Water Demand Forecasting of District Metered Areas through Learned Consumer Representations

Adithya Ramachandran*, Thorkil Flensmark B. Neergaard[†], Tomás Arias-Vergara*,
Andreas Maier*, and Siming Bayer*

*Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nurnberg, Germany
Email: *adithya.ramachandran@fau.de
[†]Brønderslev Forsyning A/s, Virksomhedsvej 20, Brønderslev, Denmark

Abstract—**Advancements in smart metering technologies have significantly improved the ability to monitor and manage water utilities. In the context of increasing uncertainty due to climate change, securing water resources and supply has emerged as an urgent global issue with extensive socioeconomic ramifications. Hourly consumption data from end-users have yielded substantial insights for projecting demand across regions characterized by diverse consumption patterns. Nevertheless, the prediction of water demand remains challenging due to influencing non-deterministic factors, such as meteorological conditions. This work introduces a novel method for short-term water demand forecasting for District Metered Areas (DMAs) which encompass commercial, agricultural, and residential consumers. Unsupervised contrastive learning is applied to categorize end-users according to distinct consumption behaviors present within a DMA. Subsequently, the distinct consumption behaviors are utilized as features in the ensuing demand forecasting task using wavelet-transformed convolutional networks that incorporate a cross-attention mechanism combining both historical data and the derived representations. The proposed approach is evaluated on real-world DMAs over a six-month period, demonstrating improved forecasting performance in terms of MAPE across different DMAs, with a maximum improvement of 4.9%. Additionally, it identifies consumers whose behavior is shaped by socioeconomic factors, enhancing prior knowledge about the deterministic patterns that influence demand.**

Index Terms—**Water Demand Forecasting, End-Consumer Clustering, Contrastive Learning, Time Series Analysis**

I. INTRODUCTION

Water, a vital resource for life, necessitates effective water demand management to ensure ongoing and safe access. Water Distrubution Networks (WDNs) sustain a consistent water supply across District Metered Areas (DMAs), serving diverse end-consumers such as households, industries, and farms, monitored by the temporal aggregation of smart meter data [1]. With increasing water scarcity, accurate forecasting that accounts for anthropogenic influences [2] and socio-psychographic factors [3] is essential. Although aggregating demand data at the DMA level smoothens statistics and highlights broad consumption trends (e.g., daily and weekly seasonality), it can mask unique consumption behaviors and individual variabilities [4]. With varying consumer behavior complicating real-time demand modeling, it is crucial to establish a level of spatial granularity or consumer categorization to capture broad and localized patterns for forecasting.

Traditionally, time series analysis has relied heavily on statistical models, which fit known distributions using parametric functions. These models, combined with techniques like seasonal decomposition and correlational analysis, clarify intrinsic and extrinsic time series dependencies and periodicities. However, recent advancements in Deep Learning (DL), particularly contrastive learning, have facilitated the non-parametric learning of complex time series data representations, enabling the discovery of meaningful patterns from previously unknown representations [5]. While contrastive learning has been applied to various time series data streams including electricity, traffic, and medical for tasks like clustering, and forecasting [6], [7], its potential for modeling end-consumer water demand data remains unexplored.

Water demand forecasting has progressed from statistical and hybrid models [8] to the extensive adoption of Recurrent Neural Networks (RNNs) and transformer-based models [9], [10]. The representation of input time features has also expanded beyond the time domain to include the time-frequency domain via Fourier Transform (FT), and Wavelet Transforms [11]. Wavelet scalograms, which provide frequency details along with time localization, facilitate multi-resolution signal analysis while addressing non-stationarity [12], and are increasingly utilized in time series applications [13]. Despite these advancements, accurately forecasting water demand remains a challenge in scenarios where consumer behaviors are not uniform. The impact of non-deterministic factors such as culture and economy complicates demand modeling, obscuring existing demand patterns and hindering efficient water supply management.

In this study, we introduce a two-stage framework for multi-step hourly water demand forecasting that leverages unique consumption behaviors identified through contrastive learning on smart meter data within a DMA as additional input features for forecasting. Although existing clustering methods primarily focus on urban household patterns and intra-household activities [14, 15, 16, 17], to our knowledge, this represents the first application of an unsupervised contrastive learning-based approach to delineate water demand patterns across varied consumer sectors, including commercial, agricultural, and residential areas. The patterns identified are subsequently utilized as features with the aggregated DMA demand data, allowing the model to capture both broad trends and specific

consumer-induced fluctuations. This approach significantly enhances forecasting accuracy, especially in districts without a dominant urban consumer demographic.

