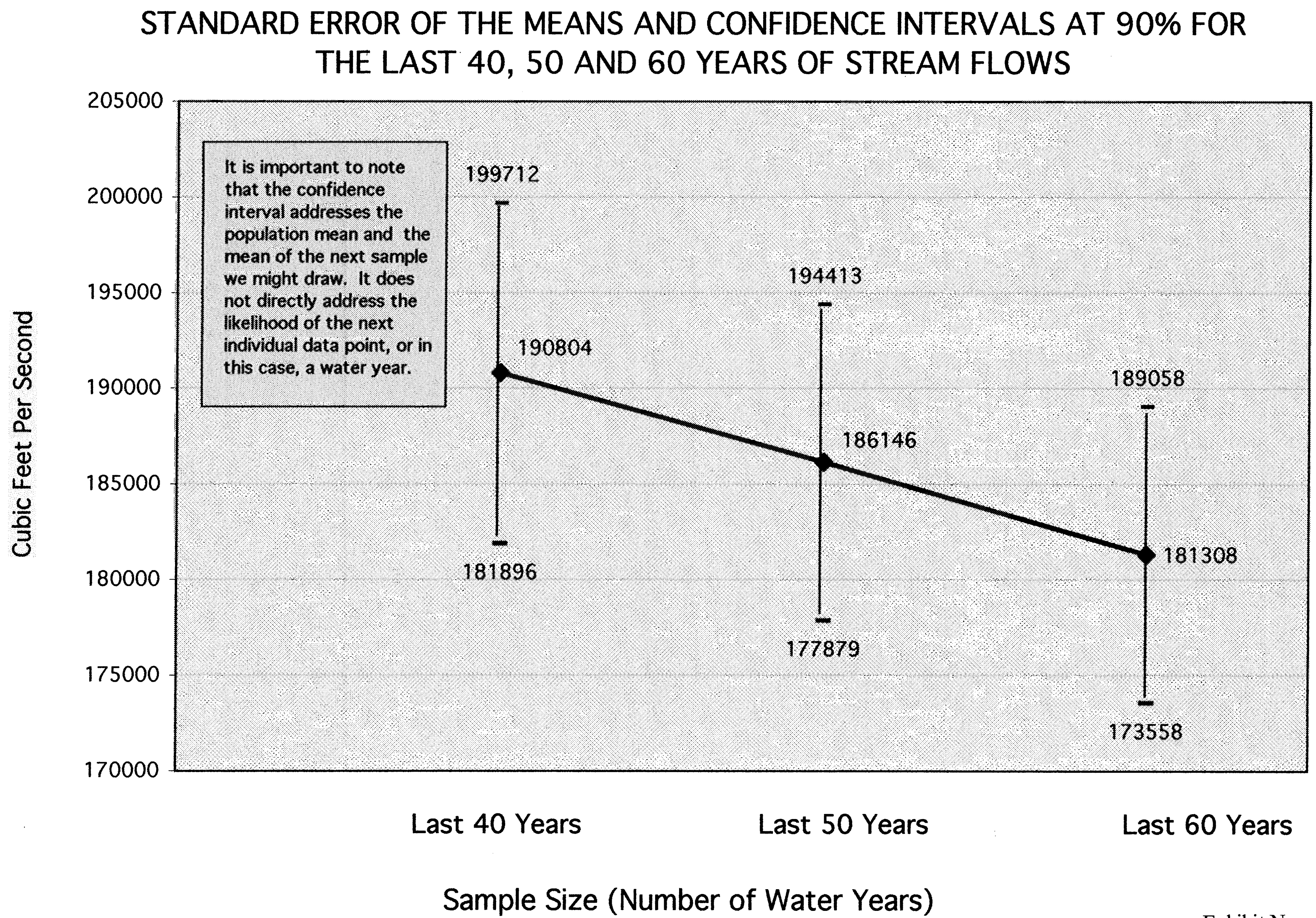
Exhibit No. ___ (TDD-5
Docket No. UE-01 ___
Page 1 of 1

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-6)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

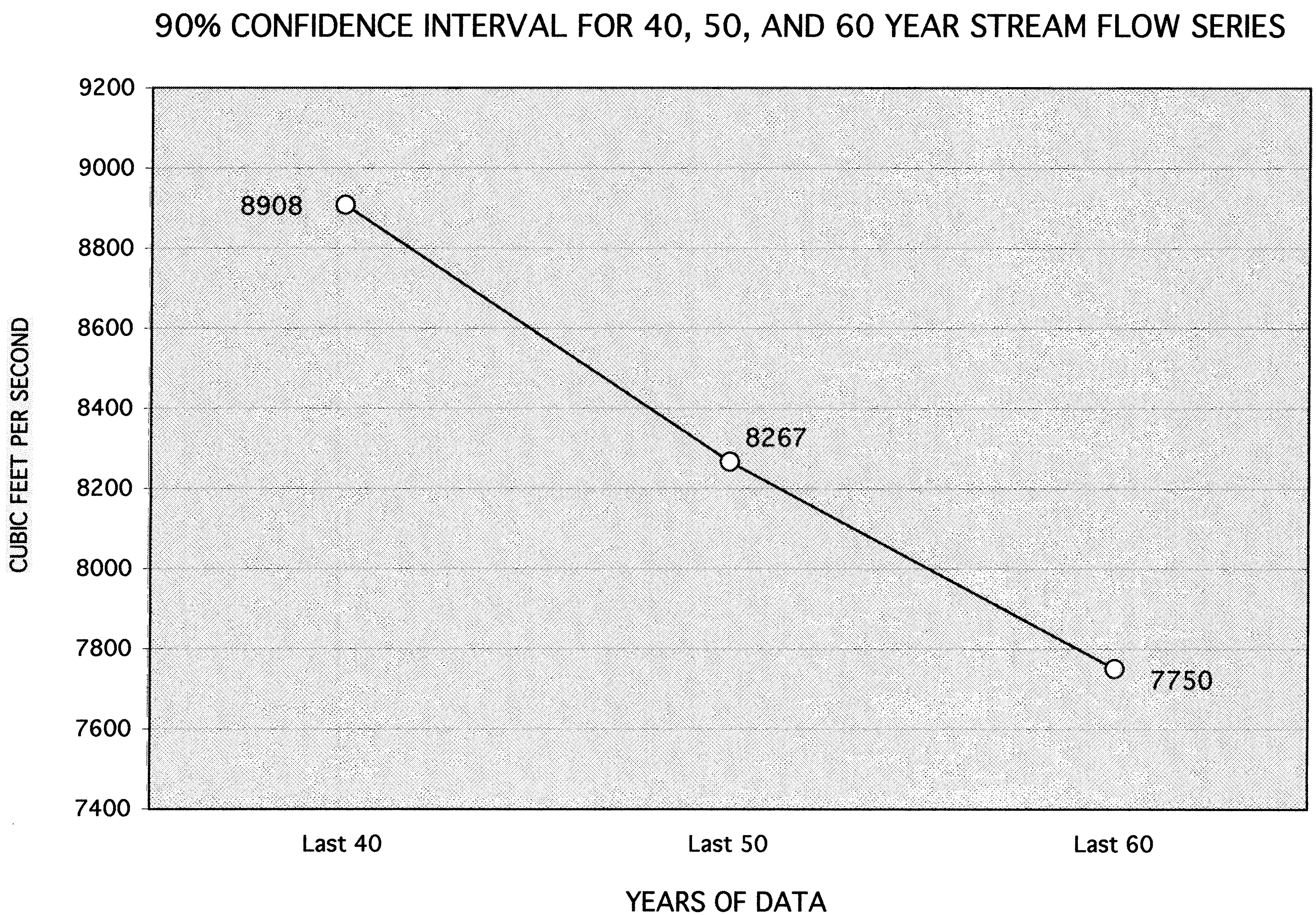

STANDARD ERROR OF THE MEANS AND CONFIDENCE INTERVALS AT 90% FOR THE LAST 40, 50 AND 60 YEARS OF STREAM FLOWS
Cubic Feet Per Second
205000
200000
195000
190000
185000
180000
175000
170000
It is important to note that the confidence interval addresses the population mean and the mean of the next sample we might draw. It does not directly address the likelihood of the next individual data point, or in this case, a water year.
199712
190804
181896
194413
186146
177879
189058
181308
173558
Last 40 Years
Last 50 Years
Last 60 Years
Sample Size (Number of Water Years)

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-7)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.


90% CONFIDENCE INTERVAL FOR 40, 50, AND 60 YEAR STREAM FLOW SERIES
CUBIC FEET PER SECOND
9200
9000
8800
8600
8400
8200
8000
7800
7600
7400
8908
8267
7750
Last 40
Last 50
Last 60
YEARS OF DATA

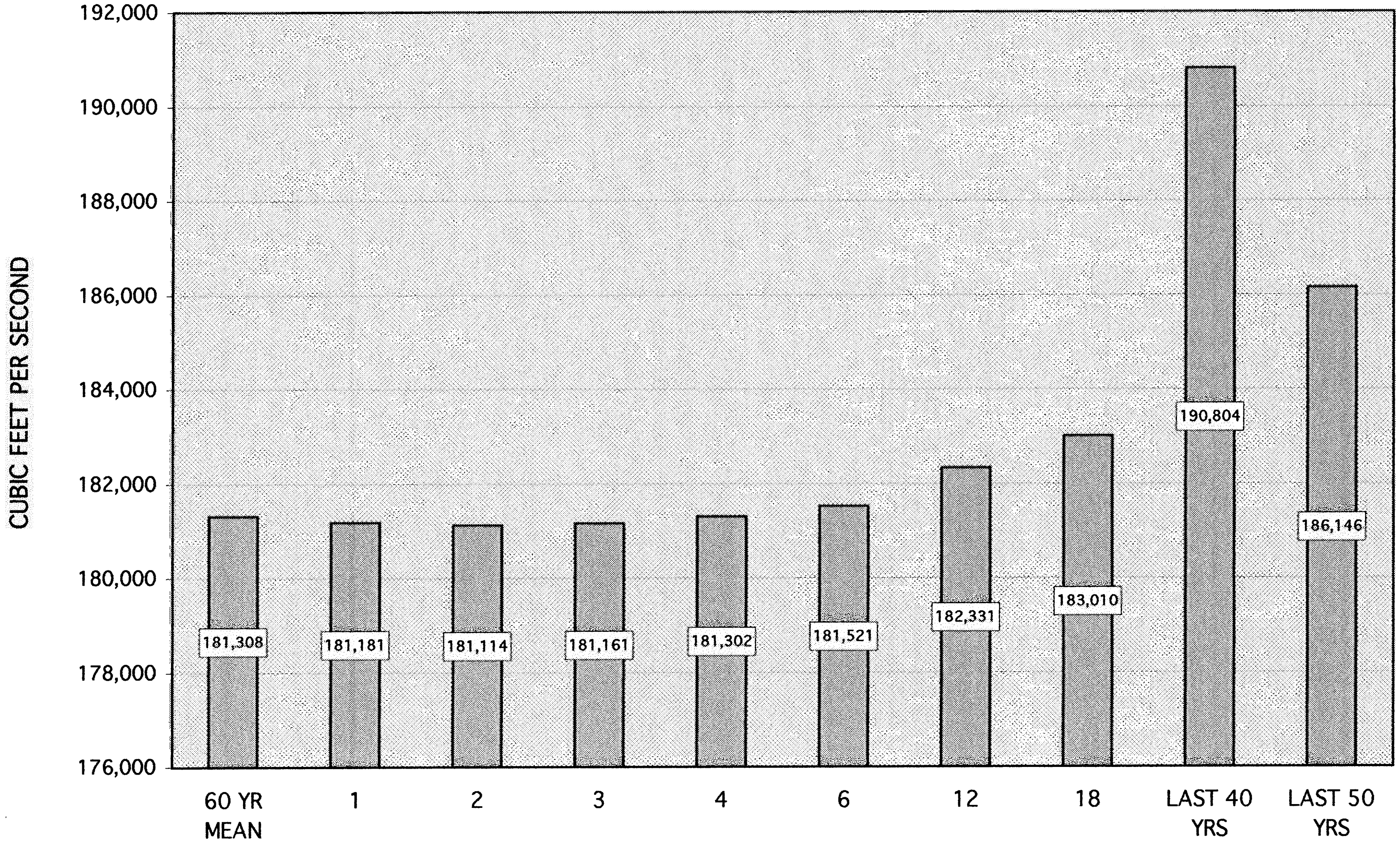
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-8)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

