Energy Disaggregation:

A Case Study of Large Number of Components

Skolkovo Institute of Science and Technology

Machine Learning | Term 3 | 2023

Professor: Evgeny Burnaev
Teaching Assistant: Petr Mokrov

the team

Bakhtiyar Kazbekov

Msc-2
Energy Systems

Arlan Zhanatbekov

Msc-2
Energy Systems

Arseniy Burov

PhD-1
Material Science

Sahar Moghimian

PhD-1
Digital Engineering

Ilia Kamyshev

PhD-2
Digital Engineering

Skoltech

Presentation Outline

- Introduction to the field
- Motivation & Problem Statement
- Related work
- Research Objectives
- Proposed Approach
- Datasets and Setup
- Results (Performance Evaluation and Comparison)

Skoltech

Introduction

Energy Disaggregation/Non-intrusive Load Monitoring (NILM)

Definition:
The process of breaking down the total energy consumption of a household into its individual appliance-level components

Goal:
Obtaining the information about appliance level energy consumption

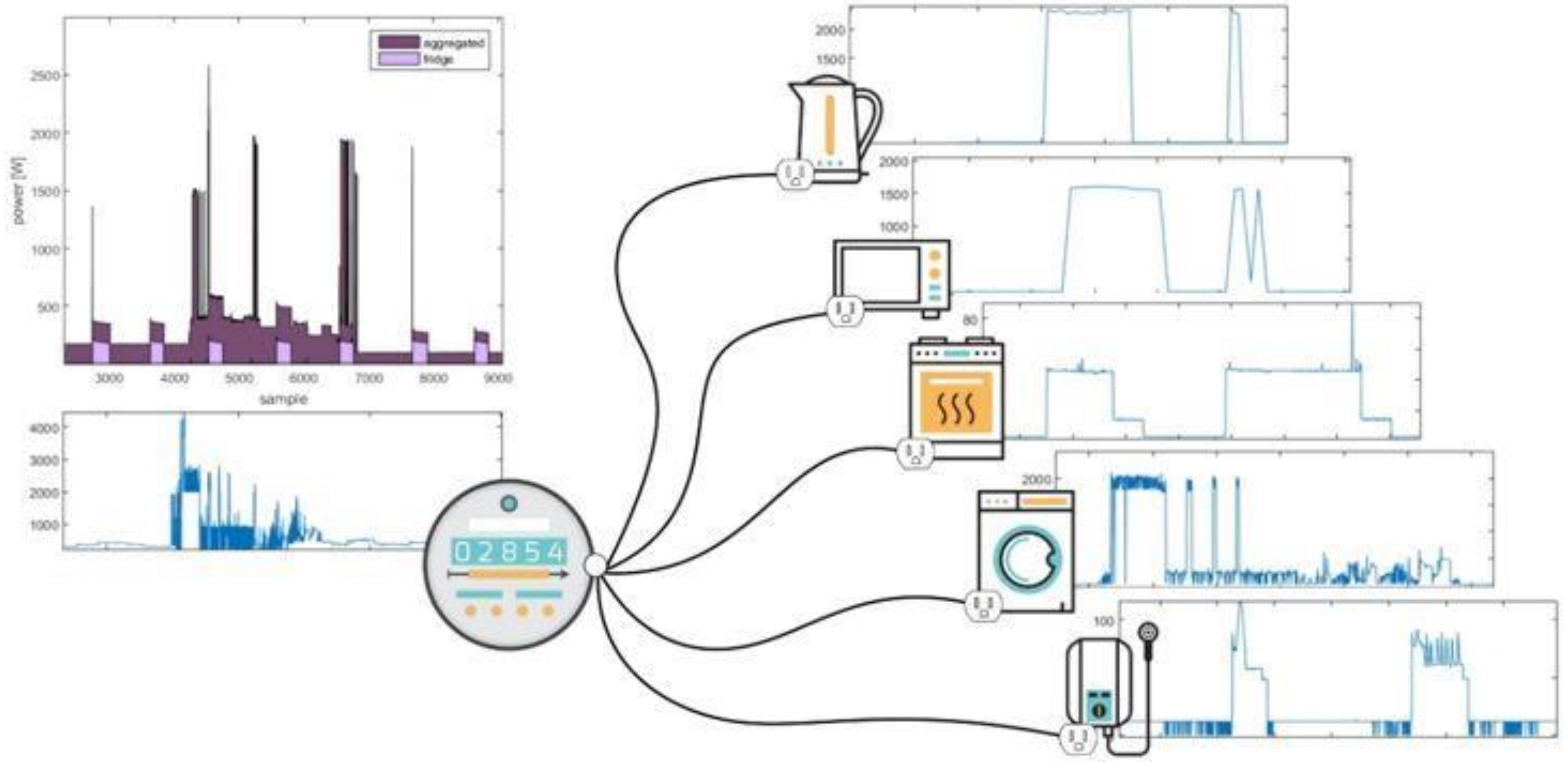


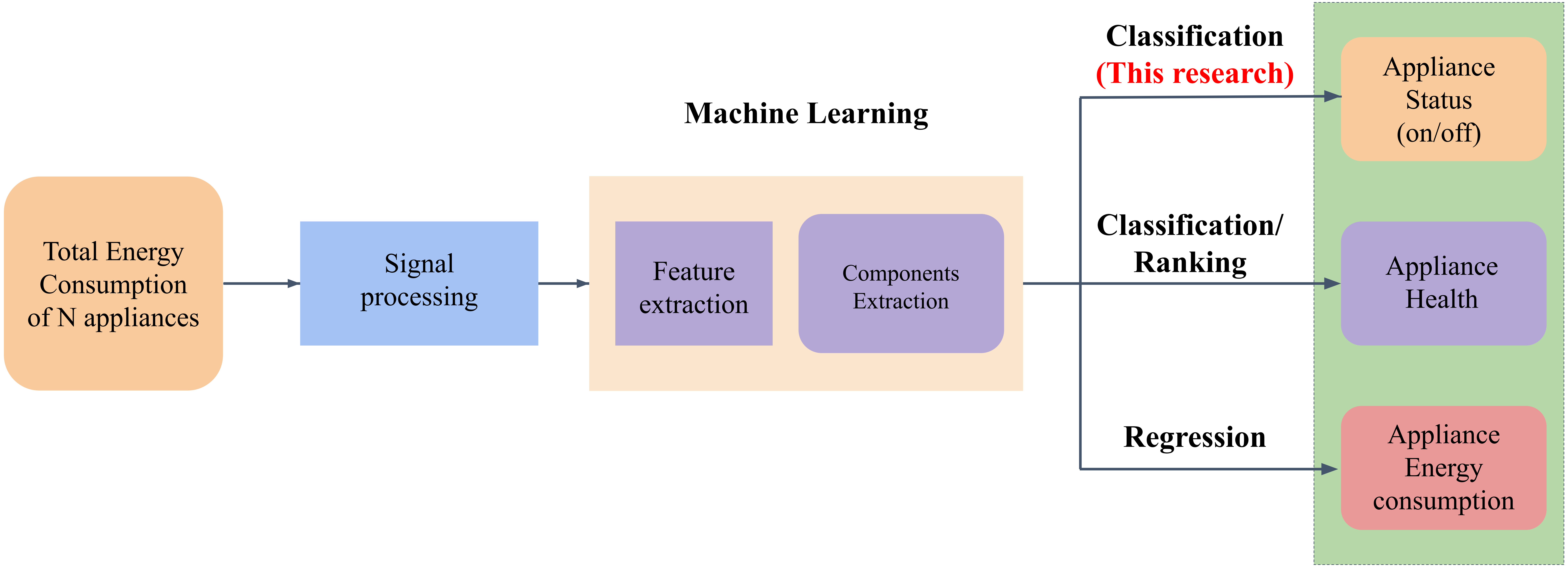
Illustration of Energy Disaggregation

Pujić, Dea & Jelić, Marko & Tomasevic, Nikola & Batic, Marko. (2020). Chapter 10 *Case Study from the Energy Domain*.

Skoltech

G. W. Hart, (1992), *Proceedings of the IEEE*, 80 (12)

Introduction



A principle scheme of Energy Disaggregation (NILM)

Motivation & Problem Statement

- Algorithm complexity

1. Typical algorithms are heavy in memory to be ported to the sensor
2. Sensitive to overfitting
3. Most of the algorithms in the research are trained on 5 most known types of appliances, while datasets contain dozens classes of appliances on average
4. No related studies on “goodness” of disaggregation with a large number of individual components presented in the aggregated signal

- Datasets

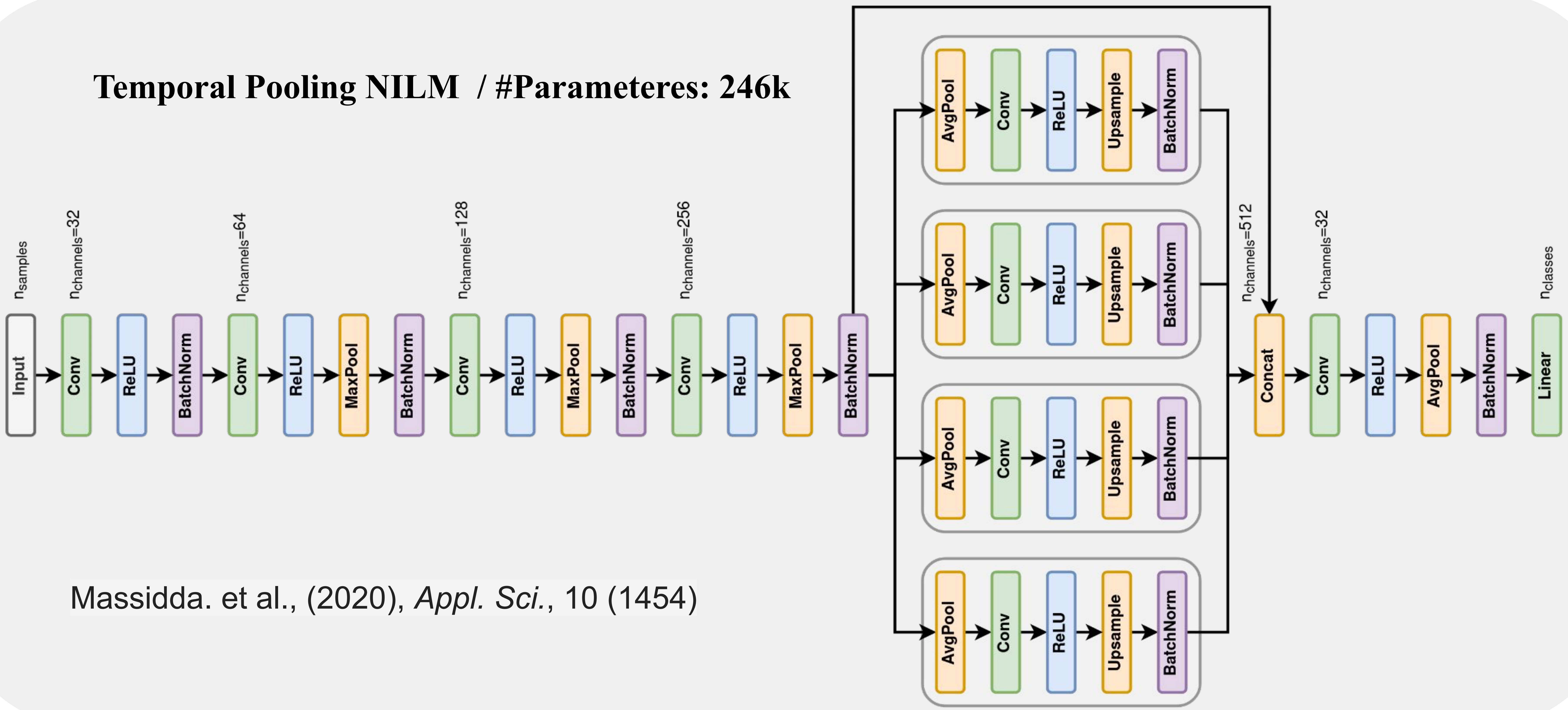
1. Limited number of combinations of different appliances is available
2. Combinations biased towards most frequently used appliances, e.g. washing machine + fridge + air conditioner
3. Only one dataset is available which follows the conventional format i.e. $\{(\mathbf{X}, y)\}^n$, the rest are raw measurements

