EXHIBIT NO. ___(JAD-1T)
DOCKET NO. UE-04____/UG-04____
2004 PSE GENERAL RATE CASE
WITNESS: JEFFREY A. DUBIN

BEFORE THE
WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION

WASHINGTON UTILITIES AND TRANSPORTATION COMMISSION, **Complainant,** **v.** **PUGET SOUND ENERGY, INC.,** **Respondent.**	**Docket No. UE-04____** **Docket No. UG-04____**

PREFILED DIRECT TESTIMONY OF JEFFREY A. DUBIN (NONCONFIDENTIAL) ON BEHALF OF PUGET SOUND ENERGY, INC.

APRIL 5, 2004

PUGET SOUND ENERGY, INC.

PREFILED DIRECT TESTIMONY OF JEFFREY A. DUBIN

CONTENTS

I. INTRODUCTION..2

II. DESCRIPTION OF HYDRO DATA AND ANALYSIS..6

III. RESULTS OF HYDROELECTRIC GENERATION ANALYSIS..10

IV. RESULTS OF HYDRO FLOW ANALYSIS..22

V. ESTIMATED FINANCIAL IMPACTS OF HYDRO FORECASTS ON PSE..25

VI. CONCLUSIONS AND RECOMMENDATION FOR SETTING THE HYDRO BASELINE FOR THE 2005 RATE YEAR..28

PUGET SOUND ENERGY, INC.

PREFILED DIRECT TESTIMONY OF JEFFREY A. DUBIN

I. INTRODUCTION

Q. Please state your name, business and address.

A. My name is Jeffrey Alan Dubin. My address is Pacific Economics Group, L.L.C. ("PEG"), 201 South Lake Street, Suite 400, Pasadena, California 91101.

Q. What is your position with PEG?

A. I am a Co-Founding Member of PEG.

Q. What are your duties as a member of PEG?

A. I actively consult with clients on demand issues, environmental issues, market issues, and antitrust policies, particularly as related to regulated industries. I specialize in microeconomic and micro-econometric modeling with an emphasis on limited dependent variable and demand analysis. Some of my current research topics include contingent valuation methods, discrete-choice econometrics, effects of welfare and entitlement programs on unemployment, energy economics and tax compliance, sampling and survey methods, valuation of intangible assets, and studies of ballot proposition voting.

Q. Do you hold any other positions?

A. I am a tenured Associate Professor of Economics at the California Institute of Technology (Caltech).

Q. What is your educational background?

A. I received my A.B. in Economics in 1978 from the University of California, Berkeley, with highest honors and great distinction. In 1982, I received my Ph.D. in Economics from the Massachusetts Institute of Technology.

Q. Please summarize your professional experience.

A. From 1982 to 1986, I was an Assistant Professor of Economics at Caltech. In 1988, I assumed my current position as a tenured Associate Professor. For over twenty years, I have taught at least two courses in econometrics per year at the undergraduate or graduate level at Caltech. I regularly use econometric methods in my empirical work and frequently rely on regression techniques.

Q. Have you published any papers or articles?

A. Yes. I have published several articles on energy and environmental issues, public utility regulation, competition and antitrust. A complete listing of my publications is included in Exhibit No. ___(JAD-2).

Q. Have you ever given expert testimony in a court or administrative proceeding?

A. Yes. A list of the proceedings in which I have provided expert testimony is also included in Exhibit No. ___(JAD–2).

Q. Who retained you for this testimony?

A. I have been retained to present testimony on behalf of Puget Sound Energy, Inc. ("PSE" or "the Company").

Q. What have you been asked to do in this case?

A. I have been asked to consider whether, from a statistical perspective, PSE's future hydroelectric generation should be estimated based on the most recent 40 years of information available (from 1948 to 1987), or whether it should be based on the full 60 years of available information (from 1928 through 1987).

Q. Please summarize your conclusions.

A. I analyzed the most currently available water flow and hydroelectric generation data for the 1928-1987 period. After analyzing the data, I conclude that there is no statistically significant trend in the 60-year hydro generation data. However, I conclude that there is a statistically significant downward trend in the 40-year (1948-1987) hydro generation data. My analysis also demonstrated a very significant correlation between natural hydro flow and hydro generation.

Further, although the hydro generation for the 60-year period is lower than for the

40-year period (836 aMW versus 867 aMW), I find that that there is no basis to conclude that the period 1928-1987 is biased or abnormal. Runs of lower water levels are expected by chance alone. I conclude that excluding the lower water observations from the 1928-1947 time period necessarily upwardly biases the forecasts of future likely water flows.

