On the Benefits of Water Consumption and Reactive Power for Eventless NILM

Justus Breyer
RWTH Aachen University
Aachen, Germany
justus.breyer@comsys.rwth-aachen.de

Paul Sonnenschein
RWTH Aachen University
Aachen, Germany
paul.sonnenschein@rwth-aachen.de

Muhammad Hamad Alizai
LUMS
Lahore, Pakistan
hamad.alizai@lums.edu.pk

Klaus Wehrle
RWTH Aachen University
Aachen, Germany
wehrle@comsys.rwth-aachen.de

Abstract

Non-Intrusive Load Monitoring (NILM) has been an active area of research for the past two decades, with the goal of inferring the consumption of individual electrical devices in a building from mains meter data. Despite substantial progress, the potential of facilitating disaggregation by incorporating additional sensor data has received limited attention. In this paper, we investigate the benefits of integrating water consumption data into eventless NILM models and compare its impact with that of reactive power. Our experiments show that appliances directly associated with water usage benefit from the inclusion of water data; however, the overall gains from reactive power are larger. The best performance is achieved by combining reactive power and water data as additional input channels alongside active power, resulting in reductions of 28% in mean absolute error and 63% in signal aggregate error in the geometric mean across all circuits.

CCS Concepts

• **Computer systems organization → Sensor networks**; • **Hardware → Energy metering**; *Smart grid*; *Sensor applications and deployments.*

Keywords

Energy disaggregation, load monitoring, non-intrusive load monitoring, algorithm performance evaluation

ACM Reference Format:
Justus Breyer, Paul Sonnenschein, Muhammad Hamad Alizai, and Klaus Wehrle. 2026. On the Benefits of Water Consumption and Reactive Power for Eventless NILM. In *ACM Sustainability Week 2026 (ACM Sustainability Week Companion '26), June 22–25, 2026, Banff, AB, Canada.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3765611.3815510

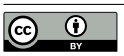
1 Introduction

The escalating impact of climate change has underscored the urgent need to reduce greenhouse gas emissions, making it one of the defining global challenges of our time [2]. Among other objectives, Non-Intrusive Load Monitoring (NILM) supports progress toward sustainable development by enabling energy-saving strategies [6, 8] through decomposing a building's total electricity consumption into appliance-level consumption [1, 11] using machine learning (ML) models. A key advantage of NILM is that it relies on only a single measurement point for electricity monitoring, which substantially reduces the installation, networking, and maintenance overhead compared to fully intrusive setups [29, 32]. Although NILM has primarily been studied in residential settings, it still faces several unresolved challenges that limit its deployment readiness despite decades of research [16].

Recent advances in NILM architectures for direct disaggregation of individual devices from aggregate mains readings have placed strong emphasis on neural networks, particularly deep learning [1]. A wide range of models, including convolutional, recurrent, and transformer-based architectures, has been proposed to reduce disaggregation error (e.g., [37, 38]). In contrast, the use of additional sensor data that may correlate with device activity has remained comparatively underexplored.

Intuitively, whole-building water consumption offers several properties that may be useful for improving energy disaggregation. First, water-related activity in households is often sparse, and simultaneous usage events occur relatively infrequently, resulting in clearer patterns. Second, water consumption can serve as an indicator of household activities such as cooking or showering, which may in turn correlate with the electricity use of devices that do not directly consume water [13]. Motivated by this intuition, prior work has examined the value of including water data in electricity disaggregation. However, these efforts have been limited to event-based [17] or device-specific [21] procedures. As a result, it remains unclear whether eventless deep learning models can directly exploit water consumption signals in a simple and architecture-agnostic manner, and how the value of such information compares with that of other auxiliary signals such as reactive power [33, 34]. To address this gap, this paper makes the following contributions:

- We introduce a general and lightweight approach for enriching eventless NILM with auxiliary sensing signals by treating them as additional input channels, allowing their integration without redesigning the core deep learning architecture.
- We provide a systematic empirical analysis of this design on real-world public data, including ablation studies that



ACM Sustainability Week Companion '26, Banff, AB, Canada

ACM ISBN 979-8-4007-2199-1/26/06
https://doi.org/10.1145/3765611.3815510

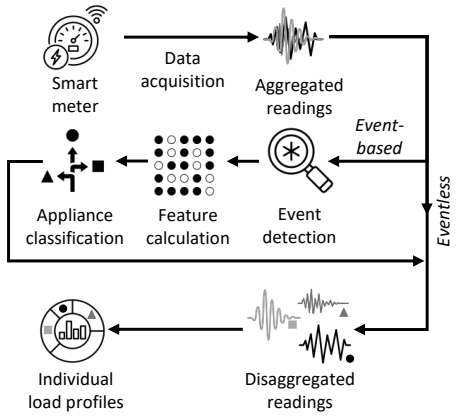


Figure 1: General architecture of NILM setups. Eventless approaches perform direct disaggregation, while event-based variants first search for individual features of appliances. The goal is the generation of per-appliance load profiles.

examine the impact of key architectural and training choices on model behavior.

- We establish the role of auxiliary modalities in eventless NILM by showing that water data offers targeted benefits for water-related loads, while reactive power contributes more consistently across circuits, and that their combination provides the strongest overall results.

2 Technical Background

To deploy a fully functional NILM setup, multiple steps need to be considered, as shown in Figure 1. There are two main approaches to the data disaggregation problem: *event-based* and *eventless* NILM [7], which differ in the required components.

Data collection is the first component to consider for *both* approaches. This data can be collected from a variety of sources, such as smart meters, power strips, or even the electrical wiring in a home. Before deploying NILM models, data collection supports model training; during deployment, it is necessary to capture information about the environment to perform energy disaggregation.

Event detection is the second component of an *event-based* NILM setup, where changes in an appliance's state, such as power-on and power-off, are usually identified through changes in overall power consumption. Event detection algorithms thereby rely on the switch continuity principle (SCP) [11], which states that, given a small enough time interval, only one appliance changes its state. Hence, these algorithms require a minimum data sampling rate to achieve correct results, as the SCP does not hold for time intervals of multiple seconds [22].

Feature calculation is the third component of an *event-based* NILM setup. Once an event has been detected, aggregated data from an area (window) around the event is used to calculate different features [4], which help distinguish between appliances or appliance states. These features can rely on preceding data transformations, e.g., Fourier or wavelet transforms, and have been thoroughly surveyed and analyzed [14, 15, 20, 31].

Classification is the fourth component of an *event-based* NILM setup. The calculated features of the event are fed into ML models for appliance or appliance state classification. The ML models can be simpler, e.g., random forests (RFs) or support vector machines (SVMs) [36], or more complex, e.g., neural networks (NNs) [30].

Energy disaggregation is the final component of an NILM setup. After event classification in the *event-based* approach, events of the same device must be matched to subsequently estimate the consumed power. With an *eventless* algorithm, the estimation of a device's consumed power is instead performed directly on the collected data, skipping event detection and classification, and typically also feature calculation. Over the years, eventless NILM has shifted from using Hidden Markov Models [19] toward NNs, in particular deep learning [37, 38].

Designing a NILM setup is a complex task, requiring careful consideration of various factors to achieve optimal performance.

