Non-Intrusive Load Disaggregation of Industrial Cooling Demand with LSTM Neural Network

Non-Intrusive Load Disaggregation of Industrial Cooling Demand with LSTM Neural Network

Simone Eiraudo*, Luca Barbierato*, Roberta Giannantonio†, Edoardo Patti*, Andrea Lanzini*, and Lorenzo Bottaccioli*
**Energy Center Lab*, Politecnico di Torino, Turin, Italy. Email: name.surname@polito.it
†*Data Office*, TIM S.p.A., Italy, Email: name.surname@telecomitalia.it

***Abstract*—As the telecommunication industry becomes more and more energy intensive, energy efficiency actions are crucial and urgent measures to achieve energy savings. The main contribution to the energy demand of buildings devoted to the operation of the telecommunication network is cooling. The main issue in order to assess the impact of cooling equipment energy consumption to support energy managers with awareness over the buildings energy outlook is the lack of monitoring devices providing disaggregated load measurements. This work proposes a Non-Intrusive Load Disaggregation (NILD) tool that exploits a literature-based decomposition with an innovative LSTM Neural Network-based decomposition algorithm to assess cooling demand. The proposed methodology has been employed to analyze a real-case dataset containing aggregated load profiles from around sixty telecommunication buildings, resulting in accurate, compliant, and meaningful outcomes.**

***Index Terms*—Non-Intrusive Load Disaggregation, Long Short-Term Memory, Cooling Load**

I. Introduction

The Information and Communication Technology (ICT) branch is experiencing an exponential growth of energy demand, being currently responsible for approximately 2% of global GHG emissions [1] and for around 7% of the worldwide electricity demand [2]. Telecommunication network and their management buildings, such as telephone switches, which are generally referred to as Central Offices, and Data Centers, are one of the main contribution of the energy consumption in the ICT sector. In fact, Central Offices, Data Centers, and their transmission network are responsible of about 2% of the global electrical demand in 2019 [3]. For instance, the Italian telecommunication sector was responsible for 3.863 GWh of electrical demand in 2018 [4], which corresponds to 1.27% of the total national electricity consumption.

Owing to its energy intensiveness and to its on-going growth, the telecommunication sector is becoming more and more of interest for what concerns energy consumption and GHG emissions reduction policies [5]. This interest is attested by institutional directives [2], [6] and by companies efforts to achieve energy savings [7], [8]. Specifically, energy consumption reduction is addressed by employing new generation Central Offices and Data Centers [9], [10]. Nevertheless, the energy consumption reduction of telecommunication buildings must be achieved by focusing as well on supporting energy efficient renovation of existing buildings. For this reason, an accurate analysis of these buildings inefficiencies and the estimation of potential savings to provide retro-fit scenarios represent a fundamental task to be addressed.

Recently, energy research community discussed many issues related to the energy efficiency enhancement of Central Offices and Data Centers, such as structural transformation, technological innovation [11], and renewable energy sourcing [12]. Nevertheless, the main issue in energy efficiency of Central Offices and Data Centers is cooling. In facts, the cooling system is responsible for impressive shares of these buildings electricity demand due to the heat generation of ICT equipment that must respect overheating constraints. In order to detect energy inefficiency, assess the potential savings related to retrofit scenarios, and enable efficiency actions, strong awareness about buildings energy outlook is needed. To this extent, smart monitoring of industrial sites is becoming more and more popular, as it allows to retrieve reliable and detailed electricity consumption measurements. On the other hand, in order to achieve effective energy consumption reduction, it is desirable to extend analysis to the largest possible number of sites, by means of proper non-intrusive analysis tool.

