American Journal of Energy and Power Engineering
2018; 5(3): 20-29
http://www.aascit.org/journal/ajepe
ISSN: 2375-3897

AASCIT American Association for Science and Technology

Electric Power System Supporting a Smart-Grid: ANN-Based Prediction of a Representative Load-Curve to Assess Power-Consumption and Tariff

Dolores De Groff*, Roxana Melendez, Perambur Neelakanta

College of Engineering and Computer Science, Florida Atlantic University, Boca Raton, USA

Email address

degroff@fau.edu (D. De Groff), rmelendez40@gmail.com (R. Melendez), neelakan@fau.edu (P. Neelakanta)

*Corresponding author

Citation

Dolores De Groff, Roxana Melendez, Perambur Neelakanta. Electric Power System Supporting a Smart-Grid: ANN-Based Prediction of a Representative Load-Curve to Assess Power-Consumption and Tariff. *American Journal of Energy and Power Engineering.* Vol. 5, No. 3, 2018, pp. 20-29.

Received: August 7, 2018; **Accepted**: August 17, 2018; **Published**: September 3, 2018

Abstract: This paper is specific to the background of *ad hoc* predictive modeling of electric power-distribution and related tariff issues by deducing objectively, a representative load-curve (RLC) *vis-à-vis* randomly-varying, daily electric-power demand in a service area. Relevantly, the method pursued uses an artificial neural network (ANN) in prescribing the said RLC within a cone-of-error (specified between a pair of stochastic bounds). Pertinent modeling and approach use a set of available (case-study) data; and, the closeness of the RLC deduced is cross-verified against relevant (existing) results derived *via* fuzzy K-mean method. The study concludes on adopting the bound-specified RLC towards formulating pertinent tariff considerations within a range of error-bar. The alternative ANN-approach proposed here has been found to produce accurate results close to FKM-results. However, it is an improvement over the FKM method in that the RLC is specified within a cone-of-error, accounting duly for the associated stochastic implications with the load-curve profile. Thus, rather than yielding a rigid RLC (yielding a rigid tariff policy), any tariff policy derived by this ANN-based method combined with stochastical bounds will have the judicious basis of technoeconomics of the utility in question. The method proposed here is novel and not hitherto done; and leads to optimal integrated planning for electricity towards load-demand *versus* tariffication decisions.

Keywords: Representative Load Curve, Artificial Neural Network, Smart-Grid, Load Distribution, Stochastical Error Bounds

1. Introduction

In modern attempts towards implementing the complex infrastructure of a smart-grid in the network of an electric-power utility facilitated in a service area, there could be different types of resources of electric energy being availed so as to meet the fluctuating load-demand posed by distinct groups of consumers across the power distribution network. Relevantly, demand (load) curves are assessed on a daily basis; and, such curves depict the statistical profile of electric-power consumption (in kW or MW) *versus* the time-of-the-day (1-to-24 hours). They implicitly decide the modeling aspects of the capital and operating expenses (CAPEX and OPEX) and hence, the tariff levied by the utility company. In this context, determination of a representative load-curve (RLC) of the aforesaid demand (load) curves is necessary so as to decide the logistics of underlying tariff to be levied on the consumer base in question. It is stressed in this study that such an RLC should also be specified within stochastic limits of a cone-of-error (rather than by a fixed-curve) commensurate with the statistical aspects of technoeconomic variables involved; and, specifications on relevant upper- and lower-bounds of the said cone-of-error are indicated for the example data analyzed here. (Relevant pursuit and suggestion indicated thereof are novel and hitherto not considered nor reported.)

1.1. Scope and Objectives

This study is pertinent to the scope of deciding the performance of an electric utility with the objectives as follows:

1. To describe a smart-grid implied infrastructure of an electric power utility in terms of deterministic features of the associated technological considerations and the stochastic profile of the power-distribution decided by the load-demand *versus* multiple power resources
2. Using the aforesaid deterministic and stochastic considerations, to evolve a model of the underlying tariff logistics
3. To assess hence, a representative load-curve (RLC) to analyze the technoeconomic details including the tariff considerations
4. To develop an artificial neural network (ANN) method to determine the RLC and specify it with pertinent statistical upper- and lower–bounds (UB and LB) of the cone-of-error
5. To formulate a compatible tariff model that can be specified within the said stochastic limits of the cone-of-error (specified by the UB and LB).

The technological aspects of power-system expansions with intended smart-grid implementations and assessing the associated economic reasonings that may impact the return-on-investment (RoI) (gained *via* tariff) are essential considerations in modeling the technoeconomics of underlying expansion scenarios. Further, inasmuch as both electric-power generation and consumer-demand are stochastic details governed by both deterministic and statistical variables involved, the smart-grids provisioned in such expansions are to be designed in order to accommodate effectively the deterministic (known) engineering details as well as the fluctuating load-distribution profiles across the service network in question. Concurrently, the cost-of-implementation (that is, the CAPEX) and the maintenance charges (namely, the OPEX) should be judiciously managed with appropriate tariff being levied on to the customers and revenues accrued thereof. Further, pertinent policy has to be logistically evolved and modeled consistent with the associated technoeconomics infested with stochastic attributes. Also, such models should be predictive (at least within a cone-of-error) *via* compatible forecast methods imposed on power generation (facilitated with multiple sources) and load-demand profiles observed on a daily basis.

Hence, proposed in this study is a strategy on pertinent stochastic modeling of the technoeconomics of a power system and deducing the associated RLC using artificial neural network (ANN) approach; the purpose thereof, is to establish the details on the tariff *versus* power-demand profiles (with due considerations on their deterministic and stochastic characteristics). Relevant simulations indicated in this study are exercises carried out on a case-study data availed from the literature [1]. In all, the proposed effort in this paper objectively focuses on forecasting tariff issues implicitly correlated to an emulated RLC (specified within the statistical limits of a cone-of-error) of a test data set *versus* electric power-demand in an electric power utility encountered with statistical attributes; and, as stated before, the objective of the proposed effort on relevant prediction is aimed at using an ANN approach.