2.1 Related Work

NILM as a field has been studied extensively and surveyed from multiple perspectives (e.g., [1, 7]). Accordingly, we focus here on work most relevant to our study, namely the combination of water and electricity data to improve NILM and water activity classification. Since both electricity and water consumption in residential environments are closely tied to human activity [13], the interaction between these two modalities has attracted growing interest, and our work builds on this emerging line of research.

The use of disaggregated energy readings to improve performance in water monitoring was proposed over a decade ago [9, 25, 35] and has shown promising results. More broadly, water monitoring approaches inspired by NILM have also been reviewed by Carboni et al. [5]. However, these approaches typically rely on water data alone and do not incorporate electricity measurements.

Jiang et al. [13] conducted a general correlation analysis suggesting that water consumption patterns can be used to infer the activity of electrical devices. However, their study relied on Bayesian networks, which are relatively uncommon in NILM, and their evaluation was based on simulated data. Liu et al. [21] proposed a multi-feature fusion approach in which a prior correlation analysis of water and gas data determines whether these additional inputs should be incorporated into a deep learning model through separate input branches. Although their method achieved strong results on real-world data, it required multiple processing steps and device-specific architectural adjustments. Keramati et al. [17] proposed an iterative approach that incorporates water data into an LSTM model to distinguish between devices with similar power levels. While their results were encouraging, the overall system design remained grounded in event-based NILM, as shown in Figure 1.

Prior work provides encouraging evidence that auxiliary sensing modalities can improve disaggregation performance. At the same time, an important gap remains. Existing studies on water-enhanced NILM have either focused on event-based formulations [17] or relied on device-specific, multi-stage architectures [21]. By contrast, the extent to which eventless deep learning models can directly exploit water consumption signals, without a separate correlation analysis or architecture tailored to individual devices, has not been systematically investigated to the best of our knowledge. Studies on

the positive impact of reactive power (e.g., [33, 34]) did not compare the improvements to other auxiliary data such as water.

3 Methodology

In order to investigate the potential of deep learning eventless NILM models to automatically extract correlation information of water use for direct energy disaggregation, we first developed a methodology to integrate the additional data into the models.

3.1 Design Principles

To ensure that differences in performances between models using only the electricity data versus models using both electricity and auxiliary data (water and/or reactive power) can be directly attributed to the data source, we emphasized a general model design oblivious of the device or circuit in question. Thus, the change in model performance cannot be assigned to a difference in model architecture, as could be the case with a specialized model design requiring a preceding analysis step as in previous approaches [21]. This design also enables a granular analysis of differences in improvement potentials regarding specific device types.

As second principle, we aimed at integrating the potential additional data sources at the edge of the network, such that the core architecture is not influenced by the additional input (besides a potential change in outcome of parameter tuning). This approach enables an easy transfer of the methodology to a variety of different deep learning architectures without the need of manual redesign. Related approaches either integrated the water data at a late point in an iterative decision procedure [17] or designed a completely new network branch that was merged into the electricity branch in a deeper layer [21].

3.2 Integration of Data Sources

When integrating additional input data into eventless NILM, we considered several options. Firstly, one could concatenate water and mains data (either alternating the input streams per sampling step or per complete window) but this would lead to an unintuitive model architecture and substantially increase complexity, as the doubled (or, for simultaneous sources of water and reactive power, tripled) input dimension would propagate downstream into the deep learning model. Simply replacing part of the input values with water and/or reactive power data would avoid increased complexity; however, it would also reduce the information from each input stream (decreasing it to $1/n$ for n input sources).

To circumvent these issues, we take inspiration from image recognition. In image recognition, networks receive 2-dimensional input via multiple overlaying channels (e.g., R, G, B) [26], which preserves the spatial context and keeps model complexity low. Only the first layer scales with the number of input channels, while deeper layers remain unaffected: adding channels increases the filter depth, but does not change the number of filter applications, since the input length and stride remain the same. As a result, the number of outputs produced by the first layer is independent of the number of channels, preserving the scale and architecture of subsequent layers. We therefore incorporate the time series of the mains reading and any additional input, be it water or reactive power or both, as separate input channels, analogous to processing a 1D image. This approach is simple yet effective, since it keeps all deeper layers intact and can be applied to any deep learning architecture, not just convolutional ones.

4 Implementation

With the general approach of integrating additional sensor data established, we detail below our proof-of-concept implementation, including the dataset selection, preprocessing, model architecture, and training procedure.

4.1 Dataset

To ensure reproducibility and comparability, we define the following requirements for the evaluation dataset:

(1) The data must be publicly available and easily accessible, ideally used by related approaches (cf. Section 2.1) or by NILM work employing comparable deep learning models.
(2) In addition to whole-building active power at a sufficiently high sampling rate, whole-building water consumption must be available, preferably at a similar temporal resolution or in a format that permits realistic down-/upsampling.
(3) Ideally, whole-building reactive power is available to enable a comparative study between water and reactive power as additional input channels.
(4) A sufficient number of appliances must be submetered to provide ground truth, and the measurement duration must be long enough to yield diverse training and evaluation samples.

Among commonly used public NILM datasets, only AMPds [23] and iAWE [3] were close to meeting these criteria. Other datasets were either collected in non-residential environments (e.g., Unicon [24]) or include gas but not water data (e.g., IDEAL [27]). The water signal in iAWE is recorded at one pulse per 10 l of water consumption, which we consider too coarse for realistic upsampling.

Consequently, we use AMPds as our primary dataset, largely due to the lack of alternative public datasets that include both electricity and water at suitable resolutions. This single-dataset choice is consistent with the closest related works [17, 21], which also rely on AMPds for similar reasons. AMPds contains water, electricity, and gas measurements as well as multiple electricity submeters (various circuit breakers) for a single Canadian household over 2012–2014, providing a long observation period. The electricity meters record multiple quantities, including active power (P) and reactive power (Q), which are required for our evaluation. While many submeters correspond to room-level circuits, AMPds also includes meters for individual appliances (e.g., clothes washer).

Water consumption is measured by a central water meter. During the recording period, the water sensor was replaced (initially pulsing every gallon, later pulsing every 0.5 l). For the subsequent period, the water meter (among other channels) is available in pulses per minute, matching the $1\,min^{-1}$ sampling rate of the electricity meters.

4.2 Data Processing

To maximize the utility of AMPds while maintaining a consistent sensor setup, we drop the initial period recorded with the first water sensor (April 1 to July 14, 2012) and retain only data recorded after the meter replacement (until March 31, 2014). From the remaining

data, we set aside 20% as a held-out *test* set for the final evaluation; this split is not used during hyperparameter optimization. The remaining data is split into *training* and *validation* sets for model selection and hyperparameter tuning.

As is common in eventless NILM with deep learning models, we apply input scaling. Rather than performing dataset-wide standardization, we use a fixed unit scaling: P is expressed in units of $1\,kW$, Q in units of $100\,VAR$, and water in units of $1\,L\,min^{-1}$. We additionally compare this unit scaling to the use of an explicit normalization layer (cf. Section 5.2) and observe no significant differences in model performance, however, our choice benefits from being generalized and device-independent.

4.3 Model Architecture

Our model is based on the Seq2Point paradigm: a convolutional neural network (CNN) predicts the target appliance's power consumption at the midpoint of an input window. Seq2Point models have been shown to outperform sequence-to-sequence variants in NILM settings [38] and have since been adopted broadly (e.g., [28]).

