Diffusion Models for Conditioned District Heating Network Time Series Generation

Frédéric Montet[1], Benjamin Pasquier[1], Michal Bryxí[1], Beat Wolf[1], Philippe Cudré-Mauroux[2], Jean Hennebert[1]

[1] *iCoSys*, HEIA-FR, *University of Applied Science and Arts Western Switzerland (HES-SO)*, Fribourg, Switzerland

[2] *University of Fribourg (UNIFR)*, Fribourg, Switzerland

E-mail: `frederic.montet@hefr.ch, benjamin.pasquier@hefr.ch, michal.bryxi@hefr.ch, beat.wolf@hefr.ch, philippe.cudre-mauroux@unifr.ch, jean.hennebert@hefr.ch`

Abstract. Modeling building behavior is important to enhance energy management and occupant comfort. Traditional simulation methods for district heating networks (DHNs) often struggle with scalability and computational efficiency. This paper explores diffusion models, a state-of-the-art deep learning approach, to generate conditioned multivariate time series data for DHNs, utilizing public building data and telemetry.

Preliminary results suggest that diffusion models have the potential to capture statistical properties of time series, especially for durations shorter than 24 hours. While conditional generation did follow the real data in terms of mean trend and variance, challenges remain in accurately reproducing peaks and extreme values. We discuss potential improvements in embedding methods, evaluation, and model architectures to enhance robustness. Our findings highlight the promising use of diffusion models for DHN simulations.

1. Introduction

Modeling building behavior is essential to enable a wide range of downstream tasks. For instance, it drives the development of energy-efficient buildings with improved energy management, minimizing environmental impact and reducing operational costs. Accurate building simulation and characterization can also improve indoor conditions by balancing temperature, lighting, and ventilation for the comfort of the occupants. Moreover, building behavior modeling can, for instance, support contingency planning for scenarios such as power outages, extreme weather adaptation, or regulatory compliance to ensure building resilience.

In building behaviour modeling, much attention has been focused on energy-related simulations. Tools such as EnergyPlus or TRNSYS exist to support modeling practitioners, but they fall short when requirements extend beyond energy-focused simulations [1, 2]. Given the inherent complexity and non-linear nature of such tasks, computations can become infeasible within a practical time frame, especially when scaled to multiple buildings or urban districts.

In computer science, the use of Diffusion Models to generate images has achieved considerable success. Based on simple text input, it is now possible to generate highly accurate images within seconds. What if building behaviour simulation, with various characteristics, could be achieved with similar ease, greatly enhancing accessibility for diverse users?

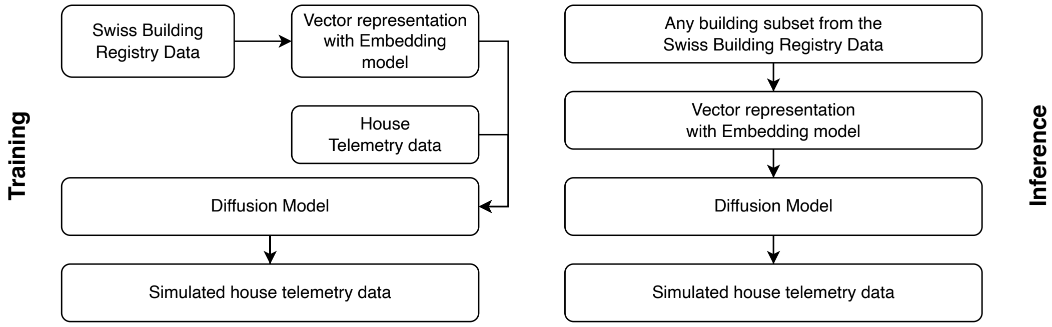


Figure 1: Flow diagram of the data for training (left-hand side) and inference (right-hand side)

Currently, the application of diffusion models to sensor data and time series remains in their early stages. While recent studies have proposed models showing promising results—particularly in tasks such as time series forecasting and imputation—relatively few have focused on time series generation [3, 4, 5, 6]. However, the ability to condition their generation process to meet various simulation conditions offers the flexibility needed for many downstream tasks where data is readily available but under-utilized. In this paper, we explore the potential of Diffusion Models, a state-of-the-art deep learning method, to efficiently model building behavior. We hypothesize that such architectures can offer scalable and accurate simulations, potentially surpassing traditional methods in flexibility and computational efficiency.

