Day-Ahead Short-Term Load Forecasting Using the Hierarchical Temporal Memory

E.N. Osegi[1*] & B.A. Wokoma
[1]National Open University of Nigeria (NOUN)
[2]Rivers State University (RSU), Port-Harcourt, Nigeria
E-mails: [1]emmaosegi@gmail.com; [1]biobele.wokoma@ust.edu.ng
Phone: [1]+234-8025838364

ABSTRACT

Short-term load forecasting is an important area in power research that requires knowledge of the future outcomes of load demand on the power system generation for short durations such as in seconds, minutes, hours or days. In this paper, an emerging state-of-the-art machine intelligence technique called the Hierarchical Temporal Memory (HTM) is employed for day-ahead short-term load forecasting. A HTM Spatial Pooler stage is used to continually form sparse distributed representations (SDRs) from a univariate load time series data, a temporal aggregator is used to transform the SDRs into a sequential bi-variate representation space and an overlap classifier makes temporal classifications from the bi-variate SDRs through time. The performance of HTM on several electrical load time series data is presented.

Keywords: Machine Intelligence, Short-term load forecasting, Sparse Distributed Representations, Time-Series, Temporal Classification.

1. BACKGROUND AND RELATED WORK

In the domain of power systems, the short term load forecasting (STLF) problem which involves the effective and accurate estimation of future load demand in hours, days or weeks has been a widely researched topic by many experts in the field. Artificial neural network techniques including variants of feed-forward back propagation algorithms such as the extreme learning machines and deep neural networks have been applied to STLF problems. Other techniques include the use of genetic algorithms including hybrid optimizations for day-ahead forecasts and Auto-regressive Moving Average (ARMA) models and its associated variants.

In **(Dudek., 2013)**, one neuron models have been proposed for forecasting electrical load time series with promising results obtained over the Exponential Smoothing (ES) and Auto-regressive Integrated Moving Average (ARIMA) models. **Saleh et al (2016)** proposed the use of Hybrid Features Selection Method (HFS) using Genetic Algorithms and Rough Sets for optimal selection of features and reliable predictions of a popular electrical load time series competition dataset (Eunite dataset).

Sudheer and Suseelatha (2015) used a combination of image processing and statistical techniques to perform day-ahead forecast in the California (USA) and Spanish electricity markets - day-ahead forecasts is one popular aspect of the STLF technique employed by power system economic managers and operators for decision making in the power markets. **Li et al (2016)** proposed an ensemble approach based on the Extreme Learning Machine (ELM) and a partial least squares regressor for aggregating the ensemble predictor outputs with wavelength pre-processing. In **(Reddy et al., 2016)**, day-ahead electricity price forecasting for Pennsylvania- New Jersey (PJM) interconnection was conducted using the back-propagation artificial neural network (ANN) and a weighted least square (WLS) technique. They utilized the WLS state estimation (WLS-SE) technique to form a better prediction of the price data fluctuations.

In **(Niu et al., 2015)**, point short-term load forecasting was carried out based on Chaos theory and a radial basis function (RBF) neural network. **Lang et al (2015)** proposed the use of random weight initialization which can help the standard feed-forward Artificial Neural Network (ANN) converge faster. However such conventional neural networks and variants thereof often times require excessive hyper-parameter tuning and cannot make online (continual) predictions.

However, in a recent study, it has been discovered that the machine intelligent based cortical learning algorithms such as the Hierarchical Temporal Memory (HTM) can indeed be a promising technique for short-term load forecasting (Anireh & Osegi, 2018a). Notwithstanding this discovery, very little work has been done in the area of applying such algorithms to electrical load time series data for forecasting in the short-term. These algorithms are modelled closely to the way the human brain operates. They learn to form sparse distributed representations, threshold coincidence maps, inculcate the notion of time and hierarchies in space and are capable of online (continual) learning **(Ahmad & Hawkins., 2015; George & Jaros., 2007; Hawkins et al., 2010; Awad & Khanna., 2015)**. Also, as earlier mentioned, the obvious challenge of extensive hyper parameter tuning in the conventional artificial neural networks and the inability of most existing neural models to perform multiple/continual predictions makes the HTM cortical learning algorithms an attractive intelligence technique **(Cui et al., 2016)**.