EFFECT ON THE 60 YEAR MEAN OF EXCLUDING HIGH AND LOW VALUES AND COMPARISON TO 40 AND 50 YEAR SERIES AVERAGES



NUMBER OF HIGH AND LOW VALUES EXCLUDED AND SERIES MEANS

Exhibit No. ___ (TDD-8)
Docket No. UE-01 ___

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-9)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

		CONTINUOUS RECORDS		1 YR ROLLING AVERAGE	
Year	Streamflow Condition	Estimate	Error	Estimate	Error
[a]	[b]	[c]	[d]	[e]	[f]
1	1				
2	0				
3	1				
4	0				
5	1				
6	1				
7	0	0.667	0.667	1.000	1.000
8	0	0.571	0.571	0.000	0.000
9	0	0.500	0.500	0.000	0.000
10	0	0.444	0.444	0.000	0.000
11	0	0.400	0.400	0.000	0.000
12	1	0.364	-0.636	0.000	-1.000
13	1	0.417	-0.583	1.000	0.000
14	1	0.462	-0.538	1.000	0.000
15	1	0.500	-0.500	1.000	0.000
16	0	0.533	0.533	1.000	1.000
17	0	0.500	0.500	0.000	0.000
18	1	0.471	-0.529	0.000	-1.000
19	0	0.500	0.500	1.000	1.000
20	1	0.474	-0.526	0.000	-1.000
21	0	0.500	0.500	1.000	1.000
22	1	0.476	-0.524	0.000	-1.000
23	1	0.500	-0.500	1.000	0.000
24	0	0.522	0.522	1.000	1.000
25	1	0.500	-0.500	0.000	-1.000
26	0	0.520	0.520	1.000	1.000
27	0	0.500	0.500	0.000	0.000
28	1	0.481	-0.519	0.000	-1.000
29	1	0.500	-0.500	1.000	0.000
30	1	0.517	-0.483	1.000	0.000
31	0	0.533	0.533	1.000	1.000
32	1	0.516	-0.484	0.000	-1.000
33	0	0.531	0.531	1.000	1.000
34	0	0.515	0.515	0.000	0.000
35	1	0.500	-0.500	0.000	-1.000
36	1	0.514	-0.486	1.000	0.000
37	1	0.528	-0.472	1.000	0.000
38	1	0.541	-0.459	1.000	0.000
39	1	0.553	-0.447	1.000	0.000
40	1	0.564	-0.436	1.000	0.000
Cumulative Error		=	-1.886		0.000

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-10)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ESTIMATES WITH 16 STRAIGHT ONES AND 18 STRAIGHT ZEROS
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY=1)

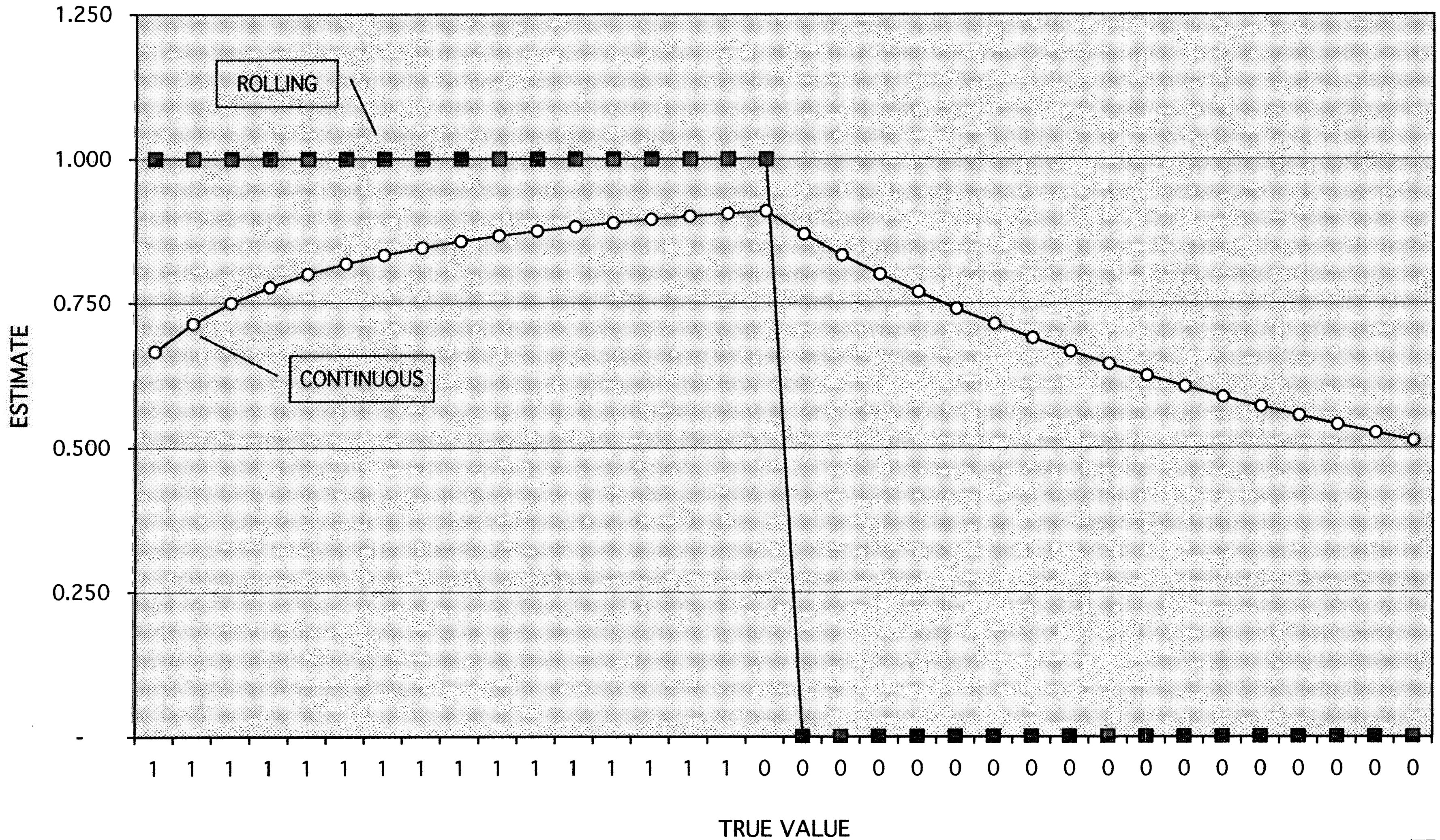


Exhibit No. ___ (TDD-10)
Docket No. UE-01 ___
Page 1 of 1

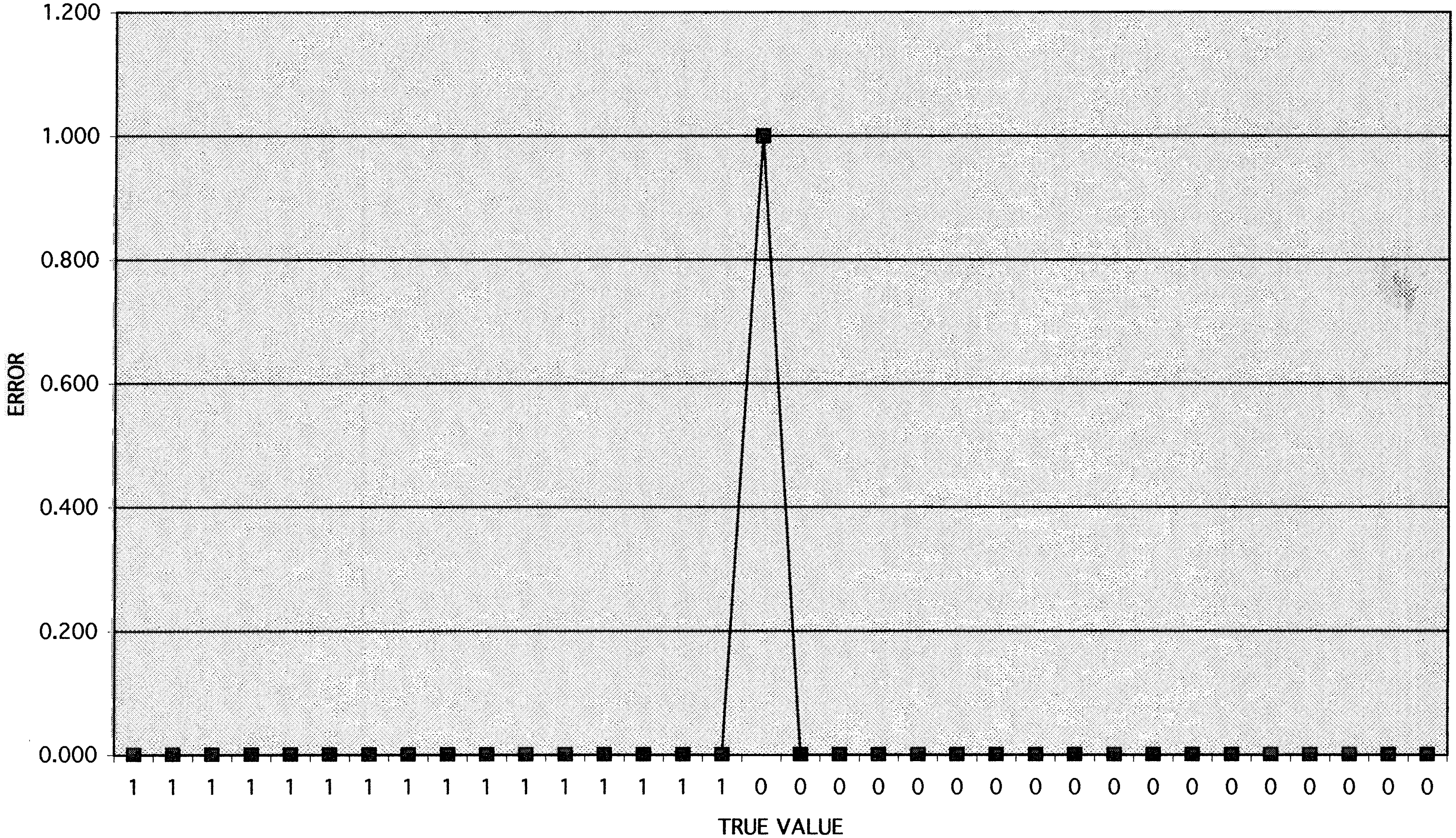
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-11)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ERRORS WITH 16 STRAIGHT ONES AND 18 STRAIGHT ZEROS
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
1-YR ROLLING AVERAGE (MEMORY =1)



