November 19, 2007

Alberta Energy and Utilities Board
640 – 5th Avenue SW
Calgary, Alberta
T2P 3G4

Attention: Jamie Cameron, Application Officer

Dear Jamie:

Re: AESO Responses to EUB Information Requests on AESO Comments on EUB Approach to POD Cost Function

In a letter from the Alberta Energy and Utilities Board (EUB) dated November 15, 2007, the Alberta Electric System Operator (AESO) was requested to answer two information requests on its comments on the Point of Delivery (POD) cost function developed by EUB staff. Responses to those information requests have been submitted via the EUB Digital Data Submission System.

In addition, the AESO would like to take this opportunity to express concern at the general direction of the information requests, and of the approach to the development of a POD cost function more broadly. Specifically, the AESO does not feel that the limited data examined in this proceeding supports the level of precision apparently being ascribed to the results of regression analysis.

The AESO asks that this context be remembered when reviewing these information responses. As noted several times by the AESO during the GTA proceeding, the appropriate tariff cannot be proven with mathematical certainty, but rather requires application of judgment to determine whether it is just and reasonable.

If you have any questions on these information responses or need additional information, please contact me at (403) 539-2465 or by e-mail to john.martin@aeso.ca.

Sincerely,

[original signed by]

John Martin
Director, Tariff Applications

Attachments

cc: Heidi Kirrmaier, Vice President, Regulatory, AESO

2500, 330 - 5th Ave SW Calgary, Alberta T2P 0L4
t (403) 539-2450 | f (403) 539-2949 | www.aeso.ca
Preamble: The AESO’s submission states:

“The AESO understands that the goal of the analysis is to determine which curve of those analyzed had the highest correlation coefficient (R2 value). Of the five curves analyzed, the AESO notes that only the power curve intercepts the y-axis at zero, which suggests there should be no fixed cost component in the POD cost function. This does not align with the evidence on the record in this proceeding, nor with the premise of including a fixed cost component in the cost function. The AESO has therefore performed additional analysis of the 48-point combined data set used by the EUB, to determine the non-zero intercept power curve that would have the highest correlation coefficient…”


Request:

(a) Please provide academic or other literature as required to demonstrate that the derivation of a non-zero intercept power function is an acceptable practice.

(b) Please provide supporting data and analysis as required to demonstrate that the non-zero intercept power function derived by the AESO for its November 5, 2007 comments (i.e. Costs = 1,836,175 × (MW + 0.3655)^{0.4342}) has been appropriately derived consistent with the methodology described in the AESO’s response to a) above.

(c) In Schedule C-1 of an excel spreadsheet attachment to its November 5, 2007 comment submission, it appears that the AESO continues to propose the following POD cost function to be the basis for its proposed POD charge:

$947,000 + $621,000/MW for the first 7.5MW +
$154,000/MW for the next 42.5MW +
$47,000/MW for all MW greater than 50 MW

Please confirm that above described POD cost function has not been derived from the non-zero intercept power function discussed in part b) above.

(d) Please provide linear approximation POD cost functions for the non-zero intercept power function referenced in question b) above for the following 3 scenarios:

(i) Breakpoints of 7.5MW, and 50 MW
(ii) Breakpoints of 7.5 MW, 17 MW, and 50 MW
(iii) Breakpoint of 7.5 MW, 17MW and 40 MW

The AESO was subsequently requested to include an additional scenario:

(iv) Breakpoints of 17 MW and 50 MW
(e) Please comment on each of the 3 linear approximations developed in question d) above, assuming in each case that the alternative becomes the approved POD cost function for both DTS POD charge and contribution policy maximum investment level purposes.

Response:

(a) Regression analysis is generally defined as a technique used “to develop an equation to express the relationship between two variables and estimate the value of the dependent variable \( Y \) based on a selected value of the independent variable \( X \).” (Mason, Robert D., and Douglas A. Lind. *Statistical Techniques in Business and Economics*. 8th ed. Boston: Irwin, 1993. p. 498) The AESO understands that an equation of any form may be proposed to express the relationship between the variables. In fact, if an equation is being sought that best expresses the relationship between the variables, it would be inappropriate to constrain that equation to passing through zero (unless such a constraint is warranted by the relationship being investigated). The coefficient of determination \( (r^2) \) is one measurement that can then be used to describe how well the proposed equation accounts for the variation of the dependent variable \( Y \).