Skoltech

Kamyshev, (2021), *arXiv*, 2106.02352

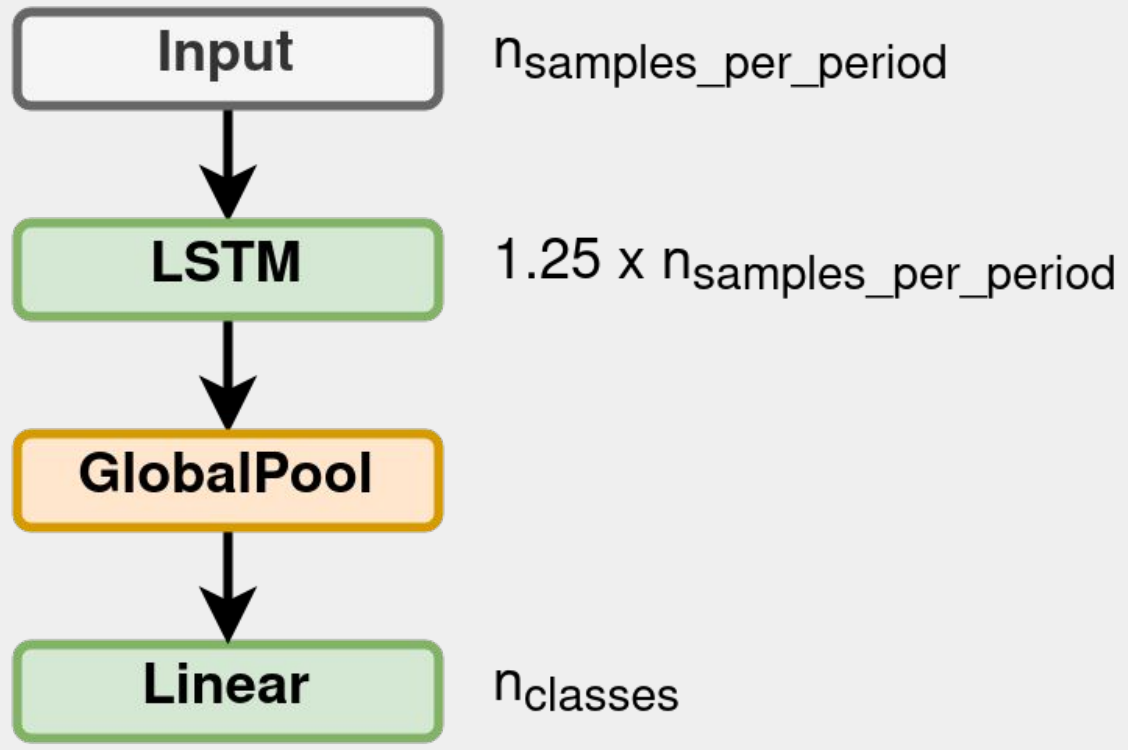
Related work

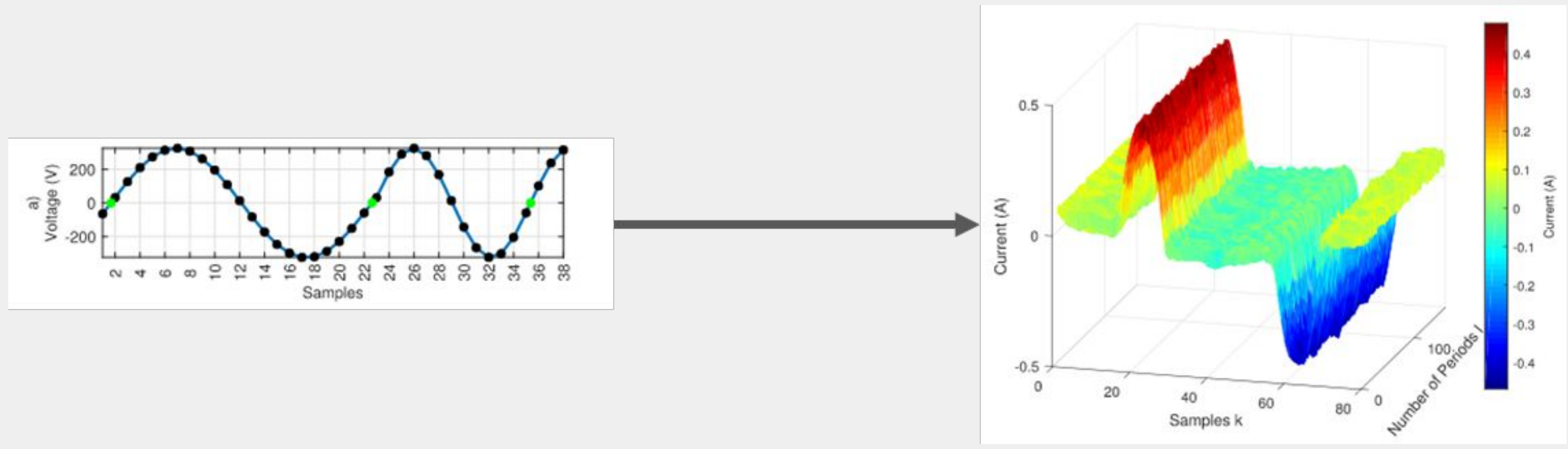
Temporal Pooling NILM / #Parameteres: 246k



Massidda. et al., (2020), *Appl. Sci.*, 10 (1454)

Held P. et al. / #Parameteres: 29k





Frequency Invariant Transformation of Periodic Signals

Help P. et al., (2018), *IEEE Transactions on Smart Grid*, 10 (5)

Skoltech

Research Objectives

- Develop a **classification** algorithm which:
  1. accounts the physics of the problem
  2. is not overfitting
  3. has lower complexity
  4. is capable of decomposing signals with large number of components

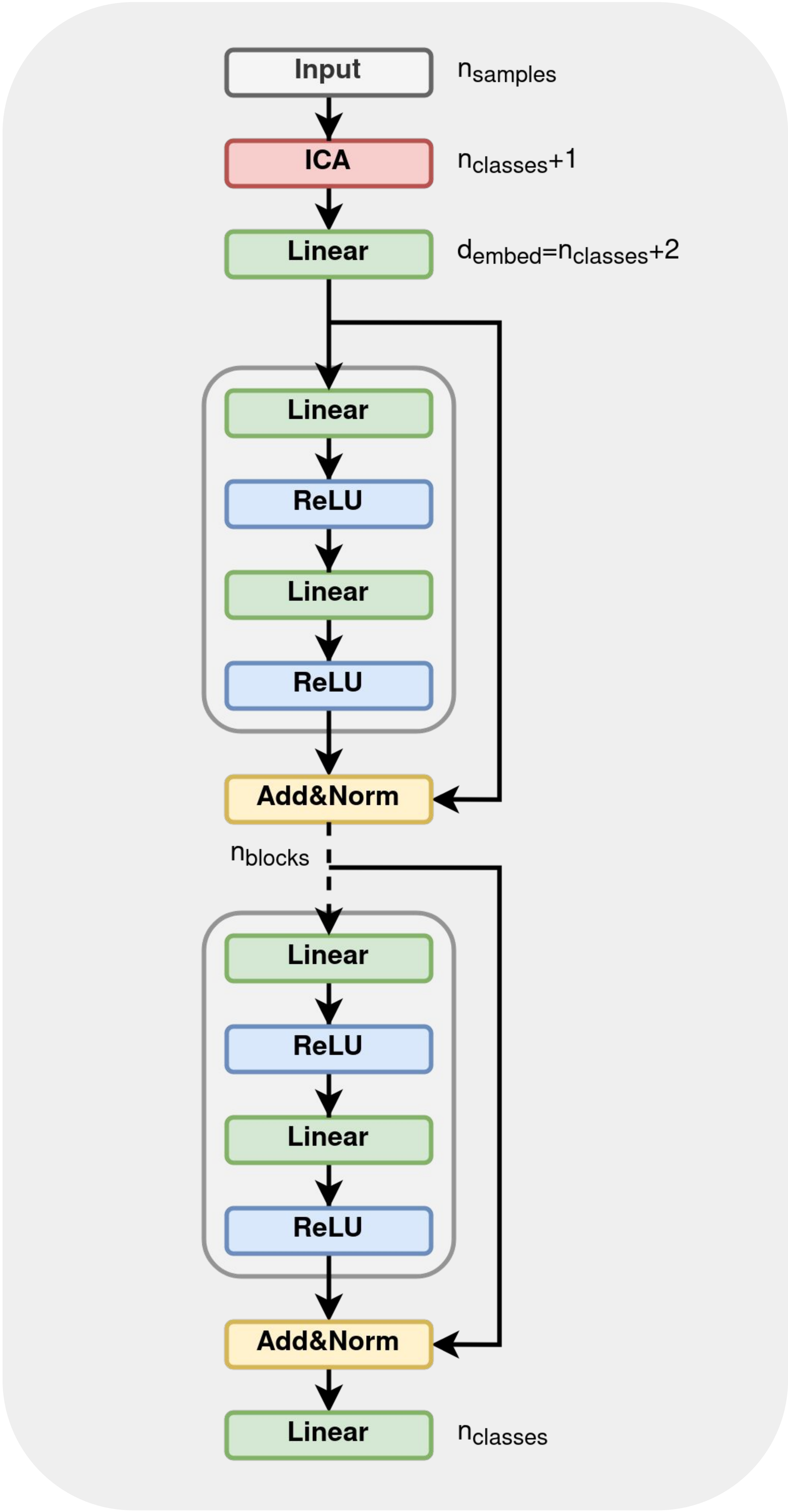
- Evaluate the impact of growing number of components on classification algorithm