For the above mentioned reasons, this paper introduces a methodological pathway to perform Non-Intrusive Load Disaggregation (NILD) of Central Offices and Data Centers energy consumption into its different load contributions. The proposed approach is applied to a real-case dataset, containing load profiles from around sixty telecommunication management buildings located in Italy. The main contribution of the proposed methodology is to enable the NILD application to load profiles characterised by low sampling rate. In facts, most of the existing methodologies take advantage of high sampling rates, which may not be available for many industrial sites. Moreover, these techniques have been rarely applied to industrial buildings. This determined a scarcity in the literature referring to this important sector. In particular, the proposed methodology, taking advantage of both specific expertise in the sector and Long Short Term Memory (LSTM), represents the first attempt to deal with the NILD challenge in the telecommunication management buildings sector.

II. Related Works

Reliable and detailed information of the cooling load of industrial sites represents a starting point for taking effective energy efficiency actions. In fact, real-time monitoring enhances

Simone Eiraudo acknowledges support from TIM S.p.A. through the PhD scholarship

optimal energy management to improve energy efficiency of industrial premises. Nevertheless, smart monitoring requires many sensors to be installed, determining high deployment and management costs. Thus non-intrusive techniques are required to analyze the energy behaviour of not sensorized buildings. These techniques called Non-Intrusive Load Disaggregation (NILD) provides accurate estimates of electrical load disaggregation from the overall electrical building consumption. NILD techniques may be designed to assess the energy demand of individual appliances or to disaggregate the total load in subcategories.

To this purpose, energy research community explored different automated analysis tools, such as Aided Integer Linear Programming [13] and Factorial Hidden Markov model [14], to assess electrical consumption of individual appliances in residential field. Yan et al. [15] exploited a model driven simplified approach, considering energy balance of buildings and linear regression models, tuned accordingly to an optimization algorithm. The goal of such a study was to distinguish among three groups of cooling load different end-uses. Assessment of cooling demand is as well addressed by Xiao et al. in [16] where authors employed a Random Forest algorithm to determine the load disaggregation due to four fundamental contributions, namely equipment, occupants, fresh air, and building envelope.

It is worth noting that the predominant approach in this branch of research is data-driven. To this extent, Machine Learning (ML) techniques are widely adopted algorithm to perform NILD. Among this diversified family of techniques, a leading role has been assumed by Artificial Neural Networks (ANN). A relatively simple Feed-Forward Neural Network (FFNN) has been used by Racines et al. [17] to assess the impact of lights, PC and other appliances electrical demand in educational buildings. Nevertheless, researchers frequently selected more complex Neural Network models. For instance, LSTM [18], [19] is a particular model of Recurrent Neural Network, which was designed with specific features, namely the forget gate and the cell internal state, to be capable of preserving long-term dependency information. LSTM was employed to disaggregate households electrical load by Xia et al. [20], to determine individual appliances consumption.

The aforementioned approaches foster a number of practical energy efficiency oriented measures, such detecting malfunctioning devices, monitoring individual loads to deepen awareness over the energy demand and providing specific renovation scenarios or suggesting energy saving actions or to improve demand side response. Finally, NILD has been employed as well to improve performance of energy demand forecasting models. For instance, Lin et al. [21] used Sparse Coding algorithm to estimate cooling sub-loads. In turns, this components were fed as input to a NN, determining enhanced short term load forecasting accuracy.

It is worth remarking that most of the attention have been addressed to the residential sector. Nevertheless, widespread of NILD is desirable as well regarding the industrial sector. This paper deals with the specific task of providing a methodology to be applied in a particularly meaningful sector, namely the Telecommunication sector. This is achieved by means of a novel pathway, that is a mixed algorithm employing both Machine Learning tools, namely LSTM neural network, and a base load decomposition block, which takes advantage of the specific domain expertise in the sector. The designed LSTM block allows the employment of low sampling rate data, representing an important step forward the NILD techniques widespread. With respect to the scarce research investigating the field to date, this methodology shall foster the state-of-the-art of NILD techiniques and enhance application of this fundamental tool to the industrial sector.

