Investigating the effects of non-uniform input data windowing on electrical load disaggregation performance

M.Sc. Thesis Seminar

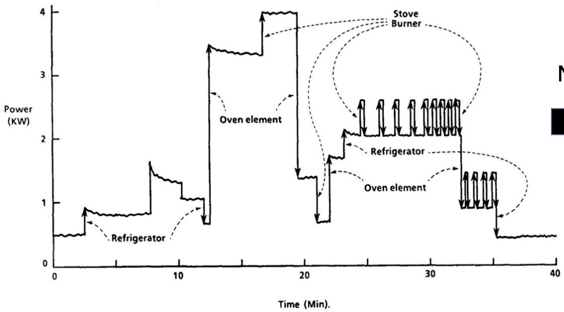
B.Sc. Abdul Hakmeh, Department of Informatics
March 2022

Agenda

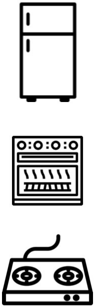
Non-intrusive Load Monitoring (NILM)

- Cost-effective centralized application of smart meters.
- Analyzing the overall electrical signal of a household and decomposing it into individual contribution of each appliance.
- Engaging the end-user in the energy management plan by providing detailed device-level consumption information.

Non-intrusive load monitoring (NILM)



NILM



500 W

1200 W

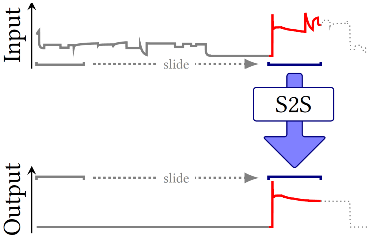
1500 W

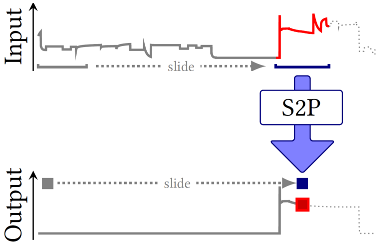
Source: Hart, G.W., Nonintrusive appliance load monitoring, 1992

TU Clausthal

State-of-the-art NILM

- Recently proposed disaggregation methods are rely on deep neural network models.
- Sequence-to-point (S2P) & sequence-to-sequence (S2S) models achieved remarkable accuracy results.





Source: Reinhardt, A. and Bouchur, M., On the impact of the sequence length on S2S and S2P learning, 2020

State-of-the-art NILM

- NILM toolkit (NILMTK) is an open source toolkit for NILM designed to enable the comparison of energy disaggregation algorithms in a reproducible manner.
- NILMTK offers reference implementations of S2P & S2S models.
- The most important configuration parameter is the window length defining the number of input samples to consider.
- Setting the parameter in most contributions can be categorized into two strategies; **one-for-all** and **per-device**.

State-of-the-art NILM

- **One-for-all**: a fixed value is assigned and only one network is trained to predict the operating conditions of all appliances.
- **Per-device**: the operational duration of the appliance under consideration is averaged. A separate network must be trained per targeted appliance.

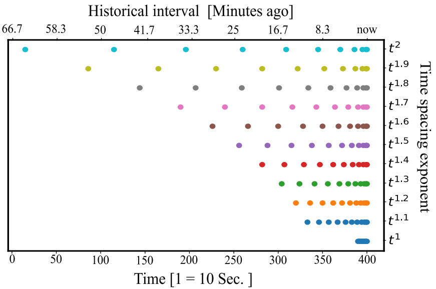
Motivation and Research Goal

- **One-for-all** and **per-device** strategies represent a trade-off between results accuracy and simplicity.
- Using the **one-for-all** strategy to maintain the simplicity while modifying the windowing technique to ensure improvement in load disaggregation performance.
- The goal is to investigate the effect of non-uniform windowing on load disaggregation without influencing the model complexity.

