Databases in Cloud – Solutions for Developing Renewable Energy Informatics Systems

Adela BÂRA[1], Iuliana ȘIMONCA (BOTHA)[2], Anda BELCIU (VELICANU)[3]
[1,2,3]The Bucharest University of Economic Studies
{bara.adela, iuliana.botha, anda.velicanu}@ie.ase.ro

The paper presents the data model of a decision support prototype developed for generation monitoring, forecasting and advanced analysis in the renewable energy filed. The solutions considered for developing this system include databases in cloud, XML integration, spatial data representation and multidimensional modeling. This material shows the advantages of Cloud databases and spatial data representation and their implementation in Oracle Database 12 c. Also, it contains a data integration part and a multidimensional analysis. The presentation of output data is made using dashboards.

Keywords:*Renewable Energy, Data Model, Databases in Cloud, Spatial Data, Data Warehouse*

1 Introduction

The main objective of the research project Intelligent System for prediction, analysis and monitoring of performance indicators of technological and business processes in the field of renewable energies (SIPAMER) is to develop a prototype of a decision support system for energy producers form renewable resources like wind and solar sources. The prototype's architecture, more detailed in [1] consists of three layers as shown in **Fig. 1**.:

- data layer that collects data from measuring devices installed in the wind/photovoltaic power plants and from SCADA/EMS systems used for monitoring the energy power produced by each turbine or photovoltaic panel;
- models layer containing the data mining algorithms with artificial neural networks (ANN) for energy production forecast and also with multidimensional OLAP models for decision support;
- interface layer with business intelligence capabilities that enables advanced and interactive analysis of key performance indicators through dashboards and dynamic reporting tools.

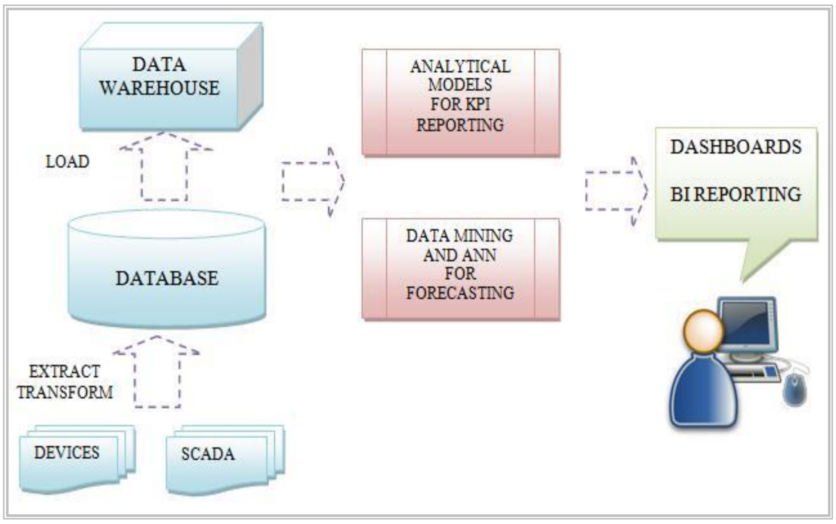

Fig. 1. SIPAMER's architecture

The producers' requirements in terms of real time analysis and accurate forecasting regarding energy production are implemented in two modules accessible through a cloud computing platform. We set up a database instance dedicated to each producer with Business Intelligence reporting tools accessible as a service via secured connection. The database instance is Oracle Database 12c that allows advanced XML and JSON integration, spatial data representation, multidimensional and OLAP analytics.

2 Cloud Databases

No matter what the business is about and what software it uses, most likely it will be faced with dealing with the concepts of Big Data, Cloud, Smart and Business Rules [2], [3].

Cloud computing is a paradigm that allows access through network to a pool of resources dynamically configurable, which are available to everyone as services. These resources can be quickly found and can be provided with minimal effort through a service provider interaction. This model's first operating principle refers to resource availability.

This innovative paradigm and also the large volumes of different types of data, which involve new data storage requirements, all have led to the cloud databases.

The database systems have as foundation the data level (which can be received as a Big Data source), represented by the database. Cloud computing offers an efficient manner of processing and managing large volumes of data, since the big data philosophy involves gathering and processing massive data sets which are so large that conventional database systems and software tools have failed to manage them.

