A synthetic dataset of French electric load curves with temperature conditioning

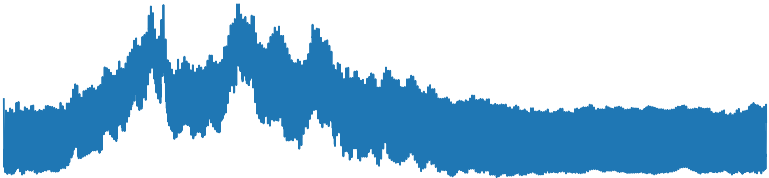T. Nabil[1]   G. Agoua[1]   P. Cauchois[2]   A. de Moliner[2]   B. Grossin[1]

April 28, 2025

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

Synthetic load curve data for power grid researchers

- **Smart Meter Data** is key to a successful **energy transition**

- Useful for: **flexibility**, **self-consumption of local renewable**, etc.

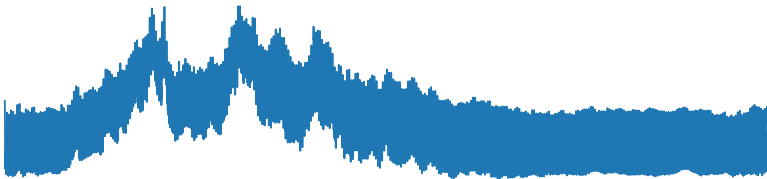- ⚡ But individual electric consumption is private → sharing ✗

Synthetic load curve data for power grid researchers

- **Smart Meter Data** is key to a successful **energy transition**
- Useful for: **flexibility**, **self-consumption of local renewable**, etc.
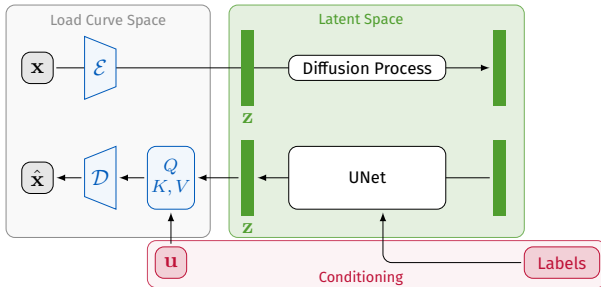- ⚡ But individual electric consumption is private → sharing ✗

Widespread use of smart meter data can be unlocked by synthetic generation
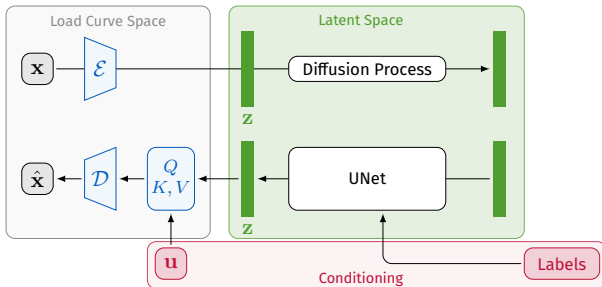
WE RELEASE A NEW **synthetic** LOAD CURVE DATASET

(i) Individual residential electric load curves

(ii) Fine-grained: one year at 1/2-hourly resolution

(iii) With labels: contracted power + time-of-use plan

(iv) Conditionally on local outdoor temperature

(v) Representative of recent (post-2022) consumer behaviors

Overall architecture: Latent Diffusion [Rombach et al., 2022]

- Load curves as 2D images instead of long 1D vectors
- Autoencoder $\mathcal{D} \circ \mathcal{E}$: 2D CNNs + Vector Quantization + compression ratio
- Diffusion: DDPM, UNet with Spatial Attention
- Conditioning by labels: concatenate to UNet

- Load curves as 2D images instead of long 1D vectors
- Autoencoder $\mathcal{D} \circ \mathcal{E}$: 2D CNNs + Vector Quantization + compression ratio
- Diffusion: DDPM, UNet with Spatial Attention
- Conditioning by labels: concatenate to UNet
- Conditioning by temperature u: x-attention in latent space z (Q) vs. u (K, V)