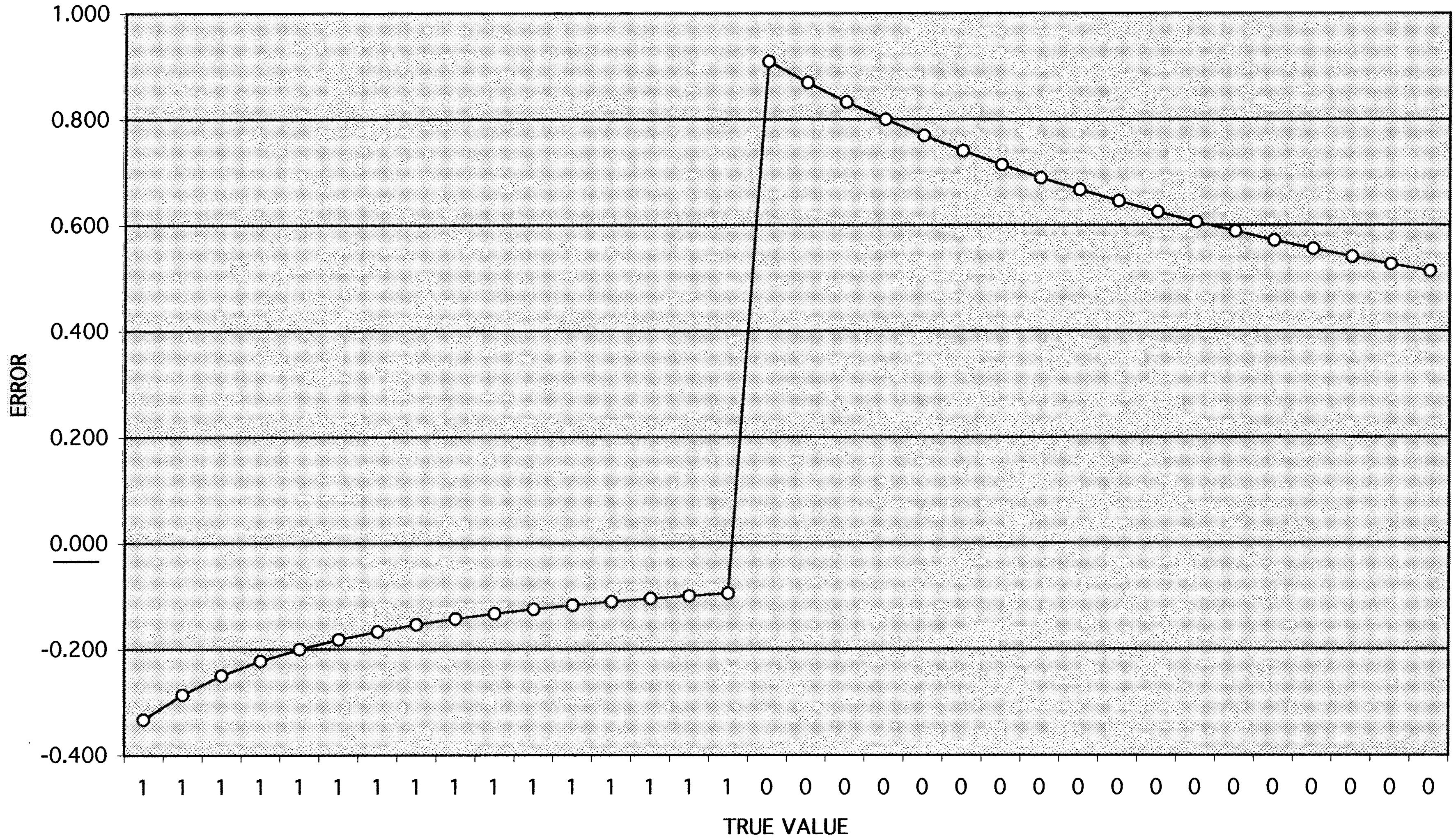
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-12)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ERRORS WITH 16 STRAIGHT ONES AND 18 STRAIGHT ZEROS EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF CONTINUOUS RECORDS (MEMORY=40)



BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-13)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ERRORS WITH 16 STRAIGHT ONES AND 18 STRAIGHT ZEROS
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY =1)

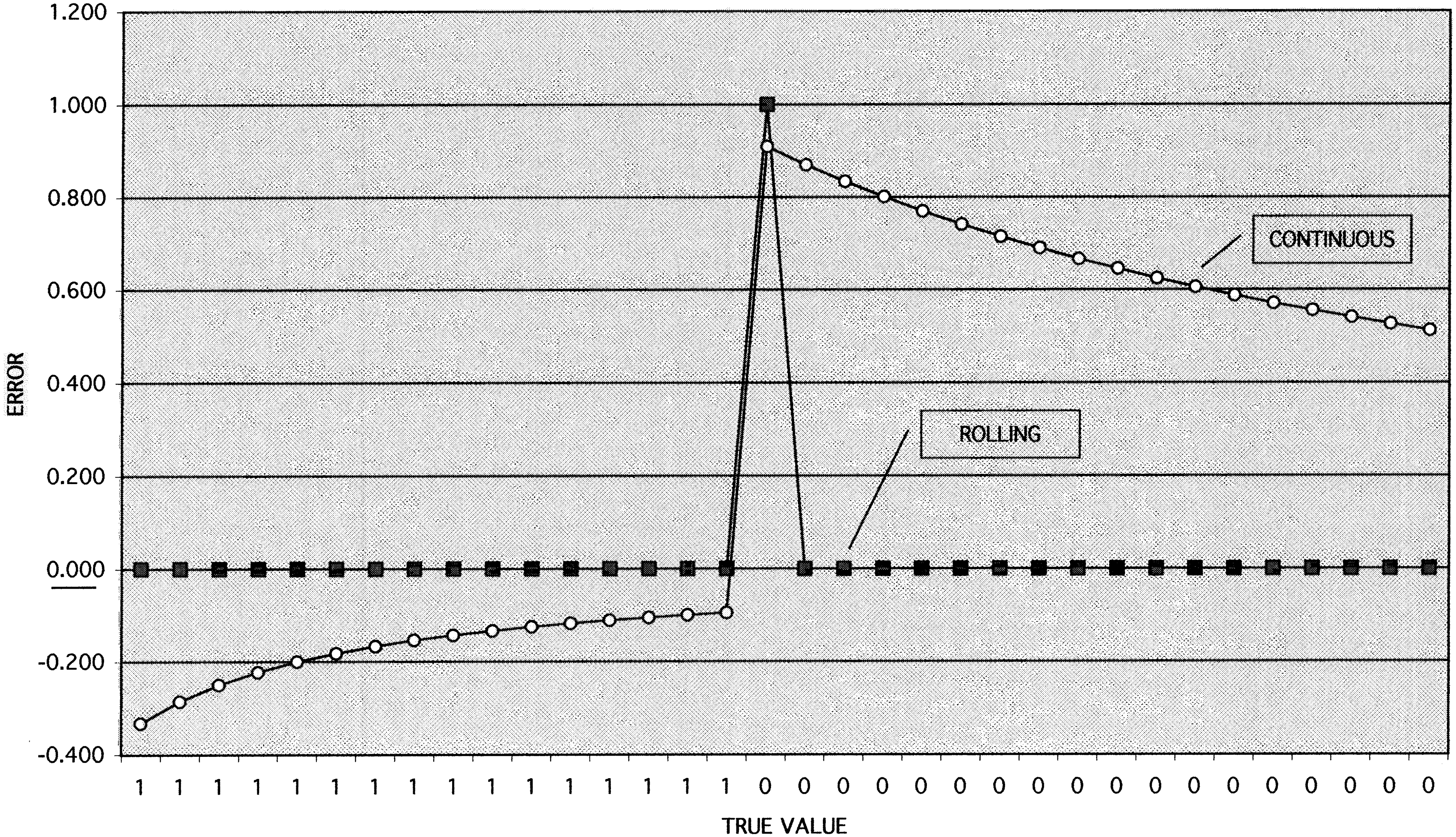


Exhibit No. ___ (TDD-13)
Docket No. UE-01 ___
Page 1 of 1

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-14)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