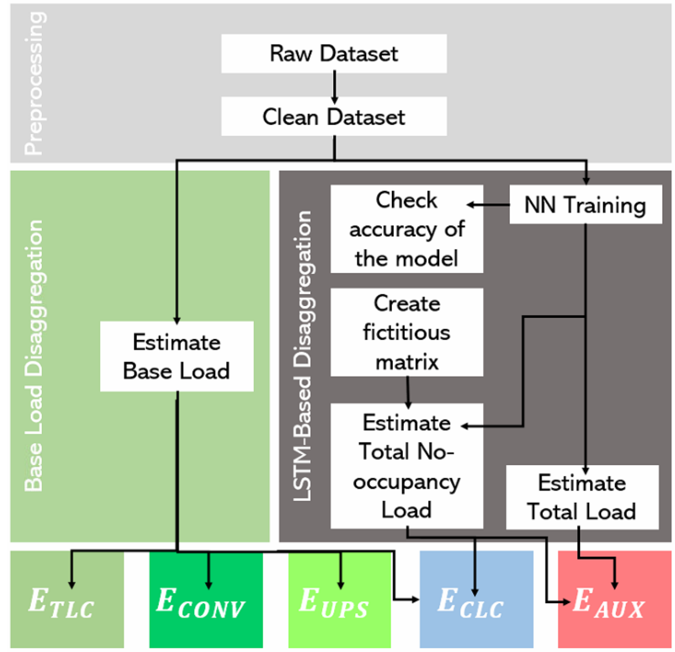

Fig. 1. The general scheme of the proposed methodology, comprehending for a literature based and a NN-based disaggregation blocks.

III. Methodology

This work is intended to support load disaggregation for telecommunication building in order to enhance energy awareness and support energy efficiency actions over these sites. Specifically, the aim of this analysis is to estimate each of the four contributions composing the telecommunication sites typical energy balance equation:

$$E_{TOT}(t) = E_{TLC}(t) + E_{DISS}(t) + E_{CLC}(t) + E_{AUX}(t) \quad (1)$$

where E_{TLC} takes into account the electrical demand from the ITC equipment, E_{DISS} represents the energy lost of the energy station for energy conversion, UPS and other required energy services, E_{CLC} is cooling system contribution to electrical load and E_{AUX} comprehends the electrical demand from the lighting system and from other services. The object of this investigation is a real-world dataset with hourly resolution. It includes the aggregated consumption time series from around sixty Central Offices and Data Centers of the most important telecommunication service provider in Italy. The dataset is integrated with weather data, specifically with temperature and solar radiation measurements. The methodological pathway foresees the following three steps, as reported in Figure 1: *i)* the Preprocessing, *ii)* the Base Load Decomposition, and *iii)* The LSTM-based Disaggregation.

A. Preprocessing

Abnormal values presence in the dataset may depend by anomaly events regarding the energy demand of the site or by measurement fallacies. Each building load profile is preprocessed in order to filter out possible outliers. This filtering step is particularly important, as it represent a preparatory step for the LSTM training. In facts, removing abnormal values and measurements noise improve the accuracy of the LSTM Neural Networks, since the training won't depend from the back-propagation of errors related to these anomalies. The task of detecting these points is achieved by employing a simple gradient-based statistical approach. This algorithm determines whether a point is a noise point in accordance to its gradient with respect to the previously recorded data. Specifically, abnormal points are defined as those featuring a gradient exceeding 3 times the standard deviation of the gradients from the whole time series.