I also examined the natural water flow data and was not able to find any statistically significant trends in the data. I conclude that the best estimate of future water flow or future hydro generation should be derived using all available data to compute a simple long-run average. There is no statistical reason to exclude the data for the period 1928-1947. Using the full 60-year data set is a revenue neutral way to set the expected hydro generation level for PSE. I further conclude that using an average based on a 40-year window is not appropriate.

Forecasts based on ARIMA models demonstrate the same result: forecasts for 2005 are lower using the 60-year data period than they would be using the 40-year data period. PSE has estimated their financial loss that results from using the truncated 40-year period rather than the full 60-year period is approximately $11 million per annum. See the testimony of Ms. Julia Ryan, Exhibit No. ___(JMR-1T). My analysis indicates that the average annual losses could be substantially larger.

Q. How is the balance of your testimony organized?

A. In Section II, I describe the hydro data and my analysis of that data. In Section III, I present the results of my hydroelectric generation analyses. In Section IV, I present the results of my hydro flow analyses. In Section V, I present the estimated financial impacts of hydro forecasts on PSE. In Section VI, I present my conclusions and recommendation.

II. DESCRIPTION OF HYDRO DATA AND ANALYSIS

Q. How did you approach the analysis?

A. I requested the most current information on hydroelectric generation and natural water flows for the 1928 to 1987 period.

Q. What is the difference between actual hydro flow, natural hydro flow, and regulated hydro flow information?

A. Actual or raw water flow data represent the flow of water as it occurred historically. Natural flows are raw flows that have been adjusted to remove regulatory effects or the presence of dams. Regulated flows are natural flows that have been adjusted to reflect the current state of regulation and the current configuration of dams and the constraints imposed on the system.

Q. Do any of these measures closely approximate or correlate to hydroelectric generation?

A. Actual flows would probably best correlate to historic generation. However, actual flows would probably not be the best indicators of current and future generation due to the siting of dams, the rules protecting salmon and other fish, irrigation, etc. As these rules and conditions have changed over time, the water available for generation has similarly changed. Regulated flows would probably best correlate to the potential for current and future generation (*i.e.*, the generation potential under current regulation and constraints). It is likely that natural flows would also closely correlate to the potential for generation, especially if the measurements are made up-stream of the generation system.

Q. What is the process by which natural flow information is adjusted to estimate Mid-Columbia ("Mid-C") megawatt hour generation for PSE?

A. Approximately every ten years, the Northwest Power Pool (NWPP) estimates the water flow that would have existed absent the siting of dams or any regulatory restrictions on water use. The NWPP attempts to produce gross water flow estimates that are close to the levels that one might expect would result from natural precipitation and snow pack melting. This is the natural flow data estimate.

Next, the natural flow data is processed to apply current flow impediments such as siting of dams, flood control, fish protection, and other constraints to determine available hydro for generation. My understanding is PSE contracts with Avo Chillingerian to model adjustments to the natural water flow estimates subject to the current regulatory constraints. This is the regulated water flow data I mentioned previously. Additionally, Avo Chillingerian estimates hydroelectric generation by estimating the optimal use of the regulated hydro flow subject to the needs for load balancing within the region. This represents the hydroelectric flows that one could have expected under current regulatory rules and constraints for the period of 1928 to1987.

Q. Is hydroelectric generation data available for the period after 1987 at this time?

A. No. Generation data will be available in about six months. However, natural flow data for the period 1928 to 1998 for the Columbia River has been released by NWPP.

Q. Please describe the hydroelectric generation information you received from PSE?

A. I received information on hydroelectric generation for the period from 1928 through 1987. PSE maintains this information in two components. In the first component, PSE estimates hydroelectric generation given current capacity and regulatory constraints for its own "Westside" hydro projects, which includes projects at Upper Baker, Lower Baker, and other small facilities. In the second component, PSE maintains hydroelectric generation data for its Mid-C projects as estimated by Avo Chillingerian. This data was provided to me for water years 1928 through 1987 on a monthly basis.

Q. How has PSE defined water year?

A. A water year consists of the months August through December of a given calendar year and the months of January through July of the next calendar year. Thus, a water year is 12 months, but crosses two calendar years.

Q. What did you do with the Westside hydro project data?

A. I combined the historical megawatt generation at Upper Baker, Lower Baker, and other PSE small hydro facilities into a combined Westside total.

