

Clemson University

**TigerPrints**

---

All Dissertations

Dissertations

---

8-2022

## Data-Driven Distributed Modeling, Operation, and Control of Electric Power Distribution Systems

Hasala Dharmawardena  
hdharma@clemson.edu

Follow this and additional works at: [https://tigerprints.clemson.edu/all\\_dissertations](https://tigerprints.clemson.edu/all_dissertations)



Part of the [Power and Energy Commons](#)

---

### Recommended Citation

Dharmawardena, Hasala, "Data-Driven Distributed Modeling, Operation, and Control of Electric Power Distribution Systems" (2022). *All Dissertations*. 3090.

[https://tigerprints.clemson.edu/all\\_dissertations/3090](https://tigerprints.clemson.edu/all_dissertations/3090)

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact [kokeefe@clemson.edu](mailto:kokeefe@clemson.edu).

# DATA-DRIVEN DISTRIBUTED MODELING, OPERATION, AND CONTROL OF ELECTRIC POWER DISTRIBUTION SYSTEMS

---

A Thesis  
Presented to  
the Graduate School of  
Clemson University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy  
Electrical Engineering

---

by  
Hasala Indika Dharmawardena  
August 2022

---

Accepted by:  
Dr. Ganesh Kumar Venayagamoorthy, Committee Chair  
Dr. Johan Enslin  
Dr. Rajendra Singh  
Dr. Yongjia Song

# Abstract

The power distribution system is disorderly in design and implementation, chaotic in operation, large in scale, and complex in every way possible. Therefore, modeling, operating, and controlling the distribution system is incredibly challenging. It is required to find solutions to the multitude of challenges facing the distribution grid to transition towards a just and sustainable energy future for our society. The key to addressing distribution system challenges lies in unlocking the full potential of the distribution grid. The work in this dissertation is focused on finding methods to operate the distribution system in a reliable, cost-effective, and just manner.

In this PhD dissertation, a new data-driven distributed ( $D^3M$ ) framework using cellular computational networks has been developed to model power distribution systems. Its performance is validated on an IEEE test case. The results indicate a significant enhancement in accuracy and performance compared to the state-of-the-art centralized modeling approach.

This dissertation also presents a new distributed and data-driven optimization method for volt-var control in power distribution systems. The framework is validated for voltage control on an IEEE test feeder. The results indicate that the system has improved performance compared to the state-of-the-art approach.

The PhD dissertation also presents a design for a real-time power distribution system testbed. A new data-in-the-loop (DIL) simulation method has been developed and integrated into the testbed. The DIL method has been used to enhance the quality of the real-time simulations. The assets combined with the testbed include data, control, and hardware-in-the-loop infrastructure. The testbed is used to validate the performance of a distribution system with significant penetration of distributed energy resources.

# Dedication

To my parents, Upali and Seetha Dharmawardena,  
for all their sacrifices, inspiration, and love.

# Acknowledgments

My advisor, Dr. Ganesh Kumar Venayagamoorthy, was a wonderful mentor, supervisor, and teacher to me throughout my doctoral degree. You were an immense source of inspiration and encouragement. Thank you for the support and guidance over the past six years. I would like to thank my incredible PhD advisory committee, Dr. Johan Enslin, Dr. Rajendra Singh, and Dr. Yongjia Song, for their time, effort, and support. The world-class leading-edge laboratory facilities provided by the Real-Time Power and Intelligent Systems Lab of Clemson University were instrumental in carrying out the research presented in this thesis. A special note of appreciation goes to the Clemson University Utility Services and Energy Management for working with the Real-Time Power and Intelligent Systems Lab to integrate weather measurements and Clemson University solar photo-voltaic plant and loads for research and education. Their support helped enhance the research that I conducted during my doctoral studies.

This research work was supported by the National Science Foundation (NSF) of the United States, under grants IIP 1312260, 1638321, 2131070, and the Duke Energy Distinguished Professorship Endowment Fund. The funding agencies and the lab are greatly appreciated.

I want to thank my colleagues at the Real-Time Power and Intelligent Systems Lab; Chirath Pathiravasam, Parani Arunagirinathan, Dulip Madurasinghe, Ali Arzani, and Pramod Herath. Working with everyone and sharing ideas over team meetings and lunch has been a pleasure. I would also like to share my appreciation for my friends and colleagues at Clemson University and beyond, whose friendships have made this journey enjoyable.

Finally, I would like to thank my life partner. Thank you for your company, support, and understanding during this important chapter of my life. Coming home every day to a gentle smile and a warm hug has been instrumental in working through the difficult phases of this long journey.

# Table of Contents

Title Page . . . . .	i
Abstract . . . . .	ii
Dedication . . . . .	iii
Acknowledgments . . . . .	iv
List of Tables . . . . .	viii
List of Figures . . . . .	ix
<b>1 Introduction . . . . .</b>	<b>1</b>
1.1 Research challenges in the Electric Power Distribution Systems . . . . .	6
1.2 Research objectives of this dissertation . . . . .	6
1.3 Main contributions . . . . .	7
1.3.1 Distribution system modeling . . . . .	7
1.3.2 Distribution system control optimization . . . . .	7
1.3.3 Distribution system real-time testbed . . . . .	8
<b>2 Transformation of the Power Distribution System . . . . .</b>	<b>9</b>
2.1 Introduction . . . . .	9
2.2 The power distribution system . . . . .	10
2.2.1 Traditional power distribution system . . . . .	10
2.2.2 Distributed Energy Resources (DER) integrated distribution grid . . . . .	12
2.2.3 New challenges . . . . .	12
2.2.4 State-of-the-art . . . . .	13
2.2.5 Smart distribution grid . . . . .	17
2.2.6 Journey from the smart distribution system to an intelligent distribution system	18
2.2.6.1 Communication . . . . .	20
2.2.6.2 Computation . . . . .	21
2.2.6.3 Controls . . . . .	22
2.3 Summary . . . . .	23
<b>3 Distribution System Modeling . . . . .</b>	<b>25</b>
3.1 Introduction . . . . .	25
3.2 Problem formulation . . . . .	29
3.2.1 The power flow problem . . . . .	30
3.3 The Distributed Data-Driven Modelling ( $D^3M$ ) framework . . . . .	31
3.3.1 Overview . . . . .	31
3.3.2 Methodology . . . . .	32

3.3.2.1	Model learning . . . . .	32
3.3.2.2	Cell structure optimization . . . . .	34
3.3.2.3	System decomposition . . . . .	35
3.3.2.4	System learning . . . . .	37
3.3.2.5	Model Fusion . . . . .	37
3.4	Case Study . . . . .	39
3.4.1	System modelling . . . . .	40
3.4.2	Dataset generation . . . . .	41
3.4.3	Model development and learning . . . . .	42
3.4.4	CCN in Operation . . . . .	43
3.5	Results and Discussion . . . . .	43
3.6	Summary . . . . .	46
<b>4</b>	<b>Distributed Volt-Var Curve Optimization Using a Cellular Computational Network Representation . . . . .</b>	<b>52</b>
4.1	Introduction . . . . .	53
4.1.1	Optimization . . . . .	54
4.1.2	State-of-the-art . . . . .	56
4.1.3	Contributions . . . . .	57
4.2	Problem formulation . . . . .	58
4.3	Distributed optimization framework (DOF) . . . . .	61
4.3.1	CCN representation of the EPDS . . . . .	62
4.3.2	Node level control . . . . .	62
4.3.2.1	DER modeling . . . . .	62
4.3.3	Cell optimization . . . . .	64
4.3.4	System optimization . . . . .	65
4.3.4.1	Graph-based cell rank evaluation . . . . .	65
4.3.4.2	Impact-based cell rank evaluation . . . . .	66
4.4	Case study . . . . .	67
4.5	Results and discussion . . . . .	67
4.6	Summary . . . . .	73
<b>5</b>	<b>A Real-Time Distribution System Simulation Laboratory Testbed . . . . .</b>	<b>76</b>
5.1	Introduction . . . . .	77
5.2	Real-time power system simulations . . . . .	78
5.3	Design of real-time distribution system test-bed . . . . .	79
5.3.1	Communication . . . . .	81
5.3.2	Data-in-the-loop simulation (DIL) . . . . .	81
5.4	Implementation of laboratory testbed . . . . .	81
5.4.1	Real-time simulator . . . . .	82
5.4.2	Implementation of DIL . . . . .	83
5.4.3	IoT . . . . .	84
5.4.4	Other devices . . . . .	84
5.5	System demonstration . . . . .	84
5.6	Summary . . . . .	86
<b>6</b>	<b>Conclusions . . . . .</b>	<b>96</b>
6.1	Introduction . . . . .	96
6.2	Section summaries . . . . .	96
6.2.1	Distribution system modeling . . . . .	97
6.2.2	Distribution system volt-var curve optimization . . . . .	97

6.2.3	Distribution system real-time testbed . . . . .	97
6.3	Future work . . . . .	97
6.4	Summary . . . . .	99
<b>Bibliography . . . . .</b>		<b>100</b>



# List of Tables

3.1	Parameters of the neural networks used for $D^3M$ of the power flow. . . . .	44
4.1	DER connected to the test case. . . . .	69
4.2	Assignment of different loadshapes to power conversion components in the case study.	69
4.3	Sequentially optimized utility values . . . . .	70
4.4	Impact factor based priority based CCN: calculated utility function values for system and cells . . . . .	71
4.5	Result from the impact factor evaluation based on available reactive power resources and $U_{sys}$ for cell reactive power injection . . . . .	72

# List of Figures

1.1	The world total energy supply by source from 1971 to 2018 [1] (y-axis is Mtoe) . . .	2
1.2	The world total electricity consumption by source from 1971 to 2018 [1] (y-axis is TWh) . . .	3
1.3	The world total electricity consumption in 1973 [1] . . . . .	4
1.4	The world total electricity consumption in 2019 [1] . . . . .	5
2.1	A passive/conventional distribution system. The power flow is uni-directional [2]. . .	10
2.2	A traditional distribution grid with high penetration of DERs results in adverse operational challenges [2] . . . . .	14
2.3	A map of available hosting capacity maintained and hosted publicly by Pacific Gas and Electric Company [3] . . . . .	16
2.4	Conceptual component level overview framework of the smart power system . . . . .	19
2.5	Cyber-physical framework of the smart distribution system . . . . .	22
3.1	Different approaches to distribution system modelling. . . . .	27
3.2	The system overview showing the generic ( $D^3M$ ) framework applied to a distribution system. . . . .	33
3.3	The flowchart for calculation of objective function for cell structure identification. . .	36
3.4	(a) The single generic cell single line diagram (b) electrical equivalent circuit (c) cell input-output relationship if cell k has a child cell (d) corresponding cell input-output relationship if cell k is a leaf node. . . . .	38
3.5	Model fusion flowchart showing the iterative algorithm used to solve for the $D^3M$ solution. Here $X$ refers to $X_{CCN}$ . The load flow estimation is for the nodal load dataset given by $LD$ . . . . .	39
3.6	A generic model of a prosumer node used in the study. The active and reactive power for each prosumer lies within the boundary shown in red. The cross represents key extreme data points. . . . .	40
3.7	The $D^3M$ based cell formulation for the modified IEEE 34 network. The chosen cells and boundary variables are shown on the single line diagram. . . . .	41
3.8	The information flow diagram for structure fusion. The cells are initialized using field data or rated load and voltage values. The information is exchanged and cell outputs are computed till the output converged to the required accuracy. . . . .	42
3.9	The computational cell structure is illustrated using cell MC3. It consist of a single hidden layer Multi Layer Perceptron network. The optimal number of hidden layer neurons, $n_h$ , as well as the weight matrices, are calculated by using historical data. .	43
3.10	Statistical distribution of the $D^3M$ solution for boundary voltage magnitude absolute percentage error for a dataset of 1000 random validation data points. The x-axis represents the voltage magnitude variables that are extracted from the solution. . . .	45
3.11	Statistical distribution of the SC1 and SC2 internal node voltage magnitude absolute percentage error. . . . .	46
3.12	Statistical distribution of the MC1 cell internal node voltage magnitude absolute percentage error. . . . .	47

3.13	Statistical distribution of the MC3 cell internal node voltage magnitude absolute percentage error. . . . .	48
3.14	Statistical distribution of the $D^3M$ solution for inter-cell power flow absolute percentage error for a dataset of 1000 random validation data points. The x-axis represents different power flow variables. . . . .	49
3.15	Statistical distribution of the $D^3M$ solution for boundary voltage magnitude absolute percentage error for the power flow estimation using the Onenet network. The same training dataset and validation dataset are applied for network training and validation across all networks. . . . .	50
3.16	Statistical distribution of the $D^3M$ solution for boundary voltage magnitude absolute percentage error for a dataset of 8 extreme validation data points chosen on the circumference of the prosumer power curve. The x-axis represents the voltage magnitude variables that are extracted from the solution. . . . .	51
4.1	The classification of optimization based on framework, implementation, objective function, and solution. . . . .	55
4.2	The volt-var curve as defined in IEEE 1547 standard. The curve is optimized by changing the six variables shown on this plot. The x values of the VVC is given by (4.3). . . . .	60
4.3	Overview of the proposed DOF. . . . .	61
4.4	The overview of the prosumer based PV inverter control system. The volt-var curve parameters ( $x$ ) is changed by the cell operator. This is the node level optimization of the DOF. . . . .	63
4.5	The implemented algorithm for cell level optimization. . . . .	65
4.6	The implemented algorithm for system level optimization. . . . .	66
4.7	The CCN representation of the IEEE 34 bus system used for distributed optimization . . . . .	68
4.8	Solar irradiation variation across the one minute time frame of interest . . . . .	68
4.9	Load variation across the optimized one minute time frame . . . . .	70
4.10	Comparison of Utility across the optimized one minute time frame . . . . .	71
4.11	Comparison of voltage across the optimized one minute time frame . . . . .	72
4.12	Comparison of voltage of 3 randomly selected nodes across the optimized one minute time frame . . . . .	73
4.13	Comparison of optimization performance. Here $T_{graph}$ and $T_{impactfactor}$ represents the cumulative time for each iteration, where time is given in the right y axis, and the $U_{graph}$ and $U_{Impactfactor}$ gives the Utility value for each iteration, where utility is shown in the left y axis. . . . .	74
4.14	The volt-var curves for cell SC3 . . . . .	74
5.1	High level overview of the various types of components/sub-system that can be tested using a real-time hardware-in-the-loop simulation . . . . .	78
5.2	Overview diagram of the real-time simulation testbed. . . . .	80
5.3	Overview diagram showing the operation of the data-in-the-loop process. . . . .	82
5.4	The overview diagram of the modeled distribution system and interfaced hardware-in-the-loop devices . . . . .	87
5.5	The overview of the components and software used for implementation of the Clemson University real-time Power and Intelligent systems lab real-time power distribution hardware-in-the-loop testbed. . . . .	88
5.6	The actual R06 PV plant generation for 24 hours between 5/1/2022 and 5/2/2022. . . . .	88
5.7	The actual Riggs hall building load for 24 hours between 5/1/2022 and 5/2/2022. . . . .	89

5.8	Verification of the data-in-the-loop real-time simulation in the simulink dashboard. The reference power is matched by the measured power output from the emulated device. . . . .	89
5.9	The data visualization platform developed for the Openhistorian data processor based on a Grafana dashboard. . . . .	90
5.10	The real-time simulation output comparison from the simulink dashboard. (a) The system voltage without active reactive power contribution from the PV plant. (b) The system voltage with active reactive power contribution from the PV plant. . . .	91
5.11	The real-time simulation output comparison for a 200 second simulation (4.24pm, 5/2/2022) where the IoT changes the operation mode from no PV, PV, and PV with vvc. The top plot shows the variation the PV node voltage and the bottom plot shows the corresponding variation in load and PV plant power. . . . .	92
5.12	The real-time simulation output comparison from 7 pm to 8 pm on 5/2/2022. Here the system is operated with standard vvc. The top plot shows the variation the PV node voltage and the bottom plot shows the corresponding variation in load, PV plant active and reactive power, as well as the net power import. . . . .	93
5.13	The real-time simulation output comparison from 7 pm to 8 pm on 5/2/2022. Here the system is operated with a fixed power factor of one. The top plot shows the variation in the PV node voltage, and the bottom plot shows the corresponding variation in load, PV plant active and reactive power, and the net power import. . . . .	94
5.14	The real-time simulation output comparison from 7 pm to 8 pm on 5/2/2022. Here the system is operated with optimized vvc. The top plot shows the variation the PV node voltage and the bottom plot shows the corresponding variation in load, PV plant active and reactive power, as well as the net power import. . . . .	95

# Chapter 1

## Introduction

The last decade has seen a significant reversal in public policy where sustainability and energy security have become major global energy policy driving forces [4]. This is a significant reversal from the last century, where the globally dominating energy policy was thematically centered around enhancing access to electricity. This shift in policy is best illustrated by the fact that in 2021 two of the world's largest global greenhouse gas emitters in the US [5] and China [6] have set official policy goals to reach net-zero emissions in 2050 and 2060, respectively.

In broader terms, a sustainable power sector energizes the economy in a manner that does not adversely impact the environment throughout the life cycle of the technology used for all sub-sectors it is involved in, starting from power generation to power distribution. Sustainable technologies are expected to have minimal impact on the local environment and the global environment in the short-term and very long-term time scales throughout their life cycle, starting from production, application, and end of use.

However, much ground needs to be covered to reach the goal of a globally sustainable energy sector, as evident by Fig. 1.1. It is clear from this plot that the total energy consumption is growing across the world. Additionally, the consumption of fossil fuels is also expected to increase proportionally. There is an increasing trend in the 'other' category, which covers the renewable sources such as wind and solar that are increasingly taking a more significant portion of the consumption, predominantly in the electricity sector. The challenge, therefore, consists of two steps, the first step is to stop the increase in fossil fuel consumption, and the second step is to decrease the total consumption of fossil fuel-based energy sources.

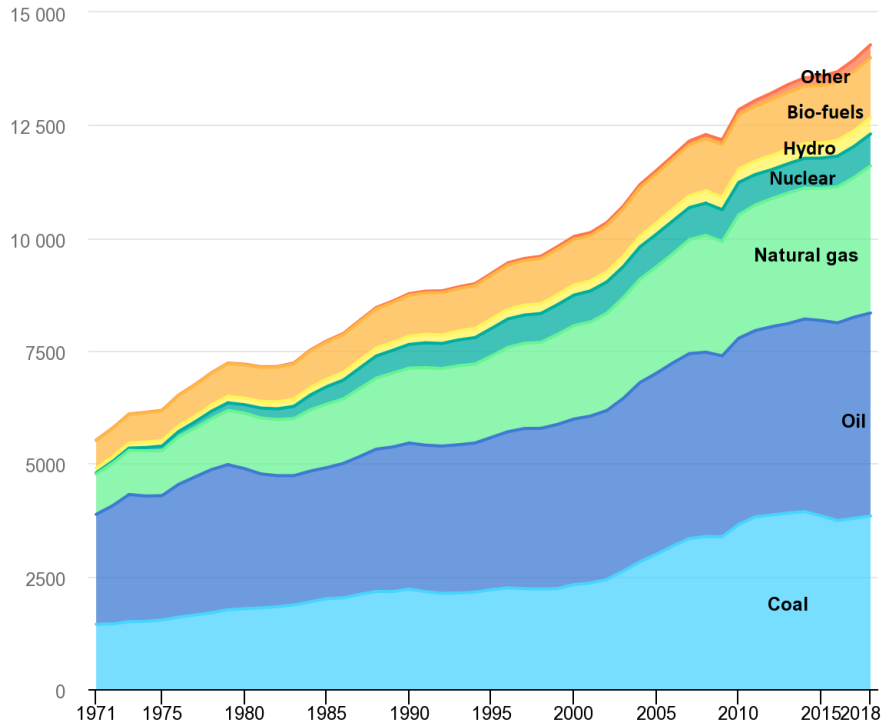


Figure 1.1: The world total energy supply by source from 1971 to 2018 [1] (y-axis is Mtoe)

When the electric sector is zoomed in, the image is less stark, as shown by Fig. 1.2. In this case, the electricity consumption is still increasing. However, fossil fuel-based generation has started showing a decreasing trend with the shortfall provided by exponentially increasing non-conventional renewable energy sources. This plot also shows the change required to achieve the net-zero goal.

When comparing the electric generation mix from 1973 and 2019, shown in Fig. 1.3 and 1.4, it is clear that it is possible to change the generation mix towards the net-zero goal. The plots show that oil-based generation has changed from 25% to 3% and non-conventional renewables from 0.3% to 11%. Even though the trends are encouraging, it is also clear that coal has stayed steady, while natural gas had an almost twofold increase in the mix. These trends show that the task at hand is very challenging but provides hope that net-zero is still an achievable goal.

Achieving a green and sustainable energy sector requires a two-prong strategy. The first strategy is to decrease the total energy consumption across all sectors. This mainly requires increasing efficiency in human activity. Optimizing transportation, housing, agriculture, and other vital sectors will play a significant role in realizing this strategy. This strategy does not necessarily focus

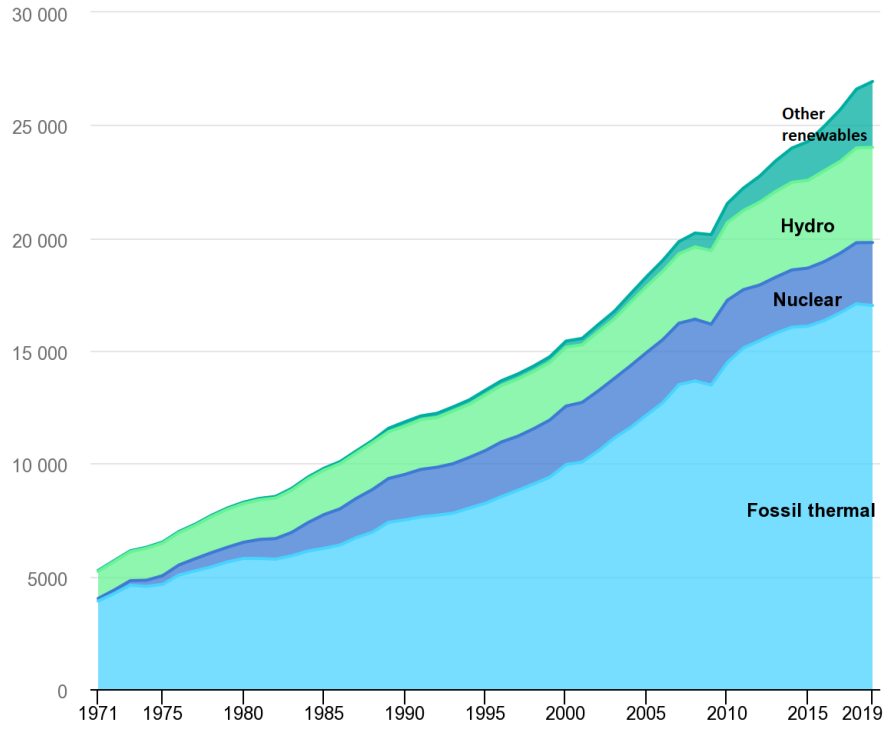


Figure 1.2: The world total electricity consumption by source from 1971 to 2018 [1] (y-axis is TWh)

on the source but on the quantity of energy used. The second strategy is focused on creating a greener energy landscape for the energy life cycle.

This involves different power generation, transmission, distribution, and demand-side management technologies. A fundamental shift in the processes involving these four areas is required to create the magnitude of changes set by policy goals. The design framework for these new technologies requires addressing any cultural differences. For example, the importance of privacy and cyber-security has become a major cultural shift observed in society. These technologies should be designed inside frameworks that provide security by architecture to be acceptable by the community. For example, large centralized corporations will have difficulty taking action related to sharing user data and control of user resources, even if it is feasible from a technological standpoint. This will also help create less infrastructure requirement, secure data flow, and less fragile and resilient systems that are hardened by design architecture against the increasing cyber-physical threats [7].

Technologies for both generation and transmission have been the focal point for greening the grid for a long time [8] [9] [10]. As a result, existing investments, policy directives, research

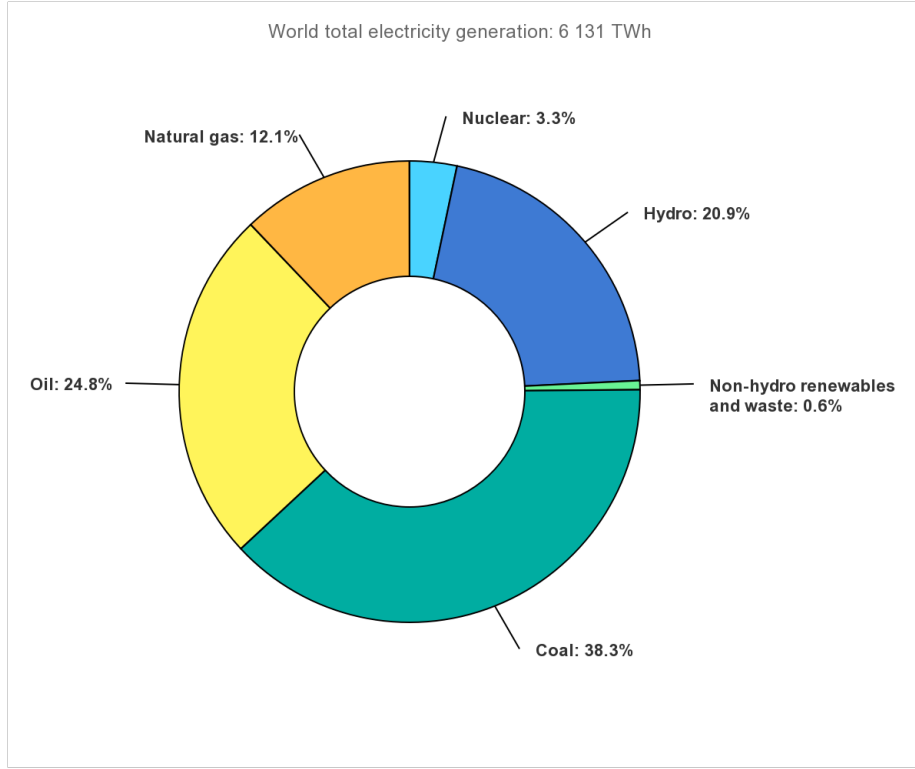


Figure 1.3: The world total electricity consumption in 1973 [1]

enhancements, and mature technology support these two parts of the power system. Unfortunately, due to the last mile nature of the distribution system, it has not been considered an important component of the grid. However, currently, the distribution grid has regained attention from a policy perspective, as well as a technological perspective [11]. With the decreasing cost of distributed energy resources and increasing penetration of distributed energy, it has come to light that the key to the next generation lies in addressing the challenges of the distribution grid. It is clear now that understanding operating, planning, and optimizing the distribution grid needs to be re-imagined and requires revolutionary work. For example, the fundamentalism assumption that the load is an uncontrollable variable or that electricity tariff should not be time-dependent needs to be changed. If we assume that the load is fixed and the economic signal that drives the demand is also fixed, it is not possible to operate a grid with high DER penetration. The costs will be high, and the reliability of this use case will be low. On the other hand, if the optimization is executed with open-ended variables, then remarkable enhancements in cost, sustainability, and reliability can be achieved [12].

The major drivers of this transformation are,



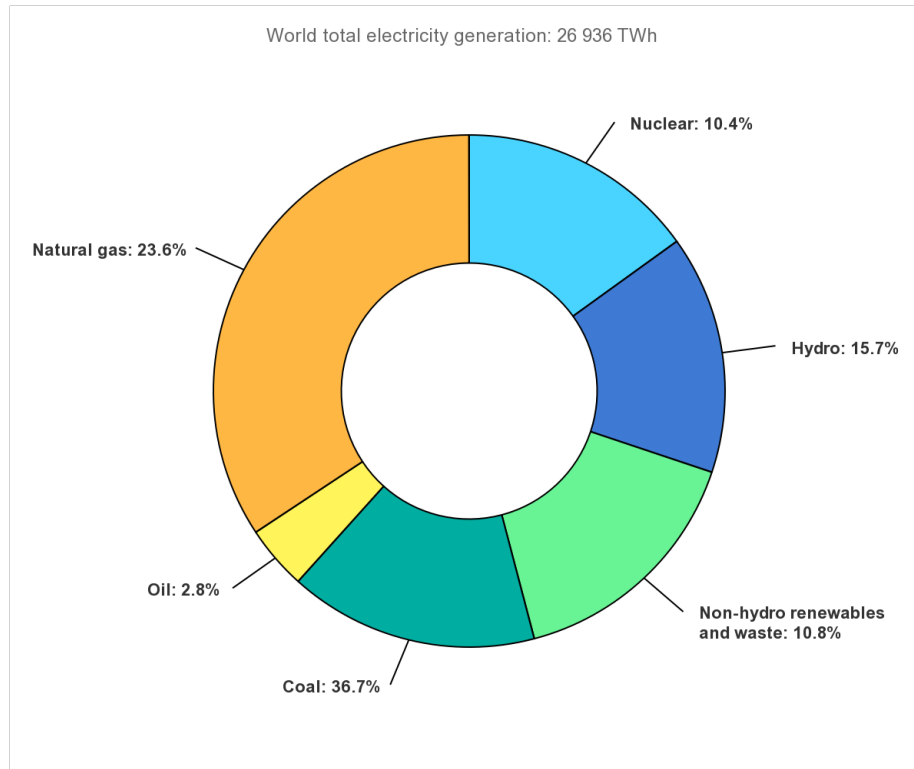


Figure 1.4: The world total electricity consumption in 2019 [1]

- Decreasing cost of non-conventional renewable energy resources
- Increasing penetration of distributed energy resources
- Increasing focus on the environmental impact of the energy sector, resulting in the integration of environmental cost in the planning process
- Increasing flexibility in the distribution network, including the ability to control household loads
- Increasing service providers who service the optimization requirements with a high level of efficiency
- Societal and cultural changes increasing the need for self-sufficiency in energy consumption
- Increasing societal need to enhance privacy and cyber-security

In 'greening the energy grid,' many essential areas in technology require innovations. Among those different options, greening the distribution grid takes prominence since it is one of the most

challenging problems with a long way to go in its transformation. This is due to reasons such as a large number of components, large investment requirements, and closeness to people’s lives. The current distribution grid has been static, and most parts operate under borrowed time. The required cost in resources and time is also significant. On the other hand, the expectation for distributed energy resource integration, demand-side management, and electrification of personal transportation pivots on the ability of the aging distribution grid to support those technologies. Technology plays a pivotal role in the distribution grid in the energy transition. This dissertation looks at new approaches to model and control the power distribution system.

This dissertation aims to blend state-of-the-art methods with artificial intelligence and cellular computational network methods to demonstrate feasible approaches to model, operate, and optimize the control of DER integrated future distribution grids. Therefore, the overarching goal of this dissertation is to contribute to the development of an intelligent power distribution system.

## **1.1 Research challenges in the Electric Power Distribution Systems**

Some of the research challenges that need to be addressed in order to advance the field are:

- Modeling a large distribution system based on predominantly data while ensuring privacy and model accuracy.
- Adapting a system model dynamically when the distribution system changes over time.
- Controlling a distribution system to be resilient and efficient.
- Operating and controlling a distribution system fairly while ensuring optimality.
- Demonstrating the application of smart control methods in a realistic power distribution system simulation platform.

## **1.2 Research objectives of this dissertation**

This dissertation aims to advance distribution system modeling, operation, and control by developing new approaches. The research objectives are to:

- Develop a distributed data-driven modeling method for power distribution systems.
- Develop a distributed volt-var curve optimization method for distributed energy resources in power distribution systems.
- Develop a real-time power distribution system testbed and demonstrate the ability to perform optimization in a real-time environment.

## 1.3 Main contributions

### 1.3.1 Distribution system modeling

A distributed data-driven distributed ( $D^3M$ ) framework to model a power distribution system based on CCN has been proposed to be used for distributed energy resource-rich power distribution system [13]. The formulation is demonstrated on a realistic power distribution system with significant voltage variations, unbalanced loads and lines, and a 100% capacity in DER penetration. The presented framework accurately models the power distribution system in a distributed data-driven manner. Additionally, it ensures the privacy and security of sensitive utility and user data used for model development.

### 1.3.2 Distribution system control optimization

A scalable distributed optimization framework for concurrent multiple volt-var curve optimization is proposed [14]. This method is based on creating a cellular computation network representation of the electric power distribution system with DERs, which allows for distributed data-driven modeling and optimization of the distribution system. A ranking method for cell prioritization for optimization has been developed. Cell prioritization for optimization based on a formulated impact ranking criteria improves the computational throughput for determining optimal volt-var curve parameters. The application of DOF on a modified IEEE 34 bus test system with 100% DER penetration has been illustrated. The demonstrated approach allows for distributed optimization of the volt-var curves of the power distribution system, allowing for enhancing voltage control of the system and thereby allowing for increased penetration of distributed energy resources.

### 1.3.3 Distribution system real-time testbed

A real-time testbed to model the core operational components of an existing power distribution system in a laboratory is proposed. The real-time testbed integrates actual data and components and allows the real-time digital model to include the cyber-physical characteristics. The implemented model includes hardware-in-the-loop, control-in-the-loop, and data-in-the-loop functionalities. A new data-in-the-loop (DIL) method is presented, which can project data from live loads and generators to the real-time simulator for more realistic simulations. The data in the loop method is implemented on this testbed, and its real-time simulation capabilities are validated. The control capabilities of the emulated PV plant are also demonstrated, and the results illustrate how this testbed can represent a comprehensive cyber-physical real-time simulation of a power distribution system.

