TENDA

Smart Energy Cluster
challenge 1

TIRONI
CRISTIAN

VINCIGUERRA
MARCO

VEDOVATI
MATTEO

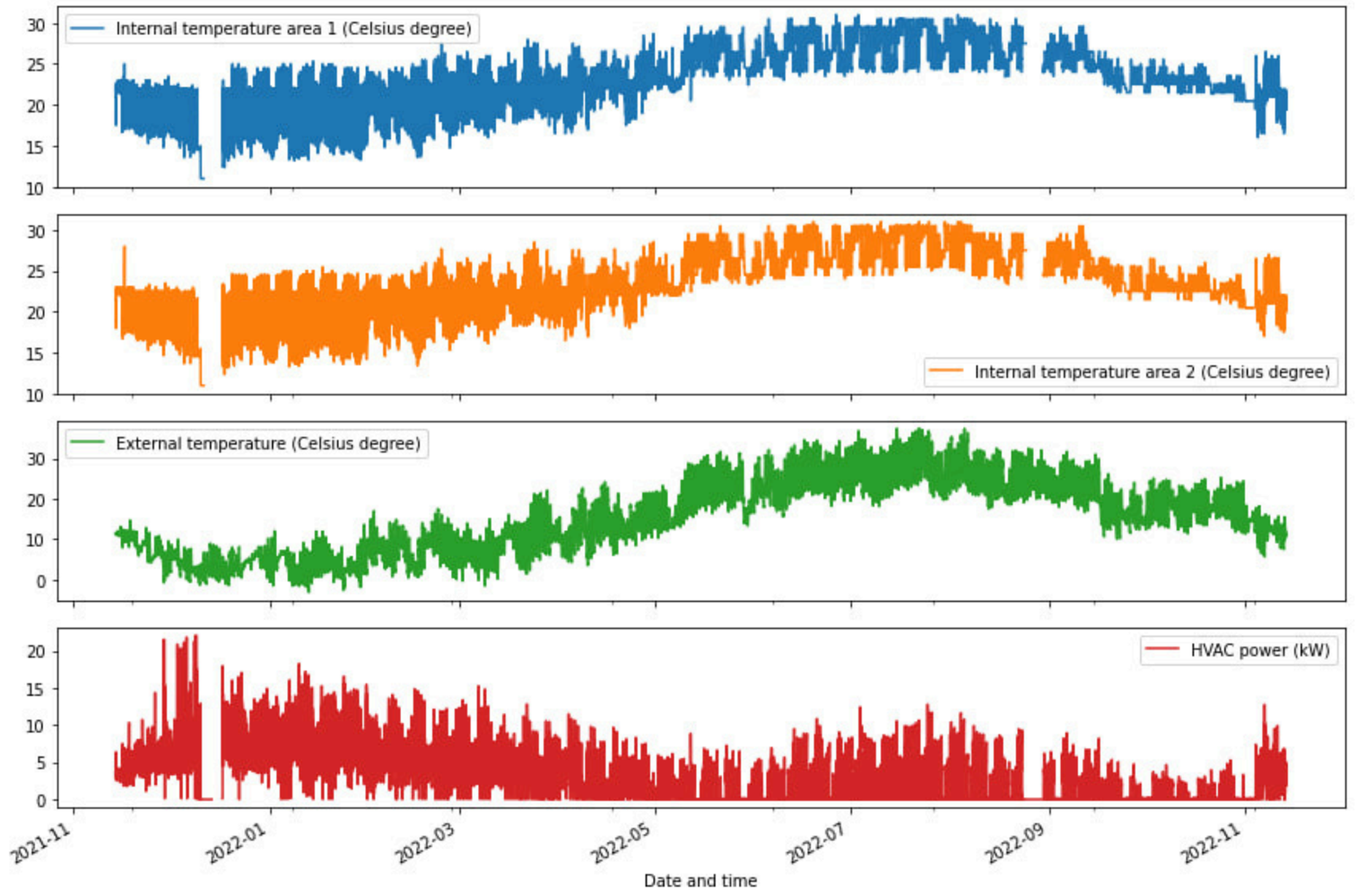
MARCHESI
GABRIELE

Datasets Info

Feature	Min Value	Max Value	Mean	Std
Internal temperature area 1 (Celsius degree)	15.500	30.522	23.732	2.930
Internal temperature area 2 (Celsius degree)	17.783	30.000	24.213	2.684
External temperature (Celsius degree)	-2.530	37.350	16.285	8.68
HVAC power (kW)	0	10.475	2.346	2.632

- 6 datasets, each one representing a building with a different size
