"HENRI COANDA"
AIR FORCE ACADEMY
ROMANIA

GERMANY

"GENERAL M.R. STEFANIK"
ARMED FORCES ACADEMY
SLOVAK REPUBLIC

INTERNATIONAL CONFERENCE of SCIENTIFIC PAPER
AFASES 2011
Brasov, 26-28 May 2011

THE APPRAISAL OF OIL PRODUCTION USING THE METHOD OF REGRESSION

Loredana-Maria PĂUNESCU

Department for Economical Mathematics and Informatics, Petroleum&Gas University,
Blvd. Bucuresti 39, Ploiesti, Romania
loredana.paunescu@yahoo.com

Abstract: *The present work presents the identification of the temporal series character of the dataset of production after an analysis of these recorded and stored data in the oil scaffolds from Romania.The exploitation development of these data regarding the production, stored in a database have the main purpose the appraisal of the moment of deposits' economic production.*

Mathematics Subject Classifications 2010: *62P30, 65Y04, 68P15, 68U35, 74G15, 93A30, 97R50.*
Keywords: *database, deposit, production, scaffold.*

1. INTRODUCTION

Data warehouse technology is a new way to use large data collections, collections that can accumulate in the activity of a firm over a longer period of time.

Formed initially as a technique for obtaining economic information useful in conducting market analysis, marketing and business development forecasts of different markets and in different contexts, data warehousing technology with other technologies developed specific (OLAP, Data Minig).

Today there are tendencies to introduce this technique and analysis of data obtained in industrial processes.

2. PRODUCTION DATA PROCESSING

The activity of oil extraction and/or natural gas is a data warehouse using their natural application because the process itself involves regular collection and storage of data which can be used not only for understanding the status of a probe at a given time but also to predict its growth and to prevent further failures or problems of maintenance.

The advantages of such an approach are given strong integration of data from points located, in terms of geographic location, distance and manually processed to obtain the additional information from the information production can be a cumbersome process and, practically imposible.

As a solution for production data analysis, we had chosen a long time, so tend to judge the well/reservoir and can be expected to achieve economic production or when production ceases.

To do this, first it was necessary to develop a data mining methods of production are stored in the data store in order to estimate the time of deposit to achieve economic production.

This method is based on the following elements: a reservoir quickly reach the point of maximum production, and most of the deposit operation is achieved by pumping the deep, while the deposit is the contraction of production; production variance analysis over a long period of time may give an indication of the trend of evolution of the probe or deposit at that time, a trend that can be extrapolated to estimate when production ceases or the probe of the probe to achieve economic production.

Of course, these estimates are much closer to reality as the period under review is higher and the current work such estimates must be corrected periodically to take into account the analysis of all factors of influence on production.

Production data in the data store have taken the character of time-series affected by "noise" because it has a sensor that produces periods and periods when it is stopped for various reasons: accidents, instrumentation, interventions, etc.

The data stored over a relatively large periods of time (5 10 years or more) they will contain information on developments in the probe and the reservoir over time.

It is known that any oil deposit in full production for a limited period of time, followed by a decline in production becoming more intensive extraction of oil from the deposit until it becomes uneconomic.

This time corresponds to achieve final recuperare. Factorul final factor determining recovery of the deposit based on estimated production, amounting to not always correspond with reality because they made assertions based on the deposit that are not verified.

Based on these aspects, we intend to develop a model for production data processing to extract new information.

If you represent a production wells over time, it will have on the interval considered as a cloud of points that may suggest a particular function of distribution (fig.1.).

Because we do not know a priori law of development of production while (size may depend on many factors), we try to apply polynomial regression.

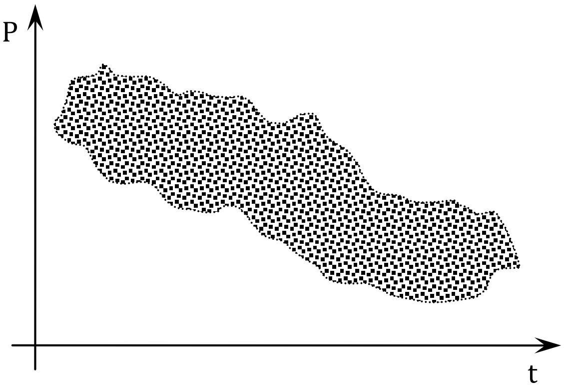

Fig.1. Distribution function for the time evolution of the production.