## Chapter 2

# Transformation of the Power Distribution System

### 2.1 Introduction

The power sector has grown with minimal and gradual changes for over a hundred years. This transformation can be accurately characterized as an evolution. However, the power sector is now undergoing a revolutionary transition fueled by rapidly evolving core and support technologies, societal and cultural changes, and energy policy changes, some of which were discussed in chapter 1.

The two visible revolutionary changes are unidirectional power flow in the distribution system changing to bi-directional power flow and passive operation and control of the distribution system changing to dynamic operation. Since the revolutionary changes are occurring in the power distribution system, it has become a pivotal part of the power sector transformation. This chapter describes the changes that have occurred thus far, currently occurring changes, and the further necessary changes in the future to complete the transformation. It will look at the core and support technologies that empower these changes and how these are helping to change the power sector. This chapter will detail the strengths and weaknesses of the different state-of-the-art technologies and the areas that need addressing for a successful transition.

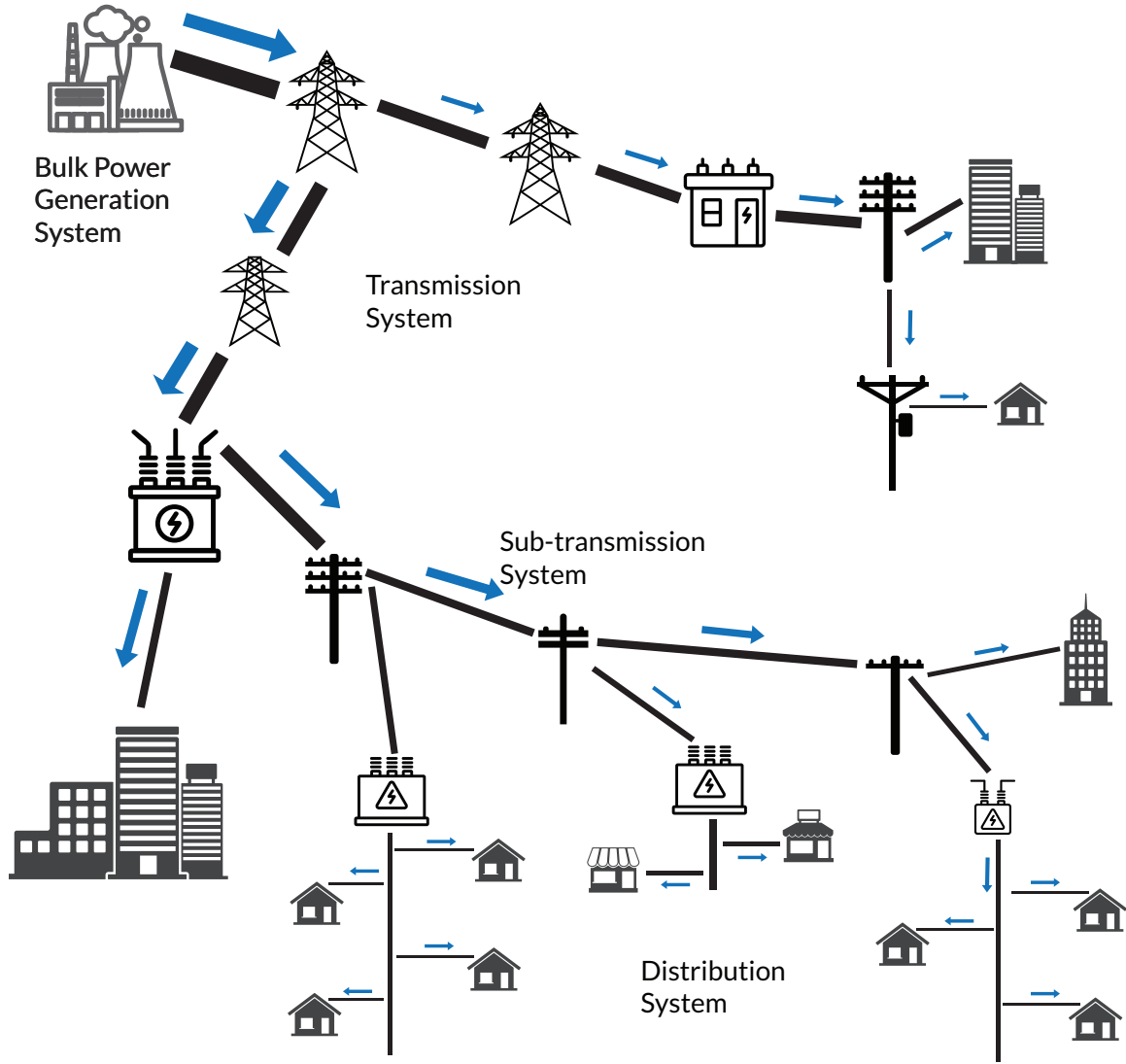


Figure 2.1: A passive/conventional distribution system. The power flow is uni-directional [2].

## 2.2 The power distribution system

### 2.2.1 Traditional power distribution system

In traditional power systems, distribution networks are predominantly operated as passive circuits with the assumption of uni-directional power flows. The typical structure where power is received in bulk from the transmission network and distributed to individual consumers (e.g., houses, businesses, factories) through a radially connected system of poles and wires (commonly referred to as feeders) is shown in Fig. 2.1.

Distribution networks often have limited observability and controllability since it is historically a passively operated infrastructure. Therefore, they are built to cope with the worst-case scenarios without external interventions. Furthermore, as new issues develop over time, the predominant solution is network augmentation (where the existing infrastructure is upgraded based on the newly determined worst-case scenarios). This conservative approach is a non-intelligent planning and operation approach. Distribution networks that are operated based on this design philosophy are passive.

Among the technical constraints considered in the day-to-day operation of distribution networks, two dominant ones are voltage and thermal limits. This means ensuring voltage limits at critical nodes (e.g., consumer connection points) and thermal limits of network assets (e.g., distribution lines and transformers) are not violated during critical periods. When distribution networks only cater to demand, in a typical passive system, these criticalities (or worst-case scenarios) correspond to periods with maximum demand since it results in the most extensive power flows and voltage drops in distribution feeders. There are traditional ways of managing voltage and thermal issues in passive distribution networks. The existing components are designed according to the expected peak demand, preventing thermal overload from a thermal capacity perspective. For voltage management, apart from sizing the conductor appropriately to minimize ohm losses and voltage drops, a few other mechanisms are in place. For instance, primary substations are equipped with on-load tap changers, which are autonomous devices that adjust the transformation ratio of the corresponding transformer (and thus the voltage) according to the variations in demand throughout the day. Furthermore, distribution transformers are often configured to provide a voltage boost to compensate for voltage drops in LV feeders. Capacitor banks and voltage regulators are sometimes installed along a feeder for voltage regulation purposes. Usually, there are two distinct voltage levels in a distribution network: medium voltage (MV) and low voltage (LV), medium voltage comes up to the residence in the US, and the drop-down is in low voltage. There are significant differences in the design philosophy for the distribution system when comparing the US and other countries, where a significant amount of wires are connected at LV. MV feeders start from primary substations and extend throughout the serviced area. Since each primary substation can service a large geographical area, transporting electricity at the MV level has technical benefits such as lower losses. Typical voltages used in MV networks can range from several kV to tens of kV; for instance, the most common voltages are 13, 25, and 36kV in the US. Large consumers such as high-rise buildings and factories

are connected directly to MV feeders for increased power capacity (peak demands over hundreds of kVA). In contrast, LV feeders are created for groups of small consumers (residential properties along a street) and interfaced with the MV feeders through distribution transformers. The North American grid brings the distribution transformer closer to the consumer, which provides a better potential for expansion of capacity, but at a higher distribution service cost.

### **2.2.2 Distributed Energy Resources (DER) integrated distribution grid**

More and more renewable generation will be connected to the distribution network to decarbonize the electricity sector. Consequently, this is changing the source of energy generation in modern power systems: from large-scale, central power stations that are located away from urban areas to small-scale, decentralized power generating units that are connected to distribution networks [2]. The power generation within the nodes of distribution networks is referred to as distributed generation (DG). In LV distribution networks, the growth of renewable DG is due to rooftop solar PV systems [15]. Since rooftops are often under-utilized areas and have minimal obstruction from sunlight, they are prime locations for solar PV systems. Depending on the available area, the installed capacity can range from multiple kW (e.g., on top of residential properties) to hundreds of kW (e.g., on top of large factories). On the other hand, in MV distribution networks, renewable DG is often observed in megawatt-scale wind farms, and solar farms [15]. This is common in rural areas as land is more readily available and utilizing the existing distribution network decreases the need for additional transmission infrastructure to connect them.

### **2.2.3 New challenges**

The major challenges created by the DER integration into the traditional grid are,

- reverse power flows in distribution feeder that was designed for unidirectional flow causing severe voltage fluctuations in the distribution system
- uncontrollability of the distributed generation causing congestion and thermal limit violations in distribution corridors
- operator having null observability to the distributed generation (decreased situational awareness), causing challenges in maintaining the optimal generation resources



The non-dispatchable nature of renewable generation creates a mismatch between generation (from DG units) and demand within the distribution networks. As a result of the reverse power flows during periods of surplus generation, technical issues such as voltage rise and distribution congestion are becoming common in renewable-rich distribution networks [16, 15], as illustrated in Fig. 2.2. In most cases, voltage and thermal limits determine the technical viability of the DG installation capacity of a distribution system. Other technical issues, such as fault currents, relate to the protection of the power system and power factor, which relates to operational capability limitations of specific power distribution system components such as transformers.

For example, in residential LV feeders, maximum PV generation typically coincides with low residential demand (due to low occupancy on weekday work hours). It results in reverse power flow towards the distribution transformers. Since the existing infrastructure is designed to accommodate voltage drop (the opposite of voltage rise caused by reverse power flows), the potential of voltage violations is increased mainly during these periods of excessive reverse power flows. To fully use the available voltage operational region (usually a 5% from the nominal), the voltage at the head of a feeder is often boosted towards the upper limit. This strategy ensures that the voltage at the feeder's end remains within the lower limit during maximum demand. However, it can cause problems when dealing with reverse power flows since it results in little space to operate from the upper voltage limit. Additionally, unlike the typical domestic residential demand profile, which is not significantly correlated, peak generation from rooftop PV systems is largely correlated in any region. As a result, there is an increased risk of system congestion in the LV feeders and the upstream MV feeder. The underlying causes of technical issues in MV distribution networks largely follow the same principles as in LV networks: a mismatch between peak generation and peak demand. Additionally, due to longer distances of rural MV feeders (which is where most renewable generation is being connected to at the MV level), the corresponding infrastructure is weaker than their urban areas (more sensitive to voltage issues due to the higher impedances, and this causes higher magnitude of voltage rise for the same amount of reverse power flows). This significantly increases the risk of technical issues from excessive reverse power flows.

#### **2.2.4 State-of-the-art**

With the proliferation of DERs in the late 1990s (driven by state subsidies, net metering, and decreasing PV costs), the operators first observed the challenge of frequency fluctuations in their

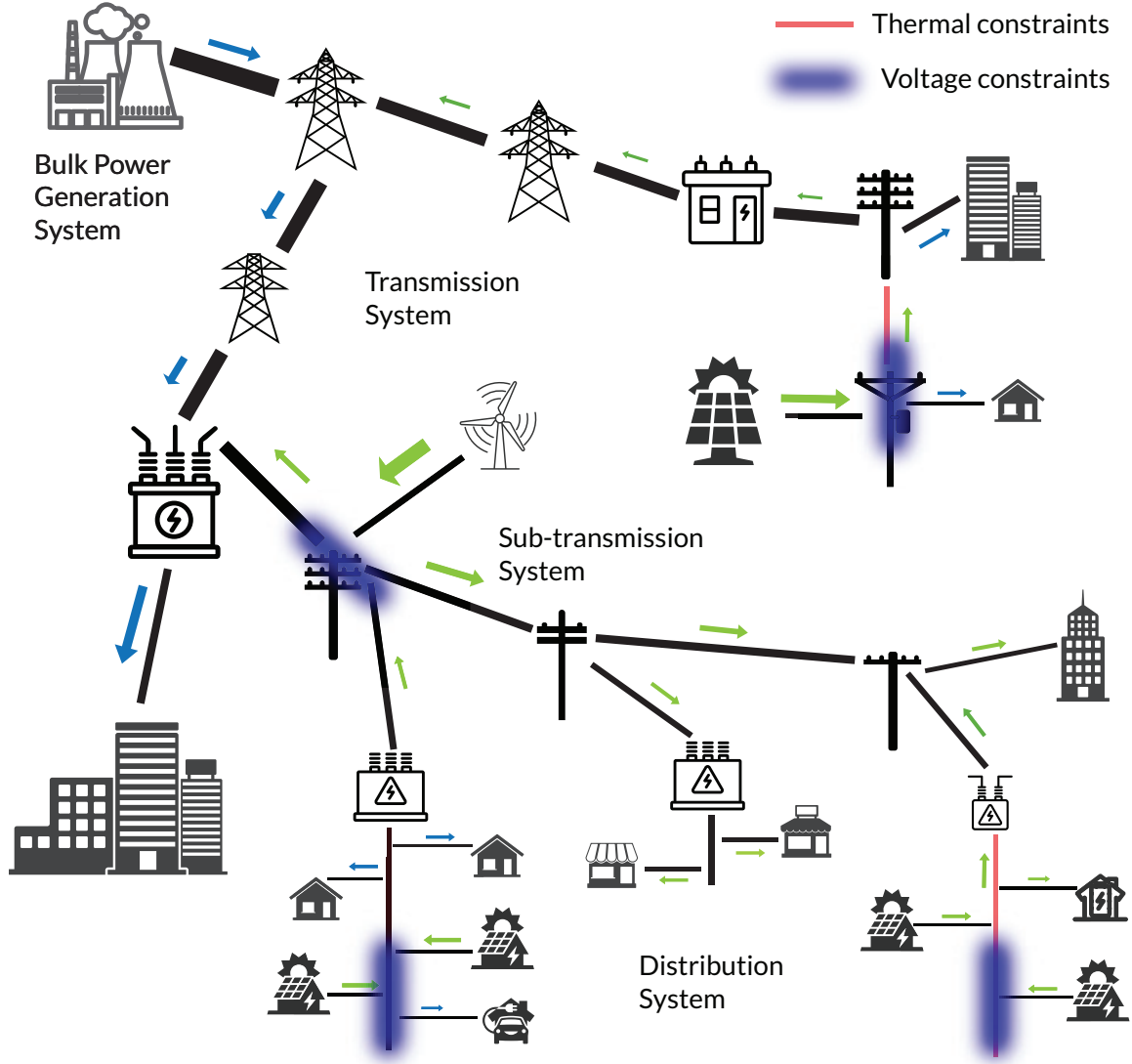


Figure 2.2: A traditional distribution grid with high penetration of DERs results in adverse operational challenges [2]

distribution feeders. In order to address this specific challenge, they came up with the idea of fully decentralized controls where the generation would react to the frequency of the system [17]. If the frequency is beyond the upper limits, the DERs were programmed to ramp down the active power generation, and at a specific limit, they were programmed to be disconnected. This helped in system frequency controls especially given that the system controller did not have observability or direct control of behind-the-meter distributed generation resources. Since frequency was a system-wide

measurement, this decentralized control approach was reasonable and only came into effect very rarely in high-impact low probability events.

However, with further penetration of DERs, the operators realized that the DERs were starting to cause voltage fluctuations in the distribution system. They observed many grid code violations when they measured the power quality in particular nodes. To address this, the first solution came up with expanding the planning and operational studies from steady-state power flow to Quasi-static time-series power flow, where the hosting capacity study became a significant component of the distribution network planning process [18]. Using the hosting capacity studies, the operators could define the boundaries for operation and ultimately come up with the maximum capacity of DERs that can be connected at any node in the system [15]. Many DSOs have started to host publicly accessible hosting capacity maps to ensure transparency, for example, Fig. 2.3. The challenge with these hosting capacity studies is that, unfortunately, it assumes homogeneous control functionality of the DER resources, which, even though compliant with IEEE 1547, results in severe underutilization of the value available from DERs.

With further penetration of DERs, many distribution systems showed large voltage fluctuations even if the hosting capacity requirements were met. To counter this challenge, the industry came up with volt-var controls for the DERs [19] [20], an innovation that emulated the operation of the generators in the conventional power generation system. Here, the flexible, reactive power capability of the inverters was used to create reactive power injection based on a generic reactive power capability curve. The IEEE 1547 standard [17] was amended in 2018 to include this capability.

However, an entirely decentralized control system that responded to voltage, which was typically a local variable, made the operation non-optimal and unfair [21]. Some DERs that are connected at high Thevenin impedance points had larger voltage variations and therefore had to contribute more reactive power while having to curtail the generation more than others. Therefore, in some critical networks, operators or owners have started to control it similar to the larger power generation system, using a Distributed Energy Resource Management Systems (DERMS) that centrally measures and controls the setpoints based on optimal power flow (OPF) algorithm. Since the OPF is computationally intensive, especially in three-phase unbalanced networks, innovations on relaxations, linearizations, and formulations, are being made to increase the potential scale of applicability [22] [23] [24] [25]. Examples such as [26, 21] have demonstrated that even simple optimization provides a significant enhancement to the conventional distribution network to provide

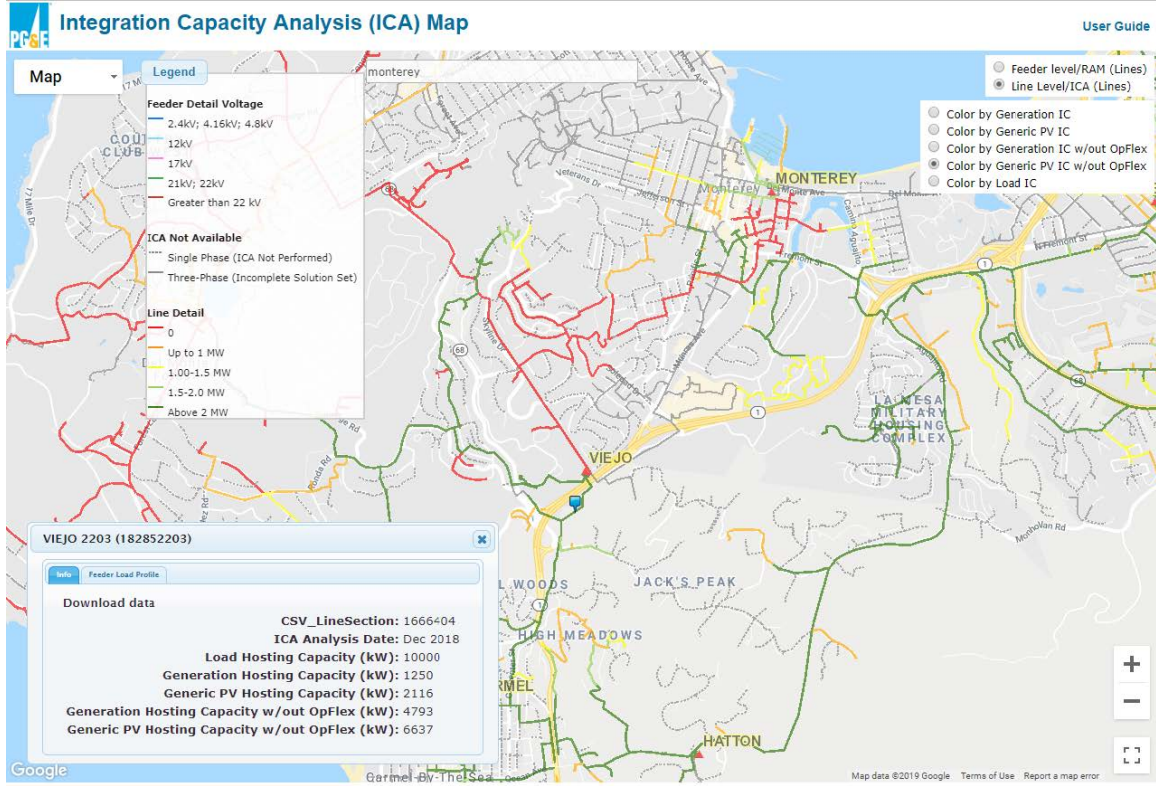


Figure 2.3: A map of available hosting capacity maintained and hosted publicly by Pacific Gas and Electric Company [3]

the capability to install higher penetrations of DG.

On the operation side, a more granular approach has been proposed in the last two years and was first presented in [27]. In this approach, the OPF study is carried and an optimal 'envelope' of the day ahead operations for DERs are calculated and sent to the DERs. The envelope defines the allowed operation region for these DERs for the day. Whereas this method could work for a small system, a centralized approach is not feasible for a generic application. This is due to the challenge of maintaining physical models of the distribution system that could consist of several thousand to tens of thousands of small nodes while gaining both the required observability of the system and the communication access/capability at the needed granularity.

Apart from the centralized control approaches, decentralized approaches based mostly on gradient-based network optimization theory has been presented in the last decade in [28] [29] [30] [31] [32] [33]. However, the work presented depends on three-phase balanced approaches to attain optimization. Additionally, the optimization requires a significant amount of data transfer and

iterations, even for a system with a single DER. Given that the number of DERs, the dynamic characteristics, and system unbalance are increasing, these methods are challenged to serve the present and future needs of the distribution grid.

The transition to active distribution networks will unlock new opportunities to support active bidirectional power flows instead of just relying on brute force reinforcement of existing assets. Furthermore, due to the increasing deployment of smart grid technologies such as communication infrastructure and smart meters [34], the observability and controllability of existing distribution networks have improved significantly, making their granular monitoring and control an increasing reality. Almost 100 million smart meters have already been installed in the US out of a total of 127 million consumers in the US [35].

### 2.2.5 Smart distribution grid

The rapid increase of DG in recent decades has changed the role of modern distribution networks at a fundamental level. Its functionality is no longer limited to serving as the last-mile power delivery to consumers. As a result, being passive throughout the life cycle from the planning stage through network augmentation may no longer be an effective norm but rather an exception. The challenges faced by the future distribution system can only be addressed by leveraging the 'smart' capabilities of new technologies and need to be addressed in a holistic manner using system engineering approaches. A 'smart' or an intelligent power system is defined by the below characteristics and functions [36] [37] [38] [39],

- has self-healing and fault tolerant capabilities:
- has plug and play capabilities
- is distributed by design
- operates with dynamic optimization
- integrates heterogeneous consumers, generators, and prosumers
- improves system reliability, power quality, security and efficiency

Therefore, an intelligent distribution system is a distributed system that allows for the connection of different sources and sinks without prior knowledge, allows for demand side and

supply side management, and optimizes dynamically, while operating optimally and efficiently. On the other hand, it is highly secure, making it difficult to attack. It can tolerate a high level of cyber and physical damage. When damaged, it can autonomously heal, providing a reliable and resilient service.

### **2.2.6 Journey from the smart distribution system to an intelligent distribution system**

In order to operate intelligently, the system needs the capabilities to monitor, store, and forecast, as well as share information; the system needs the capabilities to communicate with different entities, to calculate, learn, share understanding, schedule, take and implement decisions and well as adapt when the system environment changes. This will help the system operate in an efficient, optimal, reliable, and efficient manner.

In the context of the work presented in this dissertation, smart and intelligent distribution technologies can be differentiated based on defining characteristics. A smart system is defined by using past data (prior) to adapt its operation. For example, a set of policies that were optimized for different operating conditions and the relevant policy chosen for controls at a given time of control based on current operating conditions. This approach was presented in [40]. This provides better performance than a fixed control policy, an example of which is the standard volt-var control, which does not change based on different operating conditions. It provides satisfactory performance as long as the operating region is limited to design conditions. However, as elucidated earlier in this chapter, the power distribution system is constantly in flux and has varying spatial and temporal operating regions.

Intelligent technology is differentiated from smart technology based on its ability to learn on the go. It does not require prior knowledge to optimize since it can use available information and learn on the go using present information. The gained knowledge is also added to the current understanding of the system termed 'the well on learning.' When a system faces a previously unknown and unexpected change in the system itself or the external environment, intelligent technology provides the ability still to push the operational trajectory towards an optimal operation. Even when prior information is unavailable, the ability to learn on the go differentiates an intelligent technology from a smart one. An example of intelligent technology is [39].

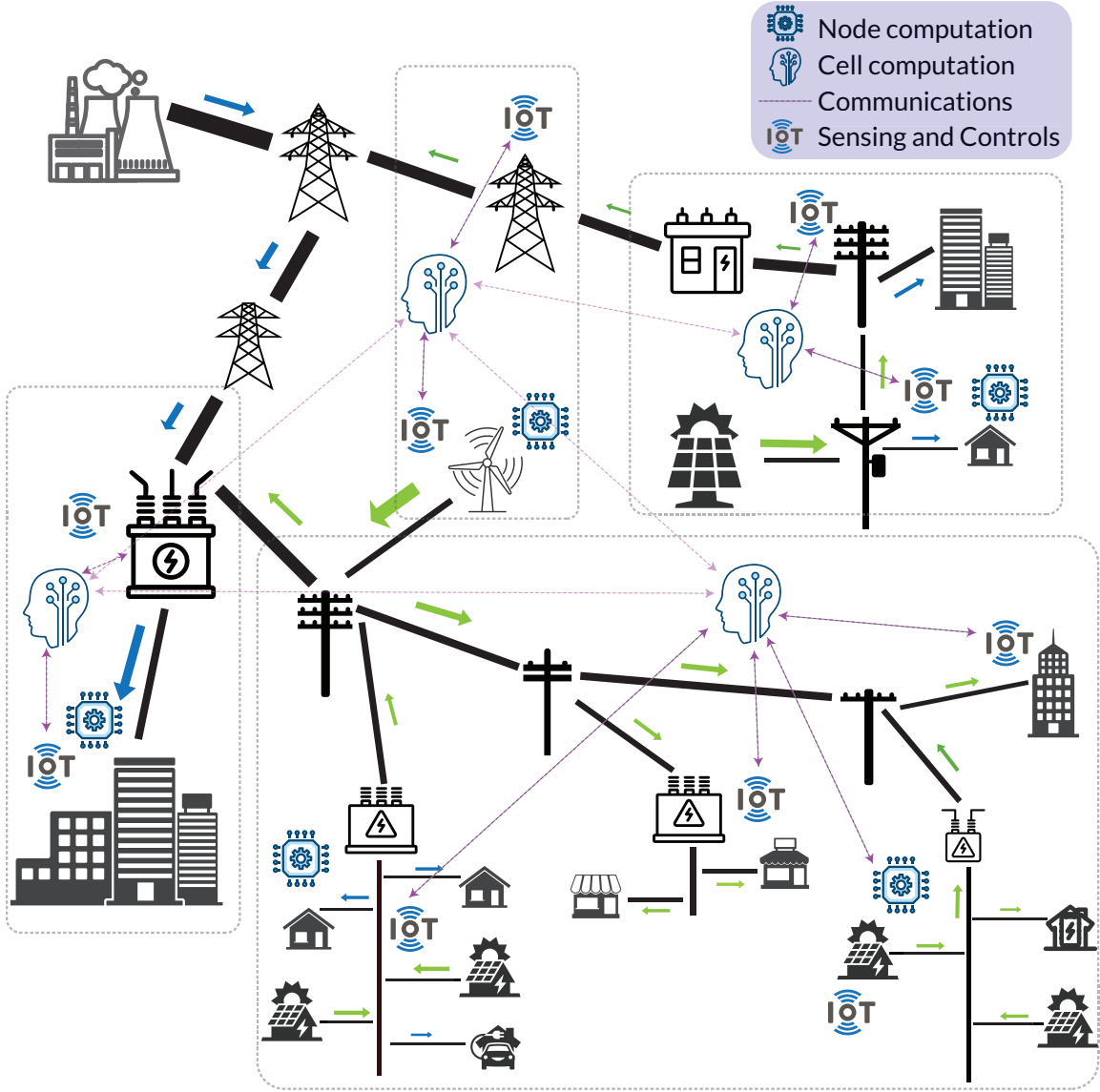


Figure 2.4: Conceptual component level overview framework of the smart power system

In order to be intelligent, it is required to leverage both cyber and physical technologies in a system thinking framework. The process for system design can be encapsulated into three main components; communication, computation, and controls, also known collectively as  $C^3$  [39]. The overview of the envisioned smart distribution system that includes the properties mentioned above is described in Fig. 2.4.

### 2.2.6.1 Communication

An intelligent distribution system requires information to be shared in a secure, reliable, and robust manner within different spatially distributed nodes. The ability to do so is critical to implementing an intelligent distribution system. The ability to share information is provided by the Communication component of the  $C^3$  framework. The dynamic distribution system will always be in flux with different loads, such as electric vehicles and HVACs connecting and disconnecting from the system. The sub-networks will also be connecting and disconnecting based on different operational conditions creating uncertainty even in the topology. The node-connected DERs will also add to the system dynamics with different active and reactive power injections based on operating conditions. Therefore, the first component of system design goes into ensuring sufficient sense-making. This has two sub-components, the first being the sensing and the second being the communication.

The current state-of-the-art provides for a plethora of options for sense-making. Micro Phasor measurement units are dedicated measuring units that connect to the distribution system. They can create phasor measurements of node voltage and link currents at a frequency up to 120 Hz [41]. Examples of other sense-making options are the integrated measurement units of components such as Advanced Metering Infrastructure, inverters, breakers, and IoT sensors that provide sensing at a granular level. Apart from these traditional sensors, visual sensors and drones that detect an adverse impact on the system such as wildfire, component thermal hot-spots, and line degradation, and sensors that provide input on the local weather conditions, also support the overall sense making mechanism.

The sensed system conditions need to be communicated in different layers to make use of them effectively. The communication shown in Fig. 2.4 using purple arrows includes the technology used for this part of the process. From a physical standpoint, a combination of technologies spanning from 2G to 6G, Ethernet, fiber, Zigbee, and satellite can be used for communication. The communication could be on dedicated or shared and public channels. The cyber part of communication technology also plays a critical part in ensuring its effectiveness. Due to the large amount of raw data created by the many sensors, at a frequency that could be up to 120 Hz, it is not possible to collect and make effective use of it at one central location. Therefore, it is vital to use different approaches to ensure that the data is processed locally and only communicated when necessary. An example of this is the phasor data concentrators and software such as OpenPDC. Therefore, an



optimal sense-making framework is a critical component of the smart distribution system.

This diagram shows that the smart distribution consists of two layers, the first being the physical layer and the second being the cyber layer. The different components of each layer and its connectivity is shown in Fig. 2.5. The cyber layer includes a significant part of the communication, control, and computation functionalities. The physical layer represents the physical equipment of the system. The distribution system shown is a radial system.

Each node of the system includes a communication hub, a prosumer, energy converters, Advanced Metering Infrastructure, and a node Energy Management System (EMS). These infrastructures are interconnected in the physical layer. The operational information from the physical layer is communicated to the cyber layer. Each node can communicate with the cell IOT in a star configuration. The cell communicators are assumed to be connected in a ring configuration to show the difference between the connections in the physical and cyber layers.

The cyber layer components collect and share information and use the computation and control components situated in that layer to share the control output with the physical layer components. The typical CPS diagram shown in Fig. 2.5 includes computation and control blocks. These functionalities are part of the cyber layer's node and cell blocks. A smart distribution system has interdependencies between the components in each layer and interdependence in components between the two layers.

### **2.2.6.2 Computation**

Computation consists of technologies that enable solutions to many complex problems that must be solved. The modeling approaches and the control approaches in the smart distribution system require computation that is fast and efficient. The computation is expected to be heterogeneous since the type of optimization is diverse and has different functions across the system. Traditional approaches based on mathematical programming and computational intelligence-based approaches such as particle swarm optimization will all have a role in the different computations required to be carried out across the system. Additionally, there will be a requirement for both distributed and central computation. The computations could be carried out in IoT-based embedded computation systems, as well as large high-performance computers. It is a critical enabling technology in realizing an intelligent distribution system.