Q. Did you treat the Mid-C projects in a similar fashion?

A. Yes, I did. However, the ownership of the Mid-C projects, including Wells, Rocky Reach, Wanapum, Priest Rapids, and Rock Island 1 and 2 has changed over time. To estimate hydroelectric generation in the rate year from March 2005 through February 2006 ("2005 Rate Year"), I asked for, and received from PSE, the most current ownership percentages that PSE has in these projects. I understand that, at present, PSE is entitled to 31.3% of the Wells Project, 38.9% of Rocky Reach, 50% of Rock Island I Project, 75% of Rock Island II Project, 10.8% of Wanapum Project, and 8% of Priest Rapids Project. By applying these percentages to the hydro generation estimates, I estimate the hydro power available to PSE under current regulatory conditions and historically available natural hydro flow.

Q. What further steps did you take to prepare the data for statistical analysis?

A. First, I determined the total hydro power historically available to PSE by combining the megawatt hours generated from their Mid-C and Westside projects. I then divide by the number of hours in a year to obtain the annual average megawatts of hydroelectric generation (aMW).

III. RESULTS OF HYDROELECTRIC GENERATION ANALYSIS

Q. Have you displayed the total estimated PSE hydro generation over the 60-year period?

A. Yes. In Exhibit No. ___(JAD-3), I show the aMWs for each of the years 1928 through 1987. The average over the entire period is approximately 836 aMW.

Q. Is there a significant trend in PSE's hydro generation during this period of time?

A. No. Although Exhibit No. ___(JAD-3) appears to show a slight upward trend, there is no statistically-significant trend. A regression analysis I performed examining PSE's 60-year hydro generation did not find a significant upward or downward trend in the hydro generation (t-statistic = 1.80).

Q. Have you also isolated the period from 1948 through 1987 and displayed that information for PSE's average hourly hydro generation?

A. Yes, I do this in Exhibit No. ___(JAD-4). I find that for the 40-year period the average level of hydro generation is 867 aMW and is therefore 31 aMW higher than the average for the 60-year period. I also determined that there was a statistically significant downward trend in the hydro generation during this period of time (t-statistic = -2.74). I show this in Exhibit No. ___(JAD-4).

Q. How do these generation estimates compare to the natural water flows?

A. I used the natural flow data measured at the Grand Coulee reservoir to make a comparison to the hydro generation data. The location of the Grand Coulee relative to the Mid-C projects is shown in Exhibit No. ___(JAD-5).

I understand that the Westside projects are not hydrologically associated with the Mid-C projects. Therefore, I separately collected natural flow information for the Westside projects and combined this with the Grand Coulee natural flow information to make a combined flow measure. Regression analysis between hydro generation and natural water flow shows very significant correlation between the two series ($R^2 = 0.80$ t-statistic = 15.3).

Q. Can you illustrate this correlation in a figure?

A. Yes. In Exhibit No. ___(JAD-6), I illustrate the degree of correlation between estimated historical generation and estimated natural flow. In this exhibit, I have simultaneously graphed total PSE generation and total water flow at Mid-C and

Westside. There is a very close connection between total PSE generation and total water flow at Mid-C and Westside. This was expected.

Q. Why is this important?

A. The finding is significant because we are analyzing a natural geological phenomenon. Hydro generation is largely a function of the available water, which itself is a function of precipitation and snowmelt. In contrast, other time-series, such as the gross domestic product (GDP) in the United States economy, or the rate of unemployment, are results of complex economic processes in our economy.

Q. Have you examined the variability and shape of the distribution of hydro generation in the 60-year period?

A. Yes. In Exhibit No. ___(JAD-7), I present a histogram of PSE's hydro generation for the 60-year period. I have performed a statistical test for normality to determine that the data are approximately normally distributed (Shapiro-Wilk normality test: p- value=0.22) or bell-shaped.

Q. Are you aware that some have argued that because the 1928-1947 period represents significantly lower levels of hydro generation (and water flow)

relative to the period subsequent to 1948, it must be a biased measure of water flow and therefore should be excluded?

A. Yes, I am aware of this argument. However, I do not believe it is correct.

Q. Why do you disagree?

A. A classic example of the flaw in this type of argument involves flipping a coin. Imagine a fair coin being flipped many times and, for the sake of argument, imagine that "heads" represents a normal water year and that "tails" represents a low water year. It is well known from common experience, and probability analysis can demonstrate, that long observed runs of heads or tails are not that uncommon. As noted by William Feller (New York: John Wiley and Sons, 1968, pp. 72) in an *Introduction to Probability Theory and Its Applications:* "…the laws governing a prolonged series of individual observations will show patterns and averages far removed for a whole population." Thus, if one were to exclude from the data a string of coin flips that initially turned up tails and only examine a subsequent time period, one could falsely conclude that the coin generating the data was biased. The same false conclusion will result if one is to exclude earlier, lower water years. Doing so is more likely to produce high water year outcomes.

