World Academy of Science, Engineering and Technology
International Journal of Cognitive and Language Sciences
Vol:15, No:9, 2021

Scenario and Decision Analysis for Solar Energy in Egypt by 2035 Using Dynamic Bayesian Network

Rawaa H. El-Bidweihy, Hisham M. Abdelsalam, Ihab A. El-Khodary

***Abstract*—**Bayesian networks are now considered to be a promising tool in the field of energy with different applications. In this study, the aim was to indicate the states of a previous constructed Bayesian network related to the solar energy in Egypt and the factors affecting its market share, depending on the followed data distribution type for each factor, and using either the Z-distribution approach or the Chebyshev's inequality theorem. Later on, the separate and the conditional probabilities of the states of each factor in the Bayesian network were derived, either from the collected and scrapped historical data or from estimations and past studies. Results showed that we could use the constructed model for scenario and decision analysis concerning forecasting the total percentage of the market share of the solar energy in Egypt by 2035 and using it as a stable renewable source for generating any type of energy needed. Also, it proved that whenever the use of the solar energy increases, the total costs decreases. Furthermore, we have identified different scenarios, such as the best, worst, 50/50, and most likely one, in terms of the expected changes in the percentage of the solar energy market share. The best scenario showed an 85% probability that the market share of the solar energy in Egypt will exceed 10% of the total energy market, while the worst scenario showed only a 24% probability that the market share of the solar energy in Egypt will exceed 10% of the total energy market. Furthermore, we applied policy analysis to check the effect of changing the controllable (decision) variable's states acting as different scenarios, to show how it would affect the target nodes in the model. Additionally, the best environmental and economical scenarios were developed to show how other factors are expected to be, in order to affect the model positively. Additional evidence and derived probabilities were added for the weather dynamic nodes whose states depend on time, during the process of converting the Bayesian network into a dynamic Bayesian network.

***Keywords*—**Bayesian network, Chebyshev, decision variable, dynamic Bayesian network, Z-distribution.

I. Introduction

ENERGY consumption has been increasing rapidly in the whole world over the past years [1] due to the technology revolutions that we have been living in. This led to unstoppable seeking for an energy source that is not just cannot be depleted by using, but also an environmentally friendly source. In this context, we consider here the conditions under which we can depend on a reliable renewable source in Egypt. Systems depending on renewable energy sources play an important role nowadays in the development and the economic growth of countries, several studies [2]-[6] presented the positive impacts of using renewable energy; whether socially, economically, politically or environmentally. It was also indicated that renewable energy will be cheaper and has more potential for growth in the upcoming decades.

Lots of researches and studies [7]-[20] showed how Bayesian Networks (BN) are considered to be a promising and depending-on tool in the field of renewable and non-renewable sources with potential applications. They are useful for time series prediction, classification, and decision making; they are also recommended when there is a lot of missing data or uncertainty in a model.

Based on our previous work [21], this study is focusing specifically on the solar energy in Egypt, as it is considered to be one of the strongest and richest countries with the solar energy.

Fig. 1 [21] shows the final previously constructed BN with its analyzed factors and relationships. Based on the ARIMA/ SARIMA models in [21] which were constructed for the quantitatively factors, the resulted forecasted values can be used to act as the threshold of the states in each node. A one-time step was added in the forecasted models that were developed in [21] for a few factors, as they were found to be one year behind; these factors are: *Oil Reserves, GDP per capita, Humidity, and Pollution & Sand Storms (which is based on Wind & Gust).* Figs. 2-9 show their updated forecasted curves, respectively. Also, the *market share of solar energy* was found to be better estimated as 10% by 2035, rather than 25%. Therefore, the final forecasted (estimated) values of our 11 quantitatively factors and the categories of the 3 qualitatively factors are shown in Table I.

II. Methodologies

According to the behavior of any normal distributed data set, it is proved that approximately 68% of this data fall within one standard deviation from the mean, while 95% fall within two standard deviations from the mean, and finally around 99% fall within three standard deviations from the mean [22]. Z-distribution is considered to be a special member of the normal distribution family; also known as standard normal distribution. It is used to find probabilities and percentiles for normal distribution data; it has a mean 0 and standard deviation 1. The number of standard deviations in which a particular value falls around the mean, whether above it or below it, is called Z-value. The basic z-value for a sample data set; is calculated as:

$$Z = \frac{(X - \mu)}{\sigma}$$

Rawaa H. El-Bidweihy is with the Department of Operations Research and Decision Support, Cairo University, Giza, Egypt (corresponding author, e-mail: r.bidweihy@fci-cu.edu.eg).

Hisham M. Abdelsalam and Ihab A. El-Khodary are with the Department of Operations Research and Decision Support, Cairo University, Giza, Egypt (e-mail: h.abdelsalam@fci-cu.edu.eg, e.elkhodary@fci-cu.edu.eg).

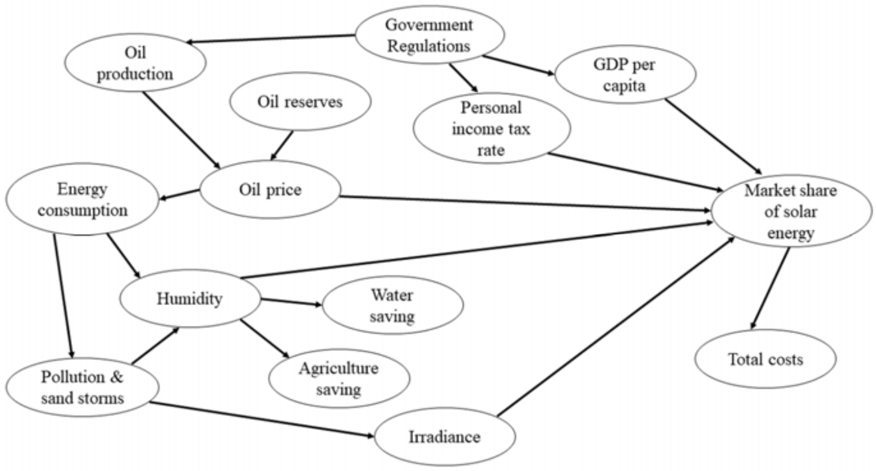


Fig. 1 Structure of the BN [21]

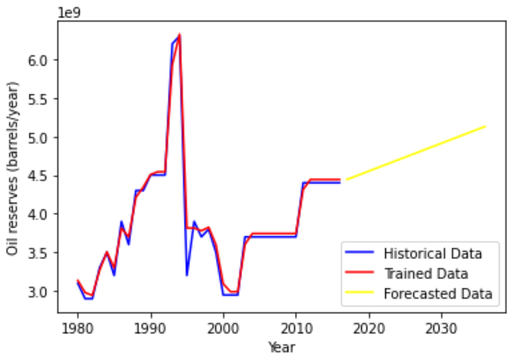


Fig. 2 Oil Reserves (barrels/year) in Egypt vs Year, using ARIMA (1, 1, 0)

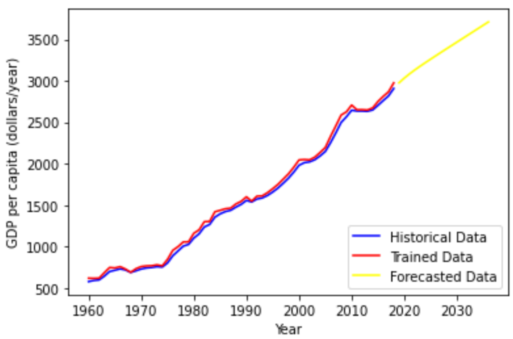


Fig. 3: GDP per Capita (dollars/year) in Egypt vs Year, using ARIMA (1, 1, 1)

Where, the percentage in front of the z-value, in the standard normal distribution table, represents the area to the left of the z-value. Fig. 20 in the appendix, shows a copy of the Z-standard normal distribution table.

In the case of a data set that does not follow a normal distribution; a different amount could be within one standard deviation. Chebyshev's inequality is used to calculate what percentage of data fall within K standard deviations from the mean [23]. The theory states that, at least $1-\frac{1}{k^2}$ of a sample data must fall within *K* standard deviations from the mean,

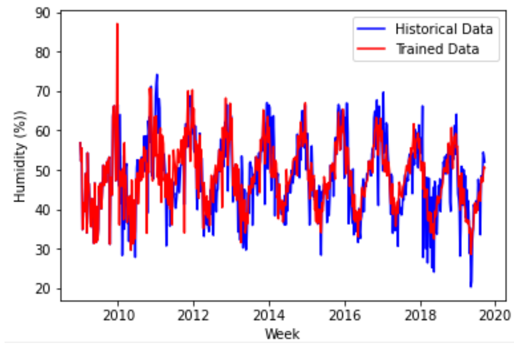


Fig. 4 Humidity (%) in Egypt vs Week, using SARIMA (1, 1, 1) (1, 1, 1) (52)

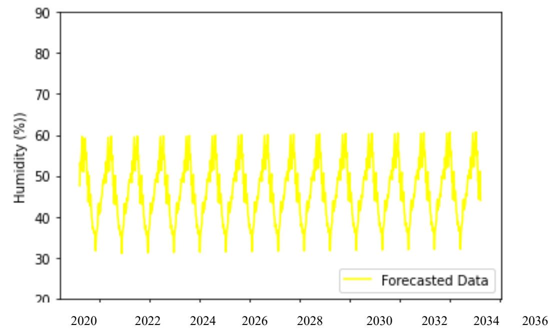


Fig. 5 Forecasted Humidity (%) in Egypt vs Week, using SARIMA (1, 1, 1) (1, 1, 1) (52)

Where *K* is any positive real number greater than one. Taking some values for example, to illustrate this theory:

- For K = 2, we have $1-\frac{1}{k^2}=1-\frac{1}{4}=\frac{3}{4}=75\%$. So, it means that at least 75% of the data values of any distribution must be within two standard deviations of the mean.
- For K = 3, we have $1-\frac{1}{k^2}=1-\frac{1}{9}=\frac{8}{9}=89\%$. So, it means that at least 89% of the data values of any distribution must be within three standard deviations of the mean.