We initialized our implementation following Zhang et al. [38]; however, in preliminary experiments the original configuration overfit on AMPds. We therefore simplified the network and then incrementally increased depth/width until additional complexity no longer yielded relevant improvements. The resulting architecture uses four convolutional layers (instead of five in [38]), followed by two instead of a single dense layer.

The input length is reduced from 599 samples to 121 samples. In the original Seq2Point setting, inputs of length 599 were used on $1\,Hz$ and $1/6\,Hz$ data; AMPds is sampled at $1\,min^{-1}$. Using 599 samples at $1\,min^{-1}$ would therefore provide a substantially longer temporal context, which can affect model behavior and comparability across sampling rates [28]. An input size of 121 samples is a compromise between maintaining a comparable temporal context and providing sufficient variability within the input window (roughly $1\,h$ in the original setting versus about $2\,h$ in our case).

Compared to the original model, we use batch normalization and spatial dropout (20%) after each convolutional layer to mitigate overfitting. We further use a kernel size (receptive field) of 3 to localize feature extraction. For the convolutional layers and the first fully connected layer, we use a sigmoid activation instead of ReLU, as it empirically improved training stability in our experiments; the impact of these and other choices is quantified in Section 5.2. The complete architecture is shown in Figure 2.

4.4 Training

For training and validation, the respective datasets are segmented into consecutive input windows of 121 measurements using a sliding window with stride 1. From the second epoch onward, we apply a data augmentation scheme that reduces memorization by perturbing the circuit composition: pseudorandomly selected circuits are replaced with the same circuit's measurements from a different pseudorandom time point. A fixed seed is used to ensure determinism and reproducibility.

Training windows are shuffled using a shuffling buffer of size 10,000 and split into batches of size 256. Before the first convolutional layer, Gaussian noise with standard deviations of $5\,W$, $0.5\,VAR$, and $0.005\,L\,min^{-1}$ is added to P, Q, and water, respectively, i.e., a noise level of 0.5%, to further reduce overfitting.

Weights are initialized with Glorot initialization [10] and optimized using Adam [18] with a learning-rate schedule. We use mean squared error (MSE) as the training loss. The impact of training-related decisions is analyzed in Section 5.2.

5 Evaluation

The implementation of our proof-of-concept was evaluated on the AMPds dataset as detailed in Section 4. To contextualize our results, we firstly provide an overview of the relevant metrics that were used, before diving into a detailed analysis of design decisions via a variety of ablation studies. Finally, the benefits of adding water, reactive power, or both sources of additional information to the models are analyzed and discussed.

5.1 Metrics

In order to compare our results against the broad field of NILM literature in general and the closest related studies in particular, we made use of several established metrics. The chosen metrics are the Mean Absolute Error (MAE) and the Signal Aggregate Error (SAE), which have also been used by Liu et al. [21] (cf. Section 2.1). Keramati et al. [17] on the other hand performed event-based NILM, which is typically evaluated as a classification instead of a regression problem and therefore uses different error metrics that are not trivially transferable. Additionally, during training the MSE was used instead of the MAE, as analyzed in Section 5.2.

Assuming a ground truth appliance consumption of p_i and a predicted consumption of $\hat{p}_i$ at a given point in time i within the observation period of N total data points, the MAE, MSE and SAE are calculated as follows:

$$MAE = \frac{1}{N}\sum |p_i - \hat{p}_i|$$

$$MSE = \frac{1}{N}\sum (p_i - \hat{p}_i)^2$$

$$SAE = \frac{\sum p_i - \sum \hat{p}_i}{\sum p_i}$$

whereby a score closer to 0 is in all cases preferred.

5.2 Architecture Performance

Before evaluating the performance of models with different input data (P with and without Q and/or water), we first established that the baseline model architecture is reasonably well optimized for the task of NILM. Thus, various parameters of the Seq2Point architecture were switched out, extended or reduced to analyze their respective impact. The evaluation was mainly performed on the dishwasher as target appliance, using P and water consumption as input data, since evaluating all circuits on all input types for all ablations was computationally infeasible. The resulting performance of the respective architectural change was measured as the average MSE and MAE over the last fifty epochs, including the standard error, as depicted in Figure 3. Usually, these types of studies are called ablations, however, since the architecture was developed using an incremental rather than a decremental approach, starting

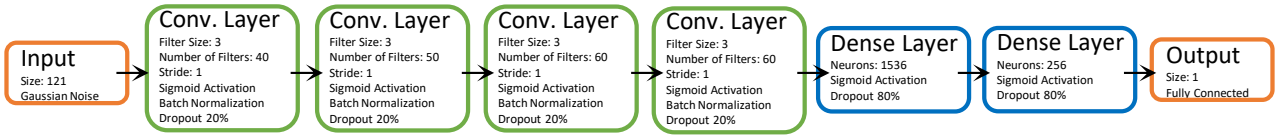


Figure 2: The architecture of the Seq2Point model, adapted from Zhang et al. [38].

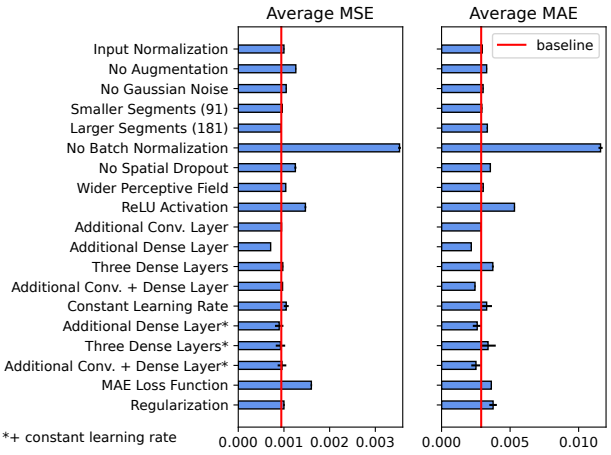


Figure 3: The impact of different architectural Seq2Point changes compared to the baseline of Section 4 with one instead of two dense layers, averaged over the performance of the final 50 epochs on the test set, including the standard error.

with a minimal amount of layers, our investigation does not fit the term in its technicality. Nonetheless, these findings serve as a rationale for certain design choices in Section 4.

5.2.1 Preprocessing. Some parameters affect the data that is used as model input for the first convolutional layers, these being the scaling, augmentation and the addition of Gaussian noise.

Common practice for deep learning architectures is to scale the model inputs to small intervalls, e.g., by using z-score normalization to scale to zero mean and a variance of one. However, even if data leakage of the test set is circumvented, this method still requires domain knowledge and a separate analysis step of some effort, and importantly does not generalize over the different devices. Instead, we opted for a simple rescaling using a fixed value for all devices (cf. Section 4). Exchanging the generalized scaling factor by z-score normalization did not result in a major impact, in fact, the models even performed slightly worse.

Data augmentation is typically used to create variety in the input-data in scenarios of sparsity to avoid overfitting. We found the augmentation to have a positive impact on the test set (as seen in Figure 3), however, the models performed worse during training, which is in line with our expectations. The addition of (Gaussian) noise follows an equal line of thought and, to a lesser extent, increases the model performance.

