Few Labels are all you need:
A Weakly Supervised Framework for Appliance Localization in Smart-Meter Series

Adrien Petralia[1,2], Paul Boniol[3,4], Philippe Charpentier[1], Themis Palpanas[2,5]

EDF R&D[1], Université Paris Cité[2], Inria[3], ENS[4], IUF[5]
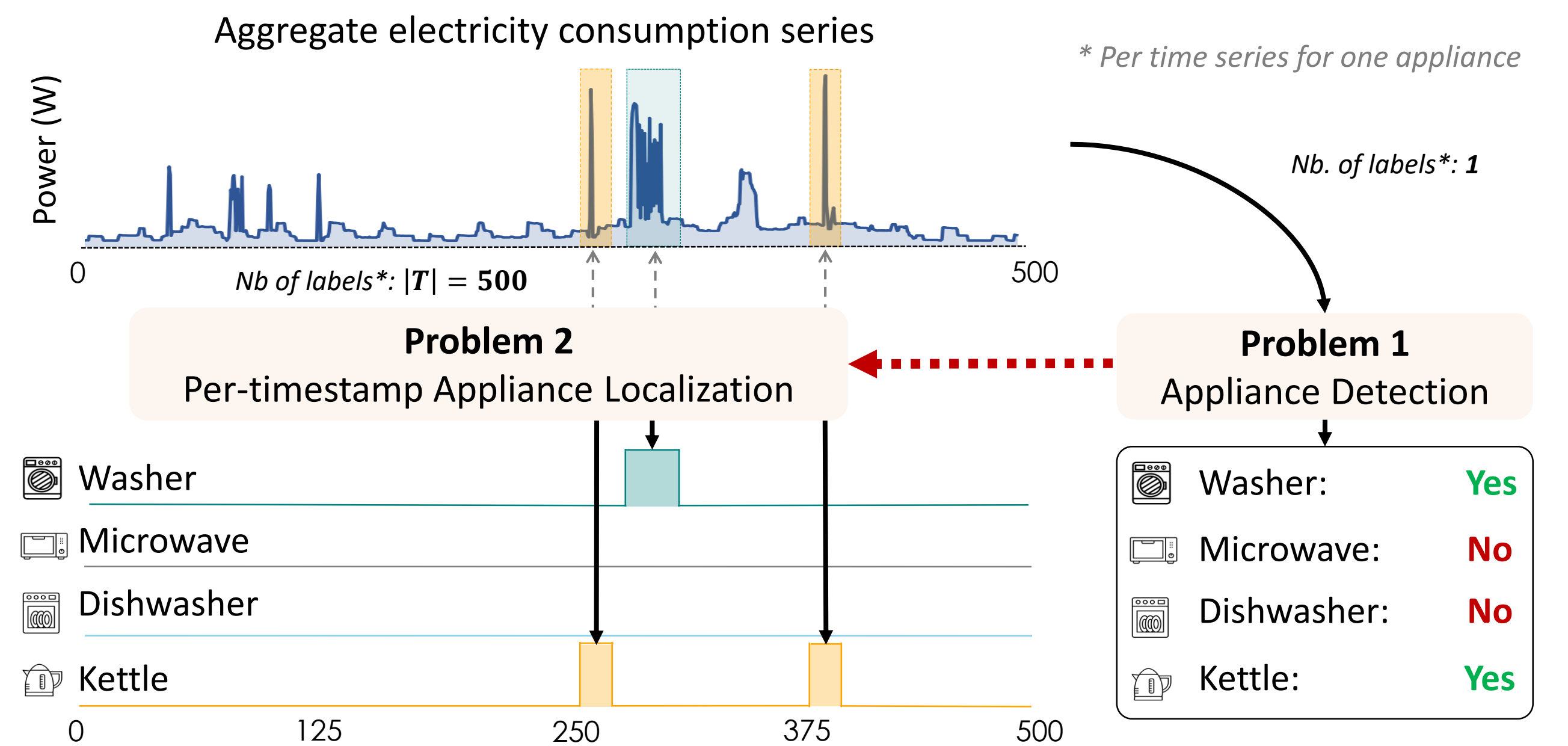
adrien.petralia@gmail.com