Oracle 12c is designed especially for cloud environment. It offers a Cloud solution, which brings new improved features like providing the database as a service in Cloud, optimizations, integration and analysis of Big Data, security and so on [4]..

Cloud databases have a number of advantages, among which saving the storing space, by using the cloud storing, is the most important. Of course, this also comes with some protection risks (integrity and security) depending on a third party's measures of security. However, if the private company is not specialized in providing database protection, probably the cloud providers do it better then than average IT companies, having a background on the subject. Another impediment can be the money spent on storing data in a Cloud database. However, they became more and more price accessible and will make it worth, by providing specialized hardware, like database machines. This is how personal hardware resources aren't a problem anymore. The only problem remains the internet dependency, which necessitates a good speed of traffic.

3 Data integration

Data layer integrates data from SCADA/EMS connectors and flat files from measuring devices that transmit temperature, atmospheric pressure, wind speed and direction, humidity and solar radiation. The integration process involves routines that extract and transform flat files into XML pattern files and import data into relational tables represented as XSD schemas. For example, the measuring data for photovoltaic panels are modeled into DATA_METEO_PV XSD schema as shown in **Fig. 2.**

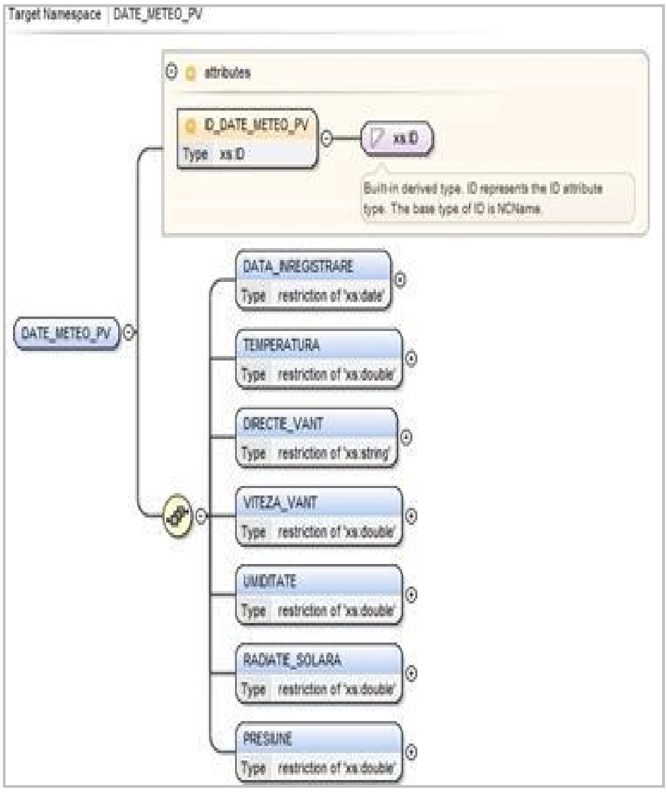

Fig. 2. XSD schema for photovoltaic panels' measuring data

The advantage of the model is that the import process is fast and flexible [5]; any change can be easily captured inside the XML patterns.
The XSD Schema for photovoltaic panels' measuring data can be implemented using specific sintax, as follows:

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSc
hema" xmlns="date_meteo_pv"
targetNamespace="date_meteo_pv">
<xs:element name=" date_meteo_pv ">
<xs:complexType>
<xs:sequence>
<xs:element name="data_inregistrare">
<xs:simpleType>
<xs:restriction base="xs:date">
</xs:restriction>
</xs:simpleType>
</xs:element>
<xs:element name="temperatura">
<xs:simpleType>
<xs:restriction base="xs:double">
</xs:restriction>
</xs:simpleType>
</xs:element>
...........
<xs:element name="presiune">
<xs:simpleType>
<xs:restriction base="xs:double">
</xs:restriction>
</xs:simpleType>
</xs:element>
</xs:sequence>
<xs:attribute name="id_pv"
type="xs:integer" use="required"/>
</xs:complexType>
</xs:element>
</xs:schema>
```

Further, we used some mapping algorithms, in order to transform XSD Schema into database objects (tables), so that uploaded data can be validated in compliance with the schema.
The database schema for each producer's instance contains relational tables for photovoltaic panels, wind turbines, meteorological recordings, production, operation and maintenance data and transactions on energy market. We also set up a dedicated instance for public actors like energy market regulators (ANRE and OPCOM) or national TSO (Transelectrica SA). The instance provide a process that imports the mandatory reports of energy producers (**Fig. 3.**) into a relational schema, aggregate them and make data available for multidimensional analysis.

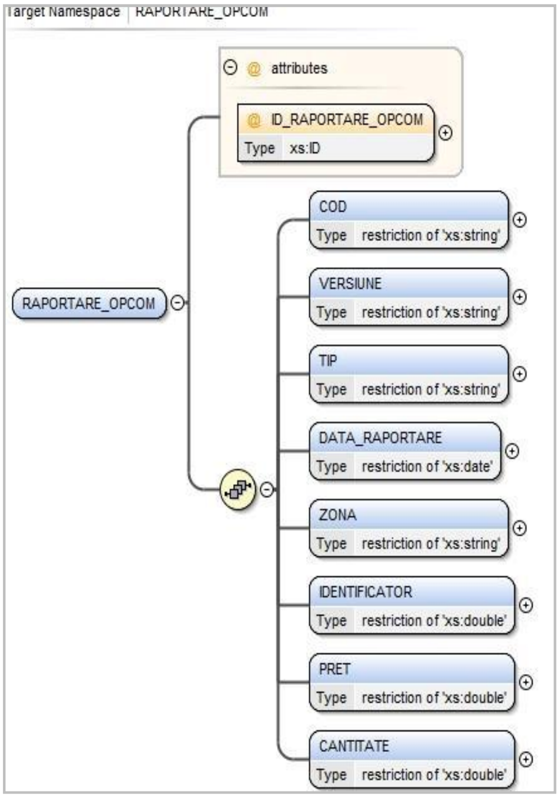

Fig. 3. XSD schema for OPCOM reporting

The XSD schema is implemented in Oracle Database 12c environment and enriched with spatial data representation in order to develop interactive charts and reports.

4 Spatial data representation

GIS (Geographic Information Systems) are based on spatial databases. Geographic Information Systems are capable of storing, manipulating, analyzing and displaying geographically referenced date towards a coordinate system (standard or user-defined), according to [6]. The most important Web GIS applications are described in [7] and the spatial analysis is the subject of [8] book.

Some domains (like transportation, energy, tourism, meteorology, etc.) need spatial representation of information on maps, by default. Because of the growing capacity of storing data in databases, in cloud and in cloud databases, we are allowed to store as much data as we need it. For example, storing all the coordinates that compound the perimeter of a lake was a problem of space a few time ago, but not anymore. Now, the only concern is how to make the analysis of spatial data go faster and return interactive results.

This is the problem that Oracle Database 12c solves quite well by grouping the spatial elements in a set of technologies that offers support starting from storing data, to analyzing and representing it. Oracle Location Technologies gather together Oracle Database Locator feature, Oracle Spatial and Graph Option and Oracle Fusin Middleware Map Viewer.

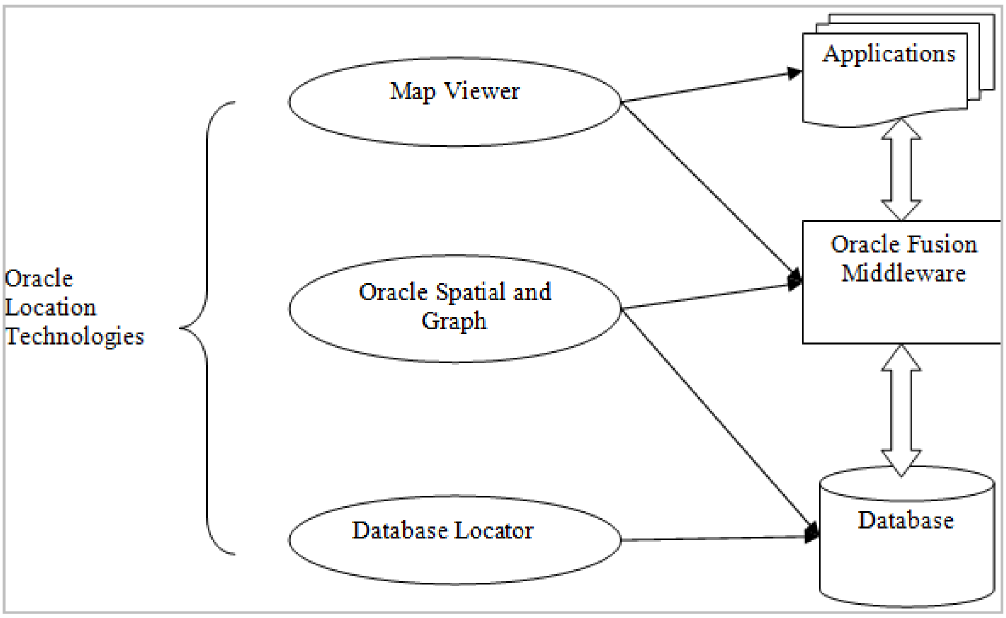

Fig. 4. Oracle Location Technologies

Fig. 4.points out the three components of Oracle Location Technologies and their interaction with the layers of an Oracle System. Database Locator allows storing spatial data in a specialized data type called SDO_GEOMETRY. Points, lines and polygons can be stored and queried using spatial operations.

Oracle Spatial and Graph (called Oracle Spatial until now) adds more spatial capabilities to Oracle Locator by using spatial and graph analysis. This type of analysis comes with the Database engine but it is used in applications through the middle layer of database system architecture. Also, the spatial queries based on spatial joins, touches, overlaps and other spatial relations, are run fifty times faster [6].

The Map Viewer develops web-based maps using data stored in Oracle Database with Spatial features. It is included in many Oracle products like Business Intelligence, SOA Suite, E-business Suite, JDeveloper and so on. Although Map Viewer it is considered to be part of the Oracle Fusion Middleware, by simplifying the complexity of an application, it can be said it provides the final application interface.

When working with SDO_GEOMETRY type in Oracle, the following steps must be followed:

- creating the spatial tables that use the spatial data
- inserting spatial data by using the SDO_GEOMETRY function, which has the following syntax:

```
  SDO_GEOMETRY(
    2003,
    NULL,
    NULL,
SDO_ELEM_INFO_ARRAY(1,sdo_etype,sdo_in
terpretation),
SDO_ORDINATE_ARRAY(x1,y1,x2,y2,…)  )
```

The combination of parameters sdo_etype, sdo_interpretation gives the type of stored object. For example, if the parameters have the values 1003, 3, they will define a rectangle. In bi-dimensional coordinates, the function SDO_ORDINATE_ARRAY() will get (xi,yi) pairs of parameters, depending on the type that was established through SDO_ELEM_INFO_ARRAY. For example, for a rectangle there will be given two pairs of coordinates.

- inserting a new record in the standard metadata view USER_SDO_GEOM_METADATA, which has the effect of defining the coordinate system, whether is a standard or a user-defined one;
- creating the spatial indexes is necessary for all spatial columns in all spatial tables. Spatial indexes are R-tree indexes and are used in requests that use spatial criteria like areas, perimeters, distance, union of spatial objects, etc.

The most common spatial operations are: create, manipulation (insert, update, delete), queries, map operations (overlapping layers, reducing the number of coordinates that define a spatial object, etc.), conversions. The spatial queries can be based on operations for determining the spatial relations or on spatial analysis operations (see [10]).

In Oracle, the spatial operators are implemented using certain functions like SDO_AREA, SDO_LENGTH, SDO_UNION, SDO_INTERSECTION, SDO_DIFFENECE, SDO_CENTROID, etc.

5 Multidimensional analysis

To develop multidimensional analysis model we designed snowflake schema and implemented in Oracle Warehouse Builder. The solution adopted for the prototype is based on cube based multidimensional model (MOLAP) implemented in the cloud computing platform. Thus, on the instance dedicated to energy producers we developed cubes for production analysis, forecasting and financial simulation and on the instance dedicated to public authorities we developed cubes for production forecasting analysis on regional/national level.

The steps for developing included the data cleansing, extract, transform and load process (ETL), mappings, validations and multidimensional objects generation: dimensions (producers, locations, regions, time, power plant, substations) and cubes (contracts, forecasting, production, weather, energy market transactions).

Data cleansing was mostly realized through data rules derived from the process of data profiling. In order to perform data profiling we need to follow some steps (detailed in [12]), which involve creating data profile objects, developing and configuring the profiles, loading the configuration parameters (such as: domains, functional dependencies, patterns etc.) and then executing the data profiling process.

We started by checking the data sources used, in order to establish common formats.

By following the steps mentioned above, we present an example of data profiling made in Oracle Warehouse Builder.

From the large SIPAMER's database schema we have selected the Investors table (see the structure in the Fig.5).

We can note the following inconsistencies in the data stored:

- tara: multiple versions of storing the same data;
- banca_investitor: some values are incorrect;
- telefon: multiple versions of storing the same data.

After running the data profile process, they will be displayed detailed statistics about the data. There will be also detected the domain values for Tara and Telefon columns. Based on these values, OWB will derive data rules which will be applied as data integrity restrictions in database tables.

The following PL/SQL function it is used to achieve country-level corrections:

```
begin
if upper (tara) in ('RO','ROU') then
return 'ROMANIA';
      elsif upper (tara) = 'SZ' then
return 'ELVETIA';
      elsif upper (tara) = 'UK' then
return 'ANGLIA';
      else                       return
upper(NVL(tara,'NECUNOSCUT'));
end if;
end;
```

Also, OWB will detect some incorrect values for the investors' bank. We need to select the cleanse strategy for the corrections, so we will choose to correct the data, according with the identified domain values.

We applied the profiling process for other tables such as bids, transactions, and beneficiary and also for the tables that gather data from meteorological devices. After the cleansing process we build in OWB the data warehouse schema (**Fig. 5.**) that is a snowflake schema with joins between dimensions and types of dimensions.

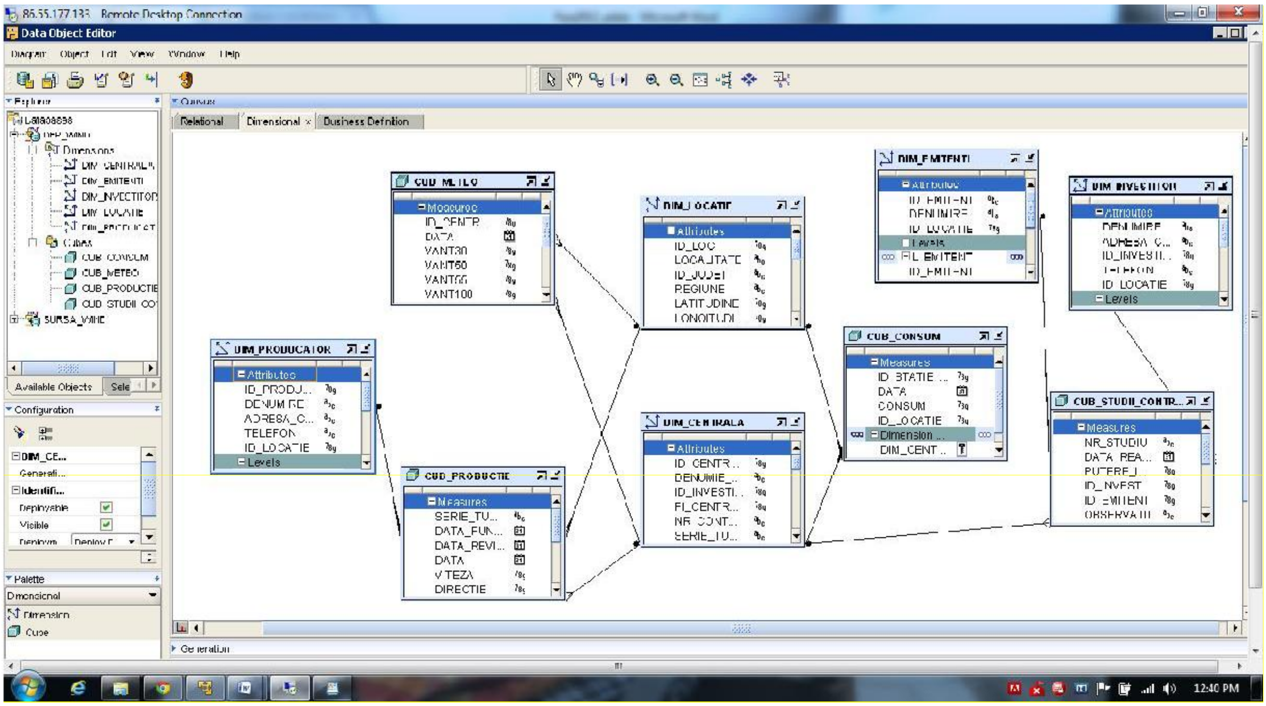

Fig. 5. Snowflake DW schema

In order to develop dashboards for multidimensional analysis, we build the following models in Oracle business Intelligence Suite: data model that contains the data locations and mappings, logical model based on the OWB data warehouse schema that contains dimensions, levels, facts and measures and presentation model that contains the attributes displayed in the dashboards and analytical reports.

6 Displaying data using dashboards

For each actor (energy producers and public authorities) we developed separate dashboards available through cloud services that offer a customized set of analytical reports regarding the production activities, forecasting for different time intervals (hours to 3-7 days), key performance indicators for energy market transactions (green certificates, intra-day, next day and balancing energy markets) as detailed in [7].

For public authorities the dashboard include a section dedicated to correlations between wind or photovoltaic power plants generation in the same region/area and aggregate regional or generation type forecasting.

Using the dashboards the decision factors have access to analytical reports from the system, such as:

- Reports for the contracts in progress in terms of installed capacity in wind or photovoltaic power plants by regions and by coupling stations;
- Reports to analyze production on different periods and in different regions;
