IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 12, 2019, accepted April 3, 2019, date of publication April 16, 2019, date of current version April 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2911543

A Nano-Biased Energy Management Using Reinforced Learning Multi-Agent on Layered Coalition Model: Consumer Sovereignty

MUHAMMAD RAMADAN BIN MOHAMAD SAIFUDDIN[1], (Graduate Student Member, IEEE), THILLAINATHAN LOGENTHIRAN[2], (Senior Member, IEEE), R. T. NAAYAGI[1], (Senior Member, IEEE), AND WAI LOK WOO[3], (Senior Member, IEEE)

[1]School of Electrical and Electronics, Newcastle Research and Innovation Institute, Newcastle University in Singapore, Singapore 609607
[2]School of Engineering and Technology, University of Washington Tacoma, Tacoma, WA 98402, USA
[3]Department of Computer and Information Sciences, Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne NE1 8ST, U.K.

Corresponding author: Muhammad Ramadan Bin Mohamad Saifuddin (m.r.bin-mohamad-saifuddin@newcastle.ac.uk)

ABSTRACT Trends in energy management schema have advanced into legislating consumer-centered solutions due to inclination interests for personal owned distributed energy resources at the low-voltage level. Thence, this paper proposes a tailorable energy manager tool that empowers Prosumer(s) in a nanostructured distribution network to take sole precedence when prosuming optimal services to the energy system. It too acts as an aggregator that attests cooperative energy management processes amongst Prosumers to enhance demand-side responses and economics. The suggested nano-biased energy manager engages multi-agent network as the basis coordinator for peer-to-peer advocacy in a decentralized environment. The agents were then programmed with reinforcement and extreme learning machine intelligence on a layered coalition model to compute joint decision-making processes with constraint relaxation relaxed decision constraints and policies. The problem formulations assure engagement of energy management in the liberalized market is sustainable, reliable, and non-discriminated. Computational validations were analyzed using MATLAB and Java agent development framework on four aggregated Nanogrids representing the residential, commercial, and industrial building. Results have shown positive eco-strategic managerial avenues where cooperative assets scheduling and bidding-abled decorum were autonomously acquired. Reduced operating costs were gained from energy trading profit margin due to strategic use/sell of electricity based on real-time tariff and conferred incentive packages but constrained within the mandatory obligation to demand-side management. The subsidiary, the inauguration of meshed communication infrastructure has shown adequate monitoring and commanding resolutions for decentralized Agent(s) to function collaboratively.

INDEX TERMS Demand-side management, multi-agent systems, adaptive scheduling, hybrid power system, stochastic processes and nanostructured power grid.

NOMENCLATURE

DER	Distributed Renewable/Energy Resource
DSO	Distributed Network Operator
ELM	Extreme Learning Machine
EM	Energy Management
JADE	Java Agent DEvelopment
LCM	Layered Coalition Model
MAN	Multi-Agent Network
MG	Microgrid
NG	Nanogrid
SHSES	Small-scaled Hybrid Sustainable Energy Source
SOC	State of Charge (Battery)
STS	Short-Term Scheduler
TSO	Transmission Network Operator
WAICS	Wide-Area Information & Communication System

The associate editor coordinating the review of this manuscript and approving it for publication was Ramakrishnan Srinivasan.

I. INTRODUCTION

As an integral part of the recent liberalization in the electricity market and digitalization for decentralized EM infrastructure, scientific communities are constantly resynthesizing stackable-ecotechnological EM solutions that radically address Consumer's expectations when embracing sustainable energy integrations [1], [2]. Indeed, at global perspective, decentralized EM schemas have already become a pivotal topic drivers involving self-suffice or standalone energy legislations, balancing electricity market participations, strengthening demand-side responses, and annex cloud database for secure data transportations [3], [4]. Conjointly, innovations in aggregating separate EMs too have shown significant importances where multi-objective optimization using evolutionary algorithms were adopted to cipher comprehensive energy administrations between DSOs, Energy Retailers, and Prosumers uniquely [5], [6]. In this sense, it calls for new Prosumers-centred business model and socio-technical connection that can resolve decentralize complexity in both technical, economical, and social dimensions.

A. RELATED WORKS AND RESEARCH GAPS

Presently, research community in decentralized EM systems have focused much dependency on DSO administrations to procure optimal coordination of supply-demand balancing at distribution level. Core services such as demand-side management, scheduling and anticipating DER penetrations, and energy market operations are some typical avenues that only DSO-TSO have the competency to superintend. At this sense, DSOs are yet placed as an energy mediator between Prosumer and Retailer where EM solutions could be biased towards local optimization approaches.

Habib *et al.* [7] proposed a coordinated strategy for optimal rate of DER utilizations based on electricity market price guidance alongside with MAN hierarchical control framework. The paper highlights proposition in devising price guidance coordination strategy for energy storage scheduling and flexible load capacity response to improve Consumer's energy consumptions based on real-time electricity tariffs. Moreover, the authors formulated an evaluation indexing that can quantify EM merits between distributed and centralized control method based on different planning schemes. Diversely in [8], Almeida *et al.* addressed few recent innovation projects involving DSO-TSO interactions that highlight optimization functionalities for larger set of problem statements during energy trading between Prosumer(s). Thus, new business model was brought forth to render strategic EM schemes that serves as a generic standard tool in the DSO toolbox. Its purpose was to provide a joined solution that deals with line congestion management, energy balancing, use of market flexibility, real-time control and supervision, and network planning. Several learning pointers were noticed from the proposed test case studies in relations to EM optimization functionalities; i) the ability to expand its decision-making search space based on previous EM experiences, ii) Adaptable- and configurable-environment when EM participants starts to aggrandize. Succeedingly in [9], Saint-Pierre and Mancarella introduced a novel framework for active EM distribution system that uses a dual-horizon rolling scheduling model to obtain optimal power flow dynamics inflicted by DER penetrations. The proposed model offers real-time operation planning realistically for DSOs to optimize and schedule available generation resources against demand capacity, incorporating Nonlinear Programming schema to cipher uncertainty elements. It too facilitates local energy reserve planning for DSO-TSO to treat load balancing mismatches under the influence of DER uncertainties in distribution network at different time horizon.

From the above mentioned methodologies, it is trivial for sole EM avocations to revolve around a centralised policy maker which can be favorable for DSO-Retailer but detrimental on Prosumer(s) as energy trading democracy at low-voltage level is still limited. Prosumer(s) are seeking new DER-installed business models that allow full ownership in personalising use of electricity while having full participation privileges during energy trading and market operation, hoping to gain good rate of return on investments. However, such undertakings can propagate predicaments when participation of Prosumer starts to escalate; i) inducing unsighted operations for DSO in times of energy crises, ii) monopolism in the electricity market and possible obsoletion of energy player(s), iii) demand-side management will be overly complex due to DER penetration and electricity tariff, and iv) power system reliability due to intermittent intentional islanding operations and DER integrations.

B. CONTRIBUTIONS

This paper unveils realization to advance EM mediation to be brought closer to Consumers granting grassroot-based (bottom-up and bilateral) administrations, transacting individualistic in-house energy manager platform that can institute either idiosyncratic or interdependent assessments. The intention is to position DSO as a neutral energy market facilitator for trading transactions while authorizing building or residential owners as the principal energy manager of its own during demand-side response settlements. Therefore, reversed obligatory role is transcend accrediting Prosumer(s) at low-voltage level to constitute scheduling- and bidding-abled model that relies on joined decision-making processes involving resource availability, demand load profiles, and feed-in against electricity price tariff in real-time. Importantly, constitution of the proposed Prosumer-centric EM must be habitable for successive integration of NG models, alleviating competency complications towards DSO's decision-making process.

Essentially, the key component for the suggested nano-biased EM for Prosumer(s) and DSO when facilitating the energy market is peer-to-peer data management where resource allocation is efficient, maintains constant connectivity, and data are encrypted. Indeed, 'one-size-fits-all' approach for data management is impracticable thus, this paper introduces a meshed-type wireless wide-area communication system that hybridizes with multi-agent network where interoperability and autonomous coordination for different operating Standards can be achieved. The Agents were then programmed with reinforcement and extreme learning machine intelligences on a layered coalition model to evoke relaxed bidding-abled EM strategy that apt in prioritizing parent exigencies while masterminding other neighboring constraints. Significantly, these Agents were primarily tasked to regulate power flow transactions based on in-house engagements to curb operating costs while maximizing energy use competencies. The proposed methodology was then tested on the nanostructured distribution network comprising 4 Testbed systems (2 residential, 1 commercial, and 1 industrial building) using MATLAB and JADE framework for MAN deployment.

The paper is organized as follow; Chapter II proposes the Testbed system comprises of four aggregated NGs at distribution level and interpreting the concept of NG engineering. Chapter III establishes the deployment of MAN in layered coalition model where Agents are defined and classified with unique allotted roles. In addition, layout of the wireless WAICS is presented for corroborative communication linkage between NGs. Chapter IV proposes control proceedings for Prosumer-centric EM system that involves supervised reinforced and extreme machine learning algorithms on Agent(s) to bid and generate global tractable solutions based on energy trading policy constraints and optimal demand-side management. Moreover, the algorithm also previsions operation's uncertainties and quantifies required spinning reserve needed to sustain during unexpected downtime. Chapter V exhibits simulation results attained from selected case studies to view the impacts of egocentric (sole NG) and altruistic (cooperative NGs) EM proceedings during demand-side operations. Chapter VI investigates operating cost efficiency and addressing uncertainty in unit commitment problems between proposed against other published EM methodologies. Lastly, Chapter VII concludes the paper.

II. DESIGNING DISTRIBUTION NETWORK INTO NANOGRID PERSPECTIVE

NG conceptions have taken its precedence towards creating an ultimate solution for building's energy awareness. It embraces incremental adjustments rather than relocating to something fundamentally new when embracing DER integrations at respective Prosumers' electrical network. Such transition enables DSO to have comprehensive EM jurisdiction for individual energy participant when dealing with Prosumers' energy consumptions or contributions during market operations. Here, a single NG is confined within a building-scale electrical network (residential or commercial building) tagged with Prosumers' identity. Each NG represents a unique EM domain that governs demand-side management through strategic scheduling of all local assets (i.e. SHSES and Appliance Loads) while creating time-based avenues for electricity market participations.

To visualize NGs deployment practicability at low-voltage level, careful selection of industrial-based electronic devices constituting present consumer's electrical networks were

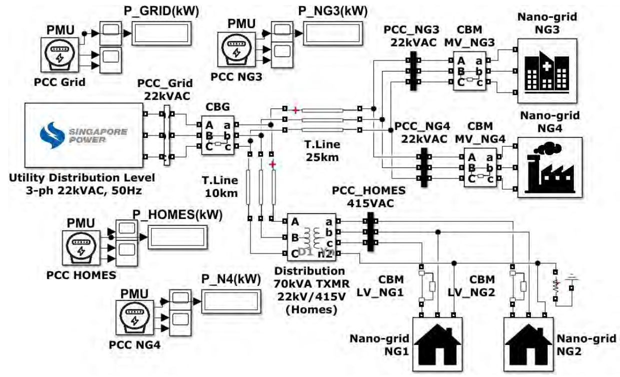


FIGURE 1. Proposed small-scaled distribution network.

TABLE 1. Power consumptions and generations for corresponding nanogrids.

Home	Appliances (24hrs)	EV Charging	PV (rooftop)	Battery storage	Other HSES
Rating (~kWh)	15.0-20.8 (Average in summer)	1.1-6.9 (Make Models & Milleage)	15.0-19.0 (Reliant on space)	10,48V (Install 2 sets)	Diesel Gen. Set
Cost (~$/kWh)	4.2-6.5	0.033/km	0.05-0.18	0.7-0.9	<0.15
Comm-ercial Building	Machine & Appliances (24Hrs)	EV (Employers)	PV (rooftop)	Battery Size Storage	Other HSES
Rating (~MWh)	0.214-0.336 (Average in summer)	0.21-0.30 (Min. 20 Charging Pt.)	0.23-0.28 (Reliant on space)	1,48V (Container based)	0.2-0.7
Cost(~$)	13.73-34.16	78-195	500k	800k	>1mil

modelled shown in Fig. 1. It consists of four peculiar NG systems involving two $25kV\ A$ Residential, a $2MV\ A$ Commercial (Hospital), and a $2MV\ A$ Industrial building denoted as $NG1$ to $NG4$ correspondingly. Using MATLAB, the mentioned NGs are integrated with relevant SHSES capacities and then coupled to the 3-ph $22kV\ AC$ primary-side of distribution network.

Detailed modelling of adopted SHSES in respective NG is exhibited in Fig. 15 of Appendix A. The installed capacity of SHSES from one NG to another may differ based on the space availability or prescribed demand load profiles, assuming that all NGs minimally inaugurated a single local power generation with energy storage system. Moreover, to better comprehend separate EM proceedings in relations to bidding strategies and energy trading, the deployed SHSES was deliberately designed with undersized capacity in contra to NG's base demand load ratings. Forcing NGs to revert it dependency on Energy Retailers and purchase electricity based on real-time tariffs. Table 1 presents the approximate data of local power usage and generation collected for typical residential and commercial building. Unfortunately, this paper does not incorporate deployment of back-up diesel generator into EM composition which can serve as a auxiliary during blackout crises. Instead, exploiting energy storage SOC as spinning reserve roles.

III. PROPOSED INFORMATION & COMMUNICATION FRAMEWORK AND DEPLOYMENT OF AGENTS

When performing decentralize EM, it is essential to extend wireless WAICS across the distribution network to establish two-way communication infrastructure for operation visibility. Likewise, connection interoperability between all monitoring devices must be synchronized under a standardized communication protocol in order to secure transmitting coordination of recorded data from one NG or appliance to another [13], [14]. Henceforth, deployment of smart metering and other remote control devices supported by WAICS can propagate close relationship synopses on on-line assets' status before undertaking desirable EM consequences.

Succeedingly, Agents of MAN are strategically positioned along respective NG to superintend parallel coordination of multiple on-line devices. These Agents are programmed to interact among themselves with certain degree of collaborative intelligence to perform joined (cooperating or competing) decision-making approach based on time-logged commands. Alternatively, due to Prosumer-centric EM domain, Agent's communication too can be confined within the NG boundary constituting uncooperative EM managerial with other NG. Nevertheless, parallel deployment of WAICS with role-assigned Agents under layered coalition model were proposed to comprehend complete EM dominance respective NG during operations.

A. WIDE-AREA WIRELESS INFORMATION & COMMUNICATION FOR SINGLE NANOGRID

To commemorate in-house or local point-to-point broadcasting service for sensory and automated actuation devices, Zigbee technology proffers smart network congestion supervision with high data rate transfer within low bandwidth spectrum. It uses a unique periodic logging of data (non-beacon and beacon mode) operations which results in low energy performance suitable for low-powered electric appliances [15], [16]. However, Zigbee solution only performs at its best for short range wireless transmitting/receiving mesh deployments targeted for in-house connection. Therefore, to establish continuous connectivity across neighboring NGs and other on-line energy player(s), advance broadband wireless access using WiMAX technology was employed to serve as a network extender for cooperative intercommunication. Together, both communication systems were integrated under a single wireless personal area network interoperated using IEEE 802.15.4 protocol platform to provide centralised data-centric communication management model. Fig. 2 presents deployment of Zigbee and WiMAX technologies across a single NG to monitor electrification signatures.