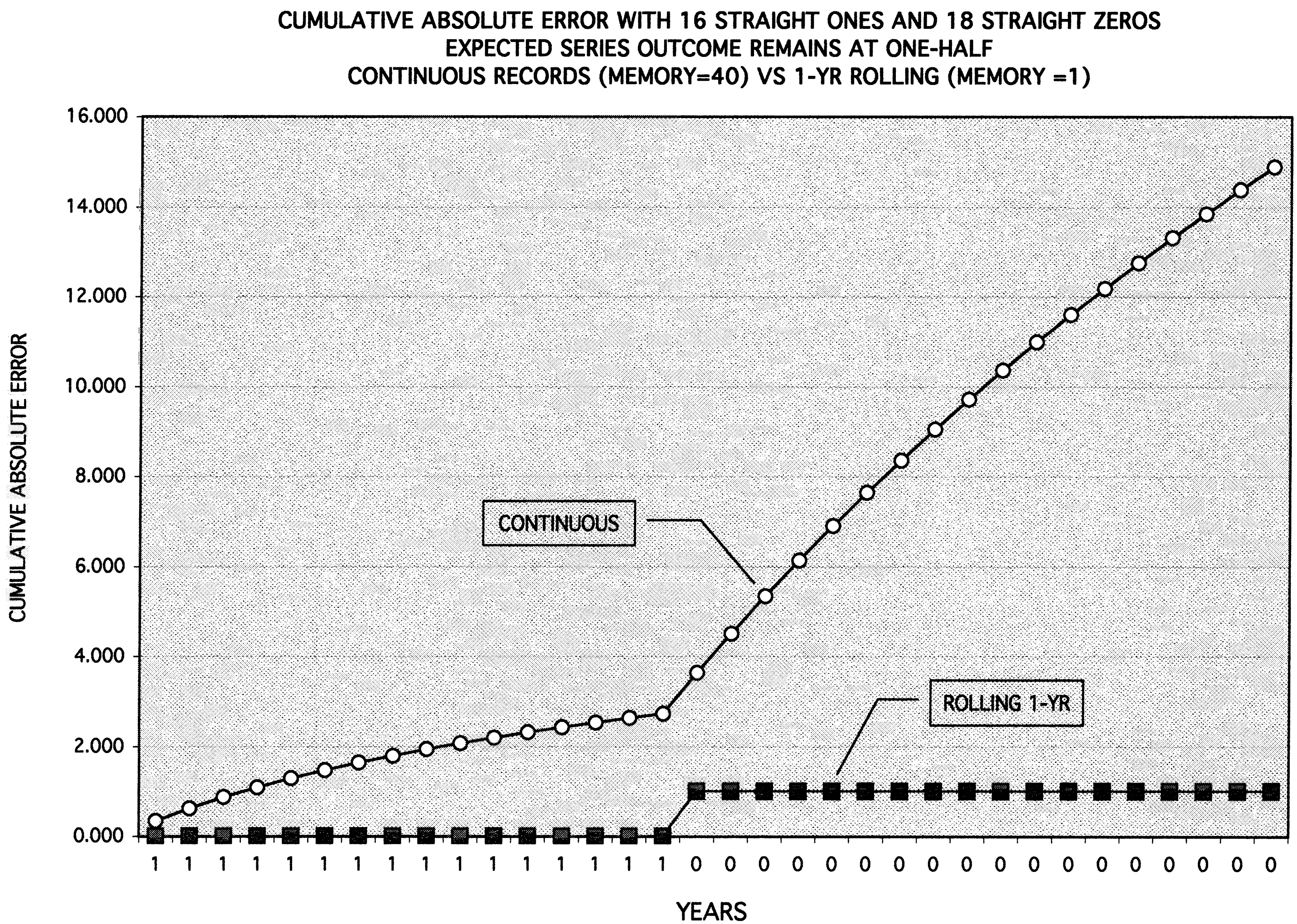


Exhibit No. ___ (TDD-14)
Docket No. UE-01 ___
Page 1 of 1

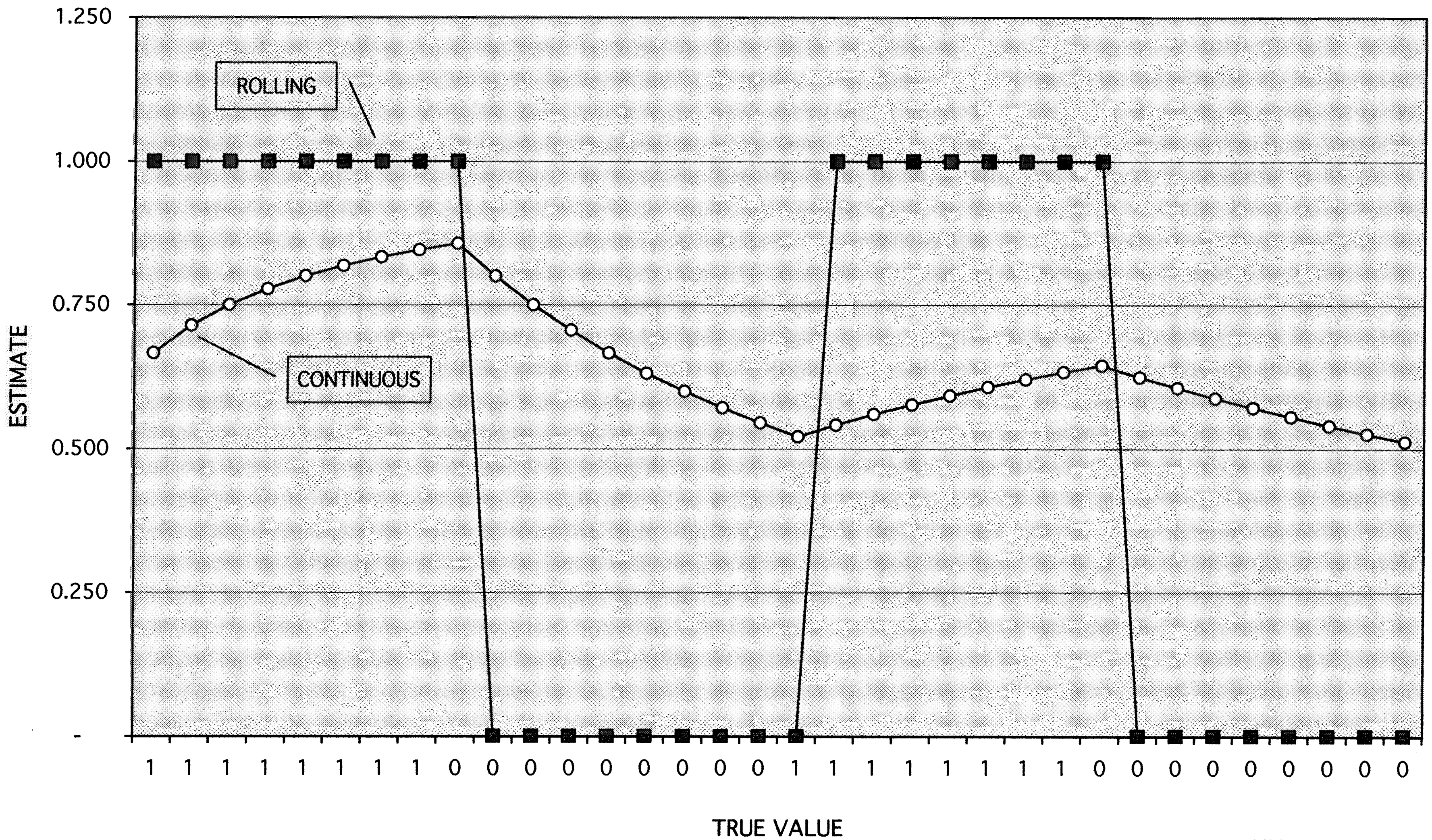
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-15)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ESTIMATES WITH 8 STRAIGHT ONES AND 9 STRAIGHT ZEROS REPEATED
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY=1)



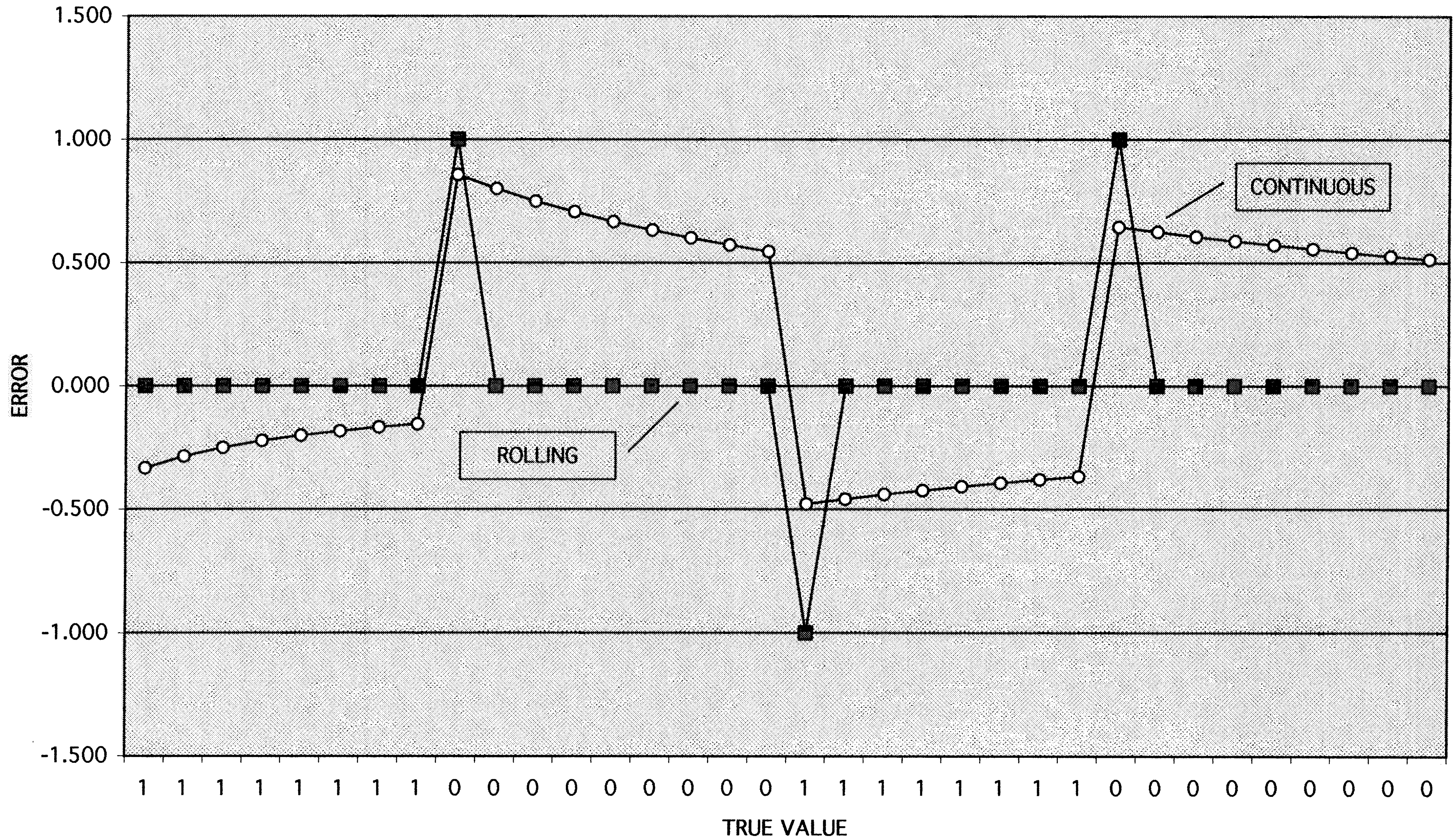
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-16)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ERRORS WITH 8 STRAIGHT ONES AND 9 STRAIGHT ZEROS REPEATED
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY =1)



