A METHOD OF DATA MINING FOR SUSTAINABLE HYDROCARBON EXPLORATION

Muhammad Shaheen*, Muhammad Shahbaz*, Zahoor ur Rehman* and Sajid Mehmood*

ABSTRACT

Sustainability needs balance in ecology, economy and social justice. The world in this century is badly suffered by unsustainable development causing recession crisis, manpower crisis etc. Energy is vital for economy of a country so it has sound impact on sustainable development. In any country energy is explored to meet consumption requirements and it is consumed in accordance with available reserves. In Pakistan, hydrocarbons are considered to be one of the major sources of energy.

The contribution of hydrocarbon to overall development will be considered sustainable only if the hydrocarbon exploration rate will conform to its consumption rate. To ensure sustainable energy development, sustainability indicators had been agreed by world experts in a summit held in South Africa in 2001. These indicators will have different values for different countries depending upon the economy, energy reserves, living standards and social conditions. Depending upon these values, a methodology is proposed in this paper to find optimal hydrocarbon exploration rate that will ensure sustainable energy development in any country. The sustainability indicators (SI) are classified into exploration SI and consumption SI. The indicators which have strong correlation with hydrocarbon exploration rate are considered to be exploration SI while the indicators which have strong correlation with consumption rate are considered to be consumption SI.

The correlation values are calculated by statistical correlation analysis. After finding exploration SI, multiple regression is used to predict optimal hydrocarbon exploration rate which will ensure sustainable energy development within the country.

Key words: *Energy Data Mining, Sustainability, Hydrocarbon Exploration, multiple regression, Correlation*

INTRODUCTION

The most accepted definition of sustainable development in the Bruntland report is “The development that meets the needs of present without compromising the ability of future generations to meet their own needs[1]. Hydrocarbons are major sources of energy in this century all over the world. Sustainability in energy sector means a balance in sociology, economy and environment by establishing conformity between the production and consumption. The consumption and production of hydrocarbons seems natural but can controlled by strategic planning. The reserves of hydrocarbons are a gift of nature and the increase in its usage seems uncontrollable now a days. The only thing which human can do is to plan for establishing proportionality between exploration and use of hydrocarbons in long term. For example if the production of natural gas in a region is 4.3 million cubic feet per day and its consumption is 2 million cubic feet per day, the planners can propose figures for growth in usage. The planners have to plan for increase in usage of energy at one side for prosperity of nation and economical growth. On the other hand they are responsible for elongating life expectancy of fossil fuel reserves to ensure consistency in prosperity and ecological sanitation. Such contradicting scenarios demand for superfluous strategic capabilities. An agreement can only be reached when energy related figures shall be made explicit to the planners. The terrific gift to the planner will be the mechanism for predicting the optimal production and consumption rate of hydrocarbon. The production and consumption of energy are usually managed by different organizations within a country. The sharing of authentic information and adopting common procedures is a key to utilize collaborative intelligence.

* *Department of Computer Science & Engineering, University of Engineering & Technology, Lahore*

The evaluation of sustainable development and classification of world countries with respect to sustainability can be supported by sustainability indicators. In 2002, at the world summit on sustainable development, the indicators for sustainable development are proposed which provide better insight into the factors that affect energy, environment, economics and social well-being. The same indicators can also be used to evaluate the past policies and help setting targets for the future[2].

Development of advanced data analysis techniques has increased the opportunities for extracting useful knowledge from data. Data mining, an emerging technology with significant advances in the type of analytical tools can contribute with its potentials to guide or at least assist the opportunist in the prediction of an optimal energy exploration rate that can better conform to the consumption. The prediction about hydrocarbon exploration rate obviously means to explore new reserves of hydrocarbon beneath earth's surface. Optimal hydrocarbon exploration rate would mean the required rate of exploration in the presence of unsustainable factors. The source to collect those factors is provided in the form of "Indicators for Sustainable Energy Development".

The rest of the paper is organized in a classical way. In Section 2, the background knowledge of hydrocarbon consumption/exploration and energy indicators for sustainable development is discussed. The utilization of advanced data analysis technologies and data mining is included in the same section. Research problem is stated and its proposed methodology is presented in Section 3 with a block diagram. The results of the experiment are structured in Section 4. The discussion on results is also included in the same section. The paper is winded up with conclusions and future work in section 5.