II. METHODOLOGY

The proposed framework is depicted in Figure 1. Let $\mathbb{S} = \{\mathbf{m}_1, \mathbf{m}_2, ..., \mathbf{m}_k\}$ be the set of k unique smart meters present in a DMA, such that $\mathbf{m}_i = [x_1^i, x_2^i, ..., x_N^i] \in \mathbb{R}^{N \times 1}$, where x_t^i is the recorded hourly water consumption (m^3/h) at time t by meter i over N hours. The total hourly consumption of a specific DMA at time t is mathematically represented as $x_t = \sum_{i=1}^{k} x_t^i$ and the downstream task of forecasting for p hours is defined as the prediction $\tilde{\mathbf{x}} = [x_t, x_{t+1}, ..., x_{t+h}]$ through a deep learning model F, conditioned to a set of input features $\mathcal{X}$ consisting of historical consumption data $[x_{t-h}, x_{t-h+1}, ..., x_{t-1}]$, where h is the historical context length, and other exogenous variables of weather, time, and demographic information.
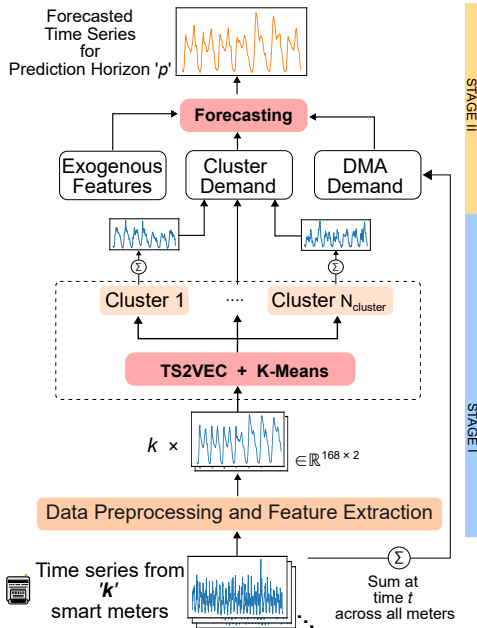
find similarities and dissimilarities across different meters. The second feature is also a load profile of $\mathbf{m}_i$, albeit for a shorter period of 12 weeks, to capture the behavior of $\mathbf{m}_i$ across different seasons of the year, and reduce the effect of stochastic temporal behavior on temporal contrasts. To capture the contrasting behavior of each consumer between weekdays and weekends, as well as across different consumer types, the required positive and negative pairs for contrastive learning includes temporal values representing consumption on both weekdays and weekends. Input features from the same consumer make up the positive pair, while negative pairs are from different users. The hierarchical loss function employed in TS2VEC also allows for a weighting factor α to emphasize the contrast between samples evenly or unevenly. The learned representations of end-consumers are clustered through K-Means clustering [18] to obtain $N_{cluster}$ consumer clusters with distinctive consumption characteristics and for a given cluster, the hourly demand of the cluster is the aggregation over all smart meters present in the cluster.
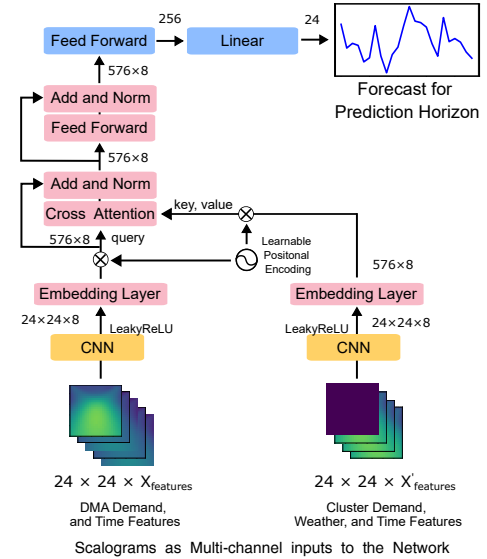

Fig. 1: The proposed framework for water demand forecasting for a heterogeneous consumer DMA.