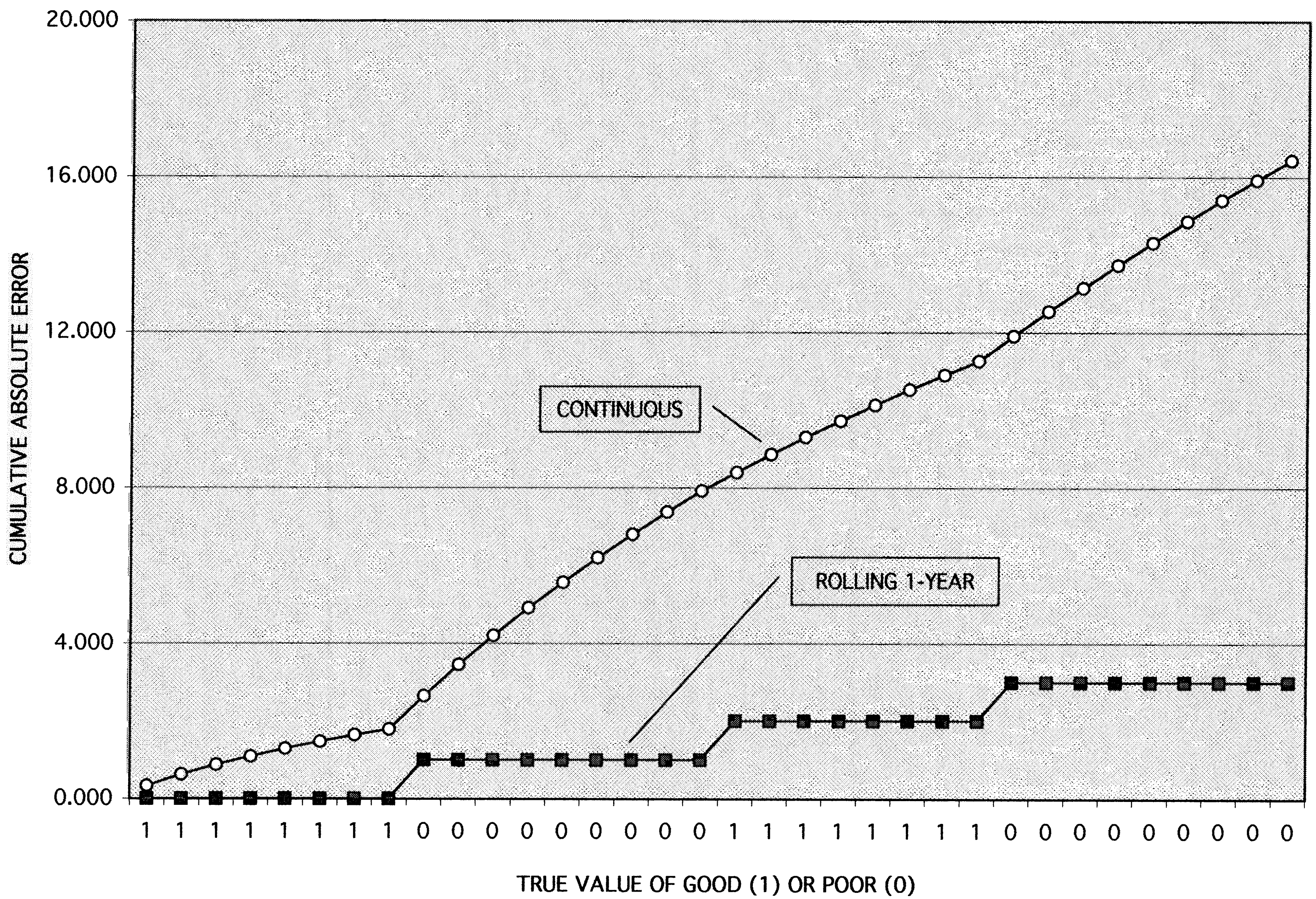
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-17)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

CUMULATIVE ERROR WITH 8 STRAIGHT ONES AND 9 STRAIGHT ZEROS REPEATED
CONTINUOUS RECORDS (MEMORY=40) VS I-YR ROLLING AVERAGE (MEMORY=1)



BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-18)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ESTIMATES WITH 2 ONES ALTERNATING WITH 2 ZEROS REPEATED
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY=1)

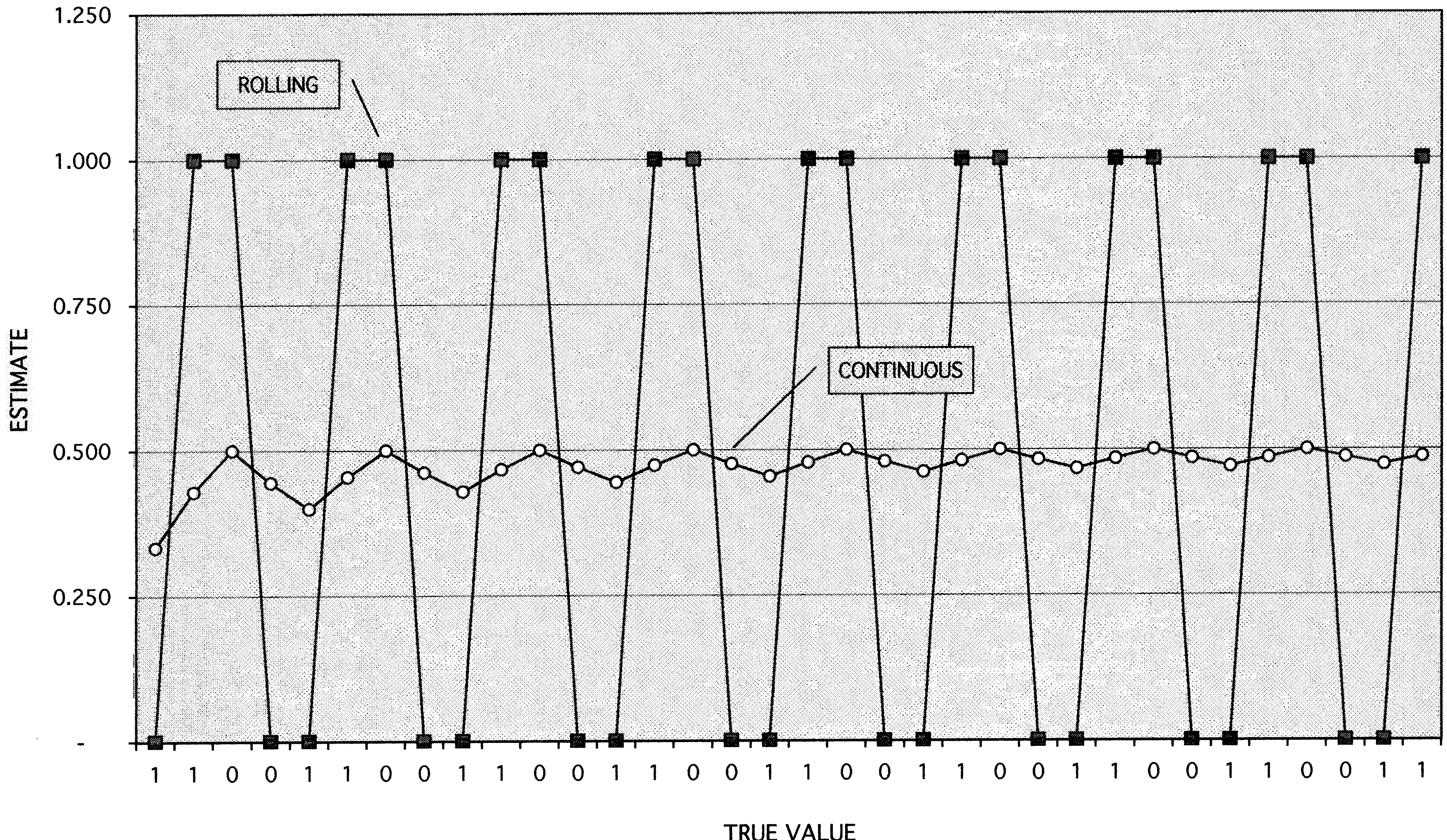


Exhibit No. ___ (TDD-18)
Docket No. UE-01 ___
Page 1 of 1

BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-19)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ERRORS WITH 2 ONES ALTERNATING WITH 2 ZEROS REPEATED
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY =1)

