Cyber Physical Systems Applications with a Case Study of Intelligent Dispatch of PV

Hugo Riggs, Mohammad Khan, Asham Amir, Fernando Barranco, Shahid Tufail, Imtiaz Parvez, and Arif I. Sarwat
Department of Electrical and Computer Engineering
Florida International University, Miami, Florida
Email: {hrigg002, mkhan135, aamir007, fbarr032, stufa001, iparv001, asarwat}@fiu.edu

***Abstract*—Increased investments into renewable energy as distributed energy sources throughout the distribution grid are a driving force behind the need to adopt smart grid advances. Issues such as intermittency of generation from these renewables have been studied concerning the stability of the grid at higher penetration levels. One need that industry meetings have pointed out is the need for forecasting models and advanced controls that mitigate the intermittency effects. Wide-scale adoption of computing nodes into the power grid such as smart meters and smart inverters have transformed the traditional grid into a cyber-physical system (CPS), as well as independent CPS nodes such as microgrids. Centralized command and control are necessary to maintain visibility over resources and management of them. With developments in cyber-physical-systems integrations more distributed controllers are potential. Data collection becomes a more challenging task as the number of intelligent devices in the field that generates data increases. New key-store based database technologies could be utilized in highly heterogeneous data environments. One of the potential distributed CPS controllers are demonstrated in this paper is a case study of a forecasting based economic dispatch controller for large scale PV power plant that is part of the energy mix. The cost in this scenario is found to be lower when the controller is used.**

***Index Terms*—Renewable energy systems; Cyber Physical Systems; Forecasting; Machine Learning; Economic Dispatch**

I. Introduction

The smart grid resources employ the infrastructure networks, data collection and analysis, distributed energy resources (DER), and machine learning to develop the powerful tools and control systems required to securely and economically utilize available DERs. Solar and wind generators have intermittency problems that can be mitigated through smart applications and battery energy storage systems. There are many mechanisms of these energy-critical infrastructures such as the dedicated network channels, the workforce, and experienced personnel, the coordination of large interdisciplinary teams, data storage solutions, etc. These infrastructures are heavily integrated with communications and use supervisory control and data acquisition (SCADA) that provide visibility and remote control for operators at the command center. The large scale of these systems leads to the generation of big datasets that require storage in dedicated data warehouses with historical requirements on multiple years of data archiving.

Photovoltaic (PV) systems have become an important energy source for the generation of electrical energy over the past years and the rapid emergence of renewable energy has resulted in a substantial increase of its demand. “natural gas additions peaked in the 2000s, and non-hydro renewables have grown since 2005.” A quote from the NREL Electricity Generation Baseline Report, about new additions of GW generators by fuel type [1]. Although the installation of a standalone PV system is costly, analysis of the design, sizing, and performance analysis are essential when choosing an optimal system that is economically attractive to investors. With the integration of PV systems to satisfy the needs of electrical power, with larger-scale grid integrations, there is a growing need to develop security and economic enhancing algorithms for control of the DERs [2].

The data collection of the PV systems’ generation plays an important role in determining the efficiency and performance of the photovoltaic system. A PV system may consist of battery storage, the components of the PV solar array, inverters, and MPPT controller. In this work, the aim is to collect the operational data of the photovoltaic system being PV power output and weather around the system and apply machine learning algorithms to predict the PV generation for an intelligent dispatch.

This paper makes contributions in the following ways: 1) It gives a perspective on the modern cyber physical system infrastructure and data handling; 2) This paper provides a case study on the ED problem with PV integrated at a large scale.

In section II the related works are summarized in the field of CPS and ML applications; section III covers infrastructure networks from the perspective of Cyber Physical Systems (CPS) of the power grid; section IV overviews the data collection challenges for these large-scale data generation systems; section V provides a case study of PV dispatch using a CPS empowered control strategy; finally, section VI ends with conclusions of the paper.