2. Method

The objective of the method introduced below is to be able to generate substation data such as water flow, supply and return temperature coming from a specific district heating network (DHN) given publicly available data about buildings. To reach this objective, the flow diagram presented in Figure 1 shows the different steps involved in the two main phases of model creation; on the left, when the model is trained and on the right when the model is used to infer data.

2.1. Buildings Under Study

We study around forty buildings connected to a DHN located in the French part of Switzerland in a village of 3000 inhabitants and at 600 meters above sea level. The dataset contains mostly residential houses of mixed types with a construction year ranging from before 1919 to 2015, with a great representation of the post-2000 era. Among those buildings, most are single or multi-family houses, with one small church. The areas of the buildings studied range from 80 to 1300 m^2 and only 13% of the energy reference area is informed. The data remains private in the context of our study, but any enquiry is welcomed so that results can be reproduced.

2.2. Using Publicly Available Data

The conditioning data comes from the Swiss Building Registry (RegBL) [7]. All samples fit a data model where most variables are optional and have heterogeneous types such as numerical and categorical values or free text [8]. In order to encode such samples in a fixed dimension, the samples are processed through a three-step encoding procedure: (1) online samples are put into a hashmap with descriptive keys and values in French, (2) those samples are fed into the OpenAI model `text-embedding-3-large` which transforms the hashmap into an embedding vector of 3072 elements. Finally, (3) six additional elements (two dates split into days, months, and year) are appended to provide timely information to the embeddings.

The use of a commercial embedding model to encode our data is a shortcut to simplify the encoding process. As this model has not been trained on RegBL samples, better embeddings might be obtained with other methods, such as a specific encoding pipeline per data type.

2.3. House Telemetry Data

The time series came from the substation part of the district heating network, i.e. the secondary, that acts as the heat exchanger in the building connected to the network. We selected four different variables which are the flow (m^3/h), power (kW), supply and return temperature (C). The collected data on the buildings from Section 2.1 starts on the 19$^{\text{th}}$ of March 2022 and ends on the 17$^{\text{th}}$ of April 2022 and has a 10-minute resolution, which represents 145'152 samples.

2.4. Diffusion Model

We use the diffusion model for time series from our prior studies *"Enabling diffusion model for conditioned time series generation"* [9]. The model can be trained to generate time series with two modes: *unconditionally*, meaning that the generated time series have some desired statistical properties, but without being specific, or *conditionally*, meaning that given a condition (i.e. the embedding vector from Section 2.2), the generated time series have specific statistical properties.

Formally, let $X_{\text{real}} = \{X_1, X_2, \ldots, X_n\}$ denote a dataset where each element $X_i = \{x_1, x_2, \ldots, x_T\}$ represents a time series of length T. The objective is to train a generative model that approximates the distribution of X_{real} in order to generate new synthetic data X_{synth}. In the *unconditional* scenario, the goal is to generate samples $X_{\text{synth}} \sim P_{\text{synth}}(X)$ such that $P_{\text{synth}}(X) \approx P_{\text{real}}(X)$, where $X_{\text{real}} \sim P_{\text{real}}(X)$. In the *conditional* scenario, a conditioning vector $C = \{c_1, c_2, \ldots, c_M\}$ of length M is introduced. The objective becomes to generate samples $X_{\text{synth}} \sim P_{\text{synth}}(X \mid C)$ such that $P_{\text{synth}}(X \mid C) \approx P_{\text{real}}(X \mid C)$, where $X_{\text{real}} \sim P_{\text{real}}(X \mid C)$.

A diffusion model generates data by simulating a two-step process: a forward process that gradually adds noise to real samples, and a reverse process that learns to remove this noise step by step, in order to reconstruct realistic data. This reverse process is learned through a neural network trained to denoise samples progressively. For time series generation, we use a recurrent neural network (RNN) with gated recurrent units (GRUs) to model the reverse process, as it effectively captures temporal dependencies inherent to sequential data. Our architecture is based on TSDiff, one of the few diffusion models specifically designed for time series modeling [6]. More specifically, the model combines two positional encoding layers, three feed-forward layers, and a GRU layer, totalling approximately 3 million trainable parameters.

To enable conditional generation, we adapt the original RNN by concatenating static conditioning vectors to each time step of the input, allowing the model to generate time series aligned with specific contextual information.