Fig. 2: The architecture of the deep learning model for water demand forecasting.

A. Consumer Clustering based on Contrastive Learning

To separate consumers with varying consumption behaviors, we adopt a contrastive learning framework TS2VEC [6], that contrasts temporally within each input sequence of $\mathbf{m}_i$ and instance-wise across different input sequences of $\mathbf{m}_i$ and $\mathbf{m}_j$ where $i \neq j; 1 \leq i, j \leq k$ for a given batch during training. The individual smart meter consumption data $\mathbf{m}_i$ is processed to extract input samples $\mathbf{S} \in \mathbb{R}^{168 \times 2}$ with two features, that are representative of its global and local seasonal behavior. The first feature is a weekly load profile of $\mathbf{m}_i$ obtained through averaging hourly consumption data across all weeks over the entire time period of the data. This feature acts as an anchor for the meter's global characteristics and helps

B. Short-term Forecasting using Wavelet Scalograms

To improve water demand forecasting for heterogeneous DMA, we employ a Convolutional Neural Network (CNN) based architecture as a baseline [11] that has been previously used for this task. The model leverages scalograms to learn feature representations from utility data. The scalograms of s scales applied at h time points are obtained for N_{feat} different features and are concatenated along the third dimension, giving an image-like representation of dimensions $s \times h \times N_{feat}$, enabling 2D convolutions to learn the relationship between different features at a given time and scale s. The adopted architecture is adapted optimally via a cross-attention layer [19] to incorporate the scalograms of the obtained demand data of the clusters from the contrastive learning process. The

resulting model is 96.53% smaller than the adopted baseline. The attention block dynamically prioritizes weights to the time and frequency information present in the historical demand of the different clusters, and their contribution to the total demand of the DMA. The latent representation of total DMA consumption and time features act as the query, while the latent representation of each cluster's consumption along with the exogenous factors such as maximum temperature, and humidity over the prediction horizon p, act as keys and values for the cross attention layer. The two input branches along with the attention block are depicted in Figure 2. The attention layer is followed by feed-forward layers with a final linear layer producing the forecasts $\tilde{\mathbf{y}}$ for the prediction horizon.

III. EXPERIMENTAL SETUP

A. Data

The water demand data for this retrospective analysis is from Brønderslev, Denmark and is outlined in Table I. Four DMAs including rural and semi-urban areas with heterogeneous end-users, a rural area with homogeneous end-users, and an urban DMA with homogeneous end-users are chosen to evaluate the effectiveness of the framework. The urban DMA (predominantly residential) serves as a control to illustrate the effects of statistical smoothing during forecasting. Geolocation of each smart meter was utilized to evaluate of clusters formed in Stage I. The qualitative assessment across all DMAs featured sparse annotations by a utility company's technical expert. Consumption data until the end of 2019 was employed for all DMAs, reserving the final 26 weeks (half an year) for testing, which represented between 72% and 78% of the total data depending on the DMA. This segmentation allowed the framework to be tested across summer, winter, and a transitional season. Moreover, 10% of the training dataset was used as validation data. Actual recorded weather data from the respective regions are used, rather than forecasts.

TABLE I: Data on the selected DMAs for the study

ID	Number of meters	Data from	Description
A	159	$2017-09-01$	Rural
B	210	$2017-09-01$	Semi-Urban
C	67	$2018-01-01$	Rural
Control	1286	$2018-03-01$	Urban

B. Stage I: End-user Behaviour Clustering

Weekly hourly load profiles (from Monday to Sunday) are utilized to learn representations of individual end-consumers. Negative pairs consist of samples from different end-consumers, while positive pairs are derived by windowing the weekly profile of a consumer into two windows: w_a (Monday to Friday) and w_b (Friday to Sunday). The TS2VEC model is trained with a hidden dimension of 64 and a kernel size of 3 for the 1D convolution layers in the 10 residual blocks, producing an output dimension of 16. A custom sampler with a batch size of 64 limits each batch to only one sample from any end-consumer. The resulting representations are clustered

using K-Means, with a maximum of four clusters (residential, industrial, commercial, and agricultural). The optimal cluster count is determined using the highest silhouette score.