Skoltech

Proposed Approach

ICAResNetFFN

- Generate synthetic classes of appliances
- Generate samples of aggregated signals once based on real and then on synthetic appliances
- Apply ICA on the previously obtained data with $n_{components} = n_{classes} + 1$
- Take unmixing matrix **U** and insert to FFN with residual connections to transform the input signal **X** i.e. $\mathbf{X'}=\mathbf{XU}^T$
- Train the FFN
- Compare the proposed approach with 2 promising algorithms: Temporal Pooling NILM, LSTM on signals passed through FIT-PS



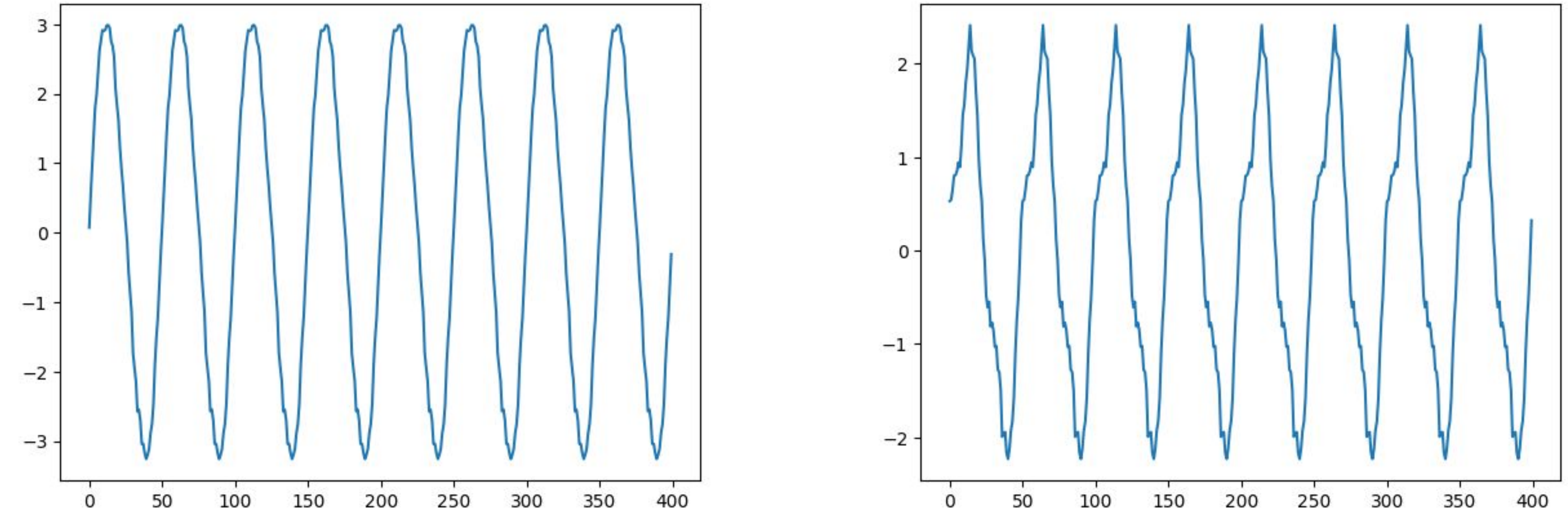
Proposed FFN/ #Parameteres: 10k

Setup

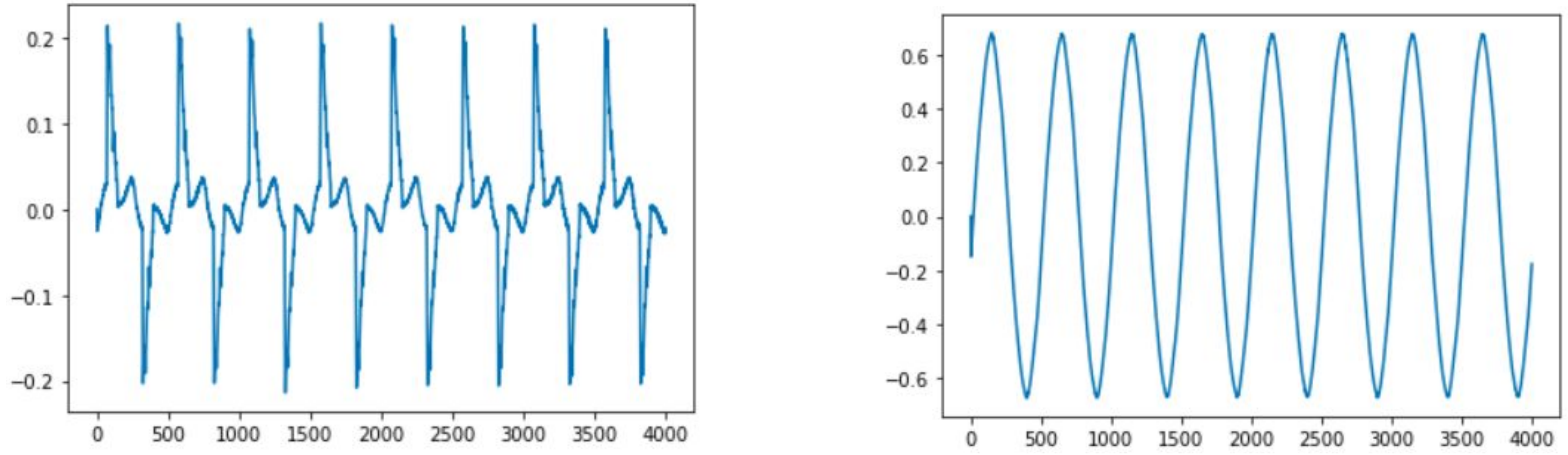
- Input data: aggregated signals / Output data: binary vector i.e. $Y=\{0,1\}^{n\_classes}$
- Min number of components = 1, Max number of components = 10 i.e. min(sum(Y)) = 1 and max(sum(Y)) = 10 respectively
- Each component may be repeated from 1 to 10 times
- Sampling rate: 3000 / Samples per signal: 400 i.e. input size Nx400
- Output size: Nx15 (N is a batch size)
- 7k train / 1k validation / 3k test samples of aggregated signals and their labels
- Number of epochs: 250
- Number of experiments: 2 (synthetic and real classes of appliances)
- Number of models: 3
- Total number of model runs: 6

Skoltech

Datasets



Synthetic classes of appliances.