II. Related Works

Cyber Physical Systems are emerging as systems that integrate cyber and physical systems into a single system with increased functionalities. They are relevant in the smart grid as more pieces become integrated from the physical and cyber realms. For example, distributed mesh networks for the advanced meter infrastructure have allowed automated data collection of utility usage instead of traditional personnel manual checking the meter. The research over CPS focuses

SoutheastCon 2021 | 978-1-6654-0379-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/SOUTHEASTCON45413.2021.9401943

978-1-6654-0379-5/21/$31.00 ©2021 IEEE

on various areas from Internet of Things, cybersecurity, to machine learning and smart grid applications to name a few [3]–[5] Deciphering an accurate model for intrusion detection is an issue that is common amongst cyber-physical security systems. Examinations of models utilizing signature detection or specification-based modeling have not returned the desired results. Behavior-based unsupervised ML methods trained using high-fidelity CPS testbeds have shown progression in the detection and definition of cyber-attacks. Fast and robust to noise, the author's model obtained low false positive rates and high precision/recall [3].

In [3]an outline of classifications for the most frequent machine learning-based attacks detected in smart grids. False data injection attacks, covert cyber-attacks, electricity theft, denial of service- are mostly generated through supervised SVM and KNN methods. Proposed methods of attack mitigation include deep belief network cyber-physical and deep Q network; these methods have shown promise in their ability to identify and mitigate attacks while maintaining the stability of the network.

Research on the models used in [6] involved reconciling the gap in data during an interruption. "Appropriate implementation of this method leads to saving time by predicting the number of interruptions. Whenever the number of interruptions is forecasted based on historical weather data, power system equipment failure rates, and aging of distribution network components, the utilities can prevent a major percent of these events by establishing preventive maintenance programs" [6] Research presented in this paper uses an artificial neural network for the forecast of a Duty Cycle of DC-DC boost converter (MPPT control). [7] In [8] a PV forecasting method is used to inform a control mechanism to prevent over-voltage at the point of interconnection. The approach used 15 second ahead rapid forecasts to derive a "virtual curtailment threshold margin along with an estimated reactive compensation" [9]The results showed that this forecasting-control mechanism can reduce the overvoltage violations at increasingly integrated PV generation sites.

III. Cyber physical Infrastructure of Power Systems

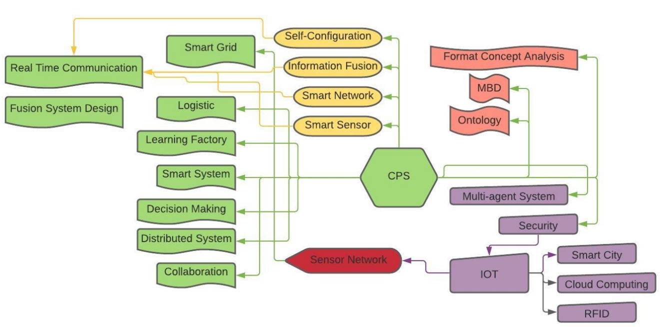

Fig. 1: Cyber Physical Systems Cross Disciplinary Roles

The cyber-physical systems are the combination of physical systems with the computational entities and communication with sensors/actuators which enable control, monitoring, and coordinating with the physical systems and their data. The cyber-physical system encompasses the intersection of the physical and cyber components. Several fields such as smart grid, autonomous vehicles, body sensor networks, medical monitoring, robotics systems, industrial control systems have adopted the practical implementation of CPS. The cyber-physical system can obtain the data from the physical world while the data analytics, computational, and management abilities from the cyber world create situational awareness and allow remote control. Although the term was coined in 2006 in the General National Science Foundation meeting, its use has been actualized much before when microprocessors were invented. The field of CPS is large, and it allows the intermingling of multiple fields like mechanical, civil, electrical, chemical, biology to collaborate and work closely with the computer science and networking field to form a system. Figure 1 shows how CPS has been used across different fields over the years.