B. Base Load Disaggregation

In this step, the preprocessed building load profiles are analysed to estimate their base load. The typical load profiles observed for Central Offices and Data Centers enlighten that the E_{TLC} contribution is tolerably constant. This behaviour is due to ITC equipment that is not affected by processed traffic [22] or other environment variables. Secondly, since the energy lost for energy conversion is proportional to the converted energy itself, it may be stated that E_{DISS} is proportional to E_{TLC}. Hence, this contribute can be assumed constant as well. As a result, Equation 1 can be written as:

$$E_{TOT}(t) = E_{TLC} + E_{DISS} + E_{CLC}(t) + E_{AUX}(t) \tag{2}$$

The two constant loads E_{TLC} and E_{DISS} represent the base load for the building load profiles time series. Load fluctuations instead are determined by variation of the the demand from lighting, or other occupancy related services, and to the cooling load. We consider that lightning and auxiliaries services determined by occupancy are not expected to affect electrical load during nights. These hypothesis hold true particularly regarding pure Central Offices and Data Centers. Since the winter weather conditions do not make it necessary for the cooling system of the sites to intervene in order to keep internal temperature below the temperature set point, the aggregated electrical demand measured during this winter night time intervals can be assumed to depend solely by the two aforementioned constant contributions. Hence, the base load can be estimated as:

$$E_{min} \approx mean(min(E_i)_d), \ i \in [1,4], \ d \in winterdays \tag{3}$$

where E_i is the energy demand from the i_{th} hourly time step from winter day d. To further distinguish within E_{TLC} and E_{DISS}, we consider other findings of recent years literature. In facts, some researches have been investigating wide dataset of telecommunication sites to provide accurate statistical data over the energy outlook of Central Officies and Data Centers. Specifically, Shebabi [23] reported the results that are elaborated and summarized in Table I, providing typical load quotas regarding the conversion losses and the energy employed for UPS. It should be pointed out that the two contributions reported in the Table, E_{CONV} and E_{UPS}, represent the two fundamental loads of the conversion losses quota introduced in the energy balance equation, E_{DISS}. Hence, we may write:

$$E_{DISS} \approx E_{CONV} + E_{UPS} \tag{4}$$

This step allow us to estimate the first two contribution of the four loads introduced in Equation 1.

TABLE I
CONVERSION LOSSES AND ELECTRICAL DEMAND BY UPS IN DATA CENTERS, ELEBORATED FROM [23]

Space Type	Typical Area	TLC	CONV	UPS
Closet	$<10m^2$	1	0.05	-
Room	$10-99m^2$	1	0.05	0.2
Localized	$50-199m^2$	1	0.05	0.2
Midtier	$200-2000m^2$	1	0.05	0.2
High-end	$>2000m^2$	1	0.03	0.1
Hyperscale	$>2000m^2$	1	0.03	-

C. LSTM-Based Disaggregation

This step represents the main innovation of this methodology. To find the optimal LSTM configuration, the preprocessed dataset is first split into a training set and a validation set following the standard procedure for Neural Network training. Hence, we considered the following inputs to be fed in the network: energy demand trend, outdoor temperature, solar radiation, day type (i.e. week day or week-end), hour of the day, and a binary variable considering whether a day was affected by the pandemic related limitations. It is worth noting that using the output variable of the model as an input normally determines better forecasting accuracy. This approach represents an auto-regressive techniques, that is tracking the dependence of the present output value from its previous values. However, LSTM auto-regressive technique is not applied to the Neural Network and it consider the solely energy consumption trend value. This allows the network to weaken the dependency of the outputs from previously recorded load values, hence limiting the impact of the eventual presence of outliers in the time serie. At the same time, a stronger dependency by the other input variables is outlined, in order to design a model capable of providing reliable reference predictions.

Furthermore, we configured the networks to have two hidden layers. The number of training epochs was limited to 600. A dropout layer was included in the network and located between the two hidden layers. This particular layer enhances

TABLE II
THE FIVE BEST ARCHITECTURES OF THE LSTM NETWORKS, RESULTING FROM THE GRID SEARCH PROCEDURE

Learning Mode	**Learning Rate**	**Learning Rate Decay**	**Look back**	**Number of neurons 1^{st} hidden layer**	**Dropout**	**Number of neurons 2^{nd} hidden layer**	**MAPE**
Mini-Batch	0.0005	$1*10^{-4}$	4	128	0.2	12	2.794
Mini-Batch	0.0002	$5*10^{-5}$	4	128	0.2	12	2.801
Mini-Batch	0.0002	$5*10^{-5}$	4	128	0.18	16	2.812
Mini-Batch	0.0002	$5*10^{-5}$	4	128	0.25	16	2.815

proper fitting of the network, as it is designed to avoid overfitting phenomena. The loss function was set as mean absolute error.

