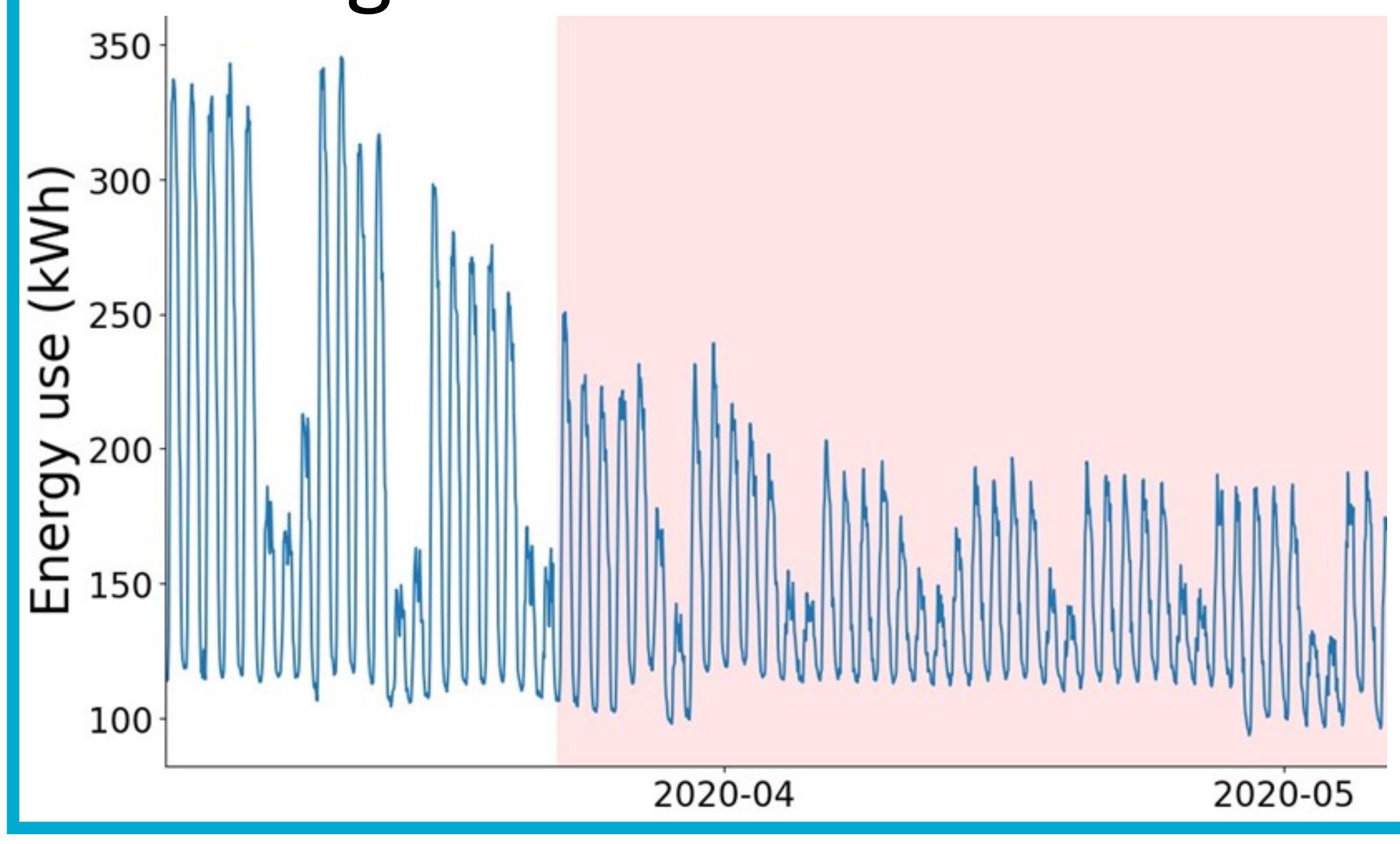
Continually learning out-of-distribution (OOD) spatiotemporal data for robust energy forecasting

Arian Prabowo (UNSW), Kaixuan Chen (UNSW), Hao Xue (UNSW), Subbu Sethuvenkatraman (CSIRO), Flora D. Salim (UNSW)

Challenge: out-of-distribution energy use forecasting



- Extreme events like COVID lock-downs, indicated in red, significantly alter data distribution.
- This also changed relationships between features e.g., remote work increases residential energy consumption during working hours.
- Traditional offline learning struggles against these changes in distributions.