A. *Cyber Physical Systems in the Smart Grid*

The concept of application of cyber-physical systems in smart grid is widely used these days to balance with the growing complexity and reliability of the smart grid system. The smart grid is the interconnection of Information and Communication Technology (ICT) with the power system networks. Especially considering growing renewable DERs integrations, the CPS helps alleviate the issues while integrating the power grid with the communication system and ensures effective interaction. The adoption of CPS in the smart grid becomes more economical with consistent results and it maintains the balance of prosumer satisfaction. The cyber-physical system has the capability to show different characteristics upon the integration of the physical power system and the cyber system. In the following list, some key attributes or abilities of CPS in the smart grid are underlined [3].

- The future performance of the physical system can be determined by making simulation models with the data we get from the physical systems.
- The consistent connection and communication interchange between the physical systems and cyber world helps ensure the timely response which is required in dynamic cooperation.
- The cyber physical system helps in information processing, parallel computation of the data thus making smart grid take accurate and timely decisions for different operations such as transmission, distribution, generation etc.
- The ability of the CPS to learn, adapt, organize and control enables it to have fault tolerance, take appropriate actions against the attackers or in emergency situations thereby ensuring security, sustainability of the smart grid.

B. *Cyber Physical Systems in Renewable Energy*

Renewable energy is now being generated at a high rate motivated by many factors including the increased demand for the cleaner generation of energy and the limited availability of non-renewable resources such as gas, coal, and oil. An estimation [10] shows there could 37% increase in the demand of the energy across the globe by 2040. The percentage of overall share of generation of renewable energy in electrical energy will increase from 23% to 28% [5]. Hence, the increased demand and the reduced availability of non-renewable energy resources necessitates the use of renewable energy resources such as solar, hydro, biomass, wind, and geothermal which are available in profusion. However, it requires the adoption of technologies to capitalize on the usage of such resources in a secured and reliable manner. Even more so, the integration of information and communication technologies (ICT) in the physical systems ensures the generation, transmission, usage, and distribution to be more efficient and reliable. One of the issues we could face is to balance the tradeoff between the requirement of the generation and the actual generation while maintaining the efficiency of the system.

IV. Data Collection and Storage in Power Systems

There are many standards, practices, and technologies in use by industry and the following subsections detail a few of these; the following provide a conceptual view of the scope of concerns surrounding data collection and storage.

A. North American Protocols

The American National Standards Institute (ANSI) provides protocol for two-way communication with meters used in North American markets; electricity meters typically operate as an 'end-device' communicating with a 'computer.' Protocol C12.19 is the standard for end device data tables in the utility industry. C12.18 describes the communication of C12.19 tables over an ANSI Type 2 optical port between the two, while C12.21 describes communication for modems.

B. European Protocols

The International Electrotechnical Commission published IEC 66107, a widely used protocol for smart meters in the European Union- instituting a half-duplex communication system. ASCII code is transmitted via serial port of either a photodiode receptor or EIA-485 modulated wires.

The European Telecommunications Standards Institute (ETSI) published the Open Smart Grid Protocol (OSGP), a globally used smart grid device network. Used in conjunction with IEC14908, the protocol controls networking standards for many smart grid applications and information.

C. Transmission Control Protocol/Internet protocol Suite

TCP/IP technology allows utilities to deploy multiple communication systems on the same network under one interface. The protocol stack maps partially to the Open Systems Interconnect (OSI) seven-layer conceptual model for networking. The physical and data link layer represents the network interface card, supporting network hardware, and the media access control protocol respectively. Network, transport, session, presentation, and application layers of the OSI model represent the Internet Protocol (IP), Transport Control Protocol (TCP), and the end application respectively. The TCP/IP solution was globally adopted and has remained the standard for Internet and many local area communication protocol stacks since the late 1980s. Many further communication standards and protocols have been built off of the existing TCP/IP stack, including the GOOSE IEEE 6800 protocol for substation automation. It is important to understand the protocol stack for smart grid engineers building communications, control, and data acquisition systems.