- Measured with intervals of 15 minutes over 1 year

Dataset Example



General Preprocessing

- **Excluded Dataset n° 6 for lack of data**
- **Filtering from 14/11/2021 to 13/11/2022 (1 Year)**

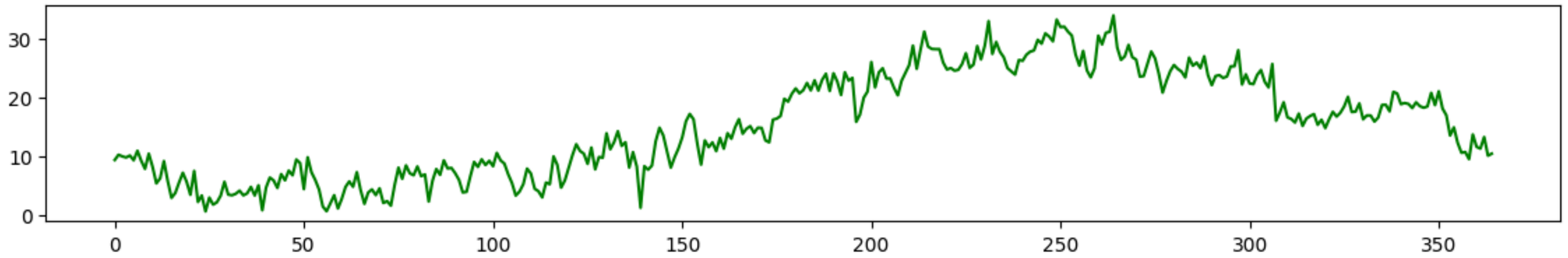
Task Division

Generating new syntetic external temperature time series

Given an external temperature time series and metadata generate the Internal temperatures and HVAC powers

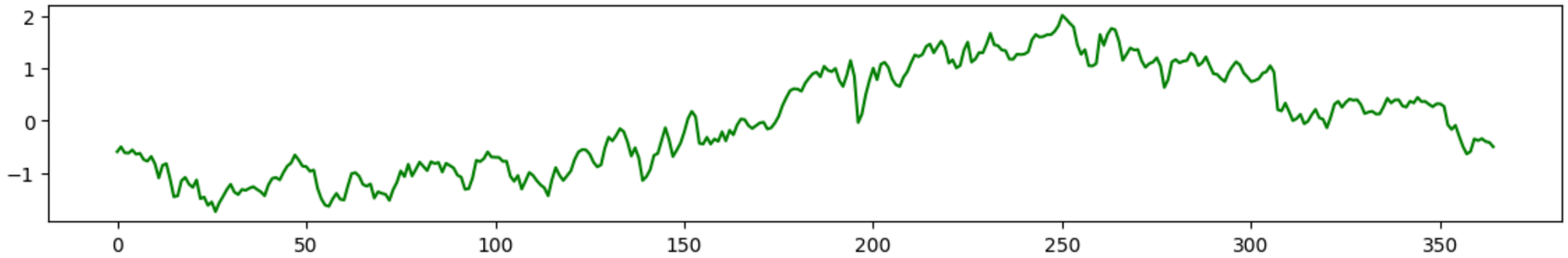
External temperature time series generation