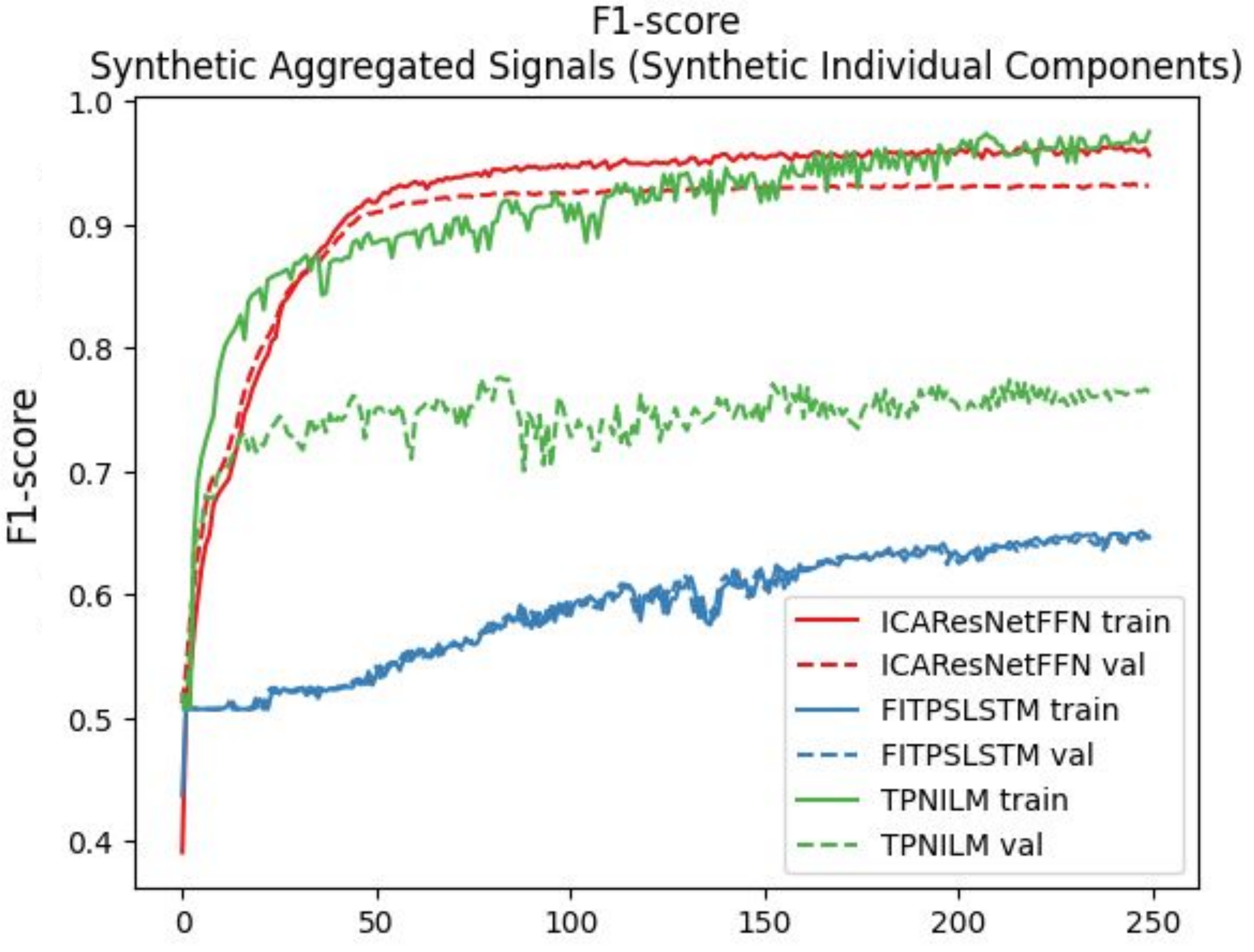


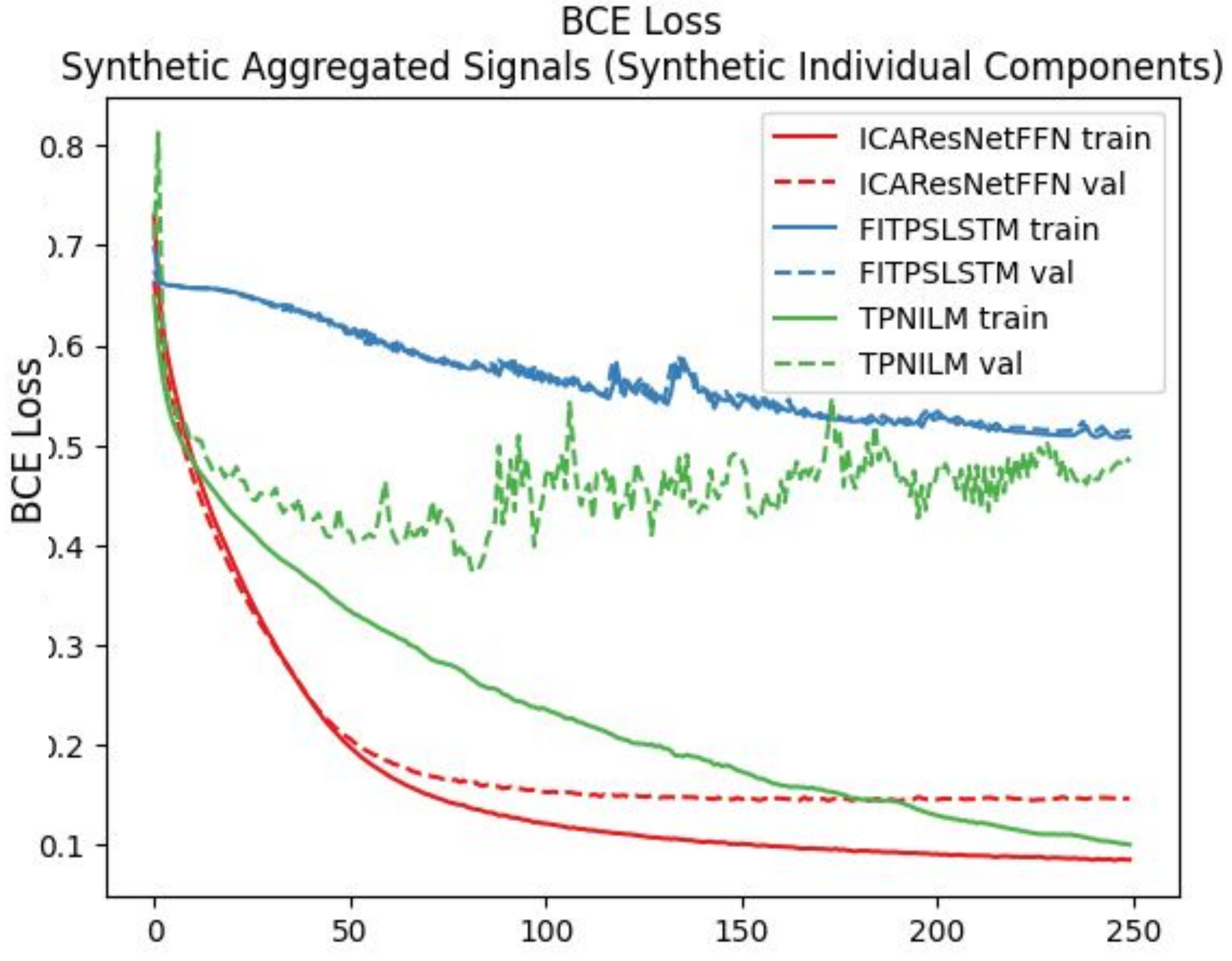
Real classes of appliances from the Plug-Load Appliance Identification Dataset (PLAID).

Medico R. et al., (2020), *Sci. Data*, 7(1)

Kamyshev et al. (2023), *PowerTech*

Results. Synthetic classes. Train





F1-score

$$F\text{-measure} = \frac{2 \cdot precision \cdot recall}{precision + recall},$$

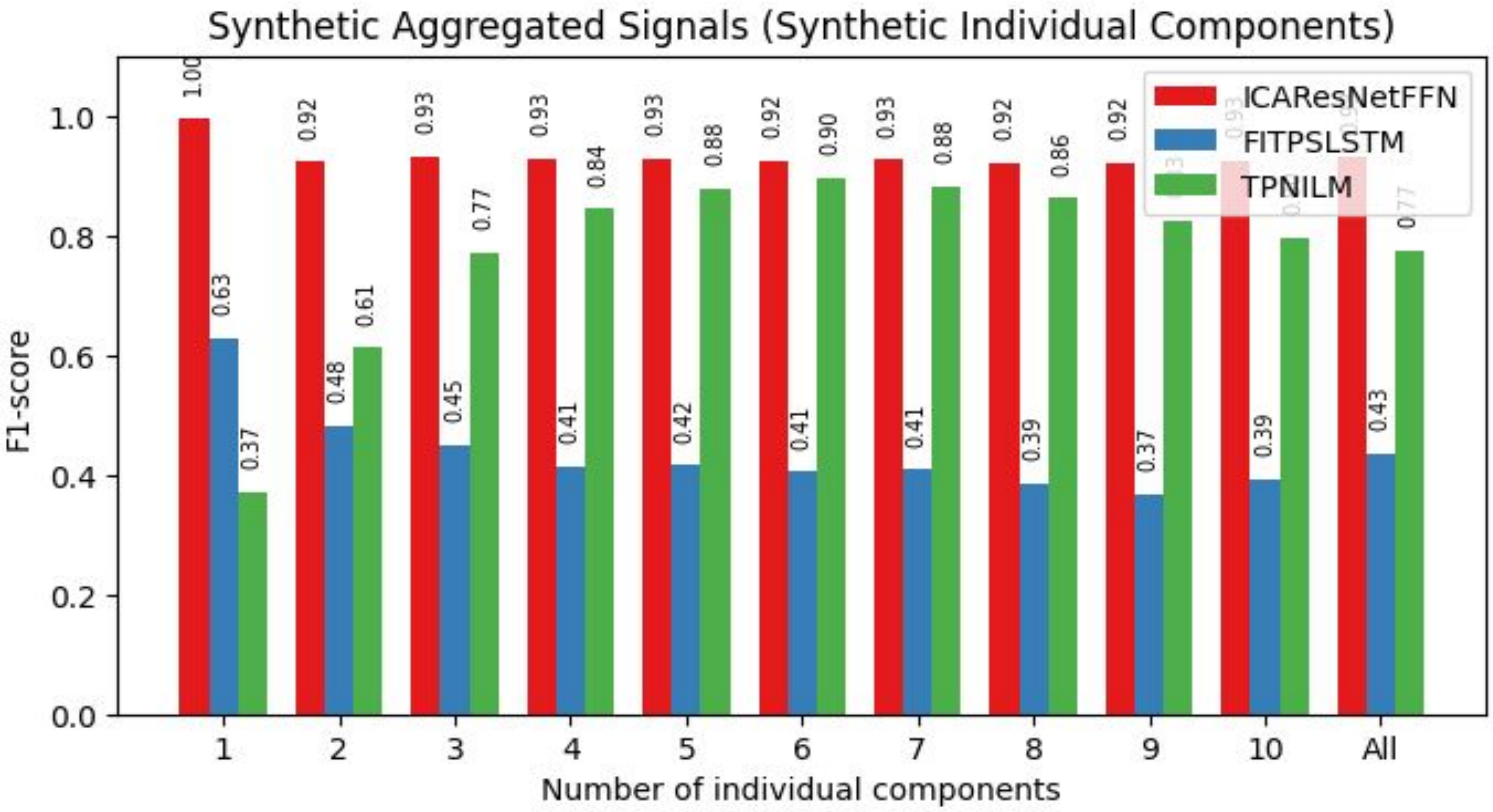
$$precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \quad recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}.$$