1. Motivation

- **Smart grid flexibility** requires **identifying appliance usage** from aggregate smart meter data

- **Existing methods** for this task need **fine-grained labels** at each time step (e.g., strongly supervised) [1]

- **Label collection** is **costly** and **intrusive** (dedicated in-home sensors)

- Appliance **detection** (yes/no) can be framed as **time series classification** [2]

- **Explainability-driven methods** use **weak labels** to **highlight relevant regions** [3]

2. Objective

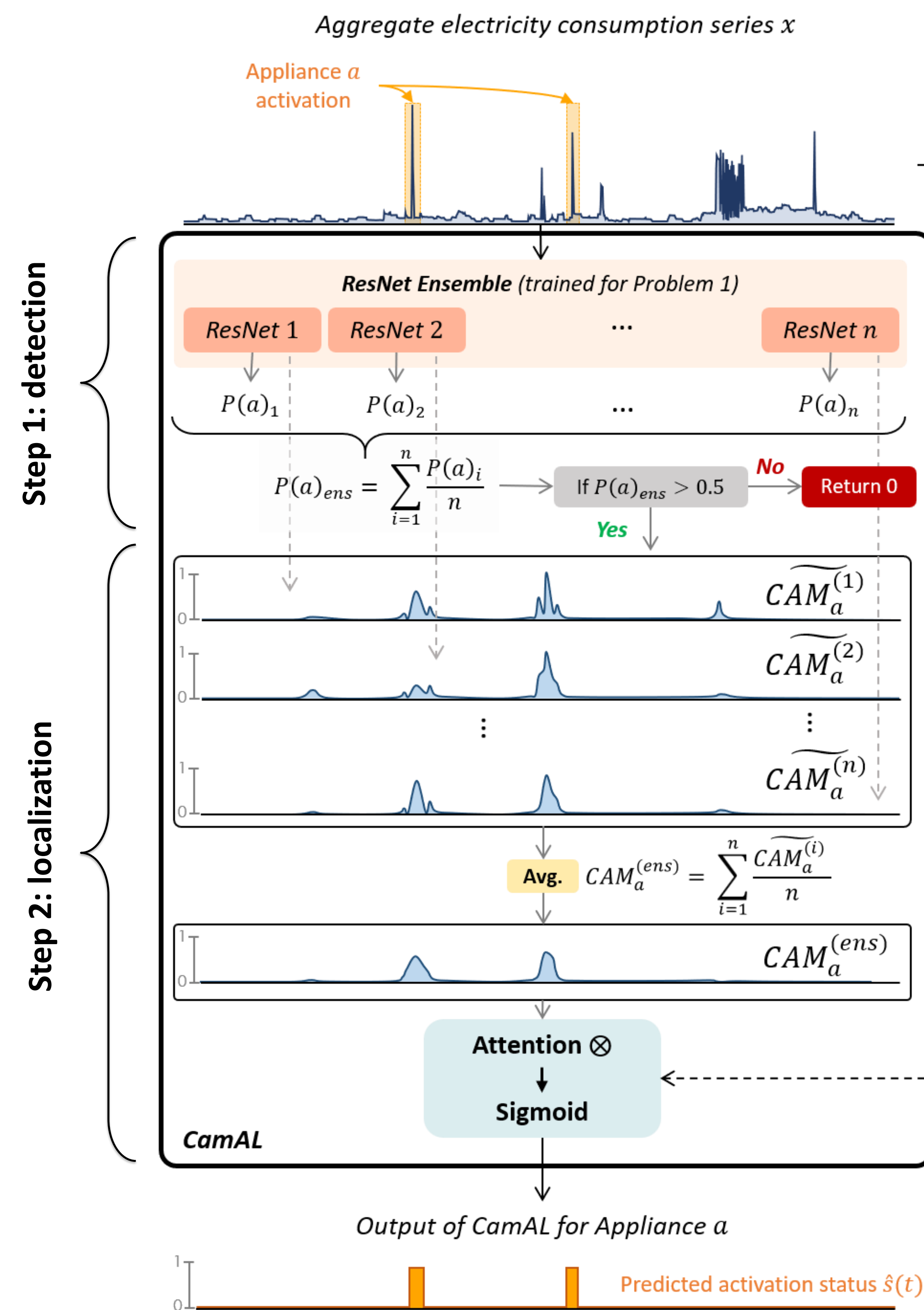**Per-timestamp Appliance Localization** based on **Appliance Detection**