1. **Normalizing the data**
2. **Calculating the year seasonality of the datasets timeseries using daily averages**
3. **Given a daily average generate the time series evolution over the specific day**
4. **Combine every day back together for the final synthetic external temperature.**

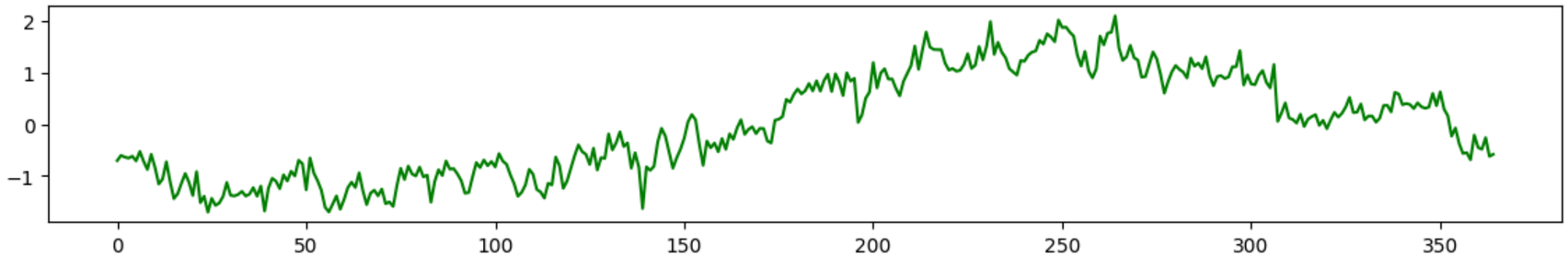


Year seasonality

The seasonalty component is calculated using seasonal decomposition

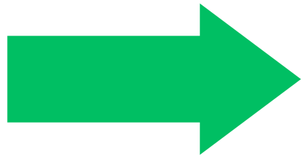


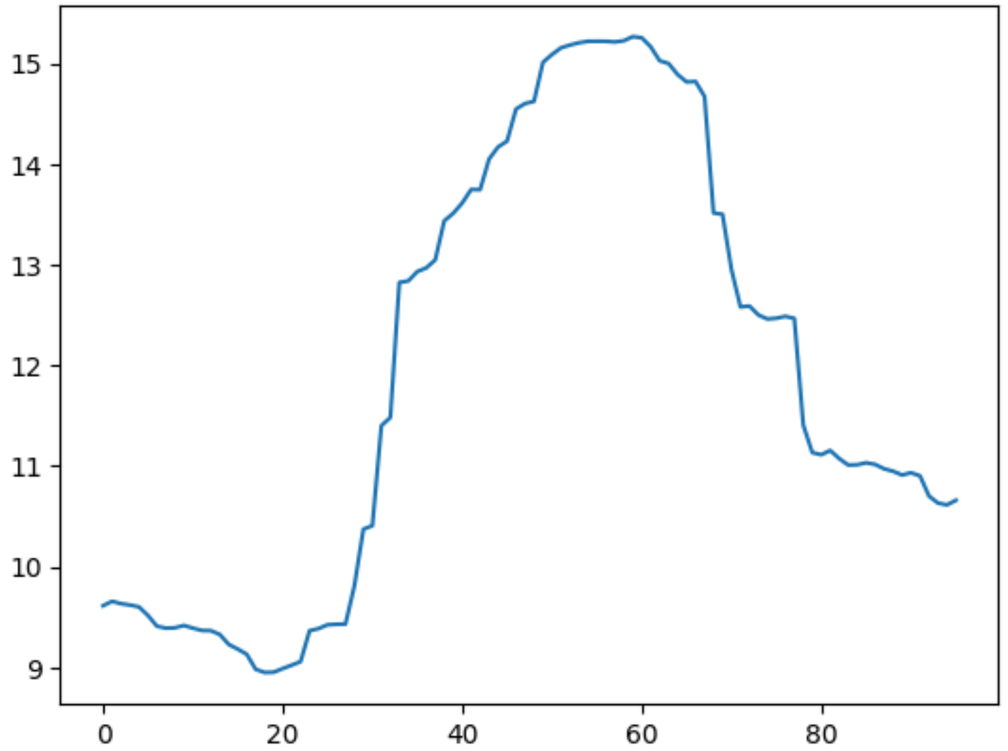
Given the residuals are not normal (Shapiro-Wilk test), the noise addition is done with residuals resampling.