D. Data Collection to Storage

Data collection or acquisition is the essential stage of a modern infrastructure system that utilizes sensors and infrastructure networks to communicate the field information to the back office. This also can apply to cyber-security resources such as intrusion detection systems (IDS) and their log data. The process of gathering this data has distinctly different requirements from the network itself, mainly involving the requirement of data storage. Large infrastructure generates large data, and this is non-trivial to handle requiring petabyte-scale storage solutions and data storage software. Data warehouses are the industry solution for large scale storage, and they are specially engineered to keep homeostasis of the environment for storage with air conditioning and humidity control. Along with database software, support software for hardware failure monitoring is run constantly so that maintainers of the Datawarehouse can switch out failing storage drives. Traditional structured query language (SQL) has been dominant in the industry as it uses fixed schemas with well-defined data tables and is a very efficient and well-matured technology. The more recent No-SQL databases have become popular in smaller projects for the ease of use of the technology requiring no fixed schema and relying on key and value mechanism for storing documents. The advantage this No-SQL has is in rapid prototyping and deployment, the ease of creating redundant nodes for the data duplication, the relaxed constraints on storage of new incoming data, and the speed of writing data to the database.

E. Data Privacy and Security

The international standards organization (ISO) issued a standard in 2019 to address data privacy with ISO 27701. This standard has information to help data managers develop privacy information management systems. A previous standard ISO 27001 is put forth for information security standardization. These standards provide a framework for the ways, respectively what data is collected that is personal data and prevention of disclosure of personal data; and that organizations keep storage accessible to only the allowed employees [11], [12].

F. Datawarehouse

Data warehouses are the industry solution for large scale storage, and these structures are engineered to keep a temperature, humidity-controlled environment, and have backup energy generators and battery backups. This environment is created to be ideal for the storage of data. Along with database software, support software for hardware failure monitoring is run constantly so that maintainers of the Datawarehouse can switch out failing storage drives. Many Datawarehouse may be used for the critical infrastructure supervisory control and data acquisition (SCADA) data storage and large companies that need to save important information such as customer data, and archives, in fact, the United States Treasury Department has addressed the importance to understand the lack of security in our existing data warehouses [13]. A storage warehouse needs to have excellent security. Much of the stored data will be sensitive and must be kept privately such as information on the customer, payment information, and location. Required updates for security tools and firewalls are a constant process for the personnel managing security at the Datawarehouse. End to end encryption is vital in securing data against man-in-the-middle type attacks. The security depends on the initial design when the infrastructure is built and what standards are adopted. The principal approach is to have a centralized location, and the maintenance will be less expensive while security is optimal. The Oracle company provides an example of these two methods [14]. Oracle has an autonomous and optimized data warehousing system that focuses on ease of management. They increase query performance by only having the systems work on relevant data and as far as maintenance goes, they have a temperature-based data optimization which is made possible with their automatic data optimization system and that system works in conjunction with Oracle Partitioning which is a very efficient and powerful way to approach the lifecycle management for large environments [15].

G. NoSQL

The use of NO-SQL has become increasingly popular and using NO-SQL has many advantages as compared to SQL. Its key characteristics are, no schema is required, and unstructured data can be stored as key-value entries. [16].NO-SQL is not a replacement for SQL, rather questions of scale and application specifics are considered in choosing the most appropriate database technology. SQL is favored for multi-row transactions, unlike NO-SQL which handles data that has no structure. NoSQL provides an easier way of storing data that may be complex. Storing data without schema also makes it easier for companies with an abundant amount of data to track and control their data without the constraints of a schema [16]. Companies that have what is called a "Session Store" are the ones who benefit the most from No-SQL since it makes it easier to manage data within the session that uses relational databases. This makes it easier to use the user's cookies to try and see who visits the sites and all these data can be used for growth purposes. When the user has a profile in the store, it also works very well for authentications and preferences [17]. Companies like Amazon are very well known for using No-SQL for user accounts and storing the data, but they also use SQL for other back end functions, this takes us back to "Which one is better?" it's all about functionality since each one is best for certain features. Other works [18]–[20] have focused on NoSQL technology used in fog computing that has its place between the edge, field data collection devices, and control center. In applications such as IoT, an intermediary database has multiple possible functionalities such as aggregation and reduction, optimization, of data streams as well as analytical processing.

