Navigating Out-of-Distribution Electricity Load Forecasting during COVID-19: Benchmarking energy load forecasting models without and with continual learning

Dr. Arian Prabowo (University of New South Wales, UNSW)
Dr. Kaixuan Chen (University of New South Wales, UNSW)
Dr. Hao Xue (University of New South Wales, UNSW)
Dr. Subbu Sethuvenkatraman (CSIRO, Energy)
Prof. Flora D. Salim (University of New South Wales, UNSW)

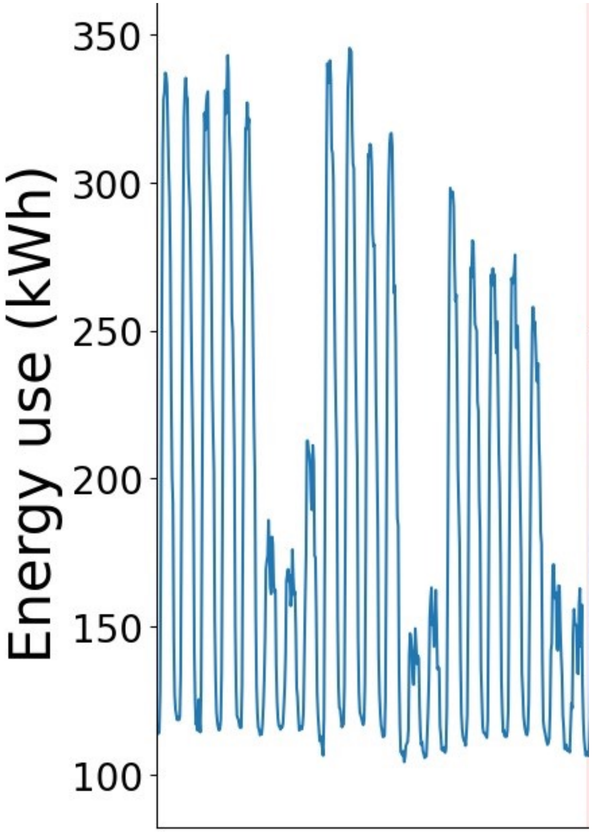
Centre for New Energy Technologies

Navigating Out-of-Distribution Electricity Load Forecasting during COVID-19: Benchmarking **<u>energy load forecasting</u>** models without and with continual learning

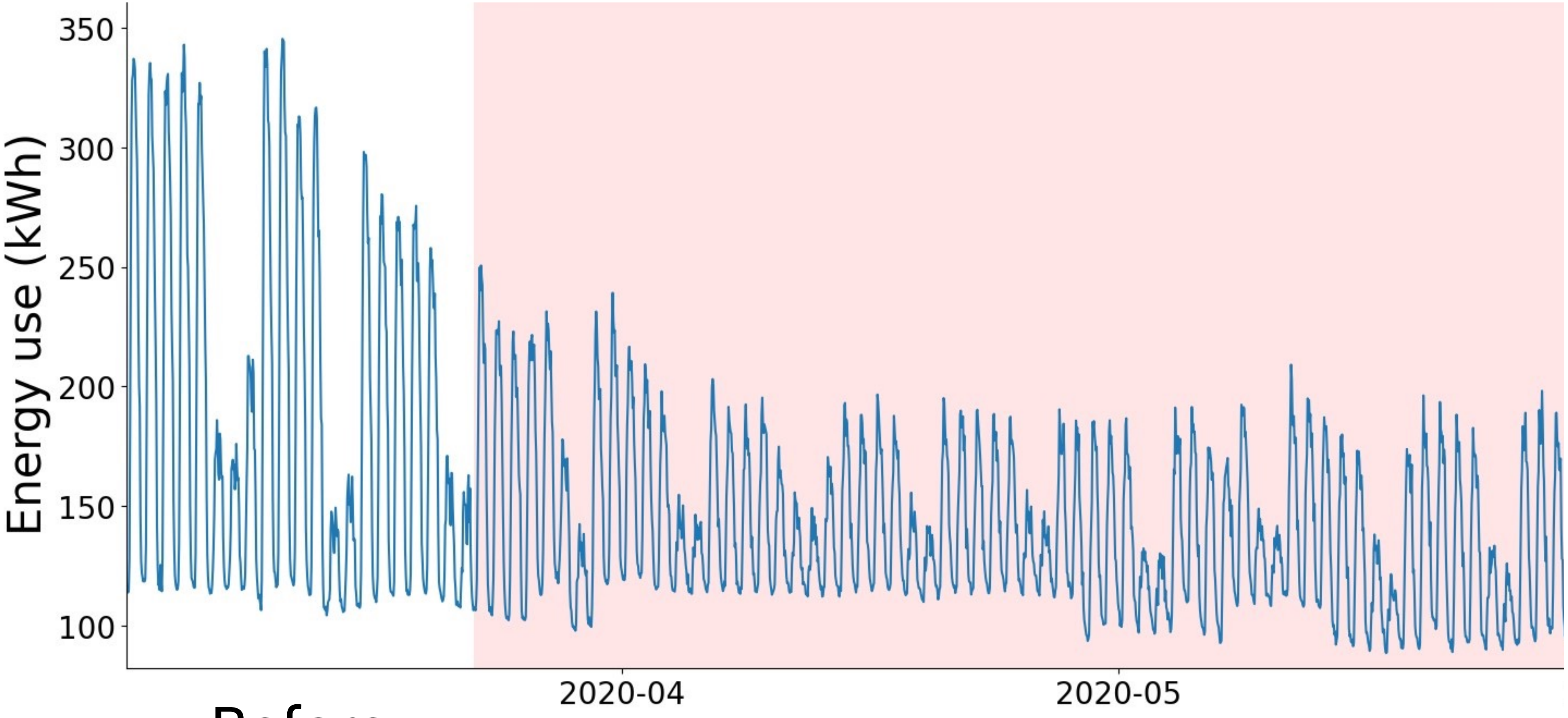
Dr. Arian Prabowo (University of New South Wales, UNSW)
Dr. Kaixuan Chen (University of New South Wales, UNSW)
Dr. Hao Xue (University of New South Wales, UNSW)
Dr. Subbu Sethuvenkatraman (CSIRO, Energy)
Prof. Flora D. Salim (University of New South Wales, UNSW)