Training details

- Training set of 17k one-year samples starting from October 2022
    - Provided by Enedis, French energy utility operating the electricity distribution network
    - Spanning 94 departments in Metropolitan France
    - Restricted to thermo-sensitive customers

- Labels: contracted power (3 classes) $\times$ ToU rate (3 classes)

- Dynamic conditioning: local outdoor temperature

- 2k samples $\rightarrow$ held-out test set

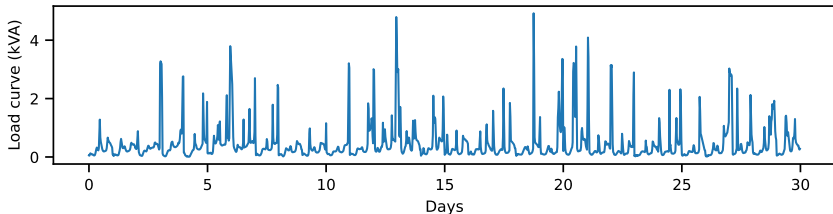LDM closely matches the distribution of real data
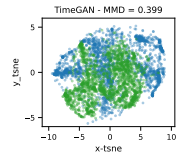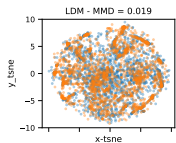
t-SNE
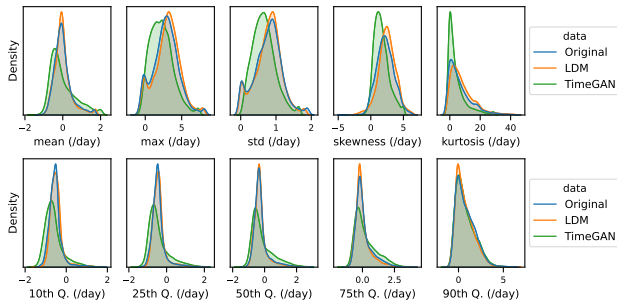
Statistics & Quantiles

Baseline: TimeGAN with optimized hyperparameters [Yoon et al., 2019]

LDM closely matches the distribution of real data

Individual samples (30-day zoom)

Table 1: Fidelity scores on the hold-out test set (*night* ToU and 6 kVA).

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

LDM samples can seamlessly replace real samples on ML tasks

Table 2: TSTR metrics. Forecasting results are averaged across horizons [48, 96, 192, 336], for a lookback of length 720. Baselines: copy from last week (forecasting).
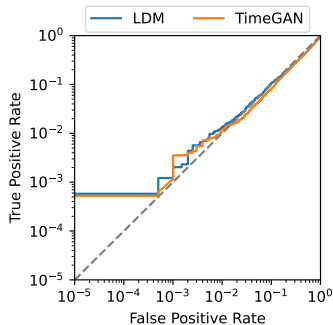
	Loss	TRTR	LDM	TimeGAN	Baseline
Forecast. MSE		0.190	**0.190**	0.209	0.306
Forecast. MAE		0.234	**0.233**	0.253	0.251

Our experiments suggest synthetic data are original, not copies
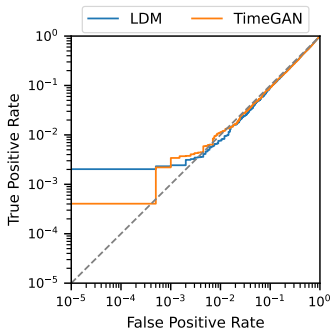
Black-box MIA

White-box MIA

Summary

- Unprecedented high-quality samples across all three dimensions
  - → Details and comprehensive evaluation: check out our paper!

- Unprecedented high-quality samples across all three dimensions
    - $\rightarrow$ Details and comprehensive evaluation: check out our paper!

- Data availability: **we publicly release a 10k sample dataset**
    - $\rightarrow$ Incl. load curves, labels & local temperatures
    - $\rightarrow$ Public url expected online by end 2025
    - $\rightarrow$ In the meantime, **contact: tahar.nabil@edf.fr for access**

THANKS

References

R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022.

J. Yoon, D. Jarrett, and M. Van der Schaar. Time-series generative adversarial networks. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/hash/c9efe5f26cd17ba6216bbe2a7d26d490-Abstract.html.