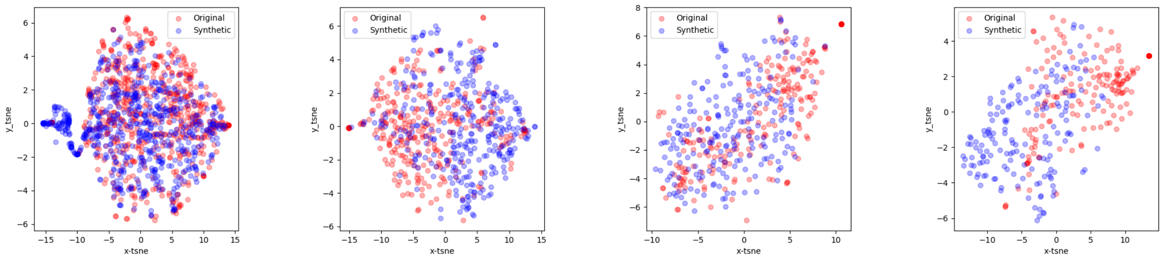
2.5. Unconditional Training

The objective of unconditional training is to assess the capacity of the model to capture the statistical properties of time series given different sequence lengths. For that purpose, the model was trained on 400 epochs with sequences of 36, 72, 108, and 144 time steps, representing durations of 6h, 12h, 18h and 24h.

Then, a two-dimensional t-SNE projection was computed to obtain a qualitative comparison of the overlap between real and synthetic samples. By observing overlap, separation, or clustering patterns, we can assess whether the generative model captures the structural and distributional properties of the real data.

2.6. Conditional Training

We also assessed the model's performance in conditioned scenarios. The training was done with 1500 epochs with a sequence length of 144, representing 24-hour samples. Then, real and



(a) 6 hours (b) 12 hours (c) 18 hours (d) 24 hours

Figure 2: t-SNE plots of the distribution of synthetic data (in red) compared to the real one (in blue). The four plots present 6h, 12h, 18h and 24h simulated samples in an unconditional scenario. The two axes don't have a direct interpretable meaning as in a traditional coordinate systems; distance between samples represent their relative similarity.

synthetic sequences were averaged per time step and plotted per month on a line plot, together with their variance. In this conditioned scenario, a spatial split was done on the building dimension: thirty buildings were used to train the model and six to test it. Testing buildings have between 1 to 3 apartments with a construction period spread between 1919 and 2015.

3. Results

The assessment of a generative model is rather abstract. Unlike prediction models where the predicted values can be compared with the real ones, in the case of generation, there are no true samples to compare with. Therefore, in addition to visualization with Principal Component Analysis (PCA) or t-distributed stochastic neighbor embedding (t-SNE), metrics to compare distributions can be used such as Mean Distribution Difference (MDD), Auto-correlation Difference (ACD) or Mean Absolute Error (MAE). In the following section, the results of the two generative scenarios are presented, first for the unconditional generation, and then for the conditional generation.

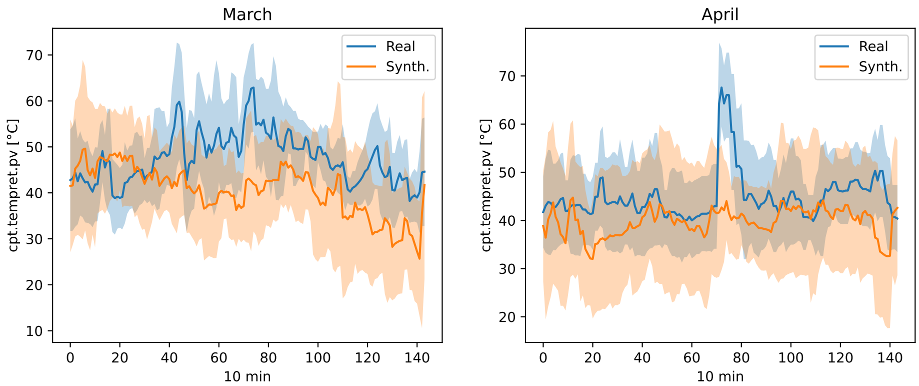
3.1. Unconditional Generation

Four models were trained for unconditional generation, each requiring approximately ten minutes. Evaluation is performed visually and is presented in Figure 2, which shows multiple t-SNE plots. From the figure, the first takeaway is the presence of a distribution overlap across all generation durations. The second takeaway is the reduction of distribution overlap as a function of the increasing length of generation. For generations until 18h, the overlap between synthetic and real data seems significant, but becomes more difficult to characterize for the 24h generation. The third takeaway is the reduction of samples as a function of the generation length. This is unsurprising since the total number of time steps remains the same across generations, but the size of the samples increases; the 24h generation has four times fewer samples than the 6h one.