B. DEPLOYMENT AND CHARACTERISATION OF AGENTS IN LAYERED COALITION MODEL

Having WAICS being established, Agents of MAN were then parallelly tagged to respective communication devices under the hierarchical-based LCM. Agents will be programmed

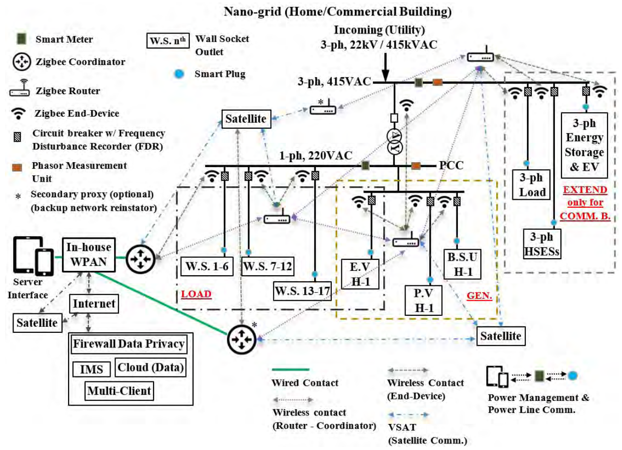


FIGURE 2. Layout of WAICS for each NG network.

with self-taught intelligences that apt in procuring non-linear decision-making processes under predefined operation commitments. Using reward- and penalty-driven learning functionality, Agents are to cipher joined decisional resolution based on high accumulative payoff from the decentralized search space while acknowledging priority constraints rendered by LCM. The functionality of LCM provides personalized layered arrangement in segregating problem areas into level of importance based on user's preferences. Fig. 3a demonstrates the deployment of Agents along the distribution network while Fig. 3b represents the proposed LCM arrangements in hierarchical order.

Constitution of Agents seen in Fig. 3a were modelled in JADE. JADE is an open source multi-agent software that establishes synergy between heterogeneous Agents in compliance to IEEE Foundation of Intelligent Physical Agents standards. It aims to cede interoperability and transact autonomous coordination with other online technologies given in their respective Standard domains [17]. It promotes software portability, instil security policies for data sharing, and endorse object oriented dataflow computing model to enhance Agent's operations when dealing with decisions. Using the inbuilt Agent Directory Facilitator, users can easily define and store Agent's associated service descriptions into Agent-to-Agent data exchange database.

1) AGENT'S CUSTOM ROLES

Following defines Agent's deployment roles involving its authorities, functionalities and communication relations:

- *Distribution Service Operator (DSO) Agent*- It schedules base and reserve power pooling in view of power delivery and exchange trends procured from aggregated NGs (i.e. export and import electricity capacity profiles, Duck Curve profile). It too acquaint energy price market and penalty for Consumers, serving as a price benchmarking during bidding process and poor EM respectively [18], [19].
- *Power Condition Monitoring (PCM) Agent*- Invigilate voltage, current and power level at secondary distribution network. PCM Agent interacts closely with DSO

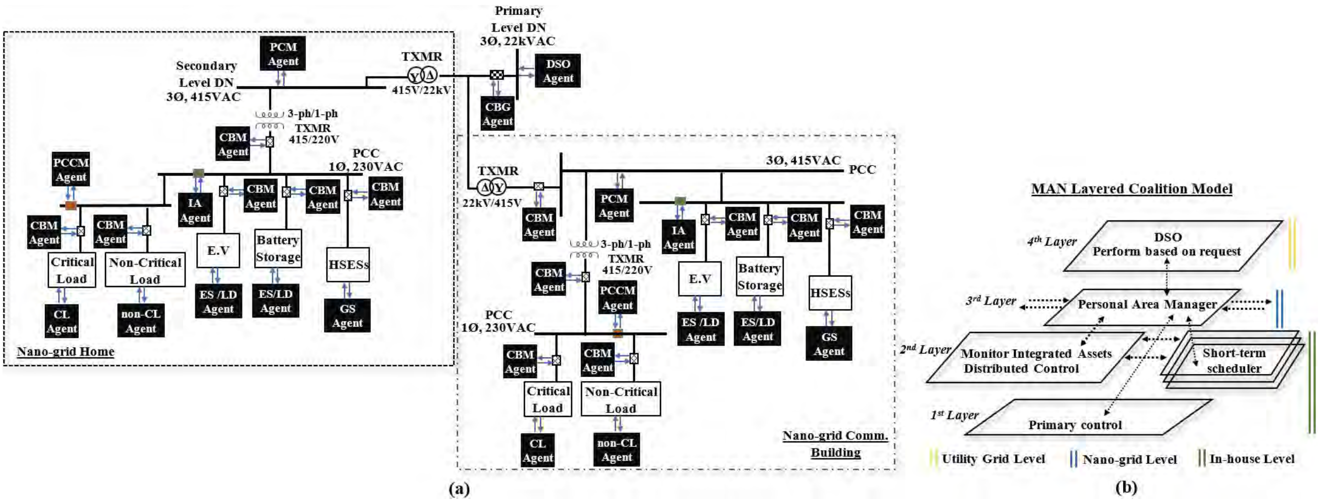


FIGURE 3. **(a) Agents of MAN deployed in NG network. (b) MAN cooperative layer model representing single NG.**

to coordinate power transactions between Utility and aggregated NGs, updating the status of the grid's quality and flagging possible threads.

- *Circuit Breaker Grid (CBG) & Monitor (CBM) Agents*- CBG Agent disengages NGs from Utility (Islanded mode) when it detects abnormality transpired from upstream while CBMs are only deployed at low-voltage level targeting a single NG or individual integrated devices to perform isolation operations.
- *Point of Common Coupling Monitor (PCCM) Agent*- Represents an identification for peculiar NG, officiating all power delivery transactions and monitoring system's stability. Program to interact with neighboring NGs to initiate bidding and negotiation strategies for better electrification policies to suit unique needs [20], [21].
- *Critical Load (CL) & non-Critical Load (non-CL) Agent*- Both Agents exemplify load entity, constantly attuning with PCCM Agent to govern grid balancing services. Diversely, for CL Agent, availability of incoming power must be fulfilled continually while non-CL allows load shedding avocations based on power availability.
- *Generation Source (GS) Agent*- Representing identification for respective integrated power generation resources (DC or AC-based technology).
- *Energy Storage/Load Device (ES/LD) Agent*- ES/LD Agent is a dual operated system that functions as a power generation or load (non-critical) interchangeable. It monitors state-of-charge level and instruct charging or discharging operations.
- *Integrated Assets (IA) Agent*- IA Agent helps to encapsulate and monitor all integrated assets (SHSES), comprehending power availability at individual integrated system while employing STS to evaluate operational status at different time intervals.

2) MAN HIERARCHY IN LAYERED COALITION MODEL

Fig. 3b illustrates the decisional task allocation at respective LCM's layers. Order of computational hierarchy is organised from bottom layer to top with the consideration of taking layer-to-layer compromising components before converging to an absolute solution [22]. Thus, the Agent's learning capacity will gain computational complexity as it advances higher into the hierarchy known as the ''n-1 coordination criterion'' effects. Such approach allows Agent's interactions to bring forth site-specific emergence of comprehensive system behaviour, opening up new possibilities for all energy participants:

- *1st Layer*- It provides primary load balancing across all online loads and monitor power consumptions of non-critical loads while ensuring critical loads are protected. Involving Agents required to address power mismatches and retrieve control operandi commands to perform switching operations based on generation availability and update power flow distributions within seconds across the network in real-time.
- *2nd Layer*- There are two governing sub-controllers, Distribution and STS. The Distribution control strategies involve in-house and Utility power flow delivery to schedule online load against generation availability with respect to electricity price market and local HSESs. The objective is to send managerial commands based on strategical planning in maximising local power production and maintaining enough reserve energy pool. To back it up, STS compliments in providing analytical information (i.e. day-ahead forecast, energy storage planning, etc.) that informs possible threads based on the estimated parameters. Measurements are recorded every second intervals.
- *3rd Layer*- The management proceedings render two unified mediator, creating a linkage between Utility

(Local Corresponder) or neighboring NGs (External Corresponder) with the local NG. The Local Corresponder imports power generation requests from Utility or vice versa. It too communicate with PCM Agent to create a bidding platform where diplomatic electricity price tariffs are tabulated in real-time. Such decisional transactions are mediated based on 30mins intervals, remapping its planning to be in sync with the wholesale electricity tariff. For External Corresponder, it focuses mainly in trading with other NGs as an alternative for power exchange and again a separate bidding proceedings will be brought upon.

- *4th Layer*- It responses closely to PCM Agent's commands to match required power delivery. To overcome oversized energy pooling from Utility, it studies demand load profiles and comprehend power mismatches trends. Such predicaments avoid Utility from issuing penalties to local NG for poor power management and improve Duck Curve crisis. Conjointly, it will broadcast wholesale electricity prices giving PCM Agents more options in purchasing their electricity either from Utility or the neighbouring NGs.

TABLE 2. Agents of MAN coordination and tasks.

1	Record load profile & notifies state changes.	2,2A	Execute power transfer, Store Utility electricity prices.
3, 3A	Display load capacity, Request for power & bid electricity cost.	4, 4A	Display surplus generation capacity, Offer electricity cost.
5	Accept/Reject Proposal.	6, 10	Fault detected status, Load Shedding.
13	Instruct ES/LD Agent to perform charging or discharging operation.	8	Update power impairity call for new assets management strategy.
9	Provide anticipated load capacities, status of coupled CBs.	11	Proposed CB switching operations and update power availability.
12	Request and update online load capacity.	7	Status of CB.
16	Request for power generation.	15	Display and forecasts local power generation.
14	Monitor ES/LD SOC level & suggest storage operations (charge/ discharge).	17	Propose load balance management (load shedding & reserve for storage).
18	Renewable power generation profile, generator's start-up time & storage SOC level.	19	Compile summary of available power generations taking battery storage & electricity prices into considerations to PCM Agent.

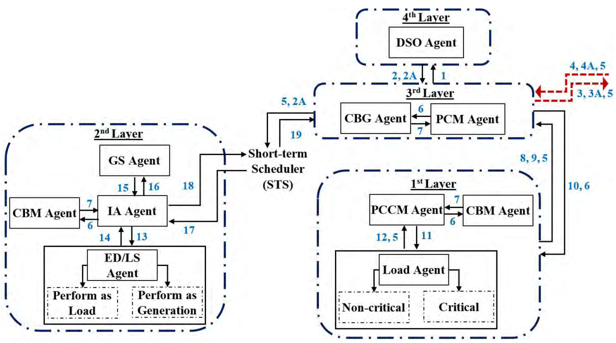


FIGURE 4. Agents of MAN in a collaborative network.

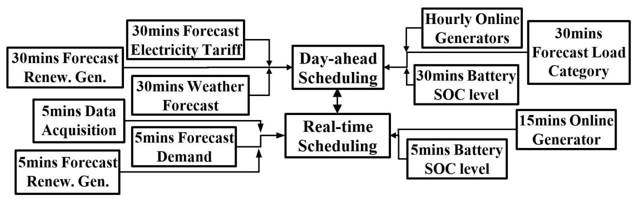


FIGURE 5. STS time-based allotment for appointed agents.

Supplementary, the LCM has a distributed cloud database for separate layers. It serves as a data access point where Agents can selectively share information across any hierarchy layer to ensure objectives are tracked and data are well managed.

IV. PROPOSED CONTROL AND MANAGEMENT INNOVATIONS

A. COORDINATION AND FUNCTIONALITY OF AGENT

A systematized codification of Agent's governance shown in Fig. 4 proposes a collaborative network that orientates dataflow interactions among Agents while Table 2 provides description of data transaction operations (exchange information) at respective Agent's roles. These messages contain user-centric directive commands that direct NG's operations based on Agent's commitments driven by data interventions. Through such collaborative assignments, the perceptive of Agent's obligatory is clearly visible when tendering ingenuity solutions. Agents in the collaborative network are constantly updating their individual findings to achieve compounding effects until all involving Agents reach convergence. All computational proceedings entail all online CL Agent's to concede prior settlement before advancing to other decision making processes.

Succeedingly, deployment of Agents for STS seen in Fig. 5 is divided into two-level scheduling approach; day-ahead and real-time scheduler. The day-ahead aids in forecasting short-term resolutions based on historical data while real-time performs ad-hoc corrections to atone day-ahead predictions. The scheduling architecture to ensure local power generation availability is sized without over- or under-fitting capacities.

Finally, appropriate allocation of Ontologies is designed to secure cooperative communication between Agents. The Ontologies provide Agents to better understand structured information that are exchanged during operations, typically packaged with Agent's ID and a target action [23]. Fig. 6 represents the proposed 'NanogridOntology' for participating Agents using JADE platform.

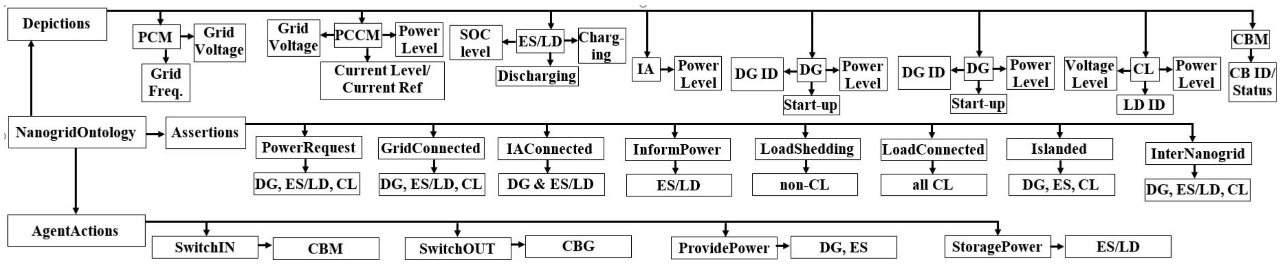


FIGURE 6. Ontology for agent's real-time communication and control protocol in NG.

B. INDIVIDUALISTIC SUPERVISED INTELLIGENCES FOR AGENT

Among indeterministic exertions instigated mainly from SHSES operations and shiftable demand load profiles, ELM intelligence serves as a cogent avenue that envisages probabilistic quantifications on uncertainties while ordering reasoning skills relevant to multi-objective constraints [24], [25]. Moreover, its data mining capacity and adaptive learning aptitude perceive influencing circumstances that can affect performance optimality in gaining desirable results. In this sense, the proposed ELM model was modelled using supervised neural network architecture is entrenched with regularized ensemble regression to enact better universal approximation abilities. Each ELM-based Agents (i.e. IA, non-CL loads, and DSO Agent) procure analytical information that serves as some prognostic initiations based on their operational profiles ruled by STS proceeding. They are programmed to cipher uncertainty into near approximate certainty throughout time, $t \rightarrow (t + n^{th})$ as shown in Table 3. Those predicted results endorses possible propositions that can strategically influence apt decisional making processes involving local NG's power generation availability versus online non-critical load adjustments to gain desirable energy trading and usage benefiting for both Prosumers and DSO-TSO.

TABLE 3. ELM-based agents computational assignments.