5.2.2 Input Segment Size. As touched upon in Section 4, the input size of our architecture is notably smaller than the 599 samples of Zhang et al.'s proposal [38], mainly due to a different sampling frequency of the input data. Generally, one can assume that an increased input size leads to a better performance since the model is provided with more context. However, the delay of the model is also increased, since the prediction of a point can only be performed, if the required subsequent measurement data is available. Hence, a reasonable tradeoff between model delay and performance should be aimed at for a practical deployment.

We tested a context of 91 and 181 minutes in comparison to the 121 minutes mentioned in Section 4. The smaller input size led to a slight decrease in performance, whereas the larger input size improved the MSE but showed significantly worse MAE on the test data. The latter may be a result of general variability, as the MAE loss is not accounted for during optimization. Overall, the context size of 121 minutes appeared to be a reasonable choice.

5.2.3 Model Architecture. Between each pair of subsequent convolutional layers, batch normalization and spatial dropouts were used in order to rescale the layer outputs and avoid overfitting. Foregoing the batch normalization has the highest impact of any changes to the model parameters, as clearly visible in Figure 3. Removing the dropout step afterwards had a lesser but still noticeable negative impact on the performance, showing that both steps are valuable to include in the model architecture. The latter finding contradicts the claims of the publication introducing batch normalization [12], which argues that dropout negatively impacts model performance when employing batch normalization.

One major difference to the initial proposal of the Seq2Point model is the size of the perceptive field in the convolutional layers. The final design used a size of three (cf. Figure 2), which is the least value possible for a symmetric perceptive field that is not simply the identity. However, testing for a larger size of five yielded a drop in performance, suggesting that most relevant features within the input series are found in very small regions of interest.

A surprising observation was made when exchanging the very common ReLU activation function with the Sigmoid function: Although ReLU is nowadays far more popular to be used in deep learning architectures, the Sigmoid activation drastically outperformed ReLU, reducing the error by more than 30% regardless of the metric. Although a reasonable explanation for this behavior could not be found, we consequently stuck to the Sigmoid activation.

5.2.4 Model Size. Since we built the model by iteratively increasing the number of layers until no further improvements could be established, we present the changes in performance when adding instead of removing further layers: An additional convolutional layer did not show a significant change in performance, however, adding a second dense layer at the end (opposed to Zhang et al.'s initial model [38]) showed a major benefit, as can be seen in Figure 3. Interestingly, the model was more capable of learning the features

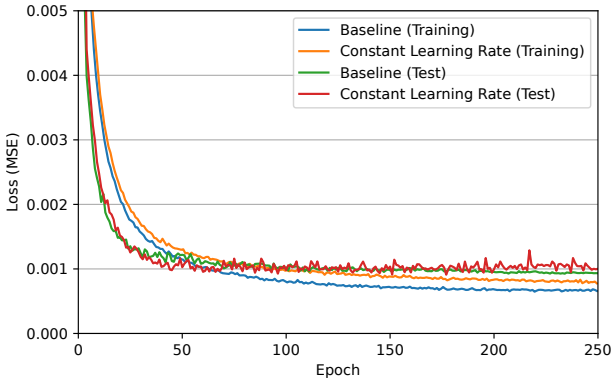


Figure 4: The effect of using a decaying learning rate during training. Notably, using a decaying learning rate leads to a worse model performance on the training data but an improved model performance of the test set, while additionally reducing the inter-epoch fluctuations.

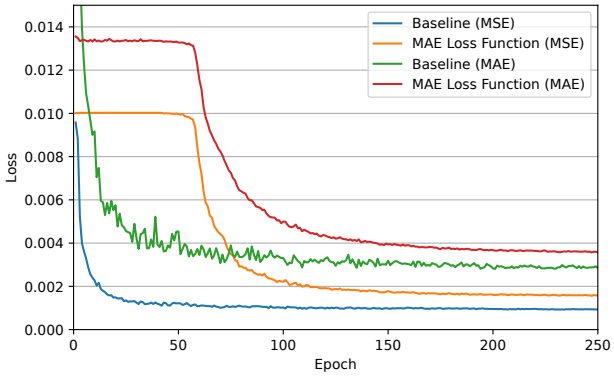


Figure 5: The effect of using MAE as loss function during training. Notably, using a MAE loss leads to a plateau during the first epochs, from which the model is unable to recover.

of the input series by first processing a number of convolutions with a small perceptive field, before finally using several iterations of dense layers to combine these features, instead of using convolutions of larger size. Adding a third dense layer did however show worse performance, similarly to the addition of another convolutional layer when using two dense layers. Consequently, we settled for four convolutional layers and two dense layers (cf. Figure 2).

As an exemplary check for the generalizability to other input types and target circuits, we compared the additional dense layer for all 80 combinations of input and target (cf. Table 1 in the Appendix). The geometric mean of the SAE improved by 2.4%, whilst the geometric mean of MAE did not change.

5.2.5 Optimization. The final category of changes are not to the model itself but the various different possibilities to modify the training process of the model instead. These include the learning rate, the loss function and regularization.

An uncommon decision is the use of a decaying learning rate in combination with the Adam optimizer. This measure was taken to reduce the noticeable fluctuations of model performance on the test set between the epochs late during the training. For this purpose, the learning rate was reduced iteratively in order to allow the model to stabilize on good parameters. As can be seen in Figure 4, not only did this measure stabilize the prediction error but it reduced the overall error as well, however, with a decrease of performance in the training data. Further, as the constant learning rate shortly outperforms the decaying learning rate early on around epoch 50, it is possible that the decay itself can be improved as it might take effect too quickly. As depicted in Figure 3, the general effect of the constant learning rate worsening performance is noticeable regardless of the number of layers in the Seq2Point model.

The choice of MSE instead of MAE as a loss function for optimization was less intuitive: Optimizing for MAE if it is one of the major evaluation metrics later on to compare related approaches against suggests itself. However, the pure use of MAE led to the loss converging to significantly worse values both in MAE and MSE. A closer investigation showed that the model using MAE was stuck at a comparatively high error for the first 60 epochs, which it was not quite capable to recover from, as shown in Figure 5. To counteract this behavior, we trained a model with MSE during the first epochs, before switching to MAE later on. When using this combination of loss functions, we observed the expected behavior of the MAE improving further than the baseline but the MSE suffering slightly. The final decision was made based on the SAE as a tiebreaker, which was better for the pure MSE training.

Lastly, regularization can force the model to conform to a metric of simplicity, while still using a complex architecture capable of modeling complex relationships. The general idea behind this is that a simpler model should be less likely to overfit and thus perform better on unseen data. However, as visible in Figure 3, the application of an L2 regularization term of strength 10^{-5} did result in worse performance on the test data. Although this could be coincidental, we settled for not including regularization.

5.3 Additional Input Evaluation

With a suitable Seq2Point architecture established in Section 5.2, we investigated the influence of additional input channels besides the active power P for the disaggregation task. For this purpose, all actual meters of the AMPds dataset [23] were used to train separate Seq2Point models, besides the aggregate data of the whole home (WHE), since the WHE was used as model input instead. The virtual meters of the main house (MHE) and the unmetered appliances (UNE) were excluded, as they were artificially created.