Centre for New Energy Technologies

Two weeks of energy use in a building.



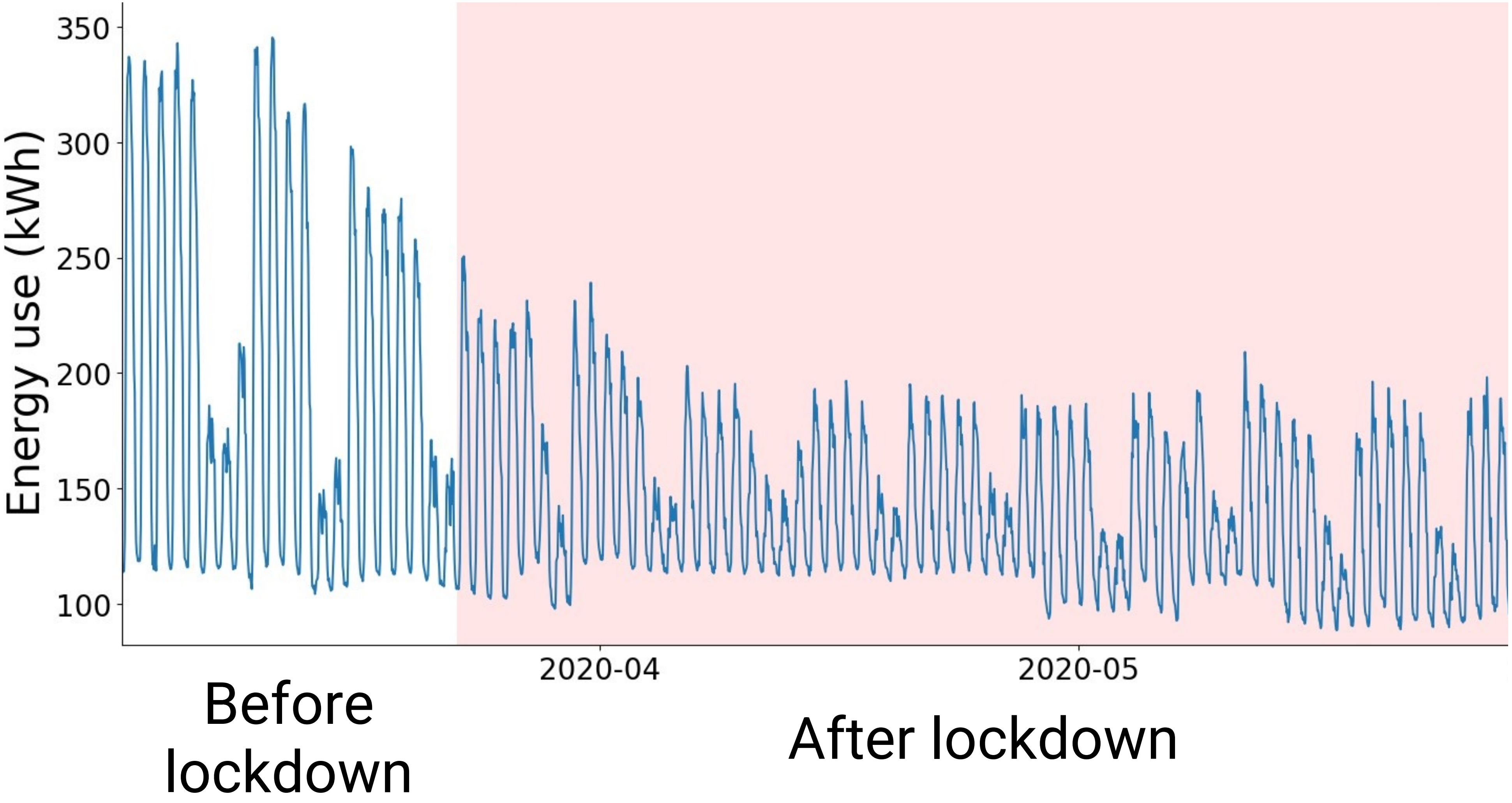
Extreme events (e.g. COVID-19) change the energy use patterns



Before lockdown

After lockdown

What is the best way to adapt to such changes?