Q. Can natural water flow exhibit long periods of low water years?

A. Yes. H. E. Hurst in his 1951 study of reservoir design capacity was one of the first to observe that precipitation can be very large or small and that periods of low or high flows can be very long. This empirical regularity of water flow data was called the "Noah and Joseph" effect by Mandelbrot and Wallis in their 1968 published study (See B. Mandelbrot and J Wallis, "Some Long-Run Properties of Geophysical Records," *Water Resources Research*, Volume 5, 1969, pp. 321-40.) The persistence of river flows is known in Hydrology as the Hurst effect. (See e.g. Bras, R. and I. Rodriguez-Iturbe, *Random Functions and Hydrology*, Massachusetts: Addison-Wesly Publishing Company, 1985, Chapter 5). Mandelbrot and Wallis extended the persistence concepts from Hydrology to Financial Economics (See e.g. Mills, T. *The Econometric Modeling of Financial Times Series*, New York: Cambridge University Press, 1999.)

Q. Do you agree that interpretation of short climate records is potentially error prone?

A. Yes. Carl Wunsch of MIT has cautioned that apparent changes in the mean levels of various phenomena, such as temperature or shifts in the return time or oscillation period of El Nino-like processes, may be illusory. (See Carl Wunsch, "The Interpretation of Short Climate Records, with Comments on the North Atlantic and Southern Oscillations," *Bulletin of the American Meteorological Society*, Volume 80, 1999, pp. 245-255.) Wunsch writes that "before concluding that one is seeing evidence for trends, shifts in the mean, or changes in oscillation periods, one must rule out the purely random fluctuations expected in stationary time series." He further states that "Claims that one has entered a new climate state because of extended excursions away from the mean need to be tested against the null hypothesis that the extreme duration is expected to occur by chance at some particular confidence interval." In our situation, we have no basis to conclude that the period from 1928 through 1947 is biased or abnormal. On the contrary, we should expect runs of lower water levels by chance alone. Indeed, my understanding is that the Pacific Northwest in recent years has observed flows comparable to those observed in the period of 1928 to 1947. To exclude these observations would however necessarily upwardly bias the forecasts of future likely water flows.

Q. Would you agree that time series econometric methods are regularly applied to hydrological processes?

A. Yes. An entire discipline devoted to such methods is called Stochastic Hydrology. The methods applied to the study of hydro flow are well developed and very similar to those used in Time-Series Econometrics. For instance, the basic Auto-Regressive statistical model is known as the Thomas-Fiering model in Hydrology and is one of most popular models of water flow (See e. g. A. Maass, M. Hufschmidt, R. Dorfman, H. Thomas, S. Marglin, and G. Fair, *Design of Water-Resource Systems*, Cambridge, Mass.: Harvard University Press, 1962).

Q. Have you conducted a statistical analysis of the difference between the first 20 years of the 60-year period and the subsequent 40 years?

A. Yes. I conducted an equality of means test to compare the hydroelectric generation in the two periods under the hypothesis that a normal distribution generated the data between 1928 and 1947 and that another normal distribution with a possibly different mean generated the data from the period from 1948 to 1987. This is the type of test that Public Counsel's witness Dr. Blackmon conducted in Docket No. UE-921262.

The statistical hypothesis is that the means of these two distributions are equal and that any measured difference is explainable by random variation alone. The estimated averages for the 20-year period and the 40-year period are, respectively, 774.8 aMW versus 867.1 aMW. (The p value for the difference in means is 0.001 – a statistically significant difference under the assumptions of the test).

Q. Is it possible with this knowledge to draw the conclusion that the observations in the two time periods result from different processes?

A. No. One can reject the hypothesis that the means are the same only under the maintained assumptions of the statistical hypothesis test (two distributions, normality, constant variance, and independent samples). However, these findings do not imply that two independent distributions with constant and unvarying means are generating these data and these results.

Q. Can you explain with an example?

A. Yes. Real Gross Domestic Product ("GDP") in our economy has grown significantly and steadily in the post-war period since 1945. One might ask whether the mean level of GDP was the same in the first 20 years, from 1945 to 1964, as it was in the second 20 years, from 1965 to 1985. One would reject the hypothesis that the mean levels of real GDP are identical in these two periods.

