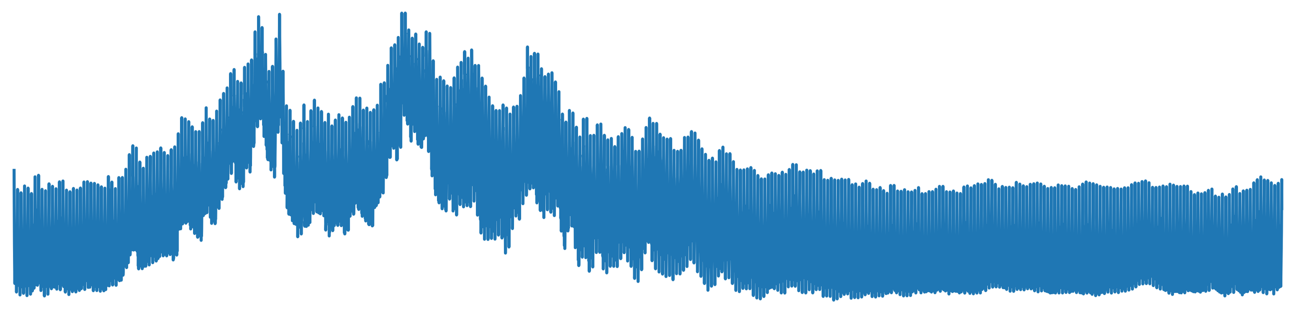
A synthetic dataset of French electric load curves with temperature conditioning

Tahar Nabil [1] Ghislain Agoua [1] Pierre Cauchois [2] Anne De Moliner [2] Benoît Grossin [1]

[1]EDF R&D, Palaiseau, France [2]Enedis, France

Context

- Individual metering data, incl. electric load curves, are crucial for smarter and flexible grids, e.g. with self-consumption of local renewables
- Secondary use of such data, e.g. sharing for innovation, is strictly limited by European privacy regulation
- How can we unlock the full potential of fine-grained electric load curves without compromising privacy?
- Recent advances in deep generative models → synthetic load curves?



Contribution

1. We release a new dataset of synthetic load curves:
   - Individual residential electric consumption
   - Fine-grained: one year at half-hourly resolution
   - With labels: contracted power and time-of-use (ToU) plan
   - Conditionally on local outdoor temperature
   - Representative of recent (post-2022) consumer behaviors
2. We show that our gen. model yields highly realistic, useful & private samples

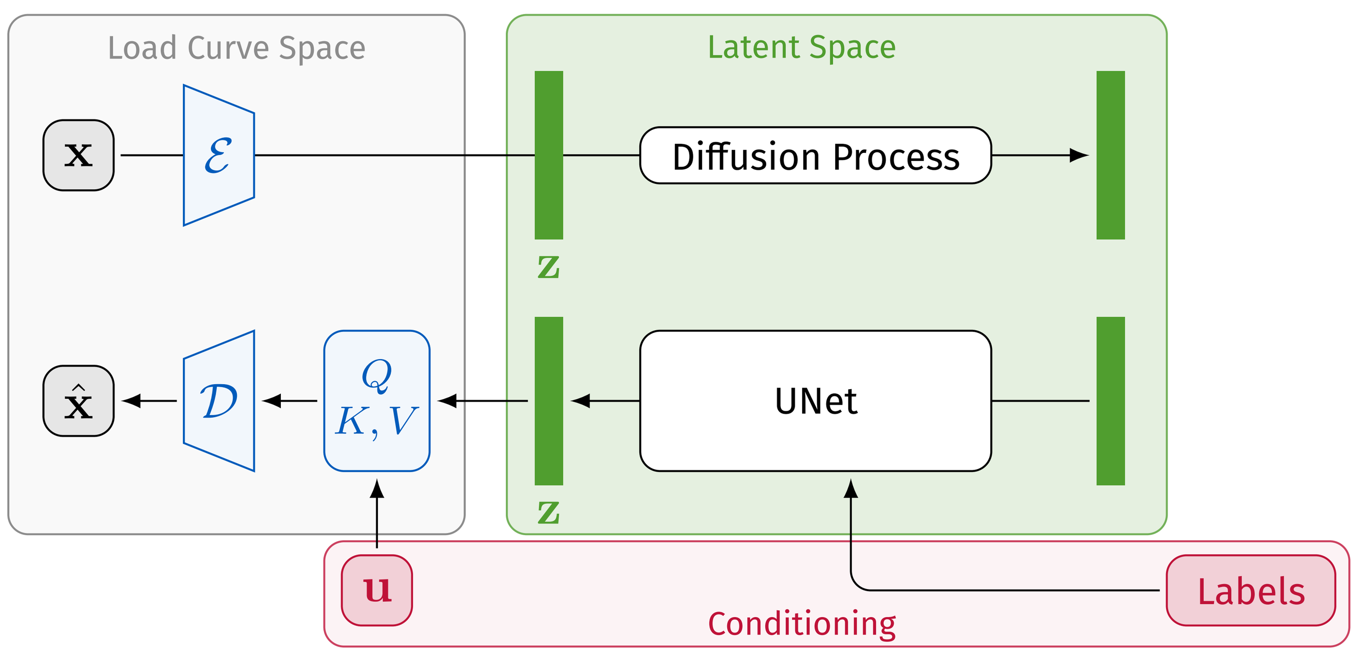
Network architecture: conditional Latent Diffusion (LDM, [2])