V. Machine Learning based PV Dispatch

The adoption of data science into business analytics has accelerated machine learning applications and workforce development. Many techniques exist and have been used in specific applications of tools that use machine learning. Models such as Artificial Neural Networks, Markov models, Support Vector Machines, and Genetic Algorithms to name a few. These powerful models can be used in classification and regression tasks that expand their application use to a multitude of possibilities. The totality of use cases for ML in CPS cannot be enumerated here, but we use the example of forecasting of photovoltaics generation in this paper as an example of a smart grid application that can be used to mitigate the side-effects of intermittency from a generator and create more reliable electrical energy dispatch solutions.

In [9] distributed generation (DG) is considered a non-dispatchable sources of energy which require short-term forecasting with error tolerances to adapt into the economic dispatch problem. In [9] they develop a microgrid level economic dispatch and security of service restorative measures. The work incorporates data models for PV and wind power as the local DG sources. The Solar Energy Technology Office (SETO) of the DOE has also expressed the industry demand for forecasting PV generation [2]. These needs are elicited in a SETO list forecasting of short-term PV production as one of the primary industry demands [2].

Integration of utility scale PV Systems into the larger power grid is an ongoing process that benefits from having the ability to forecast expected power generated from the system. The most accurate, and quick, power forecasting algorithm is the most useful in energy dispatch. Many different efforts focus on developing such forecasting algorithms [2]. As DG sources have the problem of intermittency of power that is affected by the environment, they are interesting to model and can take the form of hybrid-ML data-driven algorithms. Given a PV system the generation will be intermittent so utilities that integrate DG can benefit from improved visibility and control of the PV assets involving analytics, machine learning applications and visualizations of the system.

Deriving a reliable power output from PV Systems is a major challenge combining battery energy systems with the PV plant can help in achieving this goal. An approach that can be used is to shift the time the energy is consumed by pairing PV with energy storage while optimizing multiple objectives, e.g., when to discharge the battery and minimize its charging cycles. Having the future generation forecast and controlling the power routes, could use a combination of technologies including battery management systems and