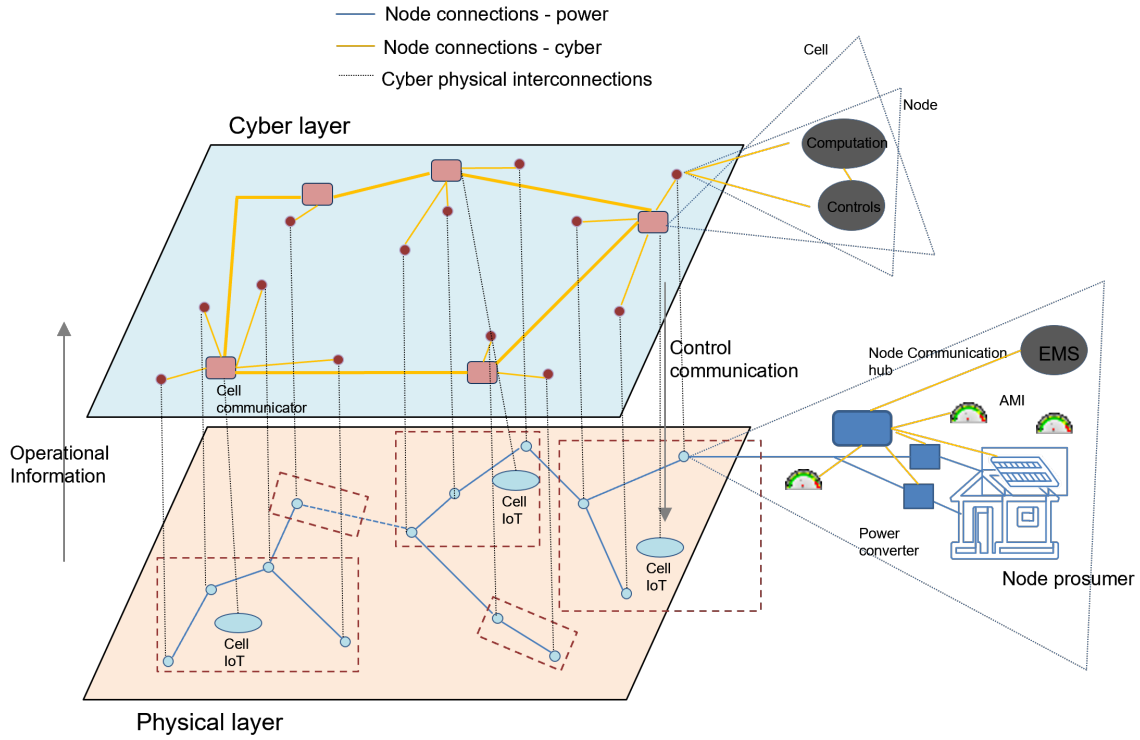


Figure 2.5: Cyber-physical framework of the smart distribution system

### 2.2.6.3 Controls

The controls are shown in orange in the conceptual framework. The computations support the smart controls of the devices to influence the operation of the different levels of the system to realize their goals. These include the adaptive and self-learning methods required to manage the fast-changing environment and provide the capability to learn, adapt, generalize, and associate the system controls. Examples of digital controls are system components such as line breakers and consumer loads such as HVAC, dishwasher, EV Charger, or Washer/Dryer. These can be controlled based on the control signals that are transmitted. Examples of equipment that are controlled using analog signals are power electronic transformers, PV inverters, D-Statcom, and batteries. These controls are shown using an Orange arrow in the conceptual framework.

The controls can be hierarchical, distributed, centralized, or a combination of them. The conceptual framework shows nodal, cell, and system-level controllers. Though cell controls are shown as centralized, the same objective can also potentially be realized using a decentralized approach. The central controller will be a virtual controller or a representation of the cell level control function.

Similarly, system-level control is shown to be decentralized, but it could also be carried out via a central controller. Each cell or node layer could have different objectives. This provides the ability to democratize the operation. It is possible to have competing objectives on different levels. The controls will then move to an equilibrium based on the many different objectives that superimpose on each other. Based on the changing operation scenarios, the controls will have to adapt and optimize in real-time and morph their characteristics such that their effectiveness is maintained in highly dynamic operations.

In combination, these technologies lead to the creation of a Computation System Thinking Machine (CSTM) [39], a framework that encapsulates all three technologies and empowers the intelligent operation of the power distribution system. It is clear that to optimally utilize the various sources of flexibility in the distribution network, advanced intelligent or 'smart' approaches must be used to operate and control power distribution networks. The 'smart distribution grid,' where the network issues are solved at the operational stage by the real-time monitoring and control of network assets and participants, is the best way forward. It helps to leverage the various sources of flexibility in real-time, and distributed controllers can optimally use the capability of installed infrastructure and provide dynamic and optimal operation.

## 2.3 Summary

The rapid growth of distributed energy resources has changed the unidirectional conventional power distribution system to a bi-directional system. This has given rise to challenges for frequency regulation, voltage control, and congestion control. In order to address these challenges, it is critical to move from evolutionary to revolutionary technological innovations supporting the distribution grid. The main evolution required is to create an intelligent distribution system. An intelligent system requires the functions and capabilities for self-healing, plug and play, distributed operation, dynamic optimization, and efficient, reliable, and resilient operation.

The current state-of-the-art consists of decentralized and centralized approaches, but these methods have flaws that make them unsuitable for generic application to a typical distribution system. A decentralized control approach will not be fair and non-optimal, while centralized options have scalability, resilience, and cost challenges. Therefore, the research gap in realizing a smart distribution grid is significant.

The smart distribution requires leveraging many different cyber-physical technologies that can be encapsulated in the  $C^3$  framework. This framework consists of communication: which refers to sensors and communication infrastructure that helps to make sense of the situation/ provides situational awareness; computation: which refers to the different technologies that help solve the optimization and control problems across different layers; and control: which provides the ability to adapt, learn and relearn, as both the environment and the system change spatially as well as temporally.

## Chapter 3

# Distribution System Modeling

The power distribution system is increasing in importance and complexity as a result of the exponential growth in the adoption of smart grid technologies. The ability to model the power distribution system is critical to ensuring a smooth transition to a sustainable power system. This study presents a distributed data-driven framework based on Cellular Computational Networks (CCN) for power distribution system modelling where the CCN framework facilitates for system decomposition. The learning in CCN is distributed and asynchronous, thus adaptive models can be developed. The computational engine of the CCN cells can be based on data-driven, physics-driven, or a hybrid approach. The CCN based distribution system modelling secures the privacy and security of the sensitive utility information, thus allowing third-party application providers access to system models and behaviours. The application of a CCN based power flow model is illustrated on a modified IEEE 34 test system. Typical results show the suitability of the new approach in modelling the sample distribution system, as well as its enhanced performance when compared to the centralized modelling approach.

### 3.1 Introduction

The power distribution system has continuously played a passive and relatively minor role in the power system since the advent of electricity as a commercial power source. However, in the last decade, the role of the power distribution system has started to increase in importance as a result of the adoption and acceptance of smart grid technologies, such as distributed energy resources, energy

storage, demand-side management, dynamic electricity pricing, advanced metering infrastructure, and electric vehicles. Additionally, there is a clear indication that extreme weather events, such as storms and forest fires are increasing. The extreme events disproportionately impact the distribution system. For example, when a storm impacts a geographical area, many assets get damaged and have to be replaced. In this kind of situation, the records of the replaced assets tend to get lost. It is clear that distribution system assets change much more frequently and are more numerous [42, 43]. As a result, the distribution system, formerly known to be a stable and predictive power system asset, is transforming into a class of power system assets that undergo frequent and unpredictable changes. With time, there is a high probability that the power distribution system physical models will be incomplete and inaccurate, causing divergence from the ground truth.

Generating and maintaining accurate models for the power system is a labor-intensive process. The transmission and distribution system has a limited number of components, is comparably static, and each component has a high impact on the system operation. On the other hand, the distribution system consists of a large number of physical assets and network nodes. Typically, the impact of a single component in a distribution system is negligible. However, when small changes occur across a significant portion of the network, the impact on the system is significant. Therefore, when developing and maintaining high-fidelity models, an automated approach with minimal human interference is preferable to a human-based manual approach. With the help of these high-quality models, the distribution networks can be optimally operated and controlled, paving the way to a fast, efficient, and economic transition from the traditional static distribution system to the smart, sustainable distribution system of the future [44, 45].

Fig. 3.1 illustrates the different options available for modeling a distribution system. The two main approaches are data-driven and first principle modeling. First principle (physical model) based approaches, which are currently dominant, have two sub-categories. The first category is exact modeling. The complete first principle model is used without any assumptions in this approach. As the size of the distribution system gets larger, developing an exact model from the first principle becomes a complex and labour consuming task. The second category is approximate modeling. Approximate modeling relaxes some of the non-linearities of the models to decrease the model computational burden and development time. There are three subcategories of approximate modeling; linear, non-linear, and hybrid. The non-linear approach includes approximations for simplification, but not up to the level of full linearisation, and therefore is more accurate over a more extensive

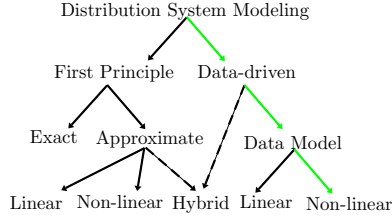


Figure 3.1: Different approaches to distribution system modelling.

operation space. When fully linearised, the model becomes computationally efficient, yet the performance (accuracy) of the model over the state space could degrade significantly. Presently, the fully linearised approaches are widely used for applications, such as distribution system optimal power flow [24].

Data-driven modeling consists of two categories. The first category is pure data model approach, where only data is used for modeling. In this case, both the structure and the parameters are learned from the data. The other approach is hybrid modeling, which uses a combination of both first principle as well as data-driven modeling [42, 43, 46]. In this case, the first principle model provides information about the structure, whereas some parameters are learned using data. The data model approach has two categories based on its internal structure. The linear and non-linear categories are based on the type of structure used in the data model. The non-linear data model approach, shown in green in Fig. 3.4 is applied in this chapter.

In addressing modeling challenges in complex systems, an increasing number of data-driven methods are being applied with success across a broad spectrum of fields [47]. Historically, the two key challenges to using a data-driven approach have been a lack of data and computation resources. However, due to the proliferation of Advanced Metering Infrastructure (AMI), [48], as well as high-end sensor technologies such as micropmus [49], there is now an abundance of data. Power system operators are challenged to find ways to use these data in an intelligent manner [50]. The other key challenge, which is finding the computational power requirement, is becoming inconsequential due to the increasing accessibility of distributed computing resources [51]. The main barriers to data-driven modeling of the distribution grid are decreasing. These supporting technologies are predicted to grow exponentially, strengthening the opportunity to use data-driven modeling techniques even further.

The Distributed Energy Resource Management Systems (DERMS) are used to ensure optimal, reliable, and resilient control and operation of the distribution grid. DERMS applications

require high-fidelity models. The coarse models available through physical modeling are not ideally suited to be used in these highly sensitive control applications [52]. A physical model is static by nature, and it is inappropriate to be used in such an environment. To reap the best value out of DERMS, the modeling should be fully automated, the model should adapt automatically, and the model should be able to learn even minute variations in the systems. A data-driven modeling approach provides a salient framework to realize these requirements.

Another increasingly prominent aspect is the need to conserve the security and privacy of the sensitive information of the distribution system since it is an interdependent critical infrastructure. If a physical model is used, the approach needs to be centralized, and the information about every part of the system is needed centrally for the modeling and simulation process. On the other hand, a data-driven model could be set up to secure all sensitive system information since it will not include explicit system information. Additionally, when a distributed approach is used, the usage of topological and parameter information can be limited to the location of interest with a minimal amount of global information exchange. This will make the grid more secure. If an anonymized model that is disaggregated from the vulnerable network data can be provided, it will open up more opportunities to openly exchange grid models. This will provide the opportunity for external service providers to be involved in system optimization and control activities [53].

A majority of the power distribution system power flow computations are currently based on centralized physical modeling approaches. However, there is a growing need for more accurate and faster computation. This requirement merits the investigation of the possibility of using a distributed approach for modeling and simulation. The work presented in [54] provides a comprehensive framework to accomplish decentralized computation. However, this approach is still based on a physical model, and the decentralization is implemented only on a computational level, leveraging parallel processing to speed up the computation.

Data-driven black-box-based modeling can be implemented as a centralized model, but the requirement for asynchronous learning and computational and data constraints limits the application space. The nature of the distribution system calls for a distributed solution, ideally where the learning is asynchronous and distributed.

These factors make the distributed data-driven approach an attractive network modeling option for distribution systems. The power of data-driven modeling is optimum when it is used side by side with a physics-based model. This will help keep the data-driven robust in regions where



historical data is minimal while increasing the situational awareness of the distribution system operator.

The main contributions of this chapter are,

- designing a data-driven distributed ( $D^3M$ ) framework to model a power distribution system based on CCN
- demonstrating the framework on the modified IEEE 34 bus test system
- demonstrating the ability of the framework to ensure privacy and security of sensitive utility and user data

The chapter is organized as follows: Section 2 gives an overview of the problem statement, Section 3 describes the methodology in detail, and Section 4 describes the simulations used to demonstrate the framework. Section 5 shares the results and the insights attained through the results. Section 6 concludes the chapter.

## 3.2 Problem formulation

Consider a radial power distribution system topologically represented by the tree graph,  $\mathcal{T}$ , consisting of the set of nodes collected in the set  $\mathcal{N}$ , with  $\mathcal{N} = \{0, 1, \dots, N\}$  (point of common coupling (root) is denoted by node 0).

Let the power injection at node  $k \in \mathcal{S}$  be represented by  $S_k$ , where  $\mathcal{S} \in \mathcal{C}^{1 \cdot M}$ . Let all the power injections be collected in set  $\mathcal{S}$ , where  $S_k \in \mathcal{S}$ . Let the complex voltage of node  $k$  be denoted by  $V_k$  be collected in set  $\mathcal{V}$ , where  $\mathcal{V} \in \mathcal{C}^{1 \cdot N}$ .

Let there exist a mapping function  $f$  where  $f : \mathcal{S} \mapsto \mathcal{V}$ . The problem then resolves to function estimation, finding an estimate for  $f$ , given by  $\hat{f}$ , when a set of  $m$  values of the  $f$  mapping is known. The general function class for this mapping is denoted by  $\mathcal{F}$ , and  $\hat{f} \in \mathcal{F}$ . There should exist in  $\mathcal{F}$  a choice for  $\hat{f}$  that can accurately reproduce a given  $\mathcal{V}$  in the output space. Different types of systems will have a different general function class  $\mathcal{F}$  since each system has a set of unique characteristics [55]. Therefore, the problem can be formulated as the minimization problem stated in (3.1).

$$\min_{\hat{f} \in \mathcal{F}} \frac{1}{M} \sum_{i=1}^M (f(\mathcal{S}) - (\hat{f}(\mathcal{S}))^2 \quad (3.1)$$

### 3.2.1 The power flow problem

The conventional approach to solving for power flow is based on the power flow equations [56]. There is a prerequisite to identifying the network parameters and topology in this approach. The system bus admittance matrix,  $Y$ , can be derived using well-known approaches if the topology and parameters of the system components are known [57]. This set of equations described in (3.2) represents the entire power flow problem. For any given  $\mathcal{S}$ , the solution of these equations will provide the corresponding  $\mathcal{V}$ . The solution is called the power flow solution, and it describes the steady-state characteristics of the network.

$$S_k^p = V_k^p \cdot I_k^{p'} = V_k^p \sum_{i=1}^n \sum_{m=1}^{p_n} (Y_{ki}^{pm} \cdot V_i^m)' \quad (3.2)$$

Here, the nodal power injection at node  $n$ ,  $n \in \mathcal{N}$ , is given by  $p_n + jq_n$ . The nodal voltage  $v_n \in V$  constitutes of voltage magnitude  $|v|$  and voltage angle  $\delta$  and represents value of  $v_n$  in polar form by  $v_n = |v_n| \angle \delta_n$ . Complex power injection at bus  $k$  phase  $p$  ( $S_k$ ) is given by (3.2), in terms of busbar voltages at all  $n$  nodes. The set of active power injections are collected in  $\mathbf{P}$  and the set of reactive power injections are collected in  $\mathbf{Q}$ .  $p_n$  is the maximum number of phases in the system and  $Y$  is the system admittance matrix. Once  $\mathcal{V}$  is calculated, it can be used to directly calculate the branch power flows.

The solution set of this power flow problem is highly non-linear. The physical model-based solution requires accurate knowledge of the system component parameters and topology. For a transmission system, this modeling approach can be effectively used due to the assets being critical, less in number, and static. However, developing and maintaining an accurate model based on the above framework for a distribution system becomes an extremely challenging task.

### 3.3 The Distributed Data-Driven Modelling ( $D^3M$ ) framework

#### 3.3.1 Overview

The proposed power flow solution framework is referred to as  $D^3M$  and is based on Cellular Computational Networks (CCN). CCN is a framework that can be used for distributed asynchronous learning and modeling of networked systems [58]. It was first introduced in 2009 [59], and since then, it has been applied in many power system applications [60] [61] [62] spanning the last decade.

A Cellular Computational Network (CCN) consists of a network of connected cells mimicking the spatial connection of the physical network. The spatial connections represent the point of exchange of boundary variables. Each cell consists of a unique computational unit that uses a learning unit to learn the system asynchronously. This approach opens up the possibility for asynchronous and distributed learning. It is this unique attribute of CCN that sets it apart from Graph Neural networks (GNN) [63], that was proposed in parallel to CCN.

A distribution network shown in the black layer in Fig. 3.2 is a tree graph. It starts with a root node and spreads across nodes with a parent-child relationship. In Fig. 3.2 Cell i is the parent of Cell j and Cell k, and the child of Cell h. Cell k is called a leaf node (it does not have a child node). The  $D^3M$  formulation uses the knowledge of this unique structure of the distribution system. As shown in the light blue layer of Fig. 3.2, the nodes of a distribution system are clustered into cells. The boundary variables that are exchanged between the cells are cell to cell power flow and parent cell boundary voltage.

In combination with local data, the exchanged data are communicated to the CCN layer (blue layer), which processes this data and uses it to learn the system power flow. The data could be collected from the AMI data at each consumer, as well as from network devices such as reclosers, disconnectors, and voltage regulators. As shown in Fig. 3.2 each of these cells is distributed and, therefore, provides the ability to be installed in separate hardware platforms stationed in localized geographic areas. For example, the Cell k could be installed in a small embedded Internet of Things (IoT) platform inside the geographic area of Cell k. This cell consumes only local data and transfers the boundary data to its parent cell. This approach minimizes bandwidth requirements and ensures the privacy of the data in Cell k. If an external data stream becomes unavailable, it can still operate

if a reasonable estimate of the unavailable boundary variable is available.

A generic cell consists of a data manager, a learning unit, and a computational unit, as shown in Fig. 3.2. The learning unit processes the data streamed from the physical layer and provides the correct inputs to the learning and computational unit. The learning unit, when active, uses this data to tune the computational unit to ensure that model accuracy is maintained within the required tolerance limits. The learning unit supports asynchronous operation, meaning all cells don't need to learn simultaneously. Instead, the cells can learn when necessary. The computational unit uses the input data to estimate the boundary variables that are provided as inputs to the child cells in regular operation.

The top layer is the application layer, which consists of local and system-wide applications of the distribution system model. The local application could be installed inside the local geographic region and provide modeling inputs to local DERMS applications such as situational awareness, local energy management, volt-var control, and other local optimization applications. On the other hand, the system-wide  $D^3M$  outputs can be used for system-wide DERMS applications. From a hardware perspective, it is possible for each component in the CCN layer and application layer to be set up in a different IoT platform. If such an approach is used, the  $D^3M$  will contribute towards reinforcing the security, reliability, and resiliency of the distribution system.

### 3.3.2 Methodology

The design of a  $D^3M$  consists of four steps. The first step is identifying the cell model, where an appropriate model for the computational unit is identified. The second step is cell structure identification, where the internal characteristics of the chosen computation unit are identified. The third step is system decomposition and cell learning. Here, the appropriate inputs and outputs, and values of the computational unit parameters for each cell are identified. In the fourth and final step, the individual models are fused to create a single  $D^3M$  model. Based on this network, CCN is executed till convergence, which estimates the local and global power flow solution.

#### 3.3.2.1 Model learning

As detailed in Section 3.2, the power flow problem given by (3.2) is a highly non-linear static Multiple Input Multiple Output (MIMO) mapping. In order to capture the complex non-linearities, a more refined structure than a linear structure is required.

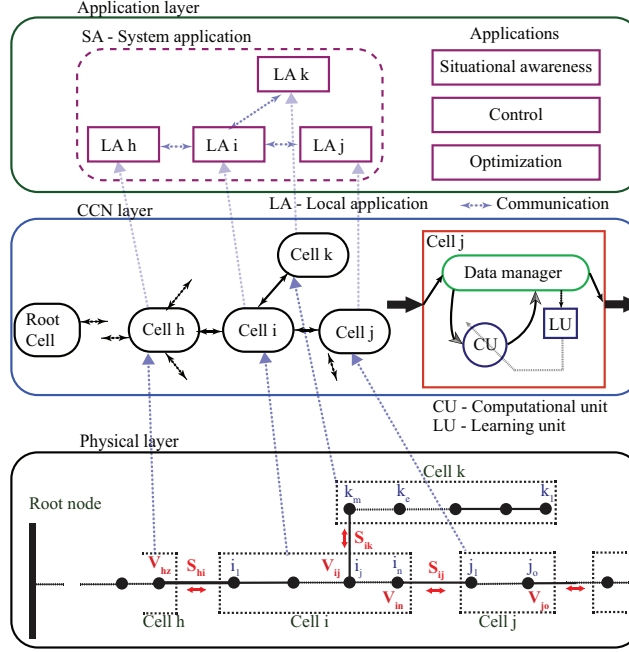


Figure 3.2: The system overview showing the generic ( $D^3M$ ) framework applied to a distribution system.

A Multi-Layer Perceptron (MLP), which is a class of feedforward artificial neural network (ANN), is well suited for approximating a highly non-linear static MIMO mapping [64]. The Stone-Weierstrass theorem [65] has shown that an MLP with one hidden layer can be used as universal approximator [66] under the constraint that a sufficient number of neurons are included in the hidden layer [67] [68].

Based on these insights, a multi-layer perceptron (MLP) with a single hidden layer is chosen as the fundamental building block for the proposed framework. The MLP,  $\mathcal{F}$ , has the structure shown in Fig. 3.9 and is a shallow network since it has only one hidden layer. This network is structurally suited to learning the input-output relationship of the system under study since it can operate as a universal approximator. The activation function on the input and output layers are linear functions and the activation function hidden layer is a hyperbolic tangent (tanh) function. The input layer includes a bias. Since the core structure is fixed, the function  $\mathcal{F}$  depends only on the number of hidden layer neurons ( $\mathcal{F}(n_h) = f$ ).

Power flow can be approximated to a linear mapping with certain assumptions [23]. Similarly, the chosen function,  $\mathcal{F}$ , can be converted to a linear mapping by appropriately selecting its parameter values. This provides further confidence in using  $\mathcal{F}$  as the estimator.

This network with  $c$  inputs and  $d$  outputs is described in (3.3) using matrix-vector notation.

$$Y = \mathbf{V} \cdot \sigma(\mathbf{W} \cdot X + \mathbf{B}) \quad (3.3)$$

Here,  $X \in \mathbb{R}^c$  is the input and  $Y \in \mathbb{R}^d$  is the output. The number of hidden layer neurons, which is a hyper parameter and fixed for  $f$ , is given by  $n_h$ . The weights matrices are  $\mathbf{W} \in \mathbb{R}^{c \times n_h}$ ,  $\mathbf{V} \in \mathbb{R}^{n_h \times d}$ , and  $\mathbf{B} \in \mathbb{R}^{n_h}$ . The activation function for the hidden layer neurons is given by  $\sigma$ .

The next step is model training which is set up as an optimization problem with the objective of minimizing the mean squared error (mse) across a dataset of  $M$  points. Therefore, model training problem is defined as finding the set of weights,  $W$ ,  $V$ , and  $B$ , such that:

$$\min \frac{1}{M} \sum_{i=1}^M (f(X)^2 - Y^2) \quad (3.4)$$

### 3.3.2.2 Cell structure optimization

Cell structure optimization refers to selecting the optimal number of hidden layer neurons to represent the system input-output mapping. With this being a data-driven approach, the selected model structure should have the ability to learn the mapping using available data. It should be complex enough to capture all significant characteristics of the mapping. It should also be able to capture any reasonable changes to the underlying physical system by adapting its parameters since real-world systems are bound to change with time. Finally, it should have a straightforward structure so that the learning is fast, generalizable, and can be guided by a minimal data set. Having more than the required parameters results in slower training and increased computational requirements. It also increases the risk of over-fitting. The ideal structure should be the best possible fit with a minimum number of parameters [69].

Therefore, a cell structure identification process is proposed, where cell structure identification refers to finding the optimal  $n_h$  which represents function  $f$  ( $f = \mathcal{F}(n_h)$ ). The identification of the optimal  $n_h$  is carried out intelligently by using the particle swarm optimization (PSO) algorithm [70]. This helps to develop the optimal cell structure to model this system.

Each particle in the swarm represents a potential solution, which is a  $n_h$  value that is an input to the utility function calculation shown in Fig. 3.3. The objective function value represents the quality of a particle. The evaluation of the objective function plays a significant role in the

structure identification process.

The input from PSO is used to create a model structure,  $Net_0$ , and it is initialized with random weights. Next, the process flow enters a training loop where the current version of the model,  $Net_p$ , is trained using a fixed training dataset,  $D_T$ , for one epoch. The resulting network,  $Net_{p+1}$ , is then checked for the stop criteria. This stop criteria consists of meeting either  $\alpha$ , which is the minimum training error,  $E_T$ , or reaching,  $p_{max}$ , which is the maximum number of iterations. If it is not reached, the network is updated with the new values for weights, and the training continues.

The utility of a particle,  $U_{nh}$  ( $U_{nh} = FIT$ ), defined in (3.5), was used to evaluate the quality of the particles. FIT represents the structural model error [71] across  $M$  data points. Here  $\bar{Y}$  is the mean value of the dataset ( $Y$ ).

$$FIT = U_{nh} = 1 - \sqrt{\frac{\sum_{k=1}^N (Y - Net_{p+1}(X))^2}{\sum_{k=1}^N (Y - \bar{Y})^2}} \quad (3.5)$$

The cell training accelerated by coarsely selecting a suitably high value for  $\alpha$ . This approach is appropriate since the sole objective is cell structure identification. The utility value is representative of the ability of a given solution to approximate the output of the independent testing dataset,  $D_{TE}$ . Once the stop criteria are reached, the resulting neural network is used to calculate the FIT value, the objective value corresponding to particle  $nh$  used in the PSO function.

In order to ensure the quality of the structure identification process, the chosen solution space for optimization should be structurally rich. Additionally, the dataset for training and testing should be reasonably representative of approximated operational space. The selected cell structure is then used in model learning detailed in section 3.3.2.1. This process uses a larger dataset and is trained to a higher level of accuracy. This step enhances the preciseness of the power flow estimations calculated using the CCN network.

### 3.3.2.3 System decomposition

Due to the nature of the problem (radial distribution network), each of the cells can be represented by a Thevenin equivalent, as shown in the generalized cell in Fig. 3.4 (a). A cell  $k$  will have parent and child cells that it is connected to. The only cell that will not have a connection to the parent cell is the root cell that includes the slack bus, and the only cells that will not have a

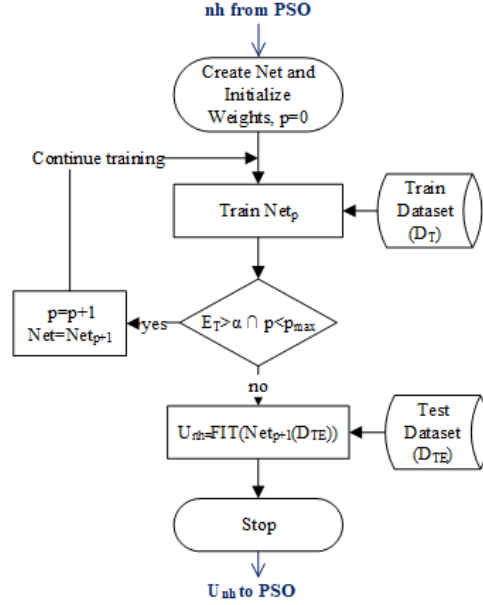


Figure 3.3: The flowchart for calculation of objective function for cell structure identification.

connection to a child cell are the cells that include an end node of a branch of the graph (leaf cell). Natural geometric boundaries can be used as the basis to form the cells.

The CCN formed for the case study, shown in Fig. 3.7, is an example of how the nodes can be grouped into cells. The number of nodes in a cell and connections to neighboring cells define the number of inputs and outputs, which define the total number of weights that need to be trained. The maximum number of nodes (cell size) that can be encapsulated in one cell depends on both the characteristics of the modeled cell and the hardware resource constraints. If the cell characteristics include a more comprehensive range of non-linearities or a more significant number of nodes, it would require more hardware resources for cell training. Therefore, it is recommended to heuristically determine the maximum number of nodes before the CCN is designed. It is possible to choose the cell size lower than the maximum based on other design requirements.

Fig. 3.4 (b) shows the circuit equivalent of a generalized cell. Here,  $V_p$  is the parent voltage, and the dotted boundary creates a Kirchoff's supernode. If the parent voltage,  $V_p$ , internal cell loads,  $S_{cell_k}$ , child branch flow  $LF_P$ , and the internal network parameters are known, the system can be solved using Kirchoff's and Ohm's laws to find the outputs for any given set inputs shown in Fig. 3.4 (c). Therefore, cell identification only depends on the network parameters. The internal structure of the computational unit can take any form, including the neural network approach proposed in



this study or a more conventional physical model-based computational structure.

#### **3.3.2.4 System learning**

The cell learning process consists of 3 stages; the data acquisition stage, the data processing stage, and the computational unit tuning stage.

The data acquisition stage aims to acquire the boundary variable data from the neighbor cells based on the CCN formulation. This stage could be implemented in a centralized approach or a decentralized approach. In a centralized approach, first, the data from all the cells will be transferred to a central database, and then the central database will segment and share the relevant formatted datasets with each cell. This requires that all cells be connected to the central data acquisition system. The cells will poll for required boundary data from neighboring cells and acquire only the required data segments in a decentralized approach. For this to be possible, the cells need a physical communication link with their neighbors. A centralized approach requires more communication infrastructure, bandwidth, and storage than the decentralized approach, where only the minimum set of information (required boundary data) is shared locally between neighboring cells. Irrespective of this choice, the communication between cells within the learning process is limited to the data acquisition phase.

The data acquisition system of the cell then processes the remote and local data into a form suitable for utilization by the computational unit. If a decentralized approach was used in the data acquisition stage, the boundary data would be appended with the local data to form the final dataset.

Finally, the learning unit operates on the cell computational unit by adjusting the neural network weights till the required estimation accuracy criteria are satisfied. Each cell computational unit takes the form shown in Fig. 3.4 (c) and (d). These computational units learn using the method described in Section 3.3.2.1. Even though each cell requires boundary data from neighboring cells to create the training data set, the cells are trained offline, and the cells learn asynchronously independently of the other cells.

#### **3.3.2.5 Model Fusion**

Model fusion is when the independently trained cells connect to form the interdependent CCN and iteratively exchange information to provide the power flow estimation for the modeled

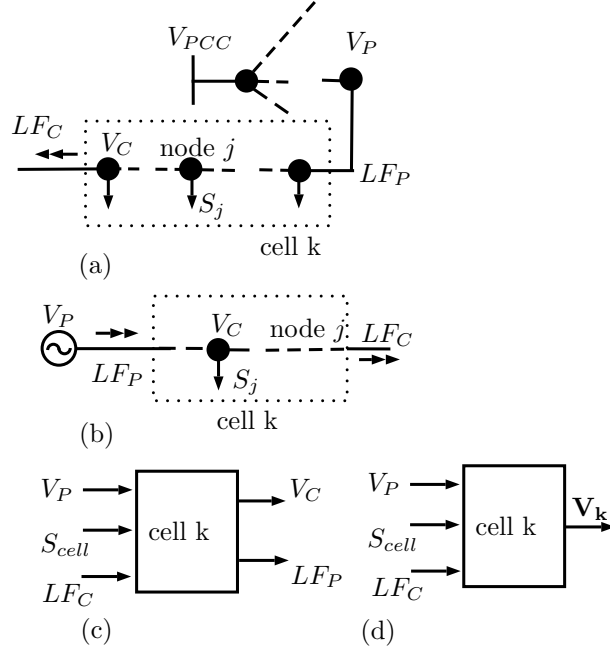


Figure 3.4: (a) The single generic cell single line diagram (b) electrical equivalent circuit (c) cell input-output relationship if cell k has a child cell (d) corresponding cell input-output relationship if cell k is a leaf node.

system.

The framework uses two types of variables. The first type is defined as CCN variables ( $X_{CCN}$ ) and constitutes of variables that integrate the cells together and transfer across cell boundaries. This set of variables is shown in Fig. 3.4 (c) and (d). The second type of variables are the internal state variables of interest ( $X_{cell}$ ), and they are the internal cell variables of interest (internal cell node voltages). They are not a direct part of the CCN framework and are independently estimated on a secondary layer, using the solution results obtained from the CCN-based estimation.

The flow chart for model fusion is shown in Fig. 3.5. The process starts with the initialization of values for inputs in the set of CCN variables  $X_{CCN}$ . Out of the two types of boundary variables, the voltages are typically initialized with available field data, while the load flows are initialized by summation of the node load values. If any data is missing, then the boundary variables are initialized by a flat start. A flat start refers to the initialization condition that uses rated voltage and load values. Then the cells are operated with available inputs and the  $X_{CCN}$  values are updated  $X_{CCN,k}(i+1) = Cell_k(X_{CCN,k}(i))$ . At each iteration, the updated boundary variables replace the last estimate, and the cell inputs and outputs are based on the CCN data flow diagram. An example

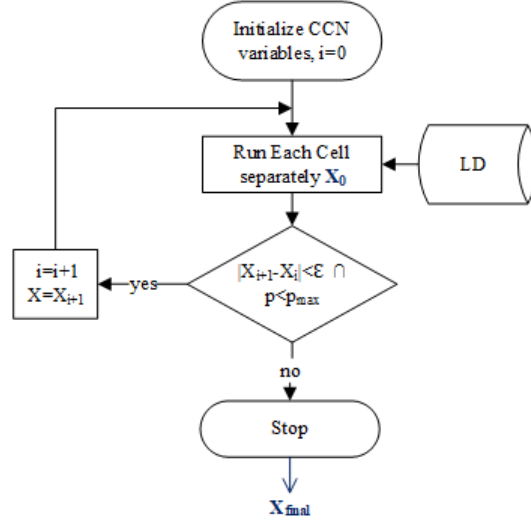


Figure 3.5: Model fusion flowchart showing the iterative algorithm used to solve for the  $D^3M$  solution. Here  $X$  refers to  $X_{CCN}$ . The load flow estimation is for the nodal load dataset given by  $LD$ .

data flow diagram that elucidates how the communication occurs, the data flow sequence, as well as the dependency between interconnected cells is shown in Fig. 3.8. The process is iterated till either all the  $X_{CCN}$  values converge to  $\epsilon$  ( $X_{CCN}(i+1) - X_{CCN}(i) \leq \epsilon$ ), or the max number of iterations are reached. Once the final CCN solution is available, these values are used as input, and a separately trained neural network is used to calculate the  $X_{cell}$  values.