1.2. Organization of the Paper

Relevantly, the details of this study are organized in this paper as follows: The next section (Section 2) provides an outlay on the power-system specifics and load-distribution considerations (expressed in terms of measured load-curves). Also, Section 2 contains the heuristics on related tariff implications versus statistical aspects of fluctuating, multiple power-generations (adopted as resources) in smart-grid networks and corresponding random load-demand variations. In Section 3, relevant tariff considerations are identified in terms of technological norms and economics with their statistical bounds; hence, the underlying model proposition along with its algorithmic description is presented. Also, details on the test-ANN architecture and the associated training and prediction phases are covered (in Section 3). Estimation of an appropriate tariff parameter consistent with the objective model proposed is the topic addressed in Section 4; relevantly, determination of this tariff parameter *versus* load-statistics is considered *via* load-curve profiles and the RLC of a typical power-system network. The required RLC details are emulated using an ANN strategy. A summary of simulations done and discussions on the results obtained *via* the test ANN are furnished in Section 5 with reference to a set of case-studies data availed from the literature [1]. Hence, the predicted RLC is indicated to lie within a pair of stochastic bounds (of a cone-of-error) close to the results obtained elsewhere on the case-study data with the RLC deduced by another method [1]. The comparison made confirms the validity of the proposed method. Lastly, concluding remarks are presented in the closure section (Section 6).

2. Power-System Load Distribution and Tariff Considerations

2.1. Electric Power Systems with Smart-Grid Infrastructure: An Overview

An avenue of modern electric-power utility includes provisioning a smart-grid network that contains multiple resources of power generation (such as thermal, hydro, solar, wind-mill, etc.) integrated or each placed in isolation on a distribution network of a service area; and, correspondingly, the consumer base may also consist of multiple types of demands posed by domestic, industrial, and commercial users.

Implementing such a smart-grid infrastructure in the power-system utility also requires concurrently a market strategy in order to set up a pricing policy matching the cost-profiles of the facilitated, multiple resources that meet different load-demands (consistent with the existence of a variety of customers); and, hence the associated tariffication policy also needs appropriate modeling. Therefore,

developing estimation procedures towards establishing a pertinent (tariff) parameter is crucial in the state-of-the-art smart power-grid systems.

2.2. Load-Curves and RLC: Deterministic and Stochastic Considerations

With reference to power-distribution networks with smart-grid provisioning, the power consumption (P expressed in kW or MW) in a service area by a given type of consumers (such as domestic, industrial or commercial) may fluctuate with time over the 24 hours of the day. In other words, the power demand P is a random variable (RV) changing instantaneously as well as, possibly in conformance to a monotonic increasing or decreasing function of time. A hypothetical depiction of P *versus* time (H in hours) is illustrated in Figure 1.

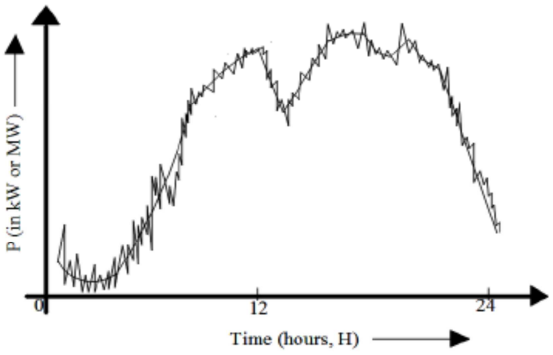

Figure 1. *P versus time (H) – a hypothetical representation of the random variation of P with respect to time.*

The variation of P with time is typically characterized by the following: (i) Noticeable peaks and troughs at designated hours; (ii) random instant-by-instant fluctuations of the load and (iii) technology-dictated plus service-type based (overall) load-demand (power-consumption) variations (growth and/or decay profiles). In essence, Figure 1 illustrates the stochastic profile of the load-curve, which is often denoted by a time-series expression of the type as follows:

$$P(t) = a_0 + a_1P_{t-1} + a_2P_{t-2} + \ldots + \xi \quad (1)$$

Equation (1) denotes P(t) in terms of its instant-to-instant values, specified as a series; here a_o is a constant intercept and, $\{a_1, a_2, \ldots\}$ is a set of the so-called auto-regressive model parameters and ξ is a random error component on each observation. Further, equation (1) implies that a set of coefficients decides the consequent elements of the series from specific time-lagged (previous) elements; and, it assumes the underlying process as stationary [2].

Typically, the stochastic aspects of equation (1) result from random fluctuations seen in power-consumption caused by endogenous technology-specific distribution system considerations; and, other causative exogeneous factors include the random patterns of electric power usage by the consumers as well as other related seasonal variations.

The depiction of load-curve in Figure 1, however, can be "smoothed" with respect to short-time (instant-to-instant) variations. However, its profile across hourly changes can be significant and retained. Likewise, variations in electric power usage over seasons are significant and they are also retained. As such, a set of load-curves {a, b, c, d} as shown in Figure 2, are usually used in power-system analyses.

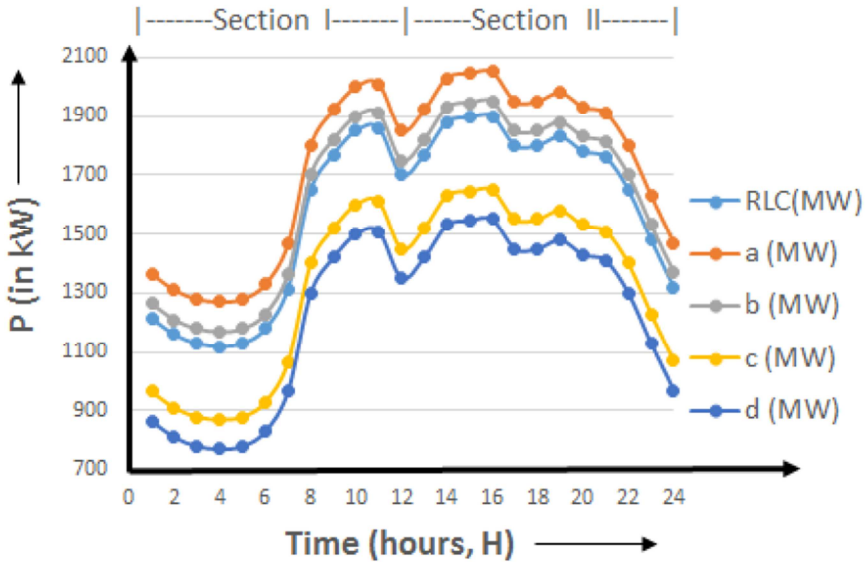

Figure 2. *A hypothetical exemplar set of load-curves {a,b,c,d} corresponding to seasonally varyingpower-demand profiles. (Each curve can be regarded as a smoothed version of the hypothetical curve in Figure 1).*