Agent	Input	Output (Ruling ranked in ascending order.)
STS	-non-CL Agent -GS Agent -Elec. Tariff -Day & Time	-Elec. Price -ES Capacity -Profiling Online Loads
non-CL	-Weather status -Day & Time -CL Agent	-Demand load Capacity -Large ΔP_{shift}
GS	-Day & Time -Weather status	Local Generation Availability
DSO	-Day & Time -Duck Curve -Elec. Tariff	-Reserve Capacity -Power Generation Pooling

The ELM-based Agent is defined using 3 separate layers (input, hidden and output layer) and interlinked using weighted lines that bridges each node from respective layers as shown in Fig. 7a. Subsequently, it uses an ensemble approach to combine multiple ELMs shown in Fig. 7b to reduce fluctuating performances and increase accuracy as compared to a single structured ELM. The input layer consists of sampled historical data abstracted from each Agent attached to an entity. Likewise, the input layer can also function in a multi-dimension criterion, taking more than a single variable (i.e. power, voltage, SOC, weather status and the list goes on). The hidden layer interlinks the input layer to the output layer, transforming data into implicit information where linguistic features from the input nodes are redefined into relational representations between input and output's objective. Here, non-linear activation function resides at every nodes in the hidden layer to control result's scalability. The function enable resultant data to be within acceptable range, procuring useful data for the subsequent computations. Finally, the output layer presents estimated solution in a single variable domain.

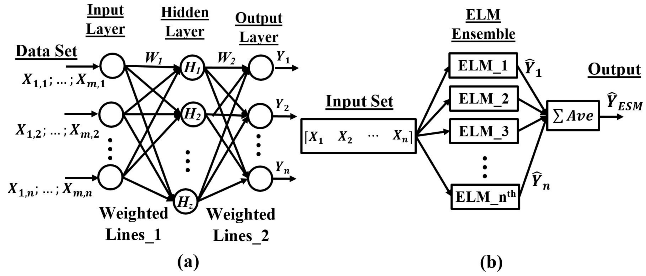


FIGURE 7. Single ELM neural network block. (b) Ensembled ELM formation.

The weighted lines interconnecting input to hidden layer and hidden to output layers attune correlation magnitude to cipher in-between exchanging of data. These weights represent best fit estimations bearing minimal error between actual and estimated results based on its learning algorithm. In ELM, single iterative estimations are only done at the output's weighted line thus, reducing computational time as compared to other gradient-descent based training schemes. Additionally, the regularised factor, I/δ, and ensemble formation were incorporated to resolve performance consistency regardless of input or hidden layer sizes. (1) to (8) demonstrate derivations of ELM:

$$H[n, z] = X[n_{samples}, m_{variables}] * W_1[m, z_{hidden\ nodes}] \quad (1)$$

$$z = ((m + 1)/2) + \sqrt{n} \quad (2)$$

$$Y[n, 1] = g(H) * W_2[z, 1] \tag{3}$$
$$g(H) = 1/(1 + e^{-H_{i,j}}) \tag{4}$$
$$W_2 = H^{\dagger} * Y \tag{5}$$
$$H^{\dagger} = (H^T * H + (I/\delta))^{-1} * H^T \tag{6}$$
$$|\hat{Y} - Y|_{min} = |(H_{out} * W_2) - Y|_{min} \tag{7}$$

constrained to:

$$\hat{Y} = \begin{cases} 0, & \hat{Y} < 0 \\ Y_{max}, & Ymax \leq \hat{Y} \\ \hat{Y}, & otherwise \end{cases} \tag{8}$$

where matrix W_1 is randomly assigned integer ($0 < x < 1$). $g(.)$ denotes a sigmoid-based activation function. X and Y are the input and actual output matrices that have equal number of rows. H expresses the hidden layer matrix formed by the number of collected samples against the calculated hidden nodes $[n \times z]$. I refers to an identity matrix while δ is a real number proportionate to z, ideally. Lastly, W_2 is the regularised output weights connecting to the output layer. Using W_2, generation of new testing set will be generated which corresponds the estimated output, $\hat{Y}$. Further evaluation is then rendered to check sanity of the predicted results using (7). It aims to gain least square error between estimated and actual output. Thereon, ensemble technique is performed on the concatenate ELMs intending to reduce root mean square deviation errors, *RMSD*, and increase accuracy resolution against the actual output, Y.

$$\hat{Y}_{ESM} = \frac{\sum_{i=1}^{N} \hat{Y}_i}{N} \tag{9}$$
$$RMSD = \sqrt{\frac{\sum_{i=1}^{N} (\hat{Y}_i - Y_i)^2}{N}} \tag{10}$$

C. REINFORCED LEARNING FOR COOPERATIVE TENDENCY IN LAYERED COALITION MODEL

Based on Prosumer's perspective towards an ideal EM bureaucracy for local NG operations, the preferred objective statement mainly dwell on; i) strategic procuration of bargained electricity based on real-time tariffs, ii) maximising incentive policies for green efforts, iii) minimizing operating costs of SHSES by securing higher feed-in tariff when selling back to grid. However, such managerial proceedings can propagate possible monopolism in the electricity market and reliability issues at low-voltage level. Prosumer(s) will try to summon low ball bidding techniques that constantly offer marginally cheaper electricity price during peace time and hike in price when the network faces interruptions. Moreover, complications in utility protection schemes will rise (i.e. false tripping of feeder, unsynchronized reclosing, and prevention of automatic reclosing) due to large operation of islanding mode. In this sense, self-centered EM can bring forth poor energy coordination and predicaments for top-level players especially DSO and Retailers. Scheduling of power generation capacity becomes uncertain and incompetent as demand shift profiles get larger due to renewable penetrations. Core operations such power system stability guarding voltage and frequency levels can potentially be afflicted due to inadequate load balancing or following management caused by renewable generation intermittency. Therefrom, the relationship between Prosumer, DSO, and Retailers must co-exist in securing demand-side management and market operations. DSO will be remodel as a neutral energy facilitator that enables competitive access to an open and accessible electricity market markets for Prosumers. To increase cooperative energy transactions at mass, DSO will present incentive packages for participants who are involved in demand-side responses. Such innovation initiates cooperative EM planning for NG to formulate optimal use of local SHSES to be both producer and consumer; enabling security, sustainability and affordability that supports operation optimizations for all energy participants.

In consequence, adaptation of reinforced learning in LCM aims to pilot Agents into endorsing swamp computing that promotes interdependency and bidding-abled EM platform across participants at low-voltage level. Negotiation protocols were exercised to sieve out some candidacy action sets based on learning payoff received as a coalition propensity. It too promotes learning of energy trading management where Agents arbitrate optimal solution from Agent's individualistic decision to gain uniform coalition payoff. Thus, probability stagnation schema was committed to ensure build-up convergence is attained during learning developments [26]–[28].

Here, deployment of Q-learning for Agents is separated into two computation tiers, Tier-1, handles EM administrations based on a single NG operation influenced by Prosumer's expectations while Tier-2, extents collaborative EM bidding across interconnecting NGs. Both Tiers will incorporate constraints required by DSO acting as a governor and policy makers during energy trading while maintaining national balance between supply and demand across distribution network in real-time. The aims are to establish fair and neutral EM which favors cooperative solutions under the governance of Prosumer(s), enhancing demand-side management. DSO will be challenged to support fair, transparent and competitive market mechanisms that will steer energy players to annex liveliest playing-field environment.

1) MODELLING Q-LEARNING FOR AGENT(S) IN LCM

The Agent's learning algorithm is quantified using Q-value where primitively it is set to be a random positive integer value. Agents will then execute respective actions, $a_1, a_2, \ldots$ based on individual state-action pairs under the influence of defined policies. Subsequently, only relevant actions will be stored under a specified state, s. When the state in environment changes s' due to Agent's actions, the Agent will then be assigned with a reward. Through such online rewarding and training system, Agents are constantly embedded with updated Q-value as it advances through the layers of LCM before it converges into a global optima expressed in (11).

$$Q(s_t, a_t)^{new} \leftarrow Q(s_t, a_t)^{current} + \alpha[re_{+1} + \gamma * Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)^{current}] \quad (11)$$

where α and γ are real integer numbers $(0 < x < 1)$ defining Agent's learning rate and temporal discount factor respectively. re_{+1} denotes acquired payoffs either reward $(+)$ or penalty $(-)$ received when moving from current state s to the next s_{+1}. $Q(s_t, a_t)^{new}$ projects the estimated action-value function based on action taken. It too observes severity of temporal difference (TD) factor, $re_{+1} + \gamma * Q(s_{i+1}, a_{i+1})$, where it depicts Agent's learning value functions under no prior knowledge $Q(s_i, a_i)$ deviate from the targeted consequence. Agents' would need to wait for payoff status to reflect reward before state-action pair values can be updated.

(a) *Cooperation Tendency*-Consider there are n^{th} number of Agents and a_i represents corresponding actions thus, $[a_1, a_2, \ldots, a_n]$ or $\overrightarrow{a}$ encapsulates the action set of peculiar Agent. $N(\overrightarrow{a})$ is the number of action set performed. φ is a boolean data type that denotes the learning stagnation or convergence status. $CP(\overrightarrow{a})$ quantifies the cooperation probability in relation to φ where larger CP offers higher cooperation tendency for Agents and vice versa. Hence, the following derivations present computations of the cooperation tendency in LCM; i) initializing $CP(\overrightarrow{a})$ randomly for involving joint action sets, ii) compute new CP after executing Agent's actions based on respective objectives, iii) equate $N'(\overrightarrow{a}) = N(\overrightarrow{a}_{+1})$, iv) assign 0 to φ if the state of Agents did not change or 1 for otherwise, and lastly v) compute CP given in (12).

$$CP(\overrightarrow{a})^{new} = CP(\overrightarrow{a}) + [(\varphi - CP(\overrightarrow{a}))/N'(\overrightarrow{a})] \quad (12)$$

where

$$\varphi = \begin{cases} 0, & \textit{Agent solution becomes stagnated} \\ 1, & \textit{otherwise} \end{cases}$$

(b) *Payoff (PO) Function*-It is a reward system that provides coalition probability comprehension from Agents' performed actions. The payoff capacity derive in (13) denotes desirable learning processes achieved by Agents and it will be reflected on respective Q-value. Thus, payoff will ultimately guide Agents to take the best possible actions from individual Q-learning algorithms. Given that the operations of Agents in NG can either be disjoint or synergetic EM, the payoff assignment for poor decision-making actions will be separated into two domains; negative and neutral net values respectively.

$$PO(\overrightarrow{s}, \overrightarrow{a}, AG) = Q_{AG}(\overrightarrow{s}, a_{AG}) * CP(\overrightarrow{a}) \quad (13)$$

where the computed payoff reflects solely for that peculiar Agent based on all actions and states implied. Let a_{AG} be the actions performed by a peculiar Agent AG. n is the number of Agents deployed and $\overrightarrow{s}$ represents the state of all joint Agents.

(c) *Candidate Policy*-The candidacy in selecting Agent's actions are influence heavily from payoffs capacity thus, Agents are forced to perform actions that are reward-driven to gain high coalition payoffs. Firstly, it selects some action sets and goes through Nash Bargaining Theorem (NBT) where candidate policies will sieved relevant action-selection criterion. The sequence in establishing the candidate policy procedure are as follows; i) use transferable utility to choose some candidate action sets for all involving Agents denoted by $A_C(\overrightarrow{s})$. ε, bounded with a threshold limit between 0 to 1, denotes the controls size of candidate action sets shown in (14), ii) compute the target policy, $\pi(\overrightarrow{s})$, given in (15) expressing which action to be taken in state $\overrightarrow{s}$ mapping deterministic action a_{AG} from state $\overrightarrow{s}$, iii) employ maxima function arguments based on NBT theorem for each action set $\overrightarrow{a} \in A_C(\overrightarrow{s})$ against the minimum of Q-value, q^*_{AG}. It uses Cartesian Sum-Product on finite state and action spaces of all Agent's belief as a probability.

$$max_{val} = \max_{\overrightarrow{a}} \left[\sum_{i=1}^{n} PO(\overrightarrow{s}, \overrightarrow{a}, AG) \right] \quad (14)$$

if $(max_{val} - \sum_{i=1}^{n} PO(\overrightarrow{s}, \overrightarrow{a}, AG)) \leq (\varepsilon * max_{val})$ then add $\overrightarrow{a}$ into $Ac(\overrightarrow{s})$.

$$\pi(\overrightarrow{s}) = arg \max_{\overrightarrow{a}} \left[\prod_{AG=1}^{n} PO(\overrightarrow{s}, \overrightarrow{a}, AG) - q^*_{AG} \right] \quad (15)$$

(d) *Personalized Q-learning using LCM*-The formulated candidate policy for each Agent endorses payoff coalition comportment and records situation of Agents in Q-Table. The Q-Table neglects taking each individual Agent's action into review. The flowchart given in Fig. 8 illustrates the Q-learning of Agents in LCM domain.

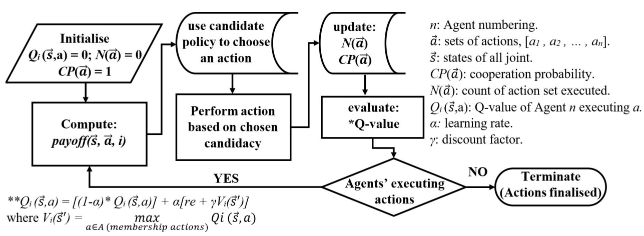


FIGURE 8. Agents' Q-learing flowchart sequence in LCM.

2) FORMULATE AGENTS' REINFORCE LEARNING AND POLICIES FOR EM

Subsequently, algorithm delegations in view to EM problems at demand-side are fused into the proposed reinforced learning to visualize how these four Agents, ELM-based

DSO and IA, PCM & PCCM, parallelly negotiate in keeping power delivery balanced. The criteria involves strategic scheduling of SHSES and non-critical loads, comprehending data abstractions from ELM-based Agents, devise economic wager to offer bidding propositions that can influence better regulatory of peak demand intervals, generate healthy power generation profiles by cultivating whole system optimization perception, and facilitate energy market arrangements for DSO-Prosumers trading chain.

