A Probabilistic Reverse Power Flows Scenario Analysis Framework

Antonin Demazy, *Member, IEEE,* Tansu Alpcan, *Senior Member, IEEE,* and Iven Mareels *Fellow, IEEE.*

Distributed Energy Resources (DER), mainly residential solar PV, are embedded deep within the power distribution network and their adoption is fast increasing globally. As more customers participate, these power generation units cause Reverse Power Flow (RPF) at the edge of the grid, directed upstream into the network, thus violating one of the traditional design principles for power networks. The effects of a single residential solar PV system is negligible, but as the adoption by end-consumers increases to high percentages, the aggregated effect is no longer negligible and must be considered in the design and configuration of power networks. This paper proposes a framework that helps to predict the RPF intensity probability for any given scenario of DER penetration within the distribution network. The considered scenario parameters are the number and location of each residential DERs, their capacity and the daily net-load profiles. Classical simulation-based approach for this is not scalable as it relies on solving the load-flow equations for each individual scenario. The framework leverages machine learning techniques to make fast and precise RPF prediction within the network for each scenario. The framework enables the Distribution Network Service Providers (DNSPs) to assess DERs penetration scenarios at a granular level, derive and localise the RPF risks and assess the respective impacts on the installed assets for network planning purpose. The framework is illustrated with scenario analysis conducted on an IEEE 123 bus system and OpenDSS and shown that it can lead to multiple orders of magnitude savings in computational time while retaining an accuracy of 94% or above compared to classical brute force simulations.

***Index Terms*—Reverse Power Flows, Intermittent Energy Source, DER-rich penetration, Machine Learning, Power System Planning, Risk Management**

I. INTRODUCTION

RENEWABLE energy technologies and policies that encourage their adoption, affordability and customer behaviour are key factors contributing to an increase in residential adoption of Distributed Energy Resources (DER) globally. For instance, in its global roadmap to 2050, the International Renewable Energy Agency (IRENA) [1] predicts an increase of global electricity demand from 20204 TWh/yr in 2015 to 41508 TWh/yr in 2050, with a renewable share jumping from 24 % to 85% in the same period and where the solar PV capacity contribution would jump from 223 GW to 7122 GW. Predictions as to how this transition may evolve may vary, but the direction is clear and driven by a desire to decarbonise the electricity market.

Increasing the share of renewable Distributed Energy Resources (DERs) to significant levels impacts the power networks in multiple ways. The effect we consider in this paper is the problem of Reverse Power Flows (RPFs)[2]. The RPFs are currents generated from DERs that are travelling upstream in power networks that were originally designed to accommodate only downstream currents. The intensity of such currents depends on the penetration, spatial distribution and individual capacity of the DERs. The phenomenon is well known and the potential consequences are discussed in the literature[3], [4], [5].

As long as the DER generations are small compared to the actual demand in the network, the aggregated intensity of the RPFs remains negligible compared to the peak load current and ratings of the assets in place to distribute it. However, if the penetration and concentration of the DERs become significant with respect to local demand, the aggregated effect of RPFs will be significant, and affect the network exploitation, security and resilience[6]. In response, it will become essential to revisit the operations of the network, and eventually even the design, to maintain a safe and reliable distribution network. The penetration rate and pace can't be predicted with certainty and therefore can be viewed as a stochastic process. Additionally, the intermittent generation profile of each DER is also stochastic with wide seasonal and daily variations. In order to maintain security and reliability in their network for any future scenario of DERs adoption and at any time, the Distribution Network Service Providers (DNSPs) must perform an impact assessment for a vast number of scenarios. Classical simulation-based approach for this is not scalable as it relies on solving the load-flow equations for each individual scenario. We present a novel framework that uses Deep Neural Networks (DNNs) as a regression technique to estimate RPFs with high computational gain compared to traditional power system simulations without forsaking accuracy.

The approach as suggested here was constructed with the following properties in mind: it is fast, reliable and scalable to a large variety of scenarios for real size distribution networks; it can deal with all feasible scenarios of DER penetration and treat them in a probabilistic manner; it can deal with all kind of individual and aggregated DER daily generation profiles and it identifies the RPFs 'hotspots' within the network in terms of RPF's intensity and probability

The predominant feature is that the nonlinear power flow equations are not required to be solved for each scenario. In fact, the solutions of these equations are embedded in a well trained DNN model that can be executed at a fraction of the computational cost of solving the power flow equations themselves (at the expense of having solved the equations during a one-off training phase) and therefore enabling the analysis of vast amount of scenarios to derive resulting probabilistic RPFs intensities at every node. Our simulations are more than 1000 time faster than the equation solving while maintaining the accuracy above 94%.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OAJPE.2020.3032902, IEEE Open Access Journal of Power and Energy

A. Literature Review

The RPF is a well-known phenomenon and their potential consequences extensively described in the literature and technical reports [3], [4], [5]. In their summary of impacts, Walling et al [6] conclude that: 'Without careful engineering, DR penetration can potentially have adverse system effects, including exposing system and customer equipment to potential damage, decrease in power quality, decrease in reliability, extended time to restoration after an outage, and potential risks to public and worker safety'. Clearly, DNSPs must appropriately prepare and plan mitigation actions on their network prior to accommodate the rapidly increasing adoption of DERs amongst consumers. In [7], [8], the authors illustrate optimised planning models for distribution network with DERs under the assumption that DNSPs can decide the location and capacity of the DERs. In the case of consumer-driven expansion such as residential solar PV systems, DNSPs must be able to assess multiple scenarios and identify the resulting potential adverse effects on their network using traditional simulation tools or specifically designed tool and models, such as [9], [10]. Simulation tools are powerful but their utilisation of real scale power networks is computationally intense and therefore might restrict the capability to assess vast number of scenarios. In [11], [12], [13], [14], machine learning techniques have been experimented to perform power flow analysis of power networks with good performances. The advantage of those techniques is the modest computational cost required to use the models to solve the load flow once they are trained. In [15], the author proposes a data-driven probabilistic power flow analysis in response to stochastic solar irradiance.