For applications such as, modeling the tariff policy consistent with power-demand in question, a representative load-curve (RLC) can be obtained for use, *in lieu* of the set {a, b, c, d}. For example, in [1], such a RLC is determined from a load-curve set pertinent to a case-study data of an electric utility system. The set of load-curves so considered is,

however, attributed with fuzzy characteristics; and, the method of K-Means is invoked to ascertain the RLC needed. Though such fuzzy considerations and K-Mean averaging could provide an average representation of the load-curves involved, yet another perspective on RLC would be to consider the stochastic aspects of the load-curves (implicitly specified by their random time-series depictions) and hence, arrive at the RLC specified with a cone-of-error confined within upper– and lower-bounds (UB and LB). A representation of the RLC under the circumstances of underlying endogeneous and exogeneous details of the set, {a, b, c, d} would be more realistic and more significant. Accordingly, pertinent tariff policy could also be established with bound-limited specifications for judiciously realistic applications.

To the best of the authors' knowledge, such RLC details prescribed within a cone-of-errors have not been determined and presented in the literature. As such, the present study is pursued and the required (bound-limited) RLC is obtained using an ANN-based method.

As stated earlier, in [1] the representative load-curve (RLC) is prescribed to establish tariff requirements; and, it conforms to a specific customer group of interest. Pertinent data used depicts the power demand (P in kW or MW) in the distribution network (availed by the test group) over the 1 to 24 hours of a day of a power utility system, in a specific case-study. The set of load-curves {a, b, c, d}, for example (as in Figure 2), would normally prevail in a power utility across the seasonal terms of a year; and, the RLC is a candidate that can be deduced from these load-curves so as to represent the net details of the associated power demands. (Typically, the RLC indicates the lowest as well as peak demand on the electric power from the consumer group under consideration over the 24 hours of the day).

The electric-power consumption activity profile *versus* time can also be adopted to decide on the tariff levied on a rational basis. That is, the utility can determine the electric price for each customer class as per the demand perceived from that class [3].

As stated earlier, the method of finding the RLC described in [1] uses the fuzzy K-Means (FKM) technique. However, the present work, is conceived to offer an alternative strategy to construct the RLC using an ANN approach; and, the RLC deduced is also specified within a bounded statistical cone-of-error. The following subsection (2.3) describes the pursuit involved.

2.3. Proposed Bound-Specified RLC Determination Using ANN Approach: An Overview

The underlying considerations and steps in constructing the RLC *via* ANN-based method are briefly outlined below and relevant details are furnished in the sections that follow

Step I: Suppose the available load-curves (measured in an electric utility distribution base) are indicated by the set {a, b, c, d}, illustrated in Figure 1. These load-curves may refer to, for example, four seasonal data collected in a service area pertinent to a designated group of electric power consumers. Shown further in Figure 3, is an exemplar set of hypothetical load-curves a, b, c, and d; and, focusing on two time-instants, say k = 10 hours and k = 14 hours, the corresponding set of P values (in kW) on the load-curves are taken as follows (as in Figure 3):

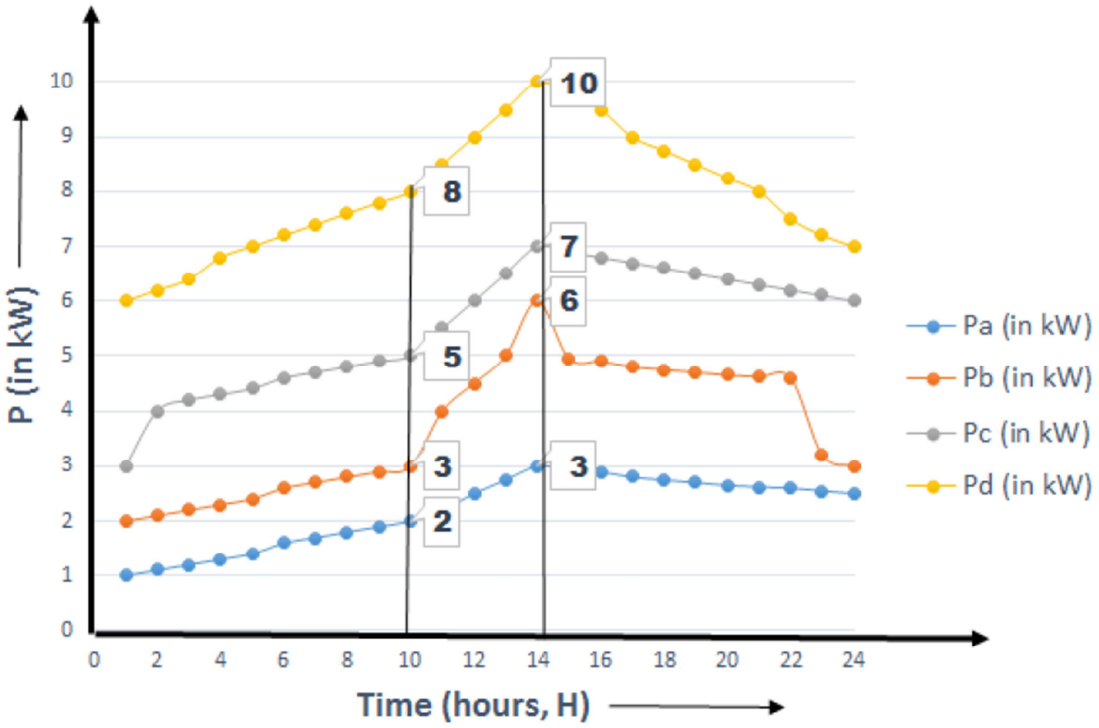

Figure 3. *Hypothetical set of load-curves {a,b,c,d} with time-markings of k = 10 hours and k = 14 hours (being considered here as examples).*

$$\text{P\_values of } \{a, b, c, d\}_{\text{at k = 10 hours}} \equiv \{2,3,5,8\}_{kW}$$

and

$$\text{P\_values of } \{a, b, c, d\}_{\text{at k = 14 hours}} \equiv \{3,6,7,10\}_{kW}$$