GAN model

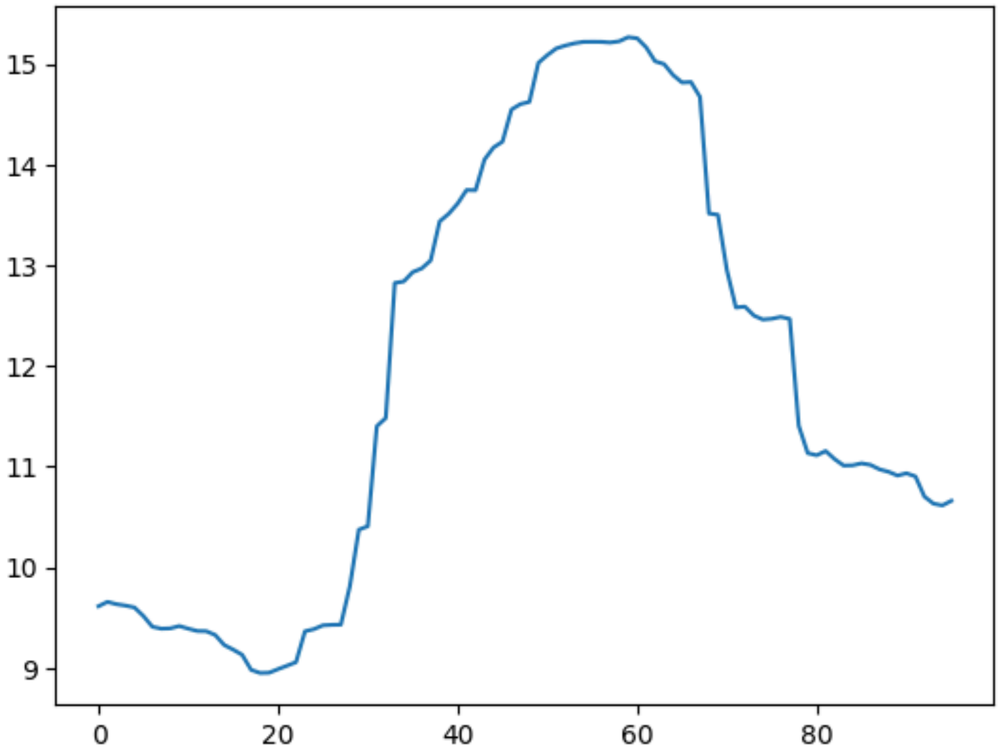
- **Daily external temp. average**



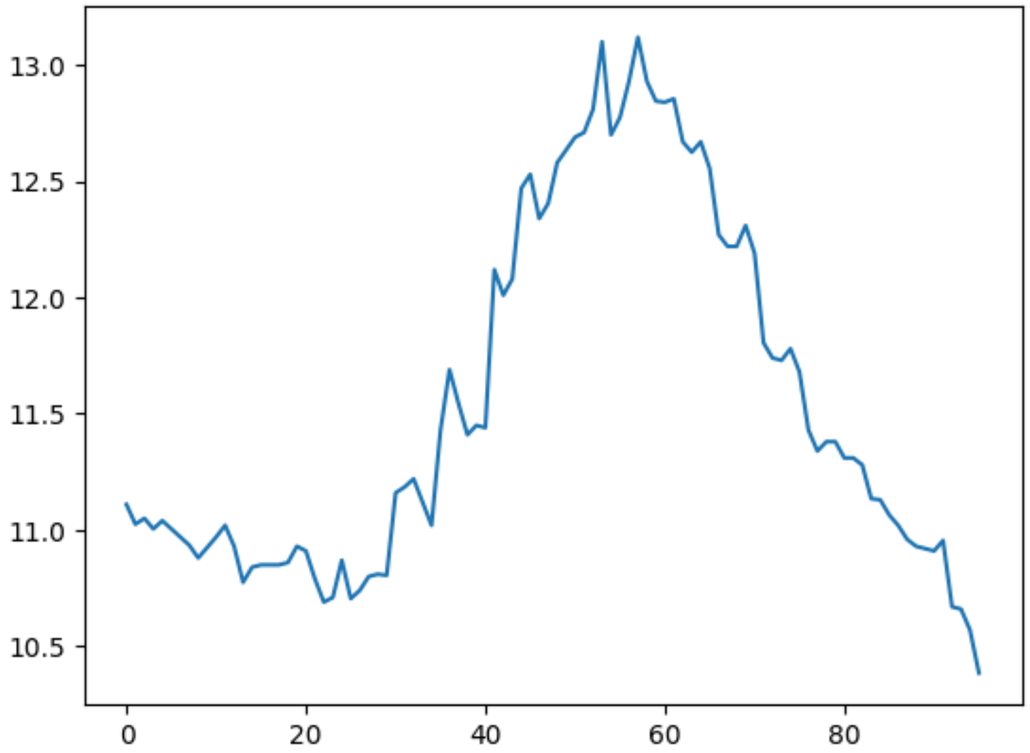


External temperature over the day

GAN model results comparrison

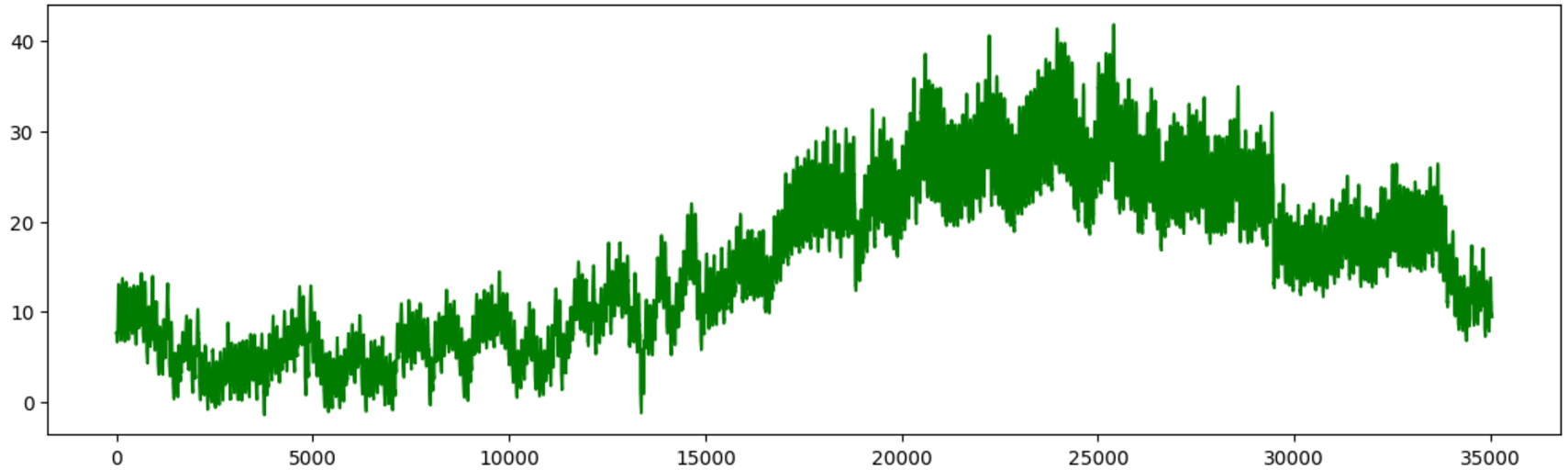