It will approximate the cloud of points with three regression functions: linear, quadratic, respectively, cubic.

Whether in the form of daily production P=F(t), expressed as discrete pairs (Pi, ti).

We consider three possible functions for the approximation, namely:

$$F_1 = a_1 \cdot t + b_1 \quad (1)$$

$$F_2 = a_2 \cdot t^2 + b_2 \cdot t + c_2 \quad (2)$$

$$F_3 = a_3 \cdot t^3 + b_3 \cdot t^2 + c_3 \cdot t + d_3 \quad (3)$$

those coefficients which are determined by the method of least squares.

Knowing the three functions for a probe, we calculate the correlation coefficient between the known values and the values *Pi* and the values *Fk (ti)* and is chosen as a model for defining such a function to probe the correlation coefficient is maximum.

$$r = \frac{n \cdot \sum_{i=1}^{n} F_i^k P_i - \left(\sum F_i^k\right) \cdot \left(\sum P_i\right)}{\sqrt{\left[\sum F_i^k - \left(\sum F_i^k\right)^2\right] \cdot \left[\sum P_i^2 - \left(\sum P_i\right)^2\right]}} \quad (4)$$

where : F_i^k - is the production in the moment ***i***, estimated with the regression function *k* (k=1,2 sau 3) P_i – is the production realized by the derrick in the moment ***i***.

Knowing the function F characterizing the production of crude oil for a longer period of time we can make predictions on how future development of the probe (Fig. 2).

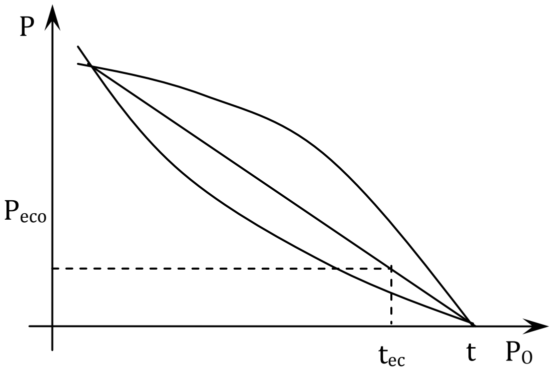

Fig. 2. The trend of development of deposit.

Extrapolating outside the range of known function will initially be able to estimate when the economic threshold is reached by operating the probe (Pec) or the point where he can no longer extract oil from the well (P0), regardless of the funds invested in it. If we look where F (t) = Pec, we obtain $t_{ec} = \frac{P_{ec} - b_1}{a_1}$ for the case with a linear function approximation. For the case $F(t) = 0$ we obtain $t_0 = -\frac{b_1}{a_1}$ all valid relation to the use of linear regression function. If we analyze the case of an aproximation parabolic function for the equasion $F(t) = P_{ec}$

we obtain :

$$t_{ec} = \min\left(\frac{-b_2 \pm \sqrt{b_2^{\,2} - 4 \cdot a_2 \cdot (c_2 - P_{ec})}}{2 \cdot a_2}\right)$$

and for the case
$F(t) = 0$
we obtain :

$$t_0 = \min\left(\frac{-b_2 \pm \sqrt{b_2^2 - 4 \cdot a_2 \cdot c_2}}{2 \cdot a_2}\right),$$

where: P_{ec} - represents the economic output of the probe, that threshold level of production that ensure production and delivery costs;

t_{ec} –is the appropriate time is to achieve economic production. Since the elapsed time measurement begins at a time considered time zero, teak is the amount estimated to achieve economic production level;

t_0 - is the estimated time until all of the probe, ie until the probe does not allow any costs associated with oil extraction operation.