Step II: As indicated in Step I, similar details for each hour (1 through 24) are also gathered with reference to the set {a, b, c, d}. Thus, the following details can be established and listed:

$$(P)_1 \equiv \{P_{1a}, P_{1b}, P_{1c}, P_{1d}\}_{kW}$$

$$(P)_2 \equiv \{P_{2a}, P_{2b}, P_{2c}, P_{2d}\}_{kW}$$

$$\vdots$$

$$(P)_k \equiv \{P_{ka}, P_{kb}, P_{kc}, P_{kd}\}_{kW}$$

$$\vdots$$

$$(P)_{24} \equiv \{P_{24a}, P_{24b}, P_{24c}, P_{24d}\}_{kW}$$

Step III: For each set $(P)_{k=1,2,\ldots,24}$, gathered and listed in Step II, relevant pseudo-replicates can be generated *via* statistical bootstrapping technique [4]. The following are indicated as examples:

Reference to the time-marking (in Figure 3) at H= 10 hours:

$$(P)_{k=10}^{a,b,c,d} \equiv \{2,\ 3,\ 5,\ 8\}$$

and, at H = 14 hours

$$(P)_{k=14}^{a,b,c,d} \equiv \{3,\ 6,\ 7,\ 10\}$$

For each of the above sets, the pseudo-replicate values can be generated *via* bootstrapping concept [4].

Thus, considering the sets $(P)_{k=10} \equiv \{2, 3, 5, 8\}$ and $(P)_{k=14} = \{3, 6, 7, 10\}$, the corresponding pseudo-replicates can be specified and listed as shown in Table 1.

Table 1. *Bootstrapping-based pseudo-replicates of the test sets $(P)_{k=10}$ and $(P)_{k=14}$ of the load-curves, {a, b, c, d}.*

Original sets	Examples of bootstrapping-based pseudo-replicates
$\{P_{k=10}\} \equiv \{2,3,5,8\}$ $\Rightarrow \{a,b,c,d\}_{10}$	$\{2,3,5,5\}$
	$\{3,2,5,8\}$
	$\{2,3,3,8\} \cdots etc.$
$\{P_{k=14}\} \equiv \{3,6,7,10\}$ $\Rightarrow \{a,b,c,d\}_{14}$	$\{3,6,6,10\}$
	$\{3,6,7,7\}$
	$\{6,7,10,10\} \cdots etc.$

Step IV: The pseudo-replicates presented in Table 1 denote examplars of shuffled values of the original sets at k = 10 and k = 14. The cardinality of each pseudo-replicate set is same as the original set; and, for each original set, the number of bootstrapped pseudo-replicates could be as large as 200 [4]. However, it can be limited to 20. That is, the number of pseudo-replicate sets for each original set could be as large as 200 or even higher; but, in this study it is limited to 20 (corresponding to 20 training iterations adopted in the test ANN as will be indicated later).

Step V: Next, considering each hour (of 1 to 24) along the time-scale, containing four values of {a, b, c, d}, corresponding 20 pseudo-replicates are generated. They can then be applied as inputs to the test ANN to train it. Upon the realization of the convergence of the test ANN (as will be described later), the so-called interconnection weights of the net would correspond to values ready to predict the representative load-curve designation at the hour or time-marking (for which the net is trained).

Thus, training the test ANN for the entire time-frame of 1 to 24 hours, the stored interconnection weights would designate the complete representative load-curve needed. The following section (Section 3) provides details on the usage of the pseudo-replicates on a test ANN.

3. Construction of the RLC: ANN-based Emulation

Considering a load-curve, as indicated earlier in Figure 2, it is divided into two major time-sections marked as I and II; and, the (empirical) description of the load-curve in those two sections can be specified in the format of time-series depicted *via* equation (1) as follows:

In the time-section I of any load-curve:

$$P(t)\big|_I = \alpha_0 + \alpha_1 P_1(t-1)\big|_I + \alpha_2 P(t-2)\big|_I + \cdots + \zeta_1 \quad (2)$$

Likewise, for the time-section II of any load-curve:

$$P(t)\big|_{II} = \beta_0 + \beta_1 P_1(t-1)\big|_{II} + \beta_2 P(t-2)\big|_{II} + \cdots + \zeta_2 \quad (3)$$

Here, the sets $\{\alpha_0, \alpha_1, \alpha_2\}$ and $\{\beta_0, \beta_1, \beta_2\}$ are empirical coefficients of time-series regressions fitted on the observed data set of load-curves in time sections I and II respectively. The cluster of load-curves conceived by varying these coefficients denotes a family of load-curves simulated using the pseudo-replicate data; and, the RLC required can be obtained as a 'representative' of this cluster of curves.

In the present study, an ANN approach is pursued to deduce the RLC as above. That is, a sample set of four (seasonal) load-curves are first chosen from the field data (to represent the set {a, b, c, d}). Next, the pseudo-replicates are generated using these four curves at each hour, 1 to 24. As stated before, these pseudo-replicates (obtained *via* bootstrapping) intuitively lead to a cluster of load-curves; and, the RLC is derived from this cluster using an ANN as described below.

In order to evolve the RLC in question, the type of test ANN adopted is described below. The test ANN being adopted corresponds to a feedforward architecture facilitated with backpropagation of the error. It consists of four input neuron units, one hidden layer (with four neuronal units) and

one output unit. The supervising teacher value corresponds to the arithmetic mean of the input set. The test ANN is illustrated in Figure 4.