GAN time series output



Real time series temperature

GAN model full results

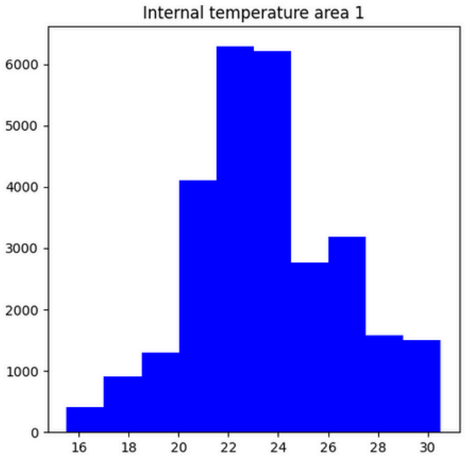


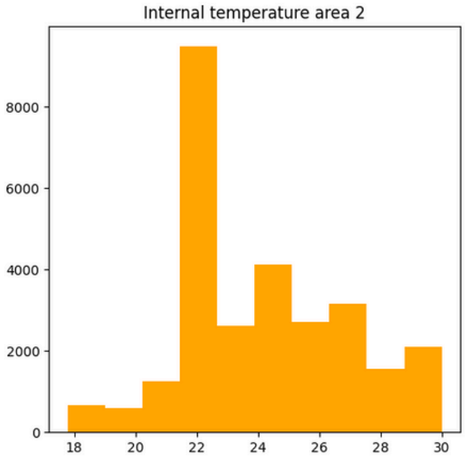
Example of an external temeprature times series generated

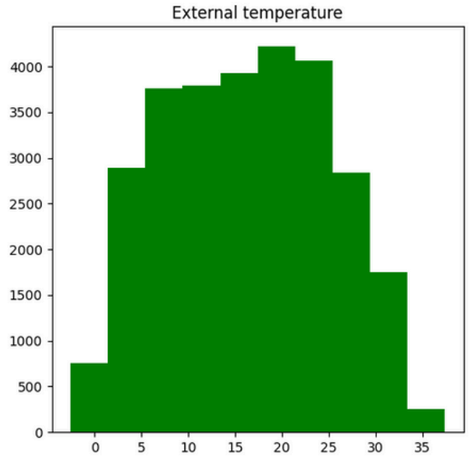
Preprocessing

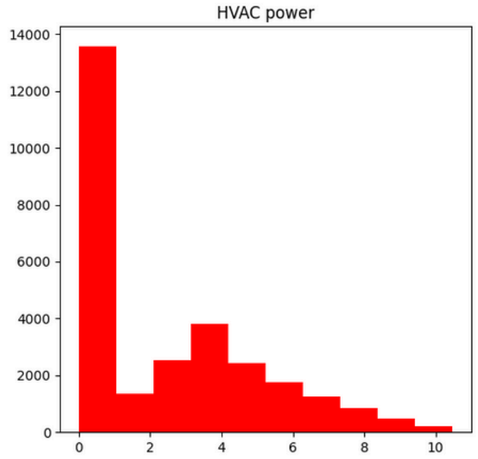
- Removed all days containing at least one NaN value
- Removed days with less than 96 measurements
- Removed outliers based on quantiles
- Added column for Italian holidays
- Retrieved building sizes