2. STATEMENT OF PROBLEM

The conventional Artificial Neural Networks (ANNs) such as those trained by the back-propagation algorithm have attained reasonable successes in the area of time series forecasting in various domains, particularly as it concerns the problem of short-term load forecasting (STLF). However, the issue still persist as to how accurately such ANNs solve the STLF problem in addition to the need to avoid excessive hyper-parameter tuning and for such neural networks to make online (continual) predictions. Thus, there is still room for improvement on existing ANN methodologies and variant schemes.

3. RESEARCH OBJECTIVE

The primary objective of this research is to determine the effectiveness of the Hierarchical Temporal Memory Spatial Pooler's (HTM-SP) continual predictive ability for day-ahead short-term load forecasting (STLF) using several open source electrical load time series datasets. In light of this objective, we will attempt to evaluate how the HTM-SP context and look-ahead parameters influence the prediction or forecast results obtained and then classify these predictions using an overlapping temporal classification (OTC) scheme; the proposed scheme will be based on a variant of an existing open source cortical learning toolbox.

4. METHODOLOGY

4.1 Hierarchical Temporal Memory

Hierarchical Temporal Memory (HTM) is specifically a constrained machine intelligence neural network technique for continual learning tasks **(Hawkins et al, 2017)**; its principle is based on the formation of Sparse Distributed Representations and then learning to make continual predictions from these representations using the theory of biology and neuroscience **(Hawkins et al., 2010; Hawkins et al., 2017)**. Most if not all conventional Artificial Neural Networks (ANNs) do not posses these important properties; one important point to note here is that most ANNs require separate training and testing dataset instead of continually learning and predicting on the training dataset.

In a typical HTM network, a Spatial Pooler (SP) stage is used to generate Sparse Distributed Representations (SDRs) of real world sensory input or synthetic sensory-like data and then a Temporal Pooler (TP) stage is used for making predictions on the SDRs formed by the HTM-SP. These SDRs are the basic data structures of any HTM neural network and capture the adaptive learning units used in the neocortex – the seat of intelligence in the brain. The idea of SDRs was based on an earlier work on the notion of sparse coding earlier proposed in **(Olshausen & Field., 1996; Olshausen & Field., 2004)**.
An instance of the HTM neuron model is as shown in Fig.1. This neuron model is inspired by neuroscience studies of activity-dependent synaptogenesis which hypothesizes that the growth and origin of the biological synapses is stimulated by an external sensory signal **(Zito & Svoboda., 2002)**. In the diagram of Fig.1, the green blobs denote proximal synapses which linearly summed to produce a feed-forward activation while a set of corresponding distal synapses are denoted by segments of blue blobs that are typically or-ed (logically summation using Boolean algebra) to give a spiking (dendritic) neuron activation when they exceed a recognition threshold (denoted by a Sigma sign). It is believed that feedback and context experiences are formed using these distal connections.

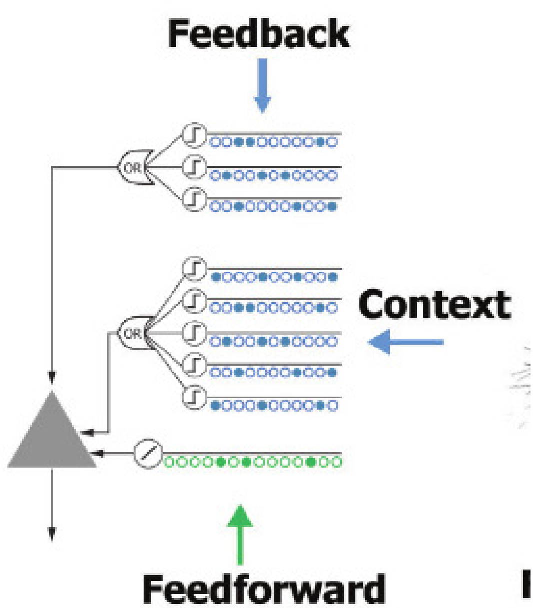

Fig.1: A typical HTM Neuron Model: adapted from (Cui et al., 2017).