Considering these constrains, we tested a number of possible configurations of the network by modifying the other hyper-parameters of the LSTM. Specifically, we performed an iterative procedure to look for the best configuration which is generally addressed to as grid-search. The variable hyper-parameters were the learning mode, learning rate, the learning rate decay over epochs, the look back, the number of neurons in the first and in the second hidden layer and the value of the dropout layer.

The LSTM has been tested applying three learning modes, namely online learning, batch learning and mini-batch learning. In online learning, the time series is handled element by element, and the connection weights are updated in accordance to the loss function calculated after each single step. This mode may provide accurate results in a limited number of epochs, nut is exposed to the risk of being stuck on local minima, that are sub-optimal models. On the contrary, the whole training set is considered at once, and weights update is performed once per epoch. In this case, the loss function decreases much slower over epochs, but the model have higher probabilities of evolving to an optimal solution. The mini-batch learning mode is a spreading solution which represent a trade-off within the other two modes.

Once the best configuration of LSTM has been selected, the aggregated load is estimated for both the training and the validation set. We will refer to this estimated load as E^*_{TOT} In parallel, a fictitious input matrix is created. This artificial input dataset maintains the real inputs which are fed into the network to retrieve total load estimation, with the exception of the binary input regarding the day type. Specifically, this value is set to 1, corresponding to week-end and holidays, for any time step of the input matrix. This input matrix is fed into the trained network to retrieve another electrical load output. The output can be described as the energy consumption the site would have if any day of the considered time series were a week-end or holiday, that is assuming that all the contributions to energy demand which depend from occupancy are removed from the load time series which is returned by the LSTM. For this reason we will refer to this output as $E^*_{TOT,NoOccupancy}$. This means that the only responsible of the fluctuations of the resulting load is the variation of the cooling system demand. It is worth remarking that E_{CLC} represents the cooling load attributable to only the ICT equipment cooling. In facts, the energy demand related to HVAC of offices or other personnel dedicated areas is included, according to this methodology, to the E_{AUX} quota. Finally, these two quotas can be obtained by considering the equations:

$$E_{CLC} = E^*_{TOT,NoOccupancy} - E_{TLC} - E_{DISS} - E_{UPS} \quad (5)$$

and:

$$E_{AUX} = E^*_{TOT} - E^*_{TOT,NoOccupancy} \quad (6)$$

IV. RESULTS

The proposed methodology is applied to a dataset of around sixty Central Offices, located in several regions of Italy. It should be remarked that the LSTM-based Disaggregation step is performed separately for each building. In fact, the best LSTM architecture may vary according to the considered time series and to its complexity. Notice that, due to the confidentiality of the analysed data, electrical load values can not be reported. For this reason, they were hide from the plots.

A. LSTM-Training and Accuracy

The optimal architecture of LSTM models was designed by performing an iterative grid search procedure. An example of a few tested configurations from the model training regarding a site is reported in Table II. All models resulted in a enhanced final accuracy when setting a rate of decay of the learning rate. The grid-search procedure enlightened that a good value for the look-back, that is the number of previous time steps considered, is four. This means that the model is considering the input variables from the last four hours in order to estimate the energy demand of the site at a certain time step. The grid

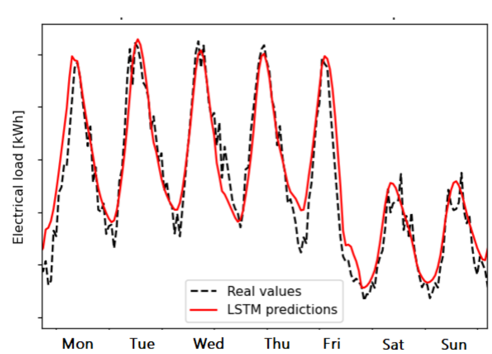

Fig. 2. Comparison within real load values and outputs of the designed LSTM model