In implementing the model fusion for this study, communication was functionally separated from the cell computation but was implemented on the same hardware platform. The information exchange was executed as an internal data exchange inside the model fusion program. If the cells are located in physically separate hardware platforms, they will depend on an external communication medium to exchange values. An appropriate state-of-the-art communication framework, such as TCP/IP on Ethernet, can be used to satisfy this requirement since the boundary data flow is a small subset of the cell data stream.

### 3.4 Case Study

This section describes the application of the proposed framework to model a distribution system. The IEEE 34 bus test system [72], which was based on a long radial rural feeder in Arizona, is used as the sample system. It has a total of 95 single-phase nodes and 134 single-phase loads.

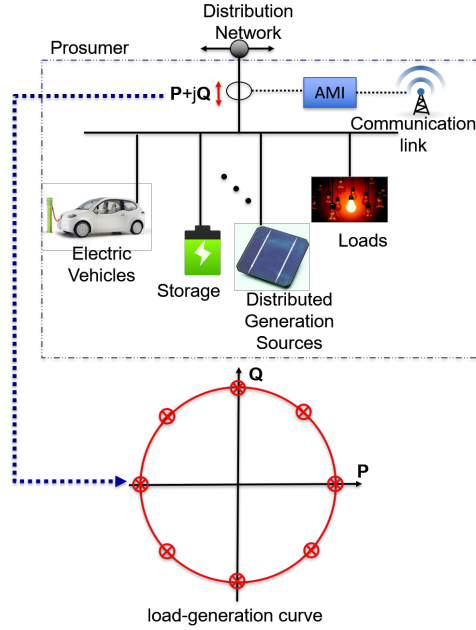


Figure 3.6: A generic model of a prosumer node used in the study. The active and reactive power for each prosumer lies within the boundary shown in red. The cross represents key extreme data points.

The network and loads are unbalanced and include single-phase, three-phase, and two-phase loads and transmission lines. Loads of the IEEE test system are modified by converting the loads to prosumers as shown in Fig. 3.6. Each node of the distribution system can include electric vehicles, battery storage, and distributed energy sources, apart from conventional loads. The prosumer power characteristic operates on four quadrants, as shown in Fig. 3.6. The prosumer connection is monitored using AMI, and the information is transmitted to the cell IoT platform via a communication network. Each prosumer has an installed capacity equal to the apparent power rating of the original load. Note that this system can represent 100% prosumer capability in both the number of installations as well as installed capacity. This means that the applied case study can easily represent a distribution grid with high penetrations of renewables.

### 3.4.1 System modelling

The system model shown in Fig. 3.7 describes the development of the CCN model of the system. Each branch is required to be on at least one cell; multiple cells in one branch are also possible. The size of each cell is based on the computation resources available. For a reasonable

IoT-based cell solution, the thumb rule limits the maximum number of nodes to 5. The data and computation resource requirements decrease when smaller cell size is used. The connection between each cell is based on the parent and child voltage and boundary load flow. Unlike the other cells, the root cell MC1 does not have a parent. The SC1, SC2, SC3, and SC4 cells are leaf nodes and do not have a child cell.

The cell SC1 provides the load flow to its parent cell MC1 and receives the information of the PCC voltage from it. The same applies to cell SC2. The cell MC2 provides its load flow to parent cell MC1 and gets the PCC voltage as an input from the parent cell. The same applies to cell MC3. The relationships that  $D^3M$  is based on are shown in the system block diagram in Fig. 3.8.

### 3.4.2 Dataset generation

A random prosumer load dataset of 86400 data points, representing a day of data gathered at a resolution of 1 second, is used to generate the dataset. The modified IEEE 34 test case is modeled in OpenDSS [18], and the system is simulated with the generated inputs to create the synthetic measurement dataset.

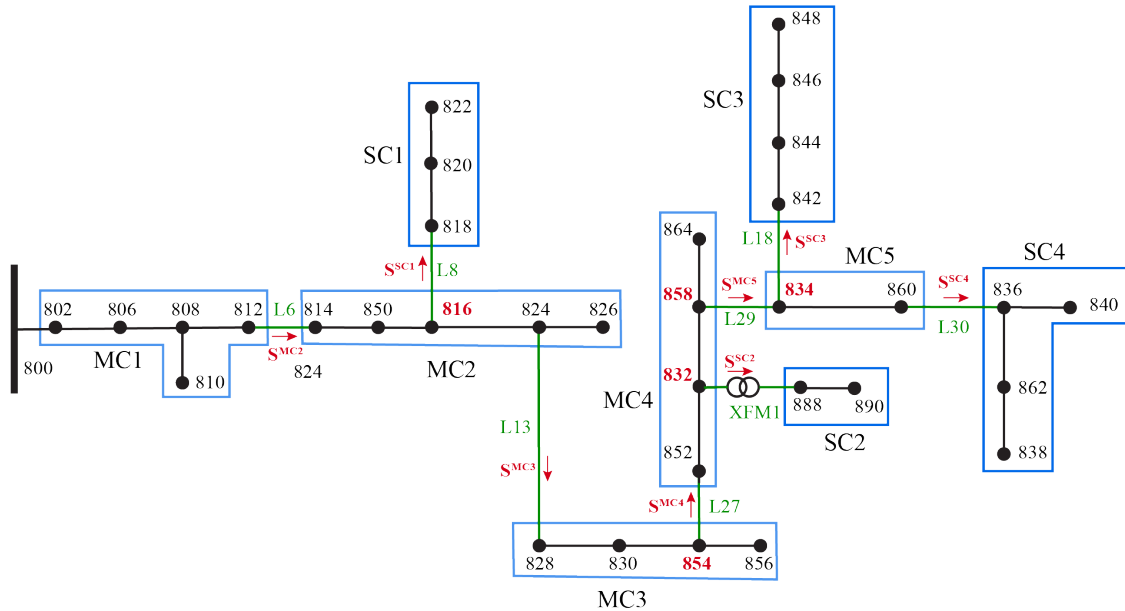


Figure 3.7: The  $D^3M$  based cell formulation for the modified IEEE 34 network. The chosen cells and boundary variables are shown on the single line diagram.

### 3.4.3 Model development and learning

The  $D^3M$  model uses a limited amount of topological information about the network. It requires knowledge about some parts of the network graph to develop the model and map the appropriate measurements to network information. However, the amount of information required is minimal compared to any other alternate option. The CCN formulation could be based on natural geographic boundaries such as a street or neighborhood. If needed, it is also possible to fragment a more extensive geographic unit into smaller sub-cells. For example, the MC5 and SC4 cells are formed as two cells, even though they can be formulated as one cell.

Choosing the base number of nodes in the hidden layer is performed based on PSO, and model performance is compared. The identified optimal values are included in Table 3.1.

The internal variables are not included in the CCN model. Instead, separate internal models that use the solved values of boundary variables as the input are used to solve for internal variables of interest.

The models are trained asynchronously based on the earlier identified optimal MLP structure. This study uses the Levenberg-Marquadt method for training. However, any suitable learning algorithm can be used.

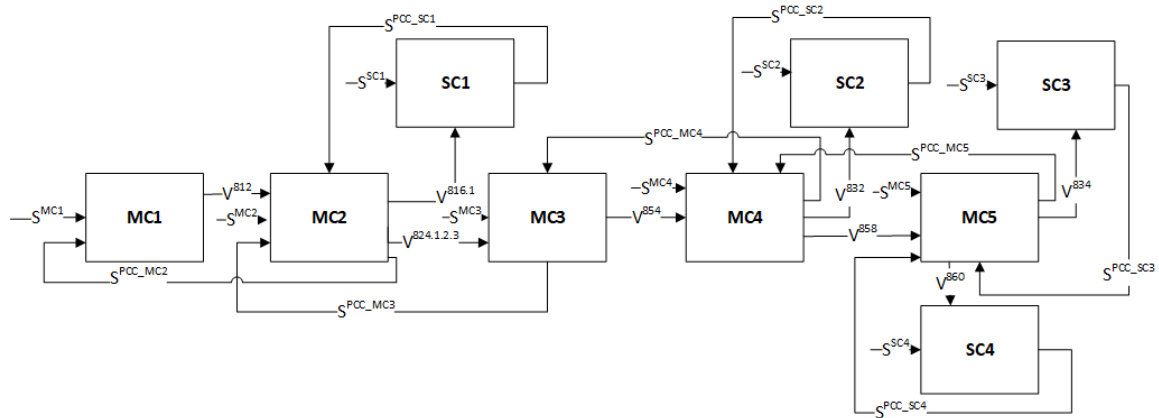


Figure 3.8: The information flow diagram for structure fusion. The cells are initialized using field data or rated load and voltage values. The information is exchanged and cell outputs are computed till the output converged to the required accuracy.

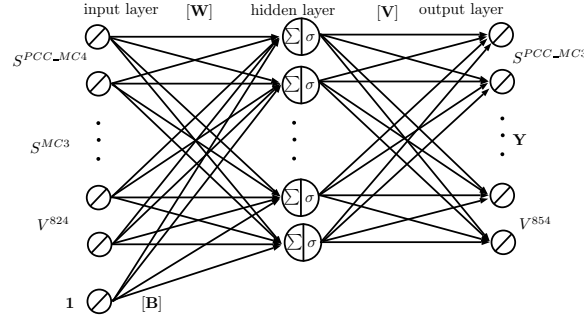


Figure 3.9: The computational cell structure is illustrated using cell MC3. It consist of a single hidden layer Multi Layer Perceptron network. The optimal number of hidden layer neurons,  $n_h$ , as well as the weight matrices, are calculated by using historical data.

### 3.4.4 CCN in Operation

The modified IEEE 34 bus system is operated based on Fig. 3.8. The  $D^3M$  variables are required to be iterated following the flow chart described in the generic framework till stop criteria are reached. The boundary variable error convergences after a maximum of 15 iterations for the system understudy.

## 3.5 Results and Discussion

The different parameters that are applied in the design of the  $D^3M$  are given in Table 3.1. Here,  $nh_H$  is the geometric mean of the neural network's number of inputs and outputs. It is a typical heuristic used to determine the number of hidden layer neurons. The  $nh_o$  value is the intelligently determining optimal number of hidden layer neurons. The values are different, highlighting the importance of determining the hidden layer in an intelligent manner in this application. The Id. time parameter is the time required to identify  $nH_o$  and has a maximum of 2 hours. The Tr. time refers to the training time of the individual cells and is also less than 2 hours for the whole group of cells. The maximum number of weights that have to be learned is 4803. The training of  $D^3M$  models was executed on an Intel Core i7 8700 computer with 16GB of RAM.

The neural network given as \*Onenet results when the system model is represented by one large network. The number of weights in this network is almost 16 times that of the largest network in the  $D^3M$  solution. The network training failed with the heuristic applied earlier. Therefore, another heuristic, which selects the number of hidden layer neurons equal to twice the number of inputs,

Cell	$nh_H$	$nh_O$	Id. Time (s)	Tr. time(s)	$n_i$	$n_o$	$W_n$
SC1	5	9	1238	1035	10	2	117
SC2	7	35	4265	2903	8	6	525
SC3	13	12	6582	2903	26	6	396
SC4	13	50	1962	2903	26	6	1650
MC1 Vm	9	29	1724	1928	22	3	754
MC1 Vd	9	12	3607	3291	22	3	312
MC2 S	15	36	3850	3905	36	6	1548
MC2 Vm	11	34	1243	4879	36	3	1360
MC2 Vd	11	15	3384	3381	36	3	600
MC3 S	15	50	6536	2906	34	6	2050
MC3 Vm	15	38	2369	1356	34	6	1558
MC3 Vd	15	52	2571	3020	34	6	2132
MC4 S	15	55	2272	2906	58	6	3575
MC4 Vm	14	34	7263	3020	58	3	2108
MC4 Vd	14	36	2677	5020	58	3	2232
MC5 S	15	31	2272	2906	58	6	2015
MC5 Vm	14	59	1103	3356	58	3	3658
MC5 Vd	14	43	7263	3020	58	3	2666
*Onenet	272	NA	NA	5020	136	80	59024

Table 3.1: Parameters of the neural networks used for  $D^3M$  of the power flow.

had to be used. The computational system could not run the training due to memory constraints. As a result, a Dell C4130 Intel Xeon E5-2680v3 processor from the palmetto cluster with 125GB had to be used to train this network. In order to provide a reasonable comparison, the run time for learning was limited to the maximum run time of the  $D^3M$  (5020s). The centralized training fails to even run on a standard computational platform. It is clear that the system decomposition and asynchronous learning both allow for optimal learning, using fewer data points and practical computational resources.

The method is first validated using a random data set of 1000 points. The random values are drawn from inside the prosumer curve shown in Fig. 3.6. The results for random validation in the form of a box plot are shown in Fig. 3.10. The Absolute Percentage Error (APE) is used as the index to compare the accuracy of the output.

The results show that the system performs well for a random dataset. For every solution, the maximum APE is less than 0.036%, the mean APE is less than 0.05%, and the third quartile is less than 0.006%. The APE boxplot of the inter-cell power flow variables is tested on the same dataset and shown in Fig. 3.14. The maximum APE has increased to 0.37%, but the mean and the third quartile are 0.05% or lower.



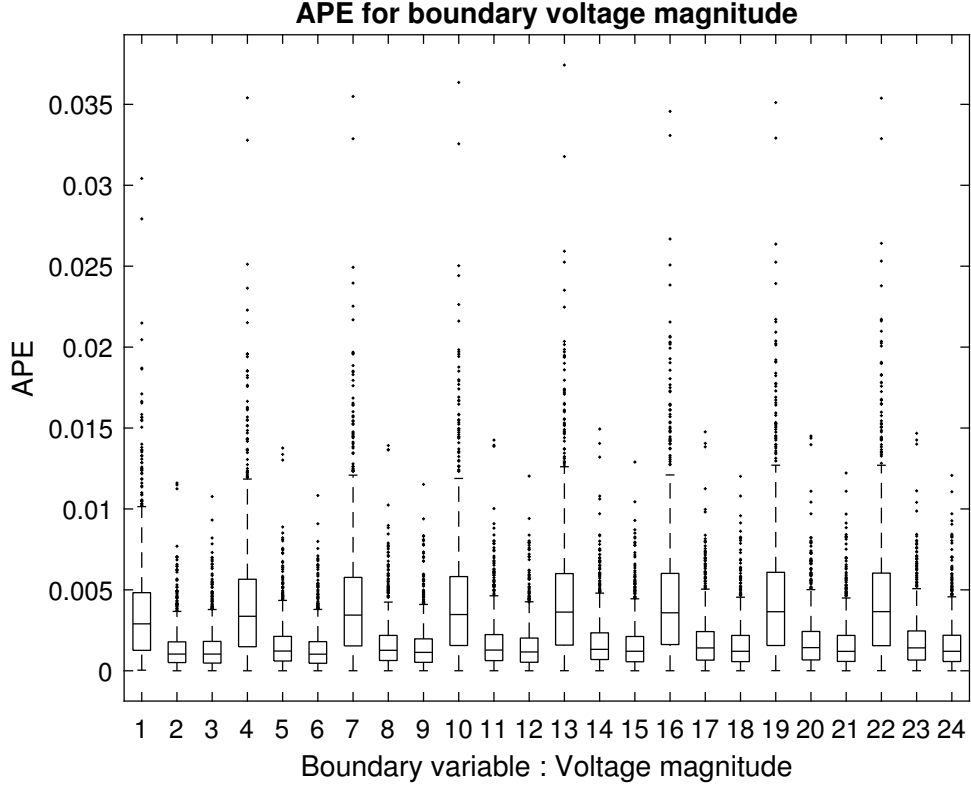


Figure 3.10: Statistical distribution of the  $D^3M$  solution for boundary voltage magnitude absolute percentage error for a dataset of 1000 random validation data points. The x-axis represents the voltage magnitude variables that are extracted from the solution.

The internal cell voltages of SC1 and SC2 are shown in Fig. 3.11. The APE of the estimates for SC1 and SC2 has an upper bound of 0.6%, which is rather high but still reasonable for typical applications. The APE of the estimates for MC1 and MC2 are shown in Fig. 3.12 and 3.13. When compared together, they have an upper bound comparable to SC1 and SC2. MC1 has an upper bound of 0.14%, and MC2 has an upper bound of 0.6%. These are comparably high but still reasonable for typical applications. However, the mean and the third quartile for these two output distributions are lower than SC1 and SC2. These statistical indices show that the proposed approach provides reasonable accuracy.

The same dataset is used to assess the performance of the \*Onenet. The results show that the Onenet has APE outliers in the range of 12%, 3rd quartile in the range of 6%, and an upper bound of the mean of 6%. These values are comparatively higher than the results from  $D^3$ . The advantage of this network is that it is not required to iterate to find the solution. Therefore, the run time per solution will be lower.

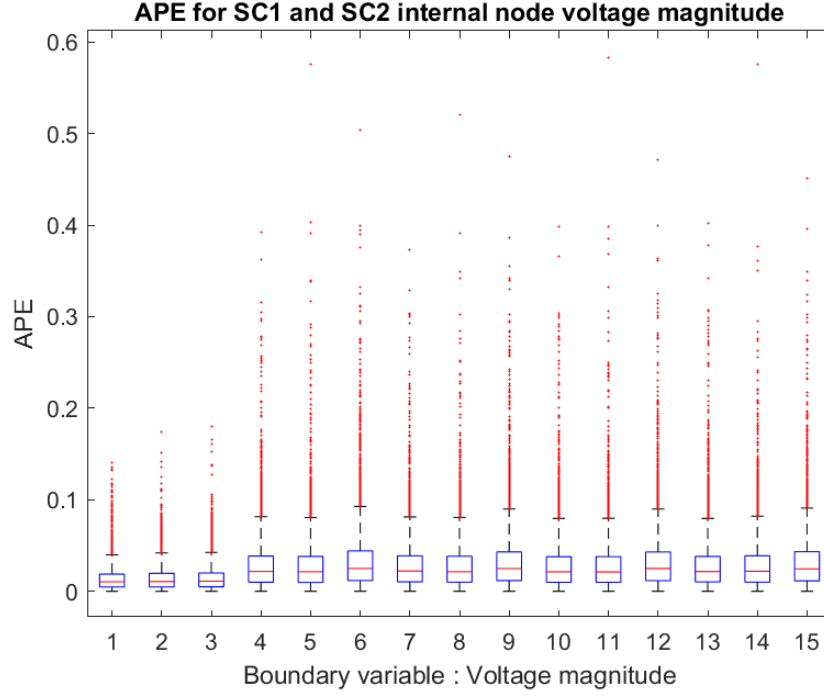


Figure 3.11: Statistical distribution of the SC1 and SC2 internal node voltage magnitude absolute percentage error.

Even if the  $D^3M$  model is accurate for random data, it is also essential to be accurate at the extreme data points. Therefore, the  $D^3M$  performance is next tested at extreme boundaries. The extreme data validation is carried out across eight data points that lie on the circumference and are marked with a circle-cross in Fig. 3.10. The output of the simulation of this extreme data set is shown in Fig. 3.16. The maximum APE of the output of this data set has increased to 0.06%, which is a 170% increase. However, the box plot shows that the mean is still less than 0.01% for all variables, while the third quartile is less than 0.02%. These values are acceptable for distribution system power flow applications, especially considering this dataset only includes extreme data. This points out the ability of the proposed computational cell to generalize over the entire state space.

### 3.6 Summary

The analysis presented for the study system shows the ability of the proposed  $D^3M$  framework to solve the power flow in a distributed and data-driven manner with a high level of accuracy. The proposed method is straightforward and powerful. The modeling is highly automated and re-

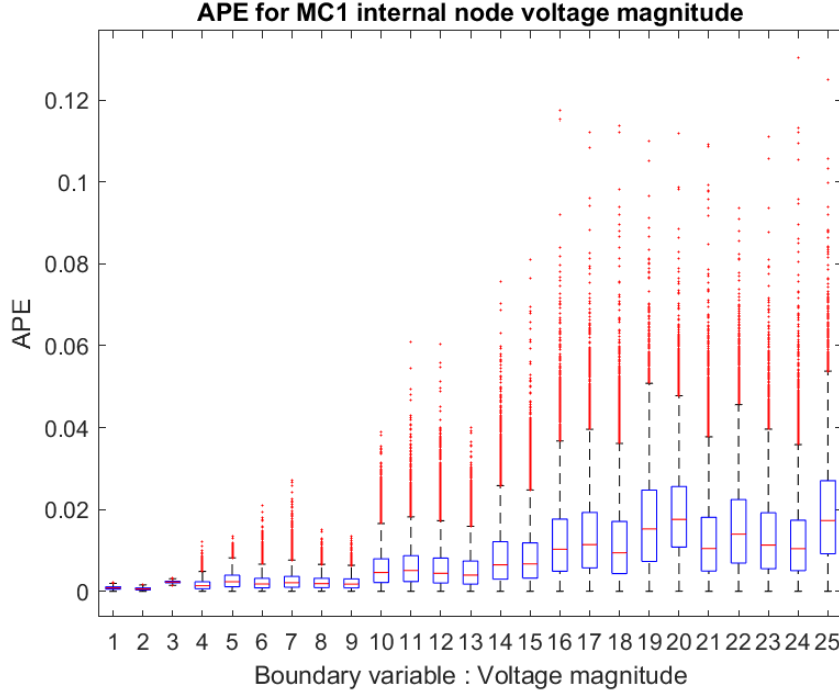


Figure 3.12: Statistical distribution of the MC1 cell internal node voltage magnitude absolute percentage error.

quires minimal manual intervention. The framework additionally ensures the privacy and security of sensitive utility information.

An AI-based computational unit is used in the presented work, and it provides an excellent ability to learn from the system data. It is clear from the results that the  $D^3M$  can operate at a high level of accuracy. Physical models can be less accurate due to approximations used in modeling and system changes between model creation and operation.

Gradual degradation due to normal decay, as well as problems such as cable splice failure, breaker contact degradation, and transformer insulation failure [73] cause significant changes in the distribution system power flow. The proposed approach can adapt the model to system changes. Assuming that the degradation is occurring over time,  $D^3M$  can learn and estimate the physical system's performance. With a physical model, it is impossible to adapt on-the-go, since the changes need to be measured, parameters have to be recalculated, and the simulation must be updated for the changes to come into effect.

The application of a CCN-based power flow model on the sample test case shows the suitability-

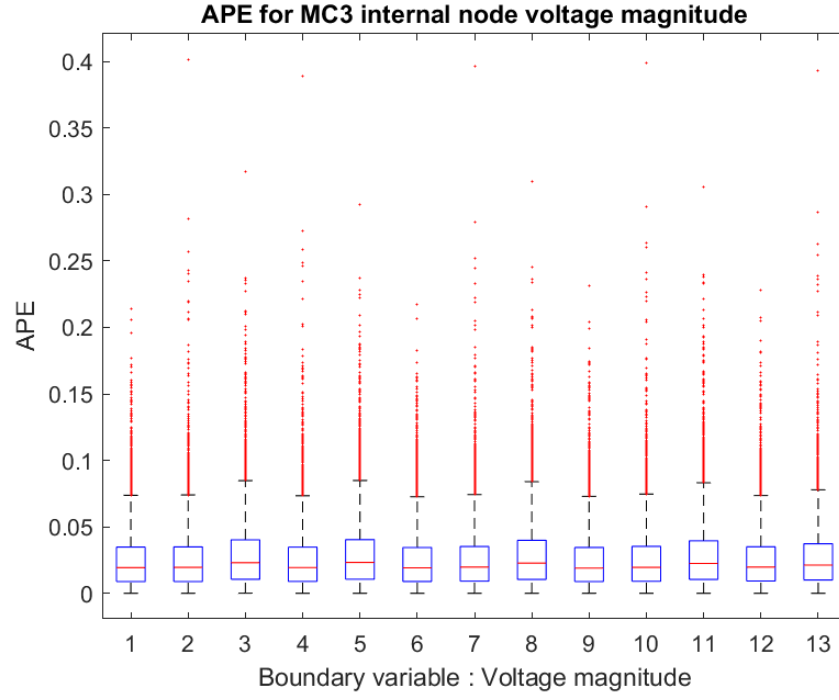


Figure 3.13: Statistical distribution of the MC3 cell internal node voltage magnitude absolute percentage error.

ity of the new approach in modeling distribution systems and its enhanced performance compared to the centralized modeling approach.

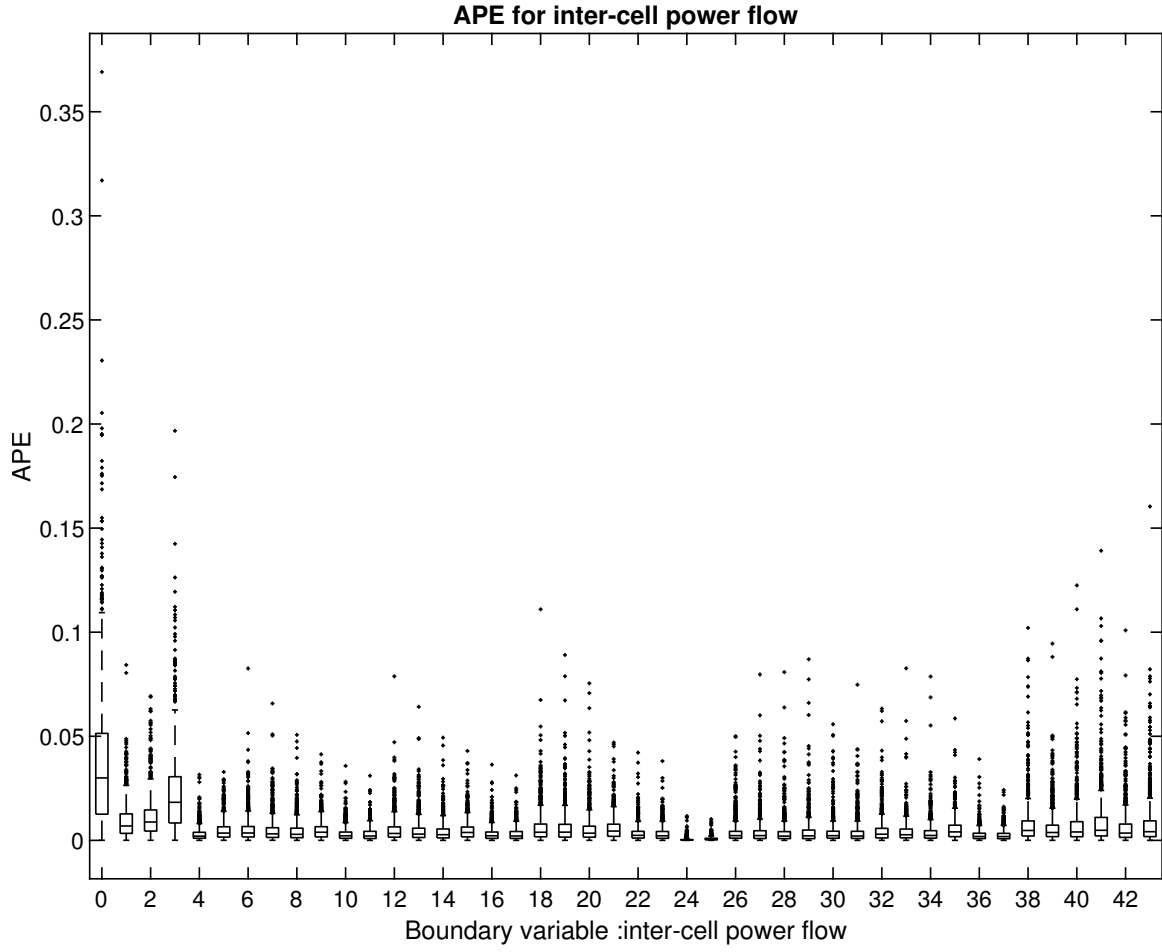


Figure 3.14: Statistical distribution of the  $D^3M$  solution for inter-cell power flow absolute percentage error for a dataset of 1000 random validation data points. The x-axis represents different power flow variables.

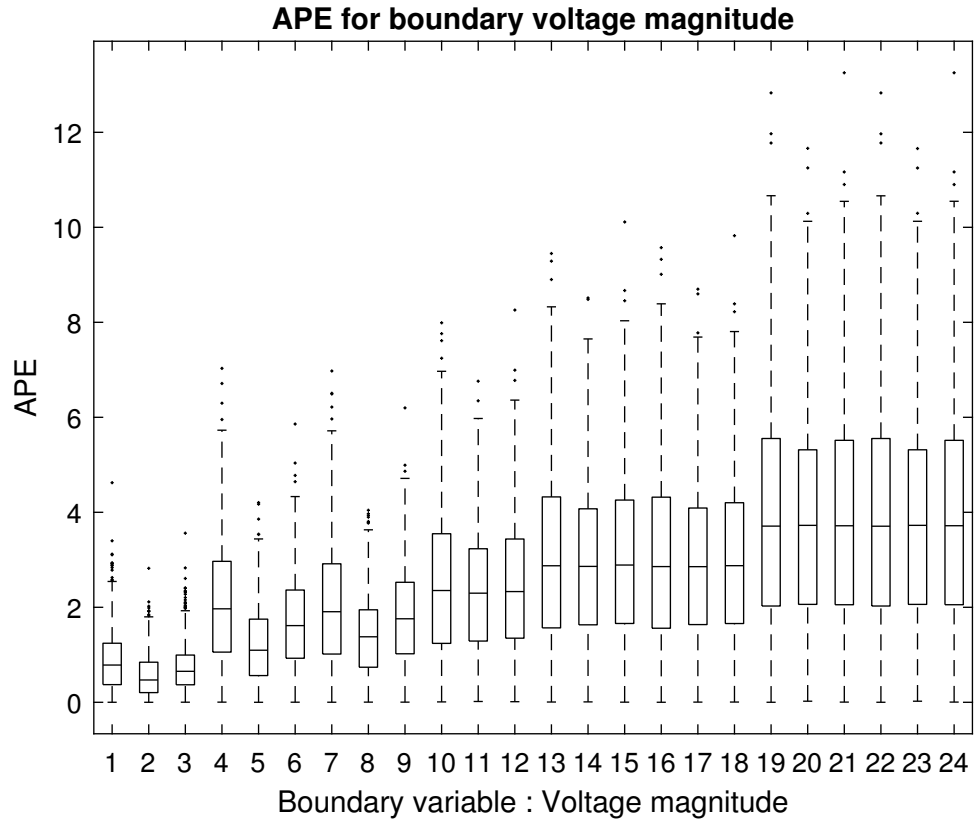


Figure 3.15: Statistical distribution of the  $D^3M$  solution for boundary voltage magnitude absolute percentage error for the power flow estimation using the Onenet network. The same training dataset and validation dataset are applied for network training and validation across all networks.

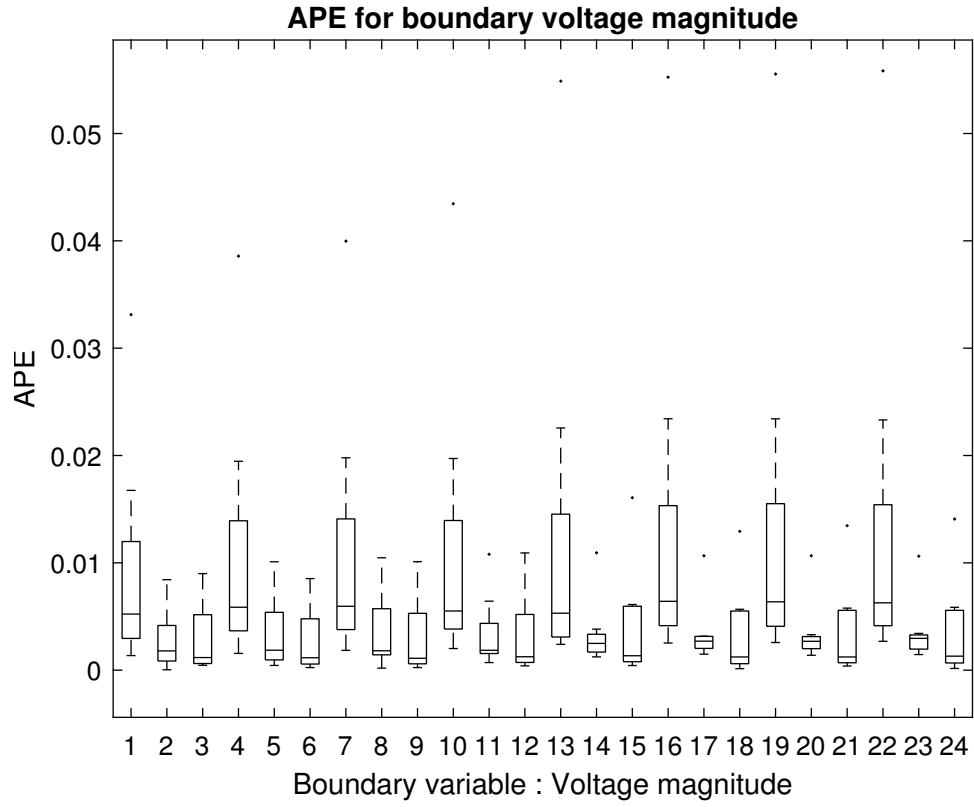


Figure 3.16: Statistical distribution of the  $D^3M$  solution for boundary voltage magnitude absolute percentage error for a dataset of 8 extreme validation data points chosen on the circumference of the prosumer power curve. The x-axis represents the voltage magnitude variables that are extracted from the solution.