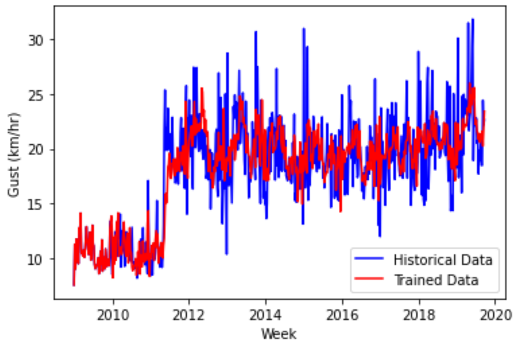


Fig. 6 Gust (km/hr) in Egypt vs Week, using SARIMA (1, 1, 1) (1, 1, 1) (52)

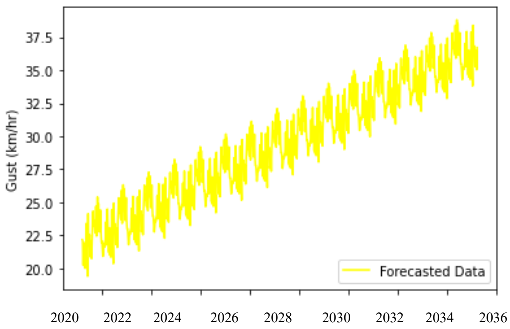


Fig. 7 Forecasted Gust (km/hr) in Egypt vs Week, using SARIMA (1, 1, 1) (1, 1, 1) (52)

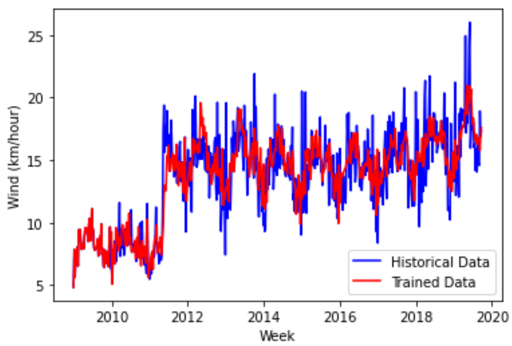


Fig. 8 Wind (km/hr) in Egypt vs Week, using SARIMA (1, 1, 1) (1, 1, 1) (52)

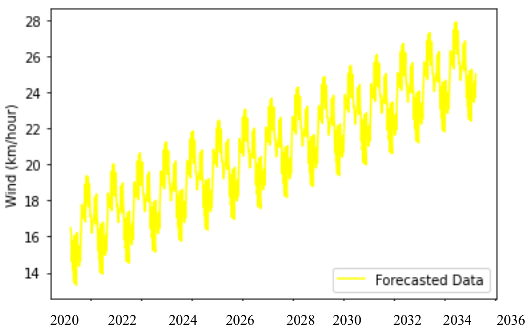


Fig. 9 Forecasted Wind (km/hr) in Egypt vs Week, using SARIMA (1, 1, 1) (1, 1, 1) (52)

TABLE I
FORECASTED VALUE/CATEGORIES OF THE FACTORS IN OUR BN

Factor	Forecasted Value/Categories
Energy Consumption	2440 kW.h/person
Oil Production	707 barrels/day
Oil Price	1 dollar/liter
Oil Reserves	5127841320 barrels/year
GDP per Capita	3710 dollars/year
Personal Income Tax Rate	24%
Humidity	50%
Pollution & Sand Storms	Wind = 25 km/hour & Gust = 37 km/hour
Agriculture Production	15%
Water Saving	25%
Market Share of Solar Energy	10% of the total energy market
Irradiance	Increased/Stable/Decreased
Total Costs	Increased/Unchanged/Decreased
Government Regulations	Changed/Unchanged

- For K = 4, we have $1-\frac{1}{k^2}=1-\frac{1}{16}=\frac{15}{16}=93.75\%$. So, this shows that at least 93.75% of any data values that follow any distribution have to be within two standard deviations of the mean [23].

Let X be a random variable with mean μ and variance σ^2, to calculate the probability of X using the Chebyshev's inequality, there are two main cases:

- Two-sided Chebyshev:

$$P(a<X<b)=P\left(\left|X-\frac{a+b}{2}\right|<\frac{b-a}{2}\right)\geq 1-\frac{\sigma^2+(\mu-\frac{a+b}{2})^2}{(\frac{b-a}{2})^2}$$
$$Where, a=-t+c, b=t+c, for\ any\ t>0$$

- One-sided Chebyshev:

$$P(X\geq\mu+a)\leq\frac{\sigma^2}{\sigma^2+a^2}\ , if\ X>\mu$$
$$(X\leq\mu-a)\leq\frac{\sigma^2}{\sigma^2+a^2}\ , if\ X<\mu$$

The Chebyshev's inequality theorem has been used in a lot of studies, journals and books [24], in order to deal with non-normal distribution data probabilities.

A Python library is a reusable code that developers could include in their programs. Unlike other programming languages like C++ or C, Python libraries are not related to a specific context. A 'library' easily describes a collection of the basic modules.

The Python standard library is a collection of syntax and semantics of the Python programming language. It is written in C, and handles the main functionality like input/output and more than 200 core modules. This library is attached to any Python executing command, but in addition to it, there are lots of other libraries that can be accessed by importing them in the code. Followed by this, we are going to introduce the main libraries that were used in our study:

A. Python Pandas

Python Pandas is considered to be the godfather of all the Python libraries; it is an essential step in the data-science. It is used in importing and reading data from spreadsheets (.csv files), also for processing and analyzing time-series data. It is capable of providing fast and flexible modules to work with structured (tabular, multidimensional… etc.) and time-series data.

B. Numerical Python (NumPy)

Numerical Python (NumPy) is usually used with Pandas for advanced data analysis. It includes math functions and a scientific computing package, which helps in facilitating numeric computations.

C. Python's Iteration Tool (Itertool)

Python's iteration tool (Itertool) is a module that provides multiple functions that work on iterators (for loop, while loop…etc.) to produce complex calculations. It works in a fast and memory-efficient technique, which results in calculating iterator algebra.

GeNIe modeler is a graphical user interface (GUI) that is used for building graphical decision-theoretical and network models. Its name and its uncommon capitalization were generated from the name Graphical Network Interface, it also allows interactive model building and learning. Structural modeling, inference, and learning engine (SMILE) is a reasoning platform engine that contains functions responsible for implementing graphical probabilistic and decision theoretic models, such as BN and Influence diagrams. GeNIe is considered to be the GUI of a SMILE engine, as if it is the front end and SMILE is the robust and running back end.

BN updating is also known as Belief updating or Bayesian inference, it is usually used to compute the impact of observing the values of a specific subset of the network variables, on the probability distribution over the other dependent variables. The techniques used in the Bayesian updating differs, based on the chosen algorithm; whereas GeNIe offers a collection of algorithms depending on the type of the variables in the BN, whether discrete, continuous, or hybrid models.

Sensitivity analysis is a technique that is used to help validate the probability parameters in BN, to check the highly effective and sensitive ones. It is usually done by analyzing the effect of small changes in the nodes' probabilities on the post probabilities of the output or the target nodes. GeNIe implements an algorithm [25] that performs simple sensitivity analysis in BN. It works by calculating a set of derivatives for the posterior probabilities of the output or target nodes over each probability of the BN, these derivatives show the importance of the parameters in calculating the posterior probabilities of the targets. If the derivative is large for a specific parameter; then a small deviation in this parameter, leads to a large difference in the posterior probabilities of the targets.

III. States of the Nodes in Our BN

Despite our published work in [21]; determining the states of each node in the constructed BN will be based on the distribution type of each factor. The historical data were scrapped from multiple sources [26]-[32].

A. Energy Consumption

Based on the collected historical data of the Energy Consumption factor [30] that was analyzed in [21], it was found that it follows a normal distribution curve with mean 1255.656671 kW.h/person and that's why we will use the Z-distribution theorem which according to its characteristics, we will divide the states of the Energy Consumption node into:

$$P(Energy\ Consumption\ \leq 2440\ kW.h/person)$$
$$P(Energy\ Consumpton\ > 2440\ kW.h/person)$$

B. Oil Production

Based on the collected historical data of the Oil Production factor [28] that was analyzed in [21], it was found that it does not follow a normal distribution curve and has mean equals to 562.3568 barrels/day and that is why we will use the Chebyshev's inequality theorem. According to its characteristics, we will divide the states of the Oil Production node into:

$$P(Oil\ Production\ \geq 707\ barrels/day)$$
$$P(Oil\ Production\ < 707\ barrels/day)$$

C. Oil Price

Based on the collected historical data of the Oil Price factor [26] that was analyzed in [21], it was found that it does not follow a normal distribution curve and has mean equals to 0.651972728 dollar/litre and that's why we will use the Chebyshev's inequality theorem. According to its characteristics, we will divide the states of the Oil Price node into:

$$P(Oil\ Price\ \geq 1\ dollar/litre)$$
$$P(Oil\ Price\ < 1\ dollar/litre)$$

D. Oil Reserves

Based on the collected historical data of the Oil Reserves factor [29] that was analyzed in [21], it was found that it follows a normal distribution curve with mean 4223796697.38724 barrels/year and that's why we will use the Z-distribution theorem which according to its characteristics, we will divide the states of the Oil Reserves node into:

$$P(Oil\ Reserves\ \leq 5127841320\ barrels/year)$$
$$P(Oil\ Reserves\ > 5127841320\ barrels/year)$$

E. GDP per Capita

Based on the collected historical data of the GDP per Capita factor [27] that was analyzed in [21], it was found that it follows a normal distribution curve with mean 1992.07204608194 dollars/year and that's why we will use the Z-distribution theorem which according to its characteristics, we will divide the states of the GDP per Capita node into:

$$P(GDP\ per\ Capita\ \leq 3710\ dollars/year)$$
$$P(GDP\ per\ Capita\ > 3710\ dollars/year)$$

F. Personal Income Tax Rate

Based on the collected historical data of the Personal Income

Tax Rate factor [31] that was analyzed in [21], it was found that it follows a normal distribution curve with mean 22.7544471666667% and that's why we will use the Z-distribution theorem which according to its characteristics, we will divide the states of the Personal Income Tax Rate node into:

$$P(Personal\ Income\ Tax\ Rate\ \leq 24\%)$$
$$P(Personal\ Income\ Tax\ Rate\ > 24\%)$$

G. Humidity

Based on the collected historical data of the Humidity factor [32] that was analyzed in [21], it was found that it does not follow a normal distribution curve and has mean equals to 47.3318368629266% and that's why we will use the Chebyshev's inequality theorem. According to its characteristics; since the mean is smaller than the forecasted value, we will divide the states of the Humidity node into:

$$P(Humidity\ \geq 50\%)$$
$$P(Humidity\ < 50\%)$$

H. Pollution & Sand Storms

As it was explained before in [3], the Pollution & Sand Storms node will be represented using two factors which are Gust and Wind. So, we have to get their states first and then merge them together in the main Pollution & Sand Storms node.