4.2 Spatial Pooling in Hierarchical Temporal Memory

In HTM, spatial pooling is performed using the notion of SDRs followed by competitive Hebbian learning rules, a Homeostatic excitability control, and an overlapping mechanism for deriving candidate or winner SDR patterns via inhibition **(Cui et al., 2017)**. SDRs are formed by activating or deactivating a set of potential synapses. These synapses are grouped into a set of mini-columns and are spread out in a hypercube based on a set of predefined rules.

Consider a group of mini-columns with a set of potential connecting logical synapses or neurons; these potential connections may be initialized accordingly as:

$$\Pi_i = \left\{ j \mid \mathrm{I}\left(x_j ; x_i^c , \gamma\right) \ \& \ \mathrm{Z}_{ij} < \rho \right\} \quad (1)$$

where,

j = HTM neuron location index in the mini-column

i = mini-column index

x_j = location of the jth input neuron (synapses) in the input space

x_i^c = location centre of potential neurons (synapses) of ith mini-column in a hypercube of input space

γ = edge length of x_j

ρ = fraction of inputs within the hypercube of input space that are potential connections

Z_{ij} = represents a uniformly distributed random number between 0 and 1

I = an indicator function

The indicator function is typically described by Eq.2:

$$\mathrm{I}\left(x_j ; x_i^c , \gamma\right) = \begin{cases} 1, & if \ \ x_j \subset x_i^c \\ 0, & otherwise \end{cases} \quad (2)$$

A set of connected synapses are described by a binary matrix, W, which is formulated by conditioning the synapses to a permanence activation rule as:

$$W_{ij} = \begin{cases} 1, & if \ \ D_{ij} \geq \theta_c \\ 0, & otherwise \end{cases} \quad (3)$$

where,

D_{ij} = independent and identically distributed (i.i.d) dendrite synaptic permanence values from the jth input to the ith mini-column

θ_c = synaptic permanence threshold

The i.i.d synapse permanence values are described by Eq.4 as:

$$D_{ij} = \begin{cases} U(0,\ 1), & if \ \ j \in \Pi_i \\ 0, & otherwise \end{cases} \quad (4)$$

Where a natural topology exists, neighbourhood mini-columns may be inhibited in accordance to the relation given in Eq.5 otherwise a global inhibition parameter is simply used.

$$N_i = \{ j \mid \| y_i - y_j \| < \phi,\ j \neq i \} \tag{5}$$

where,
y_i = is the *i*th HTM-SP mini-column
y_j = is the *j*th HTM-SP mini-column
i, j = mini-column indexes
ϕ = inhibition radius control parameter

For creating associations with input patterns, feed-forward inputs to each of the generative spatial mini-columns are computed using a matching technique called the overlap; this concept is diagrammatically illustrated in Fig.2. The overlap is computed as:

$$o_i = b_i \sum_j W_{ij} z_j \tag{6}$$

where,
b_i = is a positive boost factor for exciting each HTM-SP mini-column
z_j = input pattern vector seen by the generative HTM neuron

$$\underset{Input\ \ Vector}{[1\,0 \quad \cdots \quad 0\ 1]} \bullet \underset{Spatial\ \ Pooling}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}} = \begin{matrix} Overlap\ \ counts & & SP\ \ outputSDR \\ [8\,2 \quad \cdots \quad 10\ \ 0] & \mapsto & [1\,0 \quad \cdots \quad 1\ \ 0] \end{matrix}$$

Fig.2: An Illustrative concept of Overlap in an HTM-SP; Source: Ahmad & Hawkins (2015).

Using Eq.6, we can calculate the activation of each SP mini-column as:

$$a_i = \begin{cases} 1, & if\ \ o_i \geq \mathrm{Z}(V_i, 100 - s)\ \&\ o_i \geq \theta_{stim} \\ 0, & otherwise \end{cases} \tag{7}$$

$$V_i = \{ o_i \mid j \in N_i \} \tag{8}$$

where,
s = target activation density (sparsity)
Z = a percentile function
θ_{stim} = a stimulus threshold

The HTM-SP uses a learning rule inspired by competitive Hebbian learning for reinforcing dendrite permanence values (Cui et al., 2017). The learning rule can be calculated from the formula given in Eq.9:

$$\Delta D_{ij} = p^{+} D_{ij} \circ A^{t-1} - p^{-} D_{ij} \circ (1 - A^{t-1}) \quad (9)$$

where,
p^{+} = positive permanence value increment
p^{-} = negative permanence value increment
A^{t-1} = activation state at time step, t