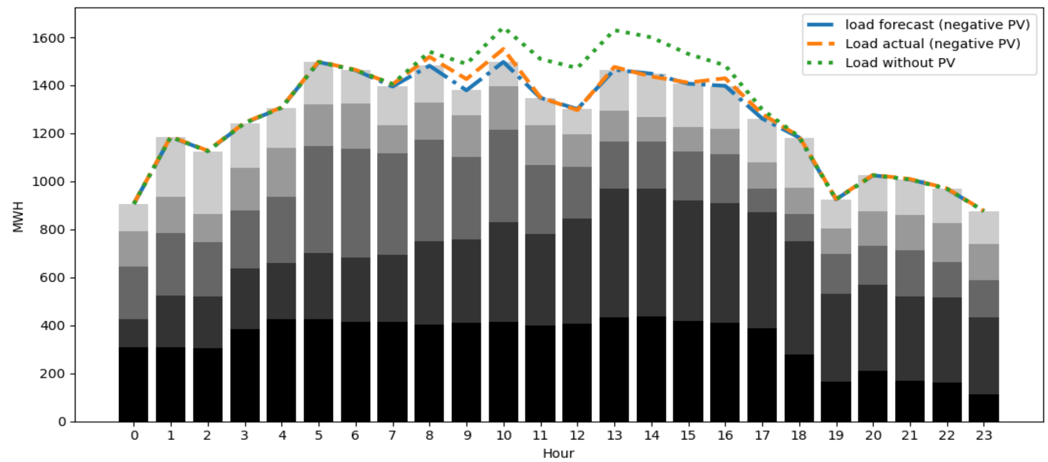

Fig. 2: The ED of the Resources showing PV Negative Load and Forecast Error

adaptive control algorithms that make use of the forecast. As PV penetration increases into the feeder level it will affect voltage and frequency fluctuations at the feeder level [21], [22], and therefore advanced control and response mechanisms should be built for PV Systems; integrate the appropriate technologies to efficiently mitigate the intermittency issues of PV power plants and potential violations. Running dispatch of PV by routing power to load or storage or back to the grid as a potential approach.

A. ML – System, Application, and Evaluation A Case Study

To develop the ML algorithm, the data is used from the real-time monitoring and reporting for the 1.4 MW PV power plant located at FIU Engineering Center (EC). This system provides minutely power recordings, the Global Horizontal Irradiance (GHI) incident on the panels. This resource has been used in other research including predictive modeling and performance analyses [23]. Machine learning models (Neural Networks and Statistical Models) are used for forecasting and imputing missing data to build the ML algorithm.

The software component is developed as part of a CPS locally or even as a remote service that provides the forecasting mechanism. The mechanism has a self-improving heuristic through online learning where the model updates every 10 minutes. The model's inputs are the PV power generation, the weather data, the irradiance, the temperature of the PV panels, and the configuration includes the PV array capacity, any energy storage at the site, and its' capacity (energy storage is simulated). Using forecast energy generation, more advanced control mechanisms are enabled.

B. Economic Dispatch of MW scale PV

Economic dispatch (ED) is the calculation of the optimal use of generators to supply the load, typically optimal in terms of the cost to produce energy, however some formulations weight security as well as cost in the generator dispatch. The problem of Economic Dispatch is well established, however with the introduction of DERs that suffer from intermittency of output power there are some further challenges to address due to the randomness of the output. Some [8] have opted to model PV as a negative load which is an approach, we follow in designing ED with PV.

In this approach the system load is based on aggregated multiple smart meters of residential loads. The power generated by the PV system is scaled to represent a larger 900+MW system. The traditional generators are scheduled to be curtailed according to the load actual considering the PV forecast as negative load. In this way a schedule can be set 24-hours ahead for generation and the error is matched by a form or multiple forms of energy storage. Spinning reserve would be ramped up automatically for the hours when the PV is expected to be generating especially around the peak zenith of the sun over the panels when large intermittencies can be experienced. The cost of the spinning reserve would be known ahead of time in the dispatch.

The cost of the traditional generators is calculated by using their alpha beta gamma cost coefficients, as the amount of generation increases so do the costs as in accordance with the law of diminishing returns as seen in Figure 3.

Costs are detailed hourly in table 1. These show the mismatch is very low until the forecasting error begins to increase mismatch at which point reserve energy systems must cover the deficit in energy only if the PV forecast is an over-estimate of the actual generation. If the forecast is and under-estimates the actual generation from the PV then the traditional generators will already be scheduled to produce more than the demand and excess can be curtailed. The same dispatch algorithm is run when PV is not considered in the dispatch to compare the total costs assuming the load forecast is very accurate and spot market buys take place only when the PV forecast is an overestimate. The spot market price is set at a fixed per MWH rate of 19$/MWH. The total cost of the dispatch disregarding the impact of PV as a negative load totals to 374,892.03 dollars and the total cost of the dispatch that