1) *PCCM Agent: Action a_1, Candidate Policy*- The focal objective is to govern online load capacity in a NG through load balancing theorem (equality constraint). Here, the algorithm engages non-CL Agent to forecast desirable online load capacity at the next time stamp. Therefrom, (16) and (17) formulate load minimization approach that proffers optimal electrification criterion where scheduling of shiftable loads are strategically shed based on generation availability. The time constraints dictated for computations were restricted to half-hour time span.

$$\min \sum_{t=1}^{n} [Pload_{online} - Pload_{objective}]^2(t) \tag{16}$$

$$Pload_{online}(t) = [Pload_{forecast} + Online - Offline](t) \tag{17}$$

where $Pload_{online}(t)$ and $Pload_{objective}(t)$ are actual and desirable load capacity at time t respectively. $Pload_{forecast}$ refers to combination of critical and non-critical load consumptions while $Online(t)$ and $Offline(t)$ are the connected and disconnected load at time t during the load shifting intervals. Individually, $Online(t)$ and $Offline(t)$ are separated into two definitions; incremental and decremental of connected load devices shifted at time t and precede t (forecasted) correspondingly as shown (18) & (19).

$$Online(t) = \sum_{i=1}^{t-1}\sum_{d=1}^{D} A_{di(t)}P_{1d} + \sum_{p=1}^{P-1}\sum_{i=1}^{t-1}\sum_{d=1}^{D} A_{di(t-1)}P_{(1+p)d} \tag{18}$$

where A_{dit} denotes the number of devices, d shifted from time t to $(t+1)$. P is the duration span of device being online, and P_{1d} & $P_{(1+p)d}$ are the power ratings consumed at first time stamp and the next, $(1+p)$ respectively at reciprocal devices d. Correspondingly, $Offline_{load}(t)$ includes initial and estimated energy consumptions from time t to i where device is fully disconnected.

$$Offline(t) = \sum_{i=t+1}^{t+del}\sum_{d=1}^{D} A_{d(t)i}P_{1d} + \sum_{p=1}^{P-1}\sum_{i=t+1}^{t+del}\sum_{d=1}^{D} A_{d(t-1)i}P_{(1+p)d} \tag{19}$$

where *del* implicates maximum time delay. The minimization statement has its constraints where selection of least one shiftable load device ($Adti > 0$; $\forall d, t, i$) and the number of load devices shifted in between time stamps i, $i+1$ cannot be more than the number of online non-CL Agents, $\sum_{t=1}^{N} A_{dit} \leq non\text{-}CL(i)$.

The policy assigned to a_1 determines the maximum demand, *MD*, capacity at respective NG. For every 15mins intervals, the $Online_{load}(t)$ will be compared against *MD* to ensure its capacity is below $max\ MD_{total}$. Note that the *MD* policy may defer based on load category (i.e. mechanical or electrical) and user-defined diversity factor percentage, *DivF*. *MD* reflects the highest permissible electricity consumptions given at any time stamp, suppressing sudden peak demand which can excites unscheduled generation. Hence, DSO will imposes penalty to Prosumer(s) due to poor management of electricity usage. For these reason, realizing maximum *MD* and minimizing demand factor *DF* can aid in shaving peak load crises.

$$\max MD_{total} = \sum_{d=1}^{D} [(PnonCL_d * DivF_d) + PCL_d] \tag{20}$$

$$DF(t) = MD(t)/Pload_{online}(t) \tag{21}$$

$$payoff(\overrightarrow{s}, \overrightarrow{a}, AG) = \begin{cases} -1, & 1.0 > DF(t) \\ 1, & 1.0 < DF(t) \end{cases} \tag{22}$$

2) *PCCM Agent: Action a_2, Candidate Policy*- In conjunction with a_1, PCCM Agent too is modelled based on a data-driven computing consequences that relies on the forecasted data prescribed by ELM-based Agents. These sampled data (5mins interval) will influence Agent's decisional performances in attaining optimal EM and assets deployments.

The policy for a_2 ensure forecasting results are not overfitted and data are not misinterpreted due to data "noise" picked as part of the model during learning processes. Therefrom, to increase predictive accuracy, training and testing sets errors are to be monitored using k-fold cross validation where it appreciates model's performances against unseen data. Explained in [29], it highlights the importance in splitting predictive model into size of training and testing sets, and shuffling of data subset selections to gain interpretation accuracy.

$$E_k(\lambda) = \sum_{i \in k^{th}\ part} (y_i - x_i\hat{\beta}^{-k}(\lambda))^2 \tag{23}$$

$$CV(\lambda) = \frac{1}{K}\sum_{k=1}^{K} E_k(\lambda) \tag{24}$$

$$payoff(\overrightarrow{s}, \overrightarrow{a}, AG) \tag{25}$$

$$= \begin{cases} 0.5, & 10 < MAPE(t) < 20\% \\ 1.0, & MAPE(t) < 20\% \\ 0, & otherwise \end{cases} \tag{26}$$

where λ is the estimated tuning parameter, E_k computes the error in predicting k^{th} which typically set at $k = 5$ or 10, $CV(\lambda)$ depicts the cross validation error through iterative procedure in changing λ, and $MAPE$ measures the percentage of error between forecast and actual values.

(c) *PCM Agent: Action a_1, Candidate Policy*- The actions delivered by a_1 is separated by two interlinked domains which involved ordering of energy based on load power mismatch and scheduling of local battery operations. Primitively, PCM Agent focuses on procuring a balanced power criterion that schedules local SHSES individually or import power from the Retailer(s) to compensate demand shift capacity at time, t.

$$\Delta P_i^{DIFF}(t) = [P_{grid} + \sum_{i=1}^{N} P_{DG,i} + P_{batt}](t) - Pload_{online}(t) \tag{27}$$

$$\Delta P_{DG,i}^{ADJ}(t) = \frac{(P_{DG,i}^{MAX} - P_{DG,i}(t)) * \Delta P_i^{DIFF}(t)}{\sum_{i=1}^{N} P_{DG,i}^{MAX} - P_{DG,i}(t)} \tag{28}$$

where $\Delta P_{DG,i}(t)$ denotes the power change in time t incurred by individual SHSES only referring to thermal or renewable power generating units. $\Delta P_i^{DIFF}(t)$ indicates the power difference between online generation capacity and load, $\Delta P_{DG,i}^{ADJ}(t)$ quantifies the power adjustments required for each local source at specific t. Note that, when exploiting local thermal generating units, asset's constraints involving valve-point loading effect needs to be considered when adopting (28). Thus, with (29) & (30), proper delegations of generating units were rendered to satisfy ΔP_i^{DIFF}.
when generating unit output increases:

$$\begin{aligned} &P_{TU,i}(t) - P_{TU,i}(t-1) \leq UR_i \\ &\min[P_{TU,i}(t-1) + UR_i(t), \\ &P_{TU,i}(t-1) + \Delta P_{TU,i}^{ADJ}(t)] \end{aligned} \tag{29}$$

when generating unit output decreases:

$$\begin{aligned} &P_{TU,i}(t-1) - P_{TU,i}(t) \leq DR_i \\ &\max[P_{TU,i}(t-1) - DR_i(t), \\ &P_{TU,i}(t-1) - \Delta P_{TU,i}^{ADJ}(t)] \end{aligned} \tag{30}$$

where P_{TU} refers to local online thermal unit generator where applicable. $UR_i(t)$ and $DR_i(t)$ are corresponding increase and decrease valve-point loading capacity at fixed interval time steps.
Inevitably, there will be instances when $\Delta P_i^{DIFF}(t)$ will not be satisfied thus, power dependency will divert to either Retailers, energy storage or mixture of both. Therefore, involvements in energy storage utilization needs strategical control measures to ensure healthy investment return is gained and substantial reserve power pooling is observed before resorting to Retailer(s). It should be noted that constant maximization of energy storage may not necessarily be an optimal solution.

$$P_{batt}(t) = [(\sum_{i=1}^{N} P_{TU,i} + P_{renew,i}) + P_{grid}](t) - Pload_{online}(t) \tag{31}$$

if $P_{batt} > 0$ AND $P_{batt} < P_{batt(max)}$ is charging.
if $P_{batt} < 0$ AND $P_{batt} > P_{batt(reserve)}$ is discharging.
where P_{renew} refers to the renewable power generation sources. Conventionally, exploitations the battery's SOC% aids in monitoring the energy threshold levels $P_{batt(reserve)}$ or $P_{batt(max)}$ regions. Based on the SOC% levels at time t, coupled batteries are to perform perpetual interchangeable operations between charging, discharging or remain disengaged.

$$SOC\%(t) = SOC\%(t-1) + (\int_0^t \frac{1}{C_{batt}} dt) \tag{32}$$

$$CG(t) = CG(t-1) + [\Delta t * \frac{\eta_{batt}^{DP}}{V_{batt}(t)} * P_{batt}(t)] \tag{33}$$

$$DCG(t) = DCG(t-1) - [\Delta t * \frac{\eta_{batt}^{DP}}{V_{batt}(t)} * P_{batt}(t)] \tag{34}$$

where $CG(t)$ and $DCG(t)$ represent the battery's corresponding SOC operation and capacity at respective time stamps, η_{batt}^{DP} refers to the battery's efficiency incorporating depreciation factor over time, V_{batt} measures the voltage level at battery terminal, and C_{batt} is the amount of electric charge passed through the battery expressed in ampere an hour.
To obtain optimal utilizations of energy storage in relation to power dependency from Retailers while securing load balancing criterion, (35) and (36) are policies that pivoted on real-time electricity tariff profiles and how Prosumer can reschedule its demand load and generating assets to transact maximum profit margin.

$$\max Profit(t) = [Revenue - Expenses](t) \tag{35}$$

$$Profit(t) = [ET * \sum_{i=1}^{N} P_{DG,i} + ET.P_{grid}](t) - [ET.P_{grid} + \sum_{i=1}^{N} bid(P_{DG,i})] \tag{36}$$

where ET refers to the wholesale electricity price and $bid(P_{DG,i})$ is the bidding price to DSO for respective local power sources i. However, deployment of local SHSES needs further exploration into defining it credibility towards the market operations and

Prosumers' expectations when bidding excess generation back to Utility. Thus, Feed-in tariff (FiT) policy was introduced to proffer cost-based compensation to SHSES producers through stable long-term agreement (15-25 years period) with DSO. The contract guarantees bidding price certainty that helps finance renewable energy investments while releasing incentives for SHSES deployments to fund FiT scheme [30]. Thus, to accelerate investments, (37) maximizes FiT policy revenue streams by predominating the wholesale electricity tariff [31].

$$LCOE = \frac{\sum_{n=0}^{N} C_n * (\frac{1+i}{1+d})^n}{\sum_{n=0}^{N} Q_n * (1-D)^n}$$

$$\begin{aligned} \max \sum_t [&(P_{DG,i} * \rho_{LCOE}) + (P_{DG,export} * \rho_{FiT}) \\ &- (P_{grid} * \rho_{wholesale}) - (Pcharge_{grid} * \rho_{wholesale}) \\ &+ (Pdischarge_{grid} * \rho_{wholesale})](\Delta t) \end{aligned} \tag{37}$$

subjected to:

$$0 \leq P_{DG,export}(t) \leq P_{DG,excess}(t) \tag{38}$$

$$P_{DG,excess(t)} = \begin{cases} P_{DG} - Pload_{online}, & P_{DG} > Pload_{online}(t) \\ 0, & otherwise \end{cases} \tag{39}$$

where C_n is the investment costs of SHSES involving installations to full deployment to maintenances, Q_n constitute energy produced in kilo-Watt hour inclusive of efficiency degradation factor, D percentage change per annum, except at Q_0 where it omits out degradation impact. N states the system's expected lifespan expressed in years while i and d denote the inflation and discount rates representing investment percentage change per annum. (37) formulates the objective function in gaining maximum profit margin when incorporating SHSES into the energy mix during operation. ρ_{LCOE}, ρ_{FiT}, and $\rho_{wholesale}$ denote the real-time operating costs of SHSES defined as liveliest cost of electricity (LCOE), the promised tariff offered by DSO based on FiT policy, and wholesale prices are actual real-time rates that Prosumer are paying for every kilo-Watt hour. $Pcharge_{grid}$ corresponds to the power rating absorbed from Retailers while $Pcharge_{excess}$ and $Pdischarge$ indicate the ordered power level that charges the battery due to excess local generation or discharging to satisfy ΔP^{DIFF} respectively. Lastly, P_{DG} refers to the available power capacity generated by local power generation units.

In addition, the objective function needs to include scheduling of charging and discharging of energy storage supported either from grid or local generations:

$$\begin{aligned} charge(t) = charge(t-1) &+ [(\eta_{batt}^{DP} * Pcharge_{excess}(t)) \\ &+ (\eta_{batt}^{DP} * Pcharge_{grid}(t)) \\ &- (\frac{Pdischarge(t)}{\eta_{batt}^{DP}})] \end{aligned} \tag{40}$$

$$\begin{aligned} Pload_{online}(t) = [&P_{grid} + P_{DG} - P_{DG,export} \\ &- Pcharge_{excess} - Pcharge_{grid} \\ &+ Pdischarge](t) \end{aligned} \tag{41}$$

$$\Delta P_i^{DIFF}(t) = [Pdischarge + P_{grid}](t) \tag{42}$$

(d) *DSO Agent: Candidate Policy*- When endorsing large penetration of local power generations at low-voltage level, DSO is constantly challenged with load balancing crises and unpredicted energy scheduling as operations of renewable system can instigate Prosumers' electricity usage profiles throughout the day. To worsen, as Retailers are singly bounded with contract agreements stating individual expectancy for power generation capacity, DSO will be trapped in a situation where power difference compensations could not be met thus leading to network failure. Sudden hike in electricity price too can be seen as energy reserve pooling is being tapped causing Prosumers to pay premium tariffs. Such phenomenon leads to demand load demography which undertakes a Duck Curve profile where the Duck's neck region represents low renewable penetrations thus inflicting high demand of energy from Retailers. Contrarily, the belly region depicts high penetration of distributed generation causing Retailers to suffer losses due to under utilized spinning reserves. Hence, DSO is requires to recognize Prosumer's baseline load trends especially at the peaks of two period: absence and high penetrations on SHSES against demand load profiles. Such criterion aids in achieving optimal scheduling of thermal generators per day to meet immediate shortfalls of power imbalance despite high uncertain fluctuations provoked by SHSES.

The role of DSO Agent is to comprehend demand load profile trends and estimate required reserve generation capacity ciphered by the ELM-based STS. Thus, implementation of state-action pairs were omitted. However, the policy ensures Duck Curve profile is guarded and constant energy utilisation is above baseline level, P_{baseL}. Large coalition payoff will be rewarded to Prosumer that aids in minimising power deviation, $P_{deviate}$, contributions. Such motives introduces gradual inclinations on demand curve profile with manageable power ordering from thermal generating units.

$$\begin{aligned} P_{deviate}(t) = P_{grid}(t) - P_{grid}(t-1) \\ P_{grid}(t) \geq P_{baseL}(24hrs) \end{aligned} \tag{43}$$

for t = 15mins intervals.

$$payoff(AG) = \begin{cases} +1.0, & P_{deviate}(t) < \sum_{i=1}^{N} UR_i(t) \\ +1.0, & P_{deviate}(t) > \sum_{i=1}^{N} DR_i(t) \\ -1.0, & otherwise \end{cases} \tag{44}$$

(e) *IA Agent: Action a_1, Candidate Policy*- Endorsing SHSES with battery storages brings forth concerns when searching for an economically solution that procures high return on investments and performance index. Prosumers are expecting maximum profit margins from SHSES operations, hoping to gain bargained operating costs. Hence, incorporating IA Agent, it governs LCOE of local SHSES integrations by evaluating it's economic assessment in contrast to the market operations. Here, formulated LCOE was further enhanced using (45) to (50) defining its combination deployment of NG's power generations and energy storage.

$$E_{DG} = E_{renew} + E_{TU} \tag{45}$$

$$LCOE_{DG+batt} = \frac{\sum_{i=1}^{N} CE_{(DG_i+batt)}}{\sum_{i=1}^{N} E_{(DG_i+batt)}} \tag{46}$$

$$CE_{DG_i+batt} = LCOE_i * E_i \tag{47}$$

$$LCOE_{DG+batt} = \frac{\sum_{i=1}^{N} LCOE_i * E_i}{\sum_{i=1}^{N} E_i} = \sum_{i=1}^{N} LCOE_i * \alpha_i \tag{48}$$

$$\alpha_i = \frac{\Delta_t * \beta_i * P_i}{\Delta_t * \beta_{DG+batt} * P_{DG+batt}} \tag{49}$$

$$\beta_i = [\frac{E_{batt,in}}{E_{batt,rated}} * \frac{\eta_i}{1+\eta_i}] \leq 50\%\ typically \tag{50}$$

where $LCOE_{DG+batt}$ quantifies the total worth of electricity produced by SHSES in dollars. CE denotes the cost of energy. E and α_i represent energy generated by respective local generation source while β_i defines the capacity factor of energy storage in relations to the charging and discharging transactions. η_i equates the efficiency of respective local generation sources.