That prediction, made for a probe, has its importance, but is not relevant to assess the deposit progress.For the whole trend of the deposit, in terms of production, production will have to examine all of the deposit period under review. Practically, considering that the deposit is exploited through the hole and I have records of production of these wells M in the period analyzed, we generate a cumulative time series.

$$P_i = \sum_{j=1}^{M} P_j^i \qquad (5)$$

where, P_j^i j is the probe output for the period *i* (in this case the period and may be a day or a longer period of time: one month or one year). Subsequently, these data are processed in a similar way, by determining a function F(t) to approximate the time evolution of the deposit (Fig. 3).

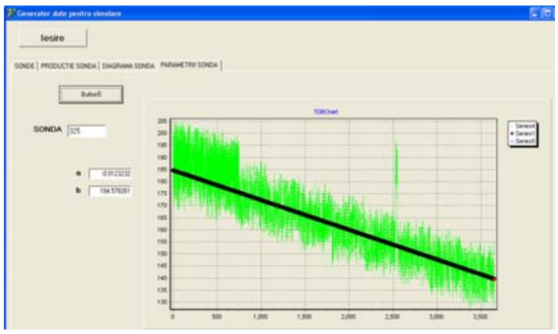

Fig.3. Evolution in time of the deposit.

The proposed analytical method is valid for wells and reservoirs in exploiting deep-sea pumping system as a reservoir from which most wells produce free rash has not entered the phase of decline.

In this case, the daily production wells depends on technical parameters and economic factors affecting production (maintenance of high oil prices may lead to limitation of production, even if the deposit could produce more).

To check the accuracy of the method noted above, a simulation was done because there were actual production data for a long enough period.

Thus, the corresponding data structure data warehouse was implemented using MS Access DBMS's, and to generate test data and analysis was done using an application development environment of Delphi.

MS Access DBMS's choice was dictated by the necessity of obtaining an application to run on machines with average performance.

Initially we tried using the Oracle DBMS, which, however, for the purposes of this work, has proved ineffective because they require special high-performance hardware to function properly in terms of speed of response to user needs.

It was taken account of the fact that in this work to implement a data warehouse itself but is developed and tested data structures characteristic of a data warehouse for demurrage and specific data processing methods.

Simulation was done by generating fictitious production database for more probes, data generated in such a way as to simulate a linear distribution, a parabolic, ie a cube.

Production data were generated as random data that falls within a given range of values.

Specified range of values was determined by analyzing the production values obtained at OMV Petrom, Buzau branch for small intervals of time (three months).

Algorithms for generating test data are presented in the sequence of Pascal code below:

```
{procedure for generating test data according to a linear variation}
procedure TForm1.Button3Click(Sender: TObject);
var sonda,an,luna,zi,zile,cod:integer;
    Pbrutmax,Pbrutmin,Pbrut,Pnet,Apa:real;
    Rmax,Rmin:integer;
    data,s:string;
begin
  val(edit1.Text,sonda,cod);
  val(edit2.Text,Pbrutmax,cod);
  val(edit3.Text,Pbrutmin,cod);
  val(edit4.Text,Rmax,cod);
  val(edit5.Text,Rmin,cod);

ADOTable2.TableName:='PRODSONDA_'+edit1.Text; ADOTable2.Active:=true;
  Randomize;
  for an:=2000 to 2009 do
   begin
     for luna:=1 to 12 do
       begin
         case luna of
           1:zile:=31;
           2:zile:=28;
           3:zile:=31;
           4:zile:=30;
           5:zile:=31;
         ..............................
.................................................
end;
```

Interface application using the above procedures is shown in Figure4.

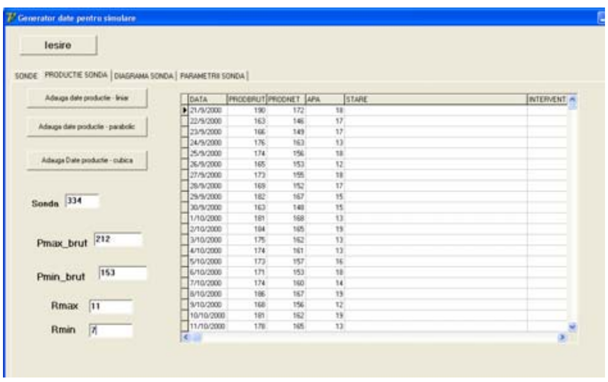

Fig. 4. Application interface for generating test data.