## Chapter 4

# Distributed Volt-Var Curve Optimization Using a Cellular Computational Network Representation

This chapter introduces the state of the art of power distribution system control, its challenges, and how they can be addressed using the proposed electric distribution system operation and control framework. The proposed framework for optimizing the power distribution systems is based on cellular computational networks and builds upon the data-driven modeling framework presented in the last chapter.

The distribution system did not historically have ways to measure voltage or communicate with other entities. Adding this functionality across thousand of small nodes was not cost-effective. However, the proliferation of different technologies such as DER and IoT has created a cyber-layer covering the power distribution system. The wide availability of sensing and communication with minimal marginal cost has enhanced the distribution system operation and control system design options. Voltage control in modern electric power distribution systems has become challenging due to the increasing penetration of distributed energy resources (DER). The current state-of-the-art



voltage control is based on static/pre-determined DER volt-var curves. Static volt-var curves do not provide sufficient flexibility to address the temporal and spatial aspects of the voltage control problem in a power system with a large number of DER. This paper presents a simple, scalable, and robust distributed optimization framework (DOF) for optimizing voltage control. The proposed framework allows for data-driven distributed voltage optimization in a power distribution system. This method enhances voltage control by optimizing volt-var curve parameters of inverters in a distributed manner based on the power distribution system’s cellular computational network (CCN) representation. The cellular optimization approach enables system-wide optimization. The cells to be optimized may be prioritized, and two methods, namely, graph and impact-based methods, are studied. The impact-based method requires extra initial computational efforts but thereafter provides better computational throughput than the graph-based method. The DOF is illustrated on a modified standard distribution test case with several DERs. The results from the test case demonstrate that the DOF-based volt-var optimization results in consistently better performance than the state-of-the-art volt-var control.

## 4.1 Introduction

The electric power system is undergoing a rapid transition from a fossil fuel-based, central system to a renewable distributed energy resource-based distributed system due to the need to enhance the security of supply, reduce cost, enhance sustainability, and battle climate change [74]. The bulk of this transformation is a result of the rise of the prosumer, a consumer that also produces and stores energy [75]. All consumers who have installed distributed energy resources such as solar PV and storage are prosumers, and they are at the center of the energy transition. A bulk of the prosumers are connected to the power distribution system. Unfortunately, the distribution grid infrastructure has been one of the most neglected infrastructures worldwide. It is the least prepared to serve the rapidly transforming needs of the energy sector. For example, in some places in the United States, the average age for the distribution of assets are between fifty to seventy years, thus past the expected lifetime of thirty years [76].

Climate change combined with digitization has accelerated the decarbonization of the energy sector, causing rapid electrification of multiple facets of human society [77]. The quality of life and the functioning of society at large is increasing in dependency on electricity. The distribution grid

is at the epicenter of transformation. Therefore, while there has been weak historic investment in the distribution grid, the highest level of quality of service, resilience, and reliability is required at the distribution grid. For example, it is critical to maintaining a high-quality voltage profile, at all times, because of the sensitivity of the customer equipment, regulatory requirements, as well as financial requirements to minimize losses in the distribution grid [78]. The need for performance is the highest at the weakest link of the power system.

The multi-directional power flow needs to be supported by an aging distribution infrastructure initially designed for unidirectional power flow, which further escalates the complexity of the problem. The stress on the distribution system and the need to improve the quality of service delivered at the distribution system have increased simultaneously, highlighting the need to address both of these challenges. Developing technologies will hasten the energy sector's transition towards a sustainable, cost-efficient, and secure future.

It is increasingly clear that ensuring voltage quality is one of the distribution system's most significant challenges. The main drivers of this voltage control challenge are the multi-directional power flow from intermittent generation and new loads with considerable magnitude and low diversity, such as Electric Vehicles. Voltage control is implemented using regulators and capacitor banks in a traditional distribution system. Since the conventional distribution system consists of unidirectional, top-down power flow with slow changes in demand, this method can be effectively used to regulate voltage in a traditional distribution grid. However, with the rise of the prosumagers, where solar PV dominates production, the variability of solar irradiation causes significant fluctuations in the power flow resulting in a degraded voltage profile [79]. Unfortunately, traditional voltage regulators are not fast enough to counter this problem since they depend on slow devices such as mechanical tap changers [80]. Further, the cost associated with an increasing number of tap changing will be significant to the operator due to the degradation of the equipment lifetime.

#### 4.1.1 Optimization

Optimization can be executed in numerous ways. The classification of optimization based on framework, implementation, objective function, and the solution is shown in Fig. 4.1 and described below.

- *Centralized versus Decentralized:* The framework for optimization is centralized if data from

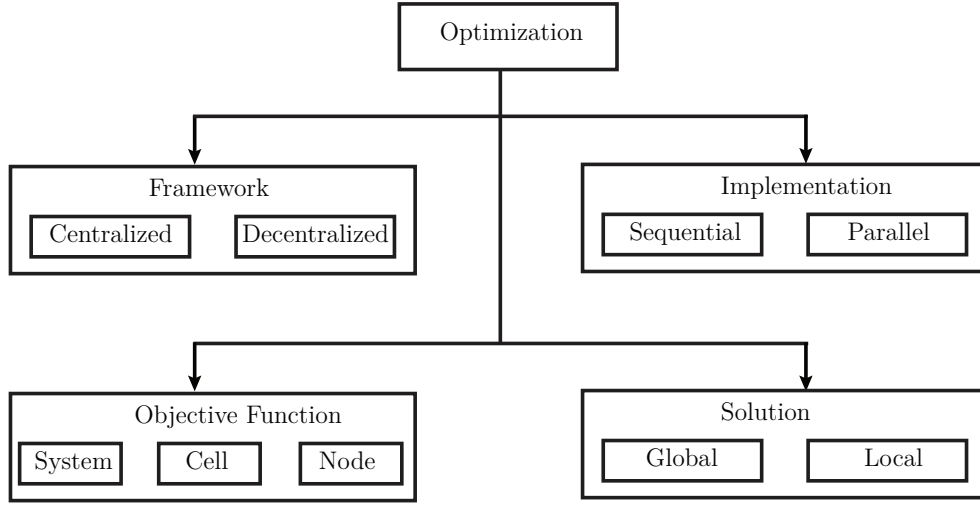


Figure 4.1: The classification of optimization based on framework, implementation, objective function, and solution.

all the nodes/buses is communicated to a central point; on the other hand, it is decentralized if the data is consumed locally. Therefore, both communication and sensing are limited to the local node.

- *System versus Cell versus Node:* An objective function can be formulated in many ways. This formulation (for a  $N$  node system) can be for system-wide optimization (all nodes), cell-wide (a set of nodes/sub-system), or node optimization (local optimization).
  - System-wide optimization (one objective function for the system) is executed in a centralized or decentralized manner. The ideal approach is to run it in a centralized way. A decentralized approach does not guarantee a system-wide global optimal solution.
  - Node optimization ( $N$  Objective functions), where each node has a unique objective function, can be executed in a centralized or decentralized manner. The best approach is a decentralized execution. A centralized approach does not guarantee a node-level optimal solution.
  - Cell-wide optimization (between 1 and  $N$  objective functions) refers to having one objective function per cell (sub-system). This is a decentralized approach, but decentralization is lower than node optimization.
- *Global versus Local:* During any optimization, based on the problem being optimized, there

could be multiple local minima and a global minimum. Local optimization seeks to find the local minimum appropriate in a local search space. Global optimization seeks to find the lowest minimum of all local minima in a global search space.

- *Sequential versus Parallel:* The optimization implementation is sequential if executed in a sequential computing environment using a sequential algorithm. If performed in a distributed environment using a parallel computing platform, it is parallel.

Different combinations of these approaches can be selected to satisfy the requirements of an application. A distributed optimization framework is advantageous when a system-wide optimization is sought while not having access to all data at a central location. The DOF decomposes the objective function at the system level into many objective functions that can be solved independently, seeking a globally optimal solution.

#### 4.1.2 State-of-the-art

The distribution system operators solved the voltage quality challenges of the distribution system by installing tap changers and capacitor banks. Whereas these conventional solutions are still in use, the modern state-of-the-art solutions are centered around power electronics devices such as d-statcom [81], edge of network grid optimizer [82] and dynamic voltage restorer [83]. However, these components are costly to install and maintain while needing support infrastructure that might not even be available to the DSO.

Instead of installing additional devices, this study navigates the possibility of optimally leveraging the flexibility of using the prosumer-owned electronic power inverter to control the distribution system voltage. This study assumes that the inverter's idling reactive power capacity is available to support voltage control [84]. The prosumers can be compensated for this ancillary service.

All optimization variants rooted in a centralized framework operate where a central unit will process the measurements, perform necessary computation, and dispatch the appropriate operation setpoints to a local controller. The local controller controls the prosumer. Each local node needs to be directly connected to the central controller. The advantage of this approach is that the central controller will have access to local measurements. Therefore, assuming perfect optimization, it has the potential to operate at global optimality. The disadvantage of this approach is that the

central controller will have to process multiple input and output data streams and solve for a large number of variables. This approach can be applied for both real-time control [85] and optimal power flow based operation [86] [87] [88]. Efforts to optimize the local reactive power control over a lengthy time horizon have also been presented in the past [89]. However, the large magnitude of optimization parameters and difficulty producing granular forecasts at the required accuracy make the central approach infeasible for most use-cases. Additionally, a failure in the communication link or the central controller would result in catastrophic failure, pointing to reliability challenges associated with these single points of failure. Therefore, this approach is not robust nor scalable and is constrained by the requirement to communicate large amounts of data and optimize at high speeds. Consequently, this approach comes with untenable implementation challenges.

All optimization variants rooted in a decentralized framework rely only on local measurements. Each node in the system will have its unique, and decisions will be taken only based on local measurements. Therefore, this does not require computation and optimization on large data sets or communication infrastructure. Consequently, local optimization is robust. However, since the local optimizer is only aware of its local environment, the results will likely be sub-optimal. This method has been extensively analyzed [79] [90] and forms the basis for the widely implemented state of the art voltage control defined in the IEEE 1547-2018 standard [91]. However, since it is a local framework, the synergies of the connected resources will not be leveraged to gain local and global optimization, and the contributions will be highly skewed based on the spatial distribution of the prosumers in the distribution feeder.

For example, in the distributed optimization framework presented in this paper, there is limited communication between nodes, and optimization is executed across a sub-set of the nodes. Similar approaches have been investigated in recent studies such as [92] [93] [94]. However, these approaches do not have a cell-wide objective function. Therefore, the number of iterations required for each solution, the assumptions for system modeling, and the constraints on the maximum possible prosumers limit their applicability.

### 4.1.3 Contributions

This paper proposes a method to optimize multiple volt-var curves of DERs in a power distribution system with distributed cell-based optimization to address the challenges mentioned above. The main contributions of this paper are:

- A scalable distributed optimization framework has been developed for concurrent multiple volt-var curve optimization. This method is based on creating a cellular computation network representation of the electric power distribution system with DERs, which allows for distributed data-driven modeling and optimization of the distribution system.
- A ranking method for cell prioritization for optimization has been developed. Cell prioritization for optimization based on a formulated impact ranking criteria improves the computational throughput for determining optimal volt-var curve parameters. Two methods, namely, graph and impact-based methods, are studied.
- The application of DOF on a modified IEEE 34 bus test system with 100% DER penetration has been illustrated. The operational results obtained with volt-var curves optimized using the DOF consistently outperform those obtained state-of-the-art volt-var curves.

The study is organized as follows: section 4.2 formulates the problem statement, section 4.3 describes the framework for distributed VVC optimization, and section 4.4 describes the simulations used to demonstrate the framework. Section 4.5 shares the results, as well as the insights attained through the results. Section 4.6 concludes the study.

## 4.2 Problem formulation

Given a power distribution system with  $N$  nodes/buses the net power injection at bus  $i$ ,  $S_{inj,i}$ , is given by,

$$S_{inj,i} = S_{inv,i} - S_{dem,i} = f_i(V, \delta) \quad (4.1)$$

Where,  $S_{inv,i} = (P_{inv,i} + jQ_{inv,i})$ , represents inverter power injection at bus  $i$  and  $S_{dem,i} = (P_{dem,i} + jQ_{dem,i})$ , represents load power consumption at bus  $i$ .  $P$ ,  $Q$ , denote active power and reactive power respectively.  $V$  and  $\delta$  represents the vector of all node voltage magnitudes, and voltage angles. The  $N$  simultaneous equations generated by applying (4.1) for all  $N$  nodes define the power flow of the system. The droop function at the bus  $i$ ,  $g_i$ , which defines the inverter reactive power dispatch at time,  $t + \delta t$ , is written as,

$$Q_{inv,i}(t + \delta t) = g_i(V_i(t)) \quad (4.2)$$

$$\begin{aligned}
V_2 &= V_1 + dV_1 \\
V_3 &= V_2 + dV_2 \\
V_4 &= V_3 + dV_3
\end{aligned} \tag{4.3}$$

The function  $g_i$  is the piece-wise linear curve shown in Fig. 4.2 and defined below. The controllable variable for one inverter,  $x_i$ , has a dimension of six as given in (4.4). The vector  $x_i$  defines the VVC of the node  $i$ .

$$g_i(V_i) = \begin{cases} Q_1 & V_i \leq V_1 \\ -\frac{Q_1}{V_2 - V_1}(V_i - V_2) & V_1 < V_i \leq V_2 \\ 0 & V_2 < V_i \leq V_3 \\ \frac{Q_2}{V_4 - V_3}(V_i - V_3) & V_3 < V_i \leq V_4 \\ Q_2 & V_4 < V_i \leq V_i \end{cases}$$

$$x_i = \{Q_{i,1}, Q_{i,2}, V_{i,1}, dV_{i,1}, dV_{i,2}, dV_{i,3}\} \tag{4.4}$$

The optimization problem is formulated as follows:

$$\min_{\mathbf{x} \in S} U_{sys} = \frac{1}{T} \sum_{i=1}^T \left( \frac{1}{N} \sum_{i=1}^N (V_i - V_{rated,i})^2 \right)^{0.5} \tag{4.5}$$

subject to the constraints given by (4.6).

$V_{rated,i}$  is the ideal value for voltage magnitude at node  $i$  and  $T$  is the time period, and  $S$  is the solution space of the problem. The overall goal is to optimize voltage quality over a time period of  $T$ . Therefore, the average root mean squared value of the voltage magnitude, defined in (4.5), was chosen as the objective function.

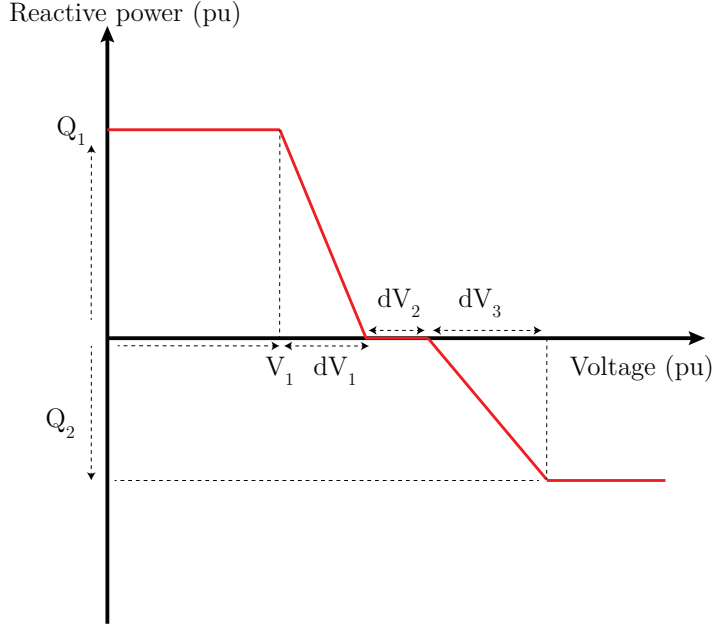


Figure 4.2: The volt-var curve as defined in IEEE 1547 standard. The curve is optimized by changing the six variables shown on this plot. The x values of the VVC is given by (4.3).

$$\begin{aligned}
Q_{1,min} &< Q_{i,1} < Q_{1,max} \\
Q_{2,max} &< Q_{i,2} < Q_{2,max} \\
V_{1,min} &< V_{i,1} < V_{1,max} \\
dV_{1,min} &< dV_{i,1} < dV_{1,max} \\
dV_{2,min} &< dV_{i,2} < dV_{2,max} \\
dV_{3,min} &< dV_{i,3} < dV_{3,max}
\end{aligned} \tag{4.6}$$

The optimal solution,  $\mathbf{x}$ , is a  $N \times 6$  matrix and takes the form given in (4.7).

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} Q_{1,1} & Q_{1,2} & V_{1,1} & dV_{1,1} & dV_{1,2} & dV_{1,3} \\ \vdots & \vdots & \ddots & \vdots & & \\ Q_{i,1} & Q_{i,2} & V_{i,1} & dV_{i,1} & dV_{i,2} & dV_{i,3} \\ \vdots & \vdots & \ddots & \vdots & & \\ Q_{N,1} & Q_{N,2} & V_{N,1} & dV_{N,1} & dV_{N,2} & dV_{N,3} \end{pmatrix} \tag{4.7}$$



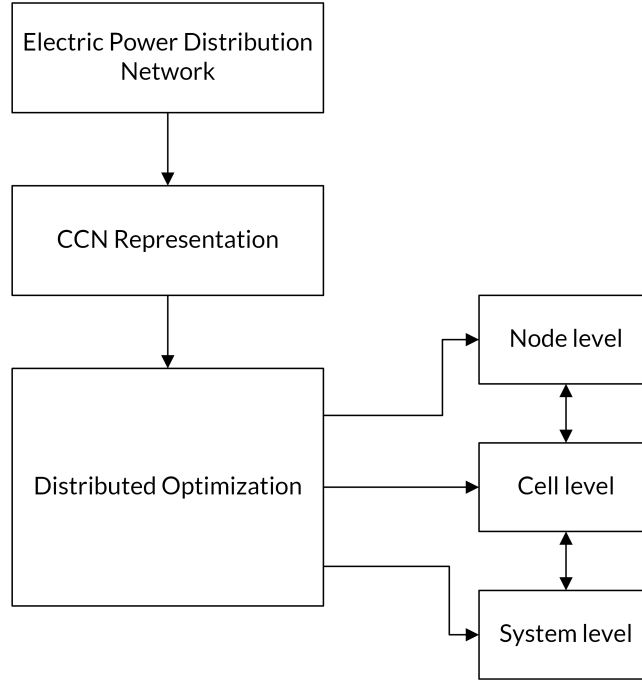


Figure 4.3: Overview of the proposed DOF.

### 4.3 Distributed optimization framework (DOF)

The structural overview of the framework is shown in Fig. 4.3. The developed implementation of the DOF has a decentralized execution, cell-wide objective function formulation, a global solution, and is implemented sequentially.

In the first step of the proposed framework, the CCN representation of the EPDS is developed based on [13]. Next, the CCN representation is used for distributed optimization based on three structural levels: node, cell, and system. The inverter devices are physically located at the nodes, and they contribute to the distributed optimization by sharing local measurements with the cell level operator and by adapting the volt-var curves based on cell operator signals.

The optimal volt-var curves for all nodes are evaluated at the cell level of the distributed optimization process. The utility function for optimization is limited to local optimization, and therefore, the optimization variables are the set of  $x$  vectors for local volt-var curves. The system-level iterative optimization algorithm uses the CCN representation to drive cell optimization.

### 4.3.1 CCN representation of the EPDS

The developed framework optimizes the volt-var curve of the inverters connected to the power distribution system, using the CCN representation of the EPDS in a distributed manner. In general, a cellular computational network consists of interconnected cells that interact with each other to realize a common goal. This approach has been used for distributed modeling, optimization, and control of power systems [95] [96] in the past. This study utilizes the cellular computational network representation for a power distribution system that was developed in [13] and applies it for distributed optimization of voltage control. The main advantage of this approach in comparison to other methods such as [89][93] is that it allows for the collection of nodes into a cell structure for efficient distributed optimization across a near-future time horizon. Based on available computational resources and the design goals, the designer has the flexibility in deciding on the level of optimization that needs to be carried out.

### 4.3.2 Node level control

The overview of the node-level control is shown in Fig. 4.4. In this method, the state-of-the-art VVC is extended by appending the ability to change the volt-var curve parameters dynamically. Therefore, the parameters of the volt-var curve ( $x$ ) are an external input to the inverter controller. The external information is provided by the cell-level optimization function and requires a communication channel. The distribution system modeling method presented in [13] is used in the selection of the cells to form the CCN representation. Since the DOF additionally requires modeling of the impact of the DER, [13] is extended to include DER control.

#### 4.3.2.1 DER modeling

The model applied for DERs in this study assumes that the consumer generation is based on solar PV plants. As described in the introduction, solar PV is the dominant DER in distribution systems, and it is expected to grow exponentially in the next few decades. If the primary energy source is changed from PV, it is still possible to use the presented approach to model any other DER by changing the generation model.

The source of energy is solar radiation which the PV cells convert to electricity [97]. Based on the PV characteristics, the maximum power dc power output from the PV plant at time  $t$ ,  $P_{DC(t)}$ ,

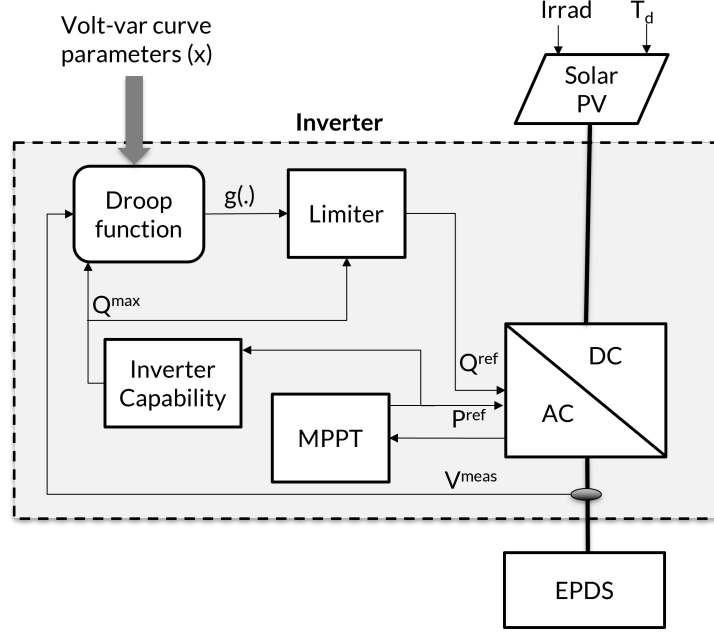


Figure 4.4: The overview of the prosumer based PV inverter control system. The volt-var curve parameters ( $x$ ) is changed by the cell operator. This is the node level optimization of the DOF.

can be determined by (4.8). Here,  $P_{pv, rated}$  is the maximum power from the PV array,  $Irrad(t)$  is the irradiation level at time  $t$ , and  $T_d(t)$  is the temperature derating factor from the power temperature characteristic curve. Here,  $P_{DC}$  is assumed to vary linearly with irradiation.

$$P_{DC}(t) = P_{pv, rated} \cdot Irrad(t) \cdot T_d(t) \quad (4.8)$$

The model of the inverter is given in Fig. 4.4. Typically the power setpoint for the inverter is based on an MPPT algorithm. For the time step of interest, it is assumed that MPPT-based reference value is equal to  $P_{ref}$ . The setpoint for reactive power,  $Q_{ref}$ , is determined based on the control loop shown in Fig. 4.4. The  $Q_{ref}$  is determined based on the inverter capability curve, the user selected priority (operation) mode (active power or reactive power priority), and the droop function (volt-var curve).

Assuming the only inverter capability constraint is current carrying capacity and inverter reactive power priority operation mode, the maximum reactive power that is available for dispatch,

$Q_{max}$ , is given by,

$$Q^{max} = \sqrt{(S^{max})^2 - (P^{ref})^2} \quad (4.9)$$

The value of  $Q^{max}$  is an input to the limiter function, and it changes with solar irradiance. The function  $g$ , which defines the volt-var characteristics of the active DER, shown in Fig. 4.2 has six independent parameters. The first two parameters define the maximum reactive power contribution (y-axis), and the last four determine the voltage for the reactive power contributions. The limiter considers the constraints and user-selected operation mode and provides the setpoints to the power electronic converter.

### 4.3.3 Cell optimization

The algorithm used for the optimization of cell  $k$  is shown in Fig. 4.5. An optimizer optimizes the selected cell. A time-series simulation that uses both the system model and expected operational data serves as an input to the optimizer. The operational data, the data set of independent load flow variables (loads and generation) for the optimized time window  $T$ , could be a near-future forecast, a persistent forecast, or a sliding window average data set. The optimizer computes the best possible set of optimization variables for the cell,  $x_{cell}(k)$ . This value is then updated in the system model.

The objective function for the optimization is the cell utility ( $U_{cell}$ ).  $U_{cell}$  takes into account the set of internal cell voltages and the boundary node voltage of the neighboring cells. The cell utility function,  $U_{cell}$ , is the root mean squared voltage and is given by (4.5). Here,  $m$  includes nodes belonging to cell  $k$  and the immediate neighbor nodes of cell  $k$ .

$$U_{cell} = \frac{1}{T} \sum_{i=1}^T \left( \frac{1}{m} \sum_{i=1}^m (V_i(\mathbf{x}) - V_{rated})^2 \right)^{0.5} \quad (4.10)$$

If a hypothetical distribution system consists of  $n$  cells, each of which has equal  $m$  volt-var curves, central optimization will require  $n \times m \times 6$  variables to be optimized. However, the proposed distributed approach will only need  $m \times 6$  parameters to be concurrently optimized. Therefore, the DOF significantly decreases the dimension of the optimization problem. The cell optimizer could be placed on a computing platform in the distribution system control center or placed on multiple

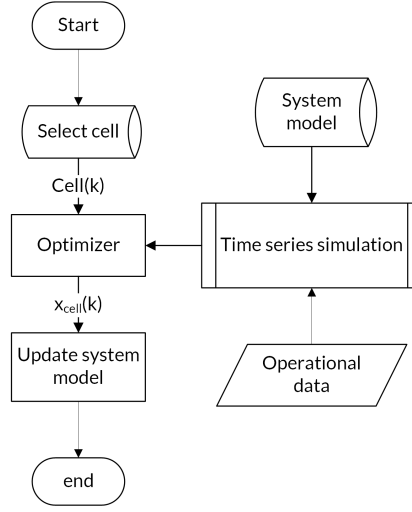


Figure 4.5: The implemented algorithm for cell level optimization.

spatially separated physical devices.

#### 4.3.4 System optimization

The system-level optimization is carried out by iteratively optimizing the cells. The priority for selecting cells for system optimization is based on an index defined as the cell rank, from the lowest to the highest CR. Two approaches for choosing the cell rank are presented in this study. The algorithm used for the optimization of cell  $k$  is shown in Fig. 4.6. The first step for system optimization is to develop the CR table. This provides the sequential order for cell optimization. Next, cell-level optimization of the cell with the current lowest priority (of non-optimized cells) is carried out. In the next step, the system objective function,  $U_{sys}$ , is evaluated, and a decision is made to iterate again or to stop the system optimization based on the stop criteria. The stop criteria for system optimization are either reaching maximum cell iterations or the required minimum value for system utility.

##### 4.3.4.1 Graph-based cell rank evaluation

In the graph-based approach, the CR is based on the graph structure of the CCN. The topology is processed, and the CR is given based on the tree traversal order from the furthestmost leaf cell to the root, covering every cell in the graph. The furthestmost leaf cell is given the lowest CR and will have the highest priority in system optimization. This approach is simple and only

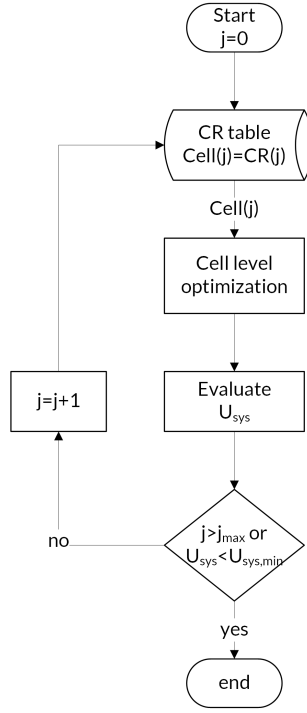


Figure 4.6: The implemented algorithm for system level optimization.

requires topology information.

#### 4.3.4.2 Impact-based cell rank evaluation

The objective of proposing the impact factor-based approach is to increase the performance of the iterative optimization. In this approach, the CR is based on a calculated impact factor that considers the ability of any cell to contribute to the optimization of the system voltage. The impact factor evaluates the combined effect of the system voltage sensitivity to reactive power and the reactive power injection capability of a cell.

The first step in the process is to find each cell's total apparent power capability. Next, the available reactive and active power are assembled to create a cellular virtual power plant representation. Each cell has a virtual reactive ( $Q_{cell}$ ) and active power load and generation. This step requires central information gathering and is used only to prioritize the optimization. Next, an artificial reactive power generation of  $\delta Q$  is injected into each cell. The  $\delta Q$  is distributed across the nodes proportional to the installed capacity of the generation. The  $U_{sys}$  sensitivity to this  $\delta Q$  is next evaluated by running an independent power flow and calculating the  $\frac{\partial U_{sys}}{\partial Q_{cell}}$  value of each cell.

The impact factor ( $IF$ ) of all the cells is then calculated by using (4.11). If the operating system point changes significantly, then a new impact factor calculation needs to be carried out since the impact factor is dependent on the power flow Jacobian, and the power flow Jacobian varies with the operating point.

$$IF = Q_{cell} \cdot \frac{\partial U_{sys}}{\partial Q_{cell}} \quad (4.11)$$

## 4.4 Case study

The case study is based on a modified IEEE 34 distribution system given in Fig. 4.7. It is a medium-sized long radial distribution system that originally existed in Arizona and had voltage control challenges due to its length. The system has 95 nodes and 68 loads. This system was modified by adding 28 PV inverters shown in Table 4.1. The system includes twenty-two three-phase PVs and eight single-phase PVs distributed across the feeder. The active power generation is limited to the load active power consumption. Therefore, this modified test case describes a system with approximately 100% DER penetration. The solar inverters are assumed to be rated for 120% of the rating of the connected PV. The irradiance curves are assigned to ensure the spatial correlation of solar irradiation and are based on four solar irradiation curves for the PVs given by Fig. 4.8. The loads are all set to four time-varying, and unique load curves are shown in Fig. 4.9. The loadshapes are assigned randomly to the loads. They are based on load time series data extracted from [98]. The PV irradiation shapes are assigned based on spatial information. The assignment of different loadshapes to power conversion components is given in Table 4.2. These time-variant loadshapes are the source of dynamics in this system. A Particle Swarm Optimization [99] based optimizer is selected for the cell optimization.

