Decision Support System Design for Photovoltaic Systems Operation and Maintenance by using Big Data Technologies

Oprea Simona Vasilica
Bâra Adela
Elefterescu Luminița
The Bucharest University of Economic Studies
simona.oprea@csie.ase.ro
bara.adela@ie.ase.ro

Abstract

Main goal of this paper is to present a decision support system (DSS) for operation and maintenance (O&M) of photovoltaic power (PV) systems which are integrated with battery systems. Such DSS essentially necessitates inclusion of Big Data analytics that will be utilized to maximize profit from power generation and consumption in a PV-battery integrated system.

Key words: Big Data, decision support systems, photovoltaic systems, operation and maintenance
J.E.L. classification: O13, C55, C82, C88

1. Introduction

One of the goals of the Energy Union (EU) strategy is to become the world leader in renewables (RES), the EU targets for 2030 define a more ambitious plan for RES integration. By the end of 2016, the worldwide total photovoltaic systems (PV) installed power is 303 GW, with 355 GW estimation for 2017 and 613 GW for 2020. Battery systems are among the promising technology that supports penetration of PV. Both operational and maintenance (O&M) of PV battery systems should be taken into account in the profit maximization problem. While operational optimization is focussing on short term (day-ahead, week-ahead), maintenance aspects focus more on mid- and long-term optimization (reduce downtime of the PV system, extend the PV lifetime, decrease cost of O&M, enhance safety and reduce risks).

Optimal O&M must strike a balance between maximising production and minimising cost, knowing that the O&M costs can be up to $40 kW/year. Also, a well-designed O&M can improve the average performance ratio of PV from 88% to 94% (NREL, 2016). O&M costs are between 30 and 70% of OPEX that represents 11-25% of the lifetime costs (Solar, 2016). A comprehensive planning and delivery of PV O&M reinforce confidence in the long-term performance and revenue capacity. O&M main activities consist of: monitoring, maintenance, reliability and management.

2. Literature review

Currently for PV O&M several types of software solutions are used, especially for monitoring large-scale PV power plants (Avisolar, 2017), (Greenbyte, 2017) based on data collected from SCADA systems. Well-known producers (Siemens, 2017), (ABB, 2017) provide more advanced solutions, including forecasting and asset management, accessible via cloud based web-services. Another complex solution that provides sensors integration and visual analytics for two types of PV, on ground and rooftops, is described in (Ecoaxis, 2017). These solutions are not developed on open and scalable platforms to be customize for different types of PV and does not currently integrate image processing or optimization algorithms to provide a complete support of O&M activities.

Several scientific papers approach the topics of big data management in case of PV O&M (Escobedo et al, 2017), (Daliento et al, 2017), (Shiva Kumar et al, 2015), including PV-battery operation optimization (Lu et al, 2005), (Weniger et al, 2014) and image processing (Mehta et al, 2017), (Ahmed et al, 2013). Apart from existing solutions, the proposed DSS prototype is an original solution, adaptable to different types of PV-battery systems that includes complex algorithms and methods to support O&M decisions.

The proposed solution is an informatics prototype for the acquisition, processing, management and analysis of large amounts of data collected from PV-battery systems from different types of sensors in order to support decisions regarding O&M. For wider implementation of the solution, our approach consists in a three-layer architecture of the DSS. Thus, the stakeholders of the prototype are: owners of PV-battery systems (prosumers, communities, micro-grids and industrial generators), O&M operators and dispatching centres. Currently, the prototype is under development, but the main solution is set up and its architecture is established.