Datasets: Energy

- 13 Building Complexes (BC)
  - Higher spatial granularity than what is common in the literature (e.g. substations)

UNSW SYDNEY

Datasets: Energy

- 13 Building Complexes (BC)
  - Higher spatial granularity than what is common in the literature (e.g. substations).
- All from Melbourne, Australia.
  - The city which experienced one of the longest and strictest COVID lockdown.

The datasets are diverse.

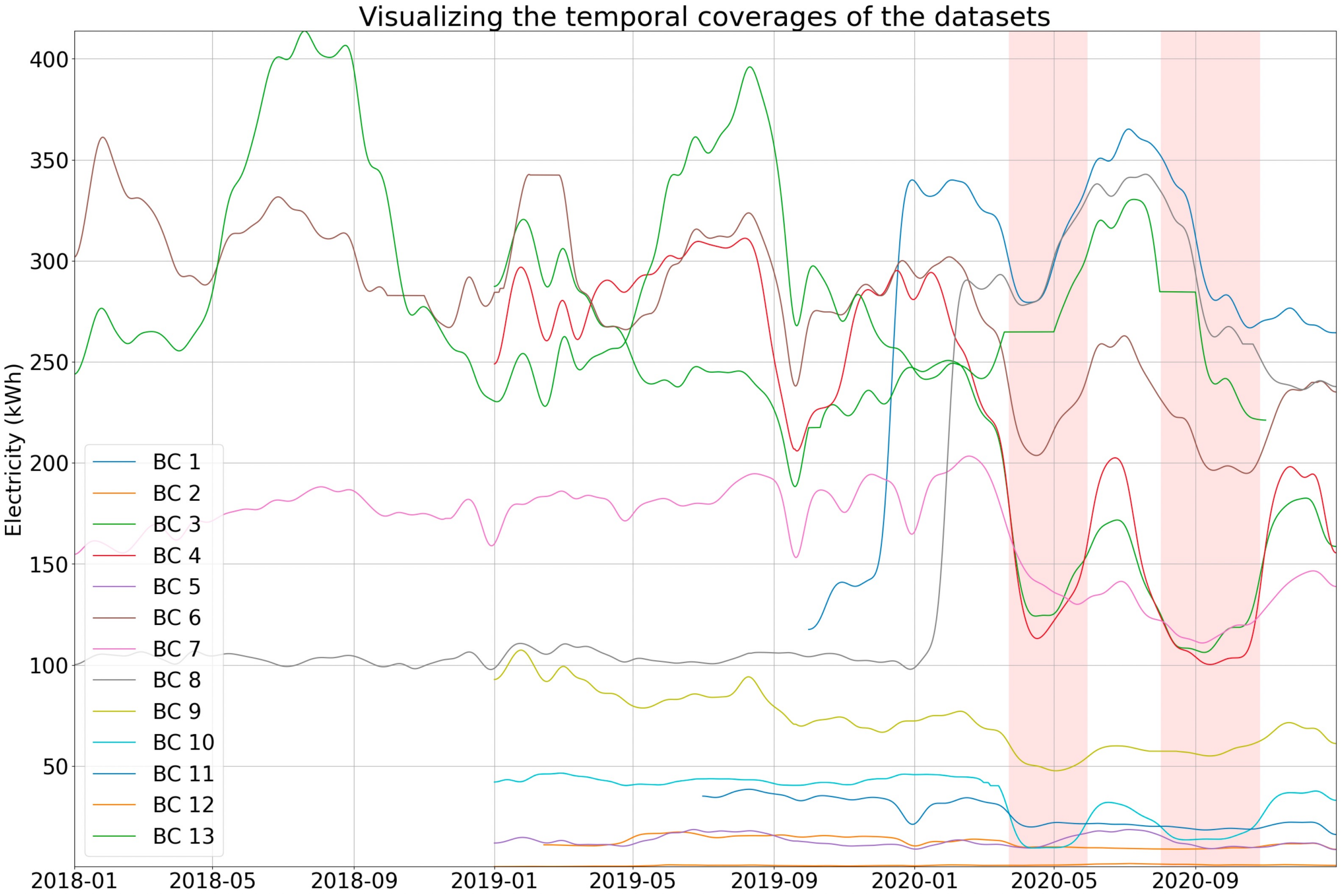
Compare the mean and median columns.

BC	start	end	duration (years)	mean (kWh)	std. (kWh)	min (kWh)	0.25 (kWh)	median (kWh)	0.75 (kWh)	max (kWh)
1	2018-08-01	2020-12-31	2.42	276.73	99.68	19.38	213.31	280.78	347.45	686.01
2	2019-02-13	2020-12-31	1.88	12.33	6.68	0.00	8.50	10.31	13.65	50.16
3	2019-01-01	2020-12-31	2.00	207.27	111.59	3.25	112.72	169.29	297.33	611.67
4	2019-01-01	2020-12-31	2.00	223.27	108.37	0.00	134.52	199.20	307.09	611.18
5	2019-01-01	2020-12-31	2.00	12.94	9.65	3.68	5.77	8.85	17.65	55.26
6	2018-01-01	2020-12-31	3.00	278.17	88.67	1.32	203.38	272.68	342.10	709.41
7	2018-01-01	2020-12-31	3.00	166.42	66.60	5.38	112.62	144.34	206.63	371.64
8	2018-01-01	2020-12-31	3.00	160.09	96.24	11.25	92.68	120.79	217.45	550.47
9	2019-01-01	2020-12-31	2.00	72.83	33.64	1.80	45.18	66.71	96.57	218.07
10	2019-01-01	2020-12-31	2.00	34.48	29.20	5.51	14.80	21.76	41.00	123.90
11	2019-07-01	2020-12-31	1.50	26.64	13.21	0.00	17.45	21.75	31.06	83.01
12	2019-01-01	2020-12-31	2.00	0.90	1.04	0.20	0.28	0.62	0.96	15.08
13	2018-01-01	2020-10-31	2.83	293.53	96.17	0.23	231.14	286.68	351.01	675.15

The dataset selection is diverse.