C. Stage II: Water Demand Forecasting

Forecasting features for water demand are identified through correlational analysis, revealing strong correlations with historical data and moderate correlations with temperature and humidity. Each feature represents an observed quantity of demand, weather, or time for 24 hours. The $X_{features}$ input features for the query branch comprise demand data with 24-hour and 168-hour lags, their seasonal components, and holiday information. Conversely, the $X'_{features}$ input features for the key/value branch are the extracted demand from each cluster with a 24-hour lag, temperature, humidity, and sine and cosine representations of the day of the week. Scalograms, with $s = 24$ and $h = 24$, are encoded using Shannon's entropy-to-energy ratio [20] to determine the optimal mother wavelet, identified as 'gaus4'.

The model architecture is depicted in Figure 2, detailing hidden layer output dimensions. LeakyReLU activation function is used in the convolutional and final feed-forward layers, while ReLU is employed in other feed-forward layers, with no pooling layers included. The final block comprises hidden dimensions of 1024, 512, and 256, concluding with a linear layer for forecasting with $p = 24$. The model employs a learning rate of 0.001, batch size of 256, ADAM optimizer, and Mean Squared Error (MSE) as the loss function, and is trained with early stopping on training and validation data set.

IV. RESULTS AND DISCUSSION

The evaluation of the water end-use behaviors involves analyzing cluster formations across the DMAs, focusing on their distinct periodicity, trend, and consumption magnitude. The results for DMA A and DMA B are depicted in Figure 3 and 4, respectively, illustrating two contrasting scenarios: DMA A represents a heterogeneous consumer profile within a rural area, while DMA B exemplifies a heterogeneous profile comprising urban and agricultural area. In all experiments, clustering through contrastive learning effectively separate varying end-user behaviors, thereby identifying the dominant demand patterns to enhance subsequent forecasting task.

In DMA A, a distinct long-term periodicity of approximately 45 days is identified primarily in two end-consumers from cluster **0**, whereas the remaining 157 consumers demonstrate only short-term periodicity. Further analysis, supported by the geolocation of smart meters and insights from the data provider's expert, reveals that this long-term periodicity within cluster **0** that has a significant difference to all other end-users is associated with a poultry farm, which typically has an average growing cycle of 6 to 8 weeks.

Conversely, in DMA B, the three distinct clusters underscore varied consumer behaviors within the DMA. Cluster **2**, comprising two end-consumer meters linked to a corporate office, displays a consumption pattern aligned with typical workweek hours, exhibiting near-zero consumption from Friday evening
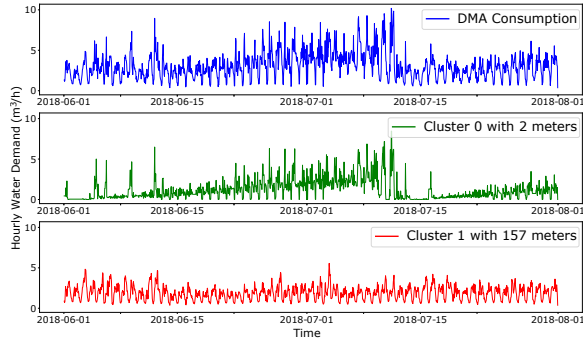
Fig. 3: Observed consumption of DMA A along with the total demand for the obtained clusters.

to Monday morning. Clusters **0** and **1** demonstrate daily periodicity, albeit with differing consumption magnitudes, and are spatially dispersed across the DMA. Cluster **1** predominantly includes small-scale commercial entities, whereas the remaining consumers in both clusters are primarily single-family residences with backyards. In contrast, DMA C lacks such dynamic consumer diversity, containing only two clusters with 14 and 53 meters respectively.
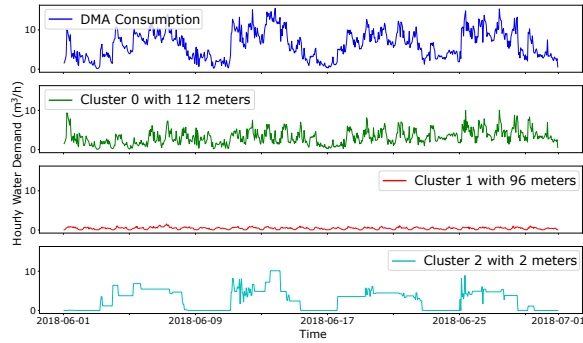

Fig. 4: Observed consumption of DMA B along with the demand for the obtained clusters.