Regression analysis is also commonly referred to as “curve fitting” and is available as a feature of specialized mathematical software. For example, the MATLAB Curve Fitting Toolbox software from MathWorks, Inc. includes curve fitting models based on exponentials, Fourier series, Gaussian, polynomials, power series, rations, sum of sines, and Weibull distribution. Similar models are provided in the Mathematica software from Wolfram Research. More information on these software applications is available from their respective websites:


The two power series models included in the MATLAB Curve Fitting Toolbox are of the form \( y = a x^b \) (zero intercept) and \( y = a + b x^c \) (non-zero intercept). The form provided in the AESO’s Comments dated November 5, 2007, \( y = a (x + b)^c \), is a variant of the non-zero intercept power series. In addition, both the MATLAB and Mathematica software allow custom models of any equation to be entered by the user. Microsoft Excel provides only the most common subset of simple regression analysis equations, compared to the almost unlimited models that can be used in the specialized mathematical software mentioned above.

(b) As discussed in part (a), any equation can be proposed to express the relationship between two variables, and those equations may contain a variety of parameters. However, a primer titled “Curve Fitting Made Easy” (Ledvij, Marko. “Curve Fitting Made Easy." *The Industrial Physicist*. Apr-May 2003: 24-27.) notes the following:

The parameter values must also make sense from the standpoint of the model. If, for example, one of the parameters is temperature in kelvins and the value obtained from fitting is negative, we have obtained a physically meaningless result. This is true even if the curve seemingly does fit the data well. In this example, we have likely chosen an incorrect model.
As explained in the AESO’s November 5 comments, a zero intercept power curve suggests there should be no fixed cost component in the POD cost function. This does not align with the evidence on the record in this proceeding nor with the premise of including a fixed cost component in the cost function. The AESO therefore proposed a non-zero intercept power curve as a more meaningful representation of the data in consideration of the purpose for which the data was being analyzed. To allow examination of the regression statistics in Microsoft Excel, the AESO transformed the non-zero intercept power curve into a linear form by taking the natural logarithms of both sides of the equation. In other words:

\[ y = a \ (x + b)^c \] is equivalent to

\[ \ln(y) = c \ln(x + b) + k, \text{ where } a = e^k \]

The Microsoft Excel LINEST function was then applied to the logarithmic (linear) form of the equation to provide regression statistics such as the coefficient of determination \( r^2 \).

However, further examination of the Microsoft Excel analysis reveals that the regression statistics it returns for the logarithmic equation \( \ln(y) = c \ln(x + b) + k \) are not directly comparable to those it would calculate for the power equation \( y = a \ (x + b)^c \). The AESO has therefore repeated its analysis using the detailed calculation which Microsoft Excel’s help page for the LINEST function describes as follows:

In regression analysis, Microsoft Excel calculates for each point the squared difference between the y-value estimated for that point and its actual y-value. The sum of these squared differences is called the residual sum of squares, ssresid. Microsoft Excel then calculates the total sum of squares, sstotal. When const = TRUE, or omitted, the total sum of squares is the sum of the squared differences between the actual y-values and the average of the y-values.... Then regression sum of squares, ssreg, can be found from: ssreg = sstotal - ssresid. The smaller the residual sum of squares is, compared with the total sum of squares, the larger the value of the coefficient of determination, r2, which is an indicator of how well the equation resulting from the regression analysis explains the relationship among the variables. r2 equals ssreg/sstotal.

The AESO first notes that the detailed calculation provide a different coefficient of determination for the EUB-proposed zero intercept power curve than that provided by the Chart > Add Trendline… option when the data is plotted in Microsoft Excel. The detailed calculation is provided in attached Schedule BR.AESO-001 (b)-A, and results in a coefficient of determination of \( r^2 = 0.45388 \). (The AESO also notes that the detailed calculation provides the same coefficients of determination for the other curve forms as those provided by the Chart > Add Trendline… option.)