Not only in magnitude, but also in distributions.

Red regions are lockdowns.



Datasets: Mobility

Installed in public space for by the City of Melbourne.

Does not capture images or videos to protect privacy.

Publicly available.

13 Baseline methods

Naïve	CopyLastHour, CopyLastDay, CopyLastWeek	
Stats.	Exponential Smoothing (**ES**)	**Statistical**
ML	Random Forest(**RF**), eXreme Gradient Boosting (**XGBoost**), Vector Autoregression (**VAR**)	**Machine Learning**
DL	Long Short-Term Memory (**LSTM**), Neural Basis Expansion Analysis for interpretable Time Series forecasting (**NBEATS**)	**Deep Learning**
OL	VAR+OL, LSTM+OL, NBEATS+OL	**(ML / DL +) Online Learning**
CL	Fast and Slow Network (**FSNet**)	**Continual Learning (subset of OL)**

UNSW SYDNEY

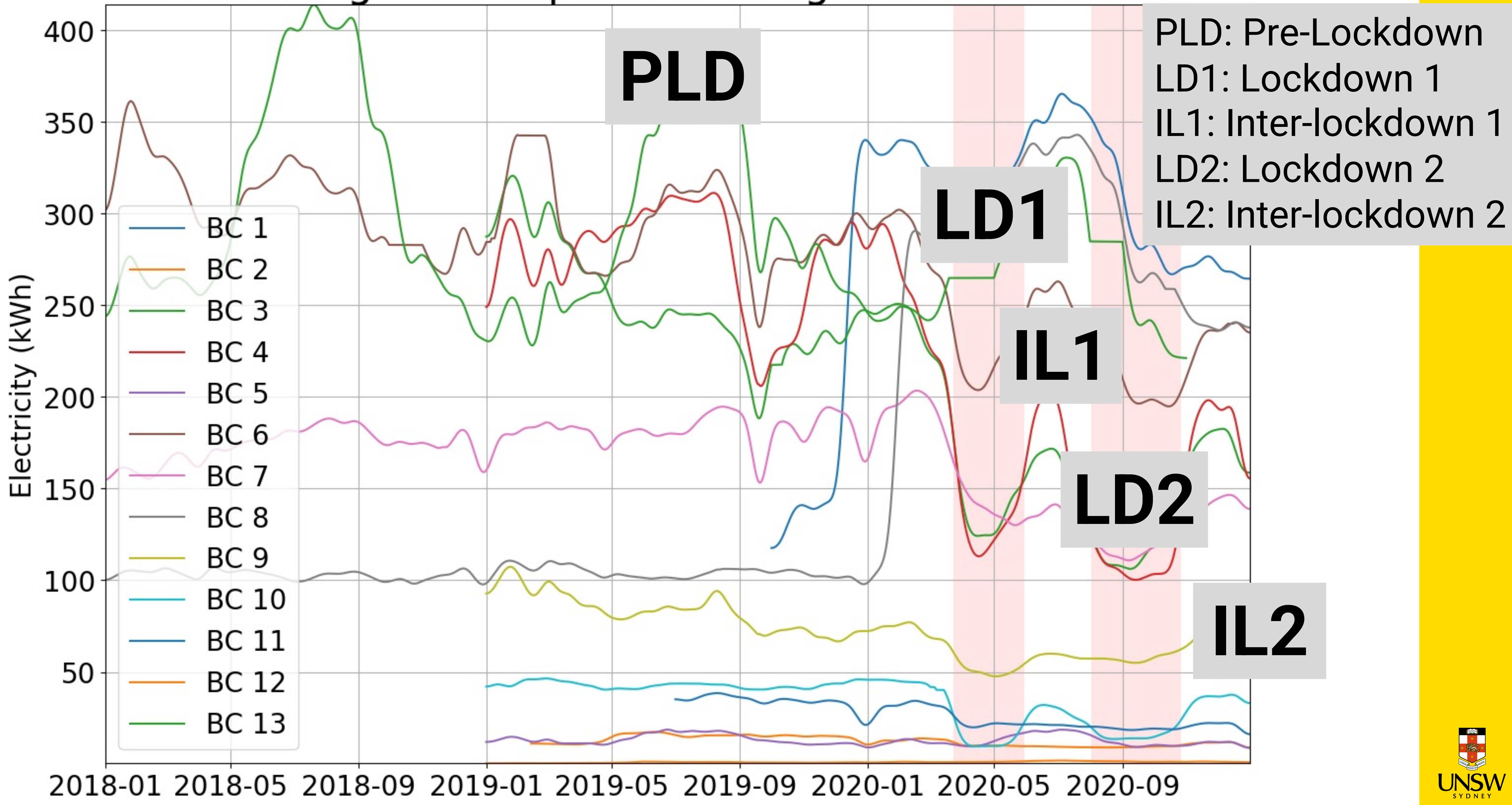
13 Baseline methods

Naïve	CopyLastHour, CopyLastDay, CopyLastWeek
Stats.	Exponential Smoothing (ES)
ML	Random Forest(RF), eXreme Gradient Boosting (XGBoost), Vector Autoregression (VAR)
DL	Long Short-Term Memory (LSTM), Neural Basis Expansion Analysis for interpretable Time Series forecasting (NBEATS)
OL	VAR+OL, LSTM+OL, NBEATS+OL
CL	Fast and Slow Network (FSNet)