Unlike previous approaches, we combine ML techniques with probabilistic analysis to infer the RPF intensity probabilities in response to a vast number of DER penetration and individual generation scenarios. Specifically, we use Deep Neural Networks (DNN) for regression purposes. The use of DNN substitutes the need for heavy computations required to solve the load flow equations for the entire system. DNN is a robust ML technique that is particularly suitable to deal with high-dimensional data [16] and the computations involved in using trained DNN are linear. Recently, in [17], the authors have explored and demonstrated the accuracy and computational advantage of Neural Networks to simulate power systems compared to conventional methods.

The next sections of the paper will illustrate the framework, discuss the methods employed to construct it and discuss the simulations conducted on the IEEE123 bus test system.

II. The probabilistic RPF Scenarios Assessment Framework

The proposed framework intends to enable a fast assessment of the RPF probability distribution in response to any future scenario of DERs penetration. The framework can manage large amount of scenarios (10^6), leading to orders of magnitude savings in computational time compared to traditional simulations while retaining appropriate accuracy. The framework identifies the 'hotspot' within the network in terms of location and probabilistic intensity of RPFs. The DNSP will leverage this information to assess the adequacy of their assets to support those RPF on a case by case basis. Those assessments will lead to identifying the most adequate reinforcement and/or protection measures that would maintain security and reliability in the network in response to the increased DER penetration scenarios.

As highlighted in the High-Penetration PV Integration Handbook [4], the RPF that are caused by non-negligible penetration of DERs such as residential solar PV embedded within the distribution network may have significant impacts on existing network's asset and their configuration. Amongst the RFP impacts, this paper will firstly consider the overload of assets (ampacity ratings of asset that are exceeded) due to aggregated peak generation values exceeding the aggregated peak load values for which the assets where originally dimensioned. Secondly, it will consider the system protection impact, that is a change in current levels, placement and coordination of protection devices to account for new sources of current that are embedded deep within the distribution network. This includes new potential islanding situation within the network that must be prevented.

The extent of those impacts would depend on the penetration and geographical concentration of installed DERs within the system. Those factors dictate the existence, intensity and probabilities of RPF at any node of the system. The adoption of solar PV over time (e.g. how many in total and where will they be located) is not strictly predictable and therefore has elements of long-term stochasticity. Additionally, the individual solar PV generation is intermittent throughout the year/day and therefore also stochastic in nature. We, therefore, propose to adopt a probabilistic approach to assessing the impacts on the network by RPF. Our proposed framework uses Monte Carlo simulations to derive probability distribution functions of the current at any node of the system for each penetration scenario, where the currents are predicted by a trained Deep Neural Network model (DNN) that models the solutions of the power flow equations for the network. The use of a DNN model to predict the currents time series in response to net-loads profiles is instrumental as it reduces significantly the computational burden of evaluating the currents in each scenario compared to solving the non-linear power flow equations. Fig.1 illustrates the end-to-end analysis framework proposed and subsection III-A illustrates the method used to train the DNN model.

Besides the grid model and configuration, the inputs to the analysis framework are the time window (eg.one full year) and the scenarios:

- Scenarios of DERs penetration, where each scenario is defined by number of DER connections, location of connections in the grid and capacity of each connected DER
- Time series scenarios of individual generation during the time window. This can be derived from solar irradiation statistics and differentiated by different geographical area of the network. Solar irradiation can significantly vary across seasons and can also be highly intermittent during the day. A vast amount of scenarios must be identified and considered to predict realistic probabilistic RPF intensities in the time window

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OAJPE.2020.3032902, IEEE Open Access Journal of Power and Energy

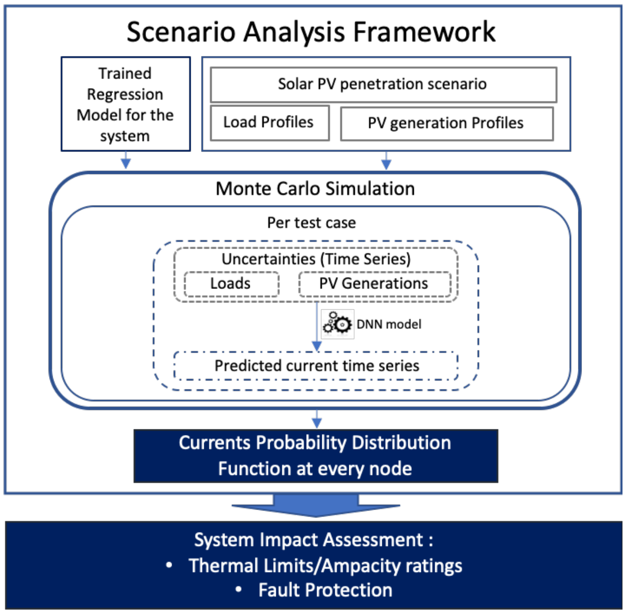


Fig. 1. DER's penetration scenario analysis framework

- Time series scenarios of individual loads during the time window. This can be derived from past data registered on the network, clustered by season where each sample of the time series is characterised by an average value and a standard deviation specific to the season
- In case of combined solar PV and storage capacity, the individual loads and generation time series would be adapted to account for the given battery charge-discharge profiles

The outputs of the scenario analysis framework are:

- Time series of predicted currents (intensity and direction) at every node of the power system for each DER penetration scenario in the selected time window. The granularity of the time series is dictated by the granularity of the daily load and generation profiles scenarios
- Aggregated probability distribution functions (pdf and cdf) of the currents at every node of the system for each scenario calculated on the selected time window

The framework enables to assess the impact of each scenario on the existing power network asset by asset. The type of impacts under consideration in this paper are:

1) The adequacy of assets ampacity rating by comparing the rating of the installed asset at any node with the predicted current probability distributions. Fig.2 illustrates a situation where an elevated concentration of high capacity residential solar PV can contribute to significant reverse currents (bigger than the peak load current) at the upstream transformer. This situation is also true at the substation level where all the RPF aggregate towards the transmission network.
2) The adequacy of existing system protection by comparing the settings of the devices installed with the predicted current probability distributions. The new current levels and direction (reverse power flows) might also affect the coordination rules and type of protection asset as discussed by Che et al in [18]. Fig.3 illustrates two types of malfunctions in protective devices (PD) that is caused by RPFs. In Fig.3a PD1 might not trip if I_{der} is large enough as it contributes to the fault current but doesn't transit via PD1. In Fig.3b, PD 2 might trip if I_{der} that contributes to the fault current is big enough and therefore unnecessarily disconnect feeder 2 from the grid (sympathetic tripping).

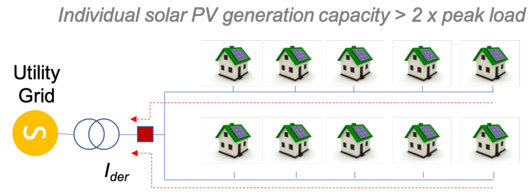
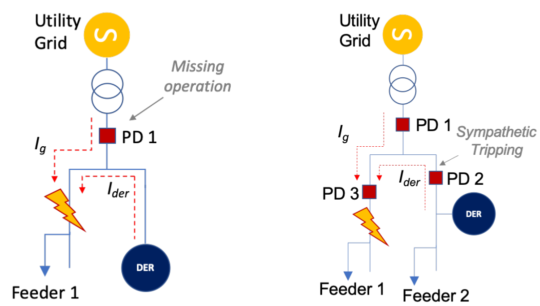


Fig. 2. Reverse power flows



(a) Malfunction of protective device (PD)

(b) Sympathetic tripping

Fig. 3. Example of DERs impact on fault detection

In both cases, the probabilistic scenario analysis provides valuable information regarding I_{der} and therefore assist the DNSPs to identify an adequate set of interventions on their asset to maintain the security and reliability of their network for each scenario.

The proposed framework is constructed upon the classic steady-state power flow analysis of a power system described by a set of $2n$ load-flow equations in $2n$ algebraic variables V_i, θ_i :

$$0 = -P_i + \sum_{k=1}^{n} |V_i||V_k|(G_{ik}\cos\theta_{ik} + B_{ik}\sin\theta_{ik}) \quad (1)$$

$$0 = -Q_i + \sum_{k=1}^{n} |V_i||V_k|(G_{ik}\sin\theta_{ik} + B_{ik}\cos\theta_{ik}) \quad (2)$$

where: n is the number of nodes in the system, P_i is the net injected real power (power generated minus power consumed) at node i, Q_i is the net injected reactive power at node i, V_i is the voltage at node i, θ_{ik} is the difference in voltage angle between node i and j, G_{ik} is the real part of the bus admittance matrix $[Y_{bus}]$ corresponding the i_{th} row and k_{th} column and B_{ik} is the imaginary part of the bus admittance

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OAJPE.2020.3032902, IEEE Open Access Journal of Power and Energy

matrix $\left[Y_{bus}\right]$ corresponding the i_{th} row and k_{th} column. The currents $\left[I\right]$ at every node are given by:

$$\left[I\right] = \left[Y_{bus}\right] \times \left[V\right] \tag{3}$$

For our analysis, the real (P_i) and reactive (Q_i) net-power at every load node will be known and given, while the currents $\left[I\right]$ are the unknown that need to be solved for. There exist no known analytical solution to this problem that must be solved using well established numerical iterative methods (e.g. Gauss-Seidel, Newton-Raphson, Fast-decoupled Method, etc.) [19]. Typically, the rate of convergence of those methods is quadratic, but in [20] it is shown that the convergence region and number of iteration can significantly vary depending on the initialisation and the actual loading condition of the system. As soon as the system's size increase, and also depending on the loading condition of the system, solving the system becomes computational intense. It is therefore poorly practical to analyse the effect of vast numbers of stochastic net-load scenarios using the brute force of solving the load flow equations for each combination of net-loads input to assess the resulting values of current flowing within the system. For that purpose, we are proposing to use a data-driven method that is based on Artificial Intelligence techniques that will offer a prediction of the currents within the system in response to any net-loads with an appropriate precision at a fraction of the computational cost, and therefore enabling to assess vast amounts of scenarios and derive the probability distribution of currents value at any node of the system. In our simulations, we experienced that solving the systems using a trained Deep Neural Network (DNN) model is $\sim$ 1250 times faster than using the established simulation tool OpenDSS.

Hornik [21] established that "multilayer feed-forward neural networks are, under very general conditions on the hidden unit activation function, universal approximators provided that sufficiently many hidden units are available.", best known as the "Universal Approximation Theorem". The theorem doesn't provide strict criteria regarding the structure of the network or number of hidden layers and neurons that would perform optimally but states that structure can be found such that it can provide an arbitrarily accurate approximation of a function. In [22], [23], [12], the authors are experimenting the use of Deep Neural Networks (DNN) to solve the Power flow equations and showing the effectiveness of the method to predict the voltages and power angles mainly on the transmission network. We are proposing to extend the method and train a regression DNN to predict the currents at any node of a distribution network in response to given net-loads using well established supervised learning techniques as described in [24], [25]. The key objective of training DNN models resides in the fact that once the model is trained and performs with adequate accuracy, it is much faster to use compared to solving the entire system. Training the DNN model using supervised learning techniques requires an adequate training data set where the true value of currents in response to net-loads are known and for the scope of this paper we are generating such training data set through system simulation.

III. Simulation results on IEEE123 bus test system

As a baseline, the IEEE123 bus test system (Fig.4) and its parameters have been used to illustrate the proposed framework.

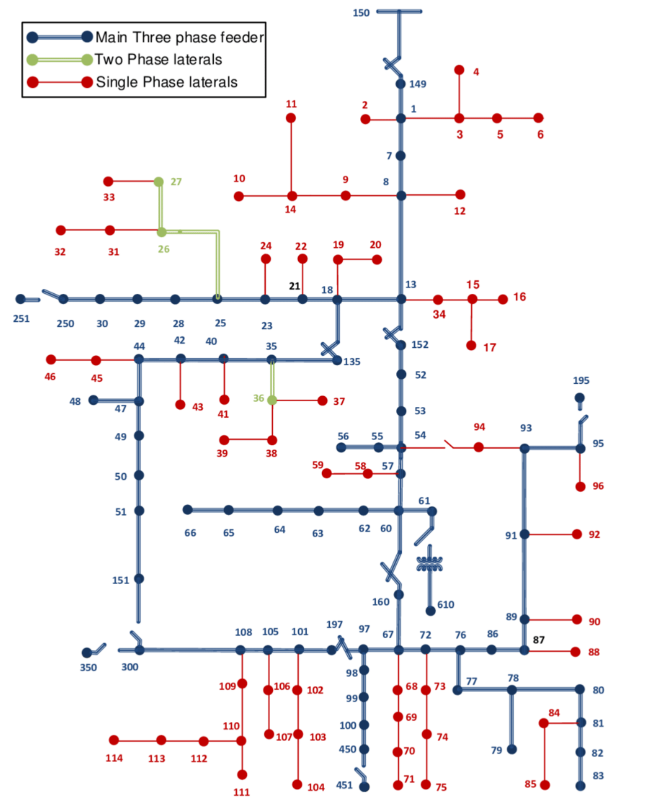


Fig. 4. IEEE 123 Bus Test Distribution Network

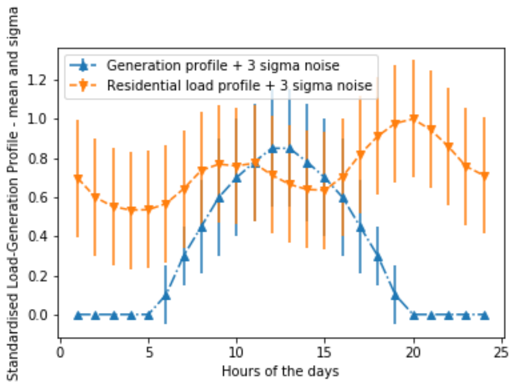
In addition to the baseline network and load parameters, individual solar PV generation units have been embedded in the system randomly in order to simulate various level of penetration (20%, 40%, 60% and 80%). It is to mention that the baseline network configuration used include some commercial loads. In accordance with recent values observed in Victoria-Australia, the average capacity of the solar PV system installed is set to $5KW$ and the average daily energy residential consumption considered of $24KWh$. Typical hourly load and generation profiles from an Australian DNSP have been considered for our simulations (Fig.5).

A. Simulation setup

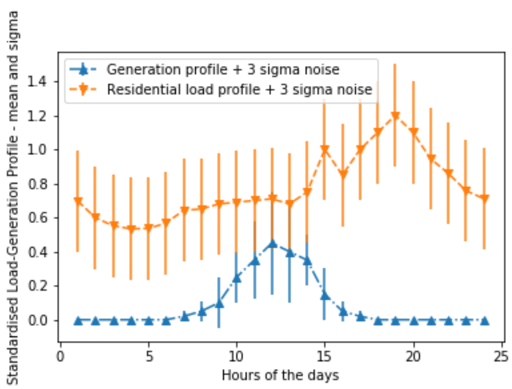
The set of data required to train the DNN model has been generated through Monte Carlo simulation for the entire system in 5 different configurations where :

- Each configuration corresponded to a scenario of solar PV penetration: 0%, 20%, 40%, 60% and 80%. The percentage of penetration corresponding to the percentage of load buses where Solar PV is installed, placed randomly
- For each configuration, multiple set of days have been simulated using the load/generation profiles illustrated in (Fig.5). Per each day, the system has been fully solved hourly between 9am and 8pm, corresponding to 12 full simulations per day

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OAJPE.2020.3032902, IEEE Open Access Journal of Power and Energy



(a) Typical summer day



(b) Typical winter day

Fig. 5. Standardised load and generation daily profiles + noise

- For every hourly simulation, the loads and generation values at every node have been randomly extracted using a normal distribution with the mean and variance illustrated in (Fig.5) capped by zero and the max nominal capacity ($5KW$ per unit) where applicable
- OpenDSS [26] integrated with Python has been used to solve the multi-phase power flow and the voltage, angles and currents of every node and hourly simulation have been recorded in a database.

The training data set generated has been used to train a DNN regression model that predicts the node's currents (Output) in the function of the node's net (generation - load) real and reactive power (Input). As result of an hyper parameter selection process, the selected model structure includes 3 hidden layers, the first layer has a number of neurons equal 4 times the number of load nodes in the system, the second has twice the number of load nodes and the last has as many neurons as the load nodes in the system. The hidden layers activation functions are $tanh$ and the output is $linear$.

We have deliberately generated four training data-sets with an increasing number of days (365, 1000, 1825, 10000) and trained 4 different models. We have tested each model on independent validation dataset and compared the predicted value with the true value calculated via OpenDSS simulation. The results shown in Table I illustrate that the Mean Absolute Percentage Error (MAPE) remains close to the 10% range that is commonly used when evaluating the rating current of assets.

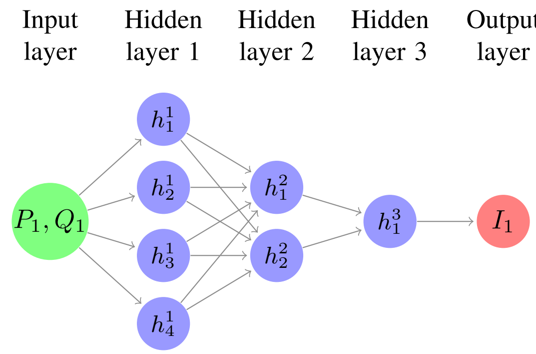


Fig. 6. DNN regression model structure - 3 hidden layers

Training data set size	Mean Absolute Percentage Error
6×10^5	5.97%
1×10^5	7.82%
6×10^4	8.57%
2×10^4	10.71%

TABLE I
MODEL ACCURACY VERSUS TRAINING SET SIZE

B. Reverse Power Flows Analysis

To appreciate and illustrate the effect of simulated increased solar PV penetration in the system, we will first use the 'true' data that was generated solving the system with OpendDSS. Fig.7 illustrates the average current for each penetration scenario at the main feeder (phase 1 of node 149 in Fig4) on a typical summer day (Fig. 5a) and where the individual solar PV capacity is set to $5kW$ with no storage or curtailing involved. Conventionally, the reverse current is positive and load current negative. Without solar PV embedded generation,

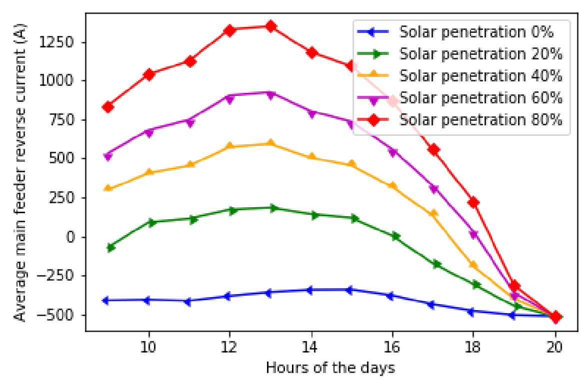


Fig. 7. Main feeder phase 1 current - "true" data

the average load current stands in the range $\sim -300 - 500A$, while in an extreme case of 80% solar PV penetration (without storage or curtailing), the reverse current exceeds $+1250A$ at 1pm. Naturally, while the day progresses, the effect of any Solar PV is fading and the currents in each scenario are converging to the same values. We observe that not only the direction of the current reverses but its intensity increases by a factor of $\sim 250\%$ during the central hours of the day. This illustrates the potential intensity of reverse power flows within

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OAJPE.2020.3032902, IEEE Open Access Journal of Power and Energy

the power system due to high penetration of solar PV and justify the need for an analysis framework that can help the DNSPs to assess impacts and make mitigation decisions to protect the network regarding fast increasing residential solar PV system penetration.

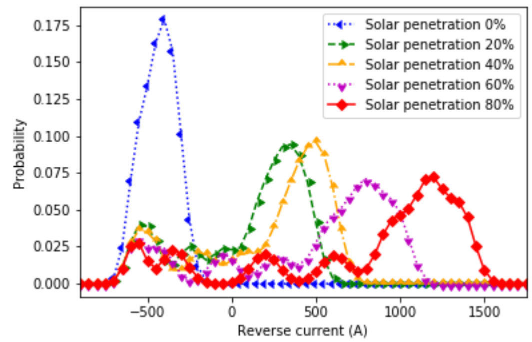
C. Computational comparison

While OpenDSS is certainly a reliable way to simulate and solve power flows, it also comes with a computing cost that restricts the number of scenarios that can be assessed. For example, on a MacBook Pro it took more than 14 hours to simulate the maximum amount of scenarios (600k) described in III-A. Using the DNN model is a good compromise between precision and speed of simulation, while the average model error in that case is close to 6%, the speed to predict 600k on the same MacBook Pro is 40 seconds, or 1250 times faster than simulating the system and solving the load-flow equations.

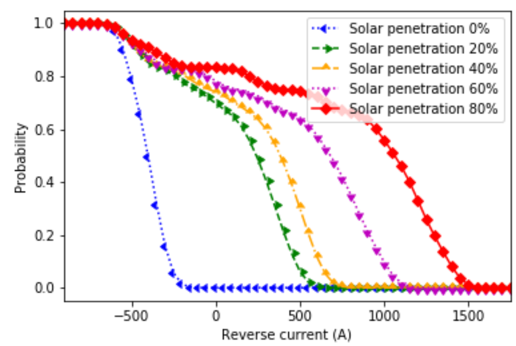
Training the model comes with its own cost and requires a training data set obtained by simulations, but it is a one-off activity which can be optimised by accurate selection of the training data set and use advanced training techniques left as future work for the scope of this paper. Nevertheless, if we accept a model error close to 10% (Tab.I), the simulation cost of creating the training data set (10^4) is 2 orders of magnitude lower than the equivalent cost of using traditional simulation techniques to execute the vast Monte-Carlo simulations (10^6).

IV. Scenario Analysis

Applying the framework illustrated in 1, a vast Monte Carlo simulation that randomly extracts the loads from a typical summer day (hourly from 9am to 8pm) for 100k days in 4 distinct solar PV penetration scenario has been performed and the currents at every node predicted using the DNN model. We have defined current buckets, ranging from $-850A$ to $+1750A$ by step of $50A$ and have assigned each predicted current in their respective bucket. Leveraging the law of big numbers, we derived probabilities from the frequency of current being in buckets for each scenario. Fig.8a shows the probability distribution of the current in the main feeder for the 4 scenarios, and Fig.8b shows the cumulative probability distribution of the current in phase 1 of the main feeder in the 4 scenarios. Remarkably, the current prediction made with the DNN shows a good sensitivity to the solar PV scenarios and consistency with the "true" data simulated (Fig.7). It also demonstrates the progressive increase of reverse current that is most likely to happen with an increased penetration of solar PV within the system. From a DNSP point of view, this node by node probability analysis is a piece of important information when associated with the thermal limits of the network asset installed at that node (e.g. Line, protection device, transformer, etc.) and the threshold current that it can safely withstand. The combination of those two information can enable to identify and map the assets within the network that are most likely to be utilised at full thermal capacity (or above) in any scenario of increased solar PV penetration.



(a) Probability distribution function (pdf)



(b) Cumulative probability function (cpf)

Fig. 8. Probabilities of current in phase 1 of the main feeder - DNN predicted data between 9am and 8pm

V. Impact Analysis

A. Ampacity Impact

To support planning decision making, the current rating of each asset must be mapped with the cumulative current probability distribution of the node where the asset is located. As an example, if the transformer located at node 149 is rated at $500A$, we know that with 80%, 60% or 40% solar penetration, there is respectively approx 75%, 60% or 30% of chance that the rating will be breached at any time between 9am and 20pm, but no risk if penetration is kept below 20% as Fig.9 shows.

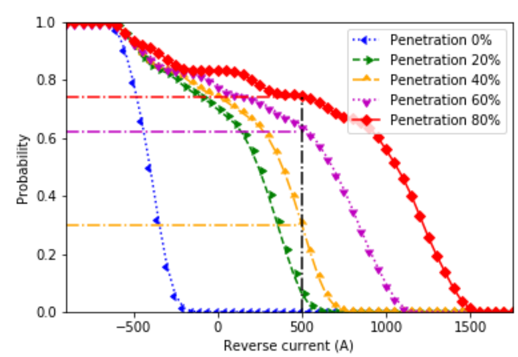


Fig. 9. Control asset rating with current probabilities

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OAJPE.2020.3032902, IEEE Open Access Journal of Power and Energy

B. Protection device

For completion, we are introducing the time dimension in the analysis and deriving for each current bucket the probability of being in that bucket for one time period (1 hour), 2 consecutive time periods or more than 2 consecutive time periods (Fig.10).

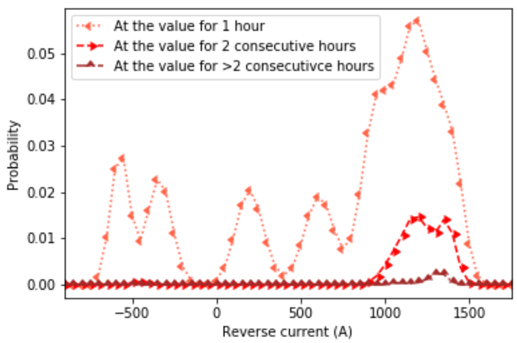


Fig. 10. Main Feeder - Probability distribution functions of being at current value for consecutive time period between 9am and 8pm

Even though for most of the times the predicted peak current would last only for 1 time period (1 hour in our simulations), we observe that the probability of having a sustained high current for more than one time period is not insignificant under our test conditions. This information is valuable when mapped to the characteristic time-current curve of protection devices embedded in the network such as circuit-breakers or transformers thermal protection relays. Furthermore, Fig.11 illustrates a typical tripping curve of protection devices such as circuit breakers that connects the time to trip to the amplitude of current relative to the current rating of the device.

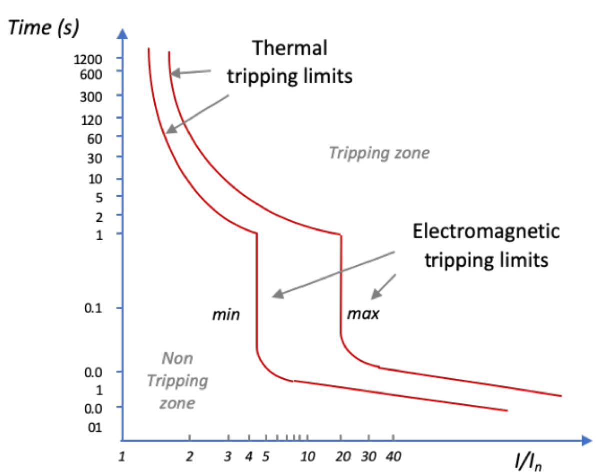
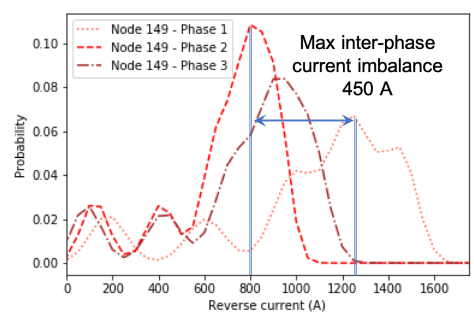


Fig. 11. Typical protection device tripping curve

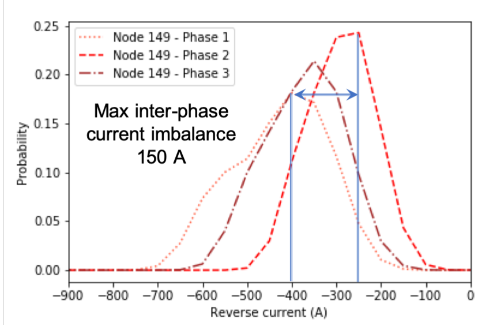
Therefore, for each protection device that is placed in the network or embedded in assets such as transformers, the probabilities of sustained current illustrated in Fig.10 and obtained as an output of the RPF scenario analysis framework are valuable to assess the risk of thermal tripping due to sustained current that is multiple of the device current rating value. Following our previous example, if the transformer located at node 149 is rated at $500A$, using the probabilities in Fig.10, we see that the probability of a current that is 250% of the transformer rating for at least 2 consecutive hours is not negligible and might trip its own protection in a scenario of 80% solar PV penetration. We see how the analysis framework enables to assess every asset located in the network against scenarios of increased penetration of intermittent DER within the system. By establishing some risk acceptance thresholds (max risk probability), the network operators can swiftly assess the rating adequacy of their asset pool and/or prioritise the assets that must be upgraded in response to increased penetration of intermittent DER (eg. solar PV).

C. Phase imbalance

Additionally, the analysis framework enables to analyse the effect of increased penetration of solar PV on the overall balance of the 3-phase system. As shown in Fig.12, for the configuration used in our simulations, we denote that the unbalance of currents between the 3 phases of the main feeder is exacerbated by an high penetration of solar PV.



(a) 80% Solar PV penetration



(b) No Solar PV penetration

Fig. 12. Main Feeder - Probability distribution functions of currents in each phase of the main feeder

Fig.12b shows that without any solar PV within the system, the unbalance gap between the 3 phases is contained and mainly due to the configuration of the IEEE 123 Test case which is slightly unbalanced. Fig.12a shows that for an elevated penetration of solar PV in the system (80% in that case), the gap between the phases widens significantly. This demonstrates the versatility of the analysis framework that enables to assess granular DERs penetration scenarios and identify precisely the impacts of the resulting RPF on installed assets on the network.

D. Storage, EV and curtailment considerations

The DNN model is generalised and enables to predict the current within the system in response to any load and generation profiles. In our simulations, for illustration purpose, we have derived the RPF probabilities from a Monte Carlo simulation using a typical summer net-load profile with Solar PV and without storage, Electric Vehicles or curtailment capability. To simulate the effect of storage, Electric Vehicle or curtailment rules, an adapted net-load profile that considers the charge and discharge cycle of the batteries would need to be used in the Monte Carlo simulation, but the core DNN model would remain similar.

VI. Conclusion and Future Work

The presented framework provides a versatile and powerful methodology to predict the intensity and probabilities of reverse power flows in any given scenario of intermittent DER penetration embedded within the distribution network. The key strengths of the proposed framework are the built-in speed and versatility to analyse the impact of any penetration and net-load profile in the system. The use of a trained DNN model substitute the requirement to solve the load flows equations for the system and enables to assess vast number of scenarios. The one-off cost of creating the training data-set for the DNN model can be 2 orders of magnitude lower than the equivalent cost of using traditional simulation techniques to assess the vast amount of scenarios. Future work required to operationalise the framework are:

- Test and optimise the DNN training methodology against large-scale networks, in particular identify an adequate sampling optimisation methodology to even reduce further the size of the training dataset without compromising the model accuracy, for instance using a mix of simulated and observed data [27]
- Perform a sensitivity analysis of the DNN model accuracy in response to local changes in parameters and test new supervised training methods to train the existing model to recognise the new configurations
- Explore the use of reinforcement learning technique to adapt the model to new conditions, such as change in system parameters and configuration.

References

[1] IRENA, “Global energy transformation: A roadmap to 2050,” International Renewable Energy Agency, techreport, 2019.

[2] C. L. Masters, “Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines,” *Power Engineering Journal*, vol. 16, no. 1, pp. 5–12, 2002.

[3] L. v. d. S. Pieter Schavemaker, *Electrical Power System Essentials*. John Wiley Sons Inc, 2017.

[4] R. Seguin, J. Woyak, D. Costyk, J. Hambrick, and B. Mather, “High-penetration PV integration handbook for distribution engineers,” NREL, Tech. Rep., jan 2016.

[5] E. Coster, J. Myrzik, and W. Kling, *Distributed Generation*, D. Gaonkar, Ed. IntechOpen, 2010.

[6] R. A. Walling, R. Saint, R. C. Dugan, J. Burke, and L. Kojovic, “Summary of Distributed Resources Impact on Power Delivery Systems,” *IEEE Transactions on Power Delivery*, vol. 23, no. 3, pp. 1636–1644, jul 2008.

[7] N. Amjady, A. Attarha, S. Dehghan, and A. J. Conejo, “Adaptive robust expansion planning for a distribution network with DERs,” *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1698–1715, mar 2018.

[8] X. Guo, H. Guo, and H.Cheng, “Coordinated planning of Distributed Energy Resources and microgrid network,” in *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*. IEEE, may 2016.

[9] M. J. E. Alam and K. M. Muttaqi and D. Sutanto, “A comprehensive assessment tool for solar pv impacts on low voltage three phase distribution networks,” in *2nd International Conference on the Developments in Renewable Energy Technology (ICDRET)*, Jan. 2012.

[10] K. Joshi and N. Pindoriya, “Advances in distribution system analysis with distributed resources: Survey with a case study,” *Sustainable Energy, Grids and Networks*, vol. 15, pp. 86–100, sep 2018.

[11] A. Augugliaro, V. Cataliotti, L. Dusonchet, S. Favuzza, and G. Scaccianoce, “Load flow solution in electrical power systems with variable configurations by progressive learning networks,” in *PowerTech Budapest 99. Abstract Records. (Cat. No.99EX376)*. IEEE, 1999.

[12] M. Fikri, B. Cheddadi, O. Sabri, T. Haidi, B. Abdelaziz, and M. Majdoub, “Power flow analysis by numerical techniques and artificial neural networks,” in *2018 Renewable Energies, Power Systems & Green Inclusive Economy (REPS-GIE)*. IEEE, apr 2018.

[13] A. Larsson, A. Germond, and B. Zhang, “Application of neural networks to the identification of steady state equivalents of external power systems,” in *2006 International Conference on Power System Technology*. IEEE, oct 2006.

[14] A. Demazy, T. Alpcan, and I. Mareels, “A quantitative risk framework for der-rich power system planning and decision making,” *IFAC-PapersOnLine*, Jul. 2020.

[15] G. E. Constante-Flores and M. S. Illindala, “Data-driven probabilistic power flow analysis for a distribution system with renewable energy sources using monte carlo simulation,” *IEEE Transactions on Industry Applications*, vol. 55, no. 1, pp. 174–181, jan 2019.

[16] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[17] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis, “Physics-informed neural networks for power systems,” *arXiv:1911.03737 [eess.SY]*, 2019.

[18] L. Che, M. E. Khodayar, and M. Shahidehpour, “Adaptive protection system for microgrids: Protection practices of a functional microgrid system.” *IEEE Electrification Magazine*, vol. 2, no. 1, pp. 66–80, mar 2014.

[19] X.-F. Wang, Y. Song, and M. Irving, *Modern Power Systems Analysis*. Springer US, 2008.

[20] J.-J. Deng and H.-D. Chiang, “Convergence region of newton iterative power flow method: Numerical studies,” *Journal of Applied Mathematics*, vol. 2013, pp. 1–12, 2013.

[21] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991.

[22] W. A. Alsulami and R. S. Kumar, “Artificial neural network based load flow solution of saudi national grid,” in *2017 Saudi Arabia Smart Grid (SASG)*. IEEE, dec 2017.

[23] A. Rathinam, S. Padmini, and V. Ravikumar, “Application of supervised learning artificial neural networks [CPNN, BPNN] for solving power flow problem,” in *IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007)*. IEEE, 2007.

[24] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (3rd Edition)*. Pearson, 2009.

[25] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[26] D. Montenegro and R. C. Dugan, “OpenDSS and OpenDSS-PM open source libraries for NI LabVIEW,” may 2017.

[27] X. Zheng, N. Wu, and J. Yi, “Development of a hybrid simulation technology to optimize parameters of multiple power system stabilizers,” in *2020 5th Asia Conference on Power and Electrical Engineering (ACPEE)*. IEEE, jun 2020.