- Reports to analyze consumption in different time periods and different points of measurement;
- Reports to analyze the evolution of weather conditions and energy predictions.

Oracle BI Dashboard integrates all reports made in a dashboard accessible by both computers and laptops and mobile devices like PDA or mobile phones. The dashboard will have the following sections: Home, Contracts, Consumption analysis, and Production analysis, Predictions (**Fig. 6.** and **Fig. 7.**).

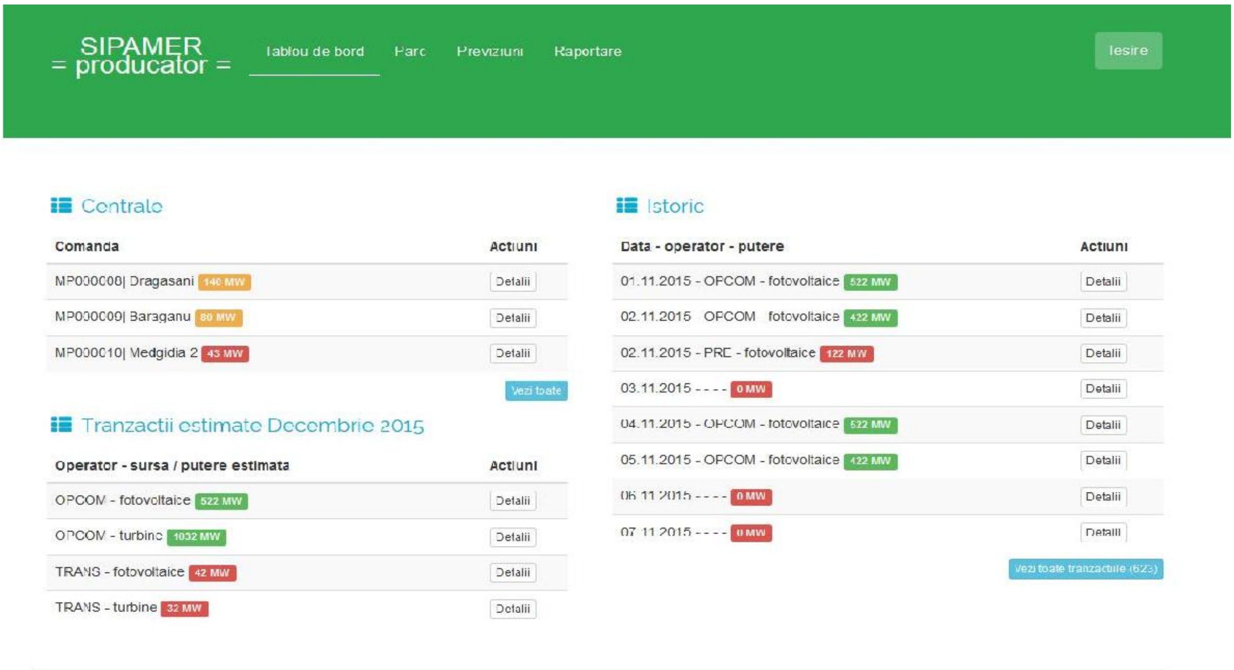

Fig. 6. Dashboard's first page

This method of presentation allows easy navigation and analyzing reports in a centralized manner, and also offers the opportunity to reconfigure reports (using OBI Answers) by modifying parameters or presentation style.

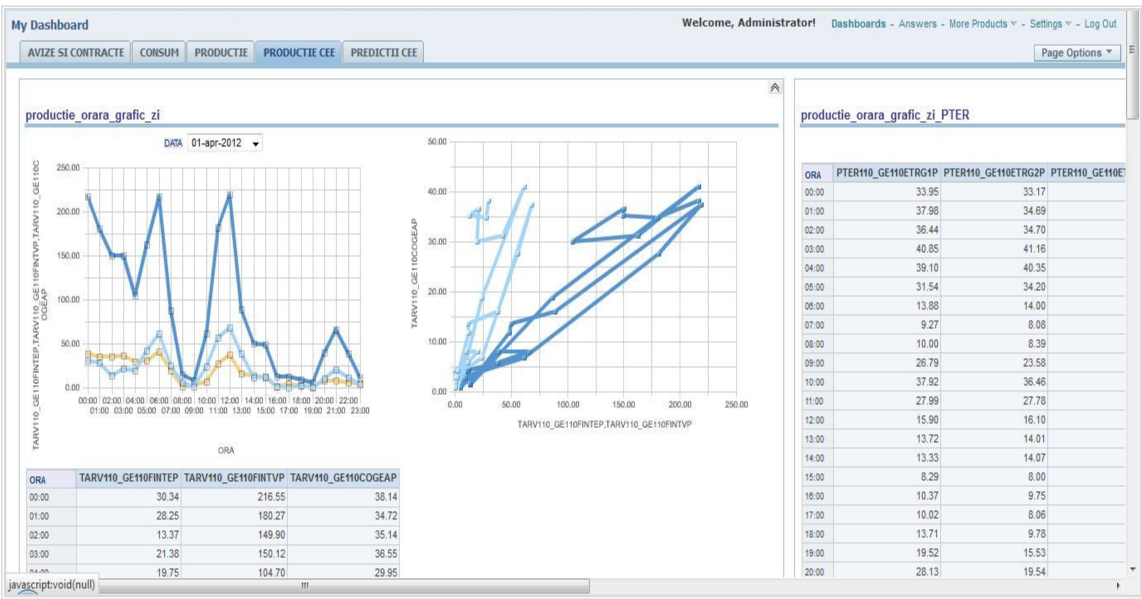

Fig. 7. Reports for analyzing the energy production, that are integrated into the dashboard

Fig. 7.presents the analysis section of the renewable production through interactive reports, charts that shows the degree of simultaneity in operation and aggregated reports for indicators regarding the production days / months and regions.

7 Conclusions

The SIPAMER prototype is a dedicated informatics solution for decision support in wind or photovoltaic power plants management both for producers and for public authorities. It is built on modules that include generation monitoring, forecasting and prediction the production based on meteorological factors and financial simulations regarding the energy market transactions. The cloud computing platform allows us to develop the prototype without infrastructure constrains, instantiating only the Oracle Database 12c for each actor and setting up the developing environment. The dashboards and monitoring tools are available to users through a portal based cloud service.

8Acknowledgment
This paper presents results of the research project: Intelligent System for prediction, analysis and monitoring of performance indicators of technological and business processes in the field of renewable energies (SIPAMER), research project, PNII –PCCA 2013, code 0996, no. 49/2014 funded by NASR.

References