Finally, boost updating in HTM-SP follows the homeostatic excitability control mechanism comparable to that observed in cortical neurons **(Davis, 2006)**. Boosting is accomplished in HTM-SP using the following model equations (Eq.10-Eq.12):

$$\overline{a}_i(t) = \frac{(T-1) * \overline{a}_i(t-1) + a_i(t)}{T} \quad (10)$$

$$< \overline{a}_i(t) > = \frac{1}{|N_i|} \sum_{j \in N_i} \overline{a}_j(t) \quad (11)$$

$$b_i = e^{-\beta(\overline{a}_i(t) - < \overline{a}_i(t) >)} \quad (12)$$

where,
$\overline{a}_i$ = time averaged activation over the last T SDR inputs,
T = an integer number denoting the number of Monte Carlo trials to obtain a reasonable activation estimate.
$a_i(t)$ = the current activity of the ith mini-column at time step t.
β = a positive parameter that controls the strength of the adaptation effect

As mentioned in (Cui et al., 2017), "such calculations have been used in previous models of homeostatic synaptic plasticity" (see Clopath et al., 2010; Habenschuss et al., 2013).

4.3Temporal Classifier

In the proposed HTM system, feed-back associations are built from the HTM Spatial Pooler (SP) SDRs using a temporal overlap classifier. The Temporal classifier uses the overlap technique which is similar to Eq.6; however predictions are made by performing a match between a set of past SDR observations (used as context) and the current SDR observation. The temporal overlaps through time are obtained using Eq.13:

$$o_{j_t} = \sum_{j_t} W_{j_t}^{sp} W_{(k-N_c):j_t}^{sp}, \quad N_c < k \le j_t, \tag{13}$$

where,

N_c = Number of past sample SDRs used as context

k = size of the temporal aggregated (bi-variate) sequence through time

j_t = temporal aggregation index number

$W_{j_t}^{sp}$ = bi-variate SDRs after temporal aggregation

4.8.1 Temporal aggregation of Sparse Distributed Representations

Temporal aggregation is used in the HTM-SP to build a cause-and-effect data sequence from the SDRs formed initially and then used for an overlapping temporal classification (OTC); such sequences have been assumed to possess a bi-variate representational requirement – indeed HTM using an OTC scheme has been proven to be very effective in certain very advanced tasks such as drug discovery (Anireh & Osegi, 2018b). In HTM-SP, adding more variables increases the degree-of-freedom for making effective overlap matches. The temporal aggregation procedure used in the forecast analysis is as follows:

Step1: Form a single-column vector matrix of length 1:*N* having a with a width of 1, where *N* represents the number of sampled sequences SDRs obtained from the HTM-SP stage. The elements in this matrix contain the indexes for temporal aggregation.

Step2: For each element in the matrix formed in Step 1 greater than 1, perform a modulus operation such that if a remainder exists for the considered element we skip that element, otherwise we select the element; this operation results in single-column vector matrix of length approximately equal to 1:*N*/2. The elements in this matrix contain the set of even indexes in the matrix obtained from Step1 at time instance, *t*. We call this set $A_{t(1)}$.

Step3: For all elements in the set $A_{t(1)}$, form a concatenation of $A_{t(1)}$ with $A_{t(1)}$ 1-step behind as $\{A_{t(1)} \quad A_{t-1(1)}\}$; this concatenation represent the temporal aggregator index set. We call this set of indexes $A_{t(agg)}$.

Step4: Using $A_{t(agg)}$ as index sequence, extract SDR patterns obtained from the HTM-SP stage in a temporal aggregated fashion and then perform overlap temporal classification through time.

4.4 System for Electrical Load Demand Forecasting

An adaptive Systems-Level architecture for day-ahead load forecasting is as shown in Fig.3. The system includes an encoder for transforming sensory input signals into a binary representation that will be suitable for spatial pooling; this is followed by the Hierarchical Temporal Memory Spatial Pooler (HTM-SP) which forms sparse distributed representations (SDRs) of the binary representations using a generative procedure and then the SDRs are temporally aggregated and fed to an overlapping temporal classification (OTC) scheme which serves as the classifier (see section 4.3); the performance of the HTM can then be interpreted in terms a metric such as the Mean Absolute Percentage Error (MAPE).