Energy is central to sustainable development by satisfactorily balancing all the three aspects i.e. social, ecological and economical. The meaning of sustainable development in hydrocarbon (specifically oil and gas) industry will differ in the sense that the production of oil and gas will contribute towards sustainable development but it is not all about. The industry will have to ensure the provision of energy services by least affecting the environment and ensuring social justice with current as well as future generations[3]. The concept of sustainable development in its acceptable form was presented by Bruntland Commission in 1987[1]. The sustainable management of energy is dependent upon three conclusive factors. (1). Exploration of hydrocarbon reserves (2). Consumption of hydrocarbon (3). Defining mechanism to ensure balance between consumption and production. The consumption of hydrocarbon is affected by increase in population at the significant side and advancements of technology. The world population is increased by 3 billion in 1950 to 6 billion today and expected to reach 9 billion in 2050[4]. The next hundred years will cause an increase in the population by more than three times. This increase in population raised the consumption exponentially. Pakistan is a developing country in South Asia. The annual consumption of natural gas in Pakistan was 231,023 MMCFD (Million cubic feet per day) in 1998 and it has been raised to 588,653 Million cubic feet per day in 2008 which is 154% of the consumption in 1998[5]. If in the same country, the consumption is compared to production, the figures are disappointing whereas Pakistan is self sufficient in natural gas. 265 exploratory and 282 development wells were drilled in the same country during the period and 82 of them resulted in a success[6]. Lyn Arscot identified three big challenges to oil and gas industry which are pollution, global climate change and biodiversity[3]. The paper considered "Conformity in consumption and exploration rate of hydrocarbon" as the most prominent issue for sustainable development.

The Enquete Commission[7] formulated four general rules for true implementation of sustainable development. One of the rules is about conformity in consumption and production of resources. Due to unbalance between these two factors, it is predicted that the depletion point of crude oil will be reached between 2010 and 2020[8]. The natural gas market is younger and its sustainability is expected for another 60 years[9]. Production-Consumption statistics taken from US Energy department are displayed in Table 1[10].

The question for making the comparison explicit and at the desktop of planner is twofold. What is the optimal exploration rate? What is the optimal consumption rate with respect to the explored reserves of hydrocarbon? Energy crisis creeks at the forehead of planners urging them to collect every possible statistical figure which can contribute in the decision whether as a variable or a constant. There are huge

Table 1: Exploration and consumption statistics (BBTU – Billion British Thermal Unit)

Year	Oil Production	Oil Consumption	Natural Gas Production	Natural Gas Consumption
1960	14,934,611	19,919,230	12,656,133	12,385,366
1970	20,401,210	29,520,695	21,665,670	21,794,707
1980	18,248,917	34,202,356	19,907,600	20,235,459
1990	15,571,185	33,552,534	18,326,155	19,603,168
2000	12,358,101	38,264,304	19,66,1518	23,823,978
2003	12,026,027	38,809,178	19,690,599	22,897,268
2007	10,802,158	39,817,588	19,817,290	23,637,890

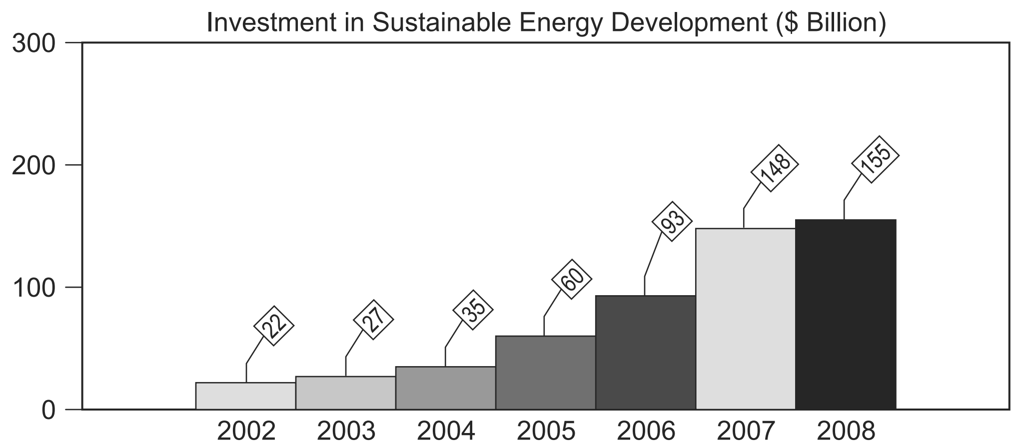

Figure 1: Investments for sustainable energy

repositories of data established to guide the decisions. Most of the repositories are explicitly available[11,12]. The need is to direct planning resources towards establishing conformity between consumption of resources and the rate of its exploration.

Data mining with its state of the art techniques attempts to make sense of information explosion embedded in huge volumes of data[13].The most reliable sphere for making predictions on the basis of hidden facts in the volume of data is the art of data mining.