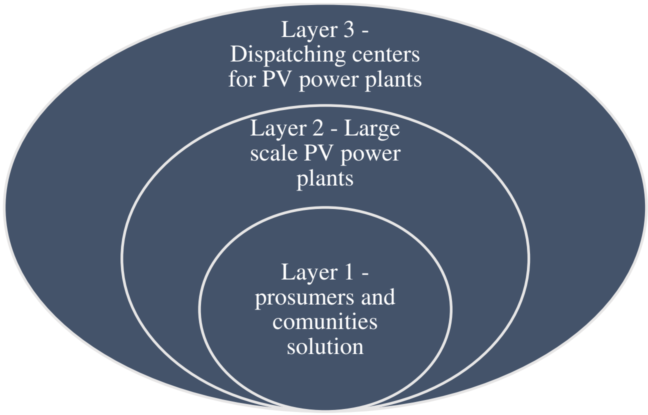
3. Big data approach

One of the main challenge for an efficient PV O&M is the amount of data that should be collected and analysed that has big data characteristics: volume, velocity, variety and veracity of data. Considering a PV of 10 MW with 50,000 solar panels (200W each), one measurement/minute of several parameters (e.g. timestamp, energy, active power, tilt, temperature, dust, irradiance, current and voltage magnitude). It gives 23 GB of data/day, or 8.4 Tb of data/year. Adding data provided by other components (e.g. weather stations, inverters, current transformers, etc.), we can estimate that a PV-battery system may produce more than 10 Tb/year. Regarding its velocity, while this amount will not normally put any stress on a server or communication network, we still have to process 50,000 individual transactions/minute, amount which is, most of the times, too high for a relational database server. As for variety, the PV may be of different type, manufacturer, generation, etc., producing data with heterogeneous formats and structures. In addition to data collected from PV-battery system, some other data sources may be added, such as thermographic cameras for failure detection or video surveillance cameras for detecting the cloud cover degree. Veracity may come from missing data due to sensors faults that require a validation and correction process to enable accurate decision regarding PV O&M. Also, O&M require real-time monitoring, diagnose and predictive analytics, optimization, forecast and advanced Key Performance Indicators (KPI) reporting. Therefore, data generated by PV-battery systems require Big Data solutions to support decisions regarding PV O&M.

4. Methodology and concept: Three - layer architecture of the DSS

The proposed informatics prototype is developed on open platforms, modular, with several tiers for Big Data management, models and analytics to provide advanced monitoring and decision support for different types and sizes of PV (residential rooftops <10kW, commercial/industrial rooftops and shade structures <1000kW; ground-mounted systems with tracking or fixed >1000kW). For wider implementation of the solution, our approach consists in a three-layer architecture of the DSS: L1-prosumers and community PV; L2-large-scale photovoltaic power plants; L3-dispatching centres for PV (Figure 1).

Figure no. 1. The three-layer architecture of the proposed prototype

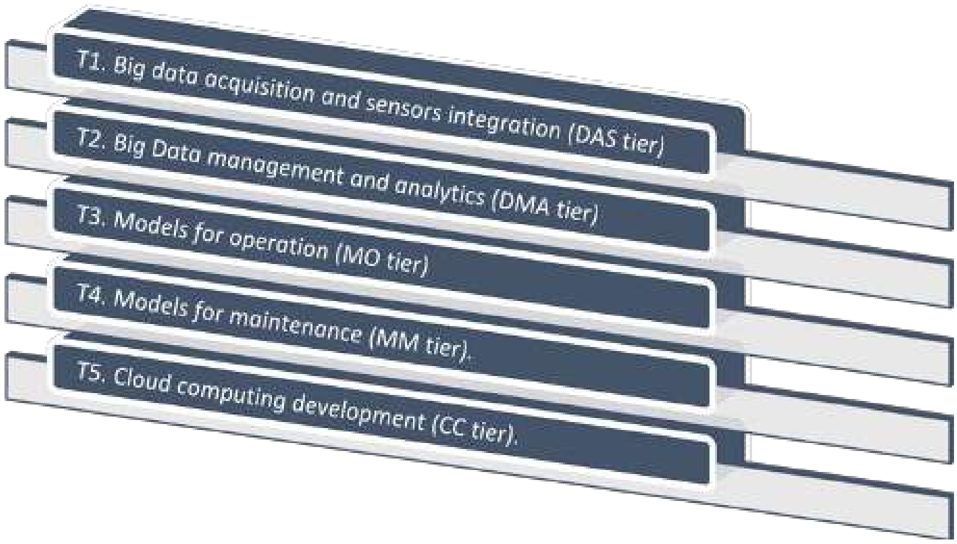


Source: Authors' contribution

By designing and developing specific models and algorithms for each of these layers and combining them into a modular cloud-based solution, we provide an original informatics prototype for decision-support regarding different PV O&M. The challenges of this concept are mainly related to massive data acquisitions and processing, designing and scaling the models and algorithms, developing Big Data analytics for diagnose, predictive analysis, KPI and real-time monitoring.

For each layer, several tiers (such as: Big Data acquisition and sensors integration, management and analytics, models for O&M and cloud computing development) will be configured and implemented (Figure 2). This approach is beyond state of the art that often address only one of the layers and doesn't provide integration of several PV O&M.