The presence of a trend in the data can manifest itself in a finding that certain subsets of the data have statistically different means, even though only one stationary process is generating the entire data series.

Q. Assuming that there were a trend in the water flow data, would you agree that only the most recent data should be examined when forecasting?

A. No. There are many ways to do forecasts. Some optimal forecasts may only require using recent data. Other optimal forecasts will require the full time-history available. The argument that, in the presence of the trend, only a recent subset of data should be used assumes that the forecast to be made in the future will be only based on an average and that because it is based on an average, it will be most distorted if earlier period information is used rather than only later period information.

Q. Could you provide an example?

A. An example is GDP. Recall that GDP is rising over time. If I were to forecast the GDP level in the near future and were constrained to make my estimate based on an average, I might rely only on the most recent data. However, nothing constrains me to use a simple average of the recent year information. Instead, I should observe the trend in the information and rely on that trend as part of my forecast. In such a way, I would use all available data and produce no bias in my estimates. This is the full information technique.

Q. Is it proper to define the estimation period by testing for changes in the mean?

A. No. I am aware that witnesses in other proceedings have testified that this is appropriate with regard to the hydro issue. In such an approach, one starts at the most recent period, say a ten-year period, and asks whether the next most recent ten years has the same average level as the initial ten years. If it does, these data are included in this approach and if it does not, one stops. Based on such a procedure, some have concluded that the 1928 through 1947 period has a significantly different average level than later periods and, therefore, is a biased addition to the estimates of the future. This is not correct. I know of no statistical tests that repeatedly examine time periods to ask whether they are the same or different before including them in a model. All information must be used simultaneously.

Q. Can one get a biased inference by using too little data?

A. Yes. By eliminating the period from 1928 through 1947, it is possible to reach significantly different conclusions about the levels of water flow and the directions in which the flow rates are going. In fact, as I have mentioned already, looking at the 60-year picture reveals a time period where flows are normally distributed with no significant trend. This is the kind of data I might expect in the presence of geological phenomena such as water flow. However, looking at the 40-year period from 1948 to 1987, I noticed that there is now a statistically significant downward trend in the estimated hydro generation that was not apparent in the full 60-year period.

Q. You just mentioned that you would expect hydro generation to be randomly distributed because of geological phenomena. Are you a geologist?

A. No. However, unlike economic time series where the behavior of the economy leads to integrated or trended time series or times-series exhibiting strong auto-correlation, it is my understanding that geological processes occur much more randomly. Having said that, I should point out that the adjustments for load management and regulation have the potential to introduce a persistence in the data that may not already be present in the underlying natural geological processes. In any case, these are empirical matters that can be tested. In fact, I have found that that there is evidence for some short-term persistence in the generation data, which means that when generation is otherwise higher in a given year, it is somewhat more likely to be higher in a subsequent year. However, that persistence is a relatively short-lived phenomenon and does not influence long-term forecasts, e.g. forecasts made using data through 1987 of likely hydroelectric generation in 2005.

Q. What statistical tests, other than testing for normality, did you perform to discover the process that generates PSE's hydroelectric flows?

A. One can analyze hydroelectric generation or hydro flows using time-series statistics. A vast literature has been developed for analyzing such processes and for doing optimal forecasting using such data. For instance, a discussion of such techniques is provided by the Nobel Prize winning author, Clive Granger in his textbook (See e.g. C. Granger, *Forecasting in Business and Economics*, New York: Academic Press, 1980 or Bras and Rodriguez-Iturbe (1985) or Vujica Yevjevich, *Stochastic Processes in Hydrology*, Colorado: Water Resources Publications, 1972).

Q. What approach did you follow?

A. I followed the Box-Jenkins approach to model building. In the Box-Jenkins approach, there are three basic steps. The first step is to identify the underlying data process and to choose a statistical model. The second step is to estimate the parameters of the model that best represents the underlying data. The final step is to do diagnostic testing to see whether further models should be considered.

Q. What tests did you do specifically to identify these time series models?

A. I used correlogram and partial correlogram methods to estimate the auto-covariances in the time-series process.

Q. What did you discover with this analysis?

A. I discovered that there was some persistence in the hydroelectric flow information and that the flow information could be modeled by a "low order" auto-regressive moving average process.

Q. Please explain briefly what an auto-regressive moving-average model is?

A. An auto-regressive moving-average model postulates that the stochastic time-series process is, in part, a distributed lag of previous values of the process and a moving average of historical random components. The degree of these auto-regressive and moving- average components is identified by the correlogram analysis.