The policy administered involves piloting excess power generation that neither was stored nor consumed by local loads. Due to poor utilisation of energy storage during high penetrations of renewable resources, owners are resulted to selling local surplus power at low prices and vice versa based on the market electricity tariff. Thus, the policy aids in managing excess power generation which are eventually sold to the grid.

$$E_{DG}^{USE}(t) = \sum_{i=1}^{N}[E_{DG,i} - E_{DG,i}^{waste}](t) \tag{51}$$

$$E_{batt,in}(t) = UI * E_{DG}^{USE}(t), \quad where\ UI \leq 1.0 \tag{52}$$

$$E_{DG+batt}(t) = [1 - UI] * [E_{batt,out}(t) + (E_{DG}^{USE}(t) - E_{batt,in}(t))] \tag{53}$$

$$payoff(\overrightarrow{s}, \overrightarrow{a}, AG) = \begin{cases} 0.0, & \sum_{i=1}^{N} E_{DG,i}^{waste}(t) > \sum_{i=1}^{N} E_{DG,i}^{waste}(t-1) \\ 1.0, & otherwise \end{cases} \tag{54}$$

where $E_{DG,i}^{waste}(t)$ denotes excess energy that is dumped to Utility due to mismatch load balance or unscheduled use of storage at time t. $E_{batt,in}(t)$ defines the energy level stored into battery while UI scales the harvested generations capacity available for charging the battery. Lastly, $E_{DG+batt}(t)$ quantifies the amount of energy generated by battery and local generation sources.

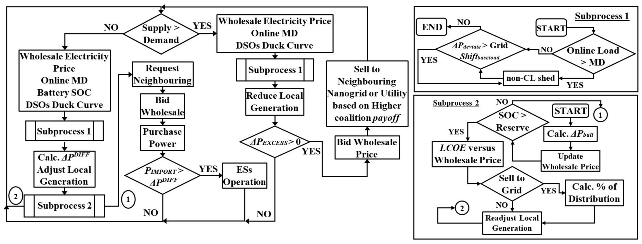


FIGURE 9. Agent's rules of engagement in gaining proposed EM at demand-side.

D. PROPOSED ENERGY FLOW MANAGEMENT

A detailed flowchart diagram illustrated in Fig. 9 interprets Agents' interactions based on proposed EM at individual NG. It transacts day-ahead and real-time asset scheduling capabilities though adaptations of hybridized ELM- and LCM-based Agents. Comprehensively, Agents are to autonomously negotiate and avoid decisional conflicts in procuring optimum power sharing transactions based on available generation resources against demand load at time of request.

V. CASE STUDIES

Investigations were done on two separate case studies: (a) Managing asset utilizations using proposed methodology on a single NG network, *NG*1, based on individualism interests, (b) Inspecting level of intelligence in piloting four aggregated NGs, *NG*1-*NG*4, based on a coalition settlements to transact power exchange among themselves. Exceedingly, observations were also laid upon to analyses how the proposed methodology can serve as a lead energy regulator in ciphering DSO's typical obligatory during grid-tied operations.

A. EM OPERATIONS OF A SINGLE NANOGRID SYSTEM (RESIDENTIAL, NANOGRID 1)

24-Hrs simulation analyses were performed on a single-phased 220V AC, 50Hz Residential *NG*1 given in Fig. 14 where real-time data of respective integrated systems were sampled at every 5mins intervals. As proposed, the 25kV A network involves a 3.5kW PV system with a 2.2kWh energy storage unit(initial 0% SOC) and a back-up diesel generator rated at 2.2kWh. Contrarily, the online demand loads were recorded in the summer where a mixture of critical, non-critical home appliances and a charging point for electrical vehicle (EV) were conceded.

In reference to Fig. 10b and 10c, all endorsed ELM-based GS and non-CL Agents were deployed to forecast available

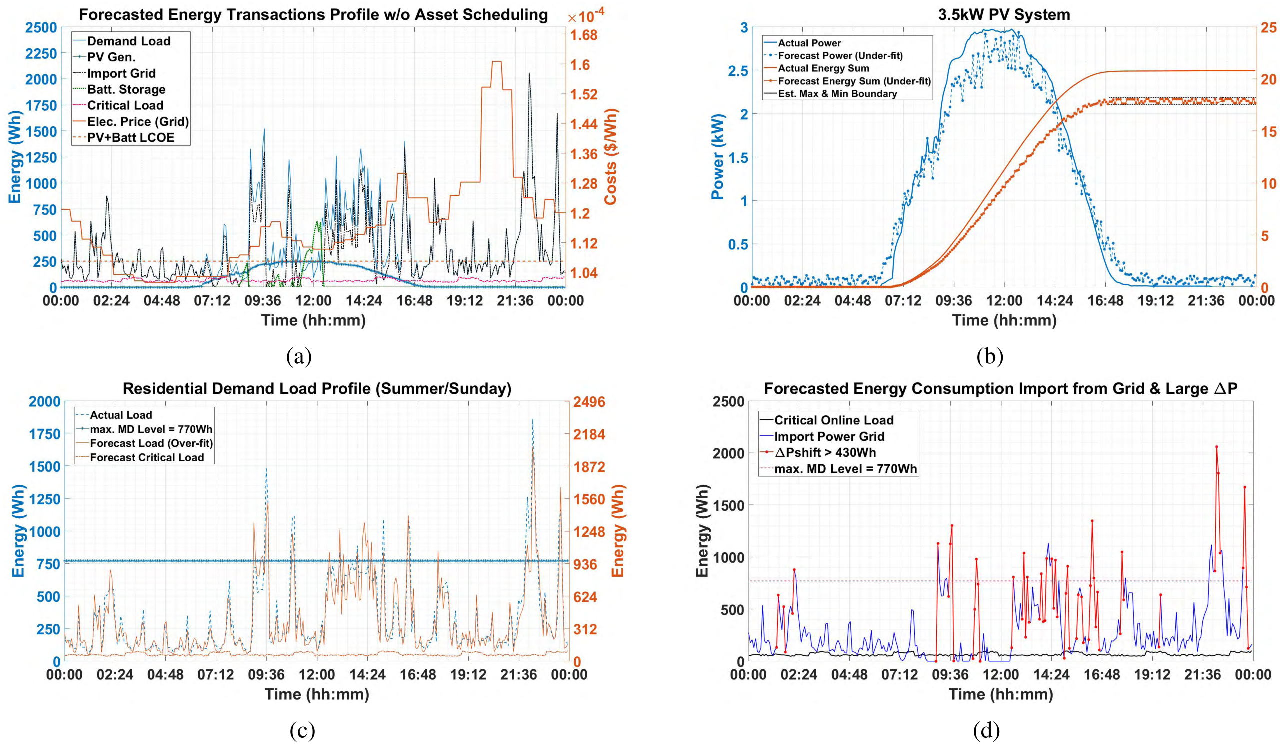


FIGURE 10. Forecast power and energy profiles for *NG*1 w/o proposed EM (24-Hrs data sampling at 5mins intervals). (a) Load balancing proceedings w/o asset scheduling against electricity price tariffs. (b) 3.5kW PV system against forecast data. (c) Online demand load against forecast data conjointly identify critical and non-shiftable appliances. (d) Flagging large energy deviation gratified by the utility.

TABLE 4. ELM-based agent's forecasting performances.

Agent	Accuracy (RMSE)	Std. Deviation (σ)
STS	Elec. Price: 2.162	0.134
2e10.8 < δ < 2e11.05	ES Capacity: 2.044	0.062
15 < ESM < 25	Load Profiling: 4.135	1.088
Computing Time: 0.323s		
non-CL (over-fit)	Load Capacity: 3.091	0.119
2e8.1 < δ < 2e8.7	Large ΔP_{shift}	0.038
15 < ESM < 25	Detection: 2.523	
Computing Time: 0.177s		
GS (under-fit)	Generation	0.065
2e12.7 < δ < 2e13.1	Availability: 2.337	
15 < ESM < 25		
Computing Time: 0.525s		
DSO	Reserve Capacity: 2.061	0.078
2e10.8 < δ < 2e11.05	Gen. Pooling: 1.915	0.111
15 < ESM < 25		
Computing Time: 0.318s		

TABLE 5. STS agent forecasting alarm report.

ΔP_{shift}	PV Op.	max MD	Batt. Charge	Elec. Price
01:20-02:10	Sunny	02:05-02:10	08:30-08:55	Tier 1
08:55-09:00	ClearSky	08:55-09:00	09:55-10:30	00:10-08:25
09:30-09:45	Sunrise:	09:20-09:40	10:55-11:00	Base:$1.086e-4
10:40-11:00	06:10	10:45-10:50	11:15-12:25	00:30-00:55(EX)
12:30-12:35	High PEN.:	12:30-12:35		Tier 2
13:00-13:20	09:30-13:40	13:05-13:15		08:30-14:20
13:50-14:10	Sunset:	13:50-13:55		Base:$1.102e-4
14:25-14:40	18:05	14:10-14:35		09:30-10:25(EX)
15:00-15:15		15:05-15:10		14:00-14:15(EX)
15:30-15:55		16:20-16:25		Tier 3
16:10-16:40		17:50:17:55		14:20-00:00
17:40-17:50		21:55-22:30		Base:$1.246e-4
19:30-19:35		23:25-23:35		20:00-21:25(EX)
22:10-22:25				
23:30-23:45				

PV generations and online non-critical demand capacities. As explained in Section IV-B, personalized ensemble units were tuned to control ELM's learning behaviour for corresponding applications. For GS Agent instances, under-fit regularization was engaged to reduce risk of overrated generation availability and vice-versa for non-CL Agent. Performances of ELM-based Agents are summarized in Table 4 based on a 2 years training and 1 year testing datasets. Subsequently, the ELM-based STS Agents will retrieve all forecast data and generate a timed-based predicament report, Table 5, suggesting plausible alarming exposures during operations. It also forecasts consumption patterns for non-critical loads, separating them from shiftable and non-shiftable loads as shown in Table 6. Primitively, critical load capacity can be determined by recognizing the network's baseline load while the non-CL are classified by distinguishing appliance's operating duration against predefined threshold of 30mins intervals.

Conversely, without implementing proposed methodology, Fig. 10a illustrates unambiguous energy transactions where the battery charges when excess power generated from PV. Mobilization of available stored energy is deployed to satisfy load balancing criterion at t. Fig. 10d evaluates large energy

TABLE 6. STS agent forecast shiftable and non-shiftable consumption patterns (30mins intervals).

(HRS) Time	(Wh) S	n-S	(HRS) TimeS	(Wh) S	n-S	(HRS) Time	S	(Wh) n-S
00:00 -	213	524	08:00 -	223	561	16:00 -	1077	2335
01:00	302	665	09:00	155	306	17:00	452	892
01:00 -	217	596	09:00 -	1550	3994	17:00 -	190	375
02:00	507	1127	10:00	1197	2822	18:00	960	2054
02:00 -	969	2211	10:00 -	409	882	18:00 -	749	1759
03:00	384	857	11:00	1022	2391	19:00	297	727
03:00 -	111	246	11:00 -	397	919	19:00 -	203	519
04:00	224	517	12:00	239	520	20:00	358	734
04:00 -	141	377	12:00 -	404	846	20:00 -	264	755
05:00	251	485	13:00	1173	2837	21:00	225	470
05:00 -	64	178	13:00 -	1213	3007	21:00 -	404	911
06:00	189	480	14:00	1397	3096	22:00	760	1954
06:00 -	188	448	14:00 -	1585	4018	22:00 -	2306	5082
07:00	211	477	15:00	1099	2690	23:00	897	2158
07:00 -	90	226	15:00 -	649	1454	23:00 -	403	959
08:00	503	1051	16:00	851	1747	23:59	924	2237

Shiftable(S): Washing Machine, Clothes Dryer, Dish Washer, Air-condition, Standby Appliances and Charging of Energy Storage or Electronic devices
non-Shiftable(n-S): Computers, Entertainment Electronics, Lighting, Electric fans, Water heater and other small kitchen or beauty appliances.

deviation drawn from Utility to compensate $\Delta P_i^{DIFF}(t)$. The attained results conclude to have poor EM that avow Duck Curve phenomenon to transpire and deployment operating costs of SHSES were not optimized. Large power demand shift instances were noticed during two peak periods where local PV generation is inactive and energy storage is depleted thus, plausible divergence in load balancing may surface due limitation in generators' ramp rates. Likewise, high operating costs was propagated firstly due to unitary dependency on Retailers after 1720*hrs* where electricity pricing starts to escalate and secondly violations of *max* MD_{total} were exposed initiating penalties to be imposed on Prosumer.

Diversely, deployment of cooperative Q-learning for LCM-based Agents were rendered to view its impacts on ciphering optimum EM in response to Prosumer's electrification interests and demand-side management regulatory. Fig. 11 presents results attained from adopting the proposed nano-biased energy manager based on an initial 0% SOC, maintaining at least 75% of Prosumer's electrification lifestyles and securing minimal baseline load of 1.2*kW* supported solely by Retailers. Fig. 12 exhibits Agent's search space and learning regression in ciphering optimal EM operations.