[1] A. Bâra, S.V. Oprea, I. Lungu, G. Căruţaşu, C. P. Botezatu - Financial simulator based on wind power plants forecast, World Renewable Energy Congress 14, WREC XIV, University Politehnica of Bucharest, Bucharest – Romania, June 8 – 12, 2015, accepted for publishing in Open Access Journal of Physics: Conference Series (JPCS), 1742-6596

[2] S.V. Oprea, I. Lungu - Informatics Solutions for Smart Metering Systems Integration, RevistaInformaticăEconomică, vol. 19, nr. 4/2015 http://revistaie.ase.ro/content/76/03%20-%20Oprea,%20Lungu.pdf

[3] Alexandra Maria IoanaFlorea, Ana-Ramona Bologa, Vlad Diaconița, RazvanBologa - Streamlining Business Processes in Academia by Building and Manipulating a Business Rules Repository, The 14th International Conference on Informatics in Economy (IE 2015), ISSN 2284-7472, ISSN-L 2247-1480

[4] Oracle White Paper - Plug into the Cloud with Oracle Database 12c, October 2015, http://www.oracle.com/technetwork/database/plug-into-cloud-wp-12c-1896100.pdf

[5] H. Huiping Cao, Y. Qi, S. Candan, M.L. Sapino - XML Data Integration: Schema Extraction and Mapping. Advanced Applications and Structures in XML Processing: Label Streams, Semantics Utilization and Data Query Technologies, IGI Global, 2010

[6] Michael Brueckner, OrasaTetiwat, Use of Geographical Information Systems for Thailand, E-leader Bangkok, 2008.

[7] Pinde Fu, Jiulin Sun, Web GIS: Principles and Applications, Esri Press, 2010, ISBN:158948245X 9781589482456.

[8] Peter Rogerson, S Fotheringham, Spatial Analysis And GIS (Technical Issues in Geographic Information Systems), 296 pg., 2007, Taylor&Francis, ISBN-10: 0748401040.

[9] Oracle White Paper - Oracle Database 12c: An Introduction to Oracle's Location Technologies, September 2014, http://download.oracle.com/otndocs/products/spatial/pdf/12c/oraspatialandgraph_12c_wp_intro_to_location_technologies.pdf

[10] Belciu, A.: Baze de date spatiale in arhitecturaorientatapeservicii, ASE Publisher, Bucharest (2014)

[11] I. Lungu, G. Căruţaşu, A. Pîrjan, S.V. Oprea, A. Bâra - A two-step forecasting solution and upscaling technique for small size wind farms located on hilly areas in Romania, Studies in Informatics and Control Journal, vol 25/issue 1, 2016, ISSN 1220-1766

[12] Oracle Warehouse Builder Data Modeling, ETL, and Data Quality Guide 11g Release 2 (11.2), http://docs.oracle.com/cd/E11882_01/owb.112/e10935/data_profiling.htm