i) Wind

Based on the collected historical data of the Wind factor [32] that was analyzed in [21], it was found that it does not follow a normal distribution curve and has mean equals to 18.1483649611563 km/hour and that's why we will use the Chebyshev's inequality theorem. According to its characteristics; since the mean is smaller than the forecasted value, we will divide the states of the Wind node into:

$$P(Wind\ \geq 25\ km/hour)$$
$$P(Wind\ < 25\ km/hour)$$

ii) Gust

Based on the collected historical data of the Gust factor [32] that was analyzed in [21], it was found that it does not follow a normal distribution curve and has mean equals to 24.9476576107678 km/hour and that's why we will use the Chebyshev's inequality theorem. According to its characteristics, we will divide the states of the Humidity node into:

$$P(Gust\ \geq 37\ km/hour)$$
$$P(Gust\ < 37\ km/hour)$$

By merging all the possible combinations of the Gust & Wind together, the final states of the Pollution & Sand Storms node will be divided as follows:

$$P(Wind\ < 25\ km/hour\ \&\ Gust\ < 37\ km/hour)$$
$$P(Wind\ < 25\ km/hour\ \&\ Gust\ \geq 37\ km/hour)$$
$$P(Wind\ \geq 25\ km/hour\ \&\ Gust\ < 37\ km/hour)$$
$$P(Wind\ \geq 25\ km/hour\ \&\ Gust\ \geq 37\ km/hour)$$

İ. Agriculture Production

According to the forecasted value of the agriculture production percentage in 2035 as shown in Table I, it was assumed that it follows a normal distribution which leads to dividing the states into:

$$P(Agriculture\ Production\ \leq 15\%)$$
$$P(Agriculture\ Production\ > 15\%)$$

J. Water Saving

According to the forecasted value of the water saving percentage in 2035 as shown in Table I, it was assumed that it follows a normal distribution which leads to dividing the states into:

$$P(Water\ Saving\ \leq 25\%)$$
$$P(Water\ Saving\ > 25\%)$$

K. Market Share of Solar Energy

According to the forecasted value of the market share of the solar energy percentage with respect to the energy market in 2035 as shown in Table I, it was assumed that it follows a normal distribution which leads to dividing the states into:

$$P(Market\ Share\ of\ Solar\ Energy\ \leq 25\%)$$
$$P(Market\ Share\ of\ Solar\ Energy\ > 25\%)$$

As for the discrete qualitative categorized factors as was illustrated in Table I, states will be divided qualitatively not numerically.

L. Irradiance

$$P(Irradiance = Increased)$$
$$P(Irradiance = Stable)$$
$$P(Irradiance = Decreased)$$

M. Total Costs

$$P(Total\ Costs = Increased)$$
$$P(Total\ Costs = Unchanged)$$
$$P(Total\ Costs = Decreased)$$

N. Government Regulations

$$P(Government\ Regulations = Changed)$$
$$P(Government\ Regulations = Unchanged)$$

IV. Independent Probabilities of the BN Factors

A. Energy Consumption

Since the historical data of the Energy Consumption factor followed a normal distribution curve; we used the Z-table for determining the probabilities of each state calculated as:

- $P(Energy\ Consumption\ \leq 2440\ kW.h/person)$

$$Z = \frac{Forecasted\ value - mean}{Standard\ Deviation} = \frac{2440 - 1255.65667093692}{706.304054398271} = 1.676818$$

$$P(Area\ \leq Z) = 0.9535$$

$$\therefore P(Energy\ Consumption\ \leq 2440\ kW.h/person) \cong 0.95$$

- $P(Energy\ Consumption\ > 2440\ kW.h/person) = 1 - 0.95 = 0.05$

B. Oil Production

Since the historical data of the Oil Production factor did not follow a normal distribution curve; we used the Chebyshev's inequality theorem for determining the probabilities of each state, calculated as:

- $P(Oil\ Production\ \geq 707\ barrels/day)$

$$= \frac{variance}{variance+(707-mean)^2} = \frac{56072.15618}{56072.15618+(707-562.3568)^2} = 0.728268352 \cong 0.73$$

- $P(Oil\ Production\ < 707\ barrels/day) = 1 - 0.73 = 0.27$

C. Oil Price

Since the historical data of the Oil Price factor did not follow a normal distribution curve; we used the Chebyshev's inequality theorem for determining the probabilities of each state calculated as:

- $P(Oil\ Price\ \geq 1\ dollar/liter)$

$$= \frac{variance}{variance+(707-mean)^2} = \frac{0.062486893}{0.062486893+(1-0.651972728)^2} = 0.340324249 \cong 0.34$$

- $P(Oil\ Price\ < 1\ dollar/liter) = 1 - 0.34 = 0.66$

D. Oil Reserves

Since the historical data of the Oil Reserves factor followed a normal distribution curve; we used the Z-table for determining the probabilities of each state, calculated as:

- $P(Oil\ Reserves\ \leq 5127841320\ barrels/year)$

$$Z = \frac{Forecasted\ value-mean}{Standard\ Deviation} = \frac{5127841320-4223796697.38724}{725712150.380782} = 1.245734$$

$$P(Area\ \leq Z) = 0.8944$$
$$\therefore P(Oil\ Reserves\ \leq 5127841320\ barrels/year) \cong 0.89$$

- $P(Oil\ Reserves\ > 5127841320\ barrels/year) = 1 - 0.89 = 0.11$

E. GDP per Capita

Since the historical data of the GDP Per Capita factor followed a normal distribution curve; we used the Z-table for determining the probabilities of each state calculated as:

- $P(GDP\ per\ Capita\ \leq 3710\ dollars/year)$

$$Z = \frac{Forecasted\ value-mean}{Standard\ Deviation} = \frac{3710-1992.07204608194}{988.389954820834} = 1.738107$$

$$P(Area\ \leq Z) = 0.9591$$
$$\therefore P(GDP\ per\ Capita\ \leq 3710\ dollars/year) \cong 0.96$$

- $P(GDP\ per\ Capita\ > 3710\ dollars/year) = 1 - 0.96 = 0.04$

F. Personal Income Tax Rate

Since the historical data of the Personal Income Tax Rate factor followed a normal distribution curve; we used the Z-table for determining the probabilities of each state calculated as:

- $P(Personal\ Income\ Tax\ Rate\ \leq 24\%)$

$$Z = \frac{Forecasted\ value-mean}{Standard\ Deviation} = \frac{24-22.7544471666667}{2.02359118706826} = 0.615516$$

$$P(Area\ \leq Z) = 0.7257$$
$$\therefore P(Personal\ Income\ Tax\ Rate\ \leq 24\%) \cong 0.73$$

- $P(Personal\ Income\ Tax\ Rate\ > 24\%) = 1 - 0.73 = 0.27$

G. Humidity

Since the historical data of the Humidity factor did not follow a normal distribution curve; we used the Chebyshev's inequality theorem for determining the probabilities of each state calculated as:

- $P(Humidity\ \geq 50\%)$

$$= \frac{variance}{variance+(50-mean)^2} = \frac{59.14741738}{59.14741738+(50-47.3318368629266)^2} = 0.892568745 \cong 0.89$$

- $P(Humidity\ < 50\%) = 1 - 0.89 = 0.11$

H. Pollution & Sand Storms

The data and calculations of the *Pollution & Sand Storms*' factor will be based on two sub-factors which are Gust and Wind.

i) Gust

Since the historical data of the Gust factor did not follow a normal distribution curve; we used the Chebyshev's inequality theorem for determining the probabilities of each state calculated as:

- $P(Gust\ \geq 37\ km/hour)$

$$= \frac{variance}{variance+(37-mean)^2} = \frac{54.85839576}{54.85839576+(37-24.9476576107678)^2} = 0.274131128 \cong 0.27$$

- $P(Gust\ < 37\ km/hour) = 1 - 0.27 = 0.73$

ii) Wind

Since the historical data of the Wind factor did not follow a normal distribution curve; we used the Chebyshev's inequality theorem for determining the probabilities of each state calculated as:

- $P(Wind\ \geq 25\ km/hour)$

$$= \frac{variance}{variance+(25-mean)^2} = \frac{25.91581515}{25.91581515+(25-18.1483649611563)^2} = 0.355689814 \cong 0.355$$

- $P(Wind\ < 25\ km/hour) = 1 - 0.355 = 0.645$

İ. Government Regulations

As for the Government Regulations factor, the probability of each state was estimated based on experts' opinions and the current unstable state of the governments to be:

- $P(Governemnt\ Regulations = Changed) = 0.8$
- $P(Governemnt\ Regulations = UnChanged) = 0.2$

V. Conditional Probabilities of the BN Factors

Our main focus is using BN and Bayesian theorem, which are based on conditional probabilities, not just independent or marginal probabilities. We have to calculate the conditional probabilities for each factor in our BN, according to the constructed relationships between factors as shown in Fig. 1, and by going through the scrapped historical data while recording how many times each combination of events has taken place.

Concerning the probabilities regarding the changing of the government regulations' state, they were assumed to be 50/50 probability, as it is considered to be a new event that did not happen before.

A. Energy Consumption

The final states of the Energy Consumption factor in the BN with an incoming arc to the node from the Oil Price factor are:

- $P(Energy\ Consumption\ \leq 2440\ kW.h/person) \mid P(Oil\ Price < 1\ dollar/litre)\ = 1$
- $P(Energy\ Consumption > 2440\ kW.h/person) \mid P(Oil\ Price < 1\ dollar/litre)\ = 0$
- $P(Energy\ Consumption\ \leq 2440\ kW.h/person) \mid P(Oil\ Price \geq 1\ dollar/litre)\ = 0.9$
- $P(Energy\ Consumption > 2440\ kW.h/person) \mid P(Oil\ Price \geq 1\ dollar/litre)\ = 0.1$

B. Oil Production

The final states of the Oil Production factor in the BN with an incoming arc to the node from the Government Regulations factor are:

- $P(Oil\ Production\ < 707\ barrels/day) \mid P(Government\ Regulations\ changed) = 0.5$
- $P(Oil\ Production\ \geq 707\ barrels/day) \mid P(Government\ Regulations\ changed) = 0.5$
- $P(Oil\ Production < 707\ barrels/day) \mid P(Government\ Regulations\ is\ unchanged) = 0.27$
- $P(Oil\ Production \geq 707\ barrels/day) \mid P(Government\ Regulations\ is\ unchanged) = 0.73$

C. Oil Price

The final states of the Oil Price factor in the BN with incoming arcs to the node from the Oil Production and the Oil reserves factors are:

- $P(Oil\ Price < 1\ dollar/litre) \mid [P(Oil\ Reserves\ \leq 5127841320\ barrels/year)\ and\ P(Oil\ Production < 707\ barrels/day)]\ =\ 0$
- $P(Oil\ Price \geq 1\ dollar/litre) \mid [P(Oil\ Reserves\ \leq 5127841320\ barrels/year)\ and\ P(Oil\ Production < 707\ barrels/day)]\ =\ 1$
- $P(Oil\ Price < 1\ dollar/litre) \mid [P(Oil\ Reserves\ \leq 5127841320\ barrels/year)\ and\ P(Oil\ Production \geq 707\ barrels/day)]\ =\ 0.92$
- $P(Oil\ Price \geq 1\ dollar/litre) \mid [P(Oil\ Reserves\ \leq 5127841320\ barrels/year)\ and\ P(Oil\ Production \geq 707\ barrels/day)]\ =\ 0.08$
- $P(Oil\ Price < 1\ dollar/litre) \mid [P(Oil\ Reserves > 5127841320\ barrels/year)\ and\ P(Oil\ Production < 707\ barrels/day)]\ =\ 0.5$
- $P(Oil\ Price \geq 1\ dollar/litre) \mid [P(Oil\ Reserves > 5127841320\ barrels/year)\ and\ P(Oil\ Production < 707\ barrels/day)]\ =\ 0.5$
- $P(Oil\ Price < 1\ dollar/litre) \mid [P(Oil\ Reserves > 5127841320\ barrels/year)\ and\ P(Oil\ Production \geq 707\ barrels/day)]\ =\ 0.5$
- $P(Oil\ Price \geq 1\ dollar/litre) \mid [P(Oil\ Reserves > 5127841320\ barrels/year)\ and\ P(Oil\ Production \geq 707\ barrels/day)]\ =\ 0.5$

D. Oil Reserves

The final states of the Oil Reserves factor in the BN with an incoming arc to the node from the Oil Production factor are:

- $P(Oil\ Reserves\ \leq 5127841320\ barrels/year) \mid P(Oil\ Production < 707\ barrels/day)\ =\ 0.89$
- $P(Oil\ Reserves\ > 5127841320\ barrels/year) \mid P(Oil\ Production < 707\ barrels/day)\ =\ 0.11$
- $P(Oil\ Reserves\ \leq 5127841320\ barrels/year) \mid P(Oil\ Production \geq 707\ barrels/day)\ =\ 1$
- $P(Oil\ Reserves\ > 5127841320\ barrels/year) \mid P(Oil\ Production \geq 707\ barrels/day)\ =\ 0$

E. GDP per Capita

The final states of the GDP per Capita factor in the BN with an incoming arc to the node from the Government Regulations factor are:

- $P(GDP\ per\ Capita\ \leq 3710\ dollars/year) \mid P(Government\ Regulations\ changed) = 0.5$
- $P(GDP\ per\ Capita\ > 3710\ dollars/year) \mid P(Government\ Regulations\ changed) = 0.5$
- $P(GDP\ per\ Capita\ \leq 3710\ dollars/year) \mid P(Government\ Regulations\ is\ unchanged) = 0.96$
- $P(GDP\ per\ Capita\ > 3710\ dollars/year) \mid P(Government\ Regulations\ is\ unchanged) = 0.04$

F. Personal Income Tax Rate

The final states of the Personal Income Tax Rate factor in the BN with an incoming arc to the node from the Government Regulations factor are:

- $P(Personal\ Income\ Tax\ Rate\ \leq 24\%) \mid P(Government\ Regulations\ changed) = 0.5$
- $P(Personal\ Income\ Tax\ Rate\ > 24\%) \mid P(Government\ Regulations\ changed) = 0.5$
- $P(Personal\ Income\ Tax\ Rate\ \leq 24\%) \mid P(Government\ Regulations\ is\ unchanged) = 0.73$
- $P(Personal\ Income\ Tax\ Rate\ > 24\%) \mid P(Government\ Regulations\ is\ unchanged) = 0.27$

G. Humidity

The final states of the Humidity factor in the BN with incoming arcs to the node from the Energy Consumption and Pollution & Sand Storms factors are:

- $P(Humidity < 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/person)\ and\ P(Gust\ < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)]\ = 0.6$
- $P(Humidity\ \geq 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/person)\ and\ P(Gust\ < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)]\ = 0.4$
- $P(Humidity < 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/person)\ and\ P(Gust\ \geq 37\ km/hour\ \&\ Wind\ < 25\ km/hour)]\ = 0$
- $P(Humidity\ \geq 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/person)\ and\ P(Gust\ \geq 37\ km/hour\ \&\ Wind\ < 25\ km/hour)]\ = 1$
- $P(Humidity < 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/person)\ and\ P(Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)]\ = 0.88$
- $P(Humidity\ \geq 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/person)\ and\ P(Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)]\ = 0.12$
- $P(Humidity < 50\%) \mid [P(Energy\ Consumption\ \leq 2440\ kW.h/$

$person)$ and $P(Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)] = 0.73$

- $P(Humidity\ \geq 50\%)|\ [P(Energy\ Consumption\ \leq 2440\ kW.h/person)$ and $P(Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)] = 0.27$
- $P(Humidity < 50\%)|\ [P(Energy\ Consumption\ > 2440\ kW.h/person)$ and $P(Gust\ < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)] = 0.5$
- $P(Humidity\ \geq 50\%)|\ [P(Energy\ Consumption\ > 2440\ kW.h/person)$ and $P(Gust\ < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)] = 0.5$
- $P(Humidity < 50\%)|\ [P(Energy\ Consumption > 2440\ kW.h/person)$ and $P(Gust\ \geq 37\ km/hour\ \&\ Wind\ < 25\ km/hour)] = 0.5$
- $P(Humidity\ \geq 50\%)|\ [P(Energy\ Consumption > 2440\ kW.h/person)$ and $P(Gust\ \geq 37\ km/hour\ \&\ Wind\ < 25\ km/hour)] = 0.5$
- $P(Humidity < 50\%)|\ [P(Energy\ Consumption > 2440\ kW.h/person)$ and $P(Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)] = 0.5$
- $P(Humidity\ \geq 50\%)|\ [P(Energy\ Consumption > 2440\ kW.h/person)$ and $P(Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)] = 0.5$
- $P(Humidity < 50\%)|\ [P(Energy\ Consumption > 2440\ kW.h/person)$ and $P(Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)] = 0.5$
- $P(Humidity\ \geq 50\%)|\ [P(Energy\ Consumption > 2440\ kW.h/person)$ and $P(Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)] = 0.5$

H. Pollution & Sand Storms

The final states of the Pollution & Sand Storms factor in the BN with an incoming arc to the node from the Energy Consumption factor are:

- *P(Gust <37 km/hour & Wind <25 km/hour)| P(Energy Consumption ≤2440 kW.h/person) =0.9*
- *P(Gust ≥37 km/hour & Wind <25 km/hour)| P(Energy Consumption ≤2440 kW.h/person) =0.004*
- *P(Gust <37 km/hour & Wind ≥25 km/hour)| P(Energy Consumption ≤2440 kW.h/person) =0.063*
- *P(Gust ≥37 km/hour & Wind ≥25 km/hour) | P(Energy Consumption ≤2440 kW.h/person) =0.033*
- *P(Gust <37 km/hour & Wind <25 km/hour)| P(Energy Consumption>2440 kW.h/person) =0.25*
- *P(Gust ≥37 km/hour & Wind <25 km/hour)| P(Energy Consumption>2440 kW.h/person) =0.25*
- *P(Gust <37 km/hour & Wind ≥25 km/hour)| P(Energy Consumption>2440 kW.h/person) =0.25*
- *P(Gust ≥37 km/hour & Wind ≥25 km/hour) | P(Energy Consumption>2440 kW.h/person) =0.25*

İ. Agriculture Production

The final states of the Agriculture Production factor in the BN with an incoming arc to the node from the Humidity factor are:

- *PAgriculture Production ≤15%| PHumidity <50%=0.3*
- *PAgriculture Production >15%| PHumidity <50%=0.7*
- *PAgriculture Production ≤15%| PHumidity ≥50%=0.7*
- *PAgriculture Production >15%| PHumidity ≥50%=0.3*

J. Water Saving

The final states of the Water saving factor in the BN with an incoming arc to the node from the Humidity factor are:

- $P(Water\ Saving\ \leq 25\%)|\ P(Humidity\ < 50\%) = 0.3$
- $P(Water\ Saving\ > 25\%)|\ P(Humidity\ < 50\%) = 0.7$
- $P(Water\ Saving \leq 25\%)|\ P(Humidity\ \geq 50\%) = 0.7$
- $P(Water\ Saving\ > 25\%)|\ P(Humidity\ \geq 50\%) = 0.3$

K. Market Share of Solar Energy

The final states of the Market Share of the Solar Energy factor in the BN are determined with incoming arcs to the node from the GDP per Capita, Personal Income Tax Rate, Oil Price, Humidity, and Irradiance factors, which will result in 96 probabilities. Due to the lack of information in getting all these probabilities from historical data and experts, they will be assumed based on historical patterns and experts' expectations.

L. Irradiance

The final states of the Irradiance factor in the BN with an incoming arc to the node from the Pollution & Sand Storms factor are:

- $P(Irradiance = Increased)|\ (Gust\ < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)\ = 0.93$
- $P(Irradiance = Stable)|\ (Gust\ < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)\ = 0.06$
- $P(Irradiance = Decreased)|\ (Gust < 37\ km/hour\ \&\ Wind\ < 25\ km/hour)\ = 0.01$
- $P(Irradiance = Increased)|\ (Gust\ \geq 37\ km/hour\ \&\ Wind < 25\ km/hour)\ = 0.05$
- $P(Irradiance = Stable)|\ (Gust\ \geq 37\ km/hour\ \&\ Wind < 25\ km/hour)\ = 0.25$
- $P(Irradiance = Decreased)|\ (Gust\ \geq 37\ km/hour\ \&\ Wind < 25\ km/hour)\ = 0.7$
- $P(Irradiance = Increased)|\ (Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)\ = 0.6$
- $P(Irradiance = Stable)|\ (Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)\ = 0.3$
- $P(Irradiance = Decreased)|\ (Gust\ < 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)\ = 0.1$
- $P(Irradiance = Increased)|\ (Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)\ = 0.01$
- $P(Irradiance = Stable)|\ (Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)\ = 0.09$
- $P(Irradiance = Decreased)|\ (Gust\ \geq 37\ km/hour\ \&\ Wind\ \geq 25\ km/hour)\ = 0.9$

M. Total Cost

The final states of the Total Costs factor in the BN with an incoming arc to the node from the Market share of the solar energy factor are:

- $P(Total\ Costs = Increased)|\ P(Market\ share\ of\ the\ solar\ energy\ \leq 25\%\ of\ the\ total\ energy\ market) = 0.8$
- $P(Total\ Costs = UnChanged)|\ P(Market\ share\ of\ the\ solar\ energy\ \leq 25\%\ of\ the\ total\ energy\ market) = 0.15$
- $P(Total\ Costs = Decreased)|\ P(Market\ share\ of\ the\ solar\ energy\ \leq 25\%\ of\ the\ total\ energy\ market) = 0.05$
- $P(Total\ Costs = Increased)|\ P(Market\ share\ of\ the\ solar\ energy > 25\%\ of\ the\ total\ energy\ market) = 0.05$
- $P(Total\ Costs = UnChanged)|\ P(Market\ share\ of\ the\ solar\ energy > 25\%\ of\ the\ total\ energy\ market) = 0.15$
- $P(Total\ Costs = Decreased)|\ P(Market\ share\ of\ the\ solar\ energy > 25\%\ of\ the\ total\ energy\ market) = 0.8$

N. Government Regulations

Concerning the Government Regulations node in the BN; since it has no parent nodes, therefore it will only depend on the independent probabilities that were discussed in the previous section. Our final BN with all the states and probabilities is shown in Fig. 10.