Hour	Production MW	Mismatch MW	Cost USD	Load MW	Spot Buy USD
1	1000	-1.32	12310.08	1001.32	0
2	1100	-1.75986	13498.83	1101.76	0
3	1200.001	-1.75861	14588.17	1201.76	0
4	1200.001	-1.75861	14588.17	1201.76	0
5	1249.998	-1.76195	15071.96	1251.76	0
6	1400	-1.76024	16608.06	1401.76	0
7	1500	-1.73688	17692.89	1501.737	0
8	1439.157	-9.10576	17005.07	1448.263	0
9	1441.743	-36.9454	17034.35	1478.688	0
10	1440.309	-46.8236	16984.36	1487.133	0
11	1405.438	-53.8208	16651	1459.259	0
12	1318.098	-1.95526	15786.64	1320.053	0
13	1309.187	4.530838	15796.72	1304.656	86.09
14	1435.041	-11.5855	17145.79	1446.627	0
15	1398.671	11.84014	16759.13	1386.831	224.96
16	1425.776	-4.67161	17034.27	1430.447	0
17	1365.004	-31.4411	16299.82	1396.445	0
18	1362.031	-18.8118	16271.01	1380.842	0
19	1196.746	-4.62449	14519.72	1201.371	0
20	999.9991	-1.76089	12448.12	1001.76	0
21	1099.999	-1.7608	13498.53	1101.76	0
22	1000.001	-1.75939	12405.58	1001.76	0
23	1000.001	-1.75939	12405.58	1001.76	0
24	900.0002	-0.87981	11390.98	900.88	0

uses the PV forecast as a negative load, to reduce scheduled traditional power with spot market purchases in the case of over forecast, in this scenario is 364,105.86 dollars. The total cost difference shows that using the PV as a negative load is beneficial, in this scenario even with an overestimate of the generation of the PV, the difference is 10,786.17 dollars less for the daily dispatch.

Economic dispatch of 24-hour ahead of MW scale PV with assumed load performed as an application of the combined capabilities of the smart grid with CPS and ML applied concepts. In this case large scale 500+ MW PV based intermittent generation is over forecasted and the error is shown in the Figure 4. This error of 300 MW or a 14% miss, this would then require to be covered by spinning reserve or by a battery energy storage system. Energy reserves can be created proportionally to the maximum error of historical forecasting results. Therefore, accelerated testing and determination of forecasting results for over estimation errors is important.

VI. Conclusion

In this paper we have briefly reviewed the related works in the areas of cyber-physical systems and machine learning applications, from areas of cyber-security to power systems. In the case study section of the paper the results of the genetic algorithm economic dispatch are summarized with the total costs, the dispatch including the PV forecast and the actual load. The results show that large scale PV can be integrated into the power grid by using error matching energy storage systems, such as risk scheduled spinning reserve. The future work envisioned is to run a similar buy expanded set of case studies on different scenarios of DERs and to compare results of power swarm optimization algorithm against the genetic algorithm in the economic dispatch, as well as comparing the accuracies of different PV generation forecasting approaches.

VII. Acknowledgment

The work published is a product of research jointly supported by the National Science Foundation Grants CMMI-1745829 and CNS-1553494, and the U.S. Department of Energy DE-OE0000779.

References

[1] J. Logan, C. Marcy, J. McCall, F. Flores-Espino, A. Bloom, J. Aabakken, W. Cole, T. Jenkin, G. Porro, and C. Liu, "Electricity Generation Baseline Report – Executive Summary," *Renewable Energy*, p. 17, 2017.

[2] "2018 SETO Portfolio Book," library Catalog: www.energy.gov. [Online]. Available: https://www.energy.gov/eere/solar/downloads/2018-seto-portfolio-book

[3] J. Zhao, F. Wen, Y. Xue, X. Li, and Z. Dong, "Cyber physical power systems: Architecture, implementation techniques and challenges," *Dianli Xitong Zidonghua/Automation of Electric Power Systems*, vol. 34, pp. 1–7, Aug. 2010.

[4] N. I. Haque, M. H. Shahriar, M. G. Dastgir, A. Debnath, I. Parvez, A. Sarwat, and M. A. Rahman, "Machine Learning in Generation, Detection, and Mitigation of Cyberattacks in Smart Grid: A Survey," *arXiv:2010.00661 [cs, eess, stat]*, Sep. 2020, arXiv: 2010.00661. [Online]. Available: http://arxiv.org/abs/2010.00661

[5] X. Yu and Y. Xue, "Smart Grids: A Cyber–Physical Systems Perspective," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1058–1070, May 2016, conference Name: Proceedings of the IEEE.