Loss Function: Binary- cross entropy

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

Skoltech

Results. Synthetic classes. Test



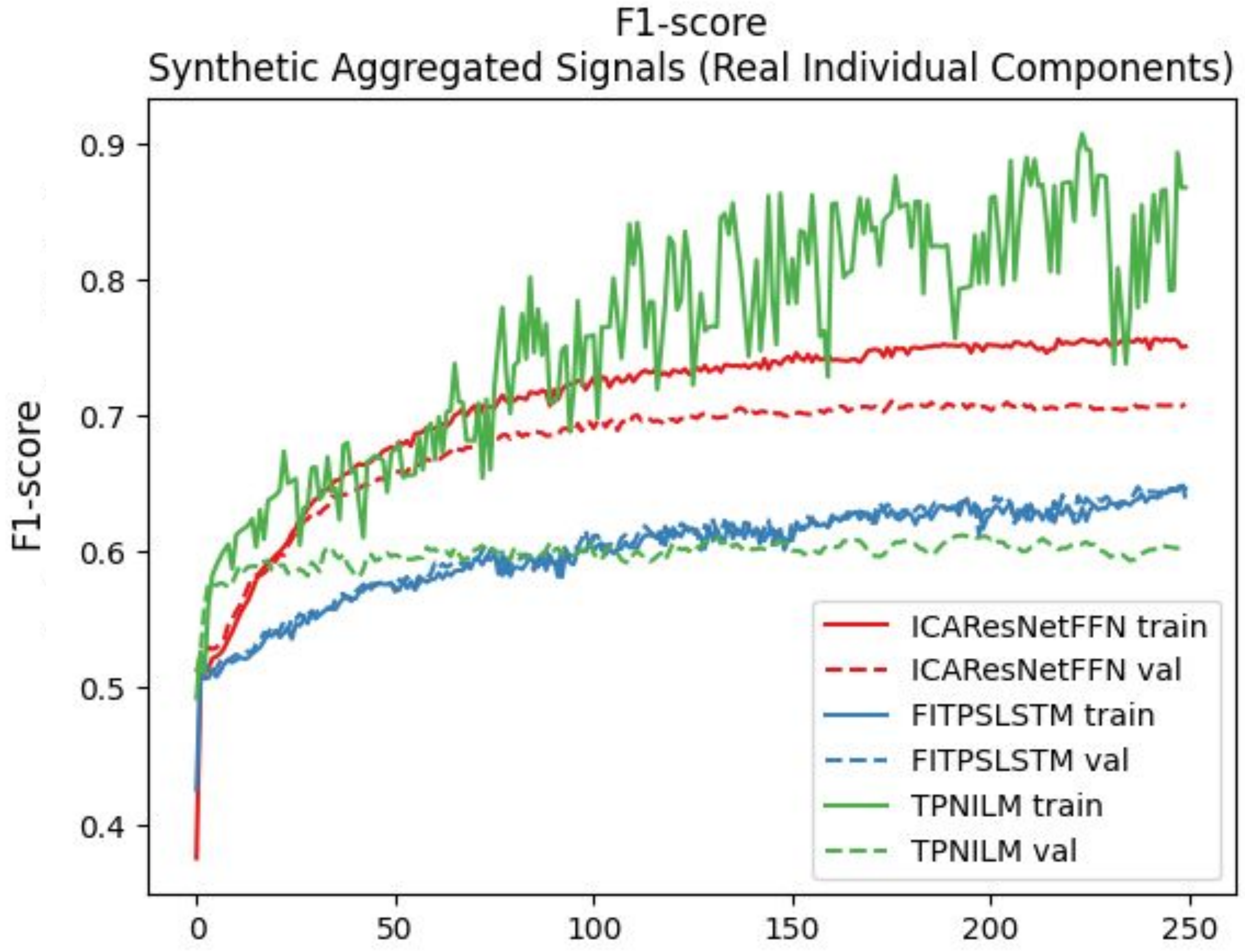
Skoltech

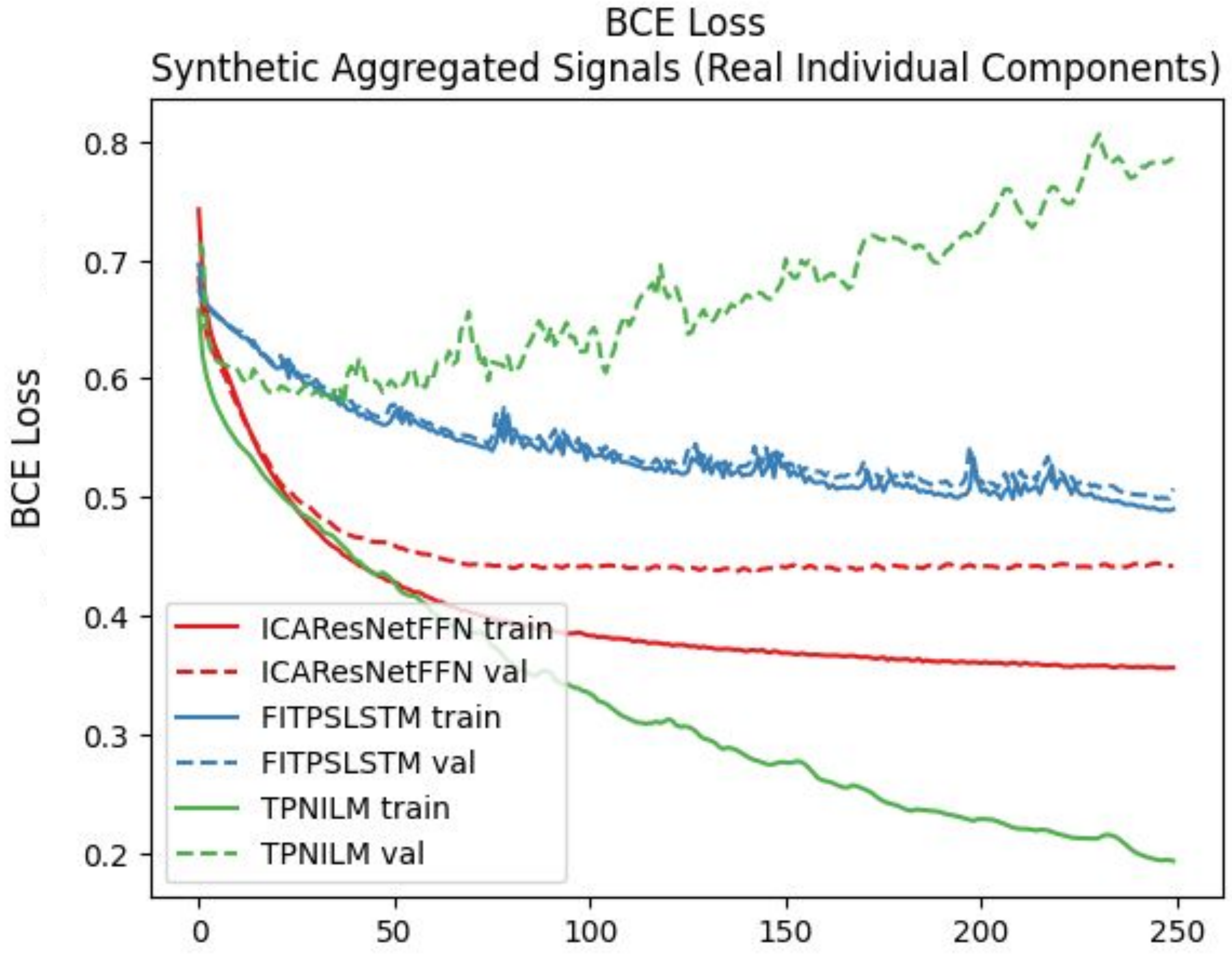
Results. Synthetic classes. Test

Model: ICAResNetFFN

Evaluation Metrics	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*All*
F1-score (samples)	0.996	0.923	0.931	0.929	0.927	0.924	0.927	0.921	0.921	0.925	0.932
F1-score (macro)	0.993	0.934	0.939	0.934	0.929	0.926	0.928	0.924	0.924	0.927	0.929
Precision (samples)	0.994	0.990	0.986	0.987	0.964	0.966	0.966	0.960	0.974	0.971	0.976
Precision (macro)	0.987	0.984	0.983	0.982	0.957	0.959	0.963	0.958	0.974	0.969	0.968
Recall (samples)	1.000	0.893	0.900	0.892	0.905	0.897	0.897	0.895	0.880	0.890	0.905
Recall (macro)	1.000	0.893	0.900	0.890	0.905	0.896	0.897	0.895	0.881	0.891	0.895
Accuracy	0.990	0.762	0.700	0.576	0.481	0.362	0.300	0.310	0.286	0.224	0.499

Results. Real classes. Train





F1-score

$$F\text{-measure} = \frac{2 \cdot precision \cdot recall}{precision + recall},$$

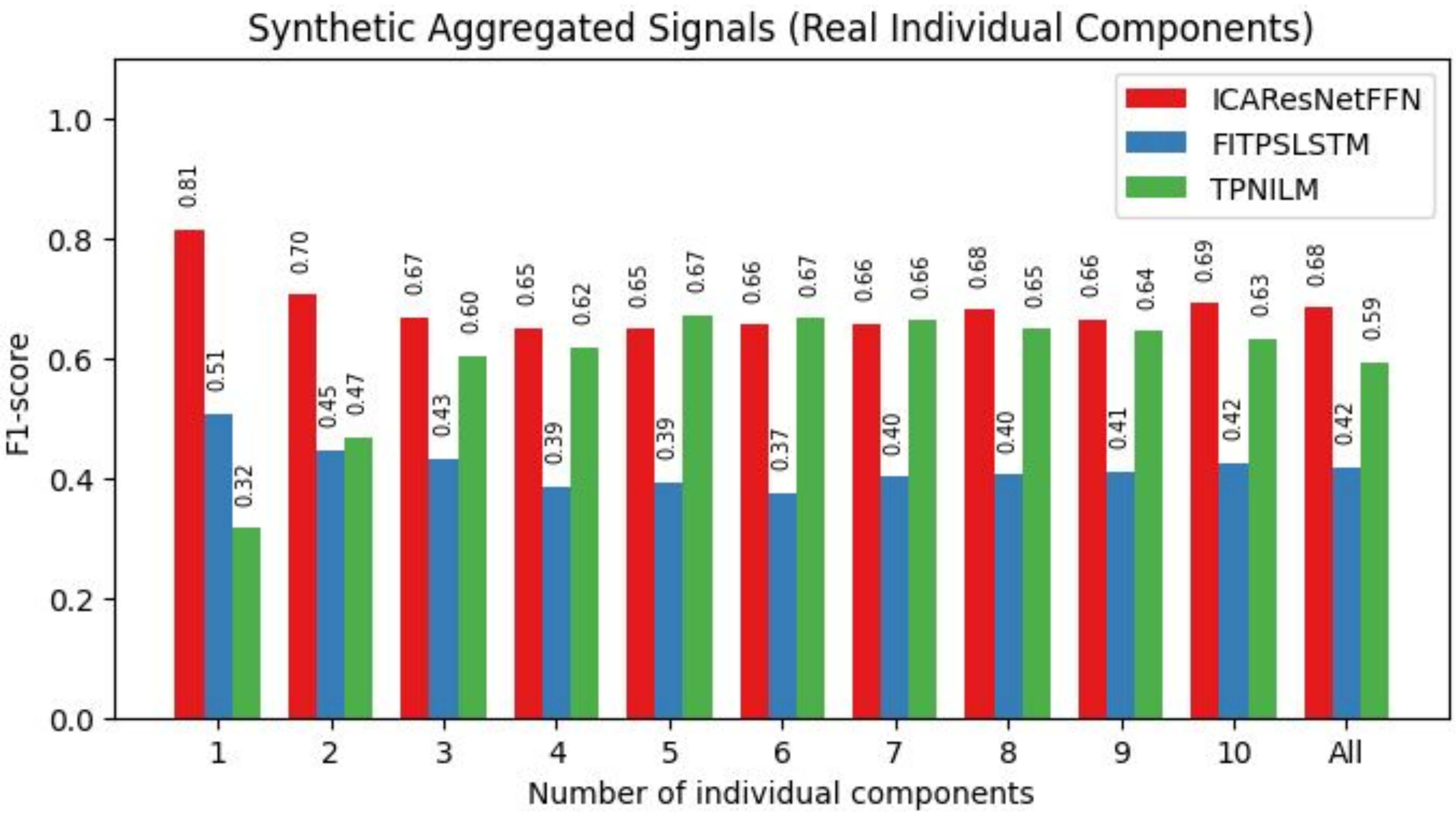
$$precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \quad recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}.$$