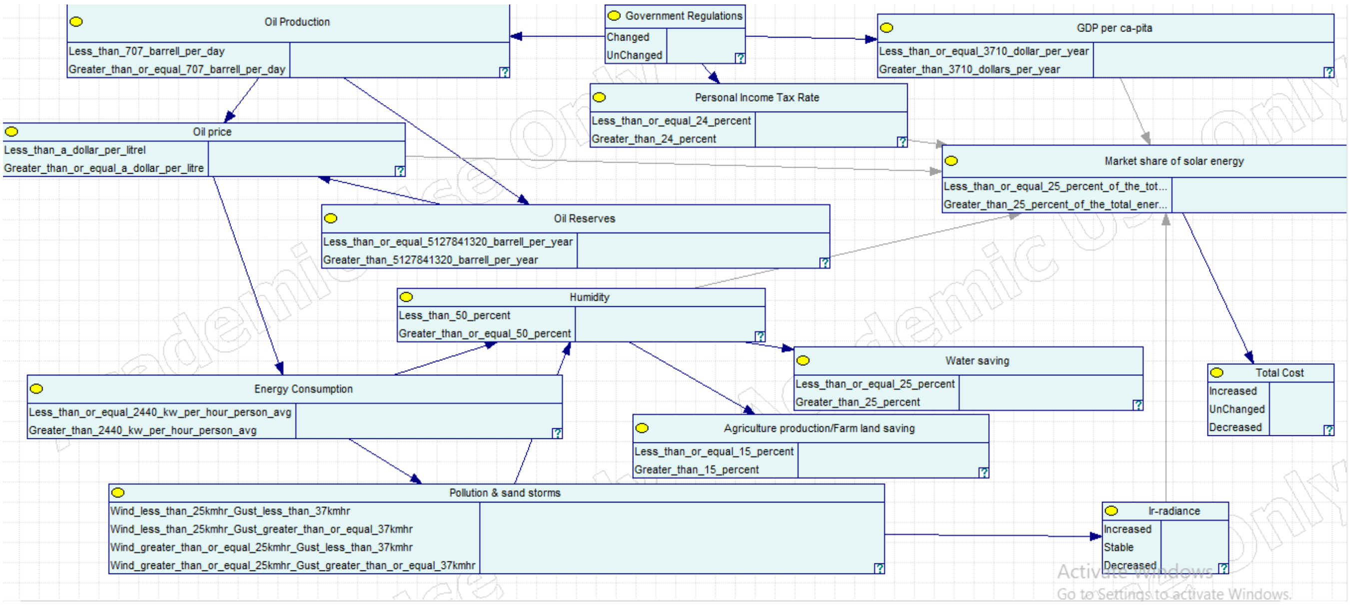


Fig. 10 Final BN with states and probabilities

VI. Results & Analysis

Based on the nature of our BN and its variables, we will be using the clustering built-in algorithm in GENIE which is known as the fastest exact algorithm in Bayesian updating probabilities, in case of not having very large or complex network as our case. Fig. 11 shows the updated conditional probabilities for each state in our BN using GENIE's clustering algorithm.

After updating the BN probabilities, we checked the arc's strength of influence between every two connected nodes. The thickness of the arcs depends on the strength of influence between the connected nodes. It is calculated from the conditional probability table of the child node, by measuring the distance between the states of the parent node and the conditional probability distributions of the child node depending on it. The distance is calculated using Euclidean rule and each arc is represented by the average of all the 2 nodes connected distances. The final strength of influences of all the arcs are shown in Fig. 12, the thickness is based on the absolute value of the average distance.

Checking different scenarios of the states that could happen in the future and their effect on the market share of the solar energy will give us detailed insights about the overall picture and this will assist in the decision-making process in the renewable energy's field, especially the solar energy.

As we have 4 target nodes, which are: *Market Share of Solar Energy, Total Cost, Agriculture Production, and Water Saving;* so, we are dealing with the other 10 nodes, which are: *Energy Consumption, Oil Production, Oil Price, Oil Reserves, GDP per Capita income, Personal Income Tax Rate, Humidity, Pollution & Sand Storms, Irradiance, and Government Regulations* that are responsible for the scenario analysis step.

Using the itertools library in Python, it was calculated that we have a total of 3072 possible combinations of states that could take place. Since, it would be hard and time consuming to check all the 3072 scenarios, so we are going to focus only on the scenarios covering the factors that have the highest direct effect on the Market share of solar energy. In order to get those factors, we checked the rate of change of the Market share of solar energy, with changing each state of each factor separately and choosing the ones with rate of change, greater than the average.

By getting the absolute total summation of the rate of changes, we get 119 as a total and we have 23 different states, thus an average of 119/23 = 5.17. Therefore, we are going to analyze all the possible combinations of the factors that have a rate of change greater than 5.17, which are 5 factors: *Energy Consumption, Pollution & Sand Storms, Irradiance, Personal Income Tax Rate, and GDP per Capita Income*. Focusing on only these 5 factors, resulted in having only 96 possible combinations instead of 3072.

A. Best Scenario

Out of the 96 implemented scenarios, the best scenarios for the Market share of the solar energy are 4 scenarios which resulted in the highest probability of the solar energy having more than 10% of the total energy market by 85% probability, and that's when:

- $Energy\ Consumption > 2440\ kw/hr/person$
- $GDP\ per\ Capita\ income > 3710\$/year$
- $Personal\ Income\ Tax\ Rate \leq 24\%$

Open Science Index, Cognitive and Language Sciences Vol:15, No:9, 2021 publications.waset.org/10012227.pdf

- *Pollution & Sand Storms is in any state*
- *Irradiance increases*

B. Worst Scenario

Out of the 96 implemented scenarios, the worst scenario for the Market share of the solar energy resulted in the lowest probability of the solar energy having more than 10% of the total energy market by only 24% probability, and that's when:

- $Energy\ Consumption\ \leq 2440\ kw/hr/person$
- $GDP\ per\ Capita\ income\ \leq 3710\$/year$
- $Personal\ Income\ Tax\ Rate\ \leq 24\%$
- $Wind\ < 25\ km/hr\ \&\ Gust \geq 37\ km/hr$
- *Irradiance decreases*

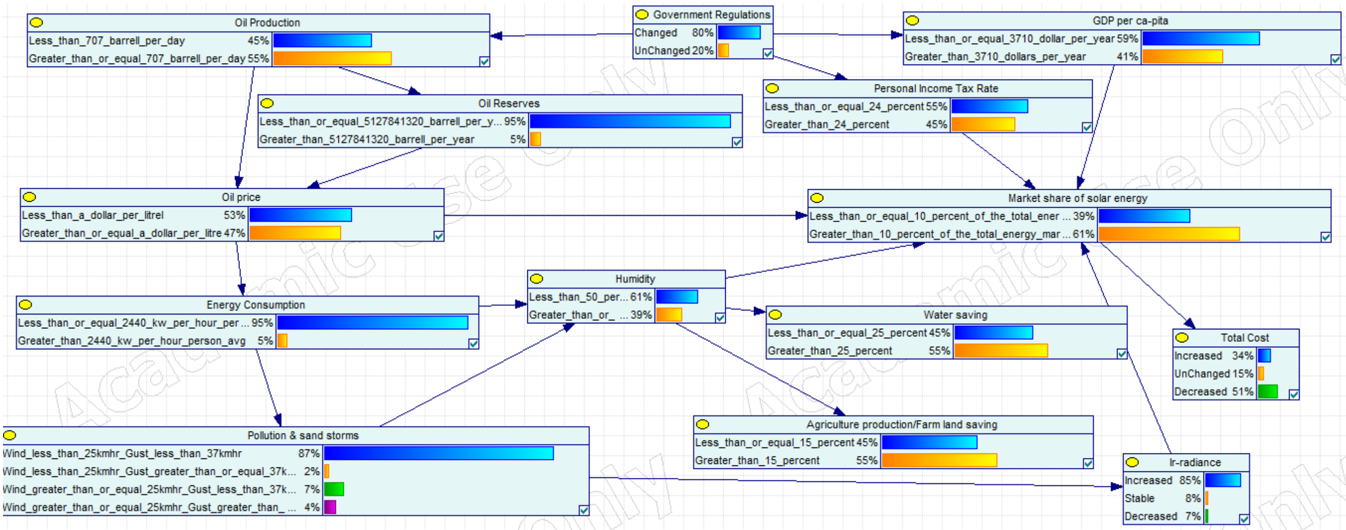


Fig. 11 Updated conditional probabilities in our BN using clustering algorithm

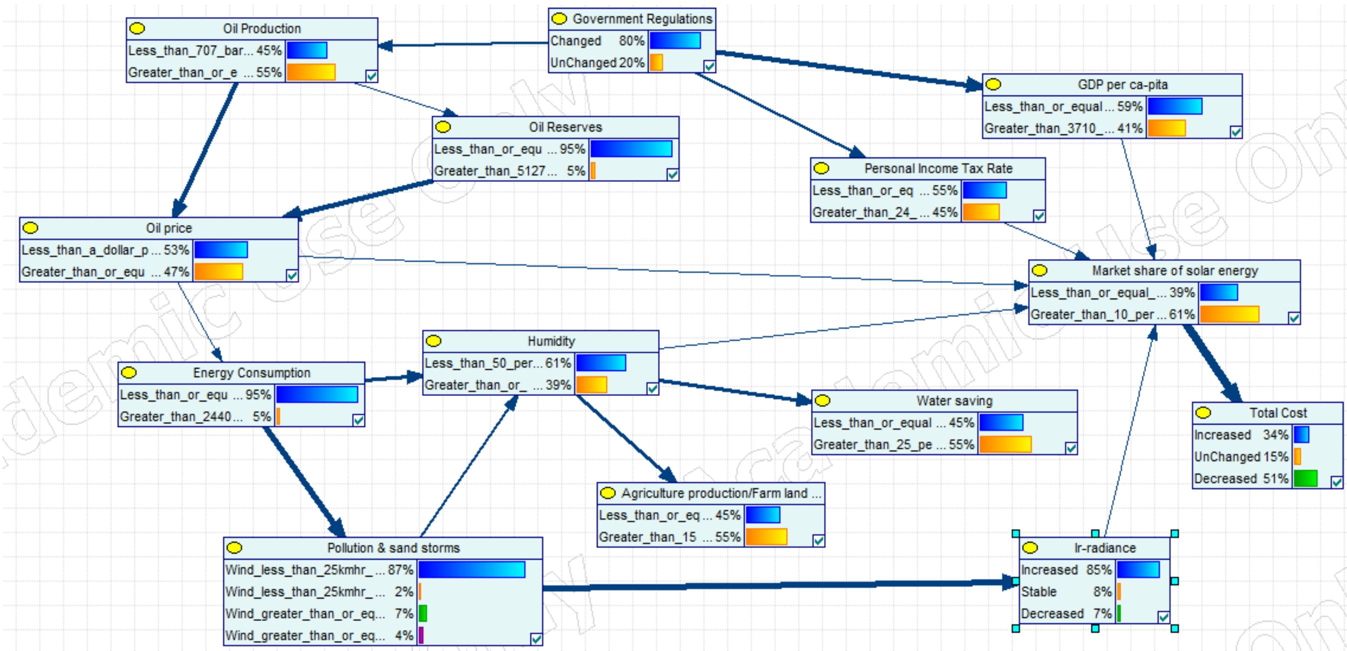


Fig. 12 Strength of Influences of Arcs

C. 50/50 Scenario

Out of the 96 implemented scenarios, there were 6 scenarios that ended up with 50/50 probability between the two states of the Market share of the solar energy. We could say that 4 of these scenarios are much better, as their impact on the *Agriculture Production and Water Saving* factors lead to 50/50 probability between their states, and that' when:

- $Energy\ Consumption > 2440\ kw/hr/person$
- $GDP\ per\ Capita\ income\ \leq 3710\$/year$
- $Personal\ Income\ Tax\ Rate \leq 24\%$
- *Pollution & Sand Storms is in any state*
- *Irradiance is stable*

The other 2 scenarios lead to 70% probability of the Agriculture Production being ≤15%, and the Water saving being $\leq$ 25%.

D. Most Likely Scenario

Out of the 96 implemented scenarios, the most likely scenarios that happened are those scenarios that resulted in 60% of the Market share of the solar energy being in the state of $\leq$ 10% of the total energy market and 40% being in the state of $>$ 10% of the total energy market. They took place 13 times out of the 96, which is about 13.5%.

Secondly, the scenarios that resulted in 35% of the Market share of the solar energy being in the state of $\leq 10\%$ *of the total energy market* and 65% being in the state of $> 10\%$ *of the total energy market*. They took place 9 times out of the 96, which is about 9.4%.

Using the sensitivity analysis add-in in GeNIe; by indicating the *Market share of solar energy, Total costs, Agriculture production, and Water saving*, as our target nodes and applying the technique, Fig. 13 shows the BN with the nodes colored based on the importance. Nodes colored in red contain factors that are the most important in calculating the posterior probabilities of the target nodes.

Fig. 14 illustrates the 10 most sensitive factors for the market share of solar energy node with a selected state, greater than 10% of the total energy market, sorted in a descending order. Each bar in the chart represents the range of change in the targeted state as the other factors changes (either 10% down or 10% up), while the color of the bar represents the direction of the factors' change, red means negative and green means positive.

From Figs. 13 and 14, we noticed that the most effective 5 factors on the target nodes are: *Humidity, Pollution & Sand Storms, Irradiance, GDP per Capita income, and Personal Income Tax Rate*, which differs only from the previous manually scenario analysis, that we have *Humidity*, instead of *Energy Consumption*.

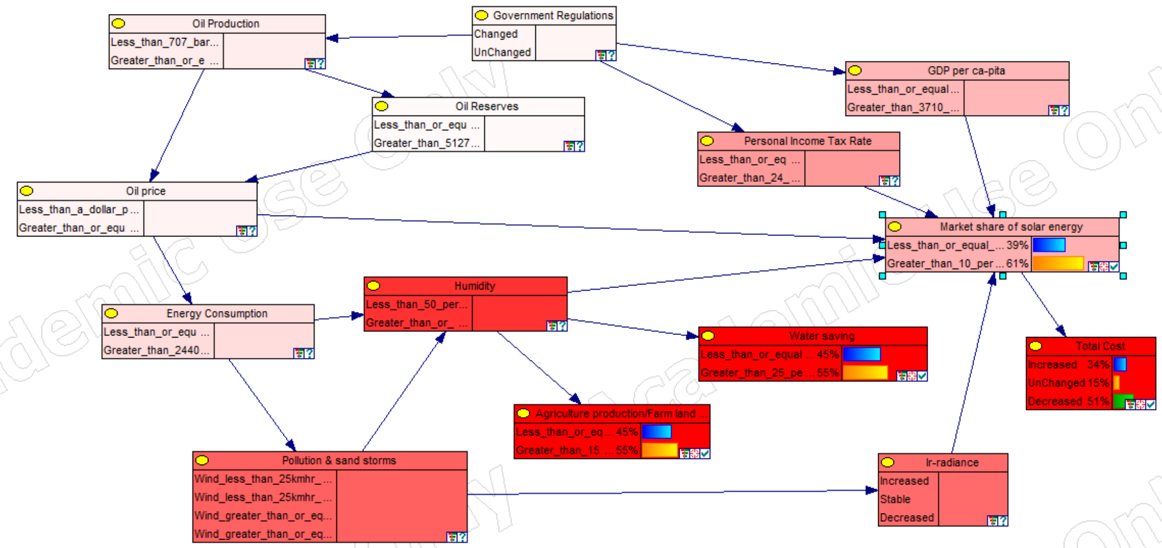


Fig. 13 Importance of nodes based on Sensitivity analysis in GeNIe

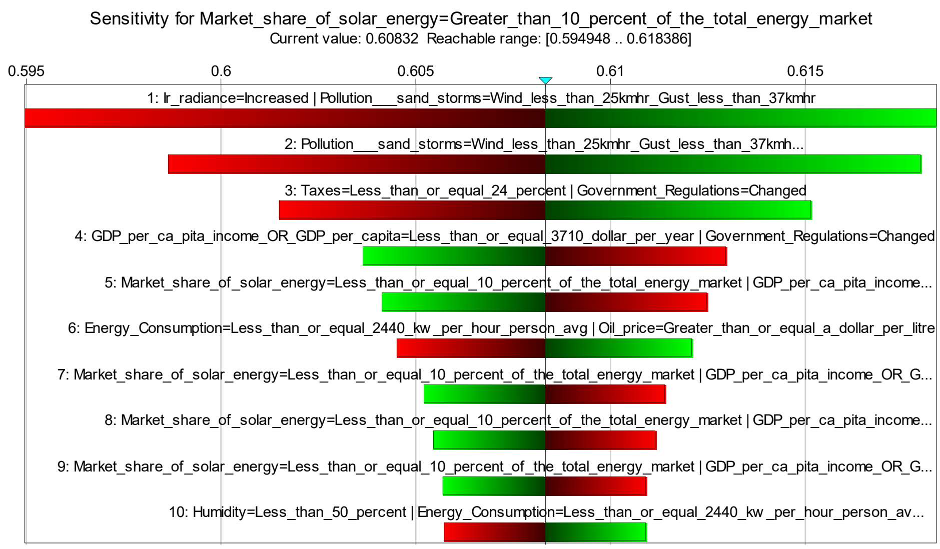


Fig. 14 The top 10 effective factors using sensitivity analysis

As in any decision problem, there are usually variables whose values can be determined or even controlled by the decision maker in the process, they are called decision variables. In our model, we can use the *Government Regulations* as our controllable (decision) variable, since it is mostly only variable that can be controlled somehow. So, by checking the impact of each state of the Government Regulations node on our 4 target nodes, there was not a major difference which proves that our target nodes are mainly depending on the other uncontrollable influential factors that can be observed with different scenarios.

In our BN model, as we are observing the *Agriculture Production* and *Water Saving* as environmental target nodes; so, by checking their best scenario that will lead to both of them having 100% probability of being in their best state, it was found that in terms of environmentally wise it is better for *Energy Consumption* to be ≤ 2440, *Humidity* to be < 50%, *Wind* to be < 25 km/hour, and *Gust* to be < 37 km/hour. This scenario will result in 62% probability of *Market Share of Solar Energy* to be greater than 10% of the total energy market, and thus 52% that *Total Costs* will decrease.

In our BN model, we have *Personal Income Tax Rate* and *GDP Per Capita* acting as the economical nodes; so, by checking their best scenario that will lead to both of them having 100% probability of being in their best state, it was found that in terms of economically wise it is better for *Oil Production* to be ≥ 707 barrel/day, *Oil Price* to be ≥ dollar/ liter, and *Government Regulations* to change. This scenario will result in 74% probability of *Market Share of Solar Energy* to be greater than 10% of the total energy market, and thus 61% that *Total Costs* will decrease.

VII. Dynamic BN

As for the dynamic BN technique in GeNIe, it works by enabling a special construct in the graph view that is called the Temporal Plate. This plate divides the graph view into four areas as shown in Fig. 15:

- Temporal Plate: It is considered to be the main part representing the dynamic model, it contains the dynamic nodes and also shows the number of time steps for which inference will be performed. In our case; *Humidity* and *Pollution & Sand Storms* will be the dynamic nodes.
- Init Conditions: It is where the anchor nodes are placed, which are nodes that have one or more children nodes in the temporal plate. In our case; the anchor node is the *Energy Consumption*.
- Term Conditions: It is where the terminal nodes are placed, which are nodes that have one or more parents in the temporal plate. In our case; the terminal nodes are: *Market share of solar energy, Total costs, Agriculture production, Water saving, and Irradiance*.
- Contemporals: The part with the static nodes, the area of the network view outside the temporal plate window.