## 4.5 Results and discussion

Simulations are conducted for the two cell ranking strategies, and the results are compared with the state-of-the-art approach. Fig. 4.10 compares the utility between CCN optimized and standard VVC across a time frame of 1 minute. The utility value for the standard VVC is significantly higher than any of the two CCN optimized values across the time frame. The impact factor-based

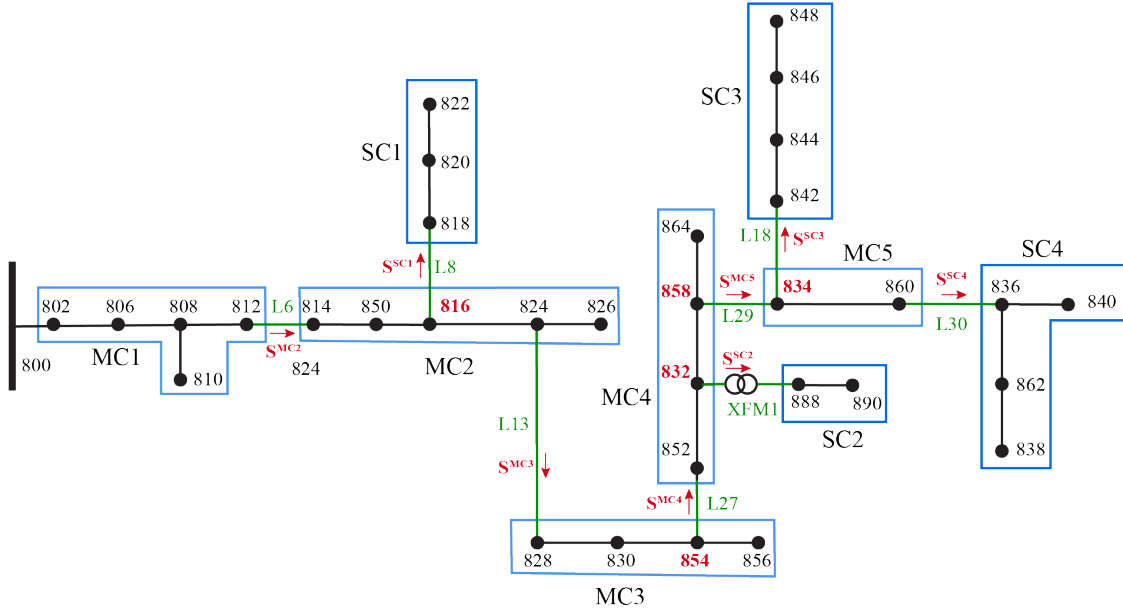


Figure 4.7: The CCN representation of the IEEE 34 bus system used for distributed optimization

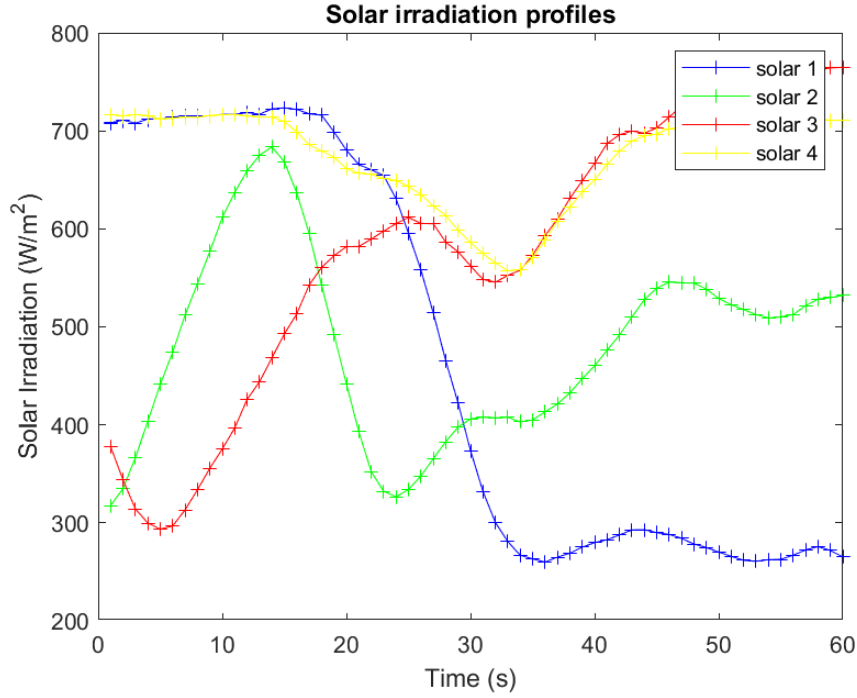


Figure 4.8: Solar irradiation variation across the one minute time frame of interest

approach shows better performance when compared to the graph-based approach. The difference in performance is comparatively significant in terms of the speed of optimization based on the



Table 4.1: DER connected to the test case.

Node	Inverter size (kVA)
802	60
806	60
808	32
810	32
816	10
818	68
820	342
822	274
824	90
826	80
828	14
830	51
834	248
836	97
838	56
840	83
842	19
844	661
846	96
848	135
852	8
854	8
856	8
858	39
860	392
862	56
864	4
890	604

Table 4.2: Assignment of different loadshapes to power conversion components in the case study.

Dynamic shape	Power conversion component
Solar 1	PV 802, 806, 808, 816, 824, 810, 818
Solar 2	PV 820, 822, 856, 826, 828, 830, 832
Solar 3	PV 864, 844, 848, 854, 858, 834, 836
Solar 4	PV 838, 860, 862, 840, 842, 846, 890
Load shape 1	Load 844, 802, 828, 854, 858, 836, 862
Load shape 2	Load 860, 830, 808, 824, 832, 838
Load shape 3	Load 890, 834, 836, 840, 864, 856
Load shape 4	Load 848, 806, 818, 820, 822, 816, 826

results shown in Table 4.3 and Table 4.4. The impact factor calculation requires an extra step of computation, as shown in Table 4.5. However, the additional computation step results in the impact factor requiring four fewer steps to arrive at the utility value of 0.0218 and therefore is well justified. Fig. 4.11 compares the corresponding mean voltage across the feeder. As expected, the mean voltage

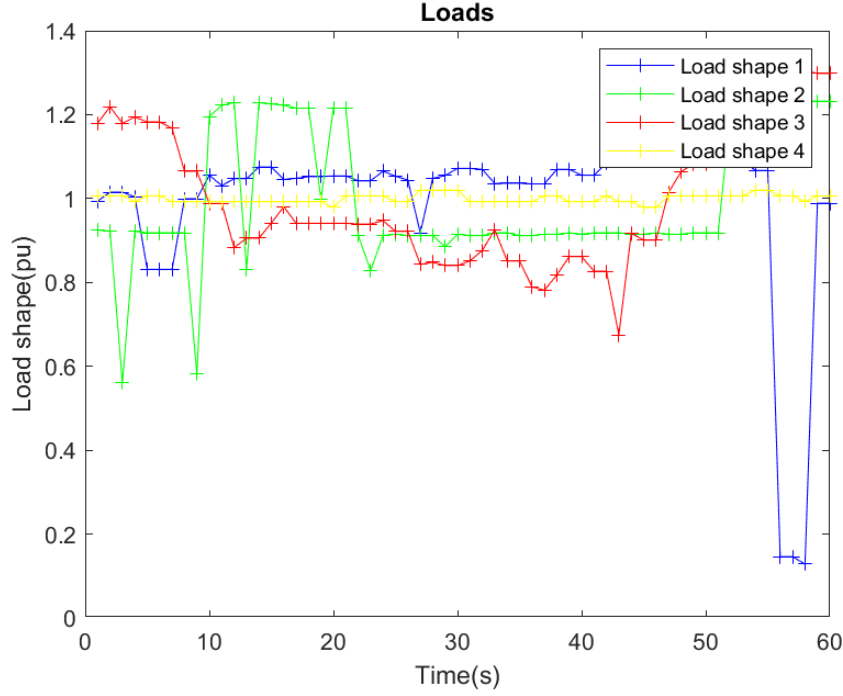


Figure 4.9: Load variation across the optimized one minute time frame

Table 4.3: Sequentially optimized utility values

No	Cell	U.sys	Unet.MC1	Unet.MC2	Unet.MC3	Unet.MC4	Unet.MC5	Unet.SC1	Unet.SC2	Unet.SC3	Unet.SC4
NA	wo VVC	0.0720	0.0567	0.0628	0.0725	0.0791	0.0830	0.0600	0.0802	0.0834	0.0831
0	Std VVC	0.0327	0.0396	0.0291	0.0289	0.0284	0.0306	0.0247	0.0227	0.0307	0.0302
1	SC4	0.0300	0.0394	0.0284	0.0259	0.0234	0.0243	0.0246	0.0187	0.0244	<b>0.0240</b>
2	MC5	0.0232	0.0368	0.0217	0.0150	0.0094	<b>0.0077</b>	0.0186	0.0088	0.0078	0.0073
3	SC3	0.0220	0.0359	0.0192	0.0116	0.0063	0.0047	0.0163	0.0070	<b>0.0047</b>	0.0047
4	MC4	0.0218	0.0358	0.0187	0.0110	<b>0.0060</b>	0.0047	0.0159	0.0068	0.0047	0.0048
5	SC2	0.0218	0.0357	0.0186	0.0109	0.0060	0.0049	0.0158	<b>0.0061</b>	0.0049	0.0050
6	MC3	0.0214	0.0353	0.0173	<b>0.0096</b>	0.0057	0.0052	0.0147	0.0061	0.0052	0.0053
7	MC2	0.0210	0.0344	<b>0.0151</b>	0.0080	0.0058	0.0058	0.0136	0.0067	0.0058	0.0063
8	SC1	0.0210	0.0342	0.0144	0.0083	0.0074	0.0074	<b>0.0101</b>	0.0078	0.0074	0.0075
9	MC1	0.0209	<b>0.0339</b>	0.0142	0.0083	0.0075	0.0075	0.0099	0.0079	0.0075	0.0077

of the CCN-based approach is significantly lower than the standard VVC. The calculated optimal power setpoints define the operation point to which the volt-var curves are shifted. The results clearly show that the local static volt-var control is insufficient to ensure optimal operation. This is due to the nature of droop-based control, where it is impossible to address the inherent steady-state error.

Figs. 4.8 and 4.9 show the source of the dynamics in the system that were applied to the loads and the PV plants. Fig. 4.12 compares the corresponding voltages of three randomly selected nodes in the system. The voltage of the CCN-based approaches is higher in quality and close to

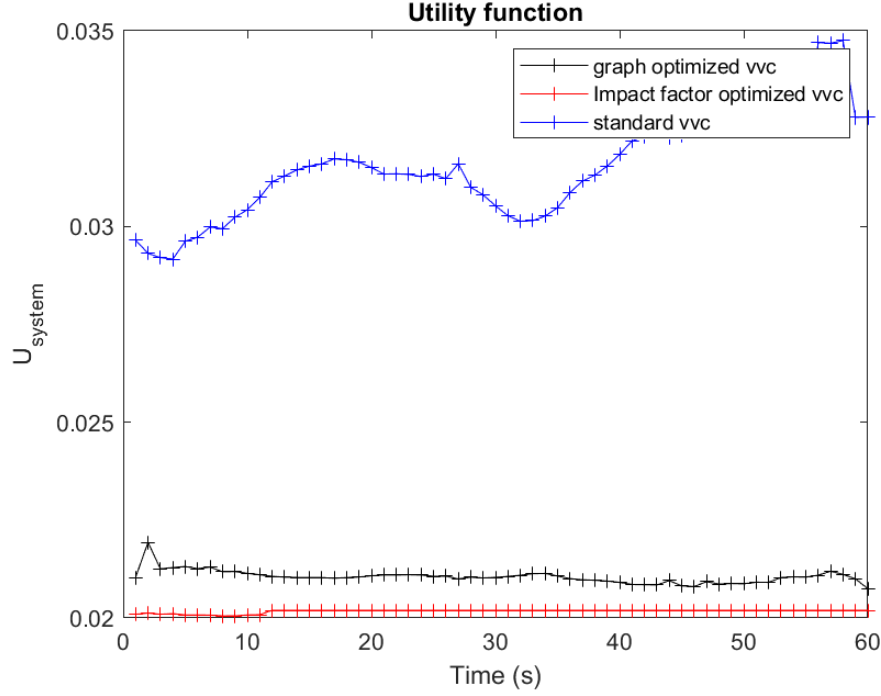


Figure 4.10: Comparison of Utility across the optimized one minute time frame

Table 4.4: Impact factor based priority based CCN: calculated utility function values for system and cells

No	Cell	U.sys	Unet.MC1	Unet.MC2	Unet.MC3	Unet.MC4	Unet.MC5	Unet.SC1	Unet.SC2	Unet.SC3	Unet.SC4
0	Std VVC	0.0327	0.0396	0.0291	0.0289	0.0284	0.0306	0.0247	0.0227	0.0307	0.0302
1	SC3	0.0217	0.0355	0.0182	0.0107	0.0059	0.0047	0.0159	0.0071	0.0047	0.0051
2	SC2	0.0216	0.0354	0.0180	0.0105	0.0061	0.0048	0.0157	0.0068	0.0047	0.0051
3	MC5	0.0216	0.0355	0.0184	0.0111	0.0063	0.0045	0.0160	0.0066	0.0045	0.0046
4	SC4	0.0216	0.0356	0.0184	0.0111	0.0059	0.0037	0.0159	0.0059	0.0037	0.0035
5	MC2	0.0212	0.0347	0.0163	0.0095	0.0061	0.0047	0.0147	0.0070	0.0047	0.0050
6	SC1	0.0212	0.0346	0.0156	0.0096	0.0072	0.0061	0.0110	0.0079	0.0061	0.0061
7	MC3	0.0209	0.0341	0.0144	0.0085	0.0070	0.0062	0.0100	0.0080	0.0062	0.0062
8	MC4	0.0209	0.0342	0.0146	0.0086	0.0068	0.0059	0.0102	0.0078	0.0059	0.0060
9	MC1	0.0208	0.0339	0.0144	0.0086	0.0069	0.0062	0.0101	0.0080	0.0061	0.0062

the desired value of 1 pu. Fig. 4.13 compares the performance of the graph-based method with the impact factor-based approach. The results show that the computation time per optimized cell is comparable for both methods, whereas the utility decreases much faster for the impact factor-based approach. Therefore, the impact factor-based approach performs better than the graph-based approach.

Fig. 4.14 compares the volt-var curves for the three operation conditions. Based on Table 4.4, SC3 is the first cell to be optimized in the impact-based method. From the volt-var curve, it is observed that node 844 has significant reactive power absorption in comparison to the other two

Table 4.5: Result from the impact factor evaluation based on available reactive power resources and  $U_{sys}$  for cell reactive power injection

Cell	$dU_{sys}/dQ$	Q	$Q \cdot dU_{sys}/dQ$	Sequence
SC3	0.0000300	2730	0.0819	1
SC2	0.0000333	1811	0.0603	2
MC5	0.0000299	1921	0.0575	3
SC4	0.0000300	765	0.0230	4
MC2	0.0000191	541	0.0103	5
SC1	0.0000149	684	0.0102	6
MC3	0.0000239	228	0.0054	7
MC4	0.0000293	161	0.0047	8
MC1	0.0000068	555	0.0037	9

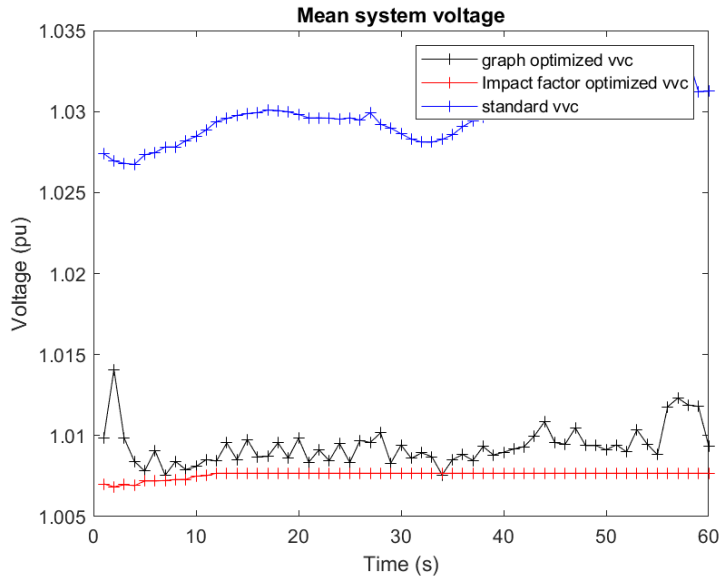


Figure 4.11: Comparison of voltage across the optimized one minute time frame

methods. Since 844 is the largest DER, it is clear that cell SC3 is facing an overvoltage condition, which is compensated by selecting the volt-var curve such that SC3 is absorbing reactive power. Instead of the adaptive optimization, if standard VVC was used, the cell's potential to contribute to voltage control of the system would have been unused. Additionally, the impact factor-based approach can accelerate system optimization since it optimizes in merit order.

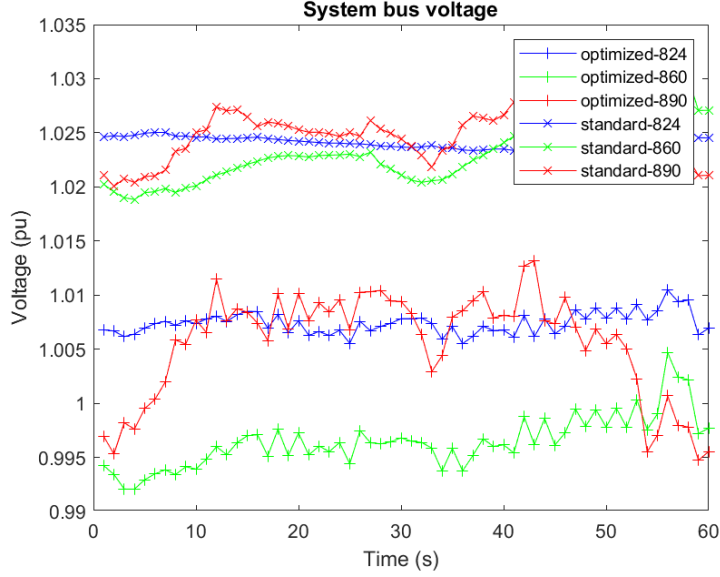


Figure 4.12: Comparison of voltage of 3 randomly selected nodes across the optimized one minute time frame

## 4.6 Summary

Due to the increasing penetration of distributed energy resources (DER), Voltage control in modern electric power distribution systems has become challenging. The current state-of-the-art voltage control is based on static and pre-set DER volt-var curves and do not provide sufficient flexibility to address the dynamic aspects of the voltage control problem in a power distribution system.

This study has presented a distributed optimization framework for volt-var curves of DERs in electric power distribution systems. The DOF is based on a cellular computational network representation of the electric power distribution system. The proposed method enhanced the system voltage control by optimizing multiple volt-var curve parameters concurrently. The method had three functional levels, node, cell, and system. The node-level used the adaptive VVC setpoints generated by the cell level. The cell level optimized the utility for the cell and generated the VVC set points for the nodes. The system-level iteratively called on cells based on the cell ranking and converged the cell parameters towards near-global optimality.

The cellular optimization approach enabled system-wide optimization. The cell optimization was prioritized using two methods, graph, and impact-based methods. The impact-based method

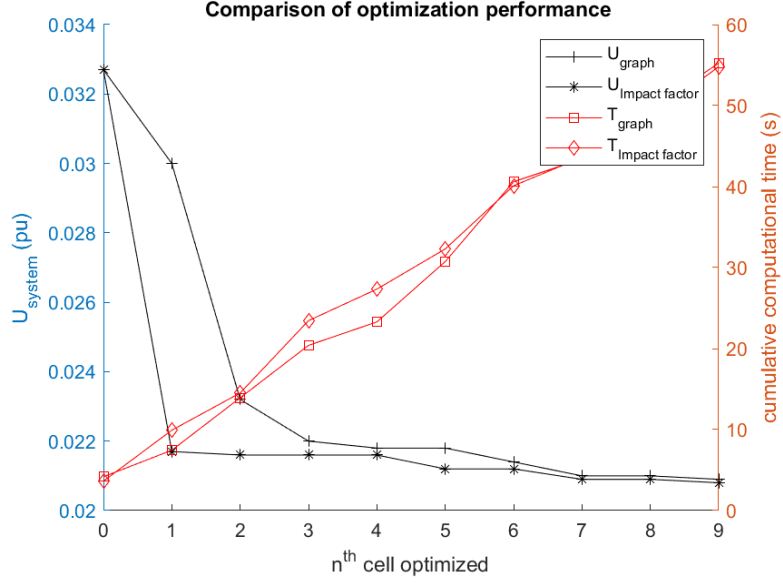


Figure 4.13: Comparison of optimization performance. Here  $T_{graph}$  and  $T_{Impact\ factor}$  represents the cumulative time for each iteration, where time is given in the right y axis, and the  $U_{graph}$  and  $U_{Impact\ factor}$  gives the Utility value for each iteration, where utility is shown in the left y axis.

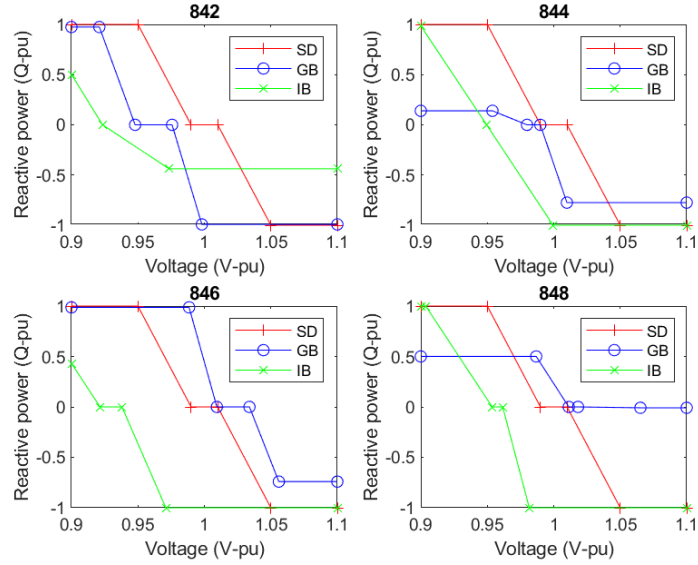


Figure 4.14: The volt-var curves for cell SC3

required extra initial computational efforts but thereafter provided better computational throughput than the graph-based method. The DOF was illustrated on a modified standard distribution test case with several DERs. The results from the test case demonstrated that the DOF-based volt-var

optimization results in consistently better performance than the state-of-the-art volt-var control. The proposed approach needed less computation and allowed for distributed scalable optimization of a power distribution system. A potential future study is to extend cell DOF from sequential to parallel implementation to allow for accelerated performance.

## Chapter 5

# A Real-Time Distribution System Simulation Laboratory Testbed

This chapter describes the design of an innovative testbed to demonstrate the real-time simulation of a distribution system. In the current state-of-the-art, studies are usually executed offline using software such as PSCAD [100], Matpower, OpenDSS, and CYME [101]. Offline studies are a vital component of power distribution system design. Still, it cannot reproduce some realistic scenarios. It does not provide the functionality to test features that need to be considered in modern power system design, such as connected grid-edge devices, communication, human-machine interfaces, and control. Demonstrating an operation and control method on a real-time system is the closest to real-world implementation. Successful implementation on a real-time system builds confidence in the field implementation. It allows for testing the operation in untested scenarios when simulated in a non-real-time system.

When conducting real-time simulations, it is important to relevant data of device characteristics to ensure that realistic scenarios close to the real world system are captured. Ideally, if an operator is trying to understand the behavior of the system by modeling the part of the system owned by the operator, it is important to be able to model, capture and integrate the characteristics of the devices that are outside of the ownership and control of the owner. In addition the emulated external device characteristics need to include spatial and temporal correlation.

The contribution of the work presented in this chapter includes the design and implementa-



tion of a comprehensive cyber-physical testbed for power distribution system real-time simulations, as well as a method that allows capturing the dynamics of devices in a real system, even if they are not controllable. In a deregulated power system, only parts of the distribution system are typically operated by a single operator. Based on the data-in-the-loop method introduced in this chapter, an operator can simulate the operation of their distribution system more realistically by projecting the real-time characteristics of the uncontrollable devices that are potentially owned, controlled, and operated by other operators. Using the data-in-the-loop approach presented in the chapter, the quality of the simulations, and the insights obtained, will be more realistic.

## 5.1 Introduction

The typical distribution system simulations are based on quasi-static time series analysis, which assumes that the system reaches a steady-state before the end of each time step. It is also assumed that there are means to get the measurements from the system, run optimizations, update field controls, and provide the output to the distribution system components. However, it is vital to consider the constraint and technology available to measure, communicate, and compute in a natural system. Additionally, the system performance considering the real-time constraints and operating conditions in terms of stability and performance degradation need to be considered, and methods to address these challenges need to be integrated into the system design. A real-time testbed helps first to understand and then address these challenges. A real-time simulation and testbed bring out more of these implementation challenges helping to close the gap between academic research to application in the industrial setting.

If an ideal real-time simulator were to be run alongside the real system, the output of both systems would match in each instance in time. Additionally, there will be means to measure the values and connect external devices to the real-time system. The response and interaction with the external devices would be the same for both the real system and the simulator. The different components that can be tested for a design using a real-time testbed are shown in Fig. 5.1. The generic real-time simulator can implement devices with controls inside the simulator yet still perform in real-time and connect and interface with external devices. The measurements from both external and internal devices can be communicated via physical communication channels to different external Internet of Things (IoT) control devices ) connected via a cyber layer. These IoT-based devices will

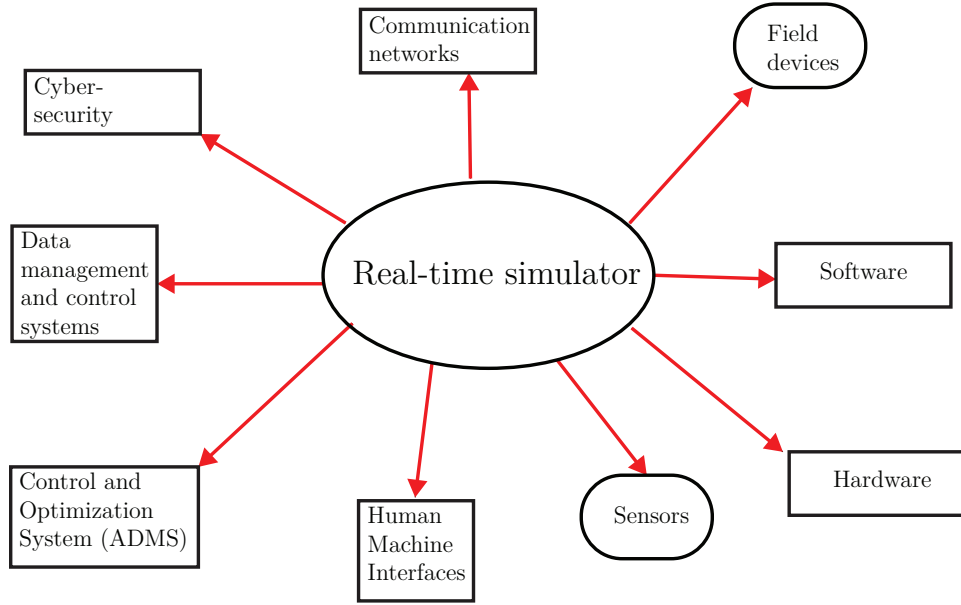


Figure 5.1: High level overview of the various types of components/sub-system that can be tested using a real-time hardware-in-the-loop simulation

communicate with the simulated devices inside the simulator and the actual devices in the course of system operation. The different types of real-time simulation systems currently available to users are analog systems [102], which are miniature versions of the simulated real system, and digital systems [103]. Digital simulators are widely used since they can be easily reconfigured to different model systems. The integration with natural controls or real hardware is possible by interfacing the digital simulator using analog to digital and digital to analog converters, among other supporting signal conditioning devices.

## 5.2 Real-time power system simulations

For a simulator to be given the 'real-time' status, the internal computation time for one simulation point should be less than that of a simulation time step. For a typical real-time simulation that represents fast dynamics such as electromagnetic transients, the typical time step is in the range of  $50\ \mu\text{s}$  [103] [104]. In the state-of-the-art of the real-time simulators, some simulators can go down to  $50\ \text{ns}$  [105], thus having the ability to simulate the fast switching transients associated with modern power electronic converters.

The bare-bones version of a real-time simulator can only execute a fast power system simula-

tion. The accelerated performance optimizes the workflow for power system studies [106]. However, it does not have any other value except for saving time. The next version of a real-time simulator can interconnect to external devices. These simulators have analog or digital interfaces that can connect to external devices. However, the simulator functionality is limited to connecting to control interfaces. This type of simulator is widely used and has a range of use cases, including testing devices such as protection relays and power system component controllers. The control-in-the-loop-based real-time simulator can be enhanced by integrating with power amplifiers and providing the ability to interface power signals/devices with the RTDS and provide the simulator the ability to test a complete power apparatus. These are highly valued real-time simulators and have a range of use cases.

Another version of the real-time simulation, which is introduced and proposed in this chapter, is called a data-in-the-loop simulation. Here, data from a live device is interfaced with the real-time simulation to simulate the real operational conditions of a live field device. The non-availability of actual system data for real-time simulation has been a significant drawback to understanding the functional characteristic of a system using simulation, especially when it comes to real-time functionalities. The data-in-the-loop simulation approach addresses this research gap. The objective of this testbed is to show the validity of the data-in-the-loop approach to integrating live devices and show the real-time operational capabilities and its practical field implementation.

### 5.3 Design of real-time distribution system test-bed

The functional overview of the real-time distribution system testbed is shown in Fig. 5.2. The designed system consists of four different subsystems. The core subsystem is the real-time simulator which consists of a real-time simulator and has integrated communication modules. The distribution system is modeled in the real-time simulator. It includes passive components such as lines, transformers, and active systems such as DER-based power generators and loads. The measurements taken from the distribution system simulation are communicated via a given protocol using a communication network to an IoT.

The Internet of Things (IoT) includes a data manager, a cell advanced distribution management system (C-ADMS), and data storage. The cell archives the system measurements, and processes, shares the data with the C-ADMS, manages the data visualization, processes and condi-

tions the data from the field devices, and exchanges information with neighboring cell IoTs. The C-ADMS runs an optimization algorithm and calculates the setpoints for the power distribution using the data provided by the data manager. The calculated setpoints are then sent to the simulated devices in the distribution system via the data manager. The inputs to the C-ADMS also consider information from the local users, who are allowed to control the local system. The data visualization is provided via a server hosted on the IoT and managed by the data manager. This allows for remote access to the cell visualizations. The data for the emulated field devices are also sent to the IoT. This data stream is conditioned for consumption and shared with the real-time simulator by the data manager. The overarching design principle revolves around spatially distributed IoT platforms, and therefore, it does not have a single point of failure or a single control room.

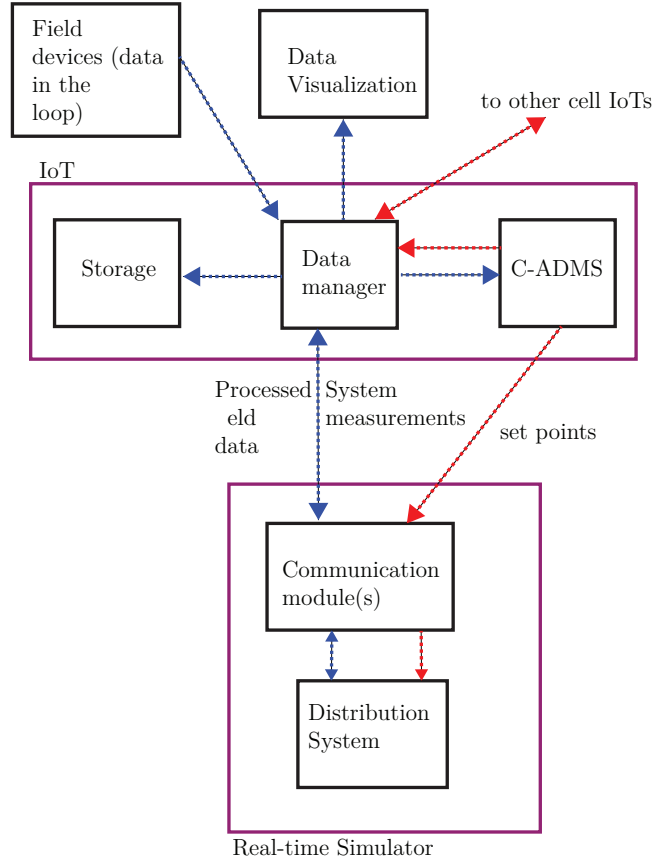


Figure 5.2: Overview diagram of the real-time simulation testbed.

### 5.3.1 Communication

Based on the overview diagram, it is clear that communication plays a significant part in the testbed. Communication modules are required for information translation at the simulator, at the data manager, and between the IoTs and the field devices. This requires a reliable communication network, but the system should also be designed to operate robustly even if a communication link fails. The network could have a heterogeneous physical layer - with options ranging from Ethernet to serial protocol and have unique application layer protocols, ranging from a raw TCP socket to protocols such as IEEE C37, DNP3, and Modbus. The scalability of the communication architecture is a significant design challenge.

### 5.3.2 Data-in-the-loop simulation (DIL)

The implemented data-in-the-loop simulation can integrate a distribution network connected live device to a real-time simulation and recreates realistic operational conditions in a distribution system simulation. The overview of the data-in-the-loop method is shown in Fig. 5.3. The emulated field device is connected to a device IoT that measures and converts the measurements to a data stream. The data is communicated to a remote IoT using a communication network. The remote IoT parses the protocol and uses a data processor to store the information while re-streaming it to the real-time simulator using a data reader and protocol parser. The communication (IO-Input/Output) module of the real-time simulator receives the data from the communication network. The data is then unpacked, parsed, and passed on to a data model. The data model uses the streaming data to recreate the impact of the emulated field device and integrates the emulated device as a power system component into the real-time simulation.