4 Feature sets

E	Energy use, HourOfDay, DayOfWeek
ET	E + Temperature
EM	E + Mobility
ETM	All of the above

Visualizing the temporal coverages of the datasets



Results for BC8

(sorted best first)
All combinations of
methods and features.

Methods	Feat.	PLD	LD1	IL1	LD2	IL2
FSNet	EM	5.26	11.79	12.50	10.45	8.10
FSNet	ET	5.67	12.46	12.86	10.89	8.58
FSNet	E	5.76	13.59	14.44	11.81	9.09
NBEATS+OL	EM	7.01	16.41	17.90	17.41	13.79
FSNet	ETM	7.05	14.48	14.43	11.45	8.87
NBEATS+OL	E	7.90	15.74	18.05	16.47	13.75
NBEATS+OL	ETM	8.02	18.36	17.84	17.74	13.28
NBEATS+OL	ET	8.18	19.39	18.02	16.90	12.78
LSTM+OL	ETM	10.21	22.17	20.99	18.31	13.22
LSTM+OL	EM	10.75	21.66	22.05	17.35	13.45
LSTM+OL	E	11.41	21.36	22.47	17.60	13.48
NBEATS	EM	11.66	76.23	92.94	73.37	53.64
NBEATS	E	11.71	87.35	106.98	83.10	61.62
NBEATS	ET	12.30	84.76	110.95	77.96	59.28
LSTM+OL	ET	12.49	38.26	33.48	23.08	14.34
VAR	E	12.70	89.64	107.16	87.43	61.36
VAR+OL	E	12.70	89.64	107.16	87.43	61.36
VAR	ET	14.73	101.85	119.63	101.82	70.37
VAR+OL	ET	14.73	101.85	119.63	101.82	70.37
NBEATS	ETM	16.45	134.57	173.04	126.76	95.06
LSTM	ET	17.82	181.81	224.16	172.39	123.94
LSTM	E	18.01	180.27	220.82	171.09	125.25
LSTM	EM	18.05	185.04	227.54	175.75	127.55
LSTM	ETM	18.52	185.39	228.29	176.47	128.00
XGBoost	E	19.12	188.56	231.48	179.20	131.04
XGBoost	ET	19.16	189.34	232.74	180.14	130.90
RF	EM	19.49	191.75	234.62	182.22	134.48
XGBoost	ETM	21.36	192.14	235.06	182.23	134.75
RF	E	21.39	192.45	235.33	182.87	134.85
XGBoost	EM	21.81	191.20	234.22	180.25	134.80
RF	ET	25.25	193.53	236.40	183.95	135.93
RF	ETM	25.42	193.73	236.61	184.14	136.14
CopyLastWeek	E	45.26	57.21	38.31	32.90	53.17
CopyLastHour	E	53.66	50.62	60.68	40.08	65.93
ES	E	53.66	50.62	60.68	40.08	65.93
CopyLastDay	E	70.80	68.29	75.00	53.54	89.29
VAR	ETM	198.03	145.25	154.38	166.35	105.84
VAR+OL	ETM	198.03	145.25	154.38	166.35	105.84
VAR	EM	200.42	146.02	151.40	167.79	101.00
VAR+OL	EM	200.42	146.02	151.40	167.79	101.00

The best method and feature combination for each BC

BC	Methods	Feat.	PLD	LD1	IL1	LD2	IL2
1	VAR	E	8.04	8.76	11.18	9.83	7.92
2	FSNet	ET	2.37	1.20	1.04	0.80	1.57
3	FSNet	EM	15.63	10.33	9.59	7.15	12.32
4	FSNet	EM	16.04	9.48	11.06	6.99	11.80
5	FSNet	ETM	1.96	1.92	2.95	1.69	1.56
6	FSNet	E	14.71	11.19	11.32	9.50	9.85
7	FSNet	E	10.21	5.69	5.95	4.08	6.21
8	FSNet	EM	5.27	11.80	12.50	10.46	8.10
9	FSNet	E	4.68	2.31	3.56	1.93	4.31
10	FSNet	E	5.47	4.11	4.68	4.25	4.88
11	FSNet	ET	4.10	2.15	2.17	1.70	2.42
12	FSNet	ETM	0.22	0.26	0.47	0.27	0.20
13	FSNet	E	16.88	15.05	15.17	12.97	12.22

FSNet in 12/13

Mobility improves in 5/13

Temp. improves in 4/13

On average,
PLD has the highest while
LD2 has the lowest error

How much better is FSNet compared to 2nd best? VAR+OL appears often.