(1) The Autoencoder $\mathcal{D} \circ \mathcal{E}$ learns to compress then reconstruct load curves

- Load curves x are reshaped as single-channel 2D images in $\mathbb{R}^{1 \times N_{days} \times 48}$
- $\mathcal{D} \circ \mathcal{E}$: 2D convolutions + Vector Quantization for latent space regularization [3]
- We condition on temperature by adding a cross-attention layer: between (i) the load curve represented by its latent code $\mathrm{z} := \mathcal{E}(\mathrm{x})$ as queries and (ii) the temperature u in patches as keys and values

(2) The Diffusion model fits the distribution of codes in the latent space of $\mathcal{D} \circ \mathcal{E}$

- Based on a Denoising Diffusion Probabilistic Model (DDPM, [1])
- We condition by static variables (contracted power and ToU) by concatenation to the input of the UNet denoiser



(3) At inference, latent codes are sampled from the reverse diffusion process, they undergo cross-attention with temperature and are decoded to new load curves by $\mathcal{D}$.

[1] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, 2020.

[2] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *CVPR*, 2022.

[3] Aaron van den Oord, Oriol Vinyals, and Koray Kavukcuoglu. Neural Discrete Representation Learning. In *Advances in Neural Information Processing Systems*, 2017.

Training details

- Training set: 17k samples, starting from Oct. 2022, spanning 94 departments in Metropolitan France & restricted to thermo-sensitive customers
- Held-out test set for evaluation: 2k samples
- Data from Enedis, the main distribution grid operator for electricity in France