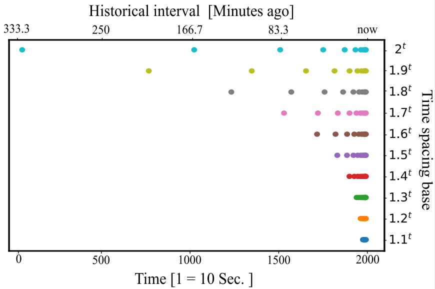
Methodology: Non-equidistant Temporal Sampling

- The core idea: splitting the traditional uniform window into two sequences according to a fraction: a high-resolution sequence and an unequally spaced sequence generated by means of non-linear algebraic functions.
- Using non-linear algebraic functions creates non-equidistant temporal space between samples.
- Allowing the input window to contain more historical information without having to adjust the window length parameter.

Time Spacing Functions



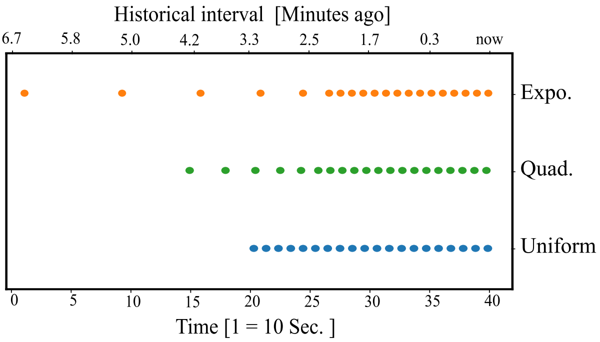
*Q*uad. function



*E*xpo. function

- Differences in:
  - Temporal distances.
  - Effective window length.
  - Samples distribution over the timeline.

Non-equidistant Temporal Sampling: Example



Windowing Parameterization

- Two parameters are associated with windowing operation.
- The spacing function value refers to the value of the exponent in the quadratic function or the base value of the exponential function.
- The history aggregation factor defines the division of the window into equidistant and non-equidistant sequences.
- An Example: given are a widow length $= 561$, spacing function value$= 2$ and history aggregation factor$= 0.05$, they create a window, that have 533 samples in full-resolution and 28 samples collected according to the function t^2.

Non-equidistant sampling variants: Inverse Function

- Proposing different variants for a fair evaluation.
- Instead of splitting the window into two parts to create a non-equidistant sampled sequence, the inverse function creates this part over the entire window in total length.
- No partitioning is used and data points are sampled according to the spacing function, resulting in a window with uneven spacing between samples.
- Targeting the entire window means setting the history aggregation factor to one.

Non-equidistant sampling variants: Linear Downsampling Function

- Linear downsampling function retains the technique of splitting the input window.
- Instead of non-equidistant temporal sampling, the sequence is performed in linear intervals, with samples collected linearly following a fixed step.
- The linear and inverse functions cover the same effective sequence length as the non-equidistant temporal sampling.

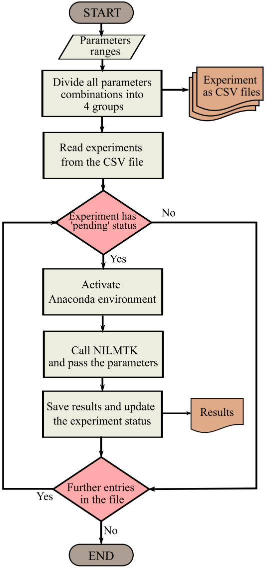
Experiment Design: Windowing Parameters Ranges

- Numerous experiments are performed to investigate the effect of the proposed methods on load disaggregation performance.
- NILMTK is used as an execution tool with setting the model configurations to the default values.
- The window length is set according to the best value (561) determined empirically in the literature on the Dutch Residential Energy Dataset (DRED).