BC	Methods	Feat.	PLD	LD1	IL1	LD2	IL2
1	VAR	E	8.04	8.76	11.18	9.83	7.92
	FSNet	EM	13.66	9.37	11.93	9.51	9.87
	Δ%		**41%**	**6%**	**6%**	**-3%**	**20%**
2	FSNet	ET	2.37	1.20	1.04	0.80	1.57
	NBEATS+OL	EM	2.91	1.40	1.20	0.98	2.02
	Δ%		**19%**	**14%**	**13%**	**18%**	**22%**
3	FSNet	EM	15.63	10.33	9.59	7.15	12.32
	VAR+OL	E	23.64	14.96	12.04	10.33	20.14
	Δ%		**34%**	**31%**	**20%**	**31%**	**39%**
4	FSNet	EM	16.04	9.48	11.06	6.99	11.80
	VAR+OL	E	23.34	16.78	16.80	10.60	20.16
	Δ%		**31%**	**44%**	**34%**	**34%**	**41%**
5	FSNet	ETM	1.96	1.92	2.95	1.69	1.56
	NBEATS+OL	EM	3.00	2.88	3.78	2.52	2.38
	Δ%		**34%**	**33%**	**22%**	**33%**	**34%**
6	FSNet	E	14.71	11.19	11.32	9.50	9.85
	VAR+OL	E	24.15	12.65	13.65	11.48	16.35
	Δ%		**39%**	**12%**	**17%**	**17%**	**40%**
7	FSNet	E	10.21	5.69	5.95	4.08	6.21
	NBEATS+OL	EM	13.81	9.48	7.13	6.66	8.29
	Δ%		**26%**	**40%**	**17%**	**39%**	**25%**
8	FSNet	EM	5.27	11.80	12.50	10.46	8.10
	NBEATS+OL	EM	7.02	16.41	17.91	17.41	13.80
	Δ%		**25%**	**28%**	**30%**	**40%**	**41%**
9	FSNet	E	4.68	2.31	3.56	1.93	4.31
	NBEATS	EM	5.47	25.73	11.18	22.39	8.39
	Δ%		**14%**	**91%**	**68%**	**91%**	**49%**
10	FSNet	E	5.47	4.11	4.68	4.25	4.88
	VAR+OL	E	8.43	5.42	5.80	6.37	7.49
	Δ%		**35%**	**24%**	**19%**	**33%**	**35%**
11	FSNet	ET	4.10	2.15	2.17	1.70	2.42
	XGBoost	EM	4.80	21.16	20.04	22.95	16.49
	Δ%		**15%**	**90%**	**89%**	**93%**	**85%**
12	FSNet	ETM	0.22	0.26	0.47	0.27	0.20
	NBEATS+OL	ETM	0.34	0.43	0.78	0.46	0.27
	Δ%		**35%**	**40%**	**40%**	**41%**	**26%**
13	FSNet	E	16.88	15.05	15.17	12.97	12.22
	NBEATS+OL	E+M	22.09	20.04	18.36	17.16	12.74
	Δ%		**24%**	**25%**	**17%**	**24%**	**4%**

When mobility improves performance, How much improvements can we expect?

BC	Feat.	PLD		LD1		IL1		LD2		IL2	
3	E	17.8796	±0.4222	11.2482	±0.3301	10.2727	±0.3535	7.1119	±0.1823	13.5849	±0.3069
	EM	15.6269	±0.3064	10.3273	±0.2541	9.5919	±0.2462	7.1529	±0.1300	12.3164	±0.2794
	ΔM	**2.2527**	**12.60%**	**0.9209**	**8.19%**	**0.6807**	**6.63%**	**-0.0409**	**-0.58%**	**1.2686**	**9.34%**
4	E	16.8746	±0.2910	10.7414	±0.2562	12.5059	±0.2301	8.0558	±0.1168	13.9173	±0.2250
	EM	16.0373	±0.2219	9.4808	±0.1886	11.0637	±0.1255	6.9857	±0.1160	11.8021	±0.2188
	ΔM	**0.8373**	**4.96%**	**1.2606**	**11.74%**	**1.4422**	**11.53%**	**1.0701**	**13.28%**	**2.1152**	**15.20%**
5	ET	2.4337	±0.0145	2.4964	±0.0238	3.4632	±0.0457	1.9876	±0.0170	1.8229	±0.0187
	ETM	1.9634	±0.0241	1.9186	±0.0380	2.9512	±0.0500	1.6898	±0.0183	1.5648	±0.0306
	ΔM	**0.4703**	**19.32%**	**0.5778**	**23.14%**	**0.5121**	**14.79%**	**0.2977**	**14.98%**	**0.2581**	**14.16%**
8	E	5.7669	±0.0793	13.5915	±0.3982	14.4428	±0.3782	11.8113	±0.2523	9.0967	±0.1716
	EM	5.2652	±0.0548	11.7971	±0.2779	12.5034	±0.2087	10.4583	±0.0536	8.1016	±0.0706
	ΔM	**0.5017**	**8.70%**	**1.7943**	**13.20%**	**1.9395**	**13.43%**	**1.3530**	**11.46%**	**0.9952**	**10.94%**
12	ET	0.2271	±0.0015	0.2715	±0.0053	0.5005	±0.0087	0.2831	±0.0043	0.2080	±0.0022
	ETM	0.2199	±0.0018	0.2589	±0.0028	0.4717	±0.0024	0.2693	±0.0019	0.2015	±0.0008
	ΔM	**0.0072**	**3.18%**	**0.0126**	**4.64%**	**0.0289**	**5.77%**	**0.0138**	**4.88%**	**0.0065**	**3.11%**

UNSW SYDNEY

Thank you. Questions?