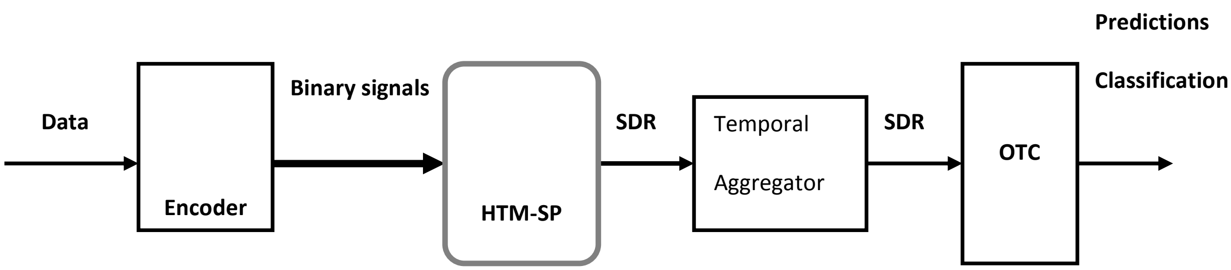

Fig.3: System-Level architecture for day-ahead forecasting using the HTM-SP

5. EXPERIMENTAL DETAILS AND DATA SOURCE

For the experiments with HTM, we have only considered the maximum daily power reading i.e. we take the maximum power for the 24-hour duration of each day; the task required here is to continually predict n-days ahead, the power demand of the power system network based on the data provided. The use of only the maximum power demands makes it difficult for the HTM-SP to make predictions but also has the advantage of dimensionality reduction as the data is then transformed to a univariate time series; we reduce this difficulty by using the temporal aggregation procedure earlier introduced in Section 4 (Sub-section 4.3.1) to form a bivariate distribution for the HTM-SP to learn from.

The experimental tests are conducted using three different electrical power demand (load) time series datasets:

The first two of these datasets comes from the Eunite Competition datasets organized by the Centre for Intelligent Technology Slovakia; it includes power readings for the years 1997 and 1998 containing a daily MW power reading for 24 hours and recorded at 30 minutes intervals; special days such as Holidays and environmental parameters such as Temperature are also provided. The datasets can be obtained from (http://eunite.org). This dataset is open source and has also been used earlier in **Saleh et al (2016).** The third dataset is based on electric load time series dataset of Polish Power System from 2002 to 2004 **(Dudek, 2013)**; this time series is similar to that of the Eunite competition dataset but with special labels for workdays or weekends. The parameters used for the HTM-SP simulations are provided in the Appendix. The software framework for simulation studies can be obtained from the MATHWORKS MATLAB central code repository; this framework is a variant of HTM-MAT, a cortical learning toolbox earlier developed in (Anireh & Osegi, 2017).

6. EXPERIMENTAL RESULTS

6.1 Experimental Results Using the Eunite datasets

The mean absolute percentage error (MAPE) values for single day-ahead forecasts using the HTM-SP on the Eunite datasets are given in Fig.4-5. This represent a fluctuation of about 0.05% to 0.10% (Fig4.4) for the Eunite 1997 competition data and 0.04% to 0.09% for the Eunite 1998 competition data.

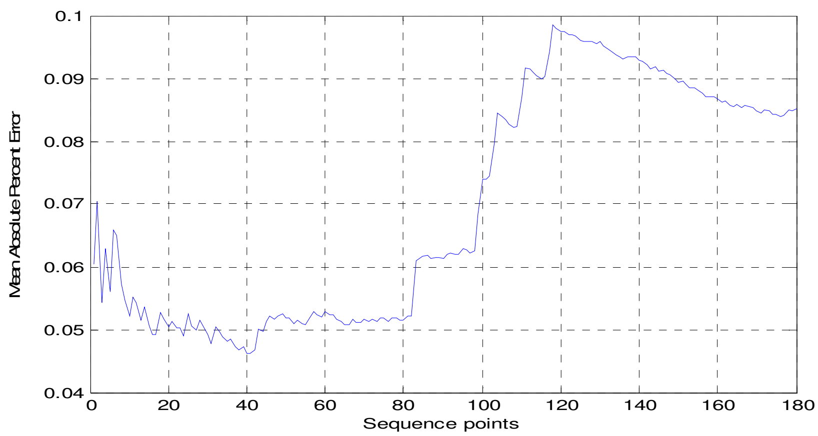

Fig.4: Error Performance using the 1997 Eunite competition dataset for 1 day-ahead forecast