The Dual-Use Customers (DUC), in discussion with the AESO regarding its November 5 comments, provided the AESO with a non-zero intercept power curve which had a higher coefficient of determination when determined through detailed calculation. The AESO has confirmed that the following equation appears to provide the highest coefficient of determination for a non-zero intercept power curve:

\[ y = 2,693,924 \ (x - 0.005715)^{0.32916}, \ r^2 = 0.47886 \]
The coefficients of determination calculation for this non-zero intercept power curve is provided in attached Schedule BR.AESO-001 (b)-B. The AESO observes that this power curve cannot be used to calculate values for \( x \) less than 0.005715, and therefore does not provide a meaningful representation of the data. However, as this is the curve which has the highest coefficient of determination, the AESO utilized this equation in response to parts (d) and (e) of this information request.

(c) Confirmed. The AESO-proposed POD cost function was initially provided on page 78 of the AESO’s Written Argument in this proceeding, filed on June 21, 2007. The development of that POD cost function was summarized in that document, and did not utilize a non-zero intercept power curve. Please refer to the response to Information Request BR.AESO-002 (a) for additional information.

(d) The AESO considers that a reasonable methodology to develop a multi-segment line from a continuous curve would involve the following steps:

- calculate the y-values resulting from the equation for the curve for all x-values within each range being examined,
- use linear regression to determine a straight line segment to represent that portion of the curve,
- given the straight line segments are generally non-continuous, average the y-values at the breakpoint between two adjacent segments to enable the creation of a series of continuous straight line segments, and
- determine the slopes between the averaged breakpoint values for each portion of the curve.

The AESO has provided the calculations for such a methodology in attached Schedules BR.AESO-001 (d) (i), (ii), (iii), and (iv). For the x-values, the AESO used the DTS contract capacities for all current points of delivery as provided in Information Response BR. AESO-003 (a)-A2 Rev 2 dated February 27, 2007 (Exhibit 214). The calculations provide the following straight line segment POD cost functions based on the requested ranges:

(i) Costs = $1,863,015 + $509,090/MW for the first 7.5 MW + $105,648/MW for the next 42.5 MW + $35,752/MW for all remaining MW

(ii) Costs = $1,964,863 + $475,079/MW for the first 7.5 MW + $150,135/MW for the next 9.5 MW + $91,926/MW for the next 33 MW + $41,394/MW for all remaining MW

(iii) Costs = $1,964,863 + $475,079/MW for the first 7.5 M + $147,458/MW for the next 9.5 MW + $102,153/MW for the next 23 MW + $47,130/MW for all remaining MW
(iv) Costs = $2,684,512
    + $275,854/MW for the first 17 MW
    + $79,206/MW for the next 33 MW
    + $41,394/MW for all remaining MW

(e) Figure 001 (e)-1 below graphically compares the straight line segment POD cost functions provided in parts (d) (i), (ii), (iii), and (iv) (the red, green, blue, and purple lines in the chart, respectively) together with the continuous non-zero intercept power curve (the dashed line) and the cost function proposed by the AESO in its Argument (the black line).

The AESO considers that the first three cost functions from part (d) are very similar, and, based on the limited data from which they were developed, neither could be considered to be more representative of the continuous power curve than the others. The fourth cost function appears less representative of the power curve. The AESO’s cost function seems least representative of the power curve, which is reasonable as it was developed using a different methodology (as described in Information Response BR.AESO-002 (a)) that did not combine data from different data sets.

As discussed in part (a) above, a coefficient of determination can be calculated for an equation of any form. The AESO has calculated that the coefficient of determination ($r^2$) is approximately 0.49 for each of the first three cost functions based on the power curve. The coefficient of determination is approximately 0.47 for the fourth cost function based
on the power curve, and approximately 0.42 for the AESO-proposed cost function when assessed against the combined data set. The calculations are provided in attached Schedules BR.AESO-001 (e)-C, -D, -E, -F, and -G.

In consideration of the similarity of the first three cost functions based on the power curve, the AESO does not consider the use of three breakpoints to provide a function which better represents the power curve; breakpoints of 7.5 MW and 50 MW appear to provide a cost function that is adequately representative.