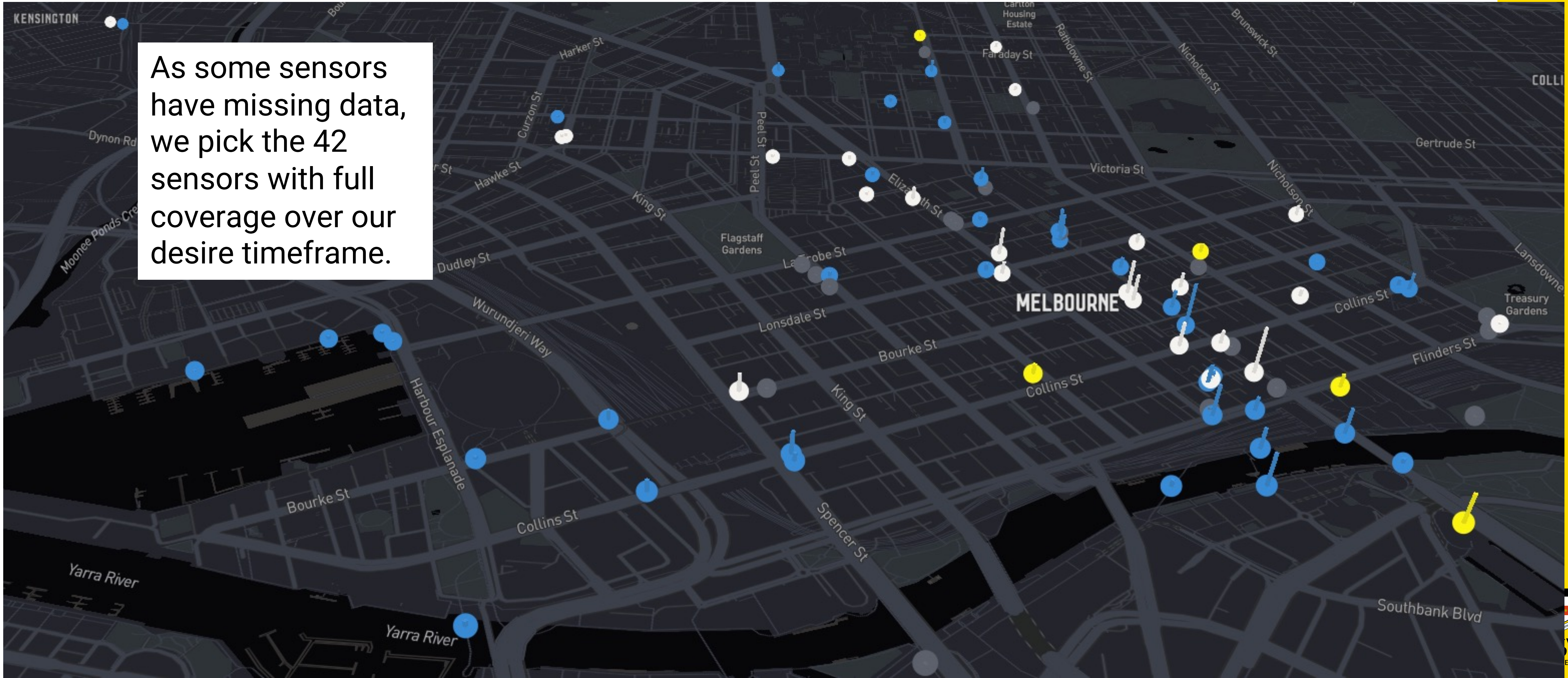
Conclusions:

1. FSNet is effective. But simple models VAR+OL is surprisingly effective. Implementing OL is more important than adopting more sophisticated DL architecture.
2. The effectiveness of mobility as a feature, is on par with temperature.
3. When mobility feature is influential during normal conditions, then it will still be useful during extreme events.

QR link to the complete results.

UNSW SYDNEY

Dataset: Pedestrian (42/79 sensors)



Experimental setup

Input
Observation window =
24 hours
E(T)(M)