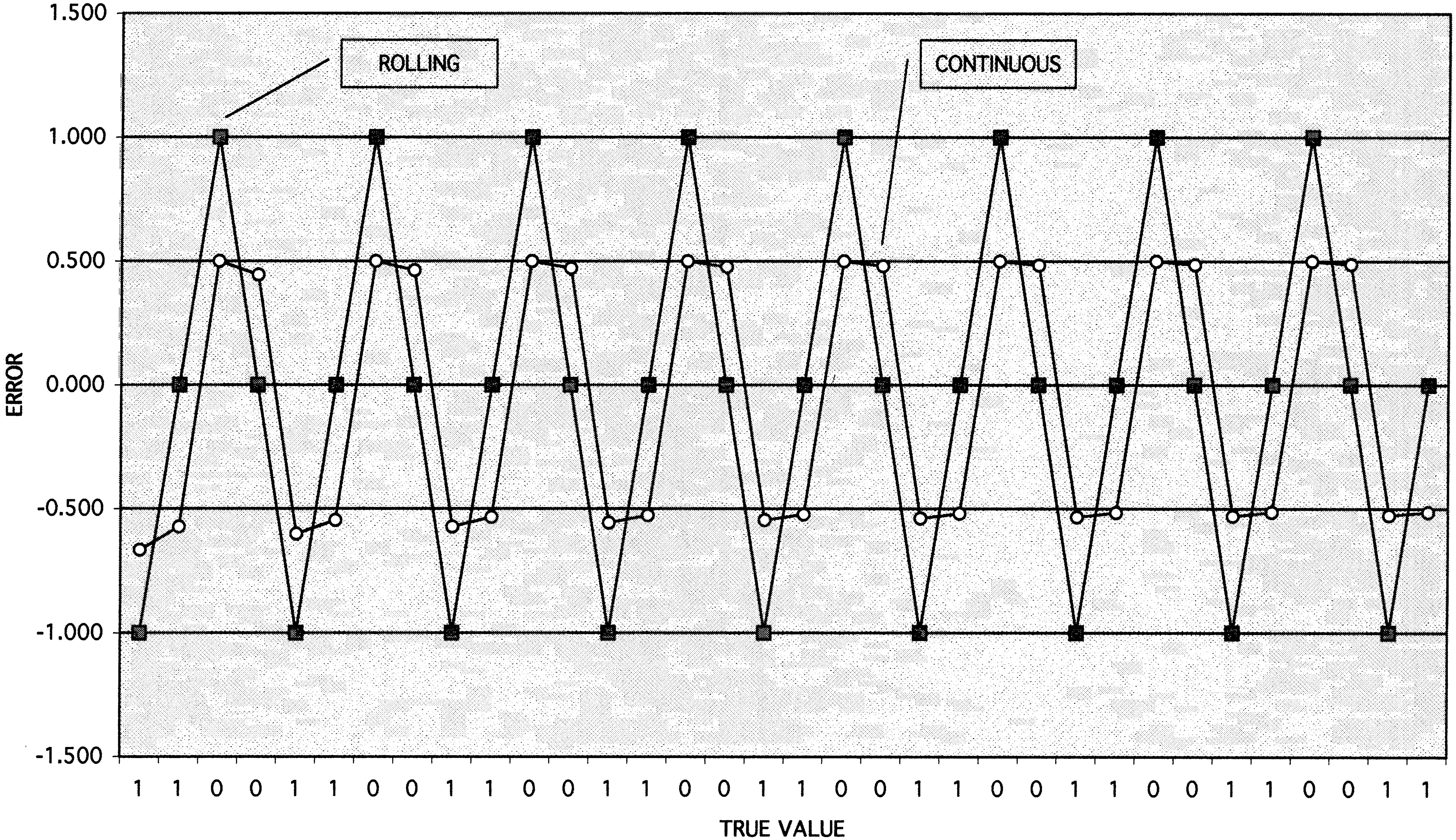


Exhibit No. ___ (TDD-19)
Docket No. UE-01 ___
Page 1 of 1

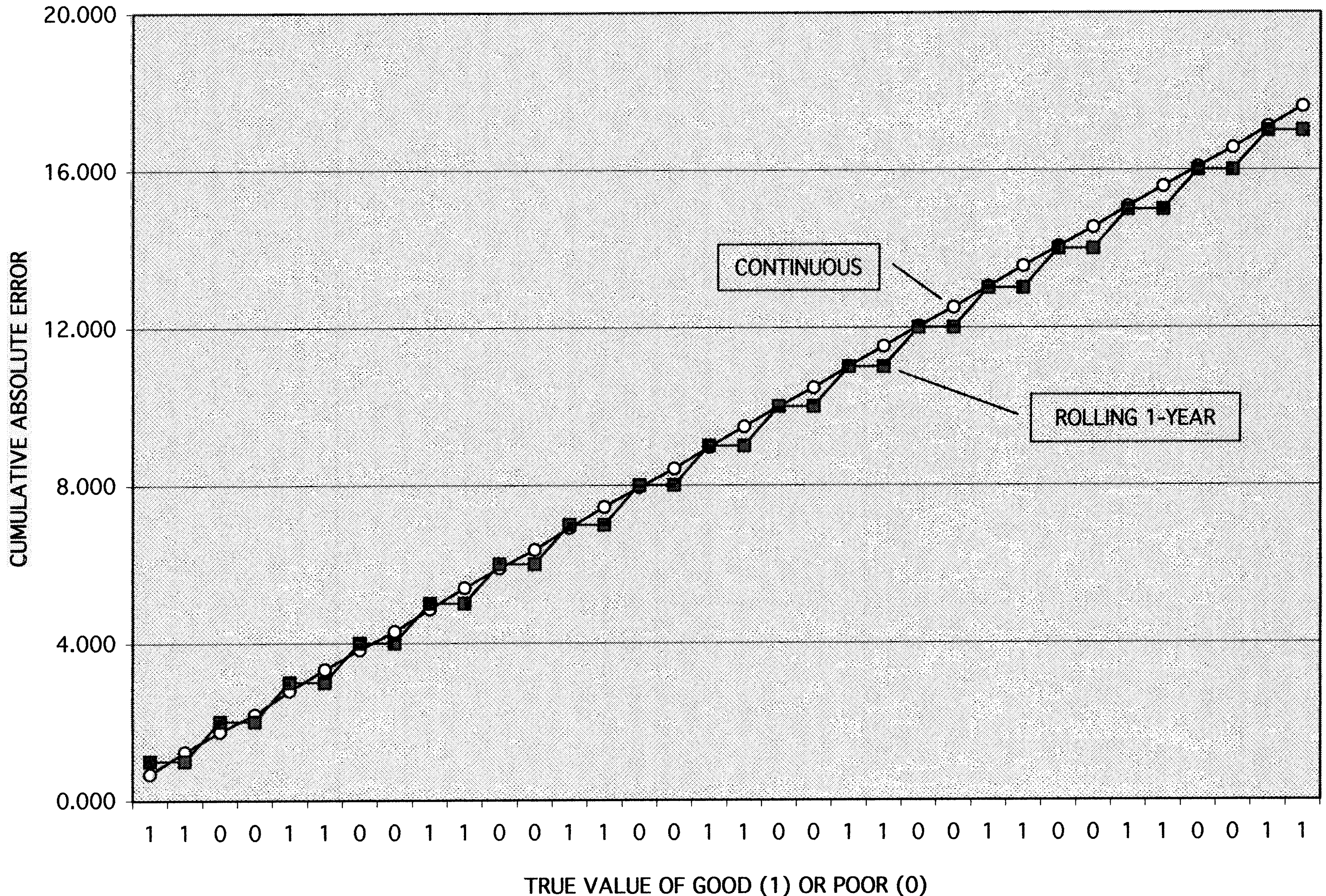
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-20)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

CUMULATIVE ERRORS WITH 2 ONES ALTERNATING WITH 2 ZEROS REPEATED CONTINUOUS RECORDS (MEMORY=40) VS I-YR ROLLING AVERAGE (MEMORY=1)



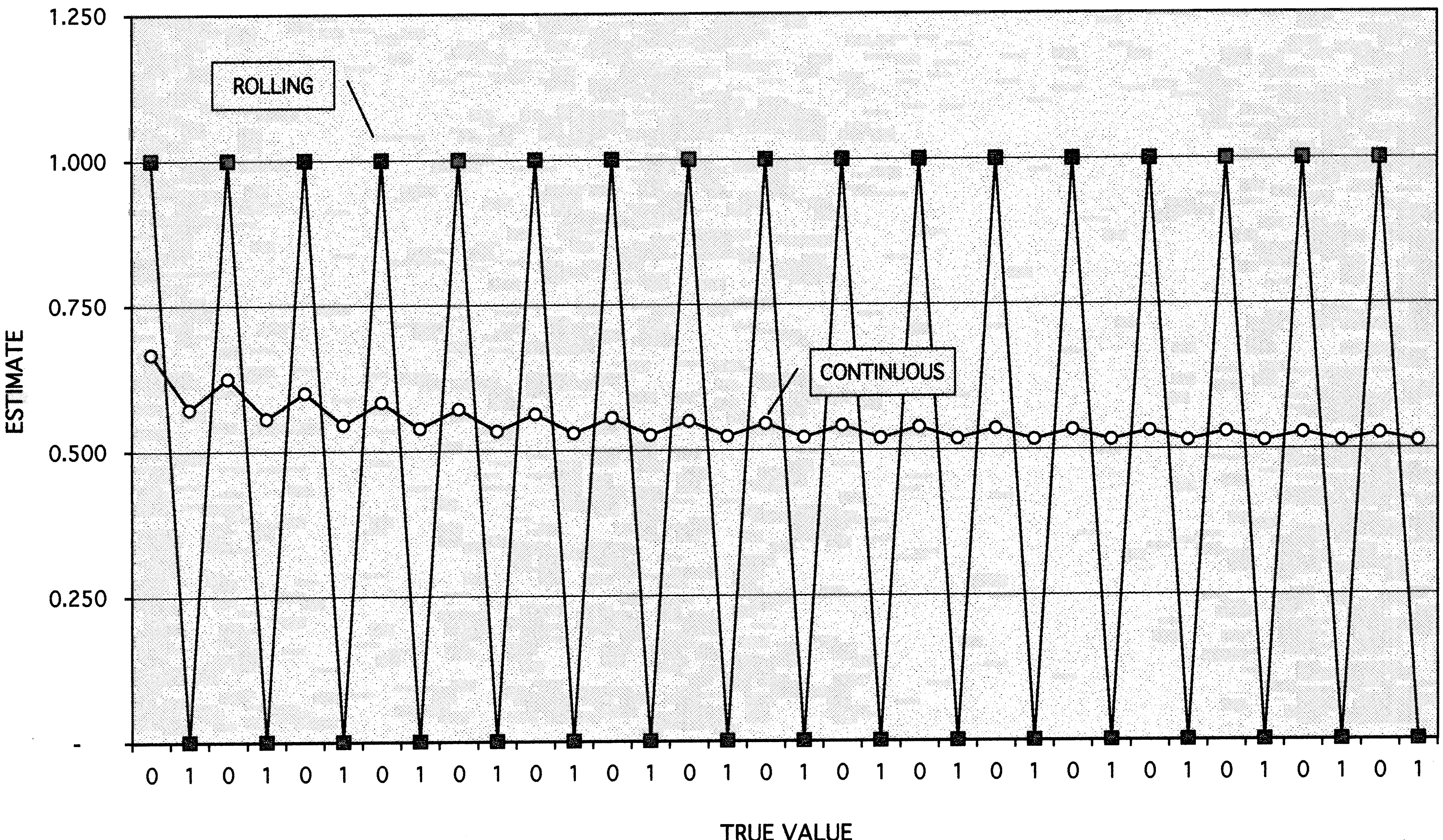
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-21)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ESTIMATES WITH ONES ALTERNATING WITH ZEROS
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY=1)



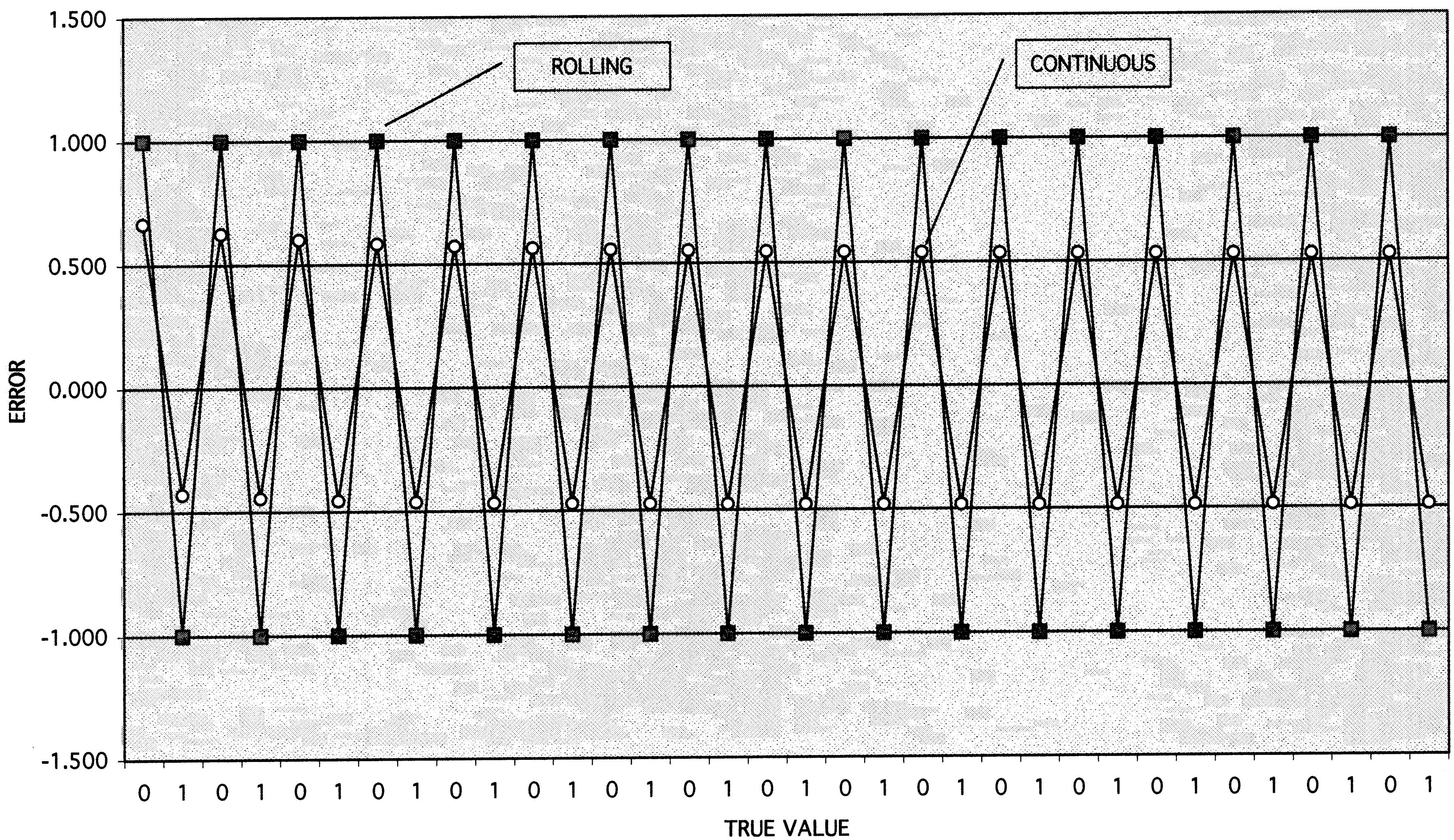
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-22)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