Data Mining in Energy Sector

Data Mining is meant to find rules for predicting the values of variables on the basis of existing variables. Its use in energy sector is rapidly growing.

Rodriguez-Ortiz G. et.al classified energy consumption using data mining technique based on adaptive resonance theory (ART) algorithm modified with Euclidean distance measure. The results of using data mining were quick retrieval and better visualization of results[14]. Graphet Inc. has developed algorithms for clustering, regression, classification and association rule mining for analyzing energy usage pattern, key indicators to predict energy usage, critical events in system and equipment and stable modes of operations[15].

Patterns in energy usage and exploration can also help developing a strategy for optimal use of energy resources and equipments. It is difficult to understand the events that cause energy usage to rise and fall, particularly when production rates are highly variable, when product mix varies, or when there are several interacting processes on a single site. To develop a plan for optimal usage of provided resources, utilization of data mining techniques can have positive effects on efficiency[16]. N. M. Maricar et.al utilized data mining in energy audit to optimize use of energy hence ensuring comfort[17]. Geoffrey K. F. Tso et.al detected patterns in domestic energy usage in Hong Kong by using data mining[18].

One of the most popular techniques for predictive modeling is Regression analysis. A multiple regression model with more than one explanatory variable may be written as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p + \xi \quad (1)$$

where Y is the output variable, β_i the regression parameters.

In this paper, least-squares method is used for parameter estimation. The estimated parameters are used in regression equation to predict continuous values of output as a linear function of input.

Regression model are popular and common in use because of its ease of interpretability and its use is limited because of its uncertainty about underlying casual mechanism[19].

Al-Garni,AZ et.al used general and stepping regression techniques for forecasting electrical energy consumption in Saudi Arabia[20].

Indicators for Sustainable energy development

Sustainable energy development is vital for eradicating poverty, improving living standards and human welfare[21]. This is understood at the very time when unsustainable development caused a lot of miseries for the world. Energy is key contributor in the scenario because most of the patterns in energy demand and supply are unsustainable. There are some listed variables which help assessing the development either sustainable or unsustainable. With the efforts of Commission on sustainable development (CSD), World Summit on sustainable development is held in Johannesburg, South Africa[22]. The indicators for sustainable development are discussed and finalized in the summit[2]. The indicators which were developed by UNDESA[23] were also considered in the discussion. UNDESA produced 58 indicators for sustainable development out of which only three are related to energy resources. International Atomic Energy Agency (IAEA) produced a set of 41 indicators for sustainable energy development[2]. The project was also registered by Commission on sustainable development (CSD).

The original set of indicators were developed by justly considering four (4) not three (3) dimensions of sustainable development. (1). Economic. (2). Social. (3). Environmental and (4). Institutional. The development of set of energy indicators was a collaborative effort of UNDESA, IEA, Eurostate and European Environment Agency (EEA)[24]. The indicators are analyzed using statistical tools to explore future developments in context of energy policy and national priorities.

PROPOSED METHODOLOGY

The conformity of hydrocarbon exploration rate with that of consumption rate needs to make both of them explicit. The paper addresses the mechanism to predict optimal hydrocarbon exploration rate only. The indicators defined in previous section are classified into Exploration Major and Consumption Major. Cor-

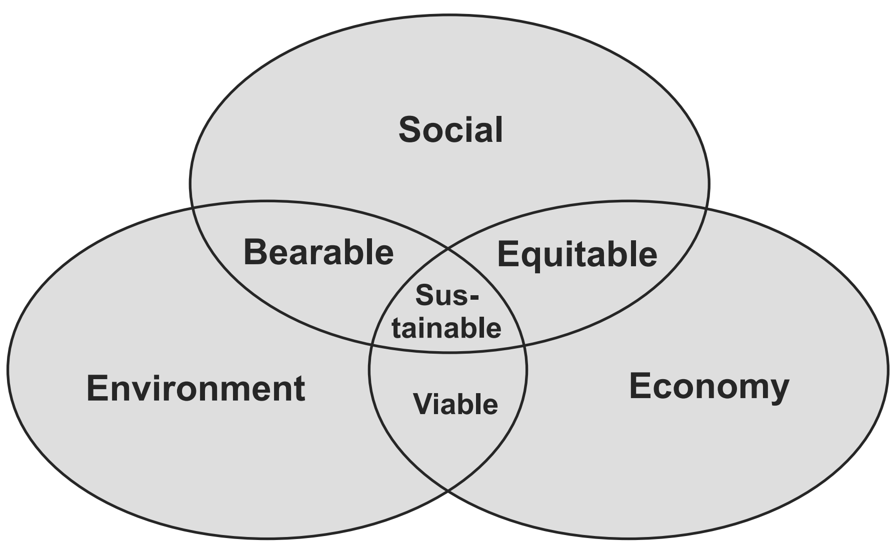

Figure 2: Sustainable Energy Development

relation analysis is used for this classification. The analysis is performed to find the quantitative estimates of underlying energy indicators. It indicates the strength of linear relationship between random variables. As mentioned earlier that 41 indicators have been proposed by IAEA/IEA. 24 out of them are selected and correlated with exploration rate and consumption rate.