Aggregate electricity consumption series

Per time series for one appliance

Nb. of labels*: 1

Nb of labels*: $|T| = 500$

Problem 2
Per-timestamp Appliance Localization

Problem 1
Appliance Detection

- Washer
- Microwave
- Dishwasher
- Kettle

Washer:	**Yes**
Microwave:	**No**
Dishwasher:	**No**
Kettle:	**Yes**

3. Proposed Approach: **CamAL**

***C**lass **A**ctivation **M**aps based **A**ppliance **L**ocalization*

- Based on an **ensemble** of **CNN classifiers** (ResNets)
- Each model trained to only **detect presence of an appliance**
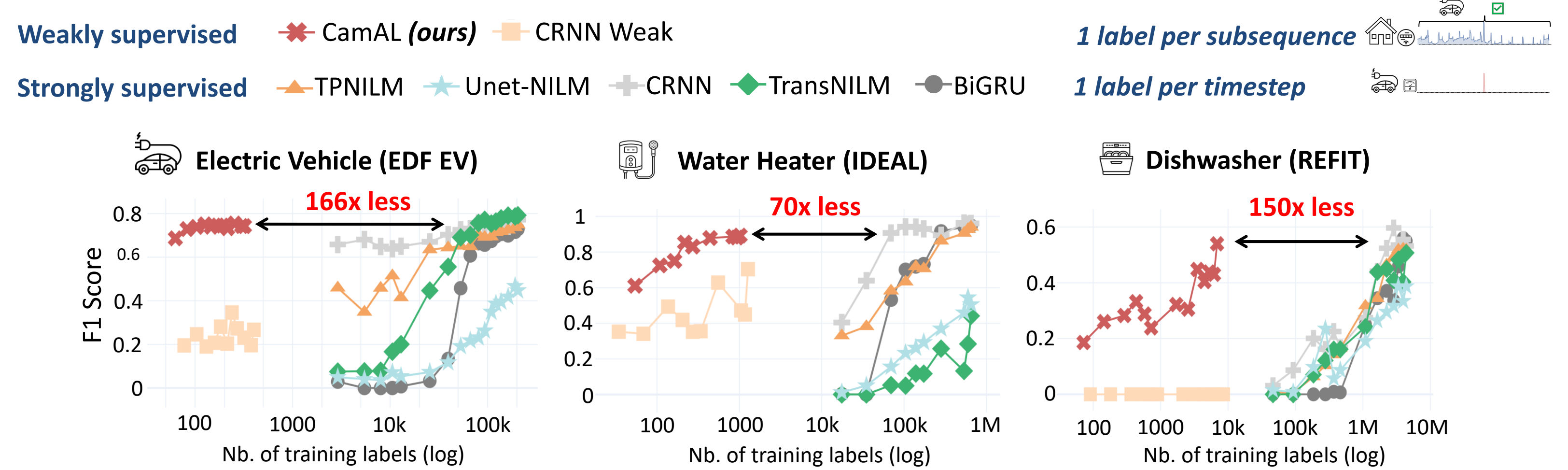- Use **explainability module** to **localize activation pattern**

Step 1: Detect appliance presence in input consumption series

Step 2: *If detected,* **extract** and **analyze** decision-related information (CAMs) to perform activation localization.

Aggregate electricity consumption series x

Appliance a activation

Step 1: detection

ResNet Ensemble (trained for Problem 1)

ResNet 1 | *ResNet 2* | ... | *ResNet n*

$P(a)_1$ | $P(a)_2$ | ... | $P(a)_n$

$P(a)_{ens} = \sum_{i=1}^{n} \frac{P(a)_i}{n}$ → If $P(a)_{ens} > 0.5$ → **No** → Return 0 / **Yes**