Figure no. 2. The Big Data tiers



Source: Authors' contribution

T1. Big data acquisition and sensors integration (DAS tier). For an efficient data integration, it is important to rapidly process data, avoiding it to become obsolete. The challenge is to collect, validate and process big data at high speed. The level of difficulty grows as the granularity increases. To manage these issues, we proposed an open framework to connect sensors with different communication protocols, controllers, data loggers and images. The framework collects and integrates data from different sources: i) *environment*: cell/module and air temperature, wind speed, humidity, elevation, plane of array and module irradiance, cloud cover, soiling; ii) *production*: active power, voltage and current magnitude, power factor, inverters, availability and temperature; iii) *external conditions*: curtailment events, loss of grid availability/quality/constraints; iv) *reliability and maintenance*: incident data (events, outages), fault codes and description, affected component and location, breaker trips, blown fuses status data; v) *components status*: breakage, wiring, loss of production, hot spots, advanced spectral imaging from visual inspection gathered with thermographic cameras or unmanned aerial vehicles (UAVs). Also, SCADA systems are among the most important data source in PV-battery systems. Small/simple or large/sophisticated, all PV-battery systems have a SCADA systems. They provide generation and load level of the PV-battery system including the switch positions. Historical generation series of PV systems along with weather data will be utilized in day-ahead generation forecast of PV systems. Data for historical generation series is extracted from SCADA of PV-battery systems. The DAS tier is implemented on an edge computing architecture to perform data processing at the edge of the network, thus reducing the communications bandwidth between sensors and the main data centre in cloud.

Through our integration framework, we collect data, store it on a distributed, scalable file system, such as HDFS (Hadoop File System) and develop query patterns with processing engines (Hive, Drill, Impala or Presto) in the DMA tier of the prototype.

T2. Big Data management and analytics (DMA tier). One of the main challenges for Big Data management is addressing data quality. Thus, the proposed solution requires a consistent data governance and information management process in place to ensure the data quality. Our approach consists in developing an Extract, Transform & Load (ETL) process that extract data from the DAS tier, place it in memory of the processing cluster by establishing a real time context in order to *extract* continuous inputs as data streams. Then, a *transformation* process is applied, statistics are calculated and records are checked for consistency. The transformed data is *loaded* into the central NoSQL database which relies on distributed system for better reliability and scalability.

Another challenge is to find meaningful information in the data and provide advanced analytics, machine learning (ML) and optimization algorithms. Our approach is to divide and split data into several subsets for enabling multidimensional modelling, spatial analysis and data mining. For multidimensional, subject oriented and historical analyses, data is aggregated and loaded into a central data warehouse (DW). Analytic solutions require data governance, data quality and stewardship that are absolutely critical and are achieved only through the DW. On top of the DW, we'll develop in the next phase of implementation a set of algorithms for analytical purposes and KPIs reporting. A modelling and simulation technique will be utilized, not only to identify failures in the system, but optimizing location and amount of sensors in the system by calculating the PV generation for each module and comparing it with actual generation. Also, we'll implement connectors in DAS tier to allow integration with on premises Energy Management Systems, Enterprise Resource Planning or financial applications already installed at PV owners.

T3. Models for operation (MO tier). For *performance monitoring,* MO tier includes a technical KPI framework that determines the following indicators: energy performance index, equipment forced outage rate, system availability, degradation rate at different component or system/frequency levels (Hill et al, 2015). The KPI framework also includes (Mokri et al, 2014): performance ratio (PR), temperature-corrected PR, yield, power performance index, operating efficiency, equipment equivalent availability factor, equipment equivalent forced-outage factor. The originality of the KPI framework consists in a comparative analysis at the PV component monitoring level (inverter, array, and module) that triggers alerts, detecting possible fault types and their impact, and provides recommendations for preventive maintenance. It provides information about the nature of discovered problems, their location and, apart from the existing solutions, the analysis is performed at DMA tier on real-time data collected at the module level (not on average data). Thus, the analysis provides a precise information on each component of PV that is used for advanced module-level diagnose and prognostic, providing a complete image of the PV health status.

T4. Models for maintenance (MM tier). This tier includes a methodology for setting up the PV maintenance plan (preventive, condition-based and corrective) based on monitoring results, diagnostic and prognostic algorithms. Scheduling and frequency of *preventive maintenance* will be scaled for each layer and is influenced by equipment type, environmental conditions, warranty terms, etc. Preventive maintenance is often carried out due to alerts provided by the PV monitoring activities, including thermal imaging inspections.