Furthermore, consistent with the quotation provided in part (b) above, the AESO suggests that the cost function which should ultimately be selected should be the one which makes the most sense based on the evidence in this proceeding. The AESO continues to believe its proposed cost function remains the most reasonable of any offered to date, but if one of the four cost functions from part (d) is required to be chosen, the AESO would consider the first, from part (d) (i) and based on 7.5 MW and 50 MW breakpoints, to be the most likely to have physical meaning.

The 7.5 MW breakpoint represents a physical quantity in that it appears to be the smallest size of service currently interconnecting to the transmission system, as discussed on page 73 of AESO Argument. Similarly, the 50 MW breakpoint represents a physical quantity in that it is the point above which service requirements frequently lead to the installation of a second transformer, as discussed on page 76 of AESO Argument. Since these breakpoints have physical relevance, the AESO suggests that they should guide the selection of a cost function in the absence of other distinguishing characteristics. As stated in point 5 of its November 5 comments, the AESO also suggests that where additional complexity is introduced by an additional breakpoint which does not substantially affect the POD charge, that additional complexity is not warranted.

The AESO has determined billing determinants and developed a DTS POD charge based on the cost functions provided in parts (d) (i), (ii), (iii), and (iv). The derivation is provided in attached Schedule BR.AESO-001 (e)-H. A comparison of the monthly POD charges resulting from the cost functions is provided in Figure 001 (e)-2 below. POD charges based on the cost functions provided in parts (d) (i), (ii), (iii), and (iv) are illustrated by red, green, blue, and purple lines in the chart, respectively, together with the POD charge resulting from the cost function proposed by the AESO in its Argument (the black line).

Similar to its views on the cost functions, the AESO considers that POD charges based on the first three cost functions from part (d) are very similar, and, based on the limited data from which they were developed, neither could be considered to be more representative of the continuous power curve than the others. As noted for the POD charge examined in its November 5 comments, the AESO observes that each of these POD charge alternatives increases bills for smaller customers, and decreases bills for larger customers, compared to the AESO-proposed POD charge (which already resulted in significant bill increases for smaller customers). The AESO recommends that considerations other than coefficient of determination, such as relevance to physical factors, simplicity, and bill impacts, should guide the choice of cost function on which the POD charge would be based.
Figure 001 (e)-2: Comparison of POD Charges From Cost Functions Based on Non-Zero Intercept Power Curve and AESO-Proposed POD Charge

Maximum investment functions based on the examined costs functions would appear similar in relation to each other as the POD charges, and have not been separately calculated.
Preamble: In its June 21, 2007 Argument Submission, the AESO utilized a logarithmic function \((y = 3.8486\ln(x) - 3.1694)\) in support of a revised interconnection project cost function. The logarithmic function was adopted by the AESO in preference to linear, polynomial, power and exponential functions.

The AESO’s November 5, 2007 comments appears to question the appropriateness of the Power function considered by Board staff because it intercepts the y-axis at zero and thus implies that there should be no fixed cost component in the POD cost function. In light of this observation, Board staff has the following questions related to the possible use of a logarithmic function as the basis for the POD cost function.

Reference: AESO Argument pp. 76-77.

Request:

(a) Please confirm that the logarithmic function proposed by the AESO in its July 21, 2007 Argument was utilized in preference to the other functional forms considered by the AESO primarily on the basis of having a higher R2 than the other functional forms derived through regression analysis for the AESO’s Argument. If this cannot be confirmed, please explain why the logarithmic function was adopted by the AESO for the purposes of its Argument.

(b) The AESO has expressed concern in its November 5, 2007 comments that the power function proposed by the Board staff may not be appropriate because of its zero-intercept function form. Does the AESO consider the logarithmic function put forth by the AESO in its Argument would be preferable to other functional forms in light of the fact that the AESO’s preferred function \((y = 3.8486\ln(x) - 3.1694)\) is associated with negative values at low values of \(x\) (i.e. DTS capacity)? Please explain.