Step 2: localization

$\widetilde{CAM}_a^{(1)}$

$\widetilde{CAM}_a^{(2)}$

$\widetilde{CAM}_a^{(n)}$

Avg. $CAM_a^{(ens)} = \sum_{i=1}^{n} \frac{\widetilde{CAM}_a^{(i)}}{n}$

$CAM_a^{(ens)}$

Attention $\otimes$

Sigmoid

CamAL

Output of CamAL for Appliance a
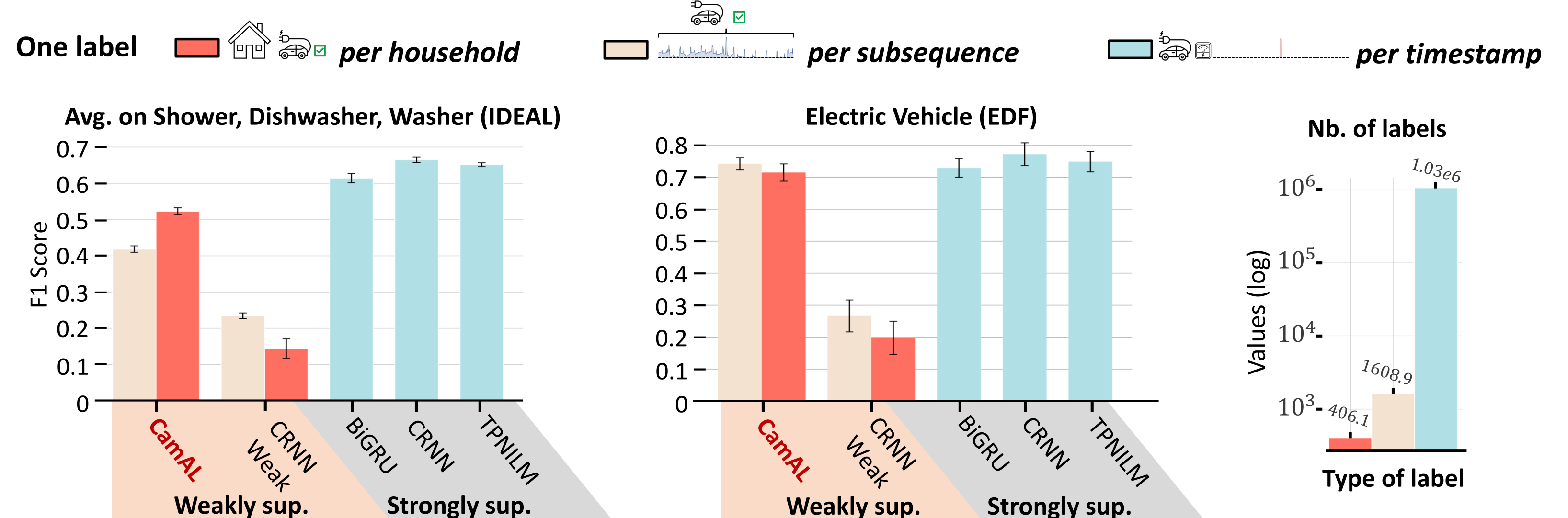
Predicted activation status $\hat{s}(t)$