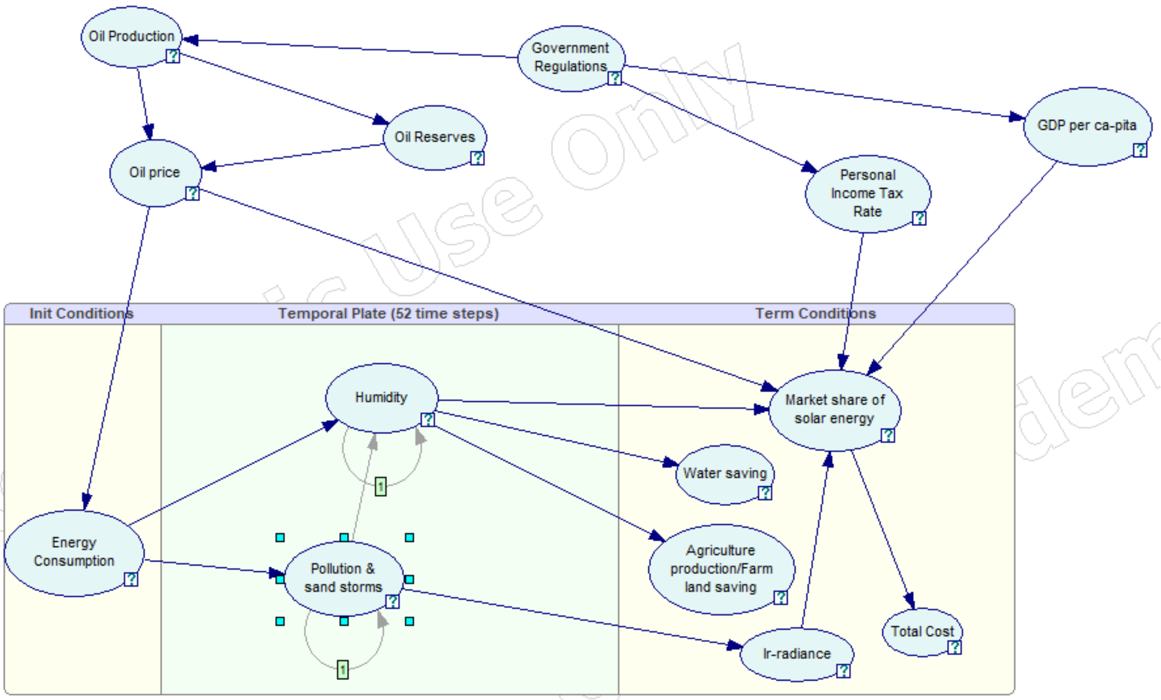


Fig. 15 Temporal Plate of our DBN in GeNIe

An arc was drawn from each dynamic node (*Humidity*, *Pollution & Sand Storms*) to itself, representing the impact of each state of the node in a given point of time, on being in another state in the next point of time. In this case, we chose order 1, which indicates that the impact has a delay of 1 week.

Concerning the number of time steps in the temporal plate, it was chosen to be 52 time steps, as we are dealing with weekly data which is needed to be used in calculating the percentage at the end of the year.

As for the dynamic probabilities of each node, the view now is different from that of the static nodes; the previously added conditional probabilities are now saved as those at (t = 0), while at (t = 1) will be depending on the previously state of the node. Using Python and the historical weekly data of *Humidity*, and *Pollution & Sand Storms*, the probabilities were calculated between each state and its previous one.

Similar to normal BN; In a DBN, we can calculate the impact of observing specific evidence on the target nodes. GeNIe converts the DBN into a BN using the unrolling technique according to the specified number of time steps (52 time steps in our case) as illustrated in Fig. 16 and updates the probabilities normally with the selected updating algorithm.

By assuming random evidence for the 52 time steps in our Humidity dynamic node using random library in Python and updating the beliefs (conditional probabilities) in our DBN using the clustering algorithm, Figs. 17-19 show the marginal posterior probability distribution as a function of time for the dynamic node (*Pollution & Sand Storms*), the posterior marginal beliefs as a contour plot with probabilities displayed with gradient colors, and the posterior probabilities as a time series plot (curved line for every state of the node) respectively.

Besides dealing with *Pollution & Sand Storms* as a dynamic target node in our DBN, we also considered the *Market share of solar energy*, *Total Costs*, *Agriculture Production*, and *Water Saving* as target nodes, to check their resulted updated probabilities after entering the evidence of *Humidity*. The results were as follows: *Market share of solar energy* had 63% probability of being > 10% of the total energy market, *Total Costs* had 52% probability of decreasing, while *Agriculture Production* > 15% and *Water Saving* > 25% with 70% probability.

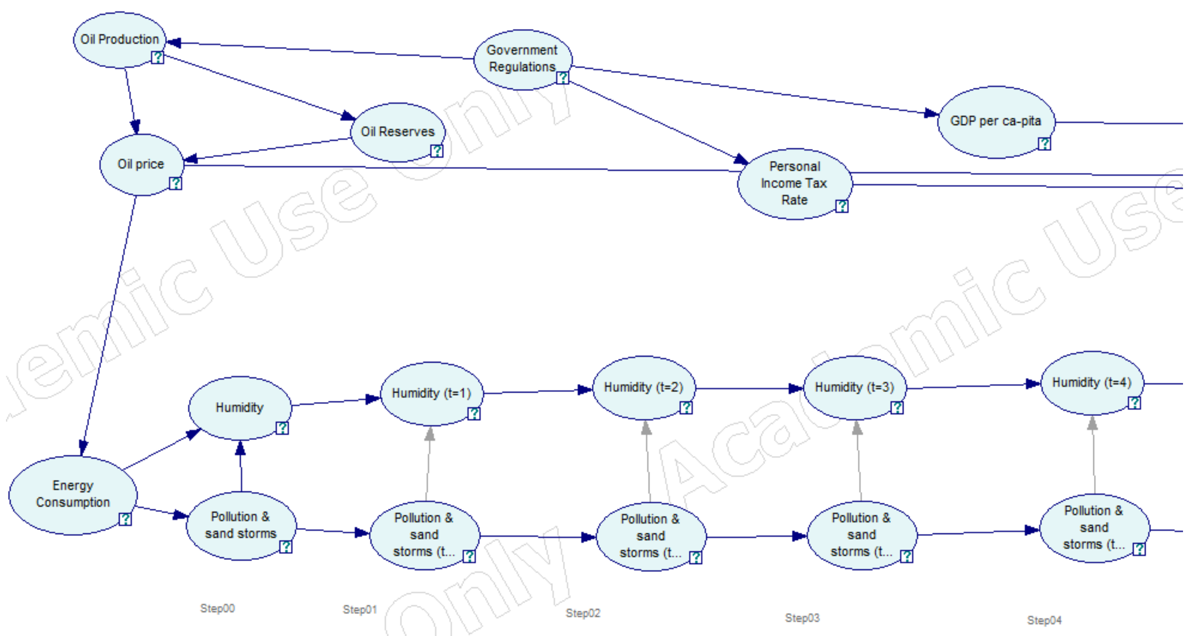
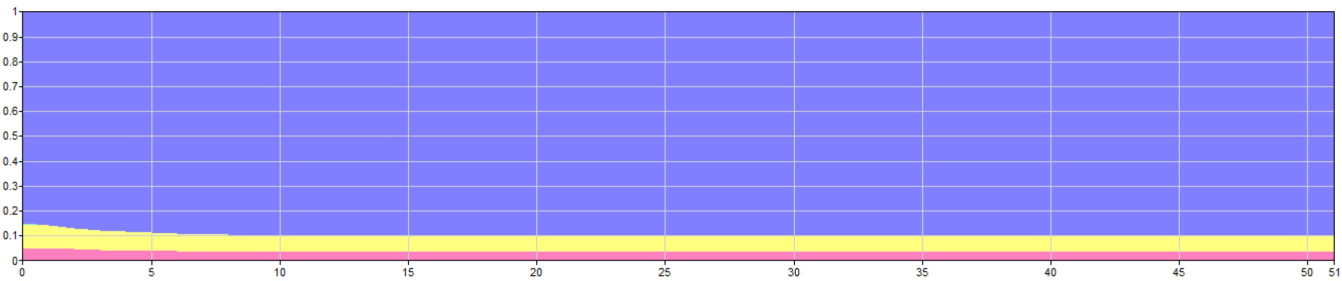


Fig. 16 Part of the unrolled DBN



Time	0	1	2	3	4
Wind_less_than_25kmhr_Gust_less_than_37kmhr	0.84769465	0.85853802	0.86862608	0.87640775	0.88244658
Wind_less_than_25kmhr_Gust_greater_than_or_equal_37kmhr	0.0096254...	0.0009402...	0.0010673...	0.0009458...	0.0008718...
Wind_greater_than_or_equal_25kmhr_Gust_less_than_37kmhr	0.095667231	0.087155249	0.083012713	0.079052562	0.075701971
Wind_greater_than_or_equal_25kmhr_Gust_greater_than_or_equal_37kmhr	0.047012659	0.05336648	0.047293878	0.043593809	0.040979572

Fig. 17 Marginal posterior probability distribution of the Pollution & Sand Storms (as a function of time) after updating beliefs

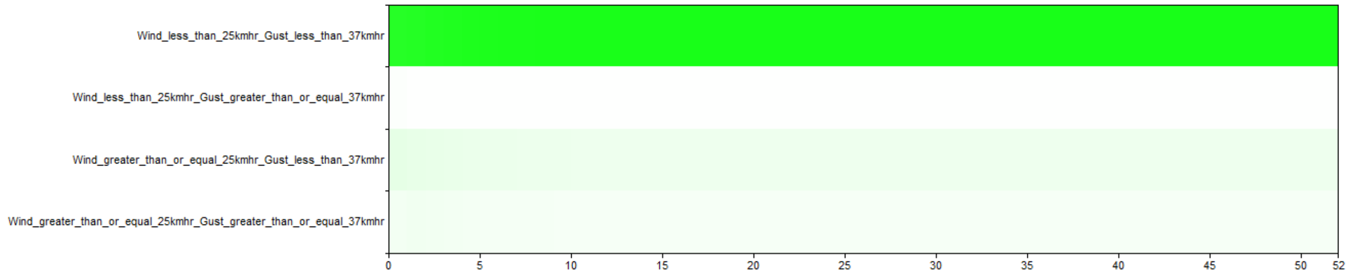
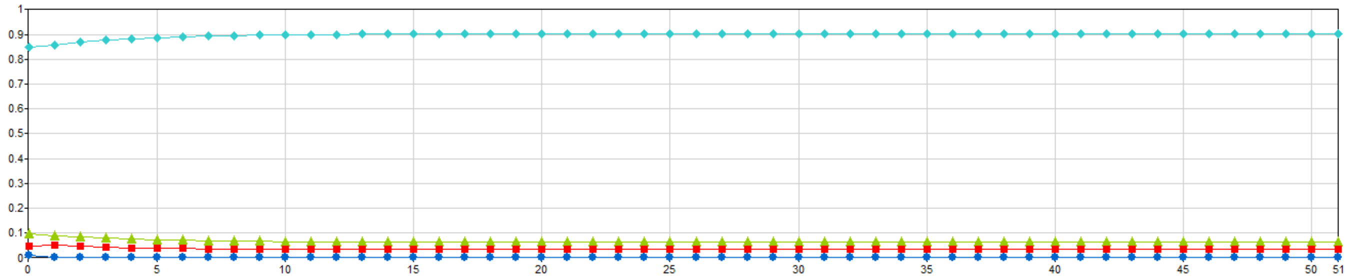


Fig. 18 Posterior marginal beliefs distribution of the Pollution & Sand Storms as a contour plot with probabilities displayed with gradient colors



Time	0	1	2	3	4
Wind_less_than_25kmhr_Gust_less_than_37kmhr	0.84769465	0.85853802	0.86862608	0.87640775	0.88244658
Wind_less_than_25kmhr_Gust_greater_than_or_equal_37kmhr	0.0096254...	0.0009402...	0.0010673...	0.0009458...	0.0008718...
Wind_greater_than_or_equal_25kmhr_Gust_less_than_37kmhr	0.095667231	0.087155249	0.083012713	0.079052562	0.075701971
Wind_greater_than_or_equal_25kmhr_Gust_greater_than_or_equal_37kmhr	0.047012659	0.05336648	0.047293878	0.043593809	0.040979572

Fig. 19 The posterior probabilities of the Pollution & Sand Storms as a time series plot (curved line for every state of the node)

VIII. Conclusion & Future Work

BN have been addressed many times in the field of energy, due to its accurate and depending results and have become a trust worthy tool in different applications. In this study continuing on [21], we have focused on deriving the states and conditional probabilities of the nodes in our constructed BN rather than collecting them, they were based on historical patterns and the use of statistical approaches such as Z-distribution and Chebyshev's inequality theorem. A few nodes were categorized qualitatively in the model, due to the lack and hardness of collecting their relative data.