Before preprocessing

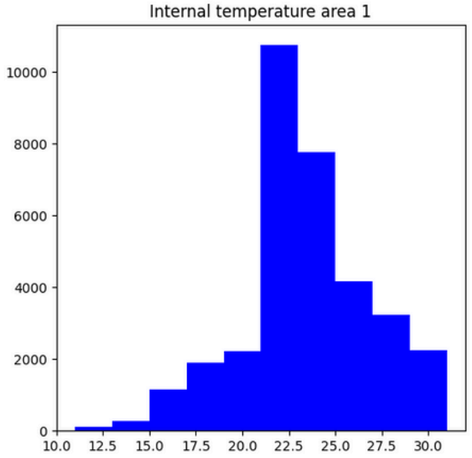


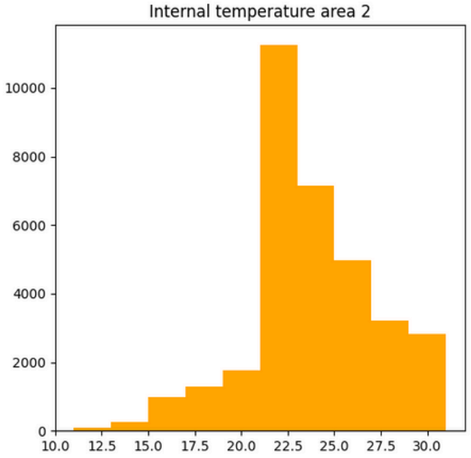


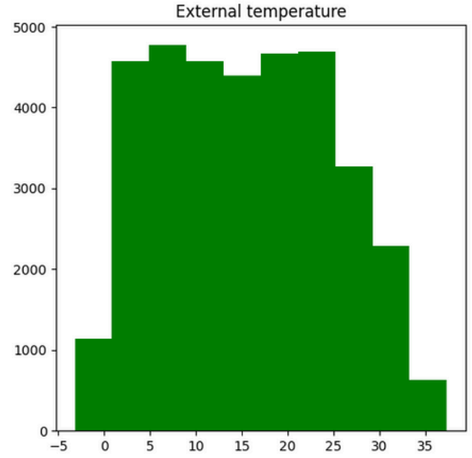


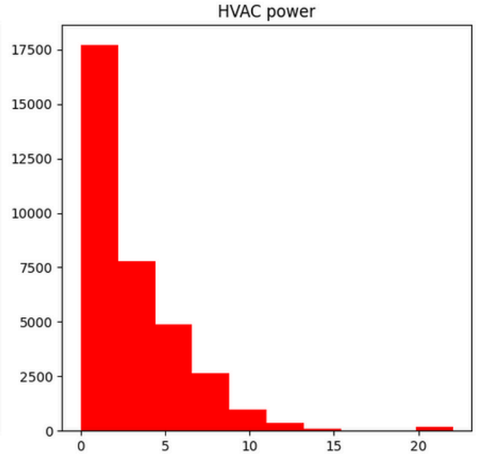


After preprocessing

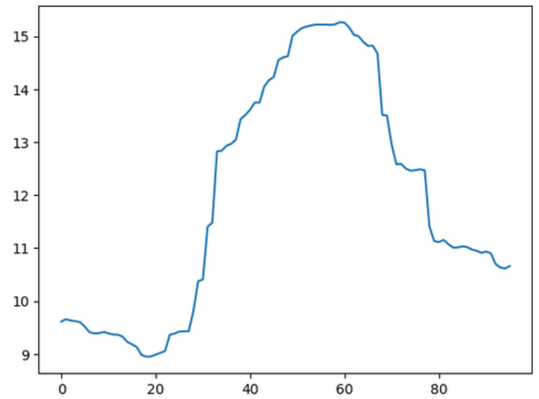








LSTM model



External temperature over the day

Holiday informations

Building size

?