[6] A. I. SARWAT, M. AMINI, A. DOMIJAN, A. DAMNJANOVIC, and F. KALEEM, "Weather-based interruption prediction in the smart grid utilizing chronological data," *Journal of Modern Power Systems and Clean Energy*, vol. 4, no. 2, pp. 308–315, Apr. 2016. [Online]. Available: https://doi.org/10.1007/s40565-015-0120-4

[7] A. Anzalchi and A. Sarwat, "Artificial neural network based Duty Cycle estimation for maximum Power Point tracking in Photovoltaic systems," in *SoutheastCon 2015*, Apr. 2015, pp. 1–5, iSSN: 1558-058X.

[8] L. Xie and M. D. Ilic, "Model predictive economic/environmental dispatch of power systems with intermittent resources," in *2009 IEEE Power Energy Society General Meeting*, Jul. 2009, pp. 1–6, iSSN: 1932-5517.

[9] M. A. Karim, J. Currie, and T.-T. Lie, "A machine learning based optimized energy dispatching scheme for restoring a hybrid microgrid," *Electric Power Systems Research*, vol. 155, pp. 206–215, Feb. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378779617304212

[10] "World Energy Outlook 2014 – Analysis." [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2014

[11] "ISO 27701 privacy information management | IT Governance Ireland Ireland." [Online]. Available: https://www.itgovernance.eu/en-ie/iso-27701-ie

[12] L. Irwin, "An introduction to ISO 27701: the international standard for data privacy," Jun. 2020. [Online]. Available: https://www.itgovernance.eu/blog/en/iso-27701-the-new-international-standard-for-data-privacy

[13] "The Cybersecurity Data Warehouse Needs Improved Security Controls," p. 27.

[14] "Security and the Data Warehouse," p. 18.

[15] "Partitioning Overview." [Online]. Available: https://www.oracle.com/database/technologies/partitioning.html

[16] DATAVERSITY, "Top 10 Enterprise Use Cases for NoSQL." [Online]. Available: https://www.slideshare.net/Dataversity/top-10-enterprise-use-cases-for-nosql-47639798

[17] "Use of NoSQL in Industry," Dec. 2018, section: Misc. [Online]. Available: https://www.geeksforgeeks.org/use-of-nosql-in-industry/

[18] B. Małysiak-Mrozek, M. Stabla, and D. Mrozek, "Soft and Declarative Fishing of Information in Big Data Lake," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 5, pp. 2732–2747, Oct. 2018, conference Name: IEEE Transactions on Fuzzy Systems.

[19] S. Yang, "IoT Stream Processing and Analytics in the Fog," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 21–27, Aug. 2017, conference Name: IEEE Communications Magazine.

[20] A. Ioualalen, "Fog Computing," in *Data Lakes*. John Wiley & Sons, Ltd, 2020, pp. 171–186, section: 8 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119720430.ch8. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119720430.ch8

[21] M. Jafari, T. O. Olowu, and A. I. Sarwat, “Optimal Smart Inverters Volt-VAR Curve Selection with a Multi-Objective Volt-VAR Optimization using Evolutionary Algorithm Approach,” in *2018 North American Power Symposium (NAPS)*, Sep. 2018, pp. 1–6.

[22] T. O. Olowu, M. Jafari, and A. I. Sarwat, “A Multi-Objective Optimization Technique for Volt-Var Control with High PV Penetration using Genetic Algorithm,” in *2018 North American Power Symposium (NAPS)*, Sep. 2018, pp. 1–6.

[23] A. Sundararajan and A. I. Sarwat, “Roadmap to prepare distribution grid-tied photovoltaic site data for performance monitoring,” in *2017 International Conference on Big Data, IoT and Data Science (BID)*, Dec. 2017, pp. 110–115.