"HENRI COANDA"
AIR FORCE ACADEMY
ROMANIA

GERMANY

"GENERAL M.R. STEFANIK"
ARMED FORCES ACADEMY
SLOVAK REPUBLIC

INTERNATIONAL CONFERENCE of SCIENTIFIC PAPER
AFASES 2011
Brasov, 26-28 May 2011

To generate test data for a probe should be specified as input the maximum gross production, ie respectively the minimum and maximum oil and minimum water ration.

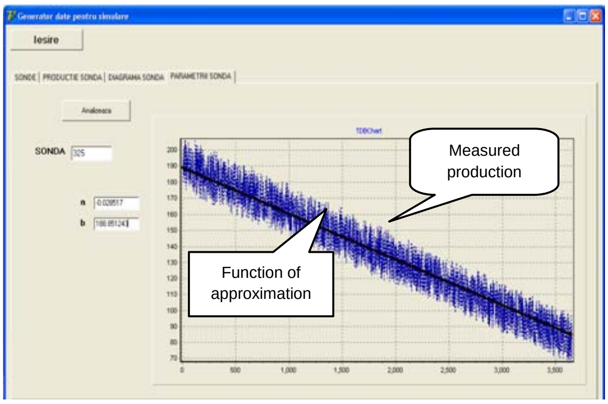

Fig. 5. The result of data analysis.

Fig.5 presents the result of production of a probe analysis using linear regression.

We note that production data distribution leading to a linear function approximation.

Obviously, the results are only a small part of what can be achieved by processing the data, but the workload involved exceeds the possibilities of a single person.

3. CONCLUSIONS & ACKNOWLEDGMENT

In conclusion, I believe that the results represent only a small part of what can be achieved by studying the problem of applying data warehouse technology in the oil industry.

Analysis of technical data can be correlated with economic analysis thus yielding new information on the efficient operation, so a certain wells and the entire deposit, or other information of interest for the decision management in the institution that exploits the deposit.

REFERENCES

1. Cristopher Adamson . *Mastering Data Warehouse Aggregates. Solutions for Star Schema Performance* , Wiley Publishing, Inc., Indinanapolis, 2006, pp. 90-124.
2. Ralph Kimball, Margy Ross. *The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling*, Second Edition, Wiley Publishing Inc., Indianapolis, 2002, pp.113-224.
3. Inmon, W.,H.. *Building the Data Warehouse*, Fourth Edition, Wiley Publishing Inc., Indianapolis, 2005,pp.123-235.
4. C. Di Mauro, J.P. Nordvik, and A.C. Lucia. *Multi-criteria decision support system and Data Warehouse for designing and monitoring sustainable industrial strategies - an Italian case study*, ISESS - International Symposium on Environmental Software Systems. www.iemss.org/iemss2002/proceedings/pdf/volume%20uno/434_dimauro.pdf
5. T. W. Liao, J.-H. Chen, E. Triantaphyllou, *Data Mining applications in Industrial Engineering: A Perspective*, Proceedings of the 25th International Conference on Computers and Industrial Engineering – New Orleans, Los Angeles, March, 1999.
6. Sebastian C. Brandt, Marcus Schluter, Matthias Jarke, A Process Data Warehouse for Tracing and Reuse of Engineering Desing Processes, The Second International Conference on Innovations in Information Technology (IIT '05).
7. Radu Platon, Mouloud Amazonz, *Application of data mining techniques for industrial process optimization*, Conmet Energy Canada, http:\\conmetenergy-conmetenergie.nrcon-rncon.go.ro.ca/industrial-

processes/industrial-systems-optimization/pulp_paper/publications/2007141.html
8. Clive Duebel, *Application of Data Mining Techniques to Industrial Processes to Improve Business Performance*, www.knowledgeprocesssoftware.com
9. ***, DataMinig, www.dmreview.com
10. ***, www.biblioteca.ase.ro/downres.php?tc=8621