ERRORS WITH ONES AND ZEROS ALTERNATING
EXPECTED SERIES OUTCOME REMAINS AT ONE-HALF
CONTINUOUS RECORDS (MEMORY=40) VS 1-YR ROLLING (MEMORY =1)



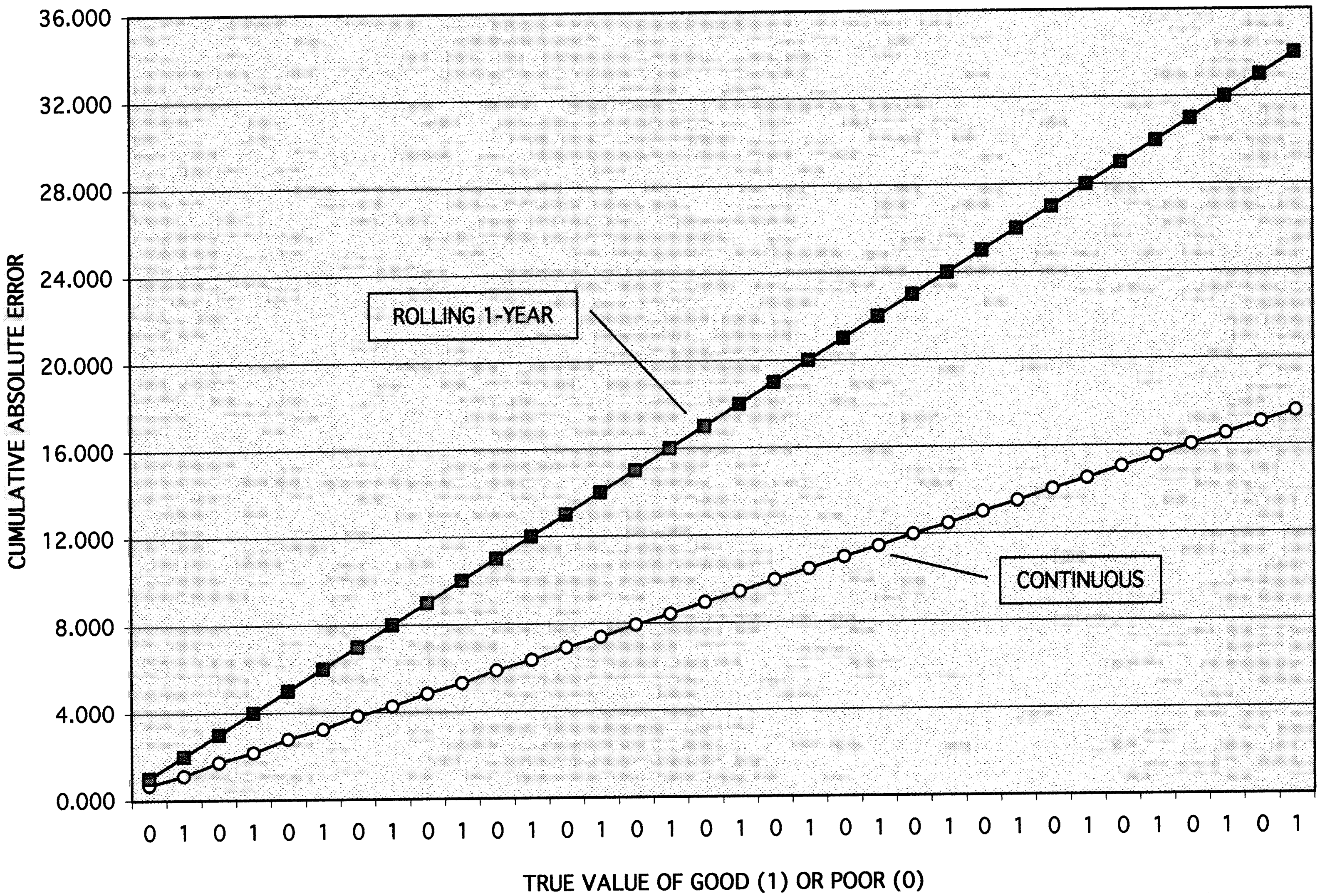
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-23)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

CUMULATIVE ERRORS WITH ONES AND ZEROS ALTERNATING CONTINUOUS RECORDS (MEMORY=40) VS I-YR ROLLING AVERAGE (MEMORY=1)



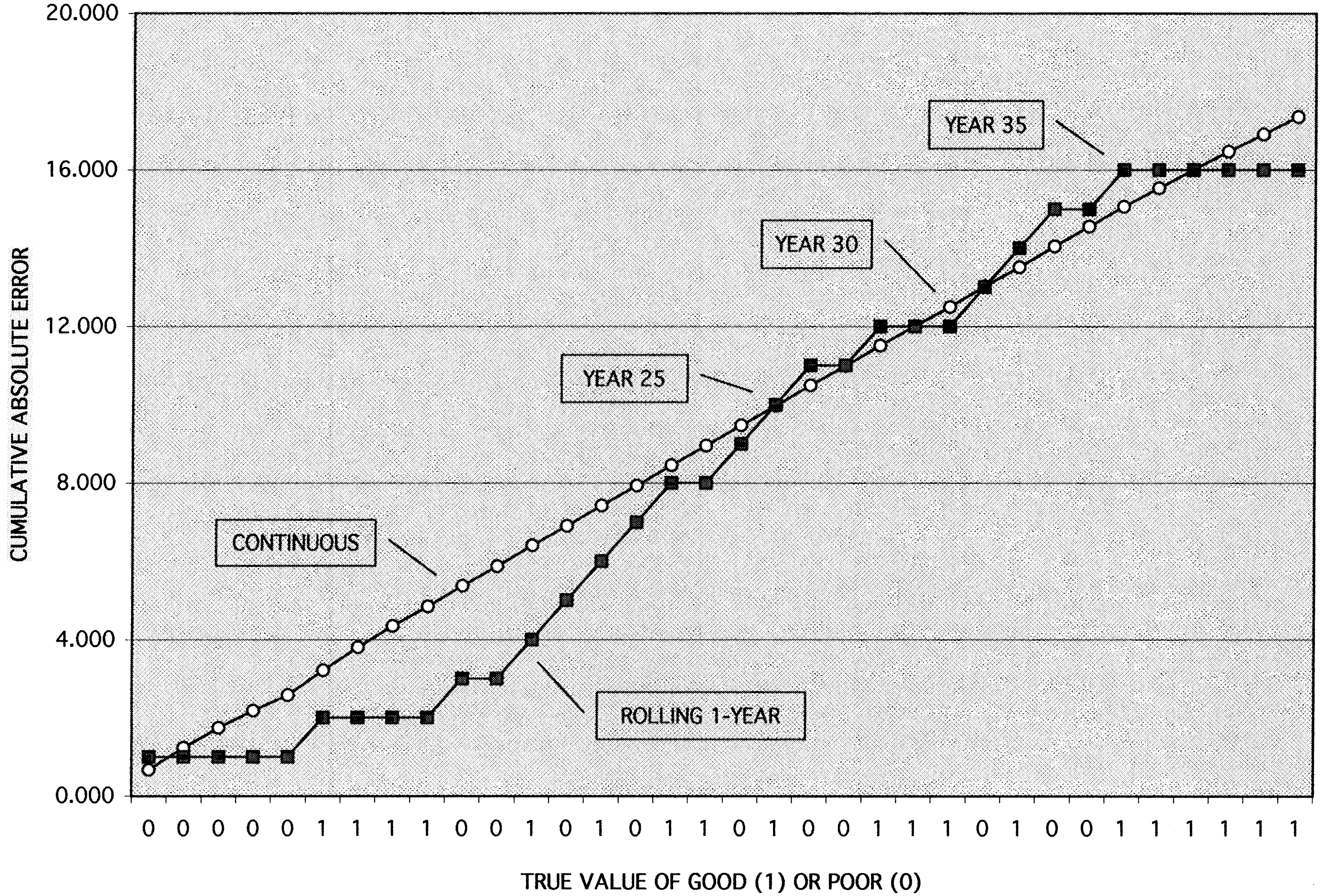
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-24)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

CUMULATIVE ABSOLUTE ERROR WITH ONES AND ZEROS PER WINTERFELD T-76 : CONTINUOUS RECORDS (MEMORY=40) VS I-YR ROLLING AVERAGE (MEMORY=1)



BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-25)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