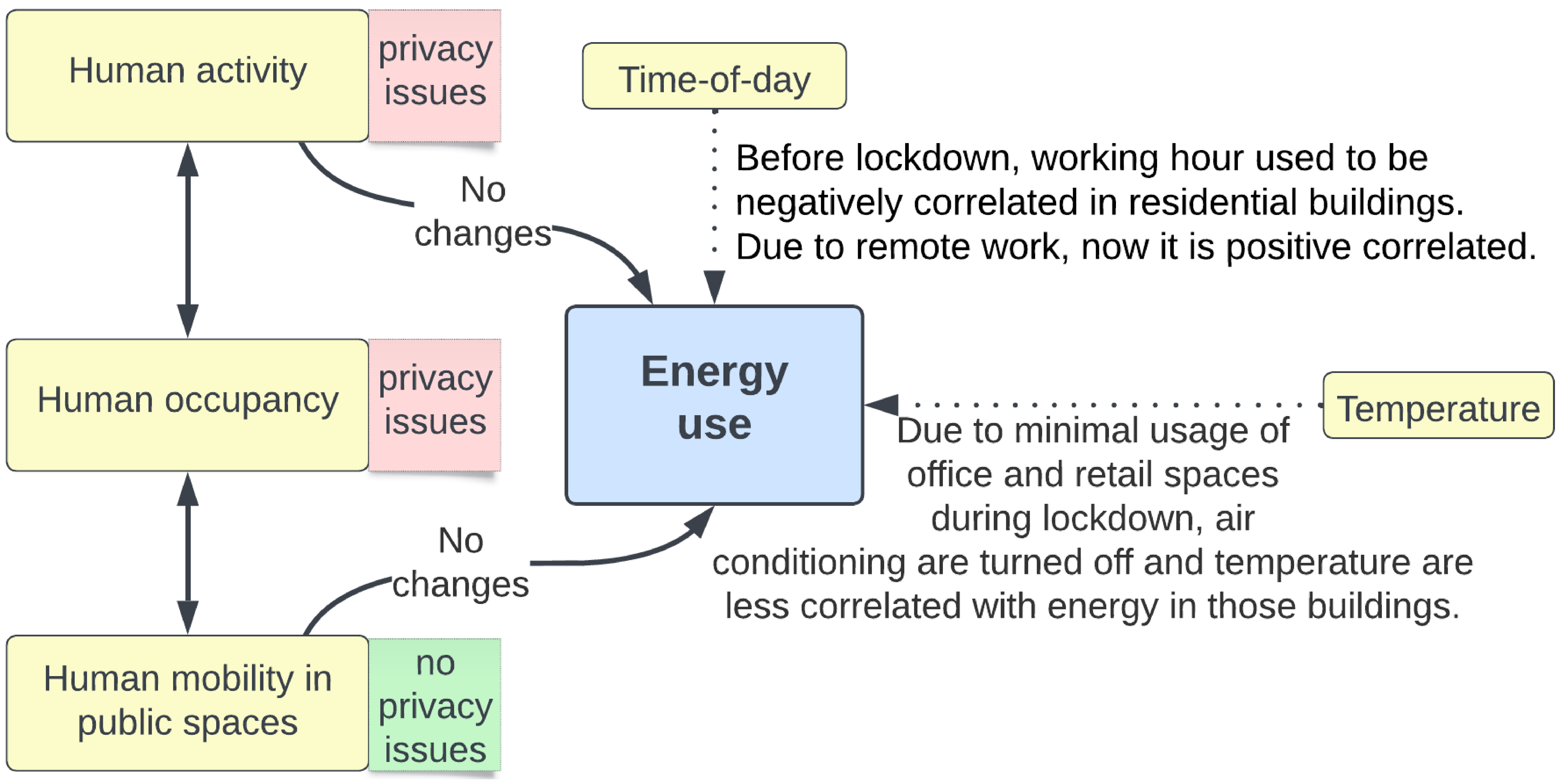
Continual Learning (CL)

Offline learning, also known as batch learning, involves training a model on a fixed dataset all at once, thus it does not adapt to new data or updates after the initial training. This approach struggles during out-of-distribution periods.