Q. Did you perform any other tests?

A. Yes. As a threshold matter, it is important to discover whether the data are generated by a random walk with drift or by a stationary time-series process. To do so I used the Dickey-Fuller test with Perron adjustments (See D. Dickey and W. Fuller, "Distribution of the Estimators for Autoregessive Time Series with a Unit Root," Journal *of the American Statistical Association*, Volume 74, pp. 427-31, 1979 and P. Perron, "The Great Crash, and the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, Volume 57, pp. 1361-401, 1989.).

Q. What conclusions did you reach?

A. I conclude that the random process generating the hydroelectric generation data is stationary and there are no concerns for so-called unit roots or for integrated random processes. A stationary time-series is randomly distributed around its mean. A non-stationary time-series will "explode" with increasing variance over time. Random perturbations are transitory with stationary series but are permanent with non-stationary series.

IV. RESULTS OF HYDRO FLOW ANALYSIS

Q. Have you also examined the natural flow data supplied to you for the period from 1928 through 1998 and various sub-periods?

A. Yes. I did so for natural flows measured at Grand Coulee. The natural flows are shown in Exhibit No. ___(JAD-8). Flows are measured in millions of acre feet (MAF) per annum.

Q. What are the average flows that you determined?

A. I found that the average flow for the period 1928 through 1998 was 78.09 MAF. I found that the average flow for the period 1928 through 1947 was 71.58 MAF, the average flow for the period from 1948 through 1987 was 81.33 MAF, the average flow for the 40 year period from 1959 through 1998 was 79.50 MAF and the average flow from 1948 through 1998 was 80.64 MAF.

Q. Did you do tests for equality of means?

A. Yes. The period from 1928 through 1947 has a statistically significant lower mean water flow level than for the period from 1948 through 1987 or when compared to the period from 1948 through 1998 (t-statistics are 2.6 and 2.5 respectively). This conclusion rests on the assumptions of the equality of means test and does imply that one period is biased. Nor can one conclude, based on this test, that there has been a geological structural shift of some type in the hydrologic process.

Q. Did you find trend or cycles in this data?

A. No. I used a regression analysis to test for structural shifts and the presence of deterministic trends. I was not able to find statistically significant trends in the water flow in the data. The trend that had been observed in generation is not present at statistically significant levels in the post 1948 period in the natural water flow data (t-statistic = -1.78). This supports the finding that the 40-year period is too short to correctly model this geological phenomenon.

Q. Did you test for stationarity in the water flow data?

A. Yes. The data are generated by stationary stochastic processes. Yevjevich (1972) has concluded that this is generally true for hydrologic series. He notes that while river basins and climate change slowly with time, "these changes are relatively small in the time span of a couple decades or centuries, man-made effects and natural disruptions excluded, so that many hydrologic series … may be considered stationary from a practical point of view."

Q. Did you examine the natural water flow data using correlogram, partial correlogram, and ARIMA analysis?

A. Yes. I did not find significant departure from a purely random or very low-order ARIMA processes in the data. This implies that there is likely to be very little (i.e. short-term) or no forecastability in the natural water flow data.

Q. Did you test for normality in the natural water flow data or outliers?

A. Yes. I could not reject the hypothesis that water flows are normally distributed (Shapiro-Wilk normality test: p-value=0.55). I did not detect statistically significant outliers.

Q. What are your conclusions based on your analysis of natural water flow data for the period 1928 through 1998?

A. The picture that emerges for natural water flow is actually quite simple. The data are trendless. The data are normally distributed. The data are not forecastable or reflect a very short-lived persistence. The best estimate of future natural water flow is consequently a simple long-run average using all available data.

Q. Does your analysis of the natural water flow data support your analysis of the hydro-electric generation that you discussed earlier?

A. Yes. As I previously mentioned, the processes are very highly correlated. The lack of a significant trend in the natural water flow data further supports my opinion that the structural shifts and trends discovered in hydro generation data for the period from 1928 through 1987 are likely to be aberrations from examining an incomplete geological record. The full available time history should be used to set hydro flow levels for the 2005 Rate Year.

V. ESTIMATED FINANCIAL IMPACTS OF HYDRO FORECASTS ON PSE

Q. Based on the modeling that you've done, have you made forecasts for the rate year beginning March 2005 ("2005 Rate Year") of expected hydro generation and have you compared those to the levels established using a 40-year average for the period 1948-1987?