## 5.4 Implementation of laboratory testbed

The proposed distribution system testbed is set up in the Real-time Power and Intelligent System (RTPIS) laboratory at the Clemson University and is centered around one of the real-time digital simulators available in the lab, manufactured by Opal-RT technologies [107]. The hardware modules are used to create the different interfaces required for interfacing with external components and integrating both the hardware and the software being tested. The modified IEEE 34 bus system shown in Fig. 5.4 is implemented in the real-time system. Fig. 5.4 also identifies the different

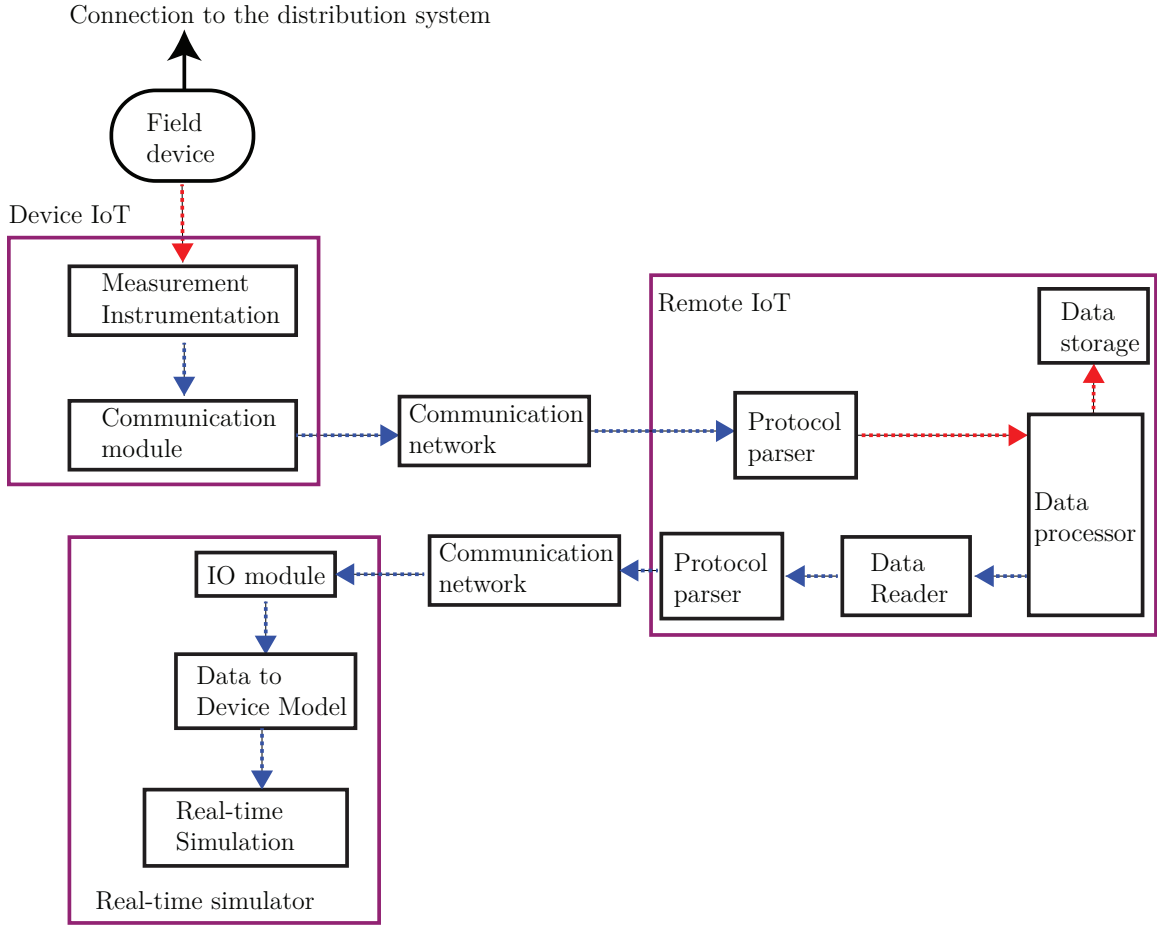


Figure 5.3: Overview diagram showing the operation of the data-in-the-loop process.

components that are used to create the testbed. The laboratory setup is used to illustrate the real-time operation of the testbed. Within the implementation of the testbed, a multitude of Open source software and platforms, vendor-specific and custom-built software is implemented and installed in the Clemson RTPIS lab as shown in Fig. 5.5.

#### 5.4.1 Real-time simulator

The Opal-RT technologies OP5607 real-time simulator is the core component of the testbed. The Artemis solver is used to solve the power distribution system modeled using the Matlab Simulink Simscape toolbox. An EMTP version of the IEEE 34 test case is modeled using the Artemis toolbox and is simulated in real-time on the OP5607 using the RTlab software. The IEEE 34 test case is modified by integrating dynamic elements into it. The Artemis toolbox solver is used to run the

system at  $50\mu s$  time steps, and the rack has two cores rated at 3.3 GHz (x86 architecture) to utilize for the simulation. The Opal-RT is connected to the cell IoT via the Ethernet-based local area network. The communication to and from the IoT is implemented by creating TCP servers and clients at the end of the IoT and the Opal-RT simulator.

## 5.4.2 Implementation of DIL

The R06 PV plant (1 MW capacity) of the Clemson University is interfaced by acquiring the micropmu data and then conditioning it to match the input to the real-time simulation. The internal generator in the simulator is controlled to match the real-time output of the R06 by using a TCP socket communication channel. Similar to the PV plant, the Riggs hall building load is captured in real-time by the micropmu, and a load is controlled inside the simulation to project the characteristics of this actual load.

The generic process for data-in-the-loop simulations is given in Fig. 5.3. The voltage and current are measured via VTs and CTs for both devices emulated in this testbed. A communication module formats and streams the data in IEEE C37.118.1-2011 protocol via the local area network. The micropmu [108] that provides this functionality is a Powerside Pcube 3 module. The physical layer of the connections is Ethernet. The data streams are parsed, processed, and stored at the IoT using the Openhistorian [109] that runs the IoT platform. The IoT platform is an Intel computer installed with a Xeon 6234 3.30 GHz processor running on a 64-bit Windows 10 OS. It includes 64GB of RAM and 16GB NVIDIA Quadro RTX 5000 GPU. The Openhistorian also calculates the active power from the extracted voltage and current phasor angles. The power values are streamed through a custom-built Python processing engine which serves the functionality of both TCP client and data access from the Openhistorian. It also performs data conditioning, where the fast data stream from the PMU is downsampled to match the system requirements. The visualization web server is a Grafana [110] dashboard that runs on the IoT platform. The Grafana client can be run either at the local IoT or at a local IoT.

The TCP data stream is received and parsed by TCP Socket servers set up at the IO interface of the Opal RT system. The data are then connected to the signals in the Simscape simulation and integrated to the simulation through a data model. An example of the data model is the PV model that emulates the R06 PV plant. The data model for a PV plant is based on the diagram given in Fig. 4.4, where the  $P^{ref}$  value from the MPPT is substituted by the value received from the IoT.

The same design method can be used to emulate other types of field devices. The data-in-the-loop simulation plots shown in Fig. 5.8 verify the accuracy of the simulated power plant by comparing the output from the PV plant running on the Opal RT real-time simulation with the actual PV plant power measurements.

### 5.4.3 IoT

Each cell has a separate interface and optimization engine. This could be either a physical or virtual separation. The IoT is the cell controller, and optimization is run on this controller. This also manages cell-level communication. IoTs communicate via TCP sockets, and visualization is using a Grafana dashboard. An external Grafana client can connect to the Grafana to stream the visualization across the network. Python scripts control data conditioning, calculation, and TCP socket-based data communication. The operation and control software will use the provided information to control and optimize processes and communicate the resulting actions to the physical and simulated controllers of the distribution system-connected devices.

### 5.4.4 Other devices

The real-time simulation could also have other devices such as integrated digital switches and HMI for operators. The system can get the status of different devices such as circuit breakers via the digital output interfaces of the simulator. The digital inputs can integrate digital controls such as circuit breaker trip or disconnection of a circuit breaker via a manual or automatic operator.

## 5.5 System demonstration

A sample view from the Grafana dashboard that provides visualization of the entire R06 PV plant active power output and the total Riggs hall load power consumption is shown in Fig. 5.9. The real-time impact of providing active grid support from the PV plant at night time with no active power generation is shown in Fig. 5.10. In this dashboard view, the Riggs Hall load stays close to 30 kW, and the PV plant's active participation is turned on to demonstrate the real-time performance. The voltage is at a low 0.898 when there is no active reactive power participation. The voltage increases to a healthy value when the PV plant changes to provide reactive power support. The system operates with the PV plant turned off time 50 seconds, at which point the PV plant



is connected. The PV plant vvc operation mode is turned on at 100 seconds. The variation of the voltage and the power of the PV plant and the load is shown in Fig. 5.10.

In order to demonstrate real-time simulation capabilities over a longer time duration, the system is simulated between 7 pm and 8 pm on 5/2/2022, and results are plotted in Fig. 5.12. This plot demonstrates how the cell SC3 changes from a net energy importer to a net energy exporter as the solar irradiation gradually decreases in the evening. The corresponding performance of the system with standard vvc shows how the active power injection supports the system voltage to fill the ancillary service gap opened by the decreasing generation of the solar power plant. During the simulated time period, the solar energy production steadily decreases, changing the node from a net power exporter to a net power importer to serve the night load. The system voltage steadily decreases, but it stays within acceptable quality limits due to the reactive power support.

The system performance is compared with the standard volt var curve control method by simulating the same dataset for operation without vvc, and the performance is plotted in Fig. 5.13. This plot demonstrates how the cell SC3 changes from a net energy importer to a net energy exporter in the evening while the PC plant operates with a fixed power factor of one, in which case the PV plant is not contributing any ancillary services to the system. During the simulated period, the solar energy production steadily decreases, thereby changing the node from a net power exporter to a net power importer to serve the night load. The system voltage steadily drops, and without reactive power contribution from the inverter, the voltage continues to be outside the typical grid code.

The system performance for optimized operation is compared with the standard volt var operation by simulating the same dataset for operation with optimized vvc, and the performance is plotted in Fig. 5.14. This plot demonstrates how the cell SC3 changes from a net energy importer to a net energy exporter while the PV plant optimally contributes to the ancillary services of the system. During the simulated period, the solar energy production steadily decreases, thereby changing the node from a net power exporter to a net power importer to serve the night load. The system voltage is optimal, even though a slight decrease in voltage is observed with the decreasing solar power generation as the day changes to night over the short one hour of operation. The real-time simulation result shows how online optimization can enhance the node voltage beyond what is possible by using a static and standard volt var curve.

## 5.6 Summary

A real-time distribution system testbed is designed and implemented in the RTPIS Lab of Clemson University. The platform includes a approach to emulating live field devices based on a data-in-the-loop simulation introduced and implemented in this chapter. The real-time testbed models a modified IEEE 34 distribution system. The testbed includes external devices and simulated power system components. The functionalities and the setup of the different components required for the testbed design are discussed, and specific vendor and custom components used for the development of the testbed in the lab are presented. The testbed is used to demonstrate the control of a PV plant. The demonstrations show the ability of this testbed to model and control a power distribution system in real-time. The testbed can be expanded by integrating Modbus or DNP3 at the TCP/IP transport layer instead of the generic TCP socket communication architecture that is being used. The scalability can be addressed by integrating ODBC based control data management system with an automation architecture based on the IEC 61850 standard [111].



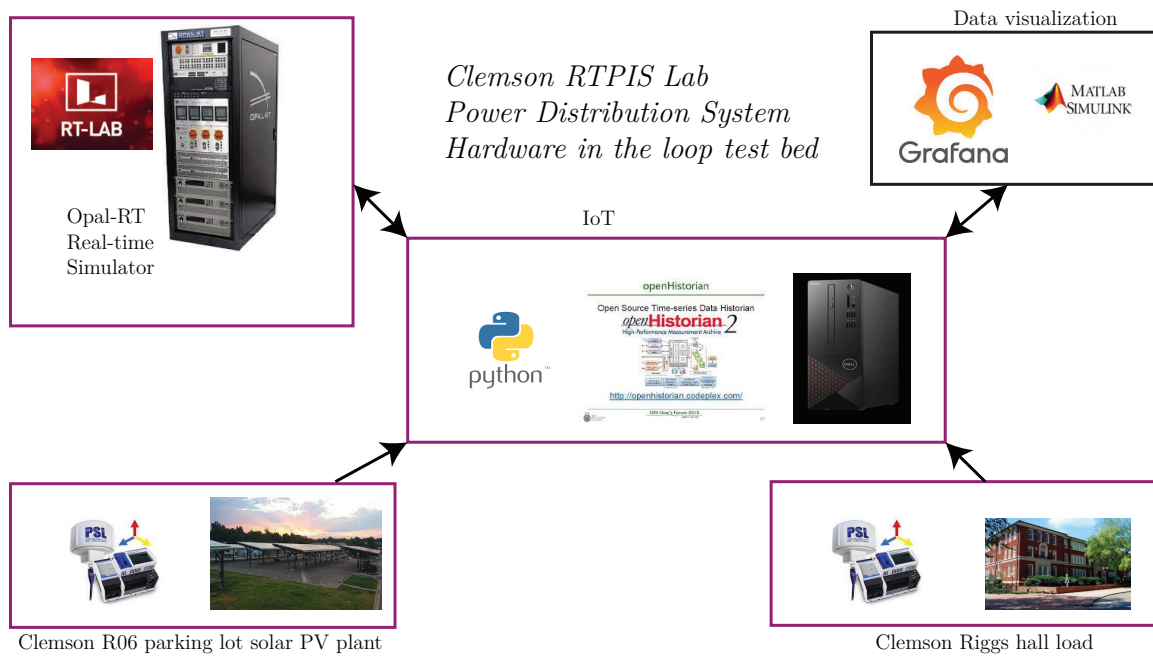


Figure 5.5: The overview of the components and software used for implementation of the Clemson University real-time Power and Intelligent systems lab real-time power distribution hardware-in-the-loop testbed.

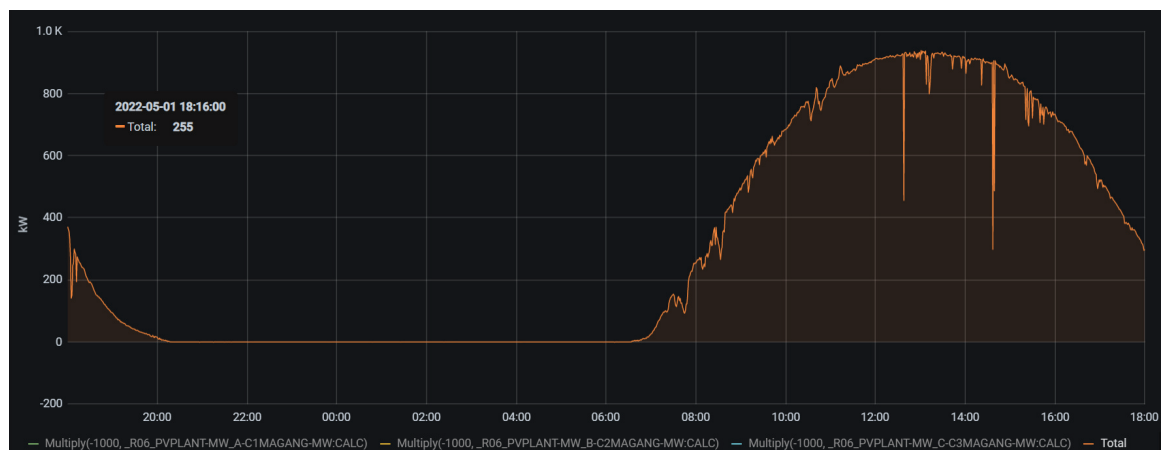


Figure 5.6: The actual R06 PV plant generation for 24 hours between 5/1/2022 and 5/2/2022.

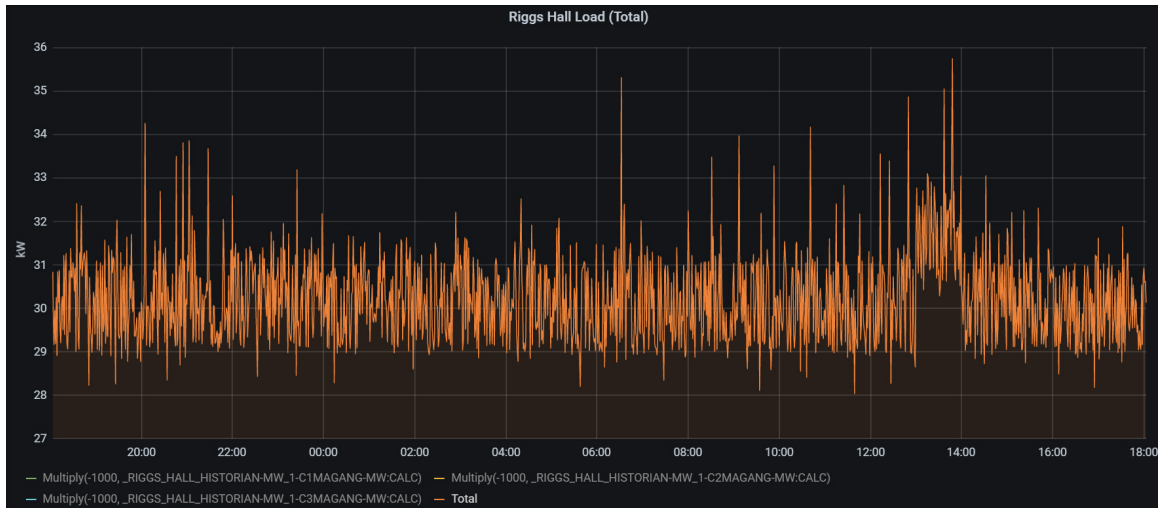


Figure 5.7: The actual Riggs hall building load for 24 hours between 5/1/2022 and 5/2/2022.

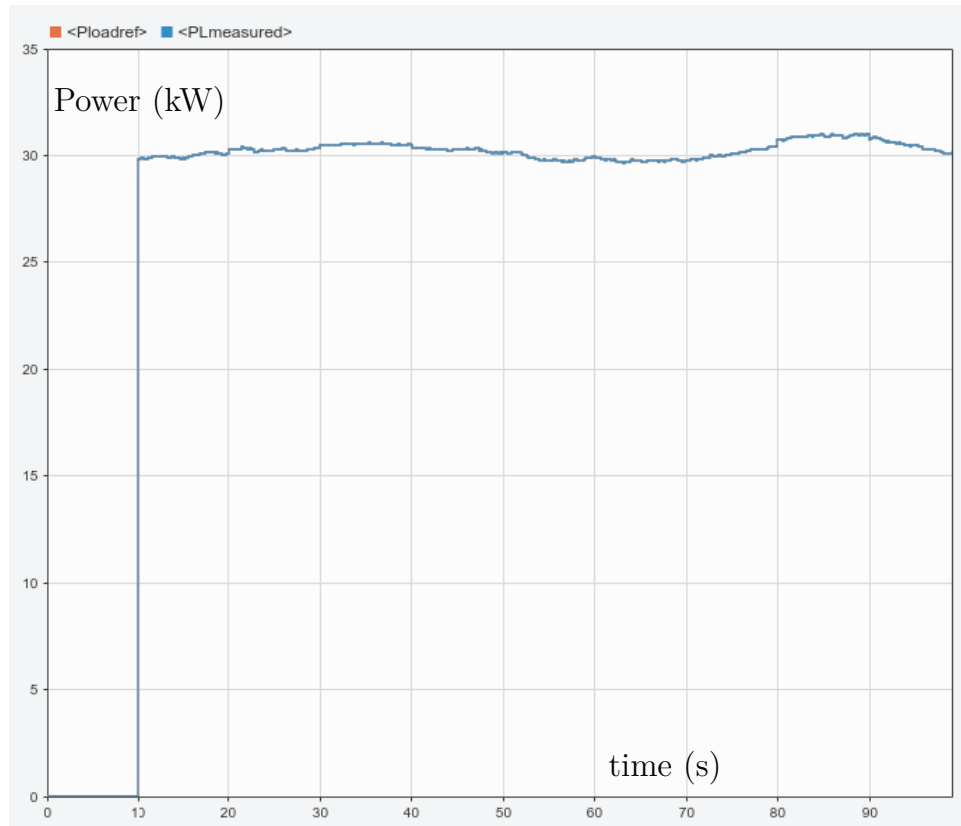


Figure 5.8: Verification of the data-in-the-loop real-time simulation in the simulink dashboard. The reference power is matched by the measured power output from the emulated device.

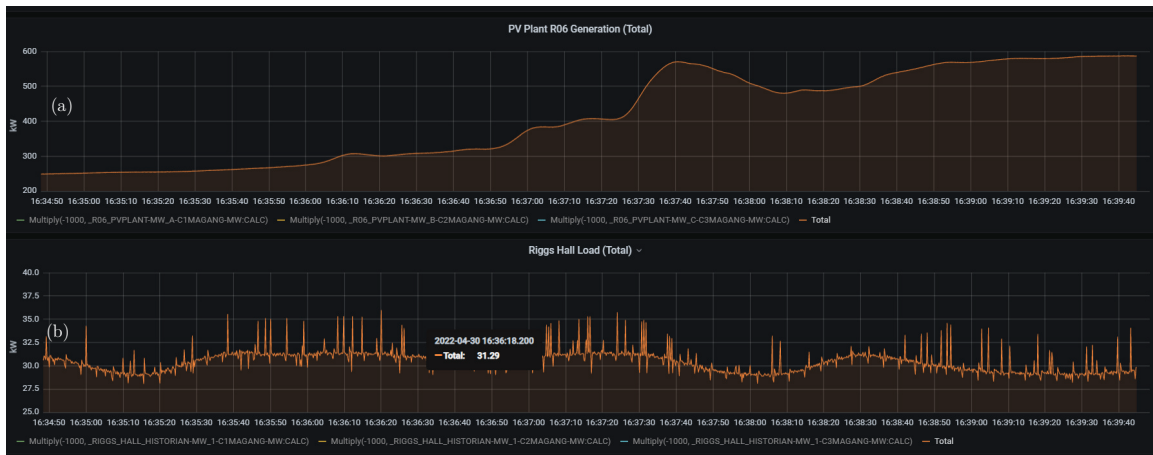


Figure 5.9: The data visualization platform developed for the Openhistorian data processor based on a Grafana dashboard.

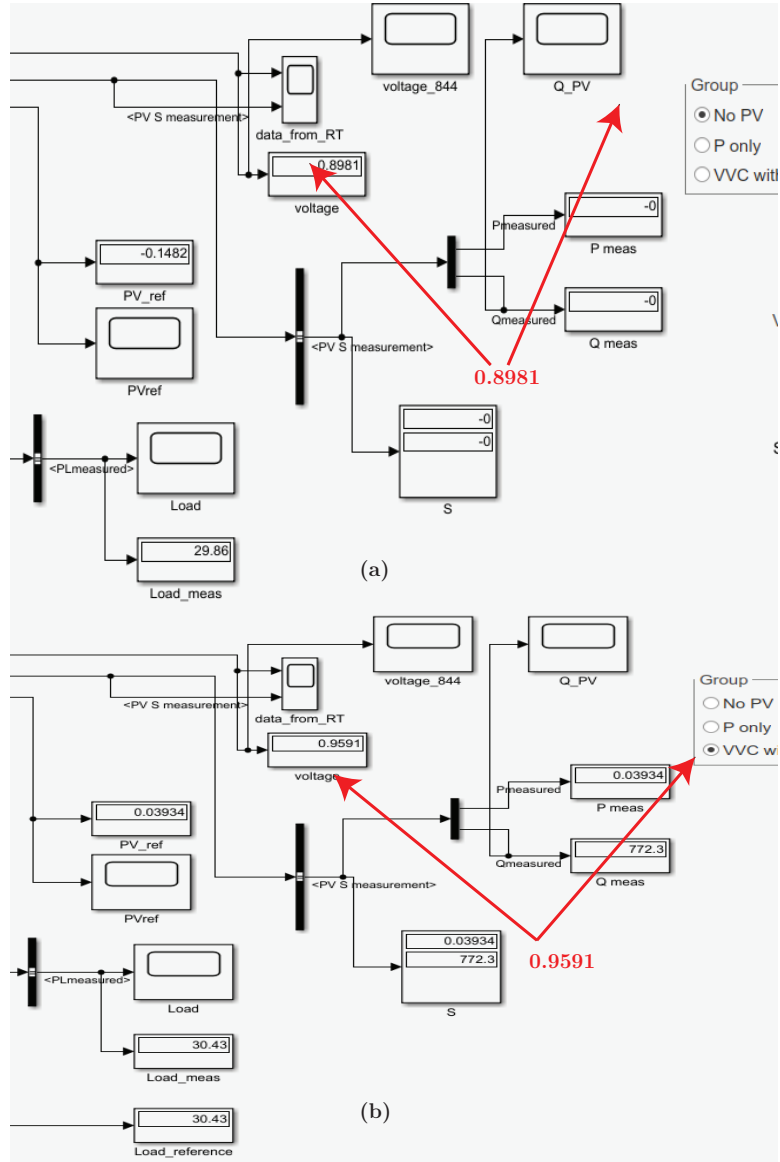


Figure 5.10: The real-time simulation output comparison from the simulink dashboard. (a) The system voltage without active reactive power contribution from the PV plant. (b) The system voltage with active reactive power contribution from the PV plant.

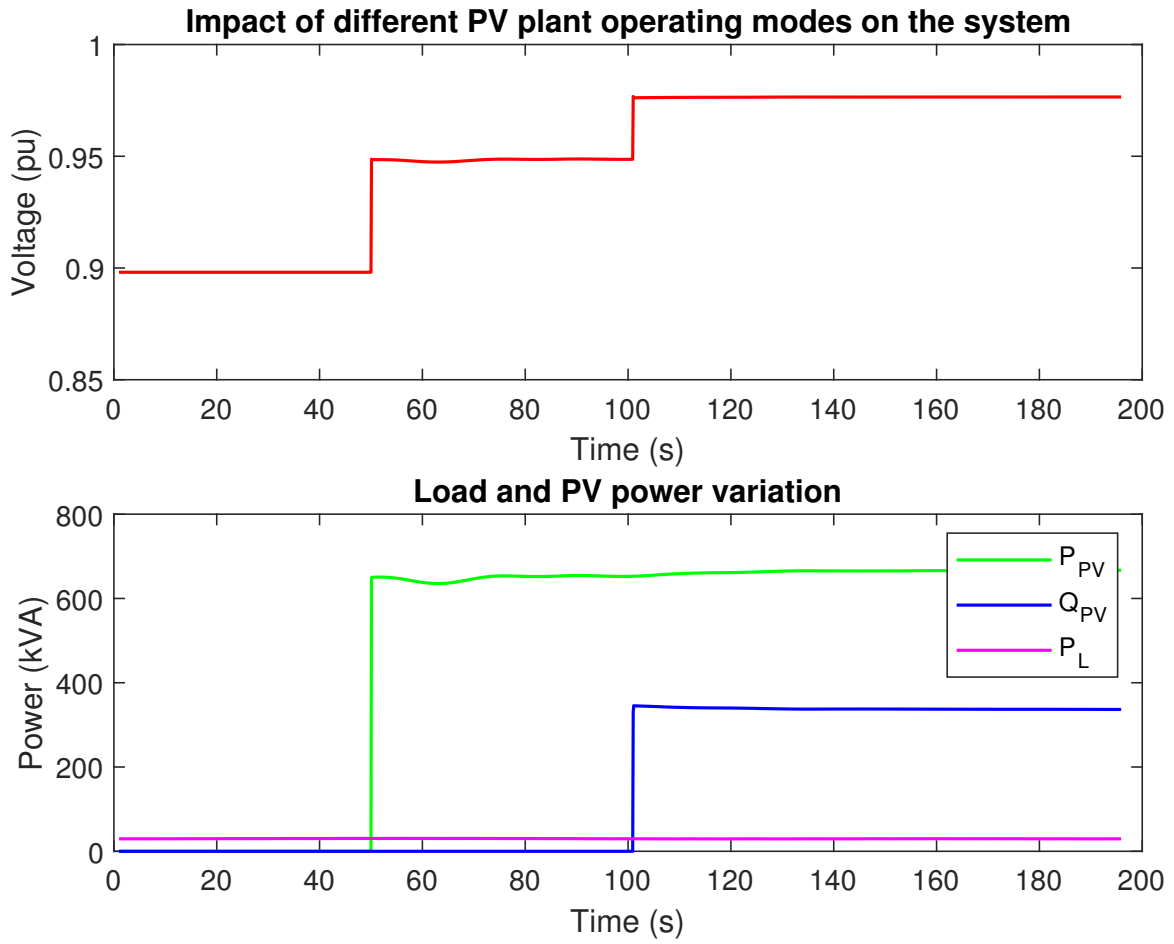


Figure 5.11: The real-time simulation output comparison for a 200 second simulation (4.24pm, 5/2/2022) where the IoT changes the operation mode from no PV, PV, and PV with vvc. The top plot shows the variation the PV node voltage and the bottom plot shows the corresponding variation in load and PV plant power.



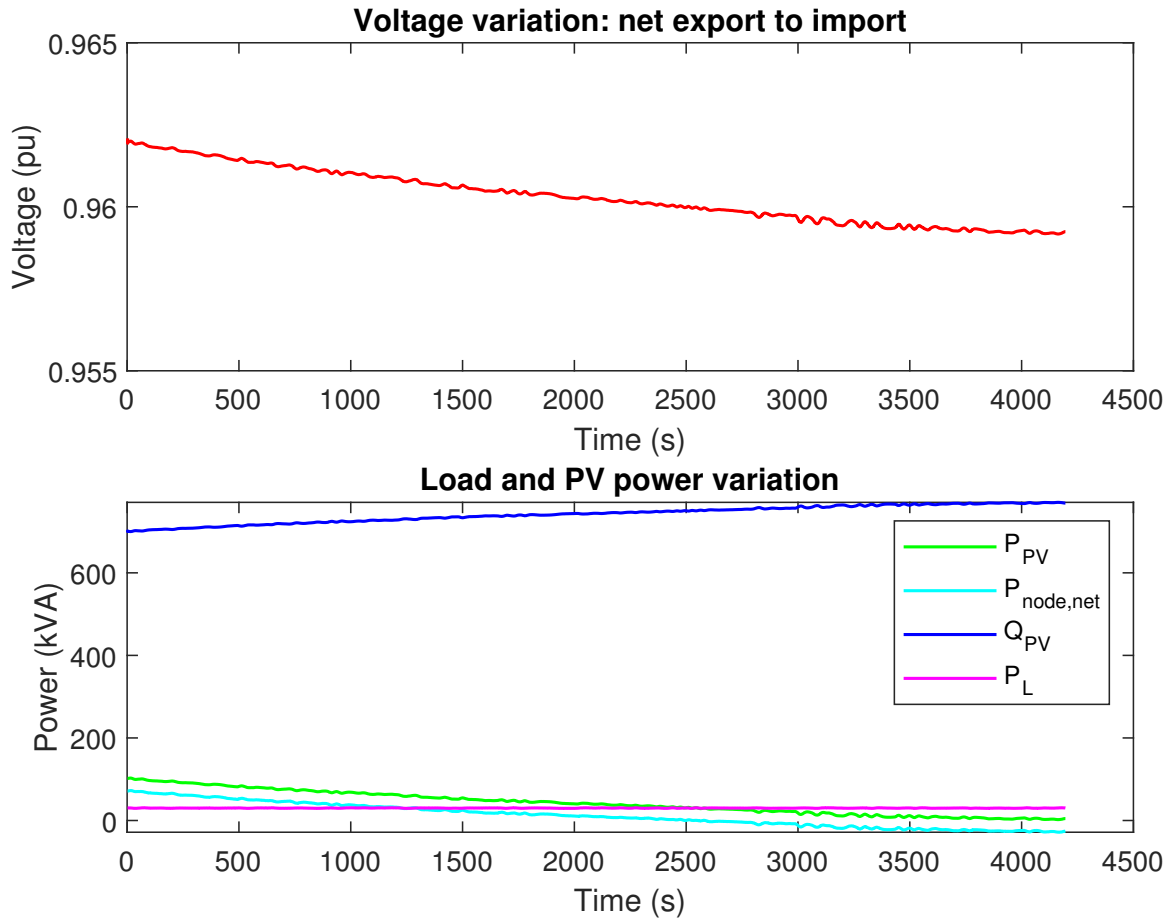


Figure 5.12: The real-time simulation output comparison from 7 pm to 8 pm on 5/2/2022. Here the system is operated with standard vvc. The top plot shows the variation the PV node voltage and the bottom plot shows the corresponding variation in load, PV plant active and reactive power, as well as the net power import.