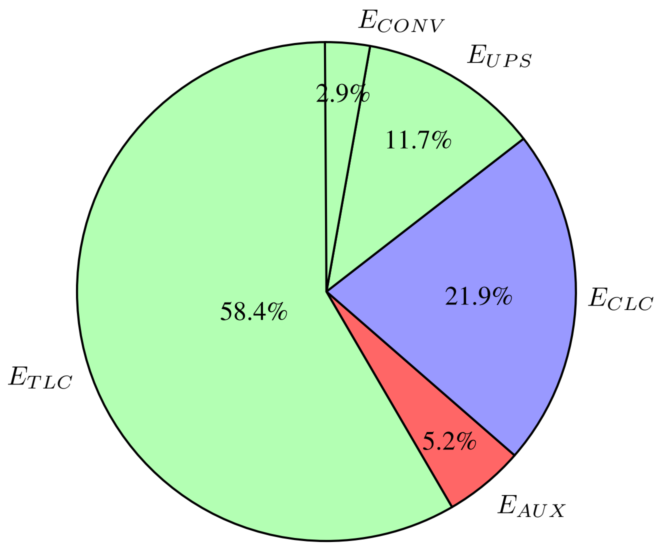

Fig. 3. Disaggregation of electricity bill from a Central Office

search procedure pointed out as well that the best performing LSTM models for this case study feature an high number of neurons in the first hidden layer, and a relatively low one in the second hidden layer. For instance, the best configuration reported in Table II is characterised by 128 and 12 in the first and second hidden layers respectively. The dropout layer s of particular interest for LSTM-training. Specifically, it enhanced reduction of the gap among the accuracy of the training set and the validation set for all the considered sites. Furthermore, in many cases the dropout layer made it possible to pop out from a local minimum, and to enhance further improvement of the model. Still, the best dropout values were in the range 0.1-0.28 for all the considered buildings.

Finally, the LSTM models generally achieved good performances, as they were good predictors of the sites electrical load. This is attested both by accurate estimated load curves shape and by low mean absolute errors with respect to real values. A comparison within the estimated total load and the recorded values is reported in Figure 2. The Mean Absolute Percentage Errors for the analyzed sites, calculated for the validation sets, ranged between 1.5% and about 10%. This variation were mostly determined by the complexity of the times series. In facts, a few time series featured strong presence of noise, while others were characterized by regular, smooth and physically rational electrical load profiles.

B. Load Disaggregation

The electrical load quotas were obtained by considering the outputs of the LSTM models and applying Equations 6 and 5, in addition to the employment of the base load disaggregation block described in III-B. Firstly, we compared the outcomes of this methodology to the values obtained by previous studies over a subset of the investigated sites. The results were compliant, even if slight differences exist regarding single consumption quotas. Generally, the quote accounting for the lighting and auxiliaries load, depending on the occupancy of sites, is not too relevant for many sites. In facts, E_{AUX} is below 10% for any building, but in some cases this quota represent less than 3% of the total energy demand of sites. On the opposite the cooling contribution resulted as a fundamental load quota, as expected. Values reported by buildings ranged between about 15% to over 25% of the total energy demand. Figure 3 reports the total consumption percentages from a Central Office for year 2019.

One of the most interesting outcome of the study is the fact that the load contributions are described for each hour of the time series. This may enlighten several meaningful issues concerning the investigated sites. For instance, Figure 4 depicts the estimated auxiliaries load over the first months of year 2020 for a Central Office. It may be easily seen that this consumption quota roughly halves at the beginning of March 2019, which was the beginning of the first national lock down due to the pandemic. A similar behaviour is reported in the whole set of buildings, with the total estimated energy demand from E_{AUX} from 2020 which is about a half than in 2019 for the vast majority of sites.