In Fig. 11a, results depict respective components of demand load profiles involving Prosumer's online shiftable loads that are continually tuned in accordance to a time-based shifting algorithm. An interval of half-hourly load classifications limits the algorithm's search space when redistributing the non-shiftable loads as not to promote high load shedding percentage. Objective is to protect Prosumer's lifestyle where 100% serviceability for shiftable loads is maintained and yet rendering bargained electricity bills. In this sense, participating LCM-based Agents were refrained from exploiting shedding solutions or rescheduling beyond time-stamp threshold (30mins) to gain superiority in operating

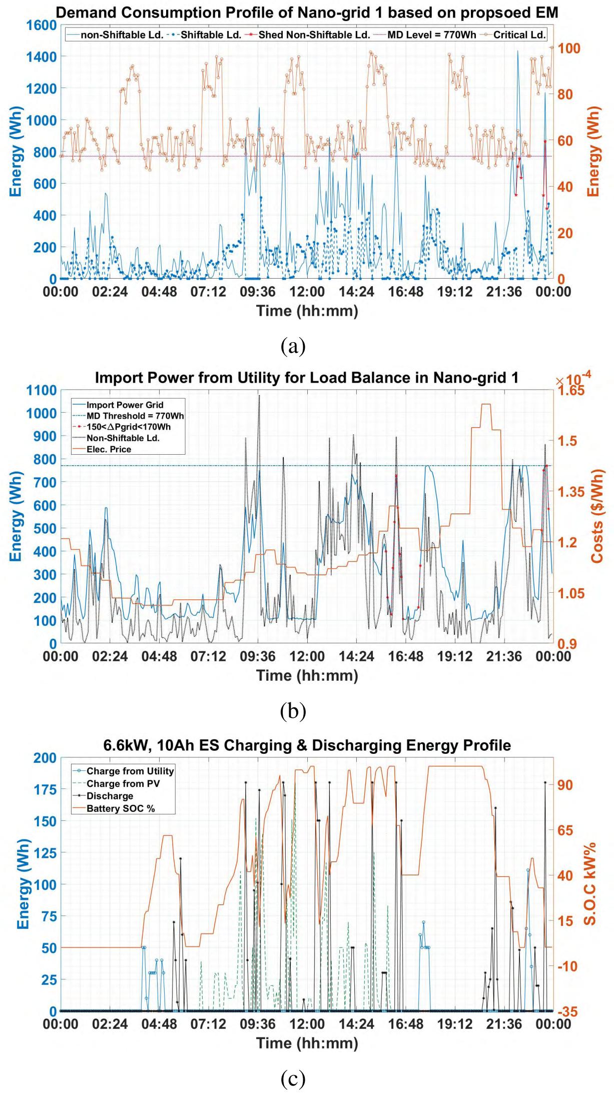


FIGURE 11. Proposed EM proceedings on *NG*1. (a) Power exchange transactions between utility and scheduling of shift-able loads. (b) Charging & discharging profile of battery storage unit with fixed LCOE of 1.22e-4 $/Wh.

costs or evade from possible grid-tied violations. Spotted two instances of shedding non-shiftable loads were directed (highlighted in red lines) for a duration of 10 and 5mins respectively primarily to conserve *max MD* criterion despite having shiftable loads suppressed and energy storage atoning at maximum discharging rate.

Fig. 11b presents energy transactions from Retailers and meeting load balance in *NG*1. It has shown significant improvements in scheduling $\Delta E^{Deviate}(t)$ capping at less than 170*Wh* for any operation time t to $(t+1)$, elevating sudden power deviation in the demand curve. In consequence, it trims overestimated spinning reserve dilemma and levelised with demand shift deviations based on generator's ramp-rate limits. Moreover, the power curve profile of import power from Retailer has complied fully to *max MD* and baseline load constraints. However, the Duck Curve phenomenon has not shown significant improvements as EM was biased towards energy market oriented operations. Greedy-based task allocations were observed when tuning demand loads and local energy storage capacity against real-time electricity tariff. In addition, Prosumer had to pay higher tariff premium by 3%, excluding incentives due to preserving *max MD* level and $\Delta E^{Deviate}(t)$ thresholds, where shiftable loads are exigently shifted into the high electricity tariff regions during the eleventh hour.

Fig. 11c exhibits operations of the 6.6kW energy storage with a 184Wh charging and 180Wh discharging rate for every 5mins. With operative report generated by STS Agent, the energy storage is strategically scheduled to anticipate charging and discharging operations to concede optimal operating costs at time, $(t+1)$. The energy storage utilizations were based on electricity tariff (3-tier price regions), PV generation (highest energy harvested) and demand load (peak order) profiles. Supplementarily, the deployment of energy storage can be exploited to compensate large $\Delta P^{Deviate}_{grid}$ capacity, compensating sudden inclination or declination of load balance. The operations of ES Agent ensures a minimal SOC threshold of 15% is maintained dictated by the assigned policy for reserve energy pooling. However, there were two admissible instances where the battery's SOC plunges below threshold at 0550hrs and 2200hrs. The first event was to discharge excessive stored energy and embrace oversized PV capacity while the second was to lead profiteering electricity price bidding and curtail energy consumptions across the day. Both cases did not surfaced any critical concerns when securing energy reserve capacity as; (a) network adopts a back-up diesel generator in times of service interruptions occurred at primary-side, (b) demand capacity begins to decline towards the baseline load level and electricity price are low.

Contrarily, through incentive programmes for *MD* regulatory, Prosumer can gain compensated operating costs despite ordering of electricity from Retailers when tariff is at its highest. Compared to conventional EM proceedings, it focuses on shifting demand loads into spotted regions where tariffs are it lowest or depriving energy consumptions based on local SHSES generation availability. Practically, such transactions will not be pragmatic as Prosumer will be restricted with their daily electricity usage and peak demand crises will still surface due to consumer's social lifestyles. Likewise, in view of SHSES penetrations, it rises grieving concerns for DSO to schedule excess generation feeding back to the grid when dealing with multiple NG operations. Therefrom, introducing *max MD* constraint can facilitate safe demand-side responses where violator will be charged with higher penalty premium and portion it out to abider as a reward/incentive. To comprehend such proceedings, Table 7 illustrates operating costs differences corresponding to *max MD* constraint.

Supplementarily, decision in determining ideal energy storage size can influence NG's operating costs. Referring to Fig. 11c, instances where acquisition of electricity from Utility for charging energy storage is not deterministically cheap. Moreover, bidding strategy between Agents gets complicated when accrediting energy storage into taking capacity size constraint into consideration. Such crises can easily be determined through monitoring Agent's cooperative reward assignments. Instances when IA Agent was forced to exhaust energy storage during low electricity prices to propel successive PV penetrations at time $(t+1)$ or limit its discharge rate when servicing peak demand load. Likewise, considerations of levelised costs affiliated to PV and energy storage need to be monitored before selling back to the Utility. Therefore, strategic sizing of energy storage capacity is vital in refining return investment profit margin and suppressing other plausible grid-tied violations as shown in Table 7.

Fig. 12a illustrates responses of Agents' Q-Learning probabilistic regression in securing NG's optimal operating costs

TABLE 7. Operating cost comparisons based on 91kWh energy consumption.

Constraints	24hrs Operating Costs (Inclu. LCOE) (≈$, Std. Dev.<0.087)
Propose $max\ MD$	10.108
W/O $max\ MD$	8.776
W/O $max\ MD$ W/ Penalty	9.918
Propose W/ $max\ MD$ W/ Incentives	9.519
Δ Pshift > 250kWh W/ Penalty	10.203
Δ Pshift < 150kWh W/ Incentive	9.931
Proposed Δ Pshift W/ Incentive	9.702
Propose ES Size: 2.2kWh (LCOE: $1.20e-4)	9.479
ES Size: 3.0kWh (LCOE: $1.636e-4)	9.221
ES Size: 3.8kWh (LCOE: $2.072e-4)	9.914
ES Size: 2.0kWh (LCOE: $1.090e-4)	9.765

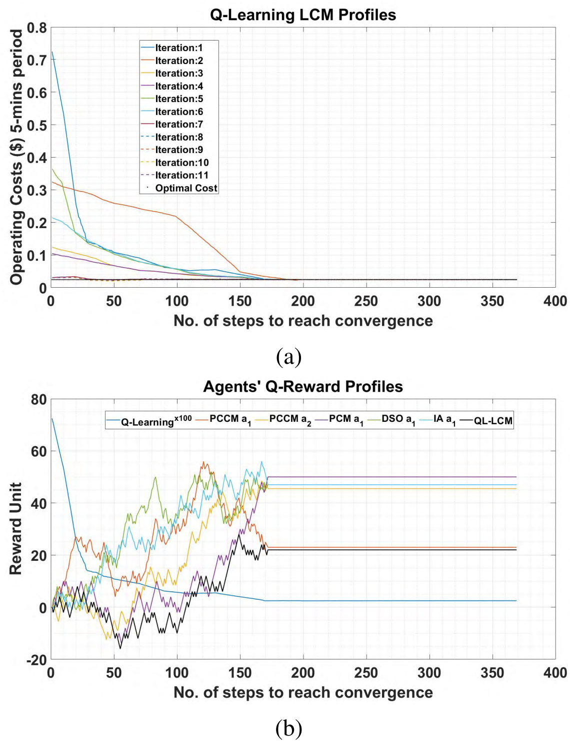


FIGURE 12. Agents' Q-learning regression and reward profiles. (a) Inspecting Q-learning convergence of agents' in LCM with multiple iterations. (b) Inspecting agents' reward based on individualism against cooperative tendency.

TABLE 8. Parameters for Q-learning.

Constants	Tuned values
ϵ	0.25
α	0.76
γ	0.9

Agent Rewarding System
Initialise each Agent & QL-LCM Reward = 0;
IF:Agent satisfy constraints = +2;
ELSE IF:Agent fails to satisfy constraints = -3;
ELSE IF:Participating Agents converged a solution)
-Agents complied Cooperative Tendency in LCM = +2;)
-Agents neglected Cooperative Tendency in LCM = 0;)
ELSE: Reward = 0;

in a cooperative tendency domain based on the predefined functions shown in Table 8. During Agents' auctioning processes, penalty issued for respective state-action pair that gratify cooperative tendency is much lower than sole EM administrations for a single NG. Such penalty resolution foster Agents to advance into cooperative tendency to meet global objective functions after prior adjustments made during sole EM proceedings. The Agent's learning regression behaves within a discrete action spaces bounded by a time-step of every 5mins intervals. It can be seen that all learning gradient at any iteration set was able to reach convergence with a standard deviation of less than 0.4, locating uniform global minima for the system's operating costs at different time stamps. The learning search space improves proportionately to the number of iterations executed at time t by preceding lesser steps to locate optimal solution. Thus, to refrain from computational intensity and time in recognising solution-optimality, it is recommended to run higher order of iteration ($>7^{th}$).

On the contrary, Fig. 12b presents the Agents' payoffs during cooperative tendency. The payoff absorbs positive and negative rewards due to unseen prior knowledge when pairing action-state against desirable consequences. The reason behind employing a negative payoff and assigning smaller positive reward is to give Agents a clear comprehension of the actions performed in a state which accelerate Agents' learning process in gaining desirable actions and computation convergence. In the case of poor decisions were made by Agents', respective Q-values will drastically decrease thus alerting Agents' to refrain from performing previous action in the future. Similarly, a sequential rewarding system was introduced for cooperative tendency of Agents. However, enforcement of negative reward on cooperative tendency was not implemented as to provide learning relaxation on Agents to gain convergence. It allows Agents to prioritize individual objective which is important for NG operations before heading forward in earning higher cooperative rewards for cooperative tendency during Agents' bidding process. Nevertheless, no reward is awarded if Agents' solution does benefit the operations of all interconnected Agents.

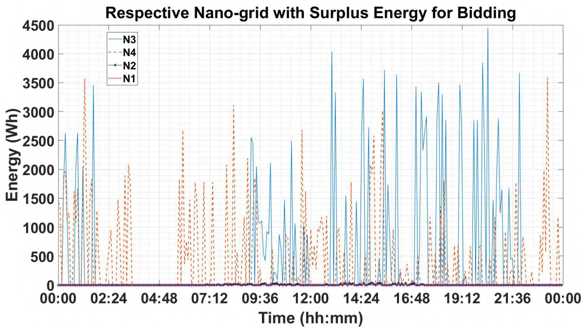
B. EM OPERATIONS OF AGGREGATED NANOGRID SYSTEMS (COMMERCIAL & RESIDENTIAL)

In this section, operations of the four aggregated NGs, $NG1$-$NG4$, as shown in Fig. 14 were deployed on-line to view performances of cooperative EM strategies on energy trading. The Agents' learning search space is now broaden up, extending respective nano-biased energy manger serviceability to other neighbouring NGs. Thus, dependency on electricity pool market for DSO and separate PCM Agents are no longer restricted as compared to a single-bounded EM operations. The role of PCM Agent is to interact closely with IA Agent in recognising surplus energy capacity excluding reserve pool available for disposal at time $(t + 1)$. Simultaneously, generating a price model for spinning reserve market.

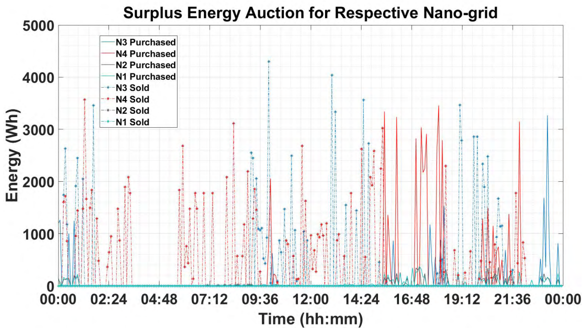
The electricity tariff oriented PCM Agents is separated into two authorities; i) classical marginal pricing (MP) determined by DSO Agent during any electrification transactions occurs between NGs and Retailers, ii) bid-as-request (BAR) pricing where forward bilateral contracts are negotiated between Prosumers. In this paper, considerations of in-depth compliance in the electricity pool policies involving competing regulations and codes of practice for Market Rules were omitted when formulating the problem statements. It assumed that the open electricity market is stable and the focus is directed on trading energy either between NG to another (PCM-to-PCM) or NG to DSO (PCM-to-DSO). With that, PCM Agents coordinate its involvements in the bidding market at time $(t + 1)$ based on the defined policies, strategically allot available energy for trading or disengages its involvements for any transactions.

The analytic studies were mainly focused on highlighting beneficiary attained in transacting superior operating costs with demand-side management and impeding Duck Curve phenomenon at low-voltage level. Fig. 13 exhibits available surplus energy ready to be auctioned at time t between NG-to-NG (limited only to commercial building) or NG-to-DSO. Fig. 13a depicts the surplus energy capacities generated at individual NG system. Fig. 13b displays successful bidding strategies directed between NGs or resorted in selling back to the Utility through DSO Agent. Finally, Fig. 13c profiles the Utility's energy generation managed by DSO where its constraints were bounded by a baseline load of 8.2kWh, $maxMD$ capped at 62.8kWh, and preserved $\Delta P_{Deviate}$ less than 3.5kWh.

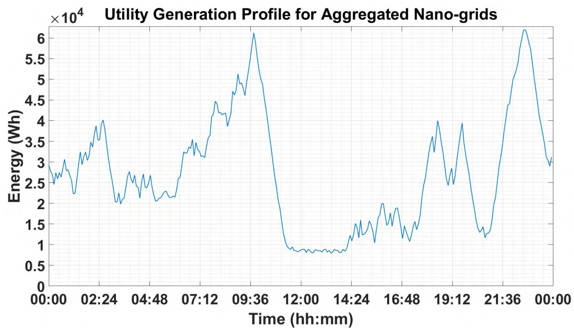
In view of the electricity pooling at respective NG, diversified bilateral trading models were introduced to inspect profit margin earned during bidding transactions. $NG3$ exploits a discharge scheduling strategy that services peak demand episodes where electricity tariff is expected to be high, taking its dominance by auctioning a marginal cheaper price that conclusively lured PCM Agents' of other NGs to divert its purchase away from Retailers. Whereas, $NG4$ prioritizes in suppressing Duck Curve phenomenon which benefits demand-side response from confronting any possible load balancing violations and relieve large deviation stress on



(a)



(b)



(c)

FIGURE 13. **Agents' bidding strategies on exploiting available surplus energies in aggregated operations. (a) Surplus energy capacity available for bidding. (b) Transactions of NG purchasing energy from each other and selling back to utility. (c) Energy generation profile at utility supplying the four aggregated NGs.**

generation units. It has a relaxed bidding strategy, auctioning its electricity at the lowest threshold rate in the hopes to break-even by wagering on incentives issued by Utility. Adversely, $NG2$ adopts oversized SHSES capacity accompanying with low electricity consumptions to increase electrification sustainability which averts itself from participating in the auctioning process. Contrarily, $NG1$ is a highly active bidder due to its undersized HSES and steep demand profile. At the bare minimal in governing the bilateral energy market, DSOs has implemented regulations that rule out possible monopolism on the bidding process. Residential(s) are restricted to perform direct bidding transactions with other residential(s) as to discourage home owners from setting large SHSES that may impose hazardous threads to the living community. Thus, DSO Agent will be assigned as a mediator when coordinating home-to-home energy trading. Whereas, commercial-building Prosumers are contracted to sell power level at *min-max* threshold of $< 1MW$ and $> 10MW$, asserting generation certainty for Utility to positively schedule its loading of generators to anticipate peak- and off-peak demands.