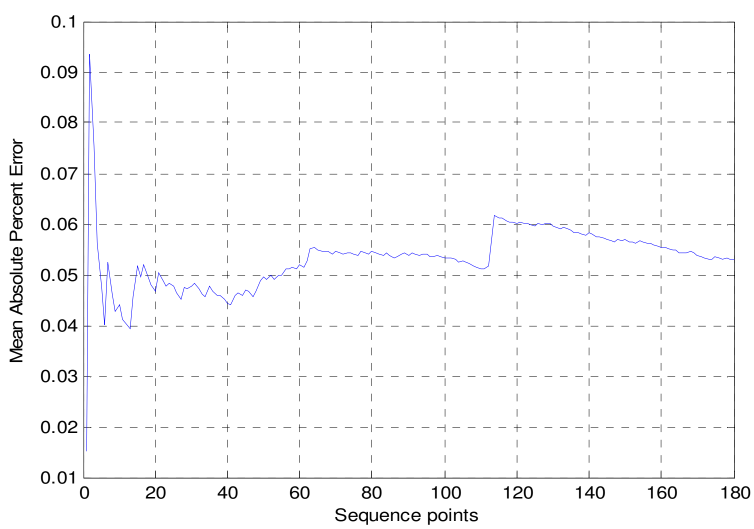

Fig.5: Error Performance using the 1998 Eunite competition dataset for 1 day-ahead forecast

6.2 Experimental Results Using the Polish dataset

In this section the MAPE values are reported for different n-step forecasts. In the first instance, MAPE values for 1 day-ahead HTM-SP predictions are reported in Fig.6. This represents a fluctuation of about 0.05% to 0.35%. In the second instance, MAPE values for 7 days-ahead forecast are as shown in Fig 7; the MAPE values fluctuate from about 0.05 to 0.35.

In Tables 1 and 2, the results of the maximum value of MAPE for the case of 7 day-ahead and 1-day ahead forecast of Polish electrical load times series data is compared to that reported in (Dudek, 2013) and (Dudek, 2015) respectively using other techniques; this comparative report clearly shows that the HTM-SP will outperform all these other algorithms.

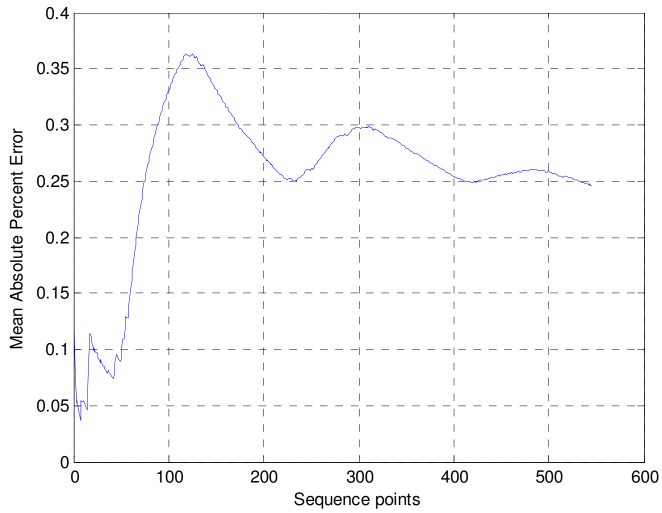

Fig.6: Error Performance using the 2002-2004 Polish power dataset for 1 day-ahead forecast