Another interesting outcome is that the lighting and auxiliaries load is represented by the weekly pattern retrieved by the model. In facts, it resulted that, in many Central Offices, E_{AUX} have higher value in the first days of the working week, while it usually has lower values and earlier decrease on Friday. This behaviour is reported by the load profile in Figure 5. Finally, the cooling load is characterised by a strong seasonal effect but daily fluctuations are evident as well. Most

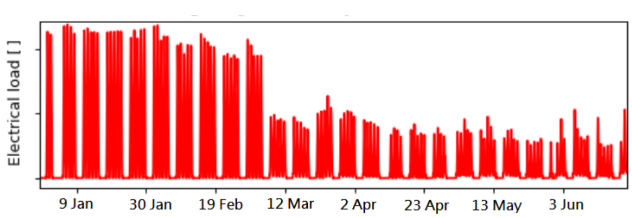

Fig. 4. Auxiliaries load over the first months of 2020, retrieved by the proposed methodology for one of the investigated TLC sites

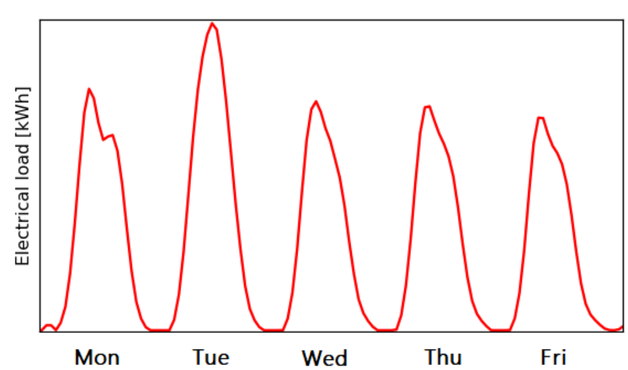

Fig. 5. Estimated auxiliaries and lighting load, retrieved by means of the proposed methodology

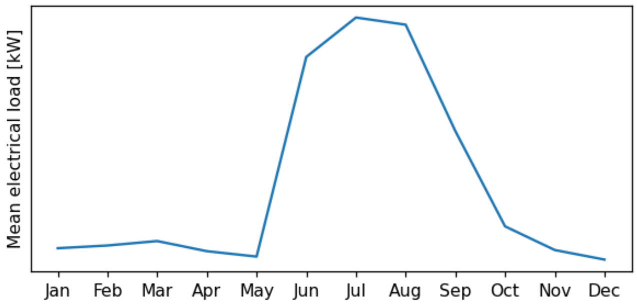

Fig. 6. Estimated cooling demand by months for a CO, retrieved by means of the proposed methodology

of the sites are characterised by a cooling load which ramp up between May and June (see Figure 6, going back to minimum consumption values in October. A few sites have longer time intervals characterised by high cooling demand. This issue is related to the geographical location of sites, with long lasting cooling season for those sites located in the warmer country regions.

V. Conclusion and future work

This work introduces a methodology to estimate the four main contributions of energy demand of Central Offices and Data Centers, namely ICT equipment load, conversion losses, cooling load, and auxiliaries and lighting load. This data-driven approach allows the design of proper and accurate LSTM Neural Network models to estimate cooling demand and auxiliaries and lighting load. The methodology, comprehending as well a base load disaggregation block relying on the specific expertise of the telecommunication sector, provides estimation of each single sub-load profile. The cooling load resulted as the most crucial issue determining energy demand fluctuations of Central Offices. This work shall be extended to an higher number of sites, in order to support energy managers and operators with fundamental information about telecommunication buildings' energy outlooks.

References

[1] "Ict sector helping to tackle climate change — unfccc." [Online]. Available: https://newsroom.unfccc.int/news/ict-sector-helping-to-tackle-climate-change

[2] M. Avgerinou, P. Bertoldi, and L. Castellazzi, "Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency," *Energies*, vol. 10, no. 10, p. 1470, 2017.