TABLE 9. **Profit margin comparisons between single-bounded and aggregated EM approach.**

	Single EM ($ per day)	Aggregated EM ($ per day)
Nanogrid 1	9.479	8.725
Incentives$_{Utility}$	0.482	0.351
LCOE losses	0.028	0.019
Nanogrid 2	9.257	8.873
Incentives$_{Utility}$	0.385	0.401
LCOE losses	0.244	0.181
Nanogrid 3	88.646	86.158
Incentives$_{Utility}$	1.273	2.706
LCOE losses	1.599	1.233
Nanogrid 4	85.691	85.829
Incentives$_{Utility}$	2.572	3.932
LCOE losses	1.828	4.720

Consequently, Table 9 gives a comparative analyses on the profit margins earned at respective NG based on assigned bidding strategy. $NG3$ has enhance its operating cost by profiting highly from other bidders during the trading process while $NG4$ suffered a small marginal loss based on a small-scaled incentives fixed at $\$0.03kWh$ across all individual services. $NG2$ managed to breaks even based on limited contributions in the bidding process to compensate induced LCOE from its SHSES while $NG1$ gained the highest profit margin in the aggregated system encompassing all grid-tied incentives and successfully bid for cheaper electricity tariff.

VI. COMPARISONS WITH OTHER METHODOLOGIES

In this section, evaluations of the proposed methodology against two others were presented to view respective superiority in addressing EM optimality for aggregated NG operations. Indeed, respective authors have unique EM models depending on the unparalleled grid components when formulating problem statements. However, they shared a common aspirations in trimming system's operation costs from strategic utilisation of local SHSES. Hoping to absorb profiting incentives from the energy market during electricity trading without violating any power system operation requirements in real-time. All proposed controllers were appraised against the 4 evaluation indices: Operation Costs Consistency (OCC), Penalty Costs imposed on grid participant (PC), Load Shifting Factor (LSF) and Market Clearing Incentives (MCI).

$$OCC = \sqrt{\frac{\sum_{i=1}^{10}(cost(t,i) - \frac{\sum_{i=1}^{10} cost(t,i)}{10})^2}{10}} \qquad (55)$$

$$PC = \frac{\sum_{i=1}^{10}(Penalty_{grid} + LCOE_{loss})(t,i)}{10} \qquad (56)$$

$$LSF = \min(\frac{shifted_{load}}{total_{load}}(t_{1hr}, 1), \ldots, \frac{shifted_{load}}{total_{load}}(t_{1hr}, 10)) * 100\% \quad (57)$$

$$MCI = \frac{\sum_{i=1}^{10}(Profit_{bid} + Incentive_{grid})(t, i)}{10} \quad (58)$$

i denotes the number of iterations performed by the discrete computation at time, t to view solution consistency, OCC calculates the standard deviation of the operating costs at time t, PC estimates the average penalty costs imposed by Utility at time t, LSF quantifies the amount of load shifted at time t_{1hr} for a duration of an hour where smaller percentage depicts greater conservation of consumer(s) lifestyle, and MCI computes the profit margin or incentives gained at time t (combination of Utility and levelised revenue from SHSES).

A. HIERARCHICAL ENERGY MANAGEMENT SYSTEM

Author Tian *et al.* [32] proposed a hierarchical EM system (HEMS) that performs two-level hierarchical optimisation method for Micro-grid community operations at distribution level. The proposed hierarchical optimisation statement is separated into two phases: (a) First stage optimisation run in the lower level EM in a scheduled period, forecasting a day-ahead of SHSES output power, charge/discharge powers of the energy storage and load demands in time intervals basis of $1hr$. Furthermore, the optimal exchanged power values between an individual Micro-grid and the upstream distribution network were deduced. (b) From the results, second stage optimisation uses linear mixed-integer programming (LMIP) is deployed to solve multi objective functions defined in eq. (16), (20), (27)-(28), (31), (35), (37) and (43) with its corresponding constraints.

Implicit logic constraints is added into the problem formulation modelling to gain greater linearized approximation.

B. SCENARIO-BASED STOCHASTIC ENERGY MANAGEMENT SYSTEM

Author Shen *et al.* [33] proposed a scenario-based stochastic EM system (SSEMS) for Micro-grid operation. It takes electricity pool market into consideration when scheduling its controllable loads to maximize profit margin. The proposed algorithm uses two level stochastic optimisation methods to address uncertainties and risk-constrained elements procured from the integrated SHSES. The first level attains information from the economic operation scheme based on the forecasted data using deviation compensator method. The second level provide solutions in scheduling the controllable units based on real-time data using Monte Carlo scenario-based. Comprehensively, to constrain the risk in misinterpreting profit margin, risk management is fused into the objective function using conditional value.

As proposed by the author in formulating the maximum profit operations, the objective function defined in (37) will be replaced in accordance to the paper while the remaining formulations remained same.

TABLE 10. Comparing different EM methodologies against proposed.

	Proposed			
	Nanogrid 1	Nanogrid 2	Nanogrid 4	Nanogrid 3
Operating Costs;OCC	\$8.725; 1.591	\$8.873; 1.858	\$85.829; 2.267	\$86.158; 1.772
PC;OCC	\$0.019;0.661	\$0.181;0.348	\$4.720;0.369	\$1.233;0.535
LSF	5.629%	4.138%	5.038%	3.812%
MCI;OCC	\$0.773;1.032	\$0.565;0.579	\$4.582;1.265	\$3.721;1.186
	HEMS [20]			
	Nanogrid 1	Nanogrid 2	Nanogrid 4	Nanogrid 3
Operating Costs;OCC	\$8.943; 4.104	\$9.004; 3.726	\$86.117; 3.883	\$86.749; 3.274
PC;OCC	\$0.397;3.161	\$0.541;2.926	\$4.828;3.374	\$5.162;3.145
LSF	21.736%	15.288%	19.638%	16.825%
MCI;OCC	\$0.894;2.590	\$0.726;2.342	\$3.638;2.689	\$4.221;2.911
	SSEMS [21]			
	Nanogrid 1	Nanogrid 2	Nanogrid 4	Nanogrid 3
Operating Costs;OCC	\$8.842; 1.240	\$9.047; 1.536	\$86.265; 1.219	\$86.769; 1.681
PC;OCC	\$0.054;0.728	\$0.318;0.819	\$4.669;0.743	\$1.382;0.725
LSF	6.522%	4.614%	4.895%	4.254%
MCI;OCC	\$0.704;0.811	\$0.497;0.725	\$3.884;0.870	\$3.176;1.132

C. PERFORMANCE AND RESULT EVALUATIONS

Analytical results comparing EM performances rendered by HEMS, SSEMS and proposed are presented in Table 10.

The results attained from employing HEMS had shown detrimental impacts on the overall operating costs due to the weak resolution in solving functions with multi-constraints in a discrete-time approach when using LMIP. Constraints involving LSF has shown great deficiency as the algorithm focuses deeply on elevating operating costs at the expense of shedding shiftable loads. It fails to comprehend strategical solutions in rescheduling within the time duration boundary ($t = 1hr$) which ultimately causes shiftable load capacity to recede. In addition, the proposed LMIP lacks in cooperative optimisation causing penalty imposed by Utility to inflate, failure to procure cohesive electricity consumptions which resultant to large $\Delta E^{Deviate}$ induced at Utility.As a result, Duck Curve phenomenon becomes conspicuous which eventually cascaded to other grid-tied violations. Nevertheless, given with such problem formulation, such episodes can be prevented by introducing cooperative strategies into LMIP as suggested in [34] where the author infused model predictive control to address overall energy system using rewarding schema.

In contrast to SSEMS, the results attained are much comparable due to its large search space in generating numerous number of scenarios representing uncertain parameters. It uses Latin hypercube sampling technique to reduce algebraic computation time without affecting accuracy of the optimized results. However, with the suggested risk management model, it heavily penalized the electricity price market to compensate weak forecasting of uncertainties under the normal distribution curve. Hence, proposed technique guarantees the profit margin figure despite succumbing to variability.

VII. CONCLUSIONS

This paper presents a reformed nano-biased EM that resolves future's predicaments in steering high penetrations of SHSES at low-voltage distribution level. It avows Prosumer sovereignty in gaining greater dominance at demand-side with DSO-Retailers adequately supporting their preferences in terms of connection, quality, security, and continuity of power supply. Such proceeding enables managerial perspective to be more apparent and habitable, relieving DSO-TSO from orchestrating nano-managing roles for individual Consumer(s) and focus more on relaxing the level-playing electricity market and establish congestion management.

The proposed EM methodology was designed to service each successive NG as a single virtual power plant aggregator that adopts layered coalition model to manage interdependent EM platform. The algorithm hybridises ELM forecasting technique with cooperative reinforced learning Agents in MAN to successfully comprehend and schedule uncertainties transpired during NG operations. These involve trading and bidding of electricity in the energy market while shifting non-critical demand loads to secure optimal operating costs at each time step based on collegial discrete solution. Indeed, the results have shown positive and comparable EM performances for both DSO-TSO and Prosumer(s) when benchmarked with other methodologies using the proposed aggregated NGs network.

However, the proposed EM operations can arise potential monopolism in the energy market during bidding of electricity. Prosumer(s) tend to model preferable energy trading requirements and yet render a balanced energy system which leads DSO-TSO astray from ill-defined competition rules. Electricity prices at Customer-end will start to inflate in proportional to the size of active Prosumers with online SHSES on the balancing market. On the contrary, tendency of Prosumer(s) shifting into the paradigm where they could be regarded as Retailers are plausible as DSOs are blinded by the safe power system analysis proceeding. The importance of Retailers fades away along with the relationship between DSO-TSO. Therefore, for future research topics, revise Market Rule implementing new policies and regulatory are essential to ensure a transparent and competitive trading environment are preserved based on an economical point of view for all energy players.

A typical reinforced learning model (RLM) is used in this paper to determine Agent's optimal reactions given in a state environment using Markov Decision Process. Influenced by awarding rewards for each executed actions, Agents' are believed to be directed towards optimality until it reaches convergence. The RLM set-up is defined by the 5-tuple (S,A,P,R,γ) where S is a set of states while A describes the set of actions. P denotes the state transition probability and R designates the reward, $R : SXA \longrightarrow R$. Lastly, γ ($\gamma \in (0, 1]$) aids in tracking which action that had procured the optimal policy function. It is tasked to learn the policy function $\pi : S \longrightarrow A$ that maps from states to actions and search for optimal policy that has the maximized sum of rewards. For an example in load balancing problem:

State: retrieve online load capacity.
Actions: Generation sources tuned their output capacity (increment or decrement).
Reward: 1 if supply = demand, 0 otherwise.

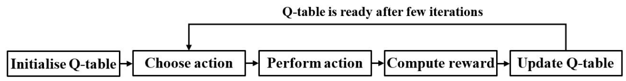


FIGURE 14. Q-learning algorithm process in composing and updating the look-up table.

TABLE 11. Q-learning algorithm for energy management.

Initialise Q(S,A),∀S, ∀A with a total discounted reward of 0
Initialise learning rate, $\alpha, \varepsilon, \gamma$
for each time step t
  for iteration,i less than X
    explore possible action set
    attain greedy action
    select action from policy π
    take action and observe reward
    update Q($S_{t,i}$,$A_{t,i}$)
    t,i $\longleftarrow$ t,(i+1)
    $S_{t,i} \longleftarrow S_{t,(i+1)}$
  end for iteration,i
end for time step,t

APPENDIX A

A. REINFORCED LEARNING OPTIMISATION

The optimal policy function, π^* can be calculated:

$$\pi^* = \arg\max_{\pi} \mathbb{E}\Big[\sum_{t\geq 0} \gamma^t R_t | \pi\Big] \tag{59}$$

To gain optimal policy, it employs Q-Learning algorithm that learns function of state-action pairs. Q-learning is basically a lookup table that calculates the maximum expected reward for action at each state, searching for the best action at each state. The Q-function exploits Bellman equation that has two input variables, state (S) and action (A):

$$Q^{\pi}(S_t, A_t) = \mathbb{E}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots | S_t, A_t] \tag{60}$$

Using (11)-(15) the Q results will be assigned into the cells of the table based on (22), (26), (44), and (54) learning payoffs incorporation with respective constraints. The Q values will be continuously updated based on the iterations perform given in a single time frame using the Q-learning process illustrated in Fig. 14 and Table 11. During the process of Agent exploring possible actions and states, an epsilon greedy strategy will be employed to size the exploration environment as shown in (11). In the initial computation, large epsilon rates are used giving Agents the freedom to choose actions randomly due to uncertainties. After few iterations, the epsilon rate will be decreased progressively as the confidence level increases in estimating the Q-values.