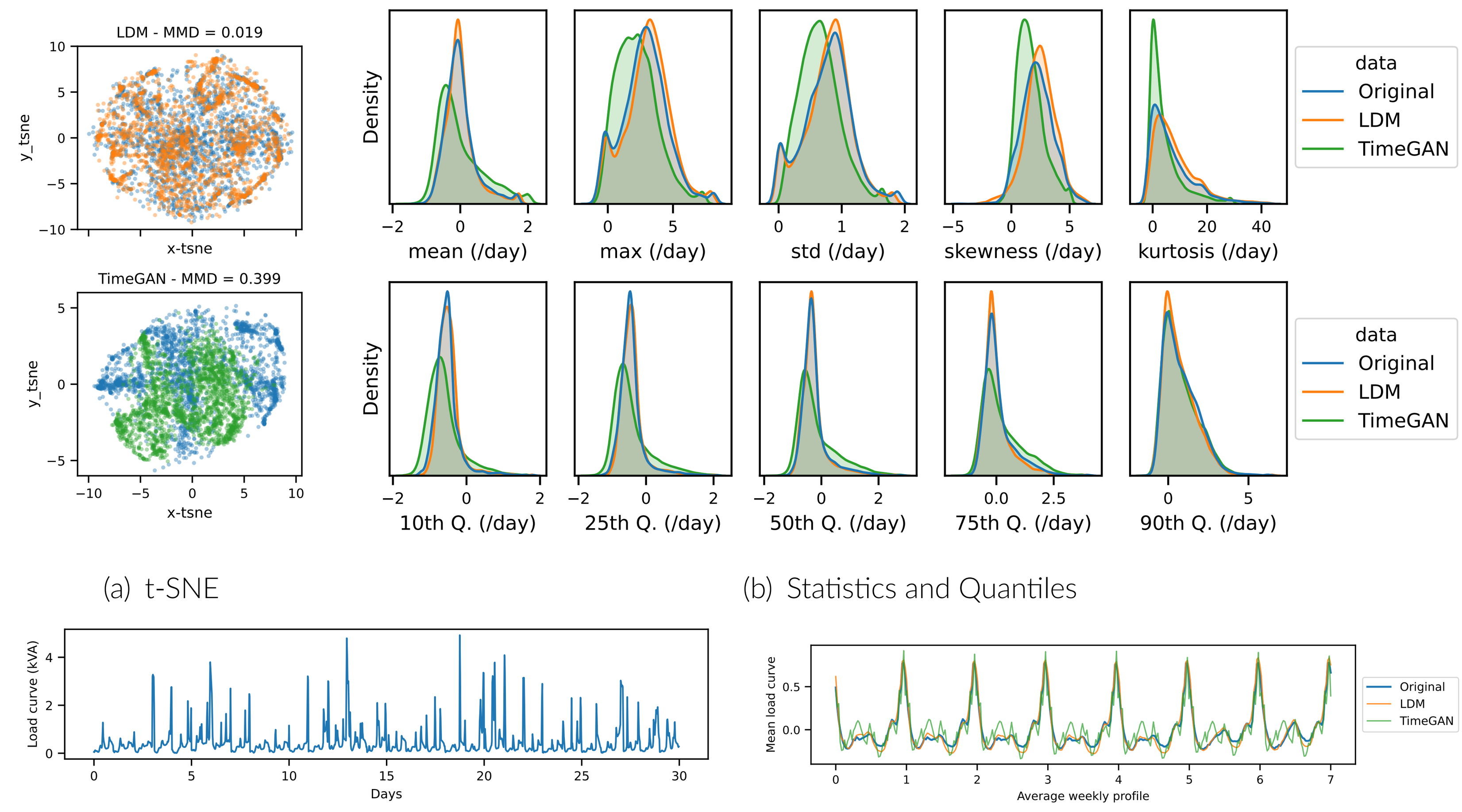
Evaluation: Fidelity & Diversity

Quantitatively, our model yields diverse high-fidelity samples, outperforming our optimized baseline (TimeGAN). This is reflected in the standard metrics for generative modelling: the discriminative, Context-FID and correlation scores.

Table 1. Fidelity scores on the hold-out test set for samples on *night* ToU and 6 kVA. Discriminative score computed on one-year load curves (D_{year}) or on averaged daily profiles ($D_{profile}$).

	D_{year} (↓)	$D_{profile}$ (↓)	Context-FID (↓)	Correlation score (↓)
LDM	**0.037**	**0.059**	**1.748**	**0.002**
TimeGAN	0.357	0.452	2.082	0.224

Multiple qualitative plots demonstrate that LDM samples are realistic consistently across diverse views of the time series: individual load curves, aggregated weekly profiles, daily statistics and quantiles.



(a) t-SNE (b) Statistics and Quantiles

(c) Individual sample (30-day zoom) (d) Weekly profiles

Figure 1. (a) t-SNE 2D projection of original (blue) and synthetic data (LDM, orange vs. TimeGAN, green). (b) to (d) are restricted to the night ToU, 6kVA category: (b) Density estimation of daily statistics and quantiles; (c) Example of a synthetic sample by LDM; (d) Mean weekly profiles.

Evaluation: Utility

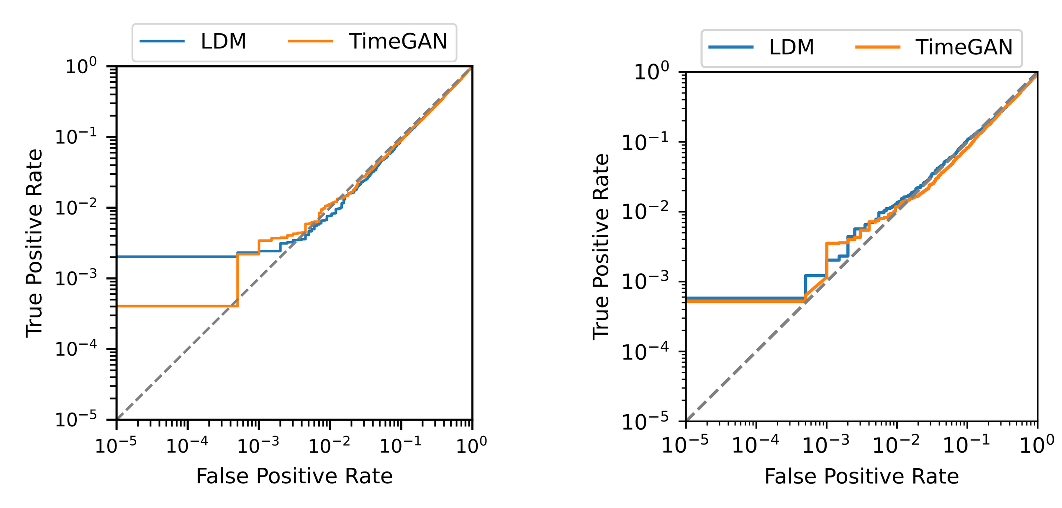
Training on LDM or real samples is equivalent for solving challenging time series tasks.

Table 2. Metrics for training on LDM or TimeGAN vs on real data (TRTR) and testing on real data. Bold: closest to TRTR. Forecasting results are averaged across horizons $[48, 96, 192, 336]$, for a lookback of length 720. Baselines: copy from last week (forecasting), majority class (classification).

Loss	TRTR	LDM	TimeGAN	Baseline
Forecast. MSE	0.190	**0.190**	0.209	0.306
Forecast. MAE	0.234	**0.233**	0.253	0.251
Classif. Acc	0.740	**0.750**	0.700	0.607
Classif. F1	0.576	**0.564**	0.537	0.252

Evaluation: Privacy

White-box or black-box membership inference attacks (MIA) fail, suggesting that the model did not overfit the training set. Similarly, three-sample MMD statistical tests give no evidence that the synthetic dataset is closer to train than to test set.



(a) Black-box (b) White-box

Figure 2. MIA ROC curves in log scale.

Data Availability

A public record of 10k synthetic one-year residential load curves is released, with their local temperatures, contracted power and ToU plans.

Data Access →

Contact: tahar.nabil@edf.fr