A. Yes, I have. In Exhibit Nos. ___(JAD-9) through ___(JAD-13), I show ARIMA models with various degrees of moving-average and auto-regressive and integration components. Exhibit No. ___(JAD-9) is based on the average level of hydro generation for the 60-year period 1928-1987. The forecast for 2005 for this model is 836.3 aMW. Exhibit No. ___(JAD-10) displays the forecasts from an ARIMA model with one order of auto-regressive component, no integration, and zero degree moving average component (ARIMA 100). As I mentioned previously, this model was introduced and popularized in Hydrology by Thomas and Fiering. This model shows that a small persistence effect *i.e.*, generation that was below average in 1987, would be forecasted to return to average levels relatively quickly, reaching those average levels by 1995. The forecast for 2005 from this ARIMA model is 831.6 aMW.

Q. Did you also increase the lag length and consider an ARIMA model with two orders for the moving-average and auto-regressive components?

A. Yes I did and the forecast for 2005 in this case is 823.5 aMW. The forecast is presented in Exhibit No. ___(JAD-11).

Q. Did you also do ARIMA models with structural shifts accounting for the possibility that the period from 1928 through 1948 was somehow different than the period after that?

A. Yes. The technique that I employed follows the Perron model, which is an ARIMA model with structural shift.

Q. Which exhibits apply to these models?

A. The ARIMA shifts are displayed in Exhibit Nos. ___(JAD-12), and ___(JAD-13). The predictions for hydro generation for 2005 are respectively, 786.9 and 696.2 aMW.

Q. Is there a pattern to these predictions?

A. Yes. All of the ARIMA models produce forecasts for 2005 that are lower than one would get using a 60-year average.

Q. What is the financial effect for PSE of using 40- versus 60-year hydro data?

A. I show the financial effect in Exhibit No. ___(JAD-14). In this exhibit, I show the aMW level established by the Commission based on the 40-year average from 1948 through 1987. That level is 867.1 aMW. I then compute the cost to PSE from setting the 40-year baseline at 867.1 aMW if, in fact, the hydro generation in the 2005 Rate Year occurs at its forecasted level. Using a 60-year average to forecast the likely level of generation in 2005 leads to roughly an $11 million loss per annum (as estimated by PSE).

Q. Do these estimates match those of PSE for the 40-year average versus the 60-year average scenario?

A. Not exactly, but the differences are not consequential to my analysis. There are some minor differences between my cost analysis and the cost estimates based on AURORA runs done by PSE. In general, my forecasts are somewhat larger than the AURORA forecasts using 40-year and 60-year information. That is a consequence of the adjustments made to estimated generation for maintenance and outages that the Company anticipates. A further adjustment occurs due to a decline in the ownership percentage in certain projects that occur in the 2005 Rate Year.

Q. How do you estimate the cost differences?

A. I express my calculations on a percentage basis relative to the base case that was performed by PSE in which they compared a 40-year average hydro level to a 60-year hydro level and calculate the additional costs needed to replace hydro generation with gas-driven generation. PSE's estimate in this base case is roughly $11.2 million. In Exhibit No. ___(JAD-14), I calculate the loss to PSE using various ARIMA forecasts for PSE's hydro generation for the 2005 Rate Year.

Q. What do you conclude from this analysis?

A. Using ARIMA models to forecast the likely level of generation in 2005 leads to large annual expected losses. For example, the ARIMA (100) model forecasts that hydro generation is actually likely to be lower than that forecasted by a 60-year average for the 2005 Rate Year. PSE would actually experience a larger estimated cost than had previously estimated in its base case runs. In fact, some of the ARIMA models with structural shift imply that hydro generation is trending so significantly downward that setting a baseline based on a 40-year average will subject PSE to an annual power cost deficit of $31 to $43 million.

VI. CONCLUSIONS AND RECOMMENDATION FOR SETTING THE HYDRO BASELINE FOR THE 2005 RATE YEAR

Q. Based on your analysis, what statistical process describes the hydroelectric generation and natural water flows for PSE?

A. Based on the Dickey-Fuller tests, the correlogram analysis, tests for normality and for structural shifts and trends, and using the full information available to me, I conclude that PSE's hydro generation and hydro flow is best characterized by a relatively low order auto-regressive moving-average model. This process shows relatively short-lived persistence and returns to the long-run average fairly quickly. Some evidence remains for negative trends in generation and flows after 1948. However, absent some external geological evidence, this trend may not be permanent and may instead be an artifact of examining an incomplete geological record.