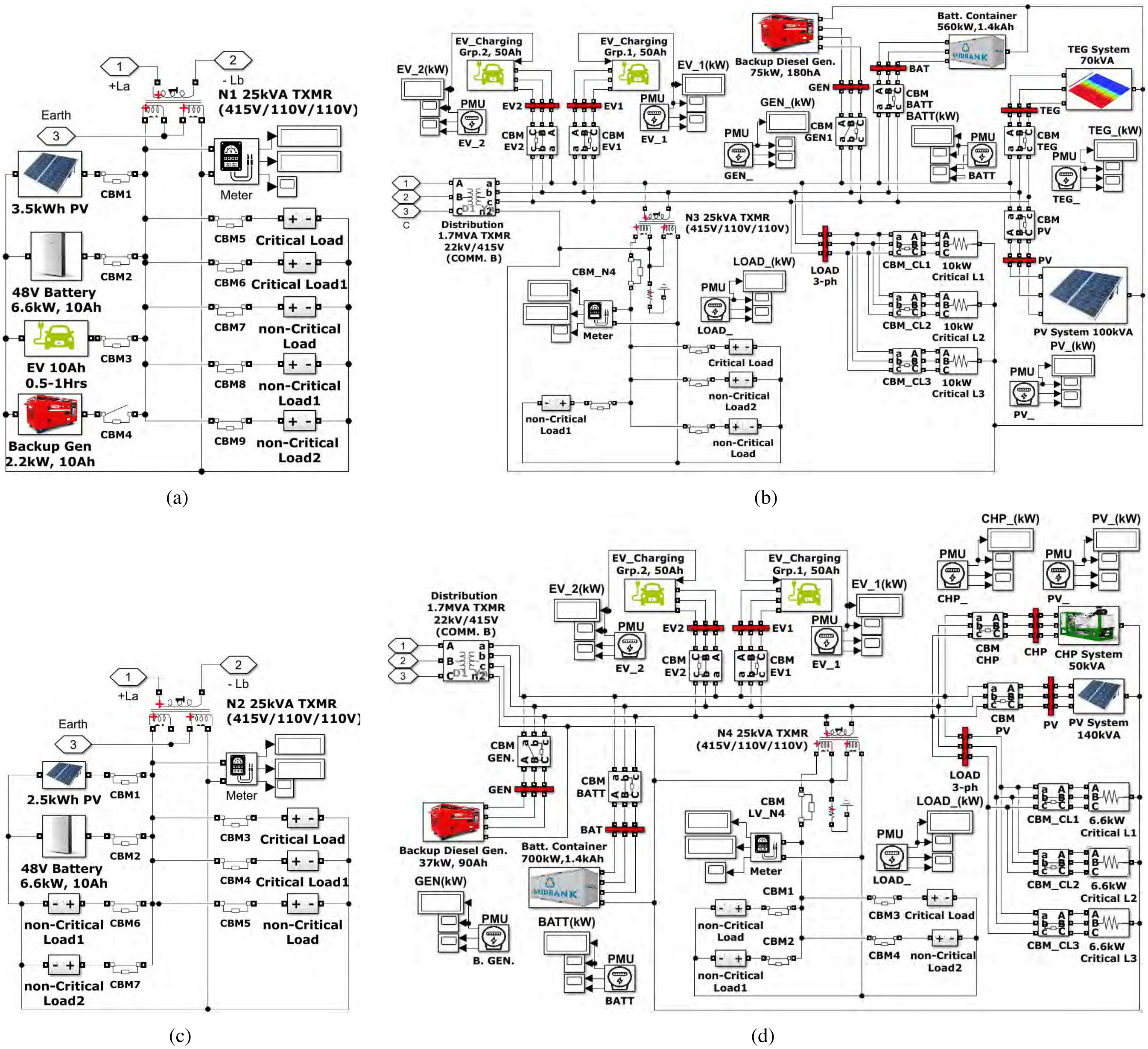


FIGURE 15. Proposed nanogrid networks coupled to 22*kVAC* medium-voltage utility grid. (a) *NG*1, 220*VAC* single-phase system for residential. (b) *NG*3, 415*VAC* three-phase system for commercial building (hospital). (c) *NG*2, 220*VAC* single-phase system for residential. (d) *NG*4, 415*VAC* three-phase system for commercial building (industry).

B. NANOGRID 1 TO 4 MODELS

All simulation analyses were based on the proposed NG models created in MATLAB shown in Fig. 15. It represents respective Nanogrid electrical systems for both residential (*NG*1, *NG*2) and buildings (*NG*3 Commercial, *NG*4 Industrial).

REFERENCES

[1] A. Kumar, A. R. Singh, Y. Deng, X. He, P. Kumar, and R. C. Bansal, "A novel methodological framework for the design of sustainable rural microgrid for developing nations," *IEEE Access*, vol. 6, pp. 24925–24951, 2018.

[2] T. Logenthiran, D. Srinivasan, and T. Z. Shun, "Demand side management in smart grid using heuristic optimization," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1244–1252, Sep. 2012.

[3] T. Morstyn, B. Hredzak, and V. G. Agelidis, "Control strategies for microgrids with distributed energy storage systems: An overview," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3652–3666, Jul. 2018.

[4] M. R. B. M. Saifuddin, T. Logenthiran, R. T. Naayagi, and W. Woo, "Apprehending fault crises for an autogenous nanogrid system: Sustainable buildings," *IEEE Syst. J.*, to be published.

[5] Y. Zhao, J. Yu, M. Ban, Y. Liu, and Z. Li, "Privacy-preserving economic dispatch for an active distribution network with multiple networked microgrids," *IEEE Access*, vol. 6, pp. 38802–38819, 2018.

[6] I. Miranda, H. Leite, and N. Silva, "Coordination of multifunctional distributed energy storage systems in distribution networks," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 3, pp. 726–735, Feb. 2016.

[7] A. Habib, A. Arshad, and R. Khan, "Distributed renewable energy under the guidance of price autonomous operation technology," *Smart Grid Renew. Energy*, vol. 8, no. 10, pp. 305–324, 2017.

[8] P. Almeida, L. Kane, M. Collins, C. Breaden, E. Davidson, and G. Ault, "Implementing optimisation functionality on network management platforms for new DSO business models," *CIRED, Open Access Proc. J.*, vol. 2017, no. 1, pp. 2779–2782, 2017.
[9] A. Saint-Pierre and P. Mancarella, "Active distribution system management: A dual-horizon scheduling framework for DSO/TSO interface under uncertainty," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2186–2197, Sep. 2017.
[10] F. Pilo, G. Mauri, B. Bak-Jensen, E. Kämpf, J. Taylor, and F. Silvestro, "Control and automation functions at the TSO and DSO interface–impact on network planning," *CIRED, Open Access Proc. J.*, vol. 2017, no. 1, pp. 2188–2191, 2017.
[11] S. Huang, Q. Wu, L. Cheng, Z. Liu, and H. Zhao, "Uncertainty management of dynamic tariff method for congestion management in distribution networks," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4340–4347, Nov. 2016.
[12] A. Smith and A. Stirling, "Innovation, sustainability and democracy: An analysis of grassroots contributions," *J. Self-Governance Manage. Econ.*, vol. 6, no. 1, pp. 64–97, 2018.
[13] E. Harmon, U. Ozgur, M. H. Cintuglu, R. de Azevedo, K. Akkaya, and O. A. Mohammed, "The Internet of microgrids: A cloud-based framework for wide area networked microgrids," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 1262–1274, Mar. 2018.
[14] M. R. B. M. Saifuddin, T. Logenthiran, R. T. Naayagi, and W. L. Woo, "Apprehending fault crises for an autogenous nanogrid system: Sustainable buildings," *IEEE Syst. J.*, to be published.
[15] S. Shoaib, X. Chen, and I. Llewellyn, "Increasing the effective range of smart meter home area network," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2898–2901, 2017.
[16] R. Ma, H.-H. Chen, and W. Meng, "Dynamic spectrum sharing for the coexistence of smart utility networks and wlans in smart grid communications," *IEEE Netw.*, vol. 31, no. 1, pp. 88–96, Jan./Feb. 2017.
[17] W. Qingyao, T. Mingkui, S. Hengjie, C. Jian, and M. K. Ng, "ML-FOREST: A multi-label tree ensemble method for multi-label classification," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 10, pp. 2665–2680, Oct. 2016.
[18] A. Alkandari, A. Sami, and A. Sami, "Proposed DSO ancillary service processes considering smart grid requirements," *CIRED, Open Access Proc. J.*, vol. 2017, no. 1, pp. 2846–2847, 2017.
[19] V. Nunes, J. P. Gouveia, A. M. Rodrigues, and T. Simão, "INSMART–towards the new distribution systems operators potential roles in low carbon future and integrated frameworks for smart cities," *CIRED, Open Access Proc. J.*, vol. 2017, no. 1, pp. 2797–2799, 2017.
[20] Z. Zhang, W. Qiao, and Q. Hui, "Power system stabilization using energy-dissipating hybrid control," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 215–224, Jan. 2019.
[21] M. Li, C. Zhang, Y. Wang, and X. Dong, "VQ sensitivity analysis considering power exchange for multi-area interconnection power networks," *DEStech Trans. Environ., Energy Earth Sci.*, 2018.
[22] Y. Han, H. Li, P. Shen, E. A. A. Coelho, and J. M. Guerrero, "Review of active and reactive power sharing strategies in hierarchical controlled microgrids," *IEEE Trans. Power Electron.*, vol. 32, no. 3, pp. 2427–2451, Mar. 2017.
[23] A. Qasim and S. A. R. Kazmi, "MAPE-K interfaces for formal modeling of real-time self-adaptive multi-agent systems," *IEEE Access*, vol. 4, pp. 4946–4958, 2016.
[24] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Raheem, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE Access*, vol. 6, pp. 33789–33795, 2018.
[25] J. Tang, F. Ni, F. Ponci, and A. Monti, "Dimension-adaptive sparse grid interpolation for uncertainty quantification in modern power systems: Probabilistic power flow," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 907–919, Mar. 2016.
[26] F. L. D. Silva, R. Glatt, and A. H. R. Costa, "MOO-MDP: An object-oriented representation for cooperative multiagent reinforcement learning," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 567–579, Feb. 2017.
[27] Z. Zhang, D. Wang, D. Zhao, Q. Han, and T. Song, "A gradient-based reinforcement learning algorithm for multiple cooperative agents," *IEEE Access*, vol. 6, pp. 70223–70235, 2018.
[28] W. Liu, P. Zhuang, H. Liang, J. Peng, and Z. Huang, "Distributed economic dispatch in microgrids based on cooperative reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2192–2203, Jun. 2018.
[29] T.-T. Wong and N.-Y. Yang, "Dependency Analysis of Accuracy Estimates in k-Fold Cross Validation," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 11, pp. 2417–2427, Nov. 2017.
[30] N. D. Doulamis, A. D. Doulamis, and E. Varvarigos, "Virtual associations of prosumers for smart energy networks under a renewable split market," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6069–6083, Nov. 2018.
[31] T. Georgitsioti, N. Pearsall, and I. Forbes, "Simplified levelised cost of the domestic photovoltaic energy in the UK: The importance of the feed-in tariff scheme," *IET Renew. Power Gener.*, vol. 8, no. 5, pp. 451–458, Jul. 2014.
[32] P. Tian, X. Xiao, K. Wang, and R. Ding, "A hierarchical energy management system based on hierarchical optimization for microgrid community economic operation," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2230–2241, Sep. 2016.
[33] J. S. Shen, C. Jiang, Y. Liu, and X. Wang, "A microgrid energy management system and risk management under an electricity market environment," *IEEE Access*, vol. 4, pp. 2349–2356, Apr. 2016.
[34] A. Parisio, C. Wiezorek, T. Kyntäjä, J. Elo, K. Strunz, and K. H. Johansson, "Cooperative MPC-based energy management for networked microgrids," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 3066–3074, Nov. 2017.

MUHAMMAD RAMADAN BIN MOHAMAD SAIFUDDIN (GS'16) received the B.Eng. degree (Hons.) in electrical power engineering from the Singapore Institute of Technology, Newcastle University, in 2015. He is currently pursuing the Ph.D. degree in electrical and electronics engineering with Newcastle University, U.K. He published several internationally referred journal paper and conference papers in relations to electrification advancements at low-voltage level: prosumer dominion and sovereignty. His research interests include prosumer-centric energy managements, design solid state-based power converters for nanotech generation (renewable), in-house fault diagnosis systems, and post-fault restoration proceedings. He is a Graduate Student Member of the IEEE Power and Energy Society. He is currently serving as a Committee Member for the IEEE PES Singapore Chapter.

THILLAINATHAN LOGENTHIRAN (S'07–A'10–M'12–SM'16) received the B.Sc. degree in electrical and electronic engineering from the University of Peradeniya, Sri Lanka, in 2005, and the Ph.D. degree in electrical engineering from the National University of Singapore, Singapore, in 2012. He is currently an Assistant Professor with the School of the Electrical Power Engineering, University of Washington Tacoma, Tacoma, USA. He published more than 100 internationally referred journal papers, book chapters, and conference papers on power system engineering and applications of computational intelligent techniques. His research interests include smart grid, micro grid, renewable energy resources, optimization, applications of computational intelligent techniques & intelligent multi-agent system for power system operation and its optimization. He is a member of the IET and an Associate Fellow of the Higher Education Academy, U.K. He received several awards and scholarships, including the First Prize Award in Siemens Smart Grid Innovation Contest and the Newcastle Teaching Award.

R. T. NAAYAGI (M'12–SM'15) received the bachelor's degree (Hons.) in electrical and electronics engineering from Bharathidasan University, Tiruchirappalli, India, in 2000, the master's degree in information technology from Alagappa University, Karaikudi, India, in 2003, the master's degree (Hons.) in power electronics and drives from Anna University, Chennai, India, in 2005, and the Ph.D. degree in electrical and electronic engineering from The University of Manchester, Manchester, U.K., in 2010.

She held various roles as a Lecturer, a Senior Lecturer, and a Professor in various institutions in India and U.K., from 2000 to 2006 and from 2010 to 2014. She has been with Newcastle University in Singapore (NUiS), since 2014, where she is currently an Associate Professor in electrical power engineering. She is a member of the Diversity Working Group and an Academic Lead for Athena SWAN Bronze Award Application with Newcastle University, U.K., and has been organizing many events to promote young professionals especially young women in engineering and technology. Her research interests include renewable energy integration and applications in smart grid, power electronics for aerospace, electric vehicle applications, low carbon electrical energy systems, and power electronic solutions to sustainability.

Dr. Naayagi received a Senior Fellow status of the Higher Education Academy, U.K., in 2016 and 2019. She is a Steering Committee Member of NU Women, U.K., and a member of the Diversity Working Group, Newcastle University, U.K., Equality & Diversity Champion for NUiS, and an Academic Lead for Athena SWAN Bronze Award Application at Newcastle University. She has received several merit certificates for her academic proficiency, including the Best Outgoing Female Graduate Award during her bachelor's and the Outstanding Master's Student Award. She was a first recipient of the Dorothy Hodgkin Postgraduate Award from the School of Electrical and Electronic Engineering, The University of Manchester, for her Ph.D., jointly sponsored by Rolls-Royce plc and the Engineering and Physical Sciences Research Council, U.K. She was also a recipient of the Woman Engineer Award from the Young Professionals Section Chennai, Institution of Engineering and Technology, U.K., in 2012, and the Newcastle University Teaching Award, in 2016. She received the IET Woman Engineer Award from the Young Professionals Section Chennai, and the Institution of Engineering and Technology, U.K., in 2012. She received the Newcastle University Teaching Award, in 2016. She is the Chair of NUiS women in science and engineering network and has been organizing many events to promote young professionals especially young women in engineering and technology. She has received several merit certificates for her academic proficiency, including the Best Outgoing Female Graduate Award during her bachelor's and the Outstanding Masters Student Award. She received the Dorothy Hodgkin Postgraduate Award from the School of Electrical and Electronic Engineering, The University of Manchester, for her Ph.D., jointly sponsored by Rolls-Royce Plc and the Engineering and Physical Sciences Research Council, U.K. She is the Chairman of the IEEE Power and Energy Society, Singapore Chapter. She is an Associate Editor of *IET Power Electronics* journal and serves as a Reviewer for the IEEE, IET, and many other international journals and conferences.

WAI LOK WOO (M'11–SM'12) received the B.Eng. degree in electrical and electronics engineering, and the M.Sc. and Ph.D. degrees from Newcastle University, U.K., in 1993, 1995, and 1998, respectively. He was the Director of research for the Newcastle Research and Innovation Institute, and the Director of operations for Newcastle University. He is currently a Professor of machine learning with Northumbria University, U.K. He has published more than 350 papers on these topics on various journals and international conference proceedings. His research interests include the mathematical theory and algorithms for data science and analytics, artificial intelligence, machine learning, data mining, latent component analysis, multidimensional signal, and image processing. He is a Member of the Institution Engineering Technology. He was a recipient of the IEE Prize and the British Commonwealth Scholarship. He serves as an Associate Editor to several international signal processing journals, including *IET Signal Processing*, the *Journal of Computers*, and the *Journal of Electrical and Computer Engineering*.