By detecting thermal variations between modules, any critical defect that is causing a reduction in module efficiency can be signalled, in addition to the proactive detection of hot spots and potential fire risks. By implementing the algorithms proposed in DAS tier, the prototype is able to identify module defects and proactively detect the following classes of array faults: i) module faults: hot spots, diode failures, full module failures, junction-box heating, cracked modules, ethylene vinyl acetate fogging, yellowing, antireflective coating degradation, acute soiling, etc.; ii) array and system faults: fuse, module-connector and inverter failures, reverse polarity wiring, major maximum power-point tracking faults; iii) racking and balance of system: major racking shifts, systemic shading, major erosion. Generally, inverter faults are the most common cause of system downtime in PV (Coleman, 2013). Therefore, the preventive maintenance of inverters should be treated as a centrally important part of the O&M plan.

Condition-based maintenance uses real-time data from DAS tier to schedule preventive measures such as cleaning, or to head off corrective maintenance problems by anticipating failures or catching them early. *Corrective Maintenance* plan will optimize the stocked spare parts in order to facilitate a rapid response in the event of equipment failure. The stocks should be justified by the benefit they bring in reducing plant downtime and avoiding revenue loss. Based on the failure rate of PV components, we will develop an optimum spare parts strategy depending on the size of the PV, local availability of replacement parts and the potential for sharing critical equipment across a number of plants under common ownership.

Fixed and variable costs of strong spare parts in the storehouse will be included in the formulation to ensure cost-benefit analysis. Maintenance and replacement costs will be compared in the optimization process. Spare part optimization is a mixed integer programming (MIP) problem. Solution of such problems necessitates decomposition techniques as described above. Optimization techniques that will be utilized in optimizing short-term scheduling of PV-battery system will also be considered in spare part optimization problem. Results of the spare part optimization problem will be validated by probabilistic failure analysis, in which optimum operational decisions will be simulated for the system.

The maintenance plan will also include a reliability model based on incident reports and PV block diagram with component failure trends and root-cause analyses. For PV maintenance and reliability purposes, we will also calculate KPIs, such as: preventive and corrective loss indicators (Oprea et al, 2017), planned and unplanned outage factor, failure rate, mean time between failures, mean time to repair (Hills et al, 2015). We will propose ML algorithms for module-level diagnostic and prognostic based on data generated by each module to predict imminent faults before they occur and make short-term O&M decisions and yield prognosis.

T5. Cloud computing development (CC tier). The prototype will be implemented on a cloud computing platform offering big data analytics through web-services for high availability and less investment, considering the following steps: i) a software development life cycle model will be adopted, taking into account the various already identified requirements; ii) development of an informatics prototype for PV O&M, targeting the users for each of the three layers, specific models and algorithms will be design and implemented for each layer; iii) extensive testing of the prototype informatics solution will be conducted at layer level.

5. Conclusion

The proposed prototype represents an original scientific approach for different scale of PV O&M activities (prosumers and communities, commercial and industrial), processing big data through models and algorithms to support semi-automatic O&M activities by using alerts, triggers, recommendations, diagnostic and prognostic faults, optimization of spare parts stock and maintenance plan. The expected impact of the prototype is to improve the PV O&M by: increasing performance of the PV and revenue, reducing unscheduled downtime, increasing the PV lifetime, decreasing maintenance costs and enhancing safety operation.

By improving O&M activities, PV will be better integrated, considering the expectations related to the rate of return investment and environmental impact due to the fact that the PV output will increase. Accurate forecast of PV output reduces the operation uncertainty, maintain power quality and increase the penetration level of PV into the power system.

Reliable estimates of the output variation are needed at energy distribution for system stability and dimensioning. Properly monitored PV plant can provide the most accurate information and hence the best estimates of uncertainty for partners in the energy ecosystem. Moreover, production uncertainty should be known in the PV production bidding in order to avoid penalties of undelivered energy. By optimizing the spare parts stocks, the downtime and costs will decrease. Also, by optimizing the PV-battery operation, the revenue will increase by improving the market strategies. The proposed solution can bring competitive advantages for PV owners over similar solutions that do not involve big data management. They will benefit from improved accuracy and reliability of big data analytics to support their decisions regarding O&M, asset management, buyout options, investments and further extension of their business as a consequence of higher bankability.