4. Experimental Evaluation

*1. How does **CamAL** perform compared to **strongly supervised approaches** according to **number of training labels**?*
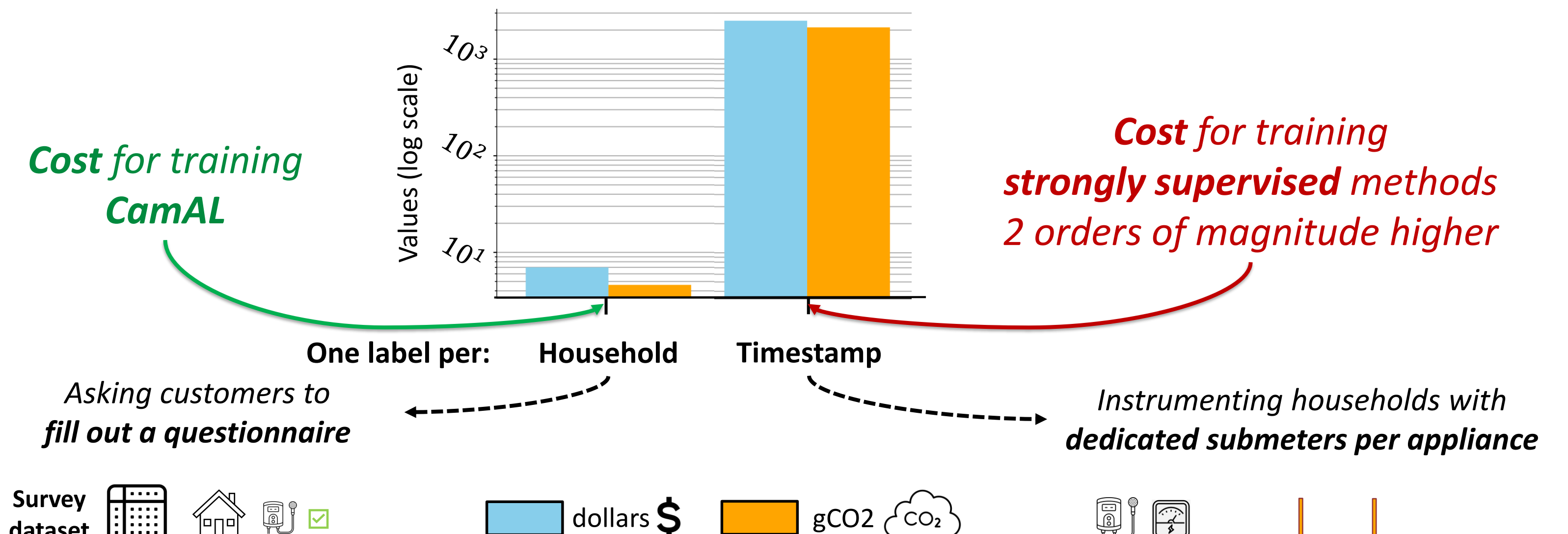
Weakly supervised CamAL *(ours)* CRNN Weak   1 label per subsequence

Strongly supervised TPNILM Unet-NILM CRNN TransNILM BiGRU   1 label per timestep

Electric Vehicle (EDF EV) — **166x less**
Water Heater (IDEAL) — **70x less**
Dishwasher (REFIT) — **150x less**

F1 Score vs Nb. of training labels (log)

- **CamAL** needs on average **145× less training labels** for the same accuracy as **strongly supervised methods** (CamAL trained with only 1 label per sub-sequence)
- **CamAL** achieves **comparable performance** on high-consumption, strategic appliances for **suppliers** (e.g., Electric Vehicle, Water Heater)

*2. Is **appliance ownership** information **sufficient** to train **CamAL**?*

One label per household   per subsequence   per timestamp

Avg. on Shower, Dishwasher, Washer (IDEAL) | Electric Vehicle (EDF) | Nb. of labels

CamAL / CRNN Weak (Weakly sup.) | BIGRU / CRNN / TPNILM (Strongly sup.)

$1.03e6$ / 1608.9 / 406.1

Yes!
- **CamAL** achieves **equivalent** or **superior** performance using only **appliance ownership information**
- **CamAL** needs up to **5200x less** training labels compared to **strongly supervised baselines**

*3. What is the **label collection cost** for **training CamAL** vs. **training strongly supervised approaches**?*

Cost** for training **CamAL

***Cost** for training **strongly supervised** methods 2 orders of magnitude higher*

Values (log scale)

One label per: Household / Timestamp

Asking customers to fill out a questionnaire

Instrumenting households with dedicated submeters per appliance

Survey dataset | dollars $ | gCO2 CO_2

Bibliography

[1] A review of current methods and challenges of advanced deep learning-based non-intrusive load monitoring (NILM) in residential context, Hasan Rafiq, Prajowal Manandhar, Edwin Rodriguez-Ubinas, Omer Ahmed Qureshi, Themis Palpanas, Energy&Building, 2024.

[2] Appliance Detection Using Very Low-Frequency Smart Meter Time Series, Adrien Petralia, Paul Boniol, Philippe Charpentier, and Themis Palpanas, ACM e-Energy, 2023.

[3] dCAM: Dimension-wise Class Activation Map for Explaining Multivariate Data Series Classification, Paul Boniol, Mohamed Meftah, Emmanuel Remy and Themis Palpanas, SIGMOD, 2022.

Github

EDF | Université Paris Cité | diNo | Inria | ENS ÉCOLE NORMALE SUPÉRIEURE | AI4EUROPE | RECITALS | DiS TwinODIS | eosc | DataGEMS