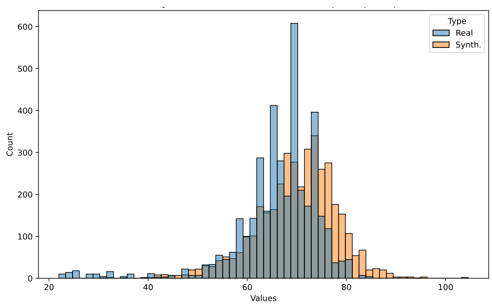
Those t-SNE projections provide qualitative insight but, without quantitative metrics (e.g., Maximum Mean Discrepancy or Wasserstein), it only confirms global alignment rather than detailed generative fidelity [10, 11]. Relying on a single manifold-learning method induces risks; therefore, additional statistical analyses would be needed to further validate generation quality.

3.2. Conditioned Generation

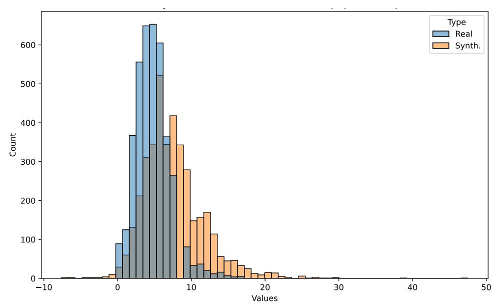
Figure 3 presents results of the conditioned generation isolated for three selected buildings, which allows us to observe the impact of the conditioning on the generated time series. In general, since the model could generate samples approximating the statistical properties of the real data



(a) Time series of the variable *'cpt.tempret.pv'*, the return temperature from the secondary for March and April. The lines are the averaged values per time step for all real and generated time series. The background in reduced opacity present the variance per time step.



(b) Distributions of the *cpt.tempaller.pv* feature, the supply temperature to the secondary.



(c) Distributions of the *cpt.puissance.pv* feature, the power provided to the secondary.

Figure 3: Visual results of interest from the conditional generation of time series. The different plots represent three different cases encountered on three different buildings. Real data are in blue and synthetic in orange.

with partial efficacy, the impact of the conditioning on the synthetic time series can be validated; however, it has limitations. Three different residential semi-detached houses with two floors are represented in the sub-figures from Figure 3. Figure 3a and 3b represent similar houses built in 2012 as part of a planned residential development. Figure 3c represents a bigger semi-detached house built before 1919.

In Figures 3a and 3b, the generated return-temperature series (dark line) closely track the real data in terms of mean trend and variance; nonetheless, thermal peaks observed in April are not faithfully reproduced, indicating that the model underestimates short-term spikes. Similarly on Figure 3b, the distribution of supply temperatures exhibits multiple modes that are absent in the synthetic output, suggesting a degree of mode collapse. For the pre-1919 house in Figure 3c, the generator under-represents the full variance, thereby failing to capture building-specific thermal dynamics.

These issues highlight the need for quantitative validation, such as RMSE targeting peak deviations or Kolmogorov-Smirnov tests on marginal distributions. Also, more detailed conditioning inputs (e.g., envelope characteristics, weather data, or dynamic occupancy indicators) could ensure that both extreme values and distinctive building behaviour are accurately modelled.

4. Discussion

The results presented in this study confirm the potential of diffusion models to flexibly generate building-related time series, complementing rather than replacing standard simulation tools.

In the *unconditional* generation scenario, the model demonstrated an ability to approximate the overall statistical properties of multivariate time series, especially for shorter horizons. Nevertheless, performance degrades as sequence length increases: as shown in Figure 2, the overlap between real and synthetic distributions diminishes at 24h, and sample sparsity further hinders reliable t-SNE embedding. This suggests that generating longer continuous sequences may require segmenting longer horizons into multiple shorter windows or employing alternative strategies to preserve distribution fidelity.

For *conditional* generation, the diffusion model successfully reproduces average trends and variances for multiple buildings, indicating that the OpenAI `text-embedding-3-large` embeddings carry enough contextual information to guide generation. However, limitations emerge in reproducing peaks and extreme values. In Figures 3a and 3b, the model fails to emulate sharp thermal spikes and multimodal supply-temperature distributions, and for the pre-1919 house (Figure 3c), the synthetic data under-represents the empirical variance. These gaps likely come from two sources: first, the low information in the current conditioning vectors (i.e., RegBL embeddings without explicit envelope or occupancy details); second, the model's difficulty in capturing rarer, high-magnitude events when trained on limited sample diversity.