Parameter	Value /range	Interval
Algorithm	[S2S , S2P]	-
Time spacing function	[non-equidistant, linear, inverse]	-
spacing function value	$\{0.1, \ldots, 2\}$	0.1
history aggregation factor	$\{0.003, \ldots, 0.15\}$	0.002

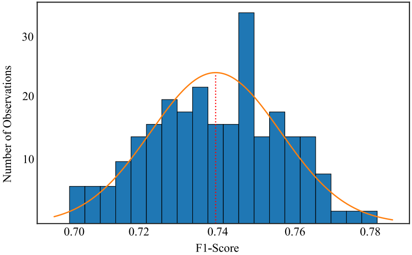
Backup slide: Execution System

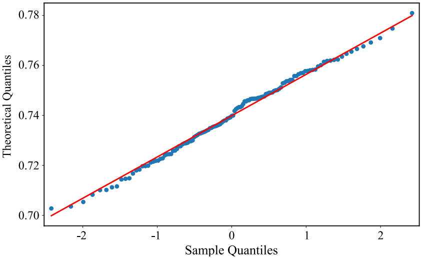
- Running a large number of experiments in parallel on multi GPU-computers.
- The system designed to automate the experiments execution.
- Automation is responsible for robustness to unexpected terminations or interruptions.



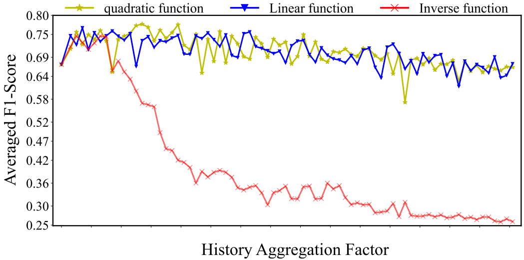
Results Margin

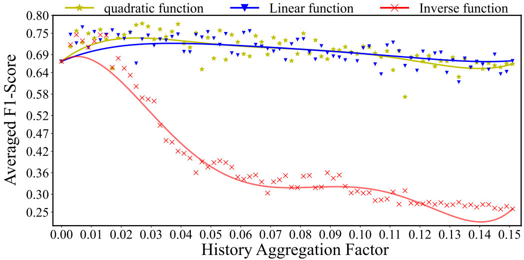
- Due to fluctuations observed in F1-score, an analysis is performed to define a results margin.
- After the margin value, the difference can be interpreted with certainty as a difference in performance caused by the function being evaluated.
- Normal distribution curve fitted to the histogram of 100 samples reports standard deviation in value of 0.01523.