5.3.1 Baseline Performance. The baseline Seq2Point model using only P as input channel showed mixed results, significantly struggling for circuits measuring either a multitude of appliances or whole rooms. The rental unit (RSE) stands out with an MAE of about $86W$, other circuits with a high MAE include the basement (BME), home office (OFE), electronics workbench (EBE) and one of the bedrooms (B2E). The second type of circuit showing high error rates were related to temperature control: The furnace fan and HVAC (FRE), heat pump (HPE) and fridge (FGE) all are related to high energy draws. Finally, the model for the TV (TVE) showed some weaknesses in its MAE.

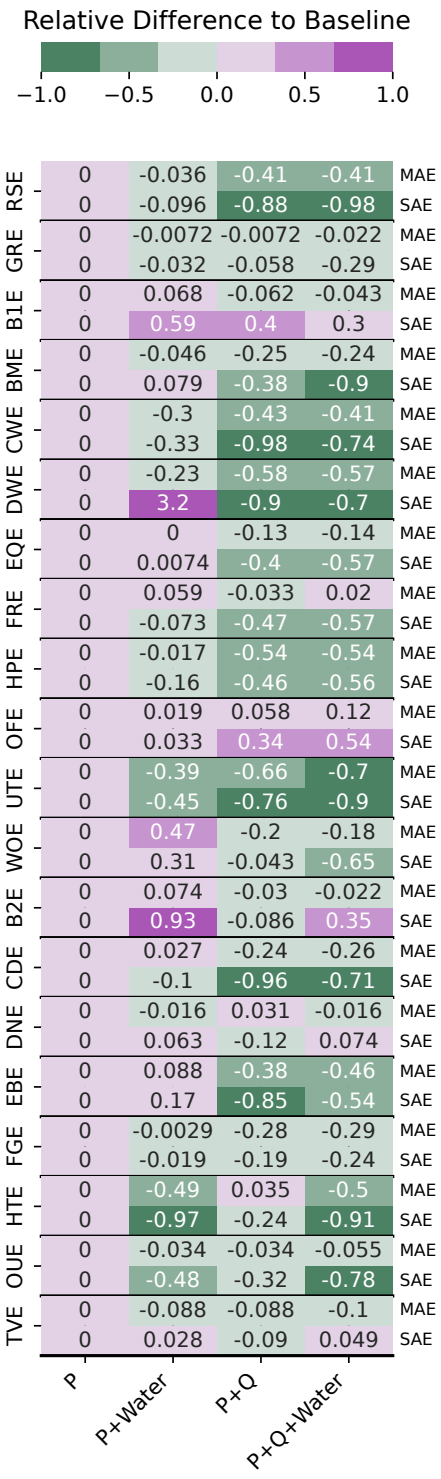


Figure 6: The impact of different input series on the performance of the Seq2Point model for all appliances and circuits of the AMPds [23] dataset. Depicted is the relative difference to the baseline performance using only Active Power (P) as input channel.

As many of the models with a larger MAE were tasked with predicting energy draw of appliances or circuits with high power demand, we normalized the MAE to investigate whether the offset of prediction was of significance in relation to the ground truth value. With an MAE normalized over the circuit's average power consumption, the temperature related circuits and TV all performed in the range of low-power consuming devices, with only the fridge still showing some weaknesses. The multi-appliance circuits still reside on the lower end of the performance spectrum, however, no longer being outliers. Instead, the other bedroom (B1E), the dining room (DNE) and the outside plugs (OUE) revealed themselves to be the most complicated circuits. A closer investigation showed that these three circuits specifically not only monitored multiple plugs but also their respective power consumption contributed less than 1% of the aggregate during the whole duration of AMPds. The sparsity and variety of data explain the struggles for the neural network to extract meaningful patterns.

The SAE generally confirmed these trends, with only the clothes washer (CWE) and the oven (WOE) showing some weaknesses in predicted total energy consumption. All numerical results, including the normalized MAE, are provided in Table 2 in the Appendix.

5.3.2 Integration of Water. Integrating the whole house water consumption generally showed a slight positive trend in model performance. However, as can be seen in Figure 6, the impact varied across the different types of circuits: Whilst many circuits showed only marginal improvements or even slightly worse performance, the water data influenced only a couple of models to a larger extent. The instant hot water unit (HTE), the clothes washer (CWE) and the dishwasher (DWE) improved by 49%, 30% and 23% in MAE, respectively, as would be expected with devices directly correlated with water usage. Similarly, the SAE of these devices could be drastically improved, with the exception of the dish washer.

Further, the error of the utility room plug (UTE) decreased by 39% and 45% (MAE and SAE, respectively), suggesting that the model is able to link water usage to household activity.

On the flip side, several models decreased in performance, with the most extreme outlier being the oven (WOE), followed by the SAEs of the bedrooms (B1E and B2E).

The overall results of adding water consumption to the Seq2Point models remain mixed, with a geometric mean improvement of MAE of just 6.5%. Given the vast differences in baseline performance, the geometric mean is preferred over an arithmetic mean to identify general tendencies.

5.3.3 Comparison to Reactive Power. An alternative option to include additional information into the NILM model is the use of reactive power (Q) [33, 34]. Generally, the functionality of a device influences its draw of reactive power, e.g., heating elements are often built as resistive components and characteristically exhibit a high draw of active and a low draw of reactive power. As some of the water related activities involve the use of heating elements (dishwasher, clothes washer, instant hot water unit), presumably parts of the water consumption can be correlated to the amount of reactive power measured at the home meter. On the other hand, the reactive power might even hold more benefits to the disaggregation task, since the capability of characterizing a device's functionality does translate to devices not correlated to the use of water, e.g., an

electric heat pump or oven will similarly exhibit low reactive and high active power draw. Further, measuring reactive power does not require advanced metering hardware but can be realistically performed by the already installed equipment.

As depicted in Figure 6, the inclusion of reactive power to the model does not only yield significant benefits to the baseline using just active power but also to the model version using active power and water. These findings confirm the intuition explained above, with almost all models benefiting in MAE besides the office (OFE), the dining room (DNE) and the hot water unit (HTE), where only slightly worse performance than the baseline was achieved. The SAE as a second metric however revealed more circuits to struggle, with the bedroom (B1E) and office (OFE) showing notably worse results than the baseline. On the flip side, the SAE of the hot water unit and dining room improved, contrary to the MAE. Interestingly, the hot water unit (HTE) showed significantly more profit when using water, whilst the dishwasher and clothes washer profited more from the inclusion of reactive power, despite all three devices being dependent on water use in general.

These latter findings, not quite confirming the intuition, led to a third variant of model, which included reactive power as well as the water consumption, in the hopes of drawing the best of both worlds. Indeed, the models using all three input channels increased the performance in terms of MAE across all circuits besides the office (OFE). In terms of SAE, the bedrooms (B1E and B2E), dining room (DNE) and TV (TVE) additionally showed weaker performance than the baseline. When comparing to using just one input channel, using all available information had no visible negative effect – in cases where one metric showed weaker results, generally the other one improved (e.g., for EBE). These tendencies are also reflected in the geometric means of the respective errors, whereby using all available input channels showed a mean improvement of 28% in MAE and 63% in SAE over the baseline.