DISTRIBUTION OF STREAMFLOWS FOR THE LAST 60 YEARS VS THE LAST 40 YEARS

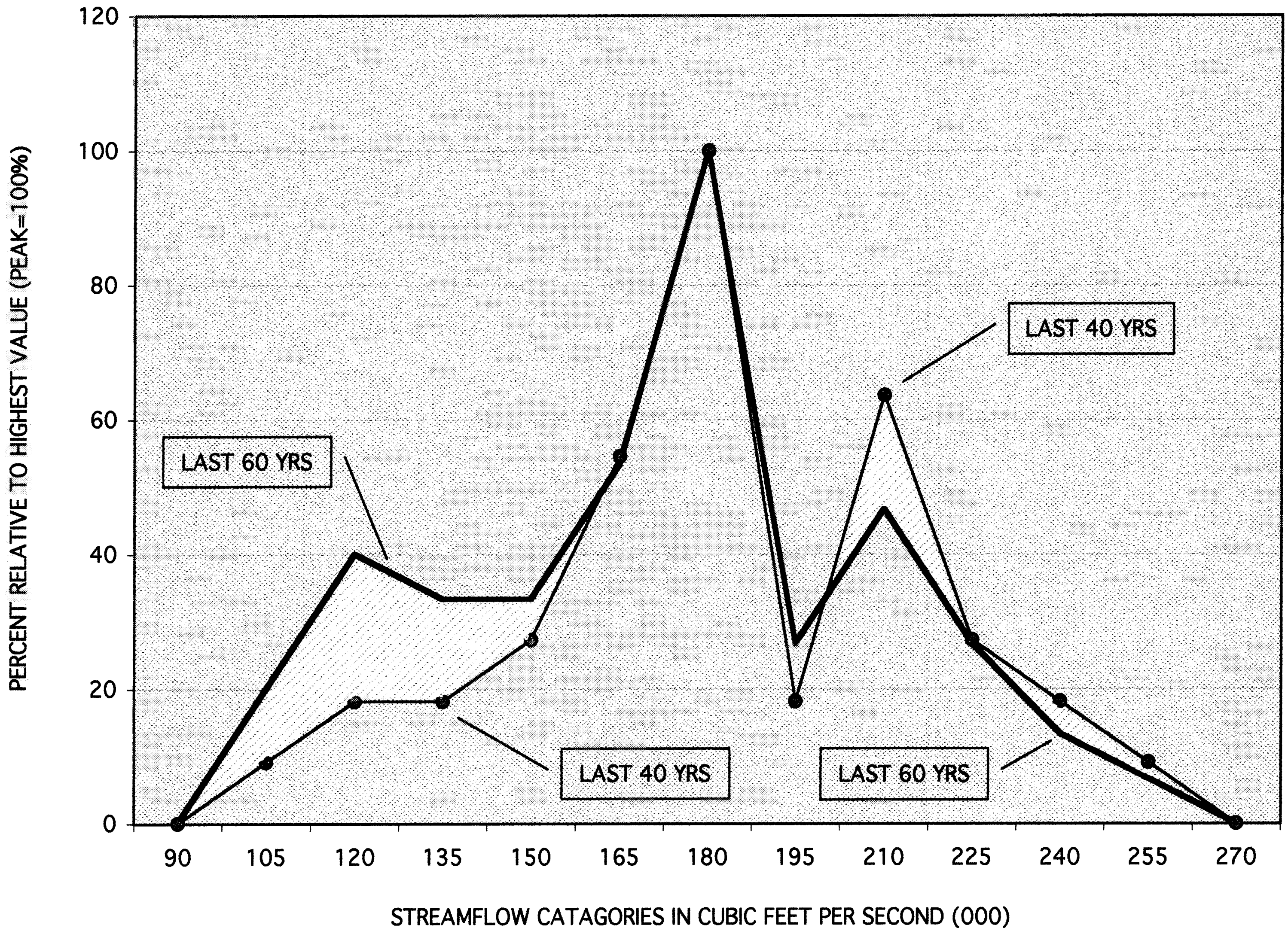


Exhibit No. ___ (TDD-25)
Docket No. UE-01 ___

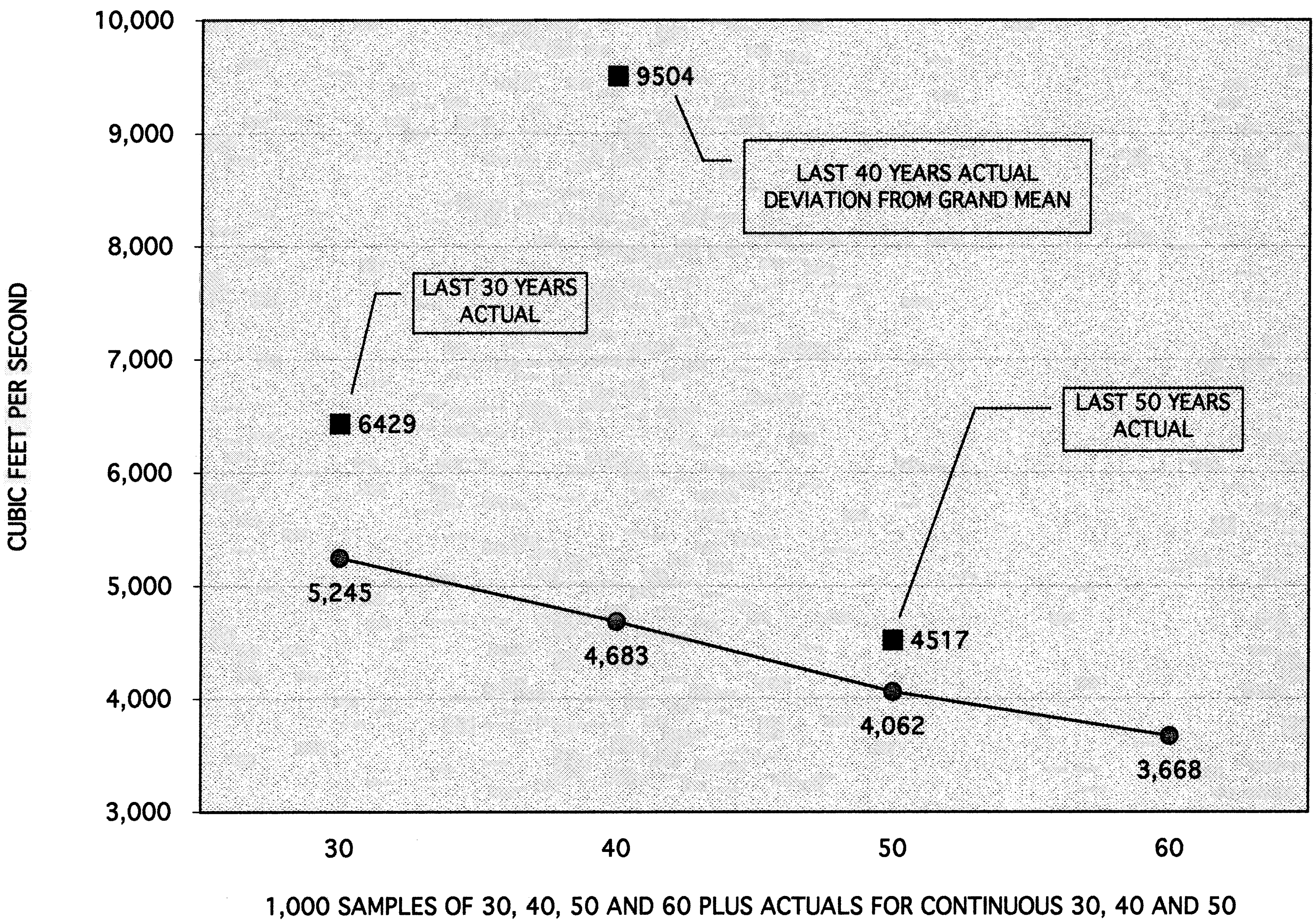
BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-26)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

AVERAGE DEVIATION FOR THE MEAN FROM THE GRAND MEAN FOR VARIOUS SAMPLE SIZES BASED ON 1,000 MEANS FOR EACH SAMPLE PLUS DEVIATION FOR THE LAST 30, 40 AND 50 CONTINUOUS WATER YEARS



BEFORE THE WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

DOCKET NO. UE-01______

EXHIBIT NO. ______(TDD-27)

WITNESS: DR. THOMAS D. DUKICH, AVISTA CORP.

RESAMPLING 90% CONFIDENCE LIMITS FOR THE MEAN AROUND THE GRAND MEAN FOR VARIOUS SAMPLE SIZES BASED ON 1,000 MEANS FOR EACH SAMPLE SIZE. ALSO SHOWS THE ACTUAL MEAN OF THE LAST CONTINUOUS 30, 40 AND 50 YEARS.

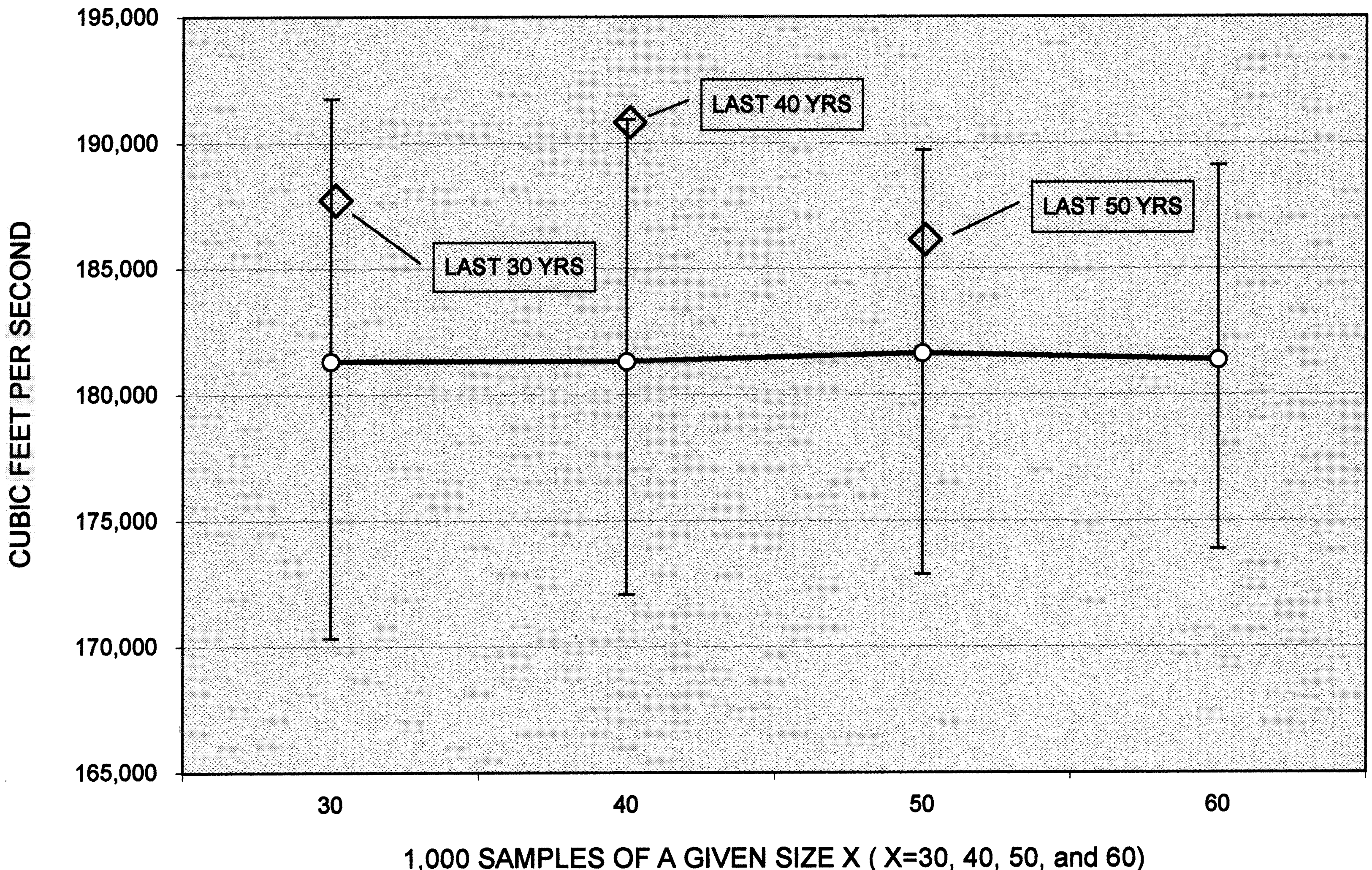


Exhibit No. ___ (TDD-27)
Docket No. UE-01 ___
Page 1 of 1