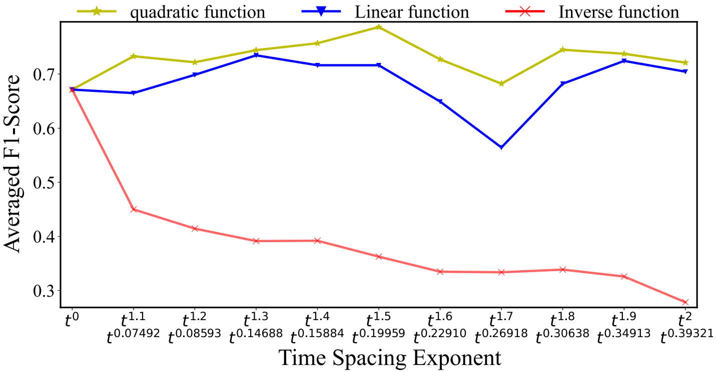


S2S Results: Quadratic Function $t^{1.2}$

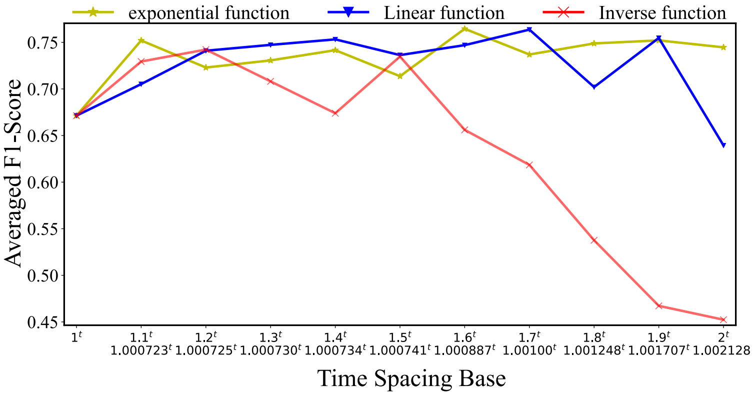




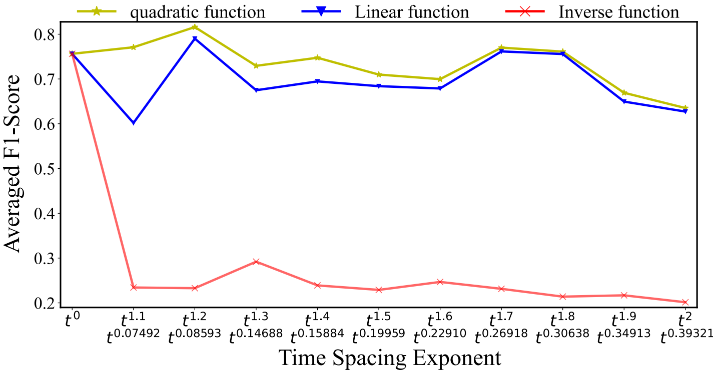
S2S Results: Quadratic Function Summary



S2S Results: Exponential Function Summary



S2P Results: Quadratic Function Summary



Conclusion

- Analyzing the results confirms a positive impact of aggregating historical information on load disaggregation.
- The promising potential of non-equidistant sampling demonstrated during the evaluation can make a step forward in the field of NILM, as collecting historical input data in a quadratic or linear fashion ensures performance improvement without compromising model complexity.

Future Work

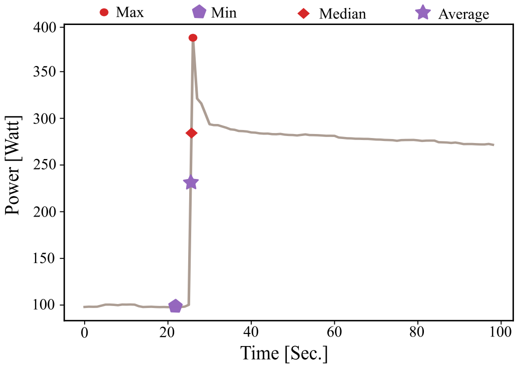
- **Short-term research** the effect of non-uniform windowing on a different dataset can be investigated, where the training and testing phases can be performed on different buildings.
- **Long-term research** since the window length used in this study has been previously defined empirically in the literature through experiments on DRED, a study can be conducted that describes a procedure for theoretically defining the appropriate window length when using a different energy dataset.

Thank you for your attention.

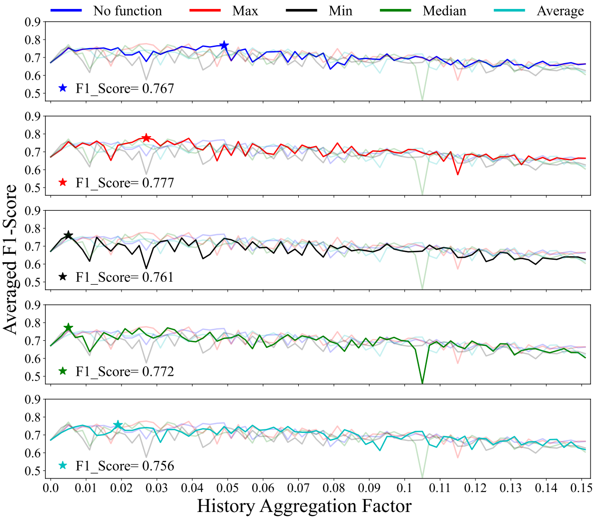
Backup slid: Data Sampling Functions

- Non-equidistant temporal sampling generates temporal gaps between sampled data points.
- It may lead to lack of information that represent the events and appliance activations that took place in this interval.
- Statistical functions are used to increase the representation of a single sample for the temporal interval generated, giving a summary of the data without listing each value.