5.3.4 Discussion. Our investigation of the benefits that additional sensor information can bring to the field of eventless NILM shows promising results: In line with the intuition, the inclusion of water data can help the models in correlating energy use with water related activity. However, the extent appears to be limited to mostly devices directly using water. Notably, the estimation of energy consumption of whole sub-circuits of the building, such as rooms, does generally not benefit from the additional information.

On the other hand, the inclusion of whole home reactive power shows more potential throughout our experiments, with the error generally decreasing more significantly. These findings stay in line with the intuition that reactive power offers information about the functionality of devices that are currently active. Combining water and reactive power on top of active power yields the best results. Crucially, the additional information does not lead to significant decrease in performance, with a few exceptions of multi-device circuits that are seldom the task of NILM algorithms, which try to infer the consumption of single devices. Hence, it appears that there is little drawback to including water and/or reactive power when tasked with predicting the consumption of single devices.

These findings generally confirm the analysis of related approaches: Keramati et al. [17] report an average increase of at least 4% in F-score of the classification task over related work not using water data. Liu et al. [21] report an average decrease of 14.5% in MAE and 43.9% in SAE over a baseline, when including gas and water data in a separate branch of per device designed deep learning architectures using correlation analyses. Our investigation showed a decrease in the geometric mean of 6.5, 24.8 and 28.2% in MAE and of 12.4, 59.1 and 63.4% in SAE when using water, reactive power or both as additional input, respectively. The reason for not directly comparing MAE and SAE scores of circuits and devices lies in Liu et al. performing their evaluation on only a subset of devices with a virtual aggregated electricity measurement, hence relying on different input for the disaggregation.

Timmermans and Henneaux [33] reported an improvement of 30 and 57% in MAE and SAE, respectively, when integrating reactive power to a Seq2Point CNN on a different dataset, staying closely in line with our averages of 24.8 and 59.1%. Valenti et al. [34] integrated reactive power to a denoising autoencoder, reporting only F-scores. Similarly to our findings, they concluded that reactive power tends to benefit the disaggregation process, however, to varying extent depending on the target circuit in AMPds.

Our presented analysis, although in-depth within its scope, does not yet allow for a generalized conclusion. The evaluation was performed on just one dataset, mostly due to a limited availability of public data including electricity, water and submetered ground truth data. Most public datasets lack in at least one of these departments, as the general research area of NILM does not depend on additional water measurements. The evaluation further was performed on just one – although established – deep learning architecture, leaving room to investigations on whether these suggested benefits hold across eventless NILM architectures.

6 Conclusion

The area of enriching NILM models with additional sensor data to increase performance has put little attention towards the potential that water consumption data holds. Given that water use is often correlated to certain household activities, we investigated the benefits that eventless Seq2Point models can gain by including whole home water consumption data without changing the internal architecture of the model. This impact of water data was compared to the benefits that the inclusion of reactive power, as an indicator of the functionality of the currently active devices, yields.

Our findings suggest that both additional meter information can lead to a reduction in regression error, however, the benefits for reactive power are higher. Including both types of data simultaneously into the NILM model on top of active power shows the best results, as the drawbacks of one type of information seldom cancel out the benefit of the other. We hence strongly recommend including reactive power into NILM models, as it does not require setting up an additional sensor, and using water data as an additional channel, if the additional overhead allows for it.

References

[1] Georgios-Fotios Angelis, Christos Timplalexis, Stelios Krinidis, Dimosthenis Ioannidis, and Dimitrios Tzovaras. 2022. NILM Applications: Literature Review of Learning Approaches, Recent Developments and Challenges. *Energy and Buildings* 261 (2022), 111951.

[2] Nigel W Arnell, Jason A Lowe, Andrew J Challinor, and Timothy J Osborn. 2019. Global and Regional Impacts of Climate Change at Different Levels of Global