Using different analysis techniques, the most influential factors that might affect the development of the market share of solar energy compared to the total energy market were determined to be *Energy Consumption*, *Pollution & Sand Storms*, *Irradiance*, *GDP per Capita*, *Personal Income Tax Rate*, and *Humidity*. Under the conditions of the best scenario, the market share of solar energy in Egypt by 2035 can be expected to exceed 10% of the total energy market by 85% probability, while in the worst scenario, this total probability drops to only 24%.

Applying policy analysis on our controllable decision factor which is *Government Regulations* showed that it does not highly affect the target nodes. Also, checking the best environmental and economical scenarios, showed which factors would affect them positively with how much changes in their probabilities.

Some factors later on were chosen to be dynamic nodes, due to their dependency on time for calculating their states' probabilities, which are: *Humidity*, and *Pollution & Sand Storms*. This led to converting the previously constructed BN into a Dynamic BN and applying additional analysis to check more accurate results on the chosen target nodes.

The target nodes that we focused on were chosen to be the *Market share of solar energy*, the *total costs* of switching to use solar energy; in order to prove that: the more we use renewable energy, the less its total costs will be. Also, we wanted to observe the impact of different scenarios on environmental conditions as *Agriculture Production* and *Water Saving*.

As for all the programming Python codes that were developed in this paper, they are uploaded on a private Github repository [33].

This study could be further extended to reach more experts in the field of energy, ensuring the analysis' results to be more robust and accurate. Future work may also concerns managing more detailed historical data for more accurate forecasted values, and collecting data for those factors that we did not have the chance to reach their data and were categorized qualitatively. Furthermore, with the help of some experts and having utility values for the states of the nodes in the constructed Dynamic BN; this model could use influence diagrams or decision trees, for strong decision making capabilities. Finally, and most importantly, our constructed model could be used to forecast the market share of solar energy in any time with the right probabilities, also it could be modified easily to cover any other type of renewable energy and check its future impact on the factors.

APPENDIX

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997

Fig. 20 Z-standard normal distribution table

REFERENCES

[1] "Global Energy Consumption" https://ourworldindata.org/energy-production-consumption.

[2] B. Sørensen, "Life-cycle analysis of renewable energy systems," Renew. Energy, vol. 5, no. 5–8, pp. 1270–1277, Aug. 1994, doi: 10.1016/0960-1481(94)90161-9.

[3] A. K. Akella, R. P. Saini, and M. P. Sharma, "Social, economical and environmental impacts of renewable energy systems," *Renew. Energy*, vol. 34, no. 2, pp. 390–396, Feb. 2009, doi: 10.1016/j.renene.2008.05.002.

[4] European Commission DG Environment News Alert Service, "Relationships between energy consumption and economic growth investigated," *Renew. Sustain. Energy Rev.*, vol. 16: 5718–5, no. 312, p. 2013, 2013, (Online). Available: http://ec.europa.eu/environment/integration/research/newsalert/pdf/312na6_en.pdf.

[5] S. Shafiei and R. A. Salim, "Non-renewable and renewable energy consumption and CO2 emissions in OECD countries: A comparative analysis," *Energy Policy*, vol. 66, pp. 547–556, Mar. 2014, doi: 10.1016/j.enpol.2013.10.064.

[6] N. J. Sheikh, D. F. Kocaoglu, and L. Lutzenhiser, "Social and political impacts of renewable energy: Literature review," *Technol. Forecast. Soc. Change*, vol. 108, no. May, pp. 102–110, 2016, doi: 10.1016/j.techfore.2016.04.022.

[7] M. Borunda, O. A. Jaramillo, A. Reyes, and P. H. Ibargüengoytia, "Bayesian networks in renewable energy systems: A bibliographical survey," *Renew. Sustain. Energy Rev.*, vol. 62, no. June, pp. 32–45, 2016, doi: 10.1016/j.rser.2016.04.030.

[8] D. Cinar and G. Kayakutlu, "Scenario analysis using Bayesian networks: A case study in energy sector," Knowledge-Based Syst., vol. 23, no. 3, pp. 267–276, 2010, doi: 10.1016/j.knosys.2010.01.009.

[9] G. de la Torre-Gea, G. M. Soto-Zarazúa, R. G. Guevara-González, and E. Rico-García, "Bayesian networks for defining relationships among climate factors," Int. J. Phys. Sci., vol. 6, no. 18, pp. 4412–4418, 2011, doi: 10.5897/IJPS11.631.

[10] Science for Environment Policy, Relationships between energy consumption and economic growth investigated, https://ec.europa.eu/environment/integration/research/newsalert/pdf/312na6rss_en.pdf

[11] K. Kim, H. Park, and H. Kim, "Real options analysis for renewable energy investment decisions in developing countries," *Renewable and Sustainable Energy Reviews*, vol. 75. Elsevier Ltd, pp. 918–926, Aug. 01, 2017, doi: 10.1016/j.rser.2016.11.073.

[12] A. Petrillo, F. De Felice, E. Jannelli, C. Autorino, M. Minutillo, and A. L. Lavadera, "Life cycle assessment (LCA) and life cycle cost (LCC) analysis model for a stand-alone hybrid renewable energy system," *Renew. Energy*, vol. 95, pp. 337–355, 2016, doi: 10.1016/j.renene.2016.04.027

[13] Z. Liu, Y. Liu, D. Zhang, B. Cai, and C. Zheng, "Fault diagnosis for a solar assisted heat pump system under incomplete data and expert knowledge," *Energy*, vol. 87, pp. 41–48, 2015, doi: 10.1016/j.energy.2015.04.090

[14] G. Li and J. Shi, "Applications of Bayesian methods in wind energy conversion systems," *Renew. Energy*, vol. 43, pp. 1–8, 2012, doi: 10.1016/j.renene.2011.12.006

[15] J. A. Carta, S. Velázquez, and J. M. Matías, "Use of Bayesian networks classifiers for long-term mean wind turbine energy output estimation at a potential wind energy conversion site," *Energy Convers. Manag.*, vol. 52, no. 2, pp. 1137–1149, 2011, doi: 10.1016/j.enconman.2010.09.008

[16] P. H. Ibargüengoytia *et al.*, "Wind power forecasting using dynamic Bayesian models," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8857, pp. 184–197, 2014, doi: 10.1007/978-3-319-13650-9_17

[17] A. S. Cofino, R. Cano, C. Sordo, and J. M. Gutierrez, "Bayesian Networks for Probabilistic Weather Prediction," *Proc. 15th Eur. Conf. Artif. Intell.*, vol. 700, pp. 695–700, 2002, (Online). Available: http://grupos.unican.es/ai/meteo/articulos/ECAI2002.pdf

[18] S. Hellman, A. McGovern, and M. Xue, "Learning ensembles of continuous Bayesian networks: An application to rainfall prediction," *Proc. - 2012 Conf. Intell. Data Understanding, CIDU 2012*, pp. 112–117, 2012, doi: 10.1109/CIDU.2012.6382191

[19] G. L. L. L. Dong, W. Zhou; Pei Zhang, "Short-term photovoltaic output forecast based on Dynamic Bayesian Network theory," *Proc. Chinese Soc. Electr. Eng.*, vol. 33, pp. 38–45, 2013.

[20] D. Gambelli, F. Alberti, F. Solfanelli, D. Vairo, and R. Zanoli, "Third generation algae biofuels in Italy by 2030: A scenario analysis using Bayesian networks," *Energy Policy*, vol. 103, no. April 2016, pp. 165–178, 2017, doi: 10.1016/j.enpol.2017.01.013.

[21] R. H. El-bidweihy, H. M. Abdelsalam, and I. A. El-khodary, "Constructing a Bayesian Network for Solar Energy in Egypt Using Life Cycle Analysis and Machine Learning Algorithms," *Int. J. Energy Power Eng.*, vol. 14, no. 10, pp. 310–319, 2020.

[22] "Empirical Rule – Financial Analysis," Investopedia. https://www.investopedia.com/terms/e/empirical-rule.asp#:~:text=The%20Empirical%20Rule%20states%20that,standard%20deviations%20from%20the%20mean.

[23] J. Mitchel, "Notes on the Chebyshev Inequality," vol. 318, no. X, p. 318.

[24] "Chebyshev Inequality," *ScienceDirect*. https://www.sciencedirect.com/topics/mathematics/chebyshev-inequality.

[25] U. Kjærulff and L. C. van der Gaag, "Making Sensitivity Analysis Computationally Efficient," no. March 1999, 2013, (Online). Available: http://arxiv.org/abs/1301.3868.

[26] "Egypt Gasoline Prices," Trading Economics. https://tradingeconomics.com/egypt/gasoline-prices.

[27] "Egypt GDP per capita," Trading Economics. https://tradingeconomics.com/egypt/gdp-per-capita.

[28] "Egypt Crude Oil: Production," CEIC Data. https://www.ceicdata.com/en/indicator/egypt/crude-oil-production.

[29] "Egypt Oil," WorldOMeter. https://www.worldometers.info/oil/egypt-oil/.

[30] "Egypt Energy Situation," EnergyPedia. https://energypedia.info/wiki/Egypt_Energy_Situation?fbclid=IwAR3U12y8Z-lrMVTR3hqp7_z5krb7LpVmDOuMYghCOnc9ol1W1IPrRFYUwWc#Production.C2.A0.

[31] "Egypt Personal Income Tax Rate," Trading Economics. https://tradingeconomics.com/egypt/personal-income-tax-rate.

[32] "Cairo Historical Weather," WorldWeatherOnline. https://www.worldweatheronline.com/cairo-weather-history/al-qahirah/eg.aspx.

[33] Rawaa El Bidweihy, "GithubRepository." https://github.com/RawaaElBidweihy/Masters-THESIS.