Chapter 4

An Automated Algorithm for Decline Analysis

Rita Aggarwala[1], Tom Alberts[2], Chris Bose[3], Adrian Driga[1], Aude Espesset[4], John Harlim[5], Jihyoun Jeon[6], Seungwon Jeon[7], Qutaibeh Katatbeh[8], Larry Kolasa[9], Melvin Leok[10], Mufeed Mahmoud[4], Rafael Meza[6], Alberto Nettel-Aguirre[1], Christina Popescu[2], Randall Pyke[9], Mariana Carrasco Teja[6],
Report prepared by Randall Pyke

4.1 Introduction

Oil and gas wells are regularly monitored for their production rates. Typically, daily production rate data is available, expressed in millions of standard cubic feet per day (MMscf/d) for natural gas wells or barrels per day (Bbl/d) for liquids, i.e., oil or water. This data reflects changing physical conditions within the oil or gas reservoir, changes in equipment (eg. failure or maintenance), activity of surrounding wells, variability in outshipping methods, and changing production rates due to economic factors. As a result, typical production rate data is noisy and highly discontinuous.

Decline analysis is a process that extrapolates trends in the production rate data from oil and gas wells to forecast future production and ultimate cumulative reserve recovery. Current software often attempts a best fit approach through all the data, but the result is erroneous in the majority of cases. A human operator with an understanding of the factors that affect the

[1]University of Alberta
[2]University of Calgary
[3]University of Victoria
[4]University of Western Ontario
[5]University of Guelph
[6]University of Washington
[7]University of Texas at Austin
[8]Concordia University
[9]Ryerson University
[10]California Institute of Technology

behavior of oil and gas wells can do a much better job of forecasting appropriately; however, it is a time-consuming process.

The goal is to find an algorithm that can be easily interfaced with standard industrial software and that incorporates some of the criteria used in the human analysis so as to perform acceptable forecasts in the majority of cases.

4.2 Proposed Solution

The proposed solution consists of three main steps: (1) Segmentation of Data, (2) Curve fitting, and (3) a Decision Process. Segmentation of Data attempts to identify intervals in the data where a single trend is dominant. A curve from an appropriate family of functions is then fitted to this interval of data. The Decision Process gauges the quality of the trends identified and either formulates a final answer or, if the program cannot come to a reliable answer, 'flags' the well to be looked at by an operator.

4.2.1 Segmentation

The input data is assumed to consist of a time series, $\{U(i)\}_{i=1}^{N}$, where $U(i)$ represents the data point (rate of oil/gas flow from the well) at time i. We assume the list is contiguous, i.e., there is a data point for each time step (the length of a time step is input by the operator). In this stage of the analysis the data set is divided into segments. Each segment will be analyzed in subsequent stages.

The end points of the segments are determined by (1) discontinuities in the data, and (2) discontinuities in the slope of the mean data (changes in trend). Two methods were developed for detecting these types of discontinuities; one method for detecting type (1) discontinuities, and one method for detecting type (2) discontinuities.

Because the data is typically very noisy, the data is smoothed a number of times. This smoothing reduces the amplitude of the oscillations of the noise relative to the amplitudes of the discontinuities in the data, making the discontinuities easier to identify.

Each smoothing operation is obtained by moving averages with a window of width three. Let $\{U_{\mathrm{S}}^{l-1}(i)\}_{i=1}^{N}$, $l \geq 1$, denote the data after being smoothed $l-1$ times. Then the next smoothed version of the data, $\{U_{\mathrm{S}}^{l}(i)\}_{i=1}^{N}$, is obtained by the formula;

$$U_{\mathrm{S}}^{l}(i) \;=\; \frac{1}{3}\Big(U_{\mathrm{S}}^{l-1}(i-1) + U_{\mathrm{S}}^{l-1}(i) + U_{\mathrm{S}}^{l-1}(i+1)\Big).$$

The number of times the data is smoothed depends on the length N of the data set. We found from experience that an appropriate value for the number of times k the data should be smoothed is $k = \lfloor \log_2(N) \rfloor$. (From now on, "smoothed data" will mean $\{U_{\mathrm{S}}^{k}(i)\}_{i=1}^{N}$.)

After the data is smoothed a data set of differences $\{U_{\mathrm{d}}(i)\}_{i=1}^{N}$ is produced, where

$$U_{\mathrm{d}}(i) = |U(i) - U_{\mathrm{S}}^{k}(i)|.$$

That is, the differences are the (absolute value of) oscillations of the data around the smoothed data. The method for detecting type (1) discontinuities analyzes these differences.

The type (1) discontinuities in the data show up as larger peaks in the differences. However, because the data is typically very noisy, and to facilitate the identification of the locations of peaks (as described below), the differences are also smoothed a number of times, precisely, $k/2$ times (where k is as above), and in the same manner as the data was smoothed, i.e., by moving averages with a window of width three. This has the effect of suppressing oscillations in the differences $\{U_{\mathrm{d}}(i)\}_{i=1}^{N}$ that are due to noise in the data, and thereby enhancing the peaks in the differences that are due to discontinuities in the data.