The reason for short listing 24 out of 41 is either non-availability of data or non-quantifiable nature of the indicator. The indicators, shown in Table 2, have finally been selected from IAEA/IEA, 2001.

Table 2: Indicators for sustainable energy development and its data type in SPSS 16.0

S. No.	Indicator for Sustainable energy development	Description	Data Type
1.	Total population	Total population	Scale
2.	GDP per capita	Values of goods produced per person	Scale
3.	End-use energy prices with and without tax	Price of energy per capita with and without tax	Scale
4.	Shares of sectors in GDP	Share of energy sector in GDP value added	Scale
5.	Distance travelled per capita	Total distance travelled per capita	Scale
6.	Energy intensity	Energy consumption per unit of GDP	Scale
7.	Energy intensity of selected energy intensive industry	Energy consumption by major industries	Scale
8.	Energy supply efficiency	Fossil fuel efficiency for energy generation	Scale
9.	Energy use per unit of GDP	Energy used per unit GDP	Scale
10.	Expenditure on energy sector	Total expenditure	Scale
11.	Energy use per capita	Use of energy	Scale
12.	Indigenous energy production	Energy production from bigger sources	Scale
13.	Net energy import dependence	Amount of energy to be imported	Scale
14.	Ratio of daily disposable income to the prices of fuels	Ratio of disposable income with energy consumption	Scale
15.	Fraction of disposable income spent on fuels	Disposable income spent on fuel	Scale
16.	Fraction of households dependent upon energy	Average number of household using energy	Scale
17.	Air pollutant emissions	Kg quantity of emissions	Scale
18.	Greenhouse emissions	Quantity of emissions	Scale
19.	Generation of solid waste	Quantity of solid waste	Scale
20.	Generation of radioactive waste	Quantity of radioactive waste	Scale
21.	Fatalities due to accidents with breakdown by fuel chains	Number of accidents at energy sites	Scale
22.	Proven recoverable fossil fuel reserves	Fuel reserves	Scale
23.	Lifetime of proven fossil fuel reserves	Predicted lifetime of reserves	Scale
24.	Rate of deforestation	Rate of forest elimination	Scale

Source: IAEA/IEA., 2001.

To find the correlation between exploration rate and Energy indicator, SPSS 16.0.0 bi-variate correlation is applied to all the indicators in turn with exploration rate and consumption rate. All the indicators are given data type "Scale" in SPSS 16.0. The correlation analysis with consumption rate is not further analyzed in this paper and left for extensions. The average value of correlation expressions for all provided year-wise inputs is taken as correlation value of that indicator. The indicators with strong correlation are extracted. The database structure defined in SPSS for these indicators is shown in Table 2. The correlation analysis equation will become like equation shown in (2) and (3). The indicators are classified on the basis of the value extracted by correlation analysis. Higher value of correlation of an indicator with exploration rate means that the exploration rate is firmly dependent upon the factor. The indicators are classified into (1). Exploration Major. (2). Consumption Major.

$$Corr(\exp\_rate, EI) = \frac{\sum_{i=1}^{n} P(\exp\_rate, EI_i}{n} \qquad (2)$$

$$P(\exp\_rate, EI) = \frac{V(\exp\_rate, EI) - V(\exp\_rate)V(EI)}{V(\exp\_rate^2 V(\exp\_rate\sqrt{V(EI^2 - V^2(EI)}} \qquad (3)$$

Exp_rate = Exploration rate

EI = Energy Indicator

Those indicators which have strong correlation value with the exploration rate will be considered

Exploration and Consumption Major. It would mean that by addressing those causes which are related to this class of indicators can be addressed to improve hydrocarbon exploration rate hence ensuring sustainable development. The indicators which have greater value of correlation with consumption rate are classified as *Consumption Major*. Multiple regression is applied to exploration Major indicators in order to predict the rate of hydrocarbon discovery which will be optimal for its conformity with hydrocarbon consumption. For example Generation of radioactive waste is an indicator for sustainable energy development which served as independent variable in *regret* function. The rate of generation of radioactive waste will not affect sustainability if the exploration rate should not fall below certain threshold. The threshold is calculated with the help of *regret* function.