TABLE II: Quantitative forecasting performance evaluation of the proposed approach, with baselines. The MAPE metric is %, while MAE quantifies error in l/h.

Model	Metric	DMA A	DMA B	DMA C	Control
ARIMA	MAPE	42.36	92.05	43.57	59.67
	MAE	50.38	107.9	67.53	86.74
LSTM	MAPE	43.80	82.41	31.49	12.98
	MAE	9.25	17.00	7.48	1.52
Wavelet CNN	MAPE	26.76	41.85	17.43	8.48
	MAE	6.34	**13.17**	4.66	0.97
Our Approach	MAPE	**21.86**	40.27	**16.95**	**8.40**
	MAE	**5.40**	13.52	**4.60**	**0.94**

The quantitative results for forecasting across all four DMAs are detailed in Table II. Across experiments, our methods outperform the baseline LSTM method greatly, while introducing further improvement of different degrees compared to the Wavelet CNN baseline although the number of parameters are reduced by 97%. Notably, DMA A, which exhibits two

distinct periodicities, in terms of behavior and time frame, across its clusters, achieved the highest performance gain of 4.9%. Conversely, DMA B and C, despite having clusters with varying behaviors, maintain weekly periodicities within each cluster, resulting in less pronounced differences compared to DMA A. The control DMA illustrates the impact of statistical smoothing due to a large number of users, where the relative gain in forecasting performance from even the baseline LSTM is minimal, starkly contrasting with the significant improvements observed in DMAs A, B, and C. The statistical baseline of ARIMA exhibits high errors in absolute and relative terms, as it fails to predict both small and large deviations.
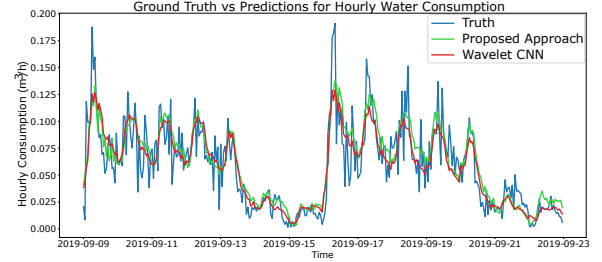

Fig. 5: Qualitative comparison between observed and predicted water demand for DMA B over two weeks.

Further analysis of DMA B through a qualitative two-week plot (Monday to Sunday) shown in Figure 5 elucidates the reasons behind the low forecasting performance noted in Table II. All models primarily capture the trend and seasonality, struggling to account for the high variance, particularly evident during weekends and the sharp increase in demand on Mondays. This variance is attributed to the corporate end-consumer in cluster **2** of DMA B, as depicted in Figure 4. This effect significantly influences the models' forecasting capabilities on weekdays. Nonetheless, the proposed approach successfully identifies the end-user responsible for these stochastic effects during working days, providing critical insights for proactive water demand management.

V. CONCLUSION

In this study, we propose and evaluate a novel framework for forecasting water demand in DMAs with heterogeneous end-consumers. Using contrastive learning, we identify user groups with varying behavioral patterns and enhance demand forecasts with features derived from these groups. We demonstrate this approach with two real-life examples of complex DMAs. We address the gap in modeling multi-faceted consumer demand beyond urban centers and elucidate the effect of statistical smoothing on large urban areas through a control group. Our method improves forecasting accuracy for DMA A by 4.9%, capturing both long- and short-term periodic patterns influenced by socio-economic factors. Additionally, clustering as a precursor for DMA B reveals end-consumers whose behavior obscures overall demand patterns, offering valuable insights for managing supply. As an extension, hybrid models combining deep learning and statistical methods could further improve forecasting for corporate consumers in DMA

B, particularly for their piece-wise linear behavior during weekends. The effect of the proposed approach could be quantified on other architectures for time series forecasting.

REFERENCES

[1] Sally Russell and Kelly Fielding. "Water demand management research: A psychological perspective". In: *Water Resources Research* 46.5 (2010). DOI: https://doi.org/10.1029/2009WR008408.