Temperature Increase. *Climatic Change* 155 (2019), 377–391.
[3] Nipun Batra, Manoj Gulati, Amarjeet Singh, and Mani B Srivastava. 2013. It's Different: Insights into Home Energy Consumption in India. In *Proceedings of the 5th ACM Workshop on Embedded Systems for Energy-Efficient Buildings*. 1–8.
[4] Justus Breyer, Jonas Koerhuis, Muhammad Hamad Alizai, and Klaus Wehrle. 2025. Practical Insights from Implementing Event-Based NILM Systems. In *Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems*. 751–756.
[5] Davide Carboni, Alex Gluhak, Julie A McCann, and Thomas H Beach. 2016. Contextualising Water Use in Residential Settings: A Survey of Non-Intrusive Techniques and Approaches. *Sensors* 16, 5 (2016), 738.
[6] K. Carrie Armel, Abhay Gupta, Gireesh Shrimali, and Adrian Albert. 2013. Is Disaggregation the Holy Grail of Energy Efficiency? The Case of Electricity. *Energy Policy* 52 (2013), 213–234.
[7] Suryalok Dash and Nirod Chandra Sahoo. 2022. Electric Energy Disaggregation via Non-intrusive Load Monitoring: A State-of-the-Art Systematic Review. *Electric Power Systems Research* 213 (2022), 108673.
[8] Karen Ehrhardt-Martinez, Kat A Donnelly, Skip Laitner, et al. 2010. Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities. American Council for an Energy-Efficient Economy Washington, DC.
[9] Bradley Ellert, Stephen Makonin, and Fred Popowich. 2015. Appliance Water Disaggregation via Non-Intrusive Load Monitoring (NILM). In *International Summit, Smart City 360°*. Springer, 455–467.
[10] Xavier Glorot and Yoshua Bengio. 2010. Understanding the Difficulty of Training Deep Feedforward Neural Networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. JMLR Workshop and Conference Proceedings, 249–256.
[11] George William Hart. 1992. Nonintrusive Appliance Load Monitoring. *Proc. IEEE* 80, 12 (1992), 1870–1891.
[12] Sergey Ioffe and Christian Szegedy. 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In *International Conference on Machine Learning*. pmlr, 448–456.
[13] Wenqian Jiang, Xiaoming Lin, Zhou Yang, Jianlin Tang, Xiuqing Lin, and Mi Zhou. 2022. Correlation Analysis Between Multivariate External Data and Power Consumption Behaviors Based on Bayesian Network. In *EMIE 2022; The 2nd International Conference on Electronic Materials and Information Engineering*. VDE, 1–5.
[14] Matthias Kahl, Thomas Kriechbaumer, Anwar Ul Haq, and Hans-Arno Jacobsen. 2017. Appliance Classification Across Multiple High Frequency Energy Datasets. In *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 147–152.
[15] Matthias Kahl, Anwar Ul Haq, Thomas Kriechbaumer, and Hans-Arno Jacobsen. 2017. A Comprehensive Feature Study for Appliance Recognition on High Frequency Energy Data. In *Proceedings of the Eighth International Conference on Future Energy Systems*. 121–131.
[16] Maria Kaselimi, Eftychios Protopapadakis, Athanasios Voulodimos, Nikolaos Doulamis, and Anastasios Doulamis. 2022. Towards Trustworthy Energy Disaggregation: A Review of Challenges, Methods, and Perspectives for Non-Intrusive Load Monitoring. *Sensors* 22, 15 (2022), 5872.
[17] Mohammad Mehdi Keramati, Elnaz Azizi, Hamidreza Momeni, and Sadegh Bolouki. 2022. Incorporating Coincidental Water Data into Non-Intrusive Load Monitoring. *Sustainable Energy, Grids and Networks* 32 (2022), 100805.
[18] Diederik P Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980* (2014).
[19] Weicong Kong, Zhao Yang Dong, Jin Ma, David J Hill, Junhua Zhao, and Fengji Luo. 2016. An Extensible Approach for Non-Intrusive Load Disaggregation with Smart Meter Data. *IEEE Transactions on Smart Grid* 9, 4 (2016), 3362–3372.
[20] Jian Liang, Simon KK Ng, Gail Kendall, and John WM Cheng. 2009. Load Signature Study—Part I: Basic Concept, Structure, and Methodology. *IEEE Transactions on Power Delivery* 25, 2 (2009), 551–560.
[21] Hang Liu, Chunyang Liu, Lijun Tian, Haoran Zhao, and Junwei Liu. 2021. Non-Intrusive Load Disaggregation based on Deep Learning and Multi-feature Fusion. In *2021 3rd International Conference on Smart Power & Internet Energy Systems (SPIES)*. IEEE, 210–215.
[22] Stephen Makonin. 2016. Investigating the Switch Continuity Principle Assumed in Non-Intrusive Load Monitoring (NILM). In *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, 1–4.
[23] Stephen Makonin, Bradley Ellert, Ivan V Bajić, and Fred Popowich. 2016. Electricity, Water, and Natural Gas Consumption of a Residential House in Canada from 2012 to 2014. *Scientific Data* 3, 1 (2016), 160037.
[24] Harsha Moraliyage, Nishan Mills, Prabod Rathnayake, Daswin De Silva, and Andrew Jennings. 2022. Unicon: An Open Dataset of Electricity, Gas and Water Consumption in a Large Multi-campus University Setting. In *2022 15th International Conference on Human System Interaction (HSI)*. IEEE, 1–8.
[25] Khoi A Nguyen, Rodney A Stewart, and Hong Zhang. 2017. Water End-use Classification with Contemporaneous Water-Energy Data and Deep Learning Network. *International Journal of Computer and Systems Engineering* 12, 1 (2017), 1–6.
[26] Simon JD Prince. 2023. *Understanding Deep Learning*. MIT press, Chapter 10.4.
[27] Martin Pullinger, Jonathan Kilgour, Nigel Goddard, Niklas Berliner, Lynda Webb, Myroslava Dzikovska, Heather Lovell, Janek Mann, Charles Sutton, Janette Webb, et al. 2021. The IDEAL Household Energy Dataset, Electricity, Gas, Contextual Sensor Data and Survey Data for 255 UK Homes. *Scientific Data* 8, 1 (2021), 146.
[28] Andreas Reinhardt and Mazen Bouchur. 2020. On the Impact of the Sequence Length on Sequence-to-Sequence and Sequence-to-Point Learning for NILM. In *Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring*. 75–78.
[29] Antonio Ridi, Christophe Gisler, and Jean Hennebert. 2014. A Survey on Intrusive Load Monitoring for Appliance Recognition. In *2014 22nd International Conference on Pattern Recognition*. IEEE, 3702–3707.
[30] JG Roos, IE Lane, EC Botha, and Gerhard P Hancke. 1994. Using Neural Networks for Non-intrusive Monitoring of Industrial Electrical Loads. In *Conference Proceedings. 10th Anniversary. IMTC/94. Advanced Technologies in I & M. 1994 IEEE Instrumentation and Measurement Technolgy Conference (Cat. No. 94CH3424-9)*. IEEE, 1115–1118.
[31] Nasrin Sadeghianpourhamami, Joeri Ruyssinck, Dirk Deschrijver, Tom Dhaene, and Chris Develder. 2017. Comprehensive Feature Selection for Appliance Classification in NILM. *Energy and Buildings* 151 (2017), 98–106.
[32] Nagender Kumar Suryadevara and Gyan Ranjan Biswal. 2019. Smart Plugs: Paradigms and Applications in the Smart City-and-Smart Grid. *Energies* 12, 10 (2019), 1957.
[33] Jean-Luc Timmermans and Pierre Henneaux. 2024. Active and Reactive Power Sequences for Energy Disaggregation with Deep Learning Models. In *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*. IEEE, 1–5.
[34] Michele Valenti, Roberto Bonfigli, Emanuele Principi, and Stefano Squartini. 2018. Exploiting the Reactive Power in Deep Neural Models for Non-intrusive Load Monitoring. In *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.
[35] Scott Vitter and Michael Evan Webber. 2018. A Non-Intrusive Approach for Classifying Residential Water Events Using Coincident Electricity Data. *Environmental Modelling & Software* 100 (2018), 302–313.
[36] Benjamin Völker, Philipp M Scholl, and Bernd Becker. 2021. A Feature and Classifier Study for Appliance Event Classification. In *International Conference on Sustainable Energy for Smart Cities*. Springer, 99–116.
[37] Zhenrui Yue, Camilo Requena Witzig, Daniel Jorde, and Hans-Arno Jacobsen. 2020. BERT4NILM: A Bidirectional Transformer Model for Non-Intrusive Load Monitoring. In *Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring*. 89–93.
[38] Chaoyun Zhang, Mingjun Zhong, Zongzuo Wang, Nigel Goddard, and Charles Sutton. 2018. Sequence-to-Point Learning with Neural Networks for Non-Intrusive Load Monitoring. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