Multiple regression is a technique that helps us predicting the value of a particular variable on the basis of values of several other variables[25]. The predictor variables used in multiple regression are the same sustainability indicators which have strong correlation with exploration rate. The procedure is depicted in the Figure 1.

The consumption rate can also be predicted by considering consumption Major and is in process which will be published soon.

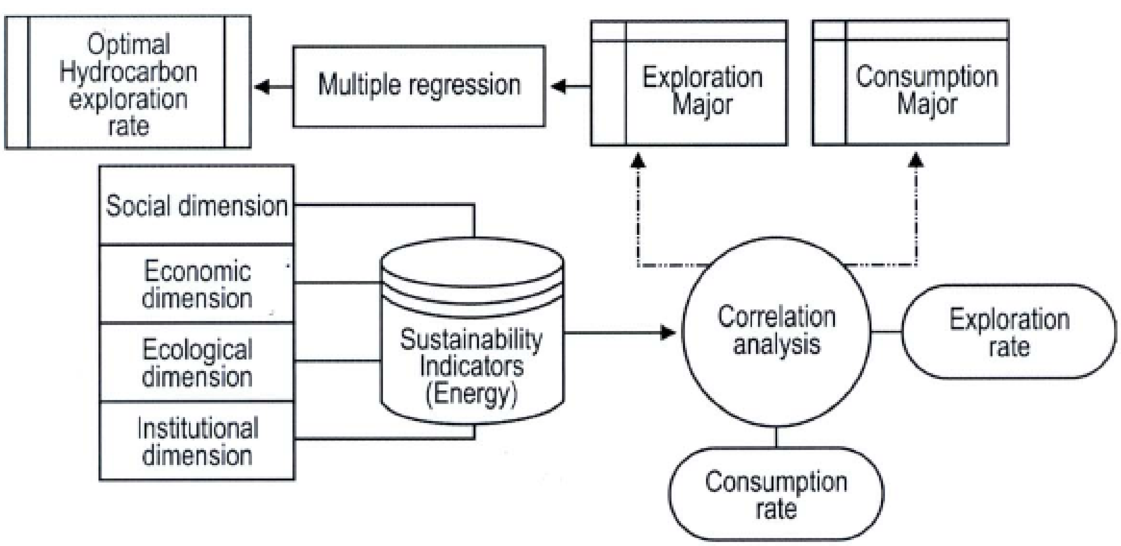

Figure 3: Mechanism for prediction about hydrocarbon exploration rate

RESULTS AND DISCUSSIONS

The database of sustainability indicators is prepared by storing last ten years data (1997-2007) of 17 countries. Indicators for sustainable energy development were 41. The data against all of them were not available therefore 24 indicators have been included which have already been listed in Table 2. The results of correlation analysis are presented in Table 3 and Figure 4. The values obtained after applying correlation analysis were shortlisted by applying *thresholding*.

The correlation is considered to be strong if the value is greater than or equal to 0.5. All those indicators with corresponding values are shown bold in Table 3.

$$O\_Exp\_rate = \alpha + \beta S_1 + \gamma S_2 + ... + \mu S_n + \varepsilon \quad (4)$$

These values are called Exploration Major. The most prominent exploration Major in Table 3 are proven recoverable fossil fuel reserve, Lifetime of proven fossil fuel reserves and quantities of greenhouse emissions. These factors show highest co varying potential with exploration rate. The results have not only selected but also graded the prominence of each indicator in relation with exploration rate. Multiple regression analysis using SPSS is applied only to the factors which have strong correlation.

$$r^2 = \frac{ExplainedVariation}{TotalVariation} = \frac{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 \sum_{i=1}^{n}\left(Y_i - \hat{Y}\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2} \quad (5)$$

In other words the indicators with boldface values are selected for multiple regression. The values of R Square and Adjusted R Square obtained from SPSS after applying multiple regression analysis to datasets is shown in Table 4. F Value is adjusted depending on the variation in R Square to predict increase in exploration rate for optimal compliance with consumption rate.