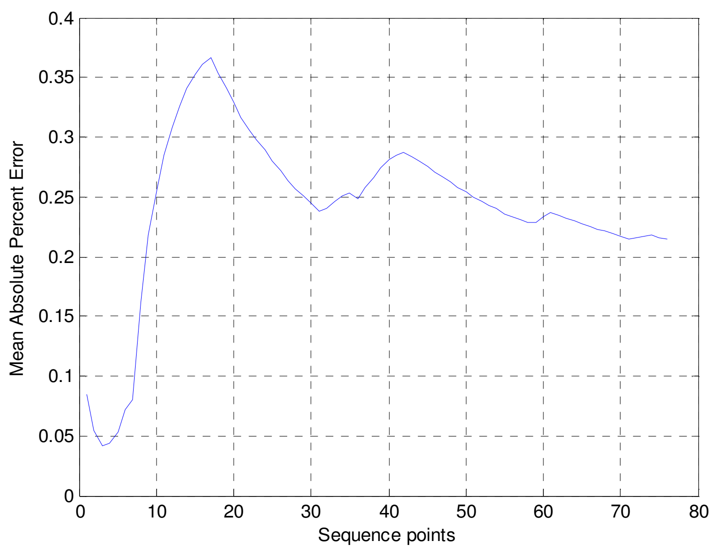

Fig. 7: Error Performance using the 2002-2004 Polish power dataset for 7 day-ahead forecast

Table1:
Reported MAPE values for 7-day ahead forecasts using Polish dataset and different techniques (Dudek, 2013)

Technique	MAPE value (%)
ANN	1.44
ARIMA	1.82
ES	1.66
Naïve	3.43
HTM-SP (proposed)	≈ 0.37

Table2
Reported MAPE values for 1-day ahead forecasts using Polish dataset and different techniques (Dudek, 2015); these values are mean value predictions for the months of January and July

Technique	MAPE value (%)
RF	1.16
CART	1.42
Fuzzy CART	1.37
ARIMA	1.91
ES	1.76
ANN	1.14
HTM-SP (proposed)	≈ 0.36

7. CONCLUDING REMARKS

Machine intelligence algorithms such as the Hierarchical Temporal Memory (HTM) based on the Cortical Learning Algorithms, presents an opportunity for industry and academic researchers in power systems to explore the possibility of using more responsive neural models for power demand forecasting. HTM can effectively learn patterns from the data using a continual learning spatial-temporal structured algorithm and help accurately and timely predict the load time series of a power system; this can save power system operators millions of naira and prevent unwarranted interruption of electrical power service. In this research, the HTM Spatial Pooler (SP) with an overlapping temporal classification (OTC) technique has been employed to the problem of short-term load forecasting. Experiments have been performed using electrical load time series datasets from the Eunite Competition and the Polish Power System. The results of these experiments indicate that the HTM-SP can continually predict the maximum load demand giving reasonable error accuracies.

However, large fluctuations in error values may result due to the effect of seasonality – these effects leads to the peaks and troughs noticed in the MAPE error response plots in Section 6 but they are not so critical as they fall under a much lower level than other techniques reported in the literature (see Table.1). Thus, continual learning/prediction algorithms can give better error accuracies even without the inclusion of helper labels such as holidays or special events. It is therefore recommended that short-term load forecasting algorithms use techniques that encourage continual learning.

Future work should focus on adapting HTM to conventional algorithms such as Genetic Programming to make the results obtained by its prediction mechanism more model expressive and investigating the possibility of other cortical-like variants that are also as effective or more accurate than the HTM, but with a much simpler neural architecture.

Also, further experimentation may be necessary including real-time/web-based open source HTM implementations to further validate the applicability of this novel machine intelligence technique and its variants.

8. CONTRIBUTIONS TO KNOWLEDGE

This research study has made three original contributions:

- The introduction of a novel classifier for univariate time series electrical load demand forecast based on an adaptive temporal aggregator overlapping temporal classification (OTC) scheme.
- The investigation of two new HTM-SP external parameters - the context and look-ahead parameters
- Adaptation of a variant of an existing open source cortical learning toolbox (HTM-MAT), for the task of short-term load forecasting.