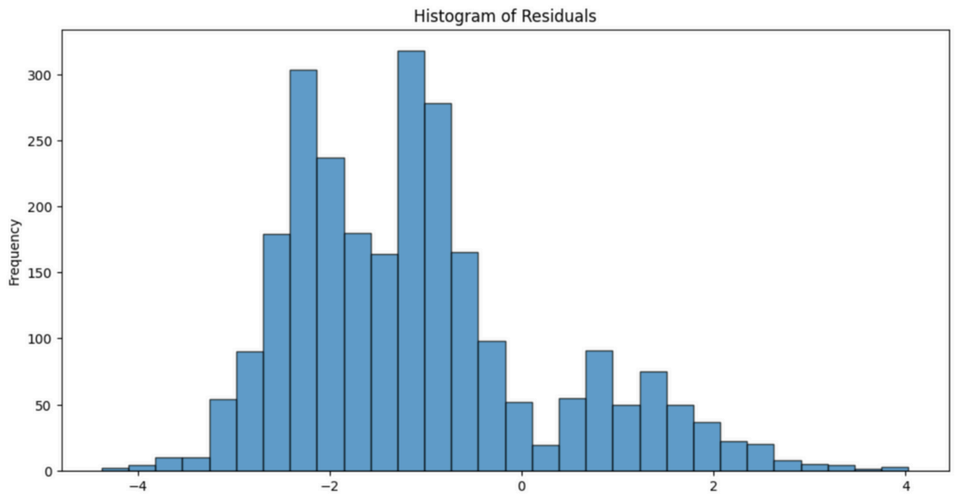
Second Model: LSTM

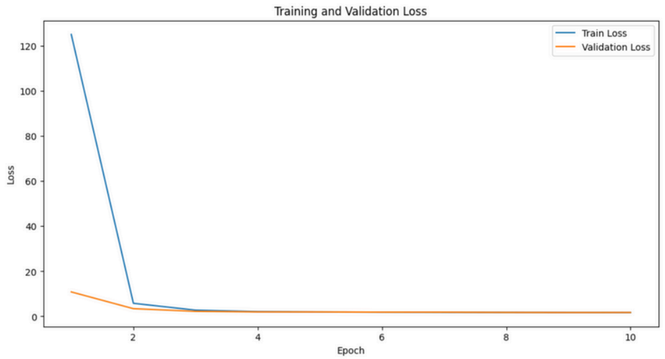
Results:

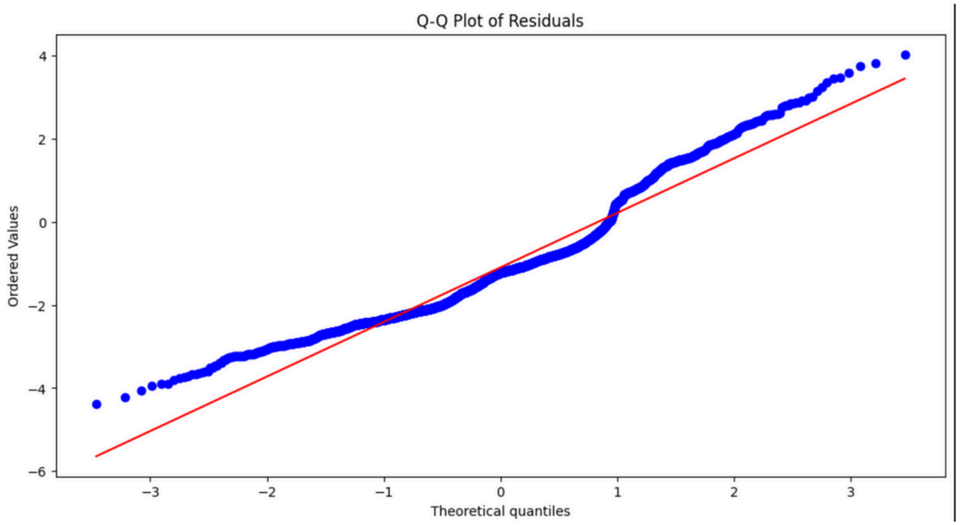
- Average train loss: 2,64
- Average test loss: 14,64

Residuals:

- Shapiro-Wilk Test Statistic: 0.9414 (data are normal)







Conclusions

Regarding the GAN model, as we saw from the previous results even though the residuals still show a small trend, it is able to generate a set of synthetic timeseries for each day of the year.

Unfortunately the LSTM model, given the achieved R^2 value of 0.28, we can’t consider as relevant the generated results. Further tuning is required to obtain better results.

Thank you