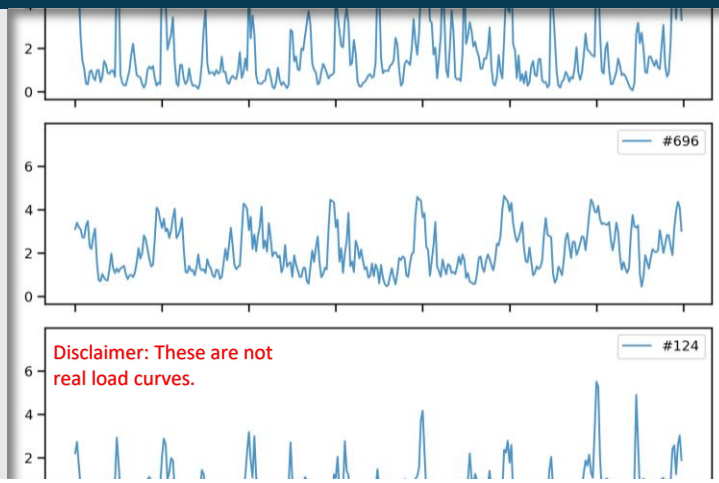
SUCCESS CASE 4.2024

DeepCourboGen

ULTRA-REALISTIC PRIVACY-PRESERVING SYNTHETIC SMART METERS DATA PRODUCED BY GENERATIVE AI



DEEPCPURBOGEN IN A NUTSHELL

DeepCourboGen is a generative artificial intelligence (GenAI) system designed to generate **realistic synthetic data for individual smart meters**.

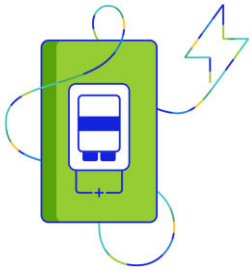
This **Privacy-Enhancing Technology**, developed by Enedis with the support of EDF's R&D team within the framework of a regulated research agreement, **enables broader applications of smart meter data without raising privacy concerns.** This will allow academics and market players to design new innovative services based on this data.

THE CHALLENGE

According to the European General Data Protection Regulation (GDPR), **residential electricity consumption data** are **personal data:** an individual intraday load curve or even a sequence of daily energy meter readings is an indirect identifier of a household. Concurrently, the challenges of the electrical grid and the ecological transition underway in France require mobilising all the resources and assets of the DSO, including data from smart meters. In response to this, Enedis has been engaged since 2017 in an R&D project to achieve **responsible (re)use of smart metering data**, ensuring strict guarantees of customer **privacy protection.**

THE SOLUTION

DeepCourboGen is an ultra-realistic synthetic smart metering data generator based on **GenAI**. This tool is a **Privacy-Enhancing Technology (PET)**. According to the OECD report '*Emerging privacy-enhancing technologies: Current regulatory and policy approaches*' published in March 2023, PETs are 'digital solutions that allow information to be collected, processed, analysed, and shared while protecting data confidentiality and



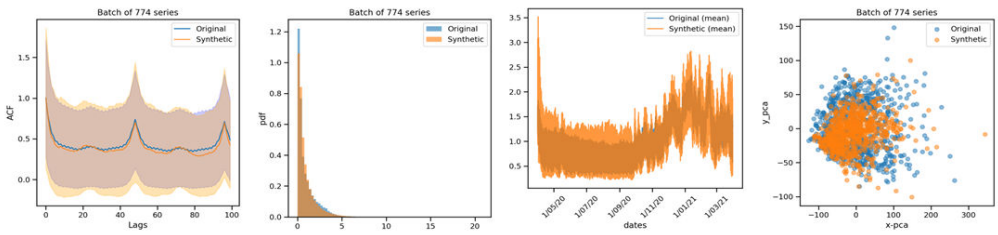
privacy'. To develop DeepCourboGen, Enedis early invested in **Deep Generative Models**. Previously, traditional approaches such as Markov models or Multi-Agent systems were proved inadequate for generating realistic individual load curves. Only **Generative Adversarial Networks** (**GANs)** allowed Enedis to achieve **high-fidelity objectives**, making it impossible to distinguish real data and synthetic data generated by DeepCourboGen.

DeepCourboGen is currently based on a TimeGAN model and focuses on residential sites.

MAIN ACHIEVEMENTS

To assess the quality of the synthetic data generated, DeepCourboGen has been evaluated based on both fidelity criteria and privacy criteria:

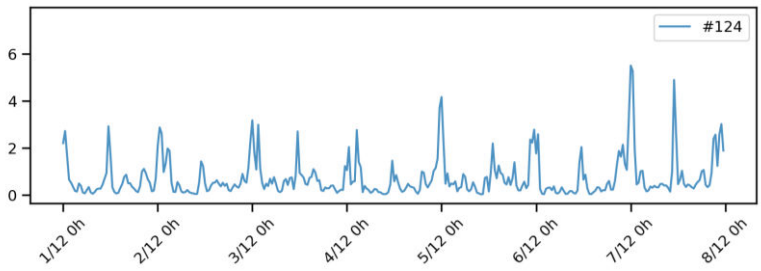
- **Fidelity**: Several standard metrics for distribution comparison were used, as well as metrics specific to individual load curve analysis. The evaluation shows good performance of the generator and the synthetic load curves show realistic characteristics without being a replica of real training data.
- **Privacy**: Both the generative model and the synthetic data were analysed for traces, even partial traces, of real data, and none were found. Notably, membership inference attacks failed, confirming the protection of real data used during DeepCourboGen training phase.



Examples of fidelity metrics, highlighting the proximity of statistics on synthetic data (in orange) with those of real data (in blue).

KEY SUCCESS FACTORS

- **Training Data**: Collecting a dataset of real load curves data from several thousand households with consent for the processing of their personal data.
- **Generative Model:** Efficiently training DeepCourboGen on the data and incorporating various generation parameters such as contracted power, temperature scenario, etc.



Example of a weekly load curve with a 30-minute time step. DeepCourboGen can generate intraday load curves for up to 1 year.

WAY FORWARD

After an initial exchange with the French Personal Data Protection Authority, Enedis is currently exploring opportunities to share synthetic datasets from DeepCourboGen, particularly aimed at the academic domain or more generally at all parties working with energy data. This open innovation initiative would allow researchers and students to work on realistic load curves and perhaps invent solutions to energy challenges.

To improve DeepCourboGen, further work includes trials with alternatives to the GAN model, notably the diffusion model, and extension of the generation capabilities to Small and Medium-sized Enterprises (SMEs).