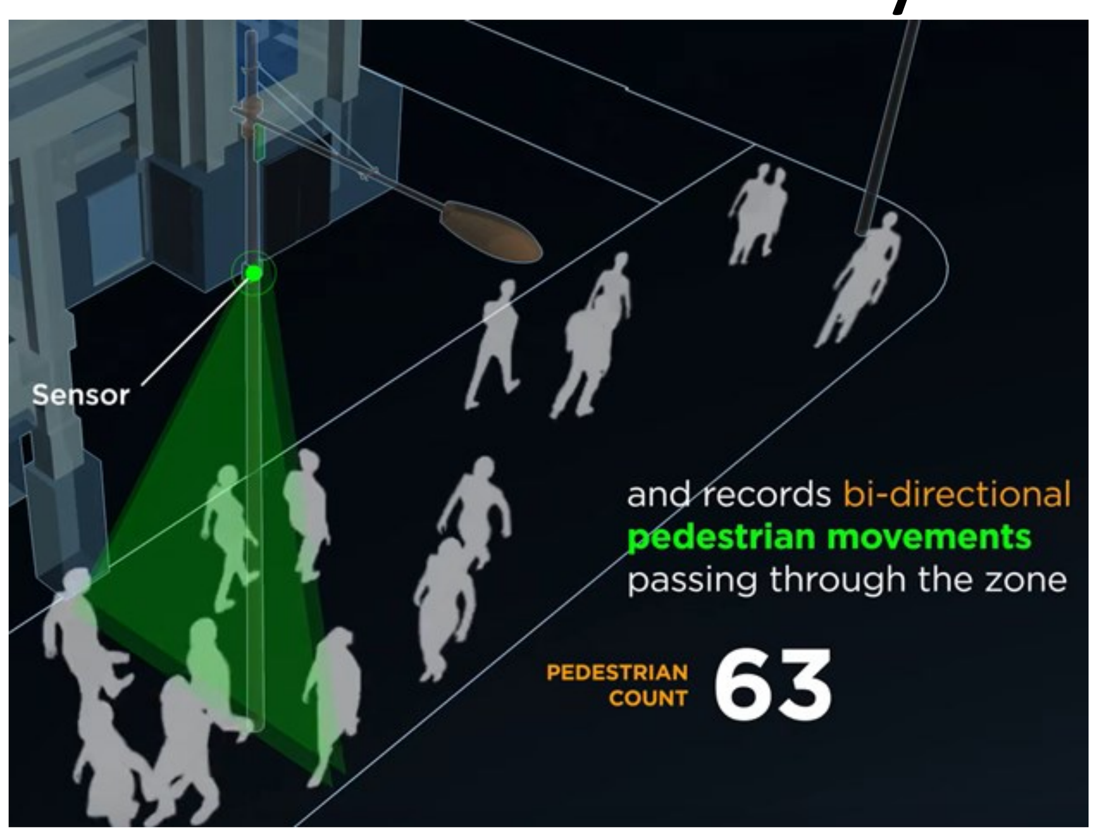
Online learning updates a model incrementally with new data points as they become available, thus it allows real-time adaptation to changing data distribution. However, it can suffers from catastrophic forgetting, where the model forgets prior knowledge when adapting to new information.

Continual learning (CL) is a specific subset of online learning that aims to overcome catastrophic forgetting. It focuses on retaining and consolidating prior knowledge.

How different features affect energy use during COVID–19 lockdowns?



Dataset: Mobility



The pedestrian counters are located in public spaces. They do not capture any photo or video, thus preserving privacy.

Results: Mobility

Table 2. Performance comparison between different contextual features. Results are average over 10 runs with different random seed. The standard deviation is shown. The algorithm used was FSNet with continual learning. +M is the improvement of adding mobility over no context, +T is the improvement of adding temperature over no context, T+M is the improvement of adding mobility over temperature only.