Backup slid:: Data Sampling Functions

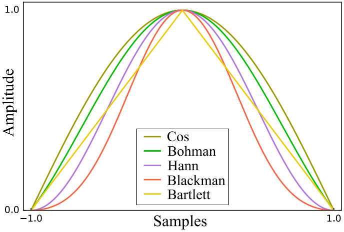


Backup slide: Sampling Functions Comparison

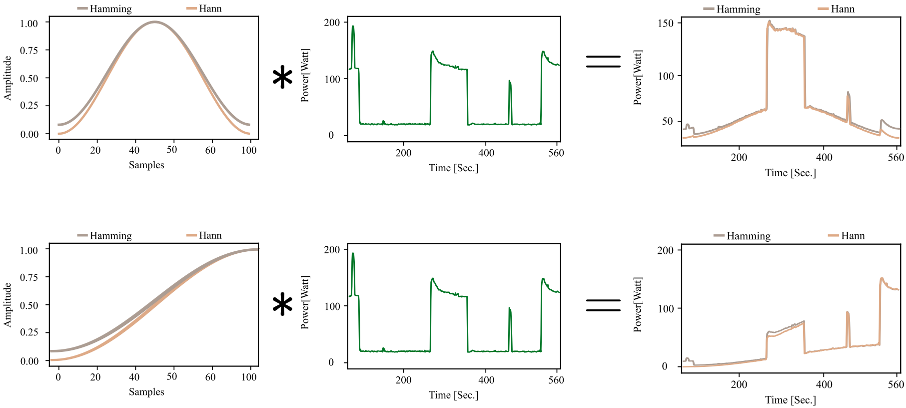


Preliminary Study: Window Functions

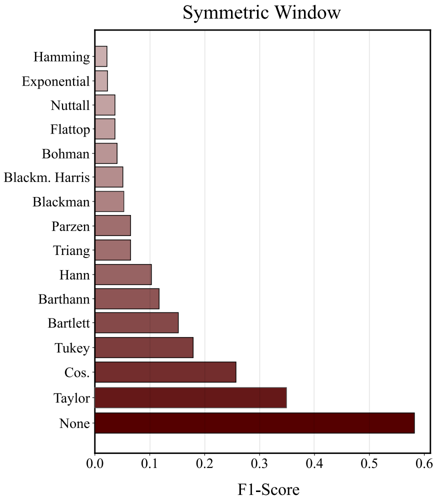
- Window functions are sequences of finite length.
- Their amplitude goes uniformly towards zero at both edges.
- They are used to merge discontinued waveform in signal processing, resulting in a continuous signal.
- They can be applied on the data input window to fade out data points on the sides reducing their effect on the learning model.
- Symmetric and Asymmetric windows are applied to target different part of the input window.