REFERENCES

1. Ahmad, S., & Hawkins, J. (2015). Properties of Sparse Distributed Representations and their Application to Hierarchical Temporal Memory. arXiv preprint. *arXiv*, *1503*.
2. Anireh, V. I., & Osegi, E. N. (2017). HTM-MAT: An online prediction software toolbox based on cortical machine learning algorithm. *arXiv preprint arXiv:1708.01659*.
3. Anireh, V.I.E., & Osegi, E.N. (2018a). The Effect of Variable Size Order Hierarchical Temporal Memory on Short-Term Load Forecasting Load Forecasting. Journal of Digital Innovations & Contemp Res. In Sc., Eng & Tech, 6(1), 39-44.
4. Anireh, V.I.E., & Osegi, E.N. (2018b). An Online Cortical Machine Learning Artificial Intelligence Technique for Drug Discovery. Toxicology Digest, West African Society of Toxicology, 2(1), 1-12.
5. Awad, M., & Khanna, R. (2015). Deep Learning. In *Efficient Learning Machines* (pp. 167-184). Apress, Berkeley, CA.
6. Clopath, C., Büsing, L., Vasilaki, E., & Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nature neuroscience*, *13*(3), 344.
7. Cui, Y., Ahmad, S., & Hawkins, J. (2016). Continuous online sequence learning with an unsupervised neural network model. *Neural computation*, *28*(11), 2474-2504.

8. Cui, Y., Ahmad, S., & Hawkins, J. (2017). The HTM Spatial Pooler—A Neocortical Algorithm for Online Sparse Distributed Coding. Frontiers in Computational Neuroscience, 11.
9. Davis, G. W. (2006). Homeostatic control of neural activity: from phenomenology to molecular design. Annu. Rev. Neurosci., 29, 307-323.
10. Dudek, G. (2013, June). Forecasting time series with multiple seasonal cycles using neural networks with local learning. In International Conference on Artificial Intelligence and Soft Computing (pp. 52-63). Springer, Berlin, Heidelberg.
11. [dataset] Dudek, G. (2015). Short-term load forecasting using random forests. In Intelligent Systems' 2014 (pp. 821-828). Springer International Publishing.
12. George, D., & Jaros, B. (2007). The HTM learning algorithms. Mar, 1, 44.
13. Hawkins, J., Ahmad, S., & Dubinsky, D. (2010). Hierarchical temporal memory including HTM cortical learning algorithms. Techical report, Numenta, Inc, Palto Alto. https://web.archive.org/web/20110714213347/http://www.numenta.com/htm-overview/education/HTM_CorticalLearningAlgorithms.pdf
14. Hawkins, J., Ahmad, S., Purdy, S., & Lavin, A. (2016). Biological and machine intelligence (BAMI). Initial online release 0.4.
15. Habenschuss, S., Puhr, H., & Maass, W. (2013). Emergence of optimal decoding of population codes through STDP. Neural computation, 25(6), 1371-1407.
16. Lang, K., Zhang, M., & Yuan, Y. (2015). Improved Neural Networks with Random Weights for Short-Term Load Forecasting. PloS one, 10(12), e0143175.
17. Li, S., Goel, L., & Wang, P. (2016). An ensemble approach for short-term load forecasting by extreme learning machine. Applied Energy, 170, 22-29.
18. Niu, D., Lu, Y., Xu, X., & Li, B. (2015). Short-Term Power Load Point Prediction Based on the Sharp Degree and Chaotic RBF Neural Network. Mathematical Problems in Engineering, 2015.
19. Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature, 381(6583), 607.
20. Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. Current opinion in neurobiology, 14(4), 481-487.
21. Reddy, S. S., Jung, C. M., & Seog, K. J. (2016). Day-ahead electricity price forecasting using back propagation neural networks and weighted least square technique. Frontiers in Energy, 10(1), 105-113.
22. Saleh, A. I., Rabie, A. H., & Abo-Al-Ez, K. M. (2016). A data mining based load forecasting strategy for smart electrical grids. Advanced Engineering Informatics, 30(3), 422-448.
23. Sudheer, G., & Suseelatha, A. (2015). Short term load forecasting using wavelet transform combined with Holt–Winters and weighted nearest neighbor models. International Journal of Electrical Power & Energy Systems, 64, 340-346.
24. Zito, K., & Svoboda, K. (2002). Activity-dependent synaptogenesis in the adult mammalian cortex. Neuron, 35(6), 1015-1017.

APPENDIX

Table A.1: HTM-SP PARAMETERS

Parameter	Value
Number of Columns	250
Initial Synaptic Permanence	0.21
Dendritic segment activation threshold	8
Boost factor	100
Synapse permanent increment	0.1
Synapse permanent decrement	0.1
Number of past sequences used as Context	2