The discontinuities in the data are identified as the largest peaks in the (smoothed) differences. This is accomplished by locating peaks in the differences that are above some threshold. The threshold is set at 15% of the mean of the (original) data. That is, a peak in $\{U_{\mathrm{d}}(i)\}_{i=1}^{N}$ is identified as a point of discontinuity in the data if the amplitude of the peak is greater or equal to 15% of the mean value of the data.

The locations of the peaks are found by identifying local maxima in the differences that are above the threshold. Here we look for points where the (finite-difference) derivative of the differences change from positive to negative. That is, if $D(i) = U_{\mathrm{d}}(i+1) - U_{\mathrm{d}}(i)$, then if $D(i) > 0$ and $D(i+1) < 0$ we record i as a point of discontinuity in the data (provided that $U_{\mathrm{d}}(i)$ is above the threshold).

Now we look for discontinuities in the slope of the (mean) data (i.e., the type (2) discontinuities). For each interval obtained above (from looking for type (1) discontinuities), a value for the derivative of the 'mean' data at that point is obtained by computing the derivative of the best-fit parabola at that point. The parabola is fitted only to the data within 15 points on either side of the point under consideration. This gives a time series of 'mean' derivative values of the data in the interval. This time series is fed into the program that finds type (1) discontinuities. The output is a list of points where the slope of the 'mean' data has a discontinuity.

The locations of the two types of discontinuities are combined into one list - the list of end points of intervals for the data. This list is passed on to the next stage in the analysis.

Wavelet analysis

Another approach to determining the segmentation of the data uses wavelets (see for example [1]). Wavelets are functions that cut up the data into different temporal and frequency components, and then study each component with a resolution matched to its scale. They are highly useful in analyzing physical situations where the signal is discontinuous. In this work, given the noisy data, wavelets are used to de-noise the data and to divide time series into segments.

The technique works in the following way. When you decompose a data set using wavelets, you use filters that act as averaging filters and others that produce details. Some of the resulting wavelet coefficients correspond to details in the data set. If the details are small, they might be omitted without substantially affecting the main features of the data set. The idea of thresholding, then, is to set to zero all coefficients that are less than a particular threshold. We then use an inverse wavelet transformation to reconstruct the data set. The de-noising is carried out without smoothing out the sharp structures. The result is a cleaned-up signal that still shows important details. In our case we used a Haar Wavelet base function to de-noise the noisy decline data and also to determine the changes in the production dynamics.

4.2.2 Curve Fitting and EUR

The previous stage in the analysis produced a list $\{a_1, \ldots, a_m\}$ of end points of intervals $I_1 = (1, a_1)$, $I_2 = (a_1, a_2), \ldots,$ $I_{m+1} = (a_m, N)$ for the data $\{U(i)\}_{i=1}^{N}$. The second stage in the analysis performs a least squares curve fitting to each of the intervals. That is, curve fitting is applied to the time series
$\{U(i)\}_{i=1}^{a_1}$, $\{U(i)\}_{i=a_1}^{a_2}$, $\ldots$, $\{U(i)\}_{i=a_m}^{N}$. The actual intervals of data used in the curve fitting are slightly smaller than these (to remove transient effects).

The class of curves used in the fitting belong to the family

$$q(t) \;=\; q_o(1 + nDt)^{-1/n} \tag{4.1}$$

where q_o, n and D are parameters. It has long been accepted within the petroleum industry that this function accurately models the uninterrupted flow rate of a well, and it can be also derived from basic physical principles.

Once we have found the best fit for the kth interval I_k, we compute a number of summary statistics;

- $q_{o,k},\ n_k,\ D_k$; least squares estimates of q_o, n and D, and
- S_k^2 = normalized variance of the data over the interval = $\sum_{i \in I_k} |U(i) - q_k(i)|^2 / l_k$ where l_k is the length of I_k and $q_k(t)$ is the curve fit for the kth interval.

The most important of these statistics is EUR_k, the Estimated Ultimate Recovery based on the best fit curve of the kth interval. It is determined by solving for the time at which the best fit curve passes below a minimum threshold rate, call this time T_k, and then computing the sum;

$$\mathrm{EUR}_k \;=\; \sum_{1}^{a_k - 1} U(i) \;\;+\;\; \sum_{i=a_k}^{T_k} q_k(i) \tag{4.2}$$

where the first sum simply represents the volume of oil or gas that has already been produced, and the second is the amount we expect to produce based on the curve fit $q_k(t)$ for the kth interval. These statistics are then passed as parameters to a weight function which decides their relevance and usefulness, and based on this we can calculate a final EUR value.

4.2.3 Decision Process, Final Estimated Ultimate Recovery (EUR), and Overall Rate Curve

Given the parameters $p_k = (n_k, q_{o,k}, D_k, \mathrm{EUR}_k, S_k^2, l_k)$, $k = 1, ..., m+1$ for the fits over the intervals $I_1, ..., I_{m+1}$, we choose a weight function $w_k = w_k(p_k)$ that indicates how 'reliable' the EUR_k is. The following conditions may be considered as rules for a reliable EUR_k;

- I_k is a long data set,
- the variance S_k^2 over I_k is small,
- $0 \leq n \leq 0.5$ (a physically plausible hyperbola), and
- I_k is a recent interval.