The regression analysis resulted in prediction about desired increase in hydrocarbon exploration rate

Table 3: Correlation analysis of sustainability indicators with (i). Exploration rate (ii). Consumption rate

ID	Indicator	Erate	Crate
E1.	Total population	0.4	0.8
E2.	GDP per capita	0.5	0.7
E3.	End-use energy prices	0.3	0.6
E4.	Shares of sectors in GDP	0.2	0.6
E5.	Distance travelled per capita	0.2	0
E6.	Energy intensity	0.3	0.5
E7.	Energy intensity (Selected)	0.3	0.5
E8.	Energy supply efficiency	0.3	0.7
E9.	Energy use per unit of GDP	0.5	0.9
E10.	Expenditure on energy sector	0.8	0.6
E11.	Energy use per capita	0.2	0.7
E12.	Indigenous energy production	0.6	0.5
E13.	Net energy import dependence	0.6	0.2
E14.	Ratio of daily disposable income to the prices of fuels	0.2	0.6
E15.	Disposable income spent on fuels	0.7	0.5
E16.	Households dependent upon energy	0	0.6
E17.	Air pollutant emissions	0.7	0.2
E18.	Greenhouse emissions	0.8	0.2
E19.	Generation of solid waste	0.2	0.4
E20.	Generation of radioactive waste	0.6	0.4
E21.	Fatalities due to accidents	0.1	0.4
E22.	Fossil fuel reserves	0.9	0.6
E23.	Lifetime of proven fossil fuel reserves	0.8	0.7
E24.	Rate of deforestation	0.7	0.3

Table 4: Multiple regression analysis to predict optimal value

ID	R^2	Adjusted R^2	Prediction
A	0.741	0.658	6%
B	0.722	0.664	5%
C	0.683	0.682	4%
D	0.650	0.584	5%
E	0.582	0.426	4%
F	0.642	0.621	4%
G	0.778	0.720	9%
H	0.882	0.642	4%
I	0.243	0.240	3%
J	0.524	0.418	4%

in order to manage sustainable development. The names of different countries are coded from A to J in Table 4. For example if the exploration rate in G is increased by 9%, the presence of unsustainable factors will have negligible effects on sustainability. If all indicators are considered in combination then the required exploration rate for G would be 9% as stated earlier. The graph in Figure 5 plotted the correlation value of significant indicators (one with bold value) with required rate of exploration if the significance of all the other indicators is assumed negligible. The statement about the plot (Figure 5) is equally logical if all other factors had been omitted. The adjusted R^2 values indicate the lack of bias in R^2, that is regression models are robust[26]. F-test is used which is a statistical test in which F-distribution is validated if the null hypothesis is true.

Threats

The work was started with certain doubts that all the indicators should not be considered. There should be a criterion to scrutinize the indicators. The doubt is confirmed after performing experiments. A few disorders have been observed in the patterns of results. The data used in the project did not show diversity in results which indicate that the sample data is taken from homogenous population. This was because of unavailability of data related to set of indicators which was either not organized in the same format or was not even available.