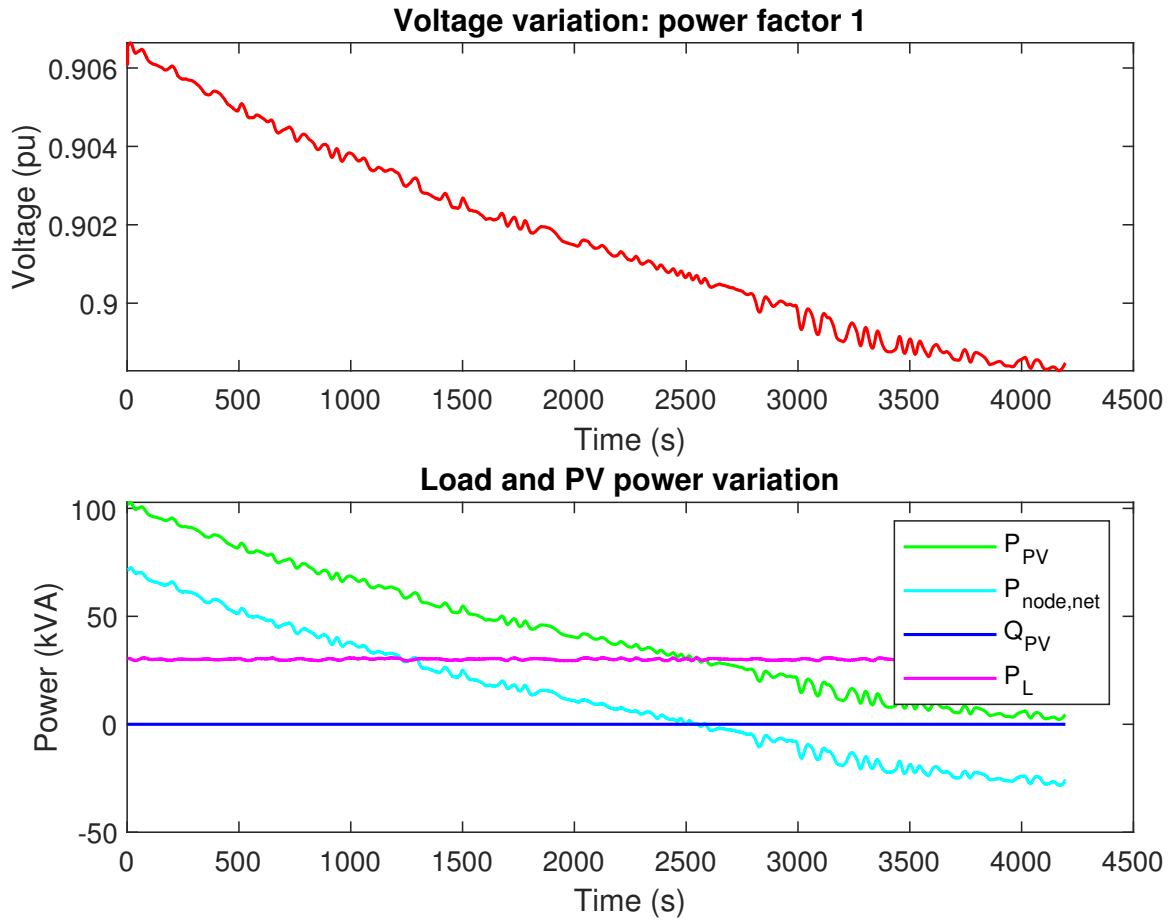


Figure 5.13: The real-time simulation output comparison from 7 pm to 8 pm on 5/2/2022. Here the system is operated with a fixed power factor of one. The top plot shows the variation in the PV node voltage, and the bottom plot shows the corresponding variation in load, PV plant active and reactive power, and the net power import.

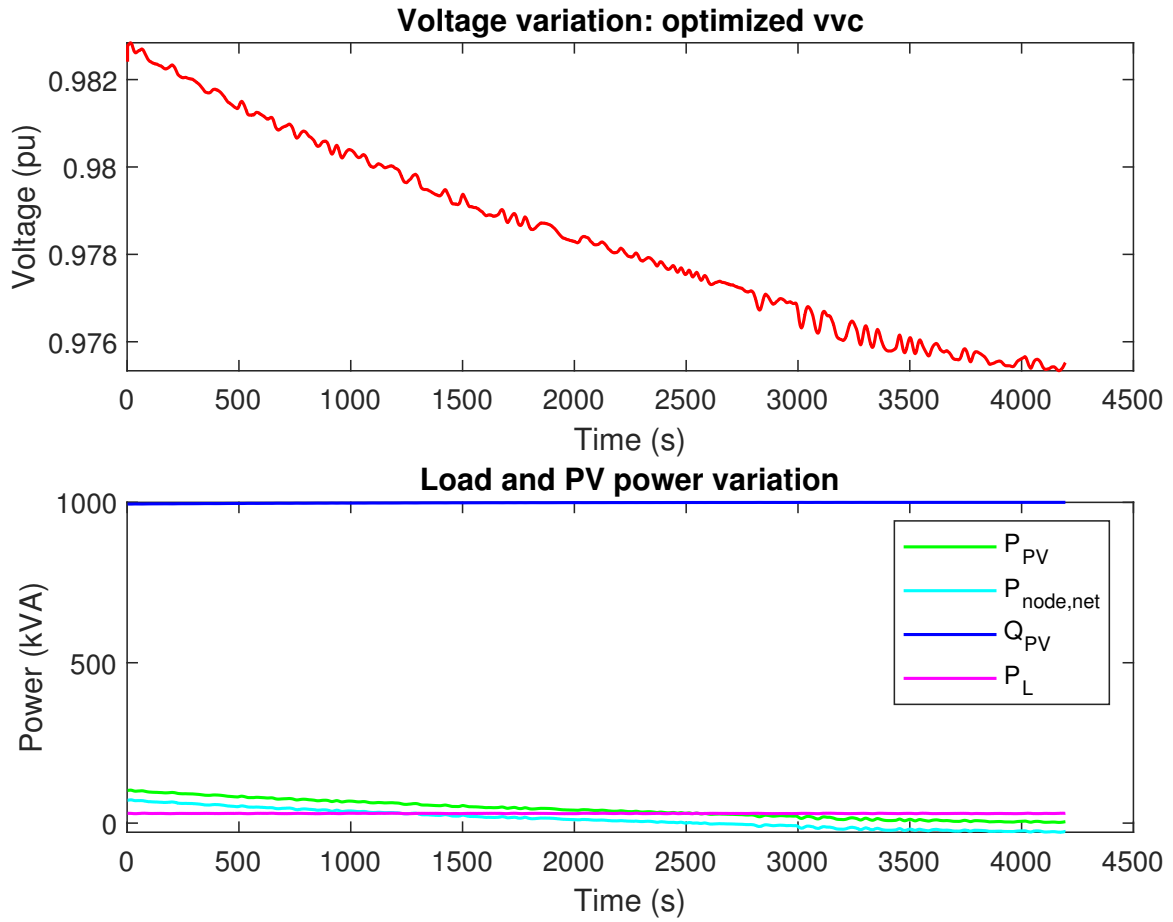


Figure 5.14: The real-time simulation output comparison from 7 pm to 8 pm on 5/2/2022. Here the system is operated with optimized vvc. The top plot shows the variation the PV node voltage and the bottom plot shows the corresponding variation in load, PV plant active and reactive power, as well as the net power import.

## Chapter 6

# Conclusions

### 6.1 Introduction

The exponential growth in the integration of active and dynamic technologies such as Distributed Energy Resources and grid-edge technologies makes it very challenging to operate the distribution system reliably while ensuring cost-effectiveness and the quality of service to the prosumers.

The power distribution grid is disorderly in design and implementation, chaotic in operation, large in scale, and complex in every way possible. Therefore, modeling, operating, and controlling the distribution system is incredibly challenging. It is required to find solutions to the multitude of challenges facing the distribution grid to transition towards a just and sustainable energy future for our society. The key to addressing distribution system challenges lies in unlocking the full potential of the distribution grid. The work in this dissertation is focused on finding solutions to a few of the challenges mentioned above.

### 6.2 Section summaries

The three main sections presented in this dissertation are distributed modeling of power distribution systems, distributed optimization of controls of power distribution systems, the development of a real-time testbed, and the verification of real time optimization by using the developed hardware, data, and optimization in-the-loop real-time simulation.

### 6.2.1 Distribution system modeling

This dissertation develops a new data-driven distributed framework to model power distribution systems. Its performance is validated on an IEEE test case. The results show a significant enhancement in accuracy and performance compared to the state-of-the-art centralized modeling approach. The modeling framework allows for distributed modeling of distribution systems using available user data in a scalable manner while ensuring the privacy and security of the data.

### 6.2.2 Distribution system volt-var curve optimization

This dissertation presents a framework to optimize multiple volt-var curves of DERs in a power distribution system with distributed cell-based optimization. The framework is validated for voltage control on an IEEE test feeder. The results show that the system has improved performance compared to the state-of-the-art approach. The optimization framework for distributed online optimization of the voltage controls allows for robust and enhanced operation of the distribution grid. This approach increases the quality of service and allows for increasing the capacity to integrate distributed energy resources into the distribution system.

### 6.2.3 Distribution system real-time testbed

The dissertation presents the design for real-time testbed that can be then implemented in the RTPIS lab. A new data-in-the-loop (DIL) method is developed and implemented to integrate actual loads and distributed generators with the simulator. The testbed is used to validate the DIL simulation capabilities and active and reactive power capabilities of a PV plant. The real-time integration assets and sub systems include communication, controls, data, and hardware-in-the-loop components.

## 6.3 Future work

There are several ways to expand the contributions of the presented  $D^3M$  framework. One option is the implementation of the  $D^3M$  in more extensive and varied distribution systems. The presented DOF is illustrated by implementing an IEEE test case representing a medium-sized rural distribution network. However, the implementation can be expanded to include larger and smaller

networks with respect to the capacity, as well as distribution grids with different characteristics, including system design and types of consumers. Another possible expansion is to include the cyber-physical components of the implementation. This allows for understanding and integrating, and addressing that relates to communication and computation infrastructure at a more granular level and helps bring the idea closer to actual implementation in the field.

The presented work can also be expanded to address the fundamental challenges of the design of the cellular computational network. This portion of the future work can be looking at the optimal design for a given system and set of hardware and software resources. The other fundamental challenge is related to increasing computing efficiency. One key advantage of the cellular computational network-based method is that it allows for parallelization. By fully utilizing this capability of cellular computational networks, it will be possible to enhance the computational efficiency of distributed modeling by parallelizing the computation algorithm.

The future work to expand the presented DOF for volt-var curve optimization includes numerous possibilities. The first possibility is to expand the implementation to more extensive and more varied distribution systems. The second possibility is to optimize for various utility functions of interest, such as network losses, fairness for access, voltage variations, or combination. The third possibility is expanding the type of active components that are included in the study since only PV is included in the presented work. Examples of potential options that can be integrated include storage, D-statcom, and grid edge devices. Even though the optimization is presented for volt-var curve optimization, the framework is designed to be generic, and any local parameters that are allowed to be optimized have the potential to be optimized based on the methodology developed in this dissertation. Validating the optimization of different control parameters and ancillary services is another potential option for future work.

There are numerous possibilities to add to the work conducted on designing and developing a power distribution system real-time testbed. The system can be expanded by adding power-in-the-loop components to the simulator, such as actual PV inverters, inverter controllers, and energy storage. The IoT infrastructure can be extended from one PC to multiple PCs and other assets in industrial systems such as cyber-security, data management, communication (Modbus and OPC server), and various types of state-of-the-art embedded hardware. The IoT computational devices can be expanded to include state-of-the-art embedded computing platforms such as Arduino and Raspberry-pi, allowing for decreased cost, space and power requirements, and increased scalability.

The realistic nature of the simulation can be enhanced by adding more DIL components into the simulation. Additionally, the real-time optimization algorithm can be expanded by increasing the number of active components.

## **6.4 Summary**

A comprehensive framework for enhancing the capabilities of the power distribution system to absorb a high level of distributed energy resources while operating at a high level of quality is presented in this dissertation. The proposed approach is verified using both offline and online simulation. The online simulation is executed on a new real-time power distribution system testbed. The presented work contributes to expanding the ability for data driven modeling and optimization of power distribution systems in a distributed and scalable manner.

# Bibliography

- [1] I. E. Agency, “Data and statistics,” 2020.
- [2] N. Roy and H. R. Pota, “Current status and issues of concern for the integration of distributed generation into electricity networks,” *IEEE Systems journal*, vol. 9, no. 3, pp. 933–944, 2015.
- [3] M. Fowlie, “Will investing big in distributed solar save us billions?,” Energy Institute Blog, Aug. 2021.
- [4] R. Secretariat, “Renewables 2021 global status report,” tech. rep., 2021.
- [5] E. D. Cartwright, “Earth day sets stage for new us climate goals,” *Climate and Energy*, vol. 37, no. 11, pp. 19–20, 2021.
- [6] H. Pollitt, “Analysis: Going carbon neutral by 2060 ‘will make china richer’,” *Carbon Brief*, vol. 24, 2020.
- [7] N. N. Taleb, *Antifragile: Things that gain from disorder*, vol. 3. Random House Incorporated, 2012.
- [8] R. Moghe, D. Tholomier, D. Divan, J. Schatz, and D. Lewis, “Grid edge control: A new approach for volt-var optimization,” in *Transmission and Distribution Conference and Exposition (T&D), 2016 IEEE/PES*, pp. 1–5, IEEE, 2016.
- [9] J. D. Jenkins, M. Luke, and S. Thernstrom, “Getting to zero carbon emissions in the electric power sector,” *Joule*, vol. 2, no. 12, pp. 2498–2510, 2018.
- [10] R. Klump, R. E. Wilson, and K. E. Martin, “Visualizing real-time security threats using hybrid scada/pmu measurement displays,” in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, pp. 55c–55c, IEEE, 2005.
- [11] C. L. Crago, “Economics of solar power,” in *Oxford Research Encyclopedia of Environmental Science*, 2021.
- [12] X. Miao and M. D. Ilić, “High quality of service in future electrical energy systems: A new time-domain approach,” *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 1196–1205, 2020.
- [13] H. Dharmawardena and G. K. Venayagamoorthy, “A distributed data-driven modelling framework for power flow estimation in power distribution systems,” *IET Energy Systems Integration*, vol. 3, no. 3, pp. 367–379, 2021.
- [14] H. Dharmawardena and G. K. Venayagamoorthy, “Distributed volt-var curve optimization using a cellular computational network representation of an electric power distribution system,” (*Under Review*) *Energies*, 2022.



- [15] M. Asano, F. Wong, R. Ueda, R. Moghe, K. Rahimi, H. Chun, and D. Tholomier, "On the interplay between svcs and smart inverters for managing voltage on distribution networks," pp. 1–5, IEEE, 2019.
- [16] R. Neal, "The use of ami meters and solar pv inverters in an advanced volt/var control system on a distribution circuit," pp. 1–4, IEEE, 2010.
- [17] "Ieee standard for interconnecting distributed resources with electric power systems - amendment 1," May 2014.
- [18] R. C. Dugan, "Reference guide: The open distribution system simulator (opendss)," *Electric Power Research Institute, Inc*, vol. 7, 2012.
- [19] J. Smith, W. Sunderman, R. Dugan, and B. Seal, "Smart inverter volt/var control functions for high penetration of pv on distribution systems," in *Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES*, pp. 1–6, IEEE, 2011.
- [20] M. Farivar, L. Chen, and S. Low, "Equilibrium and dynamics of local voltage control in distribution systems," in *52nd IEEE Conference on Decision and Control*, pp. 4329–4334, IEEE, 2013.
- [21] T. Sansawatt, L. F. Ochoa, and G. P. Harrison, "Integrating distributed generation using decentralised voltage regulation," in *IEEE PES General Meeting*, pp. 1–6.
- [22] A. Dubey, A. Bose, M. Liu, and L. N. Ochoa, "Paving the way for advanced distribution management systems applications: Making the most of models and data," vol. 18, pp. 63–75, 2020.
- [23] S. H. Low, "Convex relaxation of optimal power flow—part i: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.
- [24] W. Sheng, K.-y. Liu, and S. Cheng, "Optimal power flow algorithm and analysis in distribution system considering distributed generation," *IET Generation, Transmission & Distribution*, vol. 8, no. 2, pp. 261–272, 2014.
- [25] B. A. Robbins, C. N. Hadjicostis, and A. D. Domínguez-García, "A two-stage distributed architecture for voltage control in power distribution systems," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1470–1482, 2012.
- [26] T. T. Hashim, A. Mohamed, and H. Shareef, "A review on voltage control methods for active distribution networks," *Prz. Elektrotech*, vol. 88, pp. 304–312, 2012.
- [27] K. Petrou, M. Z. Liu, A. T. Procopiou, L. F. Ochoa, J. Theunissen, and J. Harding, "Operating envelopes for prosumers in lv networks: A weighted proportional fairness approach," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pp. 579–583, IEEE, 2020.
- [28] A. Hauswirth, A. Zanardi, S. Bolognani, F. Dörfler, and G. Hug, "Online optimization in closed loop on the power flow manifold," in *Proc. IEEE Manchester PowerTech*, pp. 1–6, June 2017.
- [29] S. Magnússon, C. Fischione, and N. Li, "Voltage control using limited communication," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1–6, 2017.
- [30] J. Annoni, E. Dall’Anese, M. Hong, and C. J. Bay, "Efficient distributed optimization of wind farms using proximal primal-dual algorithms," in *2019 American Control Conference (ACC)*, pp. 4173–4178, IEEE, 2019.

- [31] G. C. Kryonidis, K.-N. D. Malamaki, S. I. Gkavanoudis, K. O. Oureilidis, E. O. Kontis, J. M. Mauricio, J. M. Maza-Ortega, and C. S. Demoulias, "Distributed reactive power control scheme for the voltage regulation of unbalanced lv grids," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 1301–1310, 2020.
- [32] A. Maknouninejad and Z. Qu, "Realizing unified microgrid voltage profile and loss minimization: A cooperative distributed optimization and control approach," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1621–1630, 2014.
- [33] K. Turitsyn, P. Šulc, S. Backhaus, and M. Chertkov, "Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration," in *IEEE PES general meeting*, pp. 1–6, IEEE, 2010.
- [34] F. Malandra, R. Pourramezan, H. Karimi, and B. Sans<sup>2</sup>, "Impact of pmu and smart meter applications on the performance of lte-based smart city communications," pp. 1–6, IEEE, 2018.
- [35] A. Faruqui, S. Sergici, and L. Lam, "Bridging the chasm between pilots and full-scale deployment of time-of-use rates," *The Electricity Journal*, vol. 33, no. 10, p. 106857, 2020.
- [36] R. G. Pratt, P. J. Balducci, C. Gerkenmeyer, S. Katipamula, M. C. Kintner-Meyer, T. F. Sanquist, K. P. Schneider, and T. J. Secrest, "The smart grid: an estimation of the energy and co2 benefits," tech. rep., Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2010.
- [37] S. M. Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century," *IEEE power and energy magazine*, vol. 3, no. 5, pp. 34–41, 2005.
- [38] G. K. Venayagamoorthy, "Potentials and promises of computational intelligence for smart grids," in *2009 IEEE Power & Energy Society General Meeting*, pp. 1–6, IEEE, 2009.
- [39] G. K. Venayagamoorthy, "Dynamic, stochastic, computational, and scalable technologies for smart grids," *IEEE Computational Intelligence Magazine*, vol. 6, no. 3, pp. 22–35, 2011.
- [40] H. Dharmawardena, A. Arzani, and G. K. Venayagamoorthy, "Distributed voltage control for distribution feeder with photovoltaic systems," in *2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1–8, IEEE, 2018.
- [41] M. Farajollahi, A. Shahsavari, E. Stewart, and H. Mohsenian-Rad, "Locating the source of events in power distribution systems using micro-pmu data," pp. 1–1, IEEE, 2019.
- [42] R. Moghe, M. J. Mousavi, J. Stoupis, and J. McGowan, "Field investigation and analysis of incipient faults leading to a catastrophic failure in an underground distribution feeder," in *2009 IEEE/PES Power Systems Conference and Exposition*, pp. 1–6, IEEE, 2009.
- [43] T. L. Williams, Y. Sun, and K. Schneider, "Off-line tracking of series parameters in distribution systems using ami data," *Electric Power Systems Research*, vol. 134, pp. 205–212, 2016.
- [44] A. S. Meliopoulos, E. Polymeneas, Z. Tan, R. Huang, and D. Zhao, "Advanced distribution management system," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2109–2117, 2013.
- [45] I.-K. Song, S.-Y. Yun, S.-C. Kwon, and N.-H. Kwak, "Design of smart distribution management system for obtaining real-time security analysis and predictive operation in korea," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 375–382, 2012.
- [46] X. Lei, Z. Yang, J. Yu, J. Zhao, Q. Gao, and H. Yu, "Data-driven optimal power flow: A physics-informed machine learning approach," *IEEE Transactions on Power Systems*, vol. 36, no. 1, pp. 346–354, 2021.

- [47] S. Aghaei Hashjin, S. Pang, E. H. Milianni, K. Ait-Abderrahim, and B. Nahid-Mobarakeh, "Data-driven model-free adaptive current control of a wound rotor synchronous machine drive system," *IEEE Transactions on Transportation Electrification*, vol. 6, no. 3, pp. 1146–1156, 2020.
- [48] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1964–1971, 2015.
- [49] N. Duan and E. M. Stewart, "Frequency event categorization in power distribution systems using micro pmu measurements," *IEEE Transactions on Smart Grid*, 2020.
- [50] T. Zhou, B. Li, C. Wu, Y. Tan, L. Mao, and W. Wu, "Studies on big data mining techniques in wildfire prevention for power system," in *2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, pp. 866–871, IEEE, 2019.
- [51] F. Liu, G. Tang, Y. Li, Z. Cai, X. Zhang, and T. Zhou, "A survey on edge computing systems and tools," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1537–1562, 2019.
- [52] P. Schrangl, P. Tkachenko, and L. del Re, "Iterative model identification of nonlinear systems of unknown structure: Systematic data-based modeling utilizing design of experiments," vol. 40, pp. 26–48, 2020.
- [53] J. Anderson, F. Zhou, and S. H. Low, "Disaggregation for networked power systems," in *2018 Power Systems Computation Conference (PSCC)*, pp. 1–7, 2018.
- [54] D. Montenegro, G. A. Ramos, and S. Bacha, "A-diakoptics for the multicore sequential-time simulation of microgrids within large distribution systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 3, pp. 1211–1219, 2015.
- [55] P. Schrangl, P. Tkachenko, and L. del Re, "Iterative model identification of nonlinear systems of unknown structure: Systematic data-based modeling utilizing design of experiments," *IEEE Control Systems Magazine*, vol. 40, no. 3, pp. 26–48, 2020.
- [56] T. Gonen, *Electric power distribution engineering*. CRC press, 2015.
- [57] J. J. Grainger and W. D. Stevenson, *Power system analysis*. McGraw-Hill, 1994.
- [58] B. Luitel and G. K. Venayagamoorthy, "Decentralized asynchronous learning in cellular neural networks," vol. 23, pp. 1755–1766, 2012.
- [59] L. L. Grant and G. K. Venayagamoorthy, "Cellular multilayer perceptron for prediction of voltages in a power system," (Curitiba, Brazil), pp. 1–6, IEEE, 2009.
- [60] Y. Wei and G. K. Venayagamoorthy, "Cellular computational generalized neuron network for frequency situational intelligence in a multi-machine power system," *Neural Networks*, vol. 93, pp. 21–35, 2017.
- [61] L. Wu, G. K. Venayagamoorthy, and J. Gao, "Cellular computational networks for distributed prediction of active power flow in power systems under contingency," in *2019 IEEE PES Innovative Smart Grid Technologies Europe, ISGT-Europe 2019, Bucharest, Romania, September 29 - October 2, 2019*, pp. 1–5, IEEE, 2019.
- [62] M. A. Rahman and G. K. Venayagamoorthy, "A hybrid method for power system state estimation using cellular computational network," *Eng. Appl. Artif. Intell.*, vol. 64, pp. 140–151, 2017.

- [63] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, pp. 61–80, 2009.
- [64] A. Juditsky, H. Hjalmarsson, A. Benveniste, B. Delyon, L. Ljung, J. Sjöberg, and Q. Zhang, "Nonlinear black-box models in system identification: Mathematical foundations," *Automatica*, vol. 31, no. 12, pp. 1725–1750, 1995.
- [65] M. H. Stone, "Applications of the theory of boolean rings to general topology," *Transactions of the American Mathematical Society*, vol. 2, pp. 215–228, Nov. 1937.
- [66] J. A. Suykens, J. P. Vandewalle, and B. L. de Moor, *Artificial neural networks for modelling and control of non-linear systems*. Springer Science & Business Media, 1995.
- [67] Y. Makovoz, "Uniform approximation by neural networks," *Journal of Approximation Theory*, vol. 41, no. 3, pp. 375–481, 1998.
- [68] K. Hornik, M. Stinchcombe, H. White, *et al.*, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [69] L. Ljung *et al.*, "Theory for the user," in *System Identification*, Prentice-hall, Inc., 1987.
- [70] Y. Del Valle, G. K. Venayagamoorthy, S. Mohagheghi, J.-C. Hernandez, and R. G. Harley, "Particle swarm optimization: basic concepts, variants and applications in power systems," *IEEE Transactions on evolutionary computation*, vol. 12, no. 2, pp. 171–195, 2008.
- [71] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented road map," *IEEE Control Systems Magazine*, vol. 39, no. 6, pp. 28–99, 2019.
- [72] H. Dharmawardena and G. K. Venayagamoorthy, "A distribution system test feeder for der integration studies," in *2018 Clemson University Power Systems Conference (PSC)*, pp. 1–8, IEEE, 2018.
- [73] IEEE, "IEEE guide for investigating and analyzing power cable, joint, and termination failures on systems rated 5 kv through 46 kv," *IEEE Std 1511-2005*, pp. 1–21, 2005.
- [74] A. Vaughan, "The first global energy crisis," *New Scientist*, vol. 253, no. 3379, pp. 18–21, 2022.
- [75] F. Sioshansi, *Consumer, prosumer, prosumer: How service innovations will disrupt the utility business model*. 2019.
- [76] R. E. Brown, *Electric power distribution reliability, second edition*. 2017.
- [77] N. Rosetto and V. Reif, "Digitalization of the electricity infrastructure: a key enabler for the decarbonization and decentralization of the power sector," in *A Modern Guide to the Digitalization of Infrastructure*, Edward Elgar Publishing, 2021.
- [78] N. K. Roy and H. R. Pota, "Current status and issues of concern for the integration of distributed generation into electricity networks," *IEEE Systems journal*, vol. 9, no. 3, pp. 933–944, 2015.
- [79] K. Turitsyn, P. Šulc, S. Backhaus, and M. Chertkov, "Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration," in *IEEE PES general meeting*, pp. 1–6, 2010.
- [80] F. Katiraei and J. R. Agüero, "Solar PV integration challenges," *IEEE power and energy magazine*, vol. 9, no. 3, pp. 62–71, 2011.

- [81] P. V. Kishore, S. R. Reddy, and P. V. Kishore, "Modeling and simulation of 14 bus system with D-STATCOM for powerquality improvement," *Indian Journal of Scientific Research*, pp. 73–80, 2012.
- [82] R. Moghe, D. Tholomier, D. Divan, J. Schatz, and D. Lewis, "Grid Edge Control: A new approach for volt-var optimization," in *Transmission and Distribution Conference and Exposition (T&D), 2016 IEEE/PES*, pp. 1–5, 2016.
- [83] M. H. Haque, "Compensation of distribution system voltage sag by DVR and D-STATCOM," in *Power Tech Proceedings, 2001 IEEE Porto*, vol. 1, p. 5–pp, 2001.
- [84] E. Karunarathne, A. Wijethunge, and J. Ekanayake, "Enhancing PV hosting capacity using voltage control and employing dynamic line rating," *Energies*, vol. 15, no. 1, p. 134, 2021.
- [85] H. Dharmawardena, A. Arzani, and G. K. Venayagamoorthy, "Distributed voltage control for distribution feeder with photovoltaic systems," in *2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1–8, 2018.
- [86] L. Gutierrez-Lagos and L. F. Ochoa, "OPF-based CVR operation in PV-rich MV–LV distribution networks," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2778–2789, 2019.
- [87] M. Z. Liu, L. F. Ochoa, and S. H. Low, "On the implementation of OPF-based setpoints for active distribution networks," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 2929–2940, 2021.
- [88] M. N. Acosta, F. Gonzalez-Longatt, M. A. Andrade, J. L. R. Torres, and H. R. Chamorro, "Assessment of Daily Cost of Reactive Power Procurement by Smart Inverters," *Energies*, vol. 14, no. 16, 2021.
- [89] H.-J. Lee, K.-H. Yoon, J.-W. Shin, J.-C. Kim, and S.-M. Cho, "Optimal parameters of volt–var function in smart inverters for improving system performance," *Energies*, vol. 13, no. 9, p. 2294, 2020.
- [90] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov, "Options for control of reactive power by distributed photovoltaic generators," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 1063–1073, 2011.
- [91] D. G. Photovoltaics and E. Storage, "IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std*, pp. 1547–2018, 2018.
- [92] A. Maknouninejad and Z. Qu, "Realizing unified microgrid voltage profile and loss minimization: A cooperative distributed optimization and control approach," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1621–1630, 2014.
- [93] S. Magnússon, G. Qu, and N. Li, "Distributed optimal voltage control with asynchronous and delayed communication," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3469–3482, 2020.
- [94] N. Li, G. Qu, and M. Dahleh, "Real-time decentralized voltage control in distribution networks," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 582–588, 2014.
- [95] L. L. Grant and G. K. Venayagamoorthy, "Cellular Multilayer Perceptron for Prediction of Voltages in a Power System," in *2009 15th International Conference on Intelligent System Applications to Power Systems*, (Curitiba, Brazil), pp. 1–6, IEEE, 2009.

- [96] B. Luitel and G. K. Venayagamoorthy, “Decentralized Asynchronous Learning in Cellular Neural Networks,” vol. 23, no. 11, pp. 1755–1766, 2012.
- [97] R. Singh, G. F. Alapatt, and A. Lakhtakia, “Making solar cells a reality in every home: Opportunities and challenges for photovoltaic device design,” *IEEE Journal of the Electron Devices Society*, vol. 1, no. 6, pp. 129–144, 2013.
- [98] O. Parson, G. Fisher, A. Hersey, N. Batra, J. Kelly, A. Singh, W. Knottenbelt, and A. Rogers, “Dataport and NILMTK: A building data set designed for non-intrusive load monitoring,” pp. 210–214, IEEE, 2015.
- [99] Y. Del Valle, G. K. Venayagamoorthy, S. Mohagheghi, J.-C. Hernandez, and R. G. Harley, “Particle swarm optimization: basic concepts, variants and applications in power systems,” *IEEE Transactions on evolutionary computation*, vol. 12, no. 2, pp. 171–195, 2008.
- [100] O. Anaya-Lara and E. Acha, “Modeling and analysis of custom power systems by pscad/emt dc,” *IEEE transactions on power delivery*, vol. 17, no. 1, pp. 266–272, 2002.
- [101] P. Radatz, N. Kagan, C. Rocha, J. Smith, and R. C. Dugan, “Assessing maximum dg penetration levels in a real distribution feeder by using opendss,” in *2016 17th International Conference on Harmonics and Quality of Power (ICHQP)*, pp. 71–76, IEEE, 2016.
- [102] H. Doi, M. Goto, T. Kawai, S. Yokokawa, and T. Suzuki, “Advanced power system analogue simulator,” *IEEE Transactions on Power Systems*, vol. 5, no. 3, pp. 962–968, 1990.
- [103] R. Kuffel, J. Giesbrecht, T. Maguire, R. Wierckx, and P. McLaren, “Rtds-a fully digital power system simulator operating in real time,” in *Proceedings 1995 International Conference on Energy Management and Power Delivery EMPD’95*, vol. 2, pp. 498–503, IEEE, 1995.
- [104] D. Tiwari, M. Salunke, and A. Raju, “Modelling and real time simulation of microgrid using typhoon hil,” in *2021 6th International Conference for Convergence in Technology (I2CT)*, pp. 1–5, IEEE, 2021.
- [105] M. Milton, A. Benigni, and J. Bakos, “System-level, fpga-based, real-time simulation of ship power systems,” *IEEE Transactions on Energy Conversion*, vol. 32, no. 2, pp. 737–747, 2017.
- [106] Y. Yasuda, A. Yokoyama, Y. Tada, and H. Okamoto, “Study on low-cost and easy-use real-time digital power system simulator in matlab/simulink environment,” in *PowerCon 2000. 2000 International Conference on Power System Technology. Proceedings (Cat. No. 00EX409)*, vol. 2, pp. 921–926, IEEE, 2000.
- [107] E. Gómez-Luna, L. Palacios-Bocanegra, and J. E. Candelo-Becerra, “Real-time simulation with opal-rt technologies and applications for control and protection schemes in electrical networks,” *Journal of Engineering Science & Technology Review*, vol. 12, no. 3, 2019.
- [108] F. L. Grando, A. E. Lazzaretti, and M. Moreto, “The impact of pmu data precision and accuracy on event classification in distribution systems,” *IEEE Transactions on Smart Grid*, 2021.
- [109] G. P. Alliance, “openhistorian documentation.”
- [110] S. Venkatramulu, M. Phridviraj, C. Srinivas, and V. C. S. Rao, “Implementation of grafana as open source visualization and query processing platform for data scientists and researchers,” *Materials Today: Proceedings*, 2021.
- [111] Y. Rangelov, N. Nikolaev, and M. Ivanova, “The iec 61850 standard — communication networks and automation systems from an electrical engineering point of view,” in *2016 19th International Symposium on Electrical Apparatus and Technologies (SIELA)*, pp. 1–4, 2016.

# Biography

Hasala Dharmawardena (*S'08*) is a Ph.D. candidate in Electrical Engineering in the Holcombe Department of Electrical and Computer Engineering, Clemson University, Clemson, SC, USA. He received his BSc degree with first-class honors in Electrical Engineering from the University of Moratuwa, Sri Lanka, in 2010 and an MSc degree in Electrical Power Engineering from the Norwegian University of Science and Technology, Norway, in 2015. After working as an electric power engineer in the industry, he joined the Real-Time Power and Intelligent Systems Laboratory as a Research Assistant in 2017. His research interests are electric power distribution system modeling, planning, design, optimization, and artificial intelligence applications.