Model

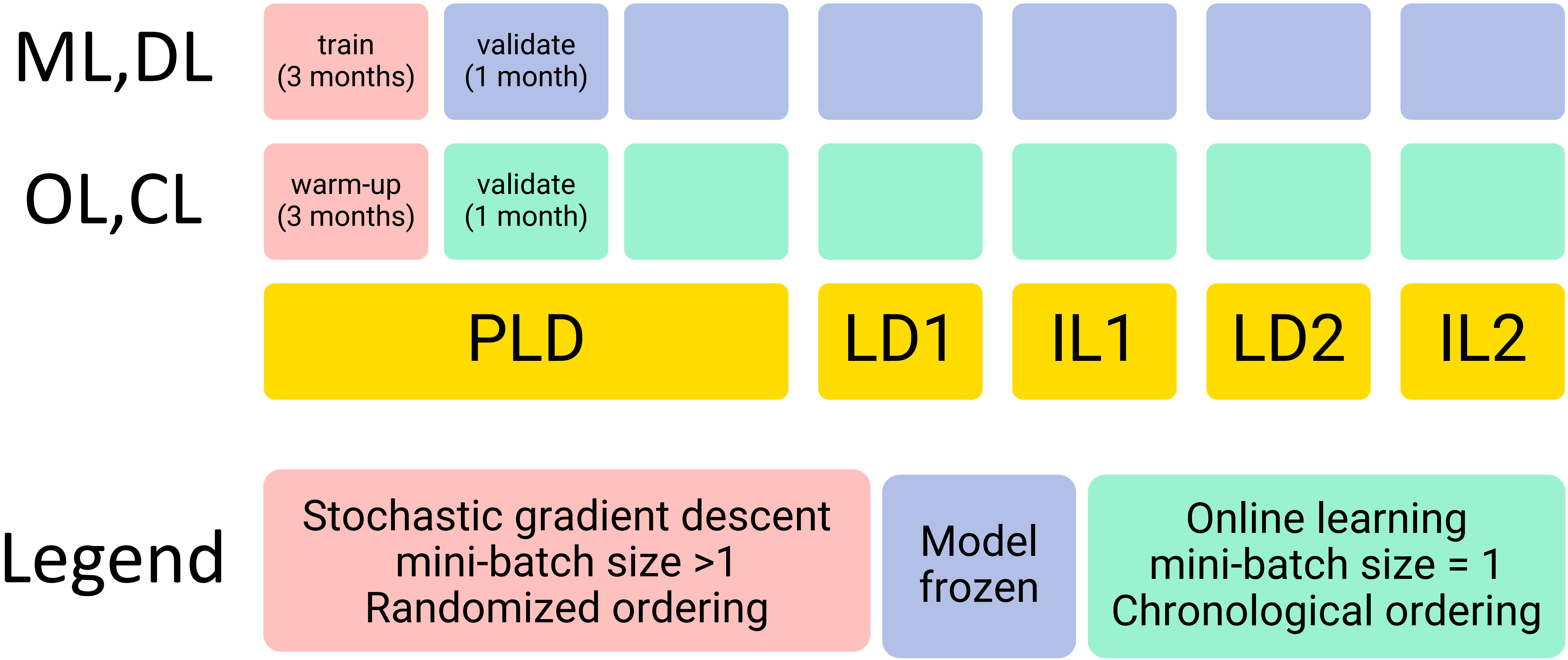
Output
Forecasting horizon =
24 hours
E

Error metric is MAE.

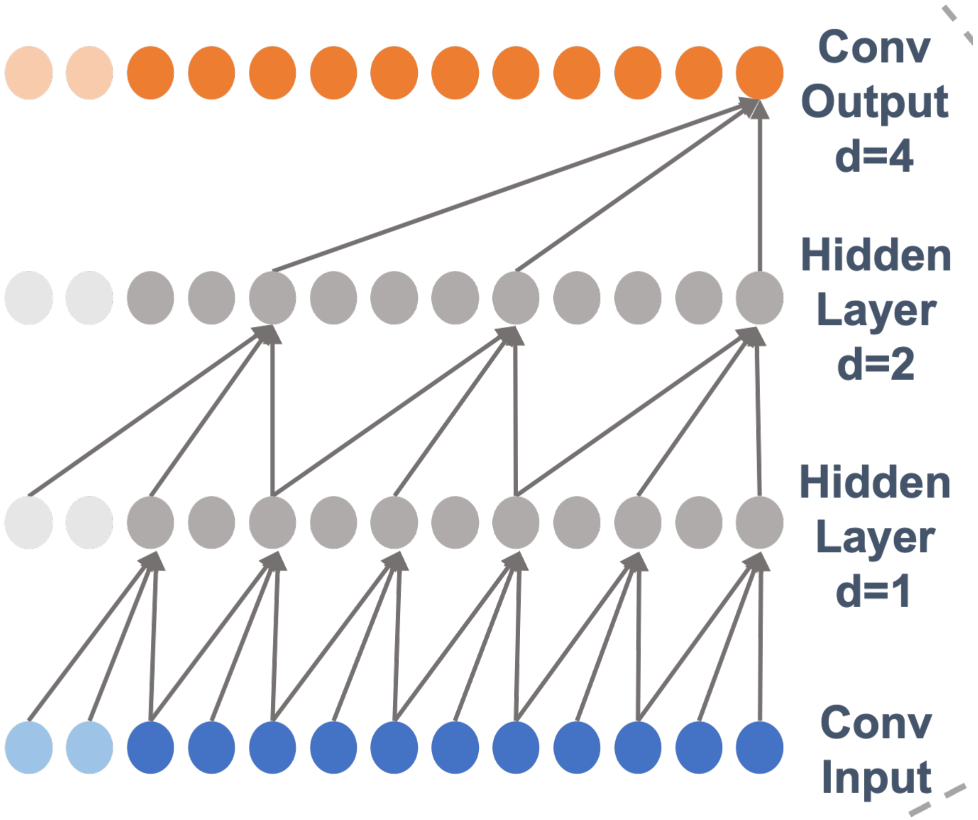
Averaged over
24 hours.

UNSW SYDNEY

Experimental setup