The weight reflects the importance of the hyperbolic curve fit over interval k in determining the final EUR. The precise formula incorporating the factors above was chosen to be;

$$w_k = \frac{\frac{(l_k - r)^+ h(n_k)}{S_k^2 (N - a_k)}}{\sum_{i=1}^{m+1} \frac{(l_i - r)^+ h(n_i)}{S_i^2 (N - a_i)}}. \tag{4.3}$$

Here, as above, l_k is the length of interval k, $m+1$ is the total number of intervals obtained from the segmentation stage of the analysis, r is an integer such that intervals of length less than or equal to r will not be used, and $(l_k - r)^+$ denotes the positive part of $l_k - r$ (i.e., $(l_k - r)^+ = l_k - r$ if $l_k - r \geq 0$, and $(l_k - r)^+ = 0$ if $l_k - r < 0$). n_k is the estimate of the model parameter n for interval k (cf. Equation (1.1)), and $h(n)$ is a function which gives the parameter n from the hyperbolic fits a separate weighting, for example,

$$\begin{aligned} h(n) &= 1 \text{ if } n \in [0,\, 0.5] \\ &\quad 2(1-n) \text{ if } n \in (0.5,\, 1] \\ &\quad 0 \text{ if } n > 1. \end{aligned}$$

(Engineering knowledge indicates that low values of the model parameter n indicate a well which is past its period of transient activity.) A similar function could instead be applied. S_k^2 is the normalized sum of squared residuals between the curve q_k and the data in the interval I_k (see above).

Once the weights have been determined, the EUR is calculated with the formula,

$$\widehat{\text{EUR}} = \sum_{k=1}^{m+1} w_k \times \text{EUR}_k \tag{4.4}$$

where EUR_k is the EUR which would be predicted using the curve fitted from interval k (cf. Equation (1.2)). The following provisions must also be implemented:

- If EUR_k is less than the amount of oil which has already been recovered, w_k is set to zero.
- If EUR_k is calculated to be infinite, w_k is set to zero.

Finally, we would like to be able to predict the amount of production between the end of the observed data and any future time t using a single, 'overall' hyperbolic curve of the form given in Equation (1.1). Engineering practice allows us to use for the estimate of the parameter n of this 'overall' curve, the convex combination of the n_k, namely,

$$\hat{n} = \sum_{k=1}^{m+1} w_k \, n_k.$$

This, along with the following two equations, will give us the parameters of a single hyperbolic curve.

For the final time point N,

$$\hat{q}_0(1+\hat{n}\hat{D}N)^{-1/\hat{n}} = q_N, \tag{4.5}$$

(cf. Equation (1.1)), the observed rate at time N. Alternatively, we could use the observed rate at the last acceptable time point used for curve fitting, or an average of the last few acceptable observed rates for the right hand side of the equation.

For the $\widehat{\text{EUR}}$,

$$\hat{q}_0\left(1+\frac{\hat{D}(\hat{n}-1)\widehat{EUR}}{\hat{q}_0}\right)^{-\frac{1}{\hat{n}-1}} = 1. \tag{4.6}$$

This is simply stating that at the time of $\widehat{\text{EUR}}$, we will be pumping at a rate of 1 barrel per day. Some other acceptable cut-off value could be used instead. Note that this equation is obtained by writing the rate function in terms of the cumulative production at time t. This transformation is easily obtained by integrating the original rate function considered, $q_t = q_0(1+nDt)^{-1/n}$ (with respect to t), to come up with a cumulative function, solving for time and substituting back into the original rate function again.

Equations (1.5) and (1.6) can be reduced to a single variable problem by solving Equation (1.5) for $\hat{q}_0$ and then substituting into Equation (1.6). Then, only $\hat{D}$ would need to be solved for numerically.

When all three parameters in the final hyperbolic curve are estimated, we can make forward predictions for rate and cumulative production at future time points.

4.3 Discussion

A more thorough testing of the algorithm presented here would include;

- More sophisticated techniques such as wavelets or neural networks (see below) could be used in the segmentation stage if the present method (of moving averages) turns out to be unreliable.

- An examination of the uncertainties in the curve fits and the final predictions made.
- Fine tuning the weighting functions w_k.
- Verification that the parameters for the final ('overall') hyperbolic curve are realistic.
- A more robust approach to the estimates obtained in stage 3; initial tests show that Equations (1.5) and (1.6) may be unstable.

One may consider using a neural network in the analysis (see for example [2]). Given sufficiently many data sets it may be sensible to avoid the choice of a heuristic weighting scheme such as given by Equation (1.3) in favour of a neural net approach where we would allow the algorithm to find good weighting schemes through training. Furthermore, it may be possible to invoke some training components in the segmentation algorithm as well. More specifically, the choice of the important thresholding parameters in the smoothing algorithm seems to be a likely candidate for a neural net approach.

Bibliography

[1] Daubechies, Ingrid, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[2] de Wilde, Philippe, Neural Network Models, Springer-Verlag, New York, NY, 1997.