(c) Please confirm that that logarithmic function derived from the 48 POD data set used to derive the power function used in support of the Board staff’s proposed POD cost function:

(i) has the following functional form:

\[ y = 1,890,801.87\ln(x) + 1,779,292.09 \]

(ii) has an R2 of 0.43

If the above cannot be confirmed, please provide the appropriate values for i) and ii) above and support all calculations.

(d) In light of the AESO’s concerns that Board staff’s proposed zero-intercept power function is not consistent with the existence of a fixed cost component in the POD cost function,
does the AESO consider that the logarithmic functional form described in part c) above to be preferable to the zero-intercept power function proposed by Board staff in spite of being associated with a lower R2 than zero-intercept power function? Please fully explain your response.

(e) Please provide linear approximation POD cost functions for the logarithmic function referenced in question c) above for the following 3 scenarios:

(i) Breakpoints of 7.5 MW, and 50 MW
(ii) Breakpoints of 7.5 MW, 17 MW, and 50 MW
(iii) Breakpoints of 7.5 MW, 17 MW and 40 MW

The AESO was subsequently requested to include an additional scenario:

(iv) Breakpoints of 17 MW and 50 MW

(f) Please comment on each of the 3 linear approximations developed in question e) above, assuming in each case that the alternative becomes the approved POD cost function for both DTS POD charge and contribution policy maximum investment level purposes.

Response:

(a) As explained in the AESO’s Argument, the AESO utilized a logarithmic function to determine only one segment of the AESO’s proposed cost function: the incremental cost for projects with large capacities. The AESO-proposed cost function was developed in three steps (summarized on pages 71-80 of the AESO’s Argument):

(i) A raw cost function was developed to appropriately represent the 30 data points for which the AESO had detailed costs and which ranged in size from 7.5 MW to 43.2 MW. This cost function was a single straight line cost function which exhibited a better coefficient of determination than a logarithmic cost function (AESO Argument, page 71) based on the same data.

(ii) For projects less than 7.5 MW, the AESO used a second set of additional data points and a minimum intercept methodology to determine a y-intercept, then interpolated a straight line function from the y-intercept to the 7.5 MW point on the straight line function determined in step (i) (AESO Argument, pages 74-75).

(iii) Finally, for projects greater than 50 MW, the AESO used a third set of additional data points and a logarithmic curve to determine an average slope for projects from 50 MW to 125 MW (AESO Argument, pages 75-77).

The AESO reiterates that a single straight line function appears to be the best fit for the best data on the record — that being the 30 greenfield data points. However, the AESO recognized that the data does not represent all possible values, and accepted that valid arguments were put forward for costs being below the average represented by this line for projects below 7.5 MW and above 50 MW. The AESO therefore utilized additional data and analysis to determine cost functions for those ranges, and used them to augment the primary cost function first determined for projects in the 7.5 MW to 43.2 MW size range.
The logarithmic function was used by the AESO only to determine the incremental cost for one segment of the cost function — that for capacities greater than 50 MW — and then only in the analysis of additional data from a consistent source, not as the basis for the entire cost function. The logarithmic function was utilized because its general shape aligned well with the concept of economies of scale. Additionally, its coefficient of determination was higher than any of the other curve types which were examined.

The AESO continues to be concerned about the appropriateness of determining a cost function by performing regression analysis on data which combines points from the greenfield data set and points from the older TFO data set, as discussed in its November 5 comments.

(b) The AESO has no “preference” for one type of curve over another, beyond the reasonable approach of the quotation provided in Information Response BR.AESO-001 (b), which suggests that the cost function should make sense based on the evidence in this proceeding.

The AESO agrees that the logarithmic function it used in the third step of its cost function development (summarized in part (a) above) provides an illogical negative cost for capacities of less than 3.3 MW. But the analysis in the third step was based on projects ranging from 7.6 MW to 122.8 MW (AESO Argument, page 76) and the logarithmic function does not provide negative costs over that range of capacities. It is therefore reasonable to accept a logarithmic function over that capacity range.

Please see part (d) below for additional comments.

(c) Confirmed. The form and coefficient of determination (r^2) values provided by the EUB are those returned by Microsoft Excel for a logarithmic regression curve on the 48-point data set. The same coefficient of determination also results from the more detailed calculation discussed in Information Response BR.AESO-001 (b).