[2] Azadeh Hemati et al. "Deconstructing Demand: The Anthropogenic and Climatic Drivers of Urban Water Consumption". In: *Environmental Science & Technology* 50.23 (2016). PMID: 27802028, pp. 12557–12566. DOI: 10.1021/acs.est.6b02938.

[3] A. Cominola et al. "Benefits and challenges of using smart meters for advancing residential water demand modeling and management: A review". In: *Environmental Modelling & Software* 72 (2015), pp. 198–214. ISSN: 1364-8152. DOI: https://doi.org/10.1016/j.envsoft.2015.07.012.

[4] Yi Wang et al. "Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges". In: *IEEE Transactions on Smart Grid* 10.3 (2019), pp. 3125–3148. DOI: 10.1109/TSG.2018.2818167.

[5] Qianwen Meng et al. *Unsupervised Representation Learning for Time Series: A Review*. 2023. arXiv: 2308.01578.

[6] Zhihan Yue et al. *TS2Vec: Towards Universal Representation of Time Series*. 2022. arXiv: 2106.10466.

[7] Pooja Basak et al. "Unveiling Consumer Behavior in District Heating Network: A Contrastive Learning Approach to Clustering". In: *SESAAU2024 – Smart Energy Systems Conference* (Aalborg, Denmark). Sept. 10–11, 2024.

[8] Satyaki Chatterjee, Siming Bayer, and Andreas Maier. *Prediction of Household-level Heat-Consumption using PSO enhanced SVR Model*. 2021. arXiv: 2112.01908.

[9] Dazhao Du, Bing Su, and Zhewei Wei. *Preformer: Predictive Transformer with Multi-Scale Segment-wise Correlations for Long-Term Time Series Forecasting*. 2022. arXiv: 2202.11356.

[10] Panagiotis I. Karamaziotis et al. "An empirical investigation of water consumption forecasting methods". In: *International Journal of Forecasting* 36.2 (2020), pp. 588–606. ISSN: 0169-2070. DOI: https://doi.org/10.1016/j.ijforecast.2019.07.009.

[11] Adithya Ramachandran et al. *Heat Demand Forecasting with Multi-Resolutional Representation of Heterogeneous Temporal Ensemble*. 2023. arXiv: 2210.13108.

[12] Manel Rhif et al. "Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review". In: *Applied Sciences* 9.7 (2019). ISSN: 2076-3417. DOI: 10.3390/app9071345.

[13] Zhongfu Tan et al. "Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models". In: *Applied Energy* 87.11 (2010), pp. 3606–3610. ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2010.05.012.

[14] Diana Arsene et al. "Clustering Consumption Activities in a Water Monitoring System". In: *2022 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*. 2022, pp. 1–6. DOI: 10.1109/AQTR55203.2022.9801952.

[15] Nicolas Cheifetz et al. "Modeling and Clustering Water Demand Patterns from Real-World Smart Meter Data". In: *Drinking Water Engineering and Science* 2.10 (Jan. 2017), pp. 75–82. DOI: 10.5194/dwes-10-75-2017.

[16] Joaquim Leitão et al. "Detecting urban water consumption patterns: a time-series clustering approach". In: *Water Supply* 19.8 (Aug. 2019), pp. 2323–2329. ISSN: 1606-9749. DOI: 10.2166/ws.2019.113.

[17] Chrysi Laspidou et al. "Exploring Patterns in Water Consumption by Clustering". In: *Procedia Engineering* 119 (2015). Computing and Control for the Water Industry (CCWI2015) Sharing the best practice in water management, pp. 1439–1446. ISSN: 1877-7058. DOI: https://doi.org/10.1016/j.proeng.2015.08.1004.

[18] David Arthur and Sergei Vassilvitskii. "k-means++: the advantages of careful seeding". In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA '07. New Orleans, Louisiana: Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035. ISBN: 9780898716245.

[19] Adithya Ramachandran et al. "A Week Ahead Water Demand Forecasting using Convolutional Neural Network on Multi-Channel Wavelet Scalogram". In: *WDSA CCWI 2024 - 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry* (University of Ferrara, Ferrara, Italy). July 1–5, 2024.

[20] "Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions". In: *Neurocomputing* 133 (2014), pp. 437–445. ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2013.12.018.