Loss Function: Binary- cross entropy

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

Skoltech

Results. Real classes. Test



Skoltech

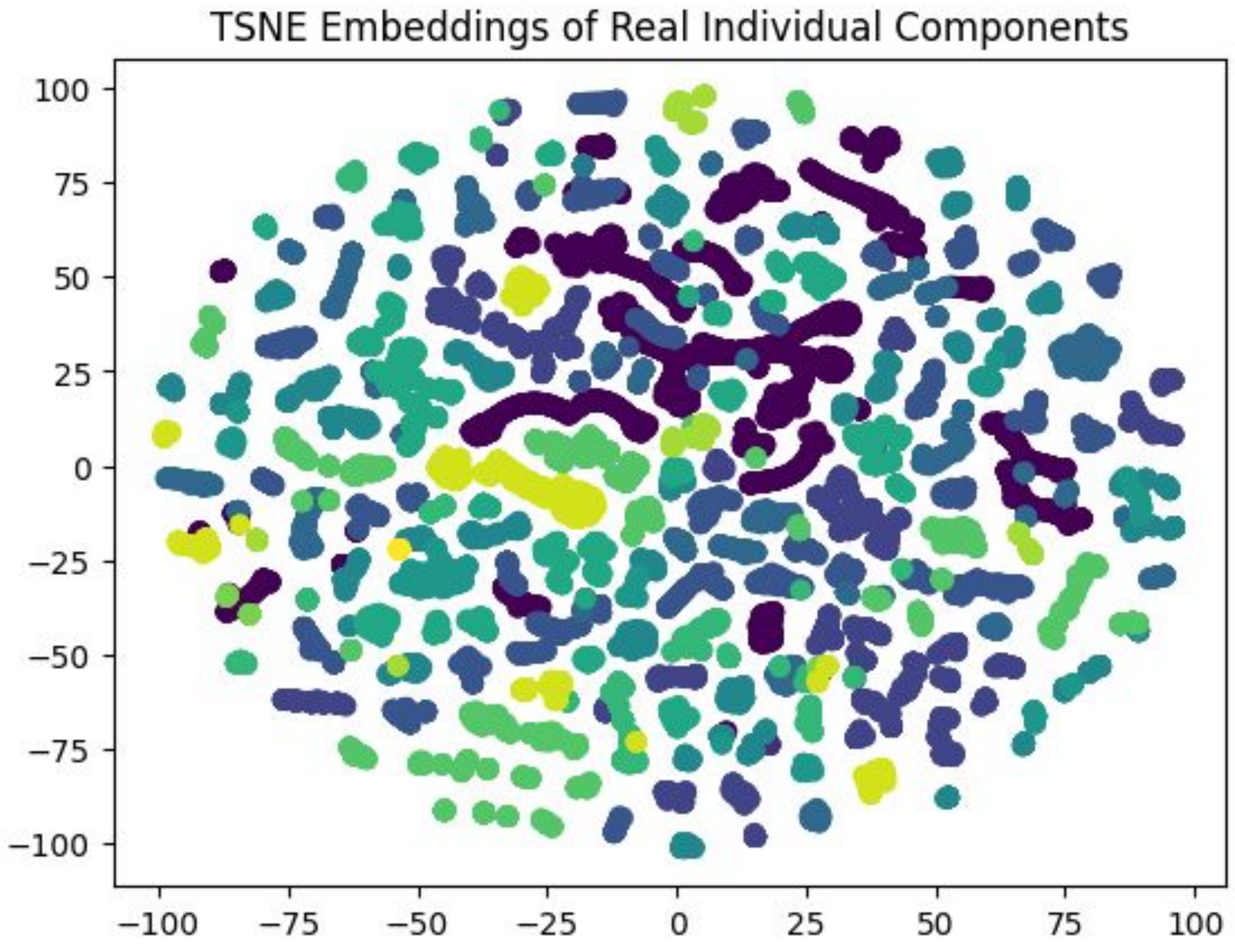
Results. Real classes. Test

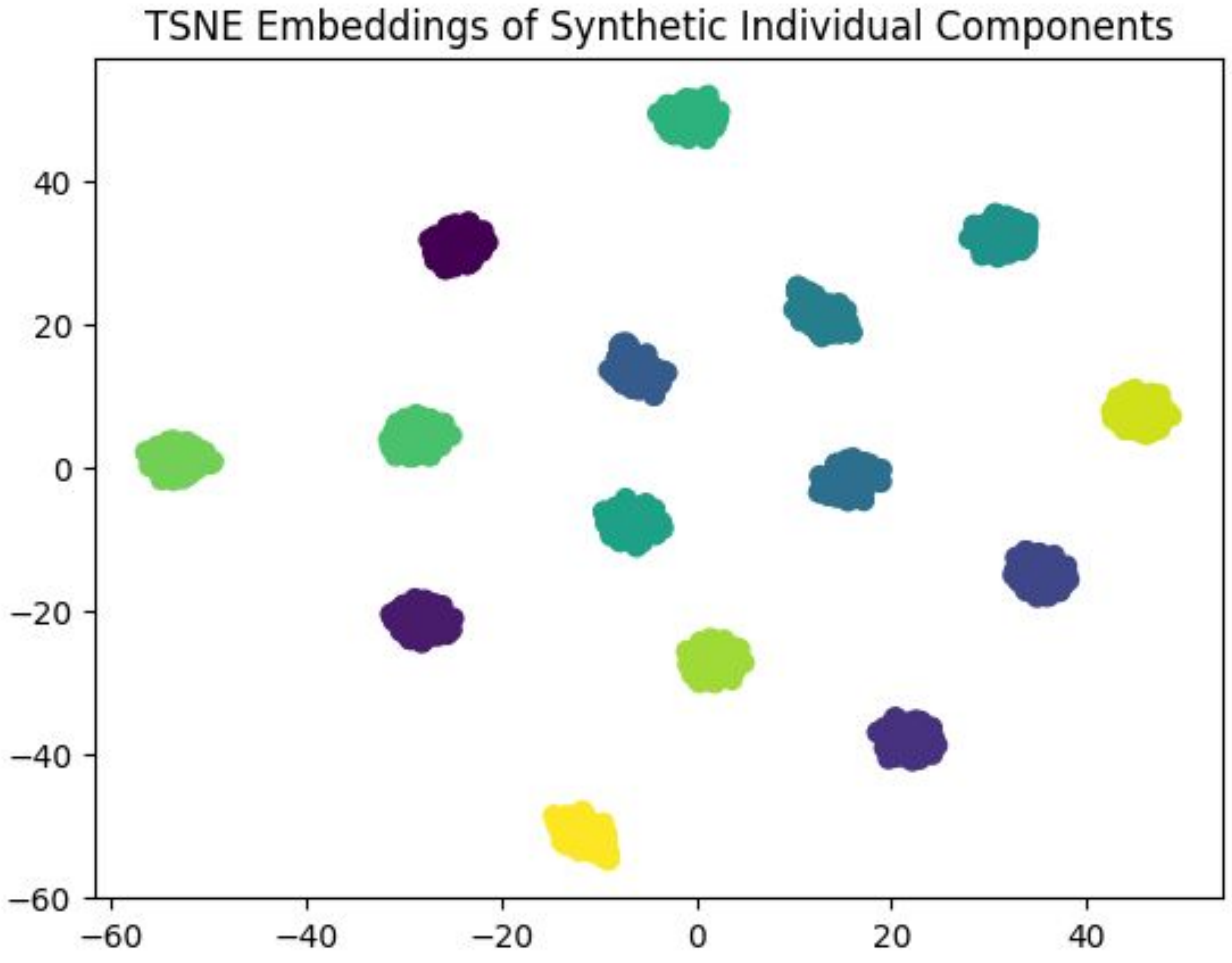
Model: ICAResNetFFN

Evaluation Metrics	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*All*
F1-score (samples)	0.814	0.704	0.668	0.650	0.647	0.655	0.657	0.680	0.664	0.690	0.683
F1-score (macro)	0.804	0.683	0.621	0.617	0.616	0.622	0.625	0.657	0.639	0.679	0.648
Precision (samples)	0.796	0.845	0.839	0.838	0.816	0.808	0.804	0.845	0.851	0.877	0.832
Precision (macro)	0.772	0.756	0.658	0.698	0.712	0.727	0.717	0.803	0.800	0.839	0.776
Recall (samples)	0.852	0.650	0.603	0.577	0.572	0.585	0.578	0.592	0.563	0.589	0.616
Recall (macro)	0.852	0.650	0.598	0.579	0.571	0.578	0.574	0.589	0.561	0.596	0.587
Accuracy	0.743	0.262	0.076	0.029	0.005	0.010	0.000	0.000	0.000	0.005	0.113