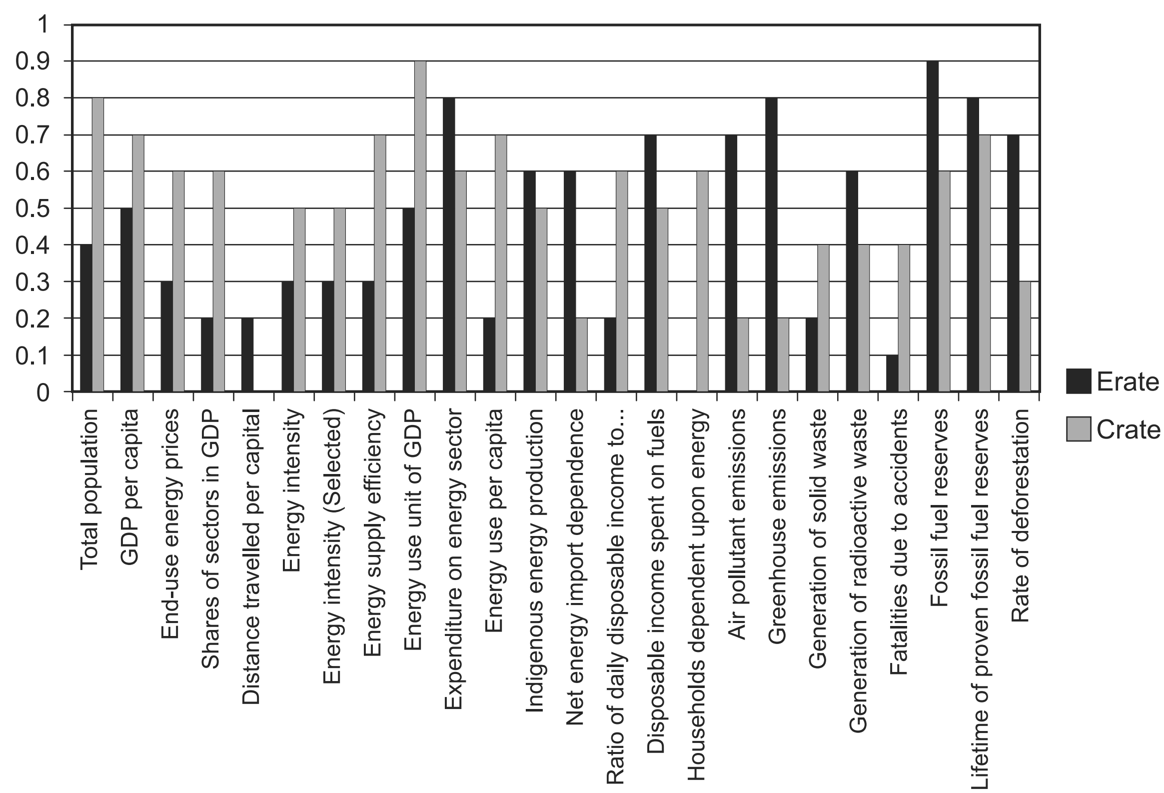

Figure 4: Correlation of sustainability indicators with (i). Exploration rate (ii). Consumption rate

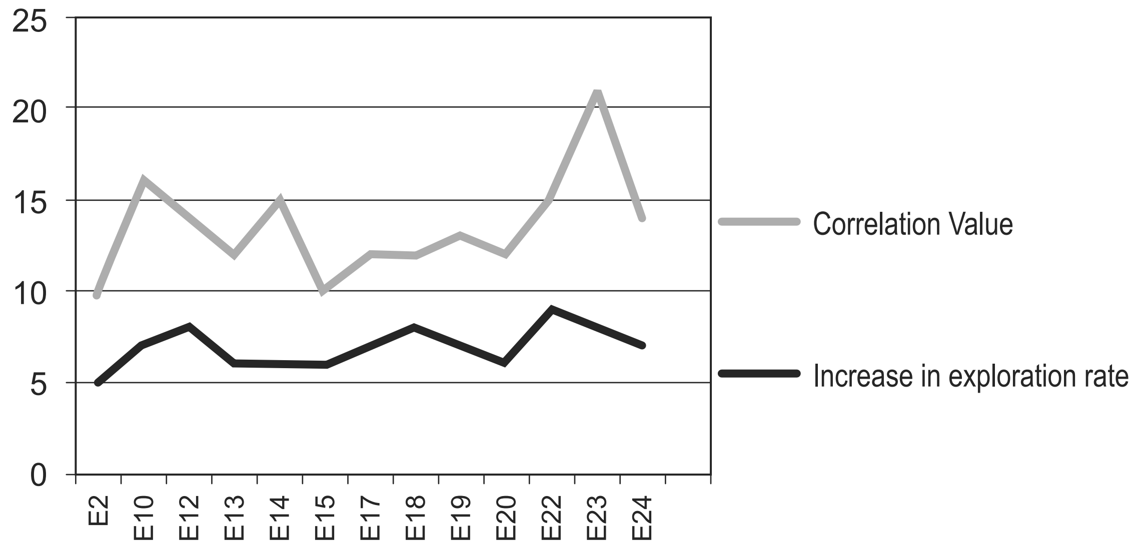

Figure 5: Required increase for individual indicators

CONCLUSION AND FUTURE WORK

In this work we addressed the issue of finding optimal hydrocarbon exploration rate. What should be the increase in hydrocarbon exploration rate that make it conformed with its consumption in a particular country for sustainable development in energy sector. It turns out that optimal exploration rate is predictable by mining the indicators for sustainable energy development. Using this approach, the exploration companies can set their targets and the policy makers can monitor both the exploration companies as well as the companies managing the consumption of energy resources. The work extends the state of the art by three ways.

i. Classifying the metrics in exploration Major and consumption Major.

ii. Gradation of indicators for sustainable energy development.

iii. Prediction of optimal hydrocarbon exploration rate for ensuring conformity of exploration rate with consumption hence assuring sustainable development.

The work never terminates. There is always more to do. The work will be extended in future in the following directions.

Consumption Indicators

The indicators which are prominent to explain consumption of hydrocarbons are not considered in this work. To establish conformance it is necessary to predict optimal consumption rate with respect to exploration as well as sustainable development. The rational value of optimal consumption rate is far ahead. Consumption of energy resources will be analyzed on different grids. The integration of these grids on single platform to predict unified value of optimal consumption rate will be a mega contribution for sustainable energy development.

Diverse Data

Neighborhood patterns were reflected in data. The data from diverse sources and of diverse nature will improve the results. To predict the exploration rate, hydrocarbon exploration factors e.g. crest complexities, rock types etc will predict about the validity of prediction made in this paper.

Mechanism for establishing conformity

The policy makers will take more interest if a procedure is discovered for establishing conformity by extracting non-conforming patterns from previous data warehouses.