[3] Cooling. [Online]. Available: https://www.iea.org/reports/cooling

[4] "Pubblicazioni Statistiche - Terna spa." [Online]. Available: https://www.terna.it/it/sistema-elettrico/statistiche/pubblicazioni-statistiche

[5] M. A. B. Siddik, A. Shehabi, and L. Marston, "The environmental footprint of data centers in the united states," *Environmental Research Letters*, vol. 16, no. 6, p. 064017, 2021.

[6] C. Koronen, M. Åhman, and L. J. Nilsson, "Data centres in future european energy systems—energy efficiency, integration and policy," *Energy Efficiency*, vol. 13, no. 1, pp. 129–144, 2020.

[7] "Data centers are more energy efficient than ever." [Online]. Available: https://blog.google/outreach-initiatives/sustainability/data-centers-energy-efficient/

[8] "To cool datacenter servers, Microsoft turns to boiling liquid." [Online]. Available: https://news.microsoft.com/innovation-stories/datacenter-liquid-cooling/

[9] M. Manganelli, A. Soldati, L. Martirano, and S. Ramakrishna, "Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy," *Sustainability 2021, Vol. 13, Page 6114*, vol. 13, no. 11, p. 6114, may 2021.

[10] J. Ni and X. Bai, "A review of air conditioning energy performance in data centers," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 625–640, 2017.

[11] S. V. Garimella, L.-t. Yeh, and T. Persoons, "Thermal Management Challenges in Telecommunication Systems and Data Centers," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 2, no. 8, pp. 1307–1316, 2012.

[12] R. Lee, D. Pinner, K. Somers, and S. Tunuguntla, "The case for committing to greener telecom networks." 2020.

[13] M. Z. A. Bhotto, S. Makonin, and I. V. Bajić, "Load Disaggregation Based on Aided Linear Integer Programming," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, no. 7, pp. 792–796, 2017.

[14] M. Aiad and P. H. Lee, "Unsupervised approach for load disaggregation with devices interactions," *Energy and Buildings*, vol. 116, pp. 96–103, 2016.

[15] C. Yan, S. Wang, and F. Xiao, "A simplified energy performance assessment method for existing buildings based on energy bill disaggregation," *Energy and Buildings*, vol. 55, pp. 563–574, 2012.

[16] Z. Xiao, W. Gang, J. Yuan, Y. Zhang, and C. Fan, "Cooling load disaggregation using a NILM method based on random forest for smart buildings," *Sustainable Cities and Society*, vol. 74, no. February, p. 103202, 2021.

[17] D. L. Racines and J. E. Candelo, "Non intrusive load identification with power and impedance obtained from smart meters," *International Journal of Engineering and Technology*, vol. 6, no. 4, pp. 1867–1876, 2014.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[19] A. Graves, "Long short-term memory," in *Supervised sequence labelling with recurrent neural networks*. Springer, 2012, pp. 37–45.

[20] M. Xia, W. Liu, K. Wang, W. Song, C. Chen, and Y. Li, "Non-intrusive load disaggregation based on composite deep long short-term memory network," *Expert Systems with Applications*, vol. 160, p. 113669, 2020.

[21] X. Lin, Z. Tian, Y. Lu, H. Zhang, and J. Niu, "Short-term forecast model of cooling load using load component disaggregation," *Applied Thermal Engineering*, vol. 157, no. February, p. 113630, 2019.

[22] M. Sorrentino, M. Bruno, A. Trifirò, and G. Rizzo, "An innovative energy efficiency metric for data analytics and diagnostics in telecommunication applications," *Applied Energy*, vol. 242, no. March, pp. 1539–1548, 2019.

[23] Shebabi, "United States Data Centers Energy usage Report 2016," vol. 4, no. 3, pp. 879–930, 2018.