6. Acknowledgment

This paper presents the scientific results of the project "Intelligent system for trading on wholesale electricity market" (SMARTRADE), co-financed by the European Regional Development Fund (ERDF), through the Competitiveness Operational Programme (COP) 2014-2020, priority axis 1 – Research, technological development and innovation (RD&I) to support economic competitiveness and business development, Action 1.1.4 - Attracting high-level personnel from abroad in order to enhance the RD capacity, contract ID P_37_418, no. 62/05.09.2016, beneficiary: The Bucharest University of Economic Studies

7. References

- Ahmed Z., Kazem H. A, Sopian K., 2013. "Effect of Dust on Photovoltaic Performance: Review and Research Status". *7th International Conference Latest Trends in Renewable Energy and Environmental Informatics*, ENINF '13, Latest Trends in Renewable Energy and Environmental Informatics 2013.
- Avisolar, 2017. *EagleSUN SCADA*, Available at: https://avisolar.com/ [Accessed 28 December 2017].
- ABB, 2017. *Aurora Vision Plant Management Platform*. Available at http://new.abb.com/power-convertersinverters/ solar/monitoring-and-communication/aurora-vision-plant-management-platform [Accessed 28 December 2017].
- Coleman T., 2013. *EPRI, PV Reliability Operations Maintenance (PVROM) Database Initiative*, Project Report, 2013. Available at: https://www.nrel.gov/pv/assets/pdfs/2014_pvmrw_12_klise.pdf [Accessed 23 March 2018]
- Daliento S., Chouder A., Guerriero P., Massi Pavan A., Mellit A., Moeini R., Tricoli P., 2017. „Monitoring, Diagnosis, and Power Forecasting for Photovoltaic Fields: A Review", *International Journal of Photoenergy,* Vol (2017), Article ID 1356851, https://doi.org/10.1155/2017/1356851
- Ecoaxis, 2017. Smart Sustainability Suite for Solar PV Power Plants. Available at http://ategroup.com/ecoaxis/ [Accessed 28 December 2017].
- Escobedo G., Jacome N., Arroyo-Figueroa G., 2017. "Big Data & Analytics to Support the Renewable Energy Integration of Smart Grids. Case Study: Power Solar Generation", *Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security (IoTBDS 2017)*, 2017
- Greenbyte, B, 2017. *Solar photovoltaics management system for industry-scale solar energy owners and operators*. Available at: http://www.greenbyte.com/bright/ [Accessed 28 December 2017].
- Hill R, Klise G., Balfour JR, 2015. Sandia Reports, SAND 2015-0587, *Precursor Report of Data Needs and Recommended Practices for PV Plant Availability, Operations and Maintenance Reporting*. Available online at: http://prod.sandia.gov/techlib/access-control.cgi/2015/150587.pdf [Accessed 23 March 2018];
- Kumar B. S., Sudhakar K., 2015. "Performance evaluation of 10 MW grid connected solar photovoltaic power plant in India", *Energy Reports*, 1, doi: 10.1016/j.egyr.2015.10.001
- Lu B., Shahidehpour M., 2005. "Short-Term Scheduling of Battery in a Grid-Connected PV/Battery System", *IEEE Transactions on Power Systems*, Vol. 20, No. 2

- Mehta S., Azad A. P., Chemmengath S. A., Raykar V., Kalyanraman S., 2017. *DeepSolarEye: Power Loss Prediction and Weakly Supervised Soiling Localization via Fully Convolutional Networks for Solar Panels,* arXiv:1710.03811v1 [cs.CV], 10 Oct 2017. Available at https://arxiv.org/pdf/1710.03811.pdf. [Accessed 28 December 2017]
- Mokri J., Cunningham J., 2014. *PV System Performance Assesment*. SunSpec Alliance. Available online at: http://sunspec.org/wp-content/uploads/2015/06/SunSpec-PV-System-Performance-Assessment-v2.pdf [Accessed 23 March 2018]
- NREL/Sandia/Sunspec Alliance SuNLaMP PV O&M Working Group, December 2016. *Best Practices in Photovoltaic System Operations and Maintenance*, 2nd Edition, Technical Report NREL/TP-7A40-67553.
- Oprea S.V., Bâra A., 2017. *Key Technical Performance Indicators for Power Plants in the book "Recent Improvements of Power Plants Management and Technology*, InTechOpen, July 7, 2017.
- Siemens, 2017. *PV Monitoring Solutions - Siemens PV Portal*. Available at: https://www.energy.siemens.com/mx/en/renewable-energy/distributed-and-hybrid-power/pv-monitoring.htm [Accessed 28 December 2017].
- Solar Bankability project, 2016. *Review and Gap Analyses of Technical Assumptions in PV Electricity Cost Report on Current Practices in How Technical Assumptions Are Accounted in PV Investment Cost Calculation*, Deliverable D3.1, 27/07/2016, Funded by the H2020 Framework Programme of the European Union.
- Weniger J., Tjaden T., Quaschning V., 2014. "Sizing of Residential PV Battery Systems". *Energy Procedia* vol. 46.

.