Discussion

- Real data is a way complex and some classes **are not separable at all**





Skoltech

Thank you

Skoltech

Extra slides

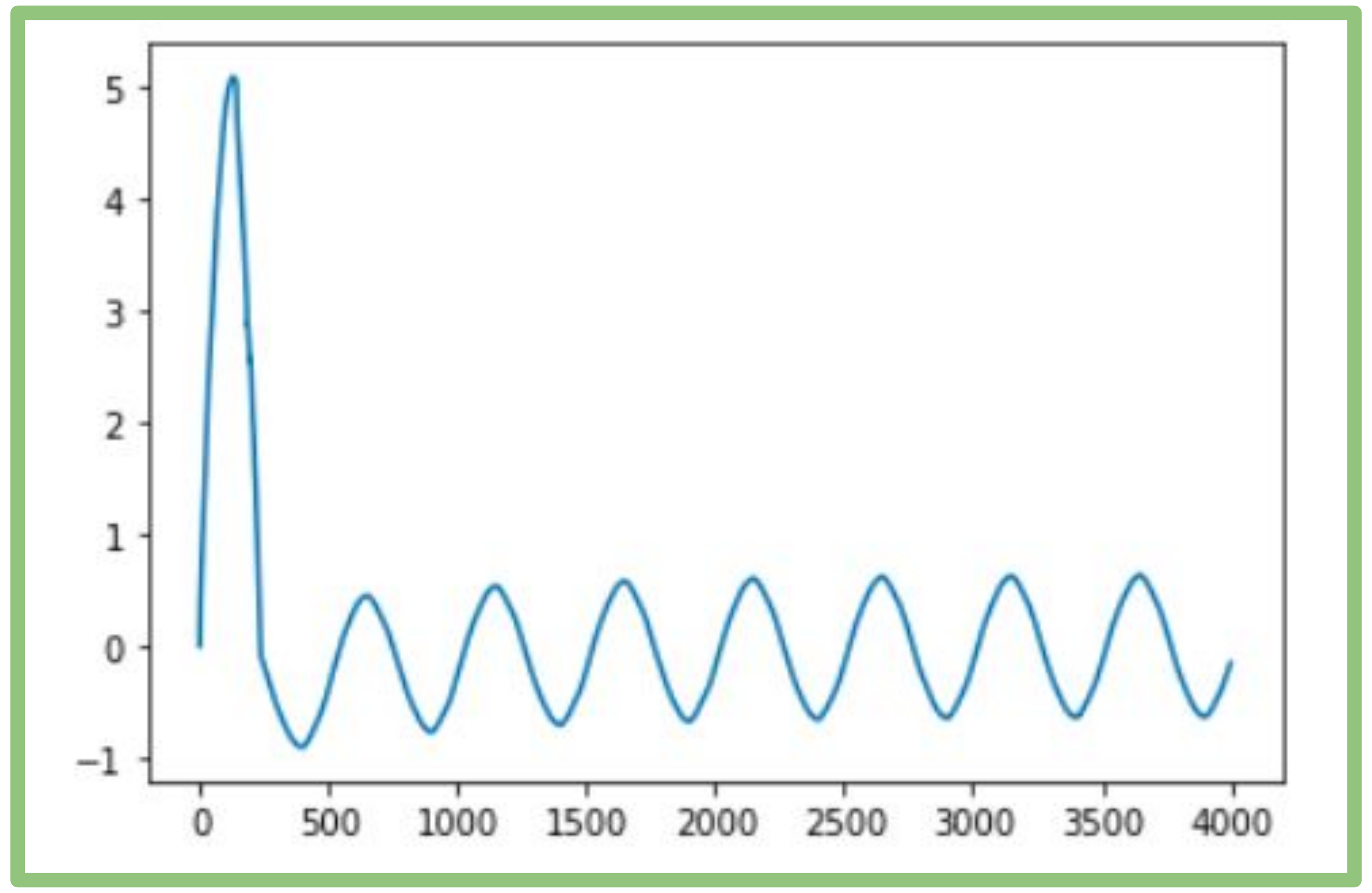
Dataset

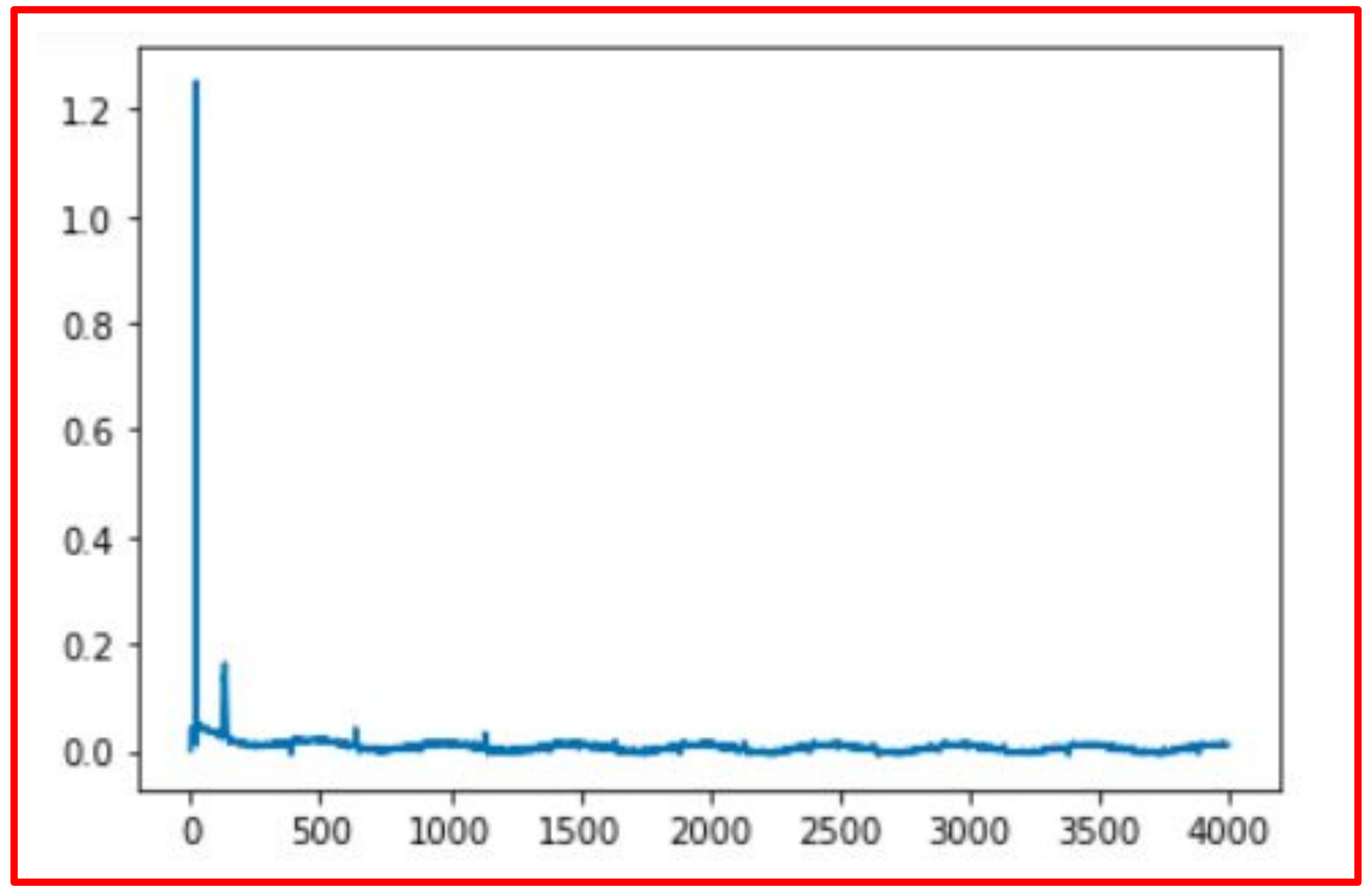
Publicly available NILM dataset:

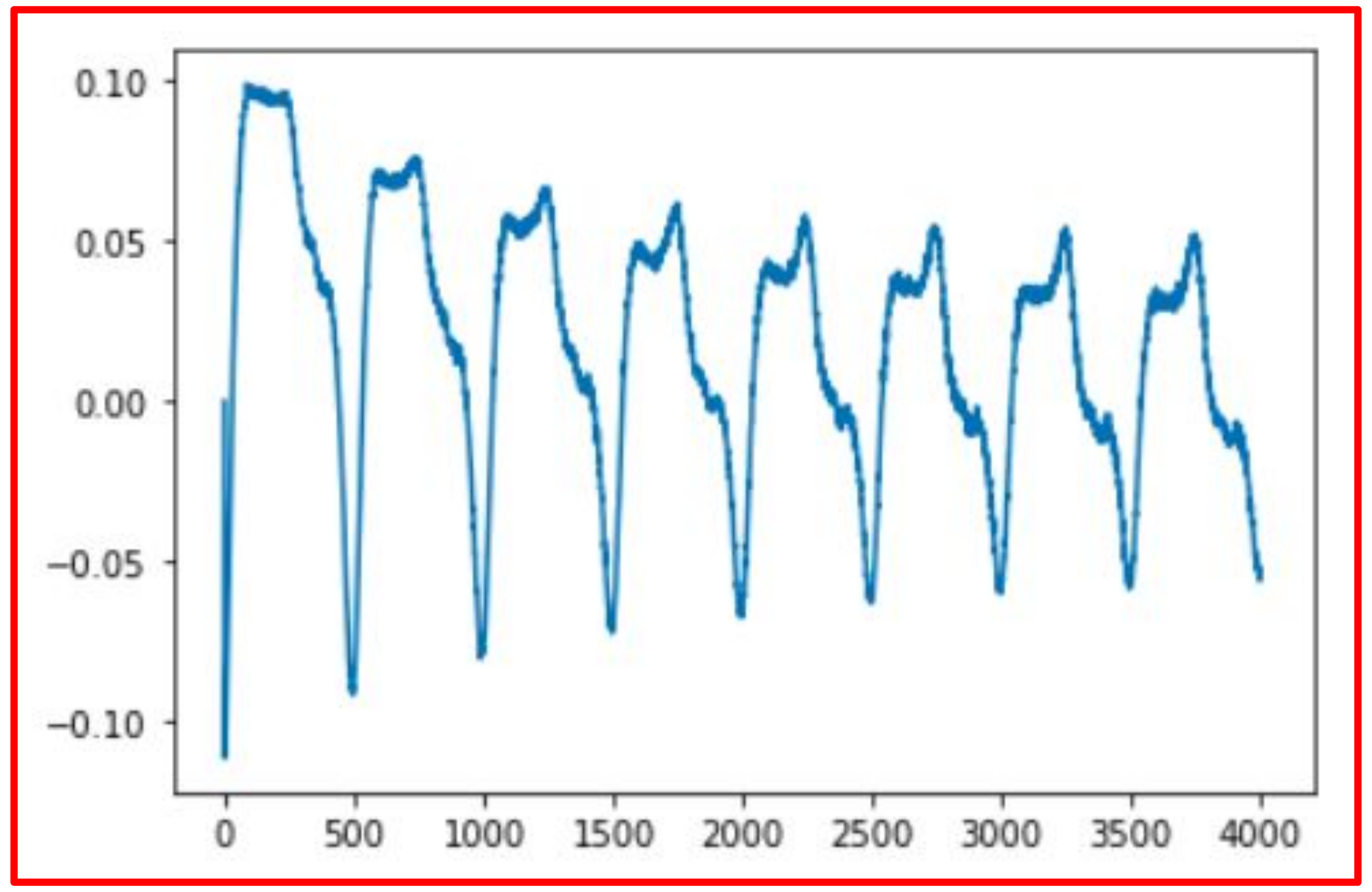
Dataset	Resolution	# Appliances	Ground truth	Access	Citations
PLAID (US)	30 kHz	12	Individual appliances	Open	~120
REDD (US)	16.5 kHz 1 Hz	20	Labelled events Submeter channels	Private	~1130
BLUED (US)	12 kHz	30	Labelled events	Private	~280
UK-DALE (UK)	16 kHz 1 Hz	40	Labelled events Submeter channels	Open	~400
WHITED (EU, US)	41 kHz	55	Individual appliances	Open	~50