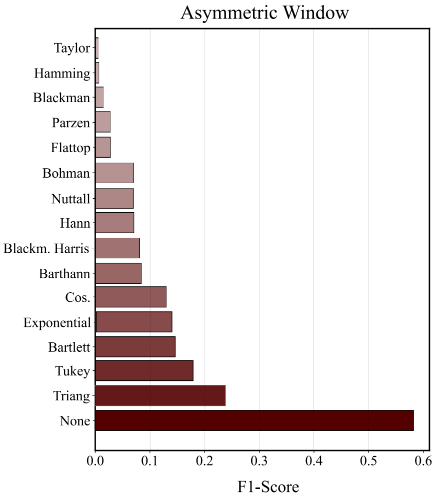


Backup slide

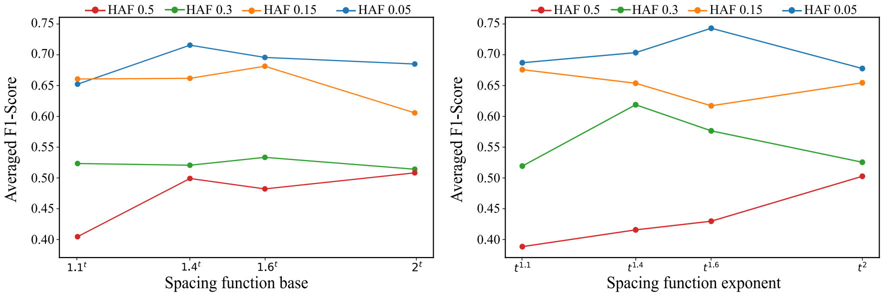


Backup slide





Backup Slide



hyperparameters S2S

(1) Input sequence with length W: Yt :t+W1

(2) 1D Convolution: filters: 30; filter size: 10

(3) 1D Convolution: filters: 30; filter size: 8

(4) 1D Convolution: filters: 40; filter size: 6

(5) 1D Convolution: filters: 50; filter size: 5

(6) 1D Convolution: filters: 50; filter size: 5

(7) Fully connected: units: 1024

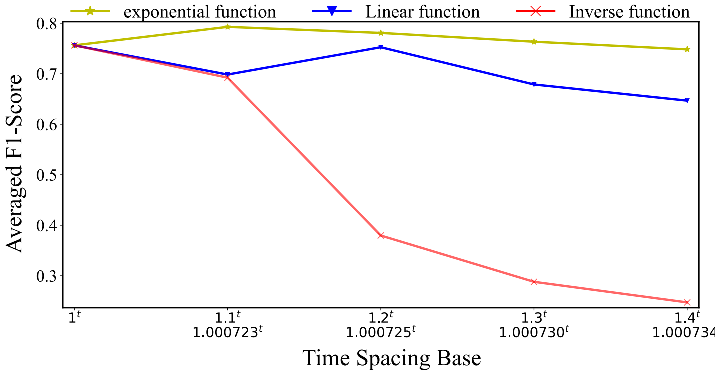
(8) Output: Number of units:W

Backup Slide

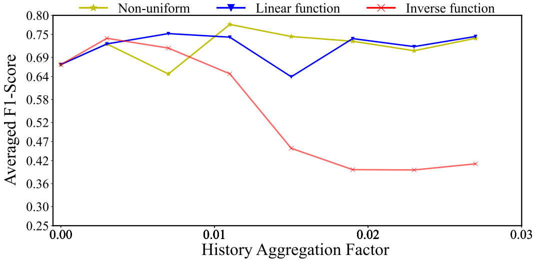
Tabelle: Data preprocessing and model parameters

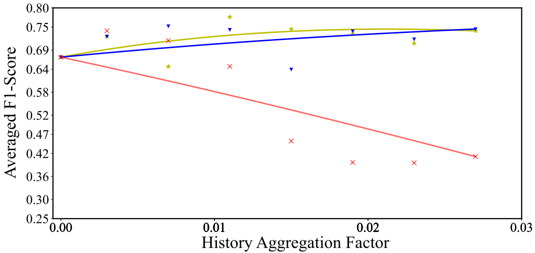
Parameter	value
σ	600
$\overline{x}$	1800
reduction factor	10
sequence length	561
Epochs	30
Batch size	512
Stride	1

Backup Slide: S2P expo.



Backup slide: Exponential function results





Backup Slide: Mathematical formulation

$$ESL_{quad} = \left(UPL + \sum_{t=1} t^{SFV} \right) * DF \tag{1}$$

$$ESL_{expo} = \left(UPL + \sum_{t=1} SFV^{t} \right) * DF \tag{2}$$

wehre:

ESL = effective sequence length
UPL = length of the uniformed part
DF = data downsampling factor
SFV = spacing function value

Backup slide: Mathematical formulation

$$ESL_{uniform} = \left(UPL + \sum_{t=1} (t + FTS) \right) * DF \qquad (3)$$

wehre:

FTS = fixed time step
ESL = effective sequence length
UPL = length of the uniformed part
DF = data downsampling factor
SFV = spacing function value

Backup slide: Execution System

- Running a large number of experiments in parallel on multi GPU-computers.
- The system designed to automate the experiments execution.
- Automation is responsible for robustness to unexpected terminations or interruptions.