Adela BÂRA (b. October 11, 1978) is Professor at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. She has graduated the Faculty of Economic Cybernetics in 2002, holds a PhD diploma in Economics from 2007. She is the author of 9 books in the domain of economic informatics, over 40 published scientific papers and articles (among which over 20 articles are indexed in international databases, ISI proceedings, SCOPUS and 10 of them are ISI indexed). She participated as team member in 5 research projects and has gained as project manager 4 research grants, financed from national research programs. Domains of competence: Database systems, Data warehouses, OLAP and Business Intelligence, Executive Information Systems, Decision Support Systems, Data Mining.

Iuliana ȘIMONCA (BOTHA) is Associate Professor at the Economic Informatics Department at the Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies of Bucharest. She has graduated the Faculty of Cybernetics, Statistics and Economic Informatics in 2006, the Databases for Business Support master program organized by the Academy of Economic Studies of Bucharest in 2008 and she holds a PhD diploma in Economic Informatics from 2012. She is author/co-author of 9 books and 44 papers published in journals or conference proceedings (4 papers published in ISI indexed journals, 14 included in international databases and 9 published in ISI proceedings). She participated as team member in 8 research projects (among which two international research programs). Her scientific fields of interest include: Database Systems, Data Warehouses, Business Intelligence, Design of Economic Information Systems.

Anda BELCIU (VELICANU) has graduated the Faculty of Economic Cybernetics, Statistics and Informatics of the Bucharest University of Economic Studies, in 2008. She has a PhD in Economic Informatics and since October 2012 she is a Lecturer. She teaches Database, Database Management Systems and Software Packages seminars and courses at the Economic Cybernetics, Statistics and Informatics Faculty. Her scientific fields of interest and expertise include database systems, e-business, elearning, spatial databases.