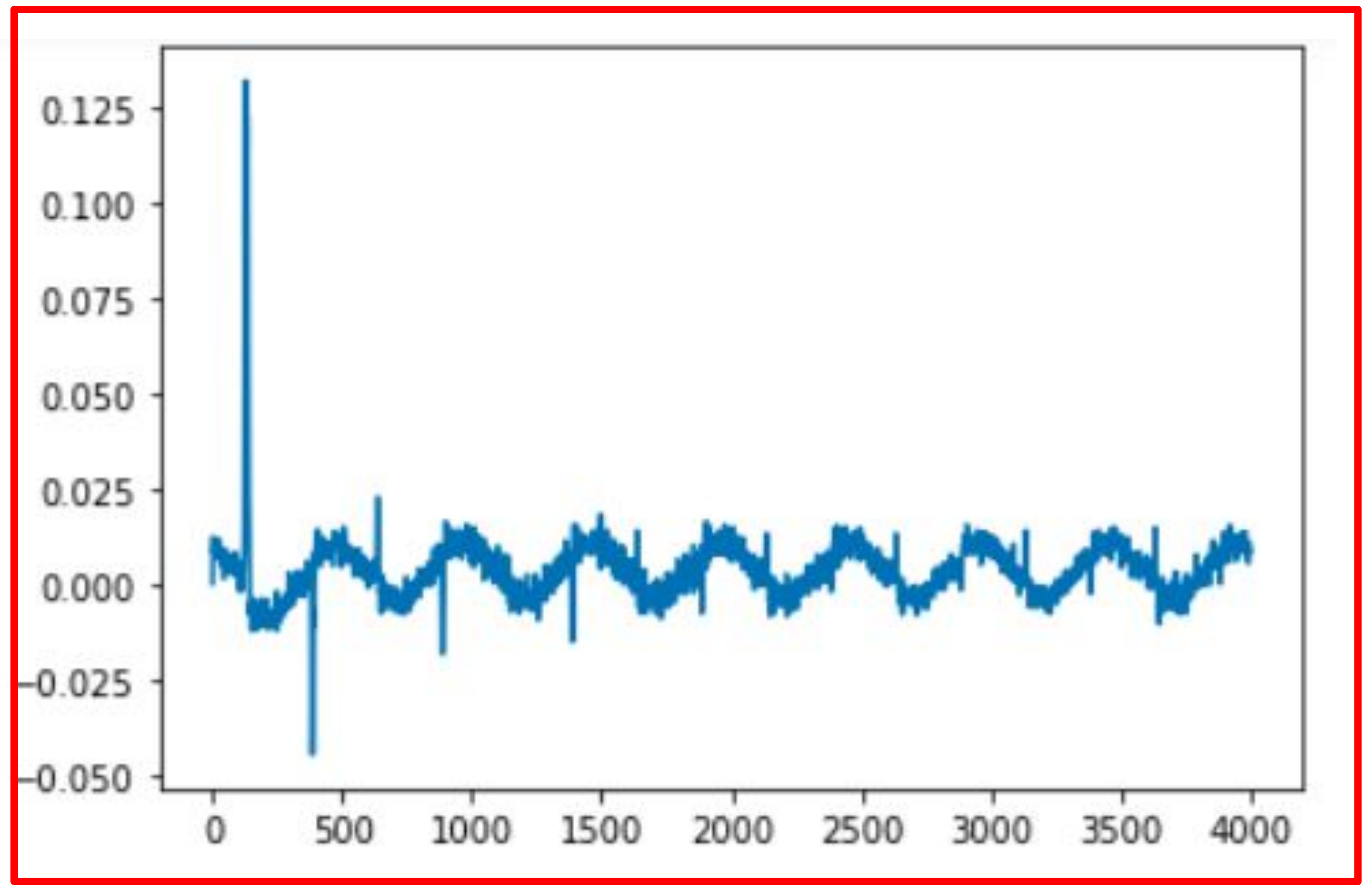
Skoltech

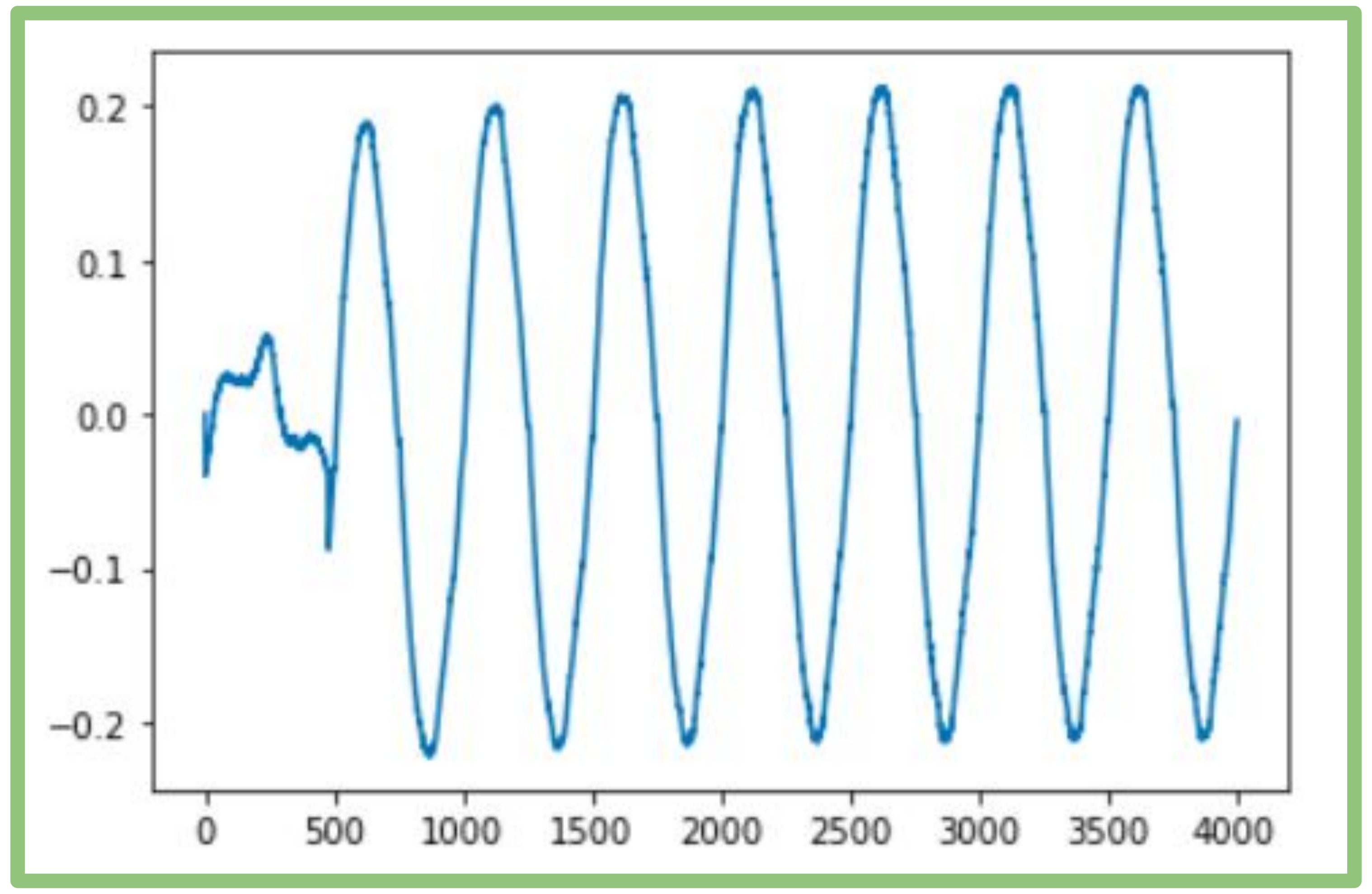
Data anomalies

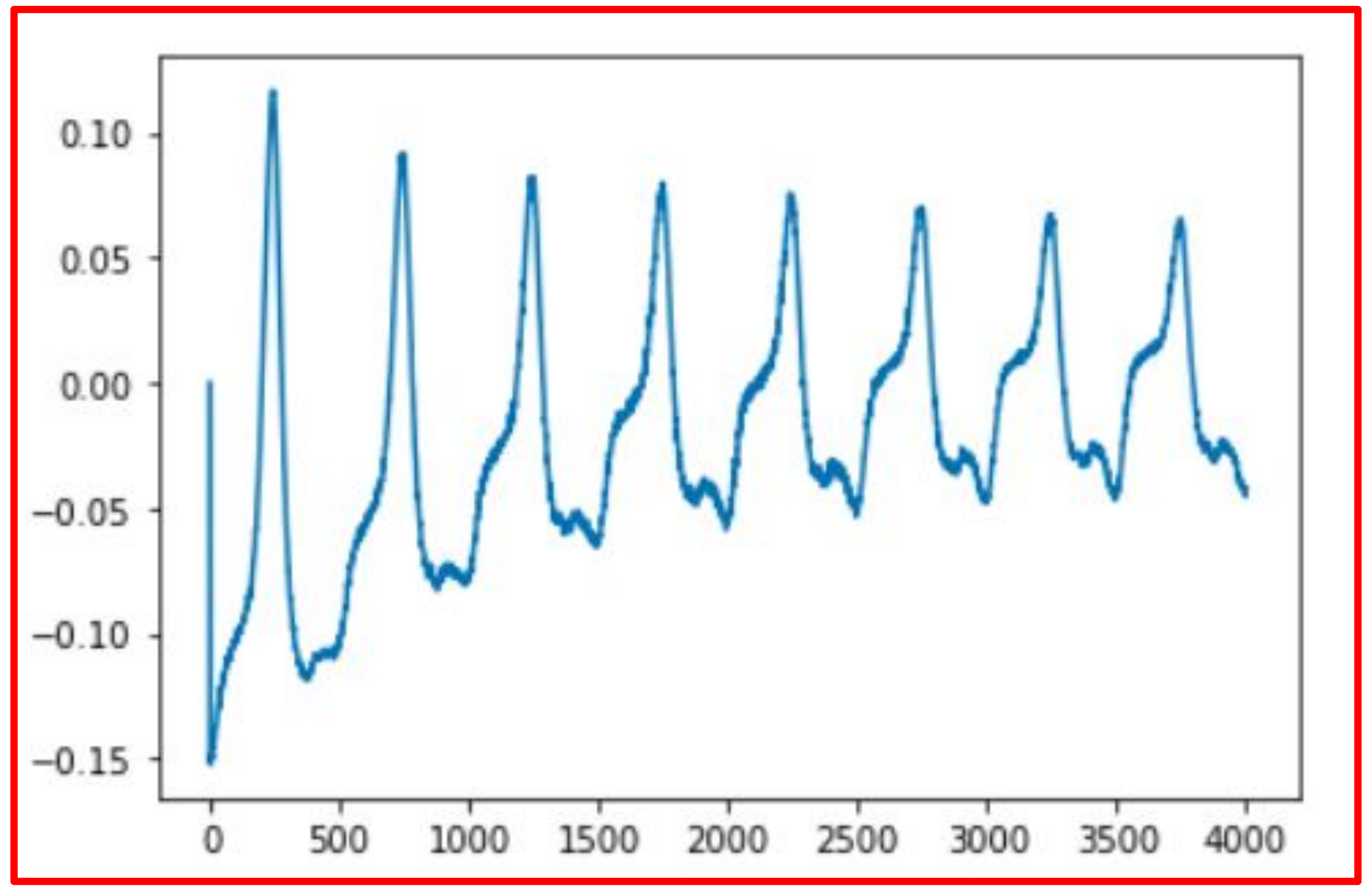












Anomalies of separated aggregated or sub-metered signals. A proper signal should have a high amplitude and preserve periodicity over most of the periods. Green color highlights signals that are fine for the models' training, whereas red color highlights the withdrawn signals.

Dataset size

	Real (PLAID)	Synthetic (generated)
Aggregate signals	27732	150000
Sub-metered signal	22581	150000

Number of separate signals extracted from the dataset PLAID with reald data and number of generated synthetic signals.

Medico R. et al., (2020), *Sci. Data*, 7(1)