Taken together, these observations underscore the importance of rigorous quantitative validation. While t-SNE (or UMAP) projections provide qualitative insight, distributional-distance metrics, such as Maximum Mean Discrepancy, and hypothesis tests on marginal distributions, such as Kolmogorov–Smirnov, are necessary to identify specific dimensions where synthetic data diverges from reality. Moreover, addressing sample imbalance, particularly the reduction of sample count for longer horizons, is critical to ensure that visual embeddings do not obscure true model behavior. To address these limitations, future work should pursue three main directions. First, *improving fidelity for peaks and extreme values* may require integrating specialized loss terms or data-augmentation techniques that emphasize tail behavior, as well as experimenting with alternative generative architectures. Second, *enhancing embedding appropriateness* can be achieved by incorporating richer conditioning inputs, such as explicit building envelope parameters, dynamic occupancy schedules, and localized weather profiles, or by exploring alternative embedding models tailored to heterogeneous building registry data. Third, *field validation and district-scale experiments* would enable assessment of model generalizability across an entire heating network: for instance, deploying a case study on a defined DHN would reveal how well the diffusion model captures inter-building interactions and network-level thermal dynamics. By systematically addressing these issues, diffusion-based conditioned generation can evolve into a robust tool for DHN simulation, supporting applications ranging from scenario planning to anomaly detection in energy management.

5. Conclusion

In this study, we demonstrated that diffusion models can generate realistic short-duration time series for district heating networks, complementing traditional simulation tools. However, fidelity degrades at longer horizons and conditional outputs seem to miss transient peaks and the complete range of observed values, which is linked to sparse conditioning information and difficulty modeling rare events. Future work should employ quantitative metrics such as MMD or Kolmogorov–Smirnov, incorporate richer embeddings (including detailed building and occupancy data), and explore architectures or loss functions that emphasize tail behavior.

References

[1] 2025 EnergyPlus The Board of Trustees of the University of Illinois, The Regents of the University of California, through Lawrence Berkeley National Laboratory (subject to receipt of any required approvals from the U.S. Dept. of Energy), Oak Ridge National Laboratory, managed by UT-Battelle, Alliance for Sustainable Energy, LLC, and other contributors.

[2] University of Wisconsin–Madison Solar Energy Laboratory 1975 *TRNSYS, a Transient Simulation Program* (Madison, Wis. : The Laboratory, 1975.)

[3] Rasul K, Seward C, Schuster I and Vollgraf R 2021 Autoregressive Denoising Diffusion Models for Multivariate Probabilistic Time Series Forecasting *Proceedings of the 38th International Conference on Machine Learning* (PMLR) pp 8857–8868 iSSN: 2640-3498 URL `https://proceedings.mlr.press/v139/rasul21a.html`

[4] Alcaraz J M L and Strodthoff N 2023 Diffusion-based Time Series Imputation and Forecasting with Structured State Space Models arXiv:2208.09399 [cs, stat] URL `http://arxiv.org/abs/2208.09399`

[5] Yan T, Zhang H, Zhou T, Zhan Y and Xia Y 2021 Scoregrad: Multivariate probabilistic time series forecasting with continuous energy-based generative models (*Preprint* `2106.10121`)

[6] Biloš M, Rasul K, Schneider A, Nevmyvaka Y and Günnemann S 2023 Modeling Temporal Data as Continuous Functions with Stochastic Process Diffusion arXiv:2211.02590 [cs] URL `http://arxiv.org/abs/2211.02590`

[7] fédéral de la statistique O Registre fédéral des bâtiments et des logements (RegBL)

[8] Office fédéral de la statistique 2025 Modèle de données et description des entités

[9] Montet F, Pasquier B, Wolf B and Hennebert J 2024 *Engineering Proceedings* **68** ISSN 2673-4591 URL `https://www.mdpi.com/2673-4591/68/1/25`

[10] Gretton A, Borgwardt K M, Rasch M J, Schölkopf B and Smola A 2012 *Journal of Machine Learning Research* **13** 723–773 URL `http://jmlr.org/papers/v13/gretton12a.html`

[11] Peyré G and Cuturi M 2020 Computational Optimal Transport (*Preprint* `1803.00567`)