(MAE)	dataset	no context	mobility only	temp. only	both	+M	+T	T+M
Pre-Lockdown	BC1	0.1591 ±0.0252	0.1587 ±0.0334	0.1595 ±0.0269	0.1516 ±0.0332	0.0004	-0.0004	0.0079
	BC2	0.1711 ±0.0085	0.1993 ±0.0385	0.1947 ±0.0391	0.1708 ±0.0068	-0.0282	-0.0236	0.0239
	BC3	0.2629 ±0.0373	0.2866 ±0.0534	0.2509 ±0.0262	0.2403 ±0.0095	-0.0237	0.0120	0.0105
	BC4	0.2706 ±0.0370	0.2516 ±0.0206	0.3142 ±0.1581	0.2776 ±0.0312	0.0190	-0.0436	0.0366
Post-Lockdown	BC1	0.1484 ±0.0318	0.1475 ±0.0464	0.1434 ±0.0283	0.1369 ±0.0355	0.0033	0.0041	0.0041
	BC2	0.1636 ±0.0085	0.1902 ±0.0371	0.1849 ±0.0381	0.1624 ±0.0063	0.0072	-0.0194	0.0053
	BC3	0.2418 ±0.0374	0.2654 ±0.0537	0.2299 ±0.0252	0.2198 ±0.0089	-0.0014	-0.0251	0.0355
	BC4	0.3236 ±0.0602	0.2943 ±0.0294	0.4134 ±0.3215	0.3282 ±0.0502	0.0293	-0.1191	0.0852

The importance of mobility data becomes particularly pronounced in the post-lockdown period, especially when considering its interaction with temperature information. This valuable insight help us better understand the dynamics of energy usage and building occupancy during out-of-distribution periods.

Results: CL

Table 3. Comparing the performance of different algorithm with or without continual learning (CL). The metric used is MAE. Results are average over 10 runs with different random seed. The standard deviation is shown.

	dataset	FSNet (no CL)	FSNet	TCN (no CL)	OGD	ER	DR++
Pre-Lockdown	BC1	0.3703 ±0.0607	**0.1583** ±0.0280	0.3668 ±0.0379	0.2056 ±0.0413	0.1820 ±0.0217	0.1696 ±0.0130
	BC2	0.6272 ±0.0914	**0.1712** ±0.0063	0.5176 ±0.0607	0.2465 ±0.0105	0.2322 ±0.0056	0.2272 ±0.0062
	BC3	0.6750 ±0.0638	**0.2462** ±0.0151	0.6500 ±0.0698	0.3308 ±0.0812	0.2862 ±0.0432	0.2726 ±0.0334
	BC4	1.0018 ±0.1053	**0.2802** ±0.0312	1.1236 ±0.1040	0.3910 ±0.0520	0.3511 ±0.0323	0.3408 ±0.0210
Post-Lockdown	BC1	0.4537 ±0.0517	**0.1429** ±0.0275	0.4179 ±0.0443	0.1797 ±0.0342	0.1589 ±0.0168	0.1482 ±0.0094
	BC2	0.6506 ±0.0994	**0.1628** ±0.0057	0.5209 ±0.0535	0.2313 ±0.0085	0.2188 ±0.0060	0.2148 ±0.0068
	BC3	0.7168 ±0.0632	**0.2255** ±0.0145	0.7083 ±0.0793	0.3014 ±0.0709	0.2636 ±0.0373	0.2518 ±0.0286
	BC4	1.8415 ±0.2765	**0.3314** ±0.0520	1.8307 ±0.2319	0.4496 ±0.0643	0.4162 ±0.0475	0.4043 ±0.0338

In our experiments, we compared four CL methods on a TCN backbone. The results revealed that CL provides performance enhancements, with its advantages becoming notably more pronounced post -lockdown. Among the four CL methods, FSNet consistently outperformed others.

Dataset: Melbourne

The dataset used in this study is sourced from four building complexes located in Melbourne, Australia. Melbourne gained notoriety for implementing one of the **world's longest and strictest COVID-19 lockdowns**. To protect privacy, we aggregated nearby buildings with different usage into building complexes (BC).

Conclusion

Leveraging **mobility** data and **CL** techniques is essential for energy forecasting during OOD periods.

Link to this paper

Link to a follow up paper with 13 BCs