REFERENCES

1. Bruntland, G.H. 1987. Our Common Future. World Commission on Environment and Development. University Press, Oxford.

2. Vera, I., Langlois, L., Rogner, H. 2005. International Atomic Energy Agency. Energy Indicators for Sustainable Development. IAEA Austria, pp. 6-20.

3. Arscott, L. 2004. Sustainable Development in the Oil and Gas Industry. Journal of Energy Resources Technology, Transactions of ASME, 126(1).

4. UN World Population Prospects., 2006. The 2006 Revision and World Urbanization Prospect, http://esa.un.org/unpp/p2k0data.asp

5. Sui Northern Gas Pipelines Ltd Pakistan., 2007. The preferred company of demanding customers, Annual Report.

6. Oil and Gas Development Corporation Pakistan., 2008. Annual Statistics. http://www.ogdcl.com/media/Facts.htm

7. Enquette-Kommission., 1993. Schutz des Menschen und der Umwelt-Bewertungskriterien und Perspektiven fur umweltvertragliche Stoffkreislaufe in der Industriegesellschaft, des 12. Deutschen Bundestages Verantwortung fur die Zukunft: Wege zum nachhaltigen Umgang mit Stoff- und Materialstromen. Economia, Bonn.

8. Hiller, K. 1999. Verfugbarkeit von Erdol, Erdol, Ergdas. Kohle 115(2), pp. 50 – 54.

9. Wellmer, F.W., Becker-Platen, J.D. 2002. Sustainable development and the exploitation of mineral and energy resources: a review, Springer-Verlag, International J Earth Sci (Geol Rundsch). 91(5), pp. 723 – 745.

10. Energy Information Administration (EIA)., 2008. Offcial Energy Statistics from US Government, US International Data, http://www.eia.doe.gov/oil_gas/petroleum/info_glance/petroleum.html

11. U.S. Department of Energy, Washington, http://www.energy.gov/

12. Netezza Corporations USA., 2009. http://www.netezza.com/data-warehouse-appliance-industries/energy.aspx

13. Mitra, S., Acharya, T. 2003. Data Mining Multimedia, Soft Computing and Bioinformatics. Wiley Publishers, pp. 1-5.

14. Rodriguez-Ortiz, G., Fernandez-Espinosa, V., Ramos-Niembro, G., Mejia-Lavalle, M. 1994. Using Data Mining Technique to classify energy demand. International Journal of Power and Energy Systems, 19 (2), pp. 168-172.

15. Graphet Inc., 2000. Energy Data Mining and analysis Toolset, http://www.graphet.com/index.php?id=24

16. Xpert Rule Software Ltd., 2008. Brewery and Power Station Pilot Sites evaluate energy efficiency technology, http://www.xpertrule.com/pages/case_eng.htm

17. Maricar, N.M., Kim, G.C., Jamal, N. 2005. Data Mining application in Industrial Energy Audit for Lightning. In. Proceedings European Power and Energy Systems.

18. Tso, G.K.F., Yau, K.K.W. 2003. A study of domestic energy usage pattern in Hong Kong. Energy, *28*, pp. 1671–82.

19. Geoffrey, K.F., K.W. Kelvin, K,W., Yau. 2007. Predicting electrical energy consumption; A comparison of regression analysis, decision tree and neural network. Energy, 32, pp. 1761-1768.

20. Al-Garni, A.Z., Zubair, S.M., Nizami, J.S. 2005. A regression model for electric energy consumption forecasting in Eastern Saudi Arabia. Energy ,*19*, pp. 1043–1049.

21. United Nations Development Program (UNDP), United Nation Department of Economic and Social Affairs (UNDESA), World Energy Council (WEC)., 2000. World Energy Assessment: Energy and the challenge of sustainability, New York.

22. United Nations (UN)., 2002. Report of the World Summit on Sustainable Development. A/CONF. 199(20), United Nations, New York.

23. United Nations Department of Economic and Social Affairs., 2002. Indicators of Sustainable Development: Guidelines and Methodologies, Second edition, New York.

24. International Atomic Energy Agency (IAEA), United Nations Department of Economic and Social Affairs (UNDESA), International Energy Agency (IEA), Eurostat, European Environment Agency (EEA)., 2005. Energy Indicators for Sustainable Development: Methodologies and Guidelines, Vienna.

25. Osborne, J.W., Waters, E. 2002. Four Assumptions of Multiple Regression that Researchers always test. Practical Assessment Research and Evaluation, *8* (2), pp. 1531-7714.

26. Naggapan, N., Ball, T., Zeller, A. 2005. Mining Metrics to predict Component Failures, Microsoft research Redmond, Washington.