(d) As stated in part (b) above, the cost function should make sense based on the evidence in this proceeding. The logarithmic function proposed in part (c) of the Information Request provides an illogical negative cost for capacities of less than 0.4 MW.

If a single continuous curve is to be used for the basis of the cost function over the entire range of potential capacities, then that curve must align with the evidence on the record in this proceeding and the premise (or model) for the POD cost function being considered. The AESO understands that all data on the record in this and prior proceedings supports a minimum cost — a positive y-intercept — associated with a point of delivery. The AESO therefore suggests that any single continuous curve used for the basis of the cost function, over the entire range of potential capacities, should similarly have a positive y-intercept.

The AESO would therefore not recommend the use of the proposed logarithmic function as the basis for the cost function over the entire range of potential DTS capacities.

(e) Using the methodology described in Information Response BR.AESO-001 (d), the AESO has calculated the following straight line segment POD cost functions based on the logarithmic equation provided in part (c) of the Information Request and the requested
ranges. Calculations are provided in attached Schedules BR.AESO-002 (e) (i), (ii), (iii), and (iv).

(i) Costs = $310,999
    + $919,964/MW for the first 7.5 MW
    + $70,592/MW for the next 42.5 MW
    + $12,385/MW for all remaining MW

(ii) Costs = $184,961
       + $877,875/MW for the first 7.5 MW
       + $91,414/MW for the next 9.5 MW
       + $63,748/MW for the next 33 MW
       + $19,080/MW for all remaining MW

(iii) Costs = $184,961
       + $877,875/MW for the first 7.5 MW
       + $88,637/MW for the next 9.5 MW
       + $73,527/MW for the next 23 MW
       + $23,834/MW for all remaining MW

(iv) Costs = $1,529,976
       + $395,310/MW for the first 17 MW
       + $33,968/MW for the next 33 MW
       + $19,080/MW for all remaining MW

(f) Figure 002 (f)-1 below graphically compares the straight line segment POD cost functions provided in parts (e) (i), (ii), (iii), and (iv) (the red, green, blue, and purple lines in the chart, respectively) together with the continuous non-zero intercept power curve (the dashed line) and the cost function proposed by the AESO in its Argument (the black line).

Similar to the comments provided in Information Response BR.AESO-001 (e), the AESO considers that the first three cost functions from part (e) are very similar, and, based on the limited data from which they were developed, neither could be considered to be more representative of the continuous power curve than the others. Furthermore, the AESO has calculated that the coefficient of determination ($r^2$) is within the range 0.45-0.47 for each of the first three cost functions from part (e). The coefficient of determination for the fourth cost function from part (e) is approximately 0.42, and for the AESO-proposed cost function is also approximately 0.42. The calculations are provided in attached Schedules BR.AESO-002 (f)-C, -D, -E, -F, and -G.

The AESO continues to suggest that where additional complexity is introduced by an additional breakpoint which does not substantially affect the POD charge, that additional complexity is not warranted.

The AESO has determined billing determinants and developed a DTS POD charge based on the cost functions provided in parts (e) (i), (ii), (iii), and (iv). The derivation is provided in attached Schedule BR.AESO-002 (f)-H. A comparison of the monthly POD charges resulting from the cost functions is provided in Figure 002 (f)-2 below. POD charges based on the cost functions provided in parts (e) (i), (ii), and (iii) are illustrated
Figure 002 (f)-1: Comparison of Cost Functions Based on Logarithmic Curve and AESO-Proposed Cost Function

Figure 002 (f)-2: Comparison of POD Charges From Cost Functions Based on Logarithmic Curve and AESO-Proposed Cost Function
by red, green, and blue lines in the chart, respectively, together with the POD charge resulting from the cost function proposed by the AESO in its Argument (the black line).

The AESO continues to recommend that considerations other than coefficient of determination, such as relevance to physical factors, simplicity, and bill impacts, should guide the choice of cost function on which the POD charge would be based.

Maximum investment functions based on the examined costs functions would appear similar in relation to each other as the POD charges, and have not been separately calculated.