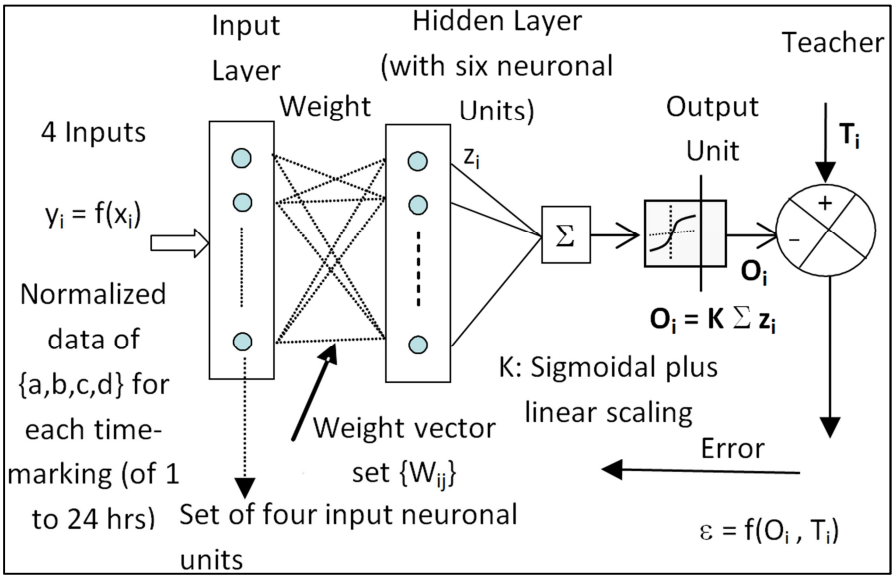

Figure 4. *Test ANN: A feedforward net with backpropagation feasibility of the error.*

Training phase of the test ANN: The training input sets correspond to pseudo-replicated data set of {a, b, c, d} values in Table 1 (obtained *via* bootstrapping as indicated earlier) for each time-markings of 1 to 24 hours; and, relevantly, the following test parameters are used:

1. Teacher value: This changes dynamically for each input. It corresponds to the arithmetic mean value of the four input values.
2. Backpropagation is effected with the mean-square error (ε) obtained at the output and applied to the weight vector matrix $[W_{ij}]$ as shown in Figure 4.

After applying the ANN training input sets (say 20 corresponding to the pseudo-replicated values of {a, b, c, d}), the converged final weight matrix $[W_{ij}]_{Final}$ (corresponding to each time-marking of 1 to 24 hours) is stored; then, each of the weight-matrices is adopted for use in the prediction phase as described below.

Prediction phase of the test ANN: At each time-marking (1 to 24), the test ANN is assigned with the converged weight-matrix $[W_{ij}]_{Final}$ deduced in the training phase; and, the subroutine RLC described below is applied to deduce the representative load-curve of the test utility. As in the subroutine RLC, by constructing the UB and LB curves of P *versus* time (hours, H) and establishing the corresponding *infimum* and *supremum* profiles, the RLC is decided by the cross-intersection of these profiles. For example, in Figure 4, considering example values at the time-instant, $H_{k=12}$, suppose the *infimum* and the *supremum* values are R_I and R_S. Then, the corresponding R is equal to the geometric mean of R_I and R_S. That is, $R = \sqrt{R_I \, X \, R_S}$. Likewise, the RLC coordinates can be established for all k = 1 to 24.

"Subroutine RLC": This refers to algorithmic details and computational routine towards deducing the required RLC. The pseudocode of the subroutine is as follows:

This *pseudocode* is written on the *'subroutine RLC'* (SRLC)

%% The following are steps detailing the computational routines

Input

→ The weight-matrix $[W_{ij}]_{Final}$ ascertained in the training phase of the ANN is recalled

← The test ANN is configured with $[W_{ij}]_{Final}$

Identify

→ Time-markings, $H_{k=1,2,\ldots 24}$ are noted

Select

→ An arbitrary set of new pseudo-replicates of {a,b,c,d} are chosen and listed

← The set of such pseudo-replicates can be arbitrarily taken, say as 10.

Test

→ Convergence of the test ANN (configured with $[W_{ij}]_{Final}$) with each pseudo-replicate set {a,b,c,d} at each time-marking, $H_{k=1,2,\ldots 24}$

← The pseudo-replicates that showed convergence (in each case of $H_{k=1,2,\ldots 24}$) are completed

→ Say, {a,b,c,d}, …, $\{a_{10},b_{10},c_{10},d_{10}\}$

Compute

→ For each time-marking, compute the average-value set, $\{a_m,b_m,c_m,d_m\}$, where

$$a_m = (a_1 + a_2 + \cdots + a_{10})/10$$

$$b_m = (b_1 + b_2 + \cdots + b_{10})/10$$

$$c_m = (c_1 + c_2 + \cdots + c_{10})/10$$

$$d_m = (d_1 + d_2 + \cdots + d_{10})/10$$

Determine

$\rightarrow [Z_m\}_{UB} = Z_m + \frac{1}{2}(0.66) L_{q\to 1/2}\left(\frac{Z_m}{2}\right)$

$\rightarrow [Z_m\}_{LB} = Z_m - \frac{1}{2}(1.66) L_{q\to\infty(say 1000)}\left(\frac{Z_m}{2}\right)$

$\leftarrow Z_m$: Represents a_m or b_m or c_m or d_m

$\leftarrow$ Lq($\bullet$): Langevin-Bernoulli function as defined in [5].

$$\rightarrow L_q(x) = \left(1+\frac{1}{q}\right)\coth\left[\left(1+\frac{1}{q}\right)x\right] - \left(1/q\right)\coth\left[\frac{1}{q}x\right]$$

$\leftarrow$ q: Order-parameter of the statistical system

$q = \frac{1}{2}$: Total disorder state

$q \to \infty$: Total order state

List

$\rightarrow$ For each time-marking $H_{k=1,2,\ldots 24}$ list $\{a_m, b_m, c_m, d_m\}$, $\{a_m, b_m, c_m, d_m\}_{UB}$ and $\{a_m, b_m, c_m, d_m\}_{LB}$

Plot: Upper- and Lower-bound (UB-LB) curves

$\rightarrow$ P versus time (hours, H) with the values of $\{a_m, b_m, c_m, d_m\}_{UB}$ and $(\{a_m, b_m, c_m, d_m\}_{LB}$ (at each time-marking $H_{k=1,2,\ldots 24}$)

$\leftarrow$ These curves depict the upper- and lower-bounds of the required RLC as illustrated in Figure 5.