Table 1: Performance of Seq2Point Models With Differing Number of Dense Layers

Circuit	Description	Dense Layers	Baseline (P) MAE	Baseline (P) SAE	P + Water MAE	P + Water SAE	P + Q MAE	P + Q SAE	P + Q + Water MAE	P + Q + Water SAE
RSE	Rental Unit	1	89.7	0.0627	86.9	0.0614	**51**	**0.00369**	**50.1**	0.00523
		2	**86.8**	**0.0522**	**83.7**	**0.0472**	**51**	0.0065	51.1	**0.0013**
GRE	Garage	1	**1.37**	**0.102**	**1.38**	**0.138**	**1.37**	**0.114**	1.37	0.123
		2	1.39	0.154	**1.38**	0.149	1.38	0.145	**1.36**	**0.11**
B1E	Northern Bedroom	1	**1.58**	0.294	**1.32**	**0.179**	**1.35**	**0.0404**	**1.23**	**0.0868**
		2	1.61	**0.186**	1.72	0.296	1.51	0.26	1.54	0.242
BME	Basement Plugs	1	21.3	0.138	21.3	0.158	18.9	0.182	16.6	0.0513
		2	**19.5**	**0.126**	**18.6**	**0.136**	**14.6**	**0.078**	**14.9**	**0.013**
CWE	Clothes Washer	1	**2.86**	**0.217**	2.75	**0.0134**	2.03	0.0135	2.18	**0.0139**
		2	3.29	0.223	**2.3**	0.149	**1.89**	**0.0037**	**1.93**	0.058
DWE	Dishwasher	1	3.65	**0.0108**	2.79	0.0741	1.76	0.0105	1.92	0.0139
		2	**3.06**	**0.0108**	**2.37**	**0.0457**	**1.29**	**0.00108**	**1.31**	**0.00322**
EQE	Network Equipment	1	**1.35**	**0.0134**	**1.54**	0.0139	**1.33**	**0.0048**	**1.33**	**0.0038**
		2	1.55	0.0136	1.55	**0.0137**	1.35	0.0081	1.34	0.0059
FRE	Furnace Fan & Theromstat	1	9.85	**0.0262**	10.3	**0.0225**	9.54	0.0188	9.94	**0.01**
		2	**9.58**	0.0287	**10.15**	0.0266	**9.26**	**0.0152**	**9.77**	0.0122
HPE	Heat Pump	1	**22.1**	**0.0541**	**22.8**	0.054	**9.19**	**0.026**	**9.28**	**0.0204**
		2	23.7	0.0616	23.3	**0.0516**	10.9	0.033	10.9	0.0268
OFE	Home Office	1	26.1	**0.272**	27	0.298	28.1	0.385	29.6	**0.422**
		2	**26**	0.274	**26.5**	**0.283**	**27.5**	**0.367**	**29.2**	**0.422**
UTE	Utility Room	1	**1.49**	**0.0278**	1.4	0.0244	**0.71**	**0.0058**	0.68	**0.0029**
		2	2.19	0.0428	**1.34**	**0.0234**	0.75	0.0104	**0.65**	0.0043
WOE	Kitchen Wall Oven	1	2.26	0.206	3.48	0.146	1.55	0.123	1.79	0.115
		2	**1.67**	**0.0703**	**2.46**	**0.0919**	**1.33**	**0.0673**	**1.37**	**0.0244**
B2E	Southern Bedroom	1	13.6	**0.154**	**14.3**	**0.263**	13.4	0.182	13.3	0.23
		2	**13.5**	0.162	14.5	0.312	**13.1**	**0.148**	**13.2**	**0.218**
CDE	Clothes Dryer	1	**2.16**	**0.0153**	**2.43**	**0.0138**	**1.68**	0.0076	**1.65**	**0.0023**
		2	3.32	0.027	3.41	0.0243	2.51	**0.0012**	2.45	0.0079
DNE	Dining Room	1	**1.29**	**0.372**	1.3	**0.362**	**1.29**	0.381	1.28	**0.38**
		2	**1.29**	0.378	**1.27**	0.402	1.33	**0.332**	**1.27**	0.406
EBE	Electronics Workbench	1	**9.74**	**0.065**	**10.5**	**0.143**	9.81	0.125	7.96	**0.046**
		2	12.5	0.336	13.6	0.393	**7.79**	**0.049**	**6.74**	0.154
FGE	Fridge	1	40.1	**0.098**	41.2	**0.097**	30.4	0.094	29.5	0.094
		2	**34.5**	0.106	**34.4**	0.104	**24.9**	**0.086**	**24.6**	**0.081**
HTE	Instant Hot Water	1	3.12	**0.0804**	1.48	0.0237	3.08	0.0893	1.48	0.0288
		2	**2.84**	0.102	**1.44**	**0.0026**	**2.94**	**0.078**	**1.41**	**0.0087**
OUE	Outside Plug	1	**0.0106**	0.298	**0.0114**	0.176	**0.011**	0.228	**0.011**	0.23
		2	0.0145	**0.288**	0.014	**0.15**	0.014	**0.196**	0.0137	**0.063**
TVE	TV & Entertainment	1	14	0.147	12.8	**0.144**	12.9	0.138	12.5	**0.143**
		2	**13.6**	**0.144**	**12.4**	0.148	**12.4**	**0.131**	**12.2**	0.151

The dishwasher with input channels P and water was used for architecture optimization

Table 2: Performance of Seq2Point Models Using Different Input Channels

		Baseline (P)			P + Water			P + Q			P + Q + Water		
Circuit	Description	MAE	NMAE	SAE	MAE	NMAE	SAE	MAE	NMAE	SAE	MAE	NMAE	SAE
RSE	Rental Unit	86.8	0.357	0.0522	83.7	0.344	0.0472	**51**	**0.21**	0.0065	51.1	**0.21**	**0.0013**
GRE	Garage	1.39	0.818	0.154	1.38	0.812	0.149	1.38	0.812	0.145	**1.36**	**0.8**	**0.11**
B1E	Northern Bedroom	1.61	2.013	**0.186**	1.72	2.15	0.296	**1.51**	**1.888**	0.26	1.54	1.925	0.242
BME	Basement Plugs	19.5	0.6	0.126	18.6	0.572	0.136	**14.6**	**0.449**	0.078	14.9	0.458	**0.013**
CWE	Clothes Washer	3.29	0.783	0.223	2.3	0.548	0.149	**1.89**	**0.45**	**0.0037**	1.93	0.46	0.058
DWE	Dishwasher	3.06	0.23	0.0108	2.37	0.178	0.0457	**1.29**	**0.097**	**0.00108**	1.31	0.098	0.00322
EQE	Network Equipment	1.55	0.038	0.0136	1.55	0.038	0.0137	1.35	**0.033**	0.0081	**1.34**	**0.033**	**0.0059**
FRE	Furnace Fan & Thermostat	9.58	0.084	0.0287	10.15	0.089	0.0266	**9.26**	**0.081**	0.0152	9.77	0.086	**0.0122**
HPE	Heat Pump	23.7	0.162	0.0616	23.3	0.159	0.0516	**10.9**	**0.075**	0.033	**10.9**	**0.075**	**0.0268**
OFE	Home Office	**26**	**0.858**	**0.274**	26.5	0.875	0.283	27.5	0.908	0.367	29.2	0.964	0.422
UTE	Utility Room	2.19	0.043	0.0428	1.34	0.026	0.0234	0.75	0.015	0.0104	**0.65**	**0.013**	**0.0043**
WOE	Kitchen Wall Oven	1.67	0.341	0.0703	2.46	0.502	0.0919	**1.33**	**0.271**	0.0673	1.37	0.28	**0.0244**
B2E	Southern Bedroom	13.5	0.758	0.162	14.5	0.815	0.312	**13.1**	**0.736**	**0.148**	13.2	0.742	0.218
CDE	Clothes Dryer	3.32	0.075	0.027	3.41	0.077	0.0243	2.51	0.057	**0.0012**	**2.45**	**0.056**	0.0079
DNE	Dining Room	1.29	1.433	0.378	**1.27**	**1.411**	0.402	1.33	1.478	**0.332**	**1.27**	**1.411**	0.406
EBE	Electronics Workbench	12.5	0.625	0.336	13.6	0.68	0.393	7.79	0.39	**0.049**	**6.74**	**0.337**	0.154
FGE	Fridge	34.5	0.625	0.106	34.4	0.623	0.104	24.9	0.451	0.086	**24.6**	**0.446**	**0.081**
HTE	Instant Hot Water	2.84	0.4	0.102	1.44	0.203	**0.0026**	2.94	0.414	0.078	**1.41**	**0.199**	0.0087
OUE	Outside Plug	0.0145	2.417	0.288	0.014	2.333	0.15	0.014	2.333	0.196	**0.0137**	**2.283**	**0.063**
TVE	TV & Entertainment	13.6	0.36	0.144	12.4	0.328	0.148	12.4	0.328	**0.131**	**12.2**	**0.323**	0.151
Geometric Mean		4.6	0.383	0.0928	4.3	0.436	0.0813	3.46	0.289	0.038	**3.303**	**0.276**	**0.034**