For these reasons, I advocate using the long-run average value based on the full information available. In my opinion, a 40-year moving average is not appropriate and is too short for a geological series such as this one. There is no reason to exclude the data from the first 20 years of this 60-year data series.

Q. Does using the full information available increase the statistical precision of the estimates?

A. Yes. Increasing the sample from 40 to 60 years will reduce the estimated standard deviation (the square root of the estimated variance) by just over 18 percent which may be useful if one is interested in the range of probable outcomes (*i.e.* the confidence interval around the forecasted mean). This result requires that the variance of the distribution be constant for all observations.

Q. Was the variance constant between the period 1928-1947 and 1948-1987?

A. No. An equality of variance test indicates that the variances in the two periods are different (under the maintained assumptions of the test). Perhaps of more importance is that an inference regarding the range of possible hydro flow or hydro generation based on the 40-year period is likely to understate the likely variation by at least 15 percent.

Q. Does excluding the first 20 years of information bias the estimated mean hydro generation?

A. Yes. As I have discussed, persistence of low or high water flow periods in hydrological observations is an empirical regularity. Sampling a part of the available information has every potential to bias estimates of mean flow or to "find" trends in the data that are in fact not at all present. Robert Pindyck calls this possibility "data snooping" and concludes that "for any long time series, it is likely that one will always find one or more 'structural changes.'" He discounts these findings when one "provides no explanation for what such a structural change means or why it occurred." (See Robert Pindyck, "The Long-Run Evolution of Energy Prices," *The Energy Journal*, Volume 20, pp. 1-27, 1999).

Q. Can you explain that further?

A. Yes. If one were to conclude that the first 20-year period (1928-1947) comes from a different geological process than the last 40-year period (1948-1987), we would be left with the inference that in a 60-year period, one-third of that time was subject to a different geological process. In statistics, this is termed a mixture model. (Examples include James Hamilton, "Analysis of Time Series Subject to Changes in Regime," *Journal of Econometrics*, Volume 45, 1990, pp 39-70 and James Hamilton, "A New Approach to the Economic Analysis of Nonstationary Time Series Subject to Changes in Regime," *Econometrica*, Volume 57, 1989, pp. 357-384.) In this case, we might say that there is a one-third chance that we are in a "low"-water period and a two-thirds chance of being in a "normal"-water period. In future years, we might expect to see the "low"-water period one-third of the time and the "normal"-water period two-thirds of the time. Indeed it is possible that such "low"-water years have returned already (see e.g. Robert Service, "As the West Goes Dry," *Science*, Vol. 303, 20 February 2004, pp. 1124-1127).

Q. What is the net result of this mixture of the two states of nature?

A. The net result for forecasting purposes is that one should combine, one-third of the time, the average level of water from the "low"-water period (first 20 years) with the average level of water in "normal"-water period (subsequent 40 years). The resulting forecast would equal the full time period forecast. This is the approach I recommend. A finding that water flows and generation in the 1928 through 1947 periods may be lower is good evidence that this had occurred and can occur again. Such information should not be rejected.

Q. Is there a best level to set the hydroelectric generation level for the 2005 Rate Year using the data that we have at hand?

A. Yes. The revenue neutral way to set the expected hydro generation level would be to use the long-run average value for the full 60 years (836 aMW) based on the evidence we have at hand. In this case, the estimated cost difference to PSE would be neither positive nor negative. It would simply be zero. Alternatively, if the Commission adopts the 40 year period from 1948 through 1987 to establish the likely hydro generation level in future years, it should acknowledge and incorporate the trends in this data. Following the trends apparent in this data for this time period compels a lower estimate of hydro generation than is estimated using the 60-year long-run average.

Q. What do you conclude with respect to risk or variability if the Commission does not set the level at the 60-year number?

A. Recall that variance measured around an assumed level is equal to the variance measured around the average level plus the square of the bias (*i.e.*, mean-squared error equals variance plus bias squared). In the presence of bias, the variance must increase. In other words, in order to produce estimates with the lowest variance, one must eliminate the bias. If the Commission sets the target hydro generation at a level different from long-run average, it, in effect, produces two deleterious effects. First, the Company, on average, starts off with a roughly $11 million underrecovery of its power costs. Second, the variance in revenues increases due to the incorrectly set baseline figure.

Q. Since you advocate using all available data for forecasting, would you revise your forecasts to use all data beyond 1987 when it becomes available?

A. Yes.

Q. Does this conclude your testimony?

A. It does at this time.

[BA040850.045 / 07771-0089