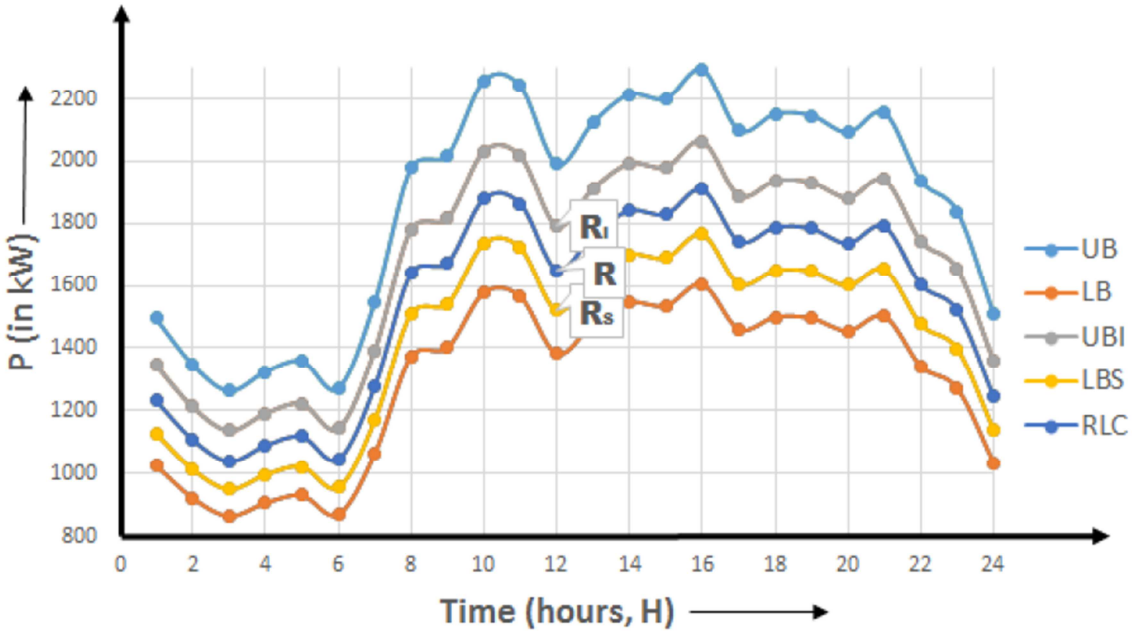

Figure 5. *Deducing the RLC from the UB and LB values of the set {a, b, c, d} from ANN-based simulations.*

Decide

$\rightarrow$ RLC with UB and LB values

$\leftarrow$ Establish the *infimum* and *supremum* profiles of the upper- and lower-bound curves (Figure 5)

$\rightarrow$ *Infimum*: 90 percentile value of UB

Supremum: 110 percentile value of LB

$\rightarrow$ Example at $H_{k=12}$ in Figure 5, R_I and R_S are *infimum* and *supremum* values

$\leftarrow$ Corresponding RLC value, R, is given by $R = \sqrt{R_I \; X \; R_S}$

$\rightarrow$ Geometric mean of R_I and R_S: Cross-intersection bound values of UB and LB determine the cone-of-error as per the algorithm.

$\rightarrow$ Algorithm cone-of-error construction

$\leftarrow$ RLC curve is decided and plotted.

End

4. Simulation Results

Evolving the RLC *via* an ANN-based approach as conceived in this study involves the following simulations and pertinent results.

4.1. Constructing the Pseudo-Replicates of the Set {a, b, c, d}

As described earlier, by choosing an electric utility system and its field-data, a set of load-curves, {a, b, c, d} is considered; and, pertinent values of {a, b, c, d} in the time-scale of k = 1 to 24 hours are noted and listed as original values. Hence, for each time-instant (k), corresponding 20 sets of pseudo-replicates are generated *via* statistical bootstrapping.

Presently, the case-study data as in [1] is considered as the original set and corresponding 20 sets of pseudo-replicates are obtained.

4.2. Training Phase Results

The training phase is done with the 20 pseudo-replicate sets of {a, b, c, d} used as inputs to the test ANN. The final weight matrix $[W_{ij}]_{Final}$ obtained at the convergence of the 20^{th} training set is then stored. Relevant weight matrix exhibits a fast-convergence as described in [6].

4.3. Prediction Phase Results

The test ANN is then organized with $[W_{ij}]_{Final}$ as its weight-matrix; and, the RLC is determined as outlined in the pseudocode presented earlier.

Shown in Figure 6 is the result on RLC deduced by the present method and shown within a cone-of-error decided by intersection bounds of the UB and LB values of the load-curves shown in Figure 5. This result corresponds to the same electric utility case-study data considered in [1]. Along with the details of Figure 6 on the proposed RLC also indicated is the FKM result due to Phan [1] shown for comparison. The efficacy of the present method could be seen by comparing the present result (within the cone-of-error) obtained against the RLC deduced *via* FKM method. It should be noted that the FKM method in [1] does not provide the span of possible bounds on the computed result on RLC. In contrast, the present study is more comprehensive in furnishing a cone-of-error result commensurate with the underlying stochastic considerations. This approach is novel and not indicated in any prior studies.

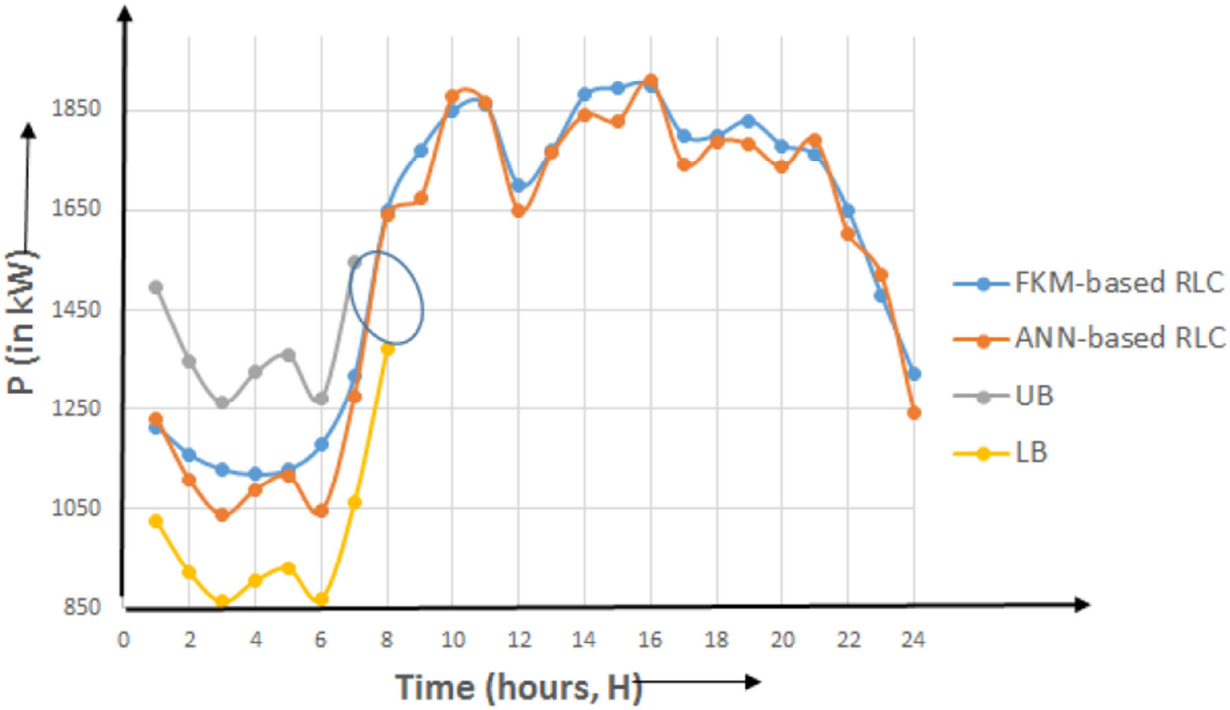

Figure 6. *RLC's obtained via FKM approach due to Phan [1] and ANN-method proposed in this study. The present study implies the RLC be specified within a cone-of-error.*

5. Discussions

The FKM versus ANN-based results on RLC indicated above are presented in Table 2 for evaluation.

Table 2. *Comparative results on RLC deduced via FKM and ANN-based methods.*

Computed Results on P (in MW)				
	FKM-based result	**ANN-based results**		
Time-marking	**RLC**	**RLC**	**RLC (UB)**	**RLC (LB)**
H_1	1215	1231	1494	1025
H_2	1160	1109	1347	922
H_3	1130	1039	1263	863
H_4	1120	1089	1324	905
H_5	1130	1118	1358	930
H_6	1180	1045	1271	868.
H_7	1315	1276	1547	1062
H_8	1650	1640	1979	1373
H_9	1770	1673	2018	1401
H_{10}	1850	1878	2258	1578
H_{11}	1860	1864	2241	1565
H_{12}	1700	1650	1991	1382
H_{13}	1770	1764	2125	1480
H_{14}	1880	1841	2215	1546
H_{15}	1895	1829	2201	1536
H_{16}	1900	1909	2294	1605
H_{17}	1800	1742	2099	1461
H_{18}	1800	1787	2152	1500
H_{19}	1830	1783	2147	1495
H_{20}	1780	1736	2092	1455
H_{21}	1760	1790	2156	1502
H_{22}	1650	1604	1937	1342
H_{23}	1480	1521	1838	1271
H_{24}	1320	1243	1508	1035

Table 3. *Percentage differences in computed results on RLC.*

Computation Method Used			
Time-marking	**FKM Result (kW)**	**ANN Result (kW)**	**Percentile Difference of the ANN-based result with respect to the FKM-based result**
H_1	1215	1231	1.34%
H_2	1160	1109	4.42%
H_3	1130	1039	8.07%
H_4	1120	1089	2.74%
H_5	1130	1118	1.07%
H_6	1180	1045	11.43%
H_7	1315	1276	3.00%
H_8	1650	1640	0.62%
H_9	1770	1673	5.48%
H_{10}	1850	1878	1.53%
H_{11}	1860	1864	0.19%
H_{12}	1700	1650	2.93%
H_{13}	1770	1764	0.33%
H_{14}	1880	1841	2.09%
H_{15}	1895	1829	3.46%
H_{16}	1900	1909	0.48%
H_{17}	1800	1742	3.20%
H_{18}	1800	1787	0.7%
H_{19}	1830	1783	2.58%
H_{20}	1780	1736	2.46%
H_{21}	1760	1790	1.72%
H_{22}	1650	1604	2.79%
H_{23}	1480	1521	2.74%
H_{24}	1320	1243	5.81%

In view of the results presented in Tables 2 and 3, the maximum percentage difference between FKM *versu*s ANN-based methods is 11.43%, and most other percentage differences were much below that. This implies that the ANN-based method of determining the RLC is valid.

Further, each difference observed in ANN-based method is also specified with error bounds (UB and LB). This is, however, absent in FKM-based approach. The bounds (UB and LB) designate a cone-of-error.

Corresponding to the RLC observed, the tariff policy can be specified in both cases; however, in case of ANN-based approach such a policy is also bound-specified.

Tariff levied in an electric power-utility is directly decided by the load-demand considerations involved.

That is, demand forecast in electric utility systems is directly dependent on technoeconomic considerations of the underlying technology investments and economic plus demographic variables (such as, GDP, population and household growth). Further, the demand profile could vary with customer type (domestic, commercial, industrial and agricultural); and, weather could also play season-based roles [7].

Relevantly, the actual electricity sales in an area (consistent with annual demand from the end-use) are tariff-specific. That is, suppose the electrical power demand per day is P_D (kW). Relevant cost-recovery (*via* tariff) should meet the RoI expectations. The energy consumption per day (E_D) can be deduced from the RLC as follows: $\left[\int_1^{24} P_D \, dP_D\right]_{RLC} = \mathcal{E}_D$ (in kWh) ; and, denoting the cost involved as ($_D$) in dollars of the ith area (i = 1,2,3,…), $[\gamma_D]_{total} = \sum_i k_i (\mathcal{E}_D)_i$ where k_i is a coefficient expressed in dollars per kWh. This coefficient is an optimal parameter decided by the utility policy towards RoI recovery. Pertinent optimization criteria can be decided as indicated for example in [8, 9, 10].

In all, RLC-based demand *versus* tariffication could form a best practice approach towards integrated planning for electric utility services.

6. Closure

As concluding remarks, the proposed study refers to deducing an RLC needed in power system planning scenario by duly considering the stochastic aspects of the associated random load-demand fluctuations. Hence, the RLC is specified within a statistically-appropriate and logically-justifiable cone-of-error (within a pair of upper- and lower-bounds), rather than as a deterministic profile (as done in earlier studies *via* FKM approach). An ANN-approach is presented thereof in formulating this alternative strategy towards deducing the RLC; and, the ANN-based RLC details so obtained are close to the FKM-based results, confirming the efficacy of the pursuit proposed.

The RLC data (and hence, the tariff policy) when indicated within a cone-of-error, duly accounts for the stochastic implications associated with the load-curve profile as decided by fluctuating consumption, endogenous distribution system considerations and technoeconomic randomness. As such, any tariff policy derived will have judicious basis of technoeconomics of the utility in question. (The FKM method in [1], however, yields only a rigid RLC and corresponding tariff policy). The statistical error-bound based RLC determination as proposed here, is novel and not hitherto been done (to the best of the authors' knowledge). Relevant approach could lead to an optimal integrated planning for electricity towards load-demand *versus* tariffication decisions. This is essentially pertinent to the contexts of overlaying smart-grids in the expansion of (or retrofitting in) existing infrastructure of an electric utility system [11, 12, 13].

References

[1] B. Phan. "Representative Load Curve and the Tariff Impact Analyzing", *American Journal of Energy and Power Engineering*, Vol. 2, No. 5, pp. 51-55, 2015.

[2] G. E. P. Box, G. M. Jenkins. *Time Series Analysis, Forecasting and Control*, Holden-Day. San Francisco, Calif. 1970.

[3] G. Chicco, R. Napoli, P. Postolache, M. Scutariu and C. Toader, "Customer Characterization Options for Improving the Tariff Offer", *IEEE Trans Power Syst.*, Vol 18, No. 1, pp. 381-387, Feb. 2003.

[4] B. Efron, "Bootstrap Methods: Another Look at the Jackknife", *Annals of Statistics*, Vol 7, No. 1, pp. 1 – 26, 1979.

[5] P. S. Neelakanta, D. De Groff, *Neural Network Modeling: Statistical Mechanics and Cybernetic Perspectives*, CRC Press. Boca Raton, Fl. 1994.

[6] D. De Groff, P. S. Neelakanta, "Faster Convergent Artificial Neural Networks", *International Journal of Computers and Technology*, Vol. 17, No. 1, pp. 7126-7132, 2018. https://cirworld.com/index.php/ijct/article/view/7106

[7] N. Alabbas, J. Nyangon, "Weather-based Long-term Electricity Demand Forecasting Model for Saudi Arabia: A Hybrid Approach Using End-use and Econometric Methods for Comprehensive Demand Analysis", *Proceedings of USAEE*, pp. 1 – 20, 2016. http://www.usaee.org/usaee2016/submissions/OnlineProceedings/9699-AlabbasN.pdf

[8] V. A. Venikov, *Cybernetics in Electric Power Systems,* Mir Publishers, Moscow, 1978.

[9] Australian Energy Market Operator (AEMO), "Forecasting Methodology Information Paper", *2012 National Electricity Forecasting Report, 2012.* http://www.aemo.com.au/-/media/Files/Electricity/NEM/Planning_and_Forecasting/NEFR/2012/Forecasting-Methodology-Information-Paper---2012-NEFR---Final.pdf

[10] B. Dong, Z. Li, S. M. M. Rahman, R. Vega, "A Hybrid Model Approach for Forecasting Future Residential Electricity Consumption", *Energy and Buildings*, Vol. 117, No. 1, pp. 341-351, 2016. https://www.sciencedirect.com/science/article/pii/S0378778815302735?via%3Dihub

[11] "IEEE P2030 Draft-guide for Smart-grid Interoperability of Energy technology and Information Technology Operation With the Electric Power System (EPS), and End-use Applications and Loads", *Connectivity Week*, PAR Approved, under IEEE SCC21, March 19, 2009.

[12] N. Phuangpornpitak, S. Tia, "Opportunities and Challenges of Integrating Renewable Energy in Smart Grid System", *Energy Procedia*, Vol. 34, 282-290, 2013.

[13] M. S. Thomas, J. D. McDonald *Power System SCADA and Smart Grids*, CRC Press: Taylor & Francis group, Boca Raton, FL. 2015.

Biography

Dolores De Groff, Ph.D., is an Associate Professor of Computer and Electrical Engineering & Computer Science at Florida Atlantic University in Boca Raton. She received her Ph.D. in Electrical Engineering from Florida Atlantic University in 1993. She has received several teaching awards from the same institution, coauthored a book in the neural network area and authored/coauthored other journal articles and conference papers.

Roxana Melendez holds a Master of Electrical Engineering from Florida Atlantic University and is a Ph.D. candidate at the same institution. She received her Bachelor of Electrical Engineering from the Universidad del Norte, Colombia and Specialist in Engineering Project Management from Pontificia Universidad Javeriana, Colombia. She was a full time faculty at Palm Beach State College in the area of Engineering Technology and has other teaching and industry experience.

Perambur Neelakanta is a Professor of Computer and Electrical Engineering & Computer Science at Florida Atlantic University in Boca Raton, Fl., USA. He received the Ph.D. (electrical engineering) degree in 1975 from the Indian Institute of Technology (IIT), Madras, India. He had served earlier in various academic institutions and corporations in India, (West) Germany, Malaysia and Singapore. Has has published extensively (in excess of 150 journal papers) and has authored ten books. His research interest includes: Electric Power Systems, Telecommunications (Wireline and Wireless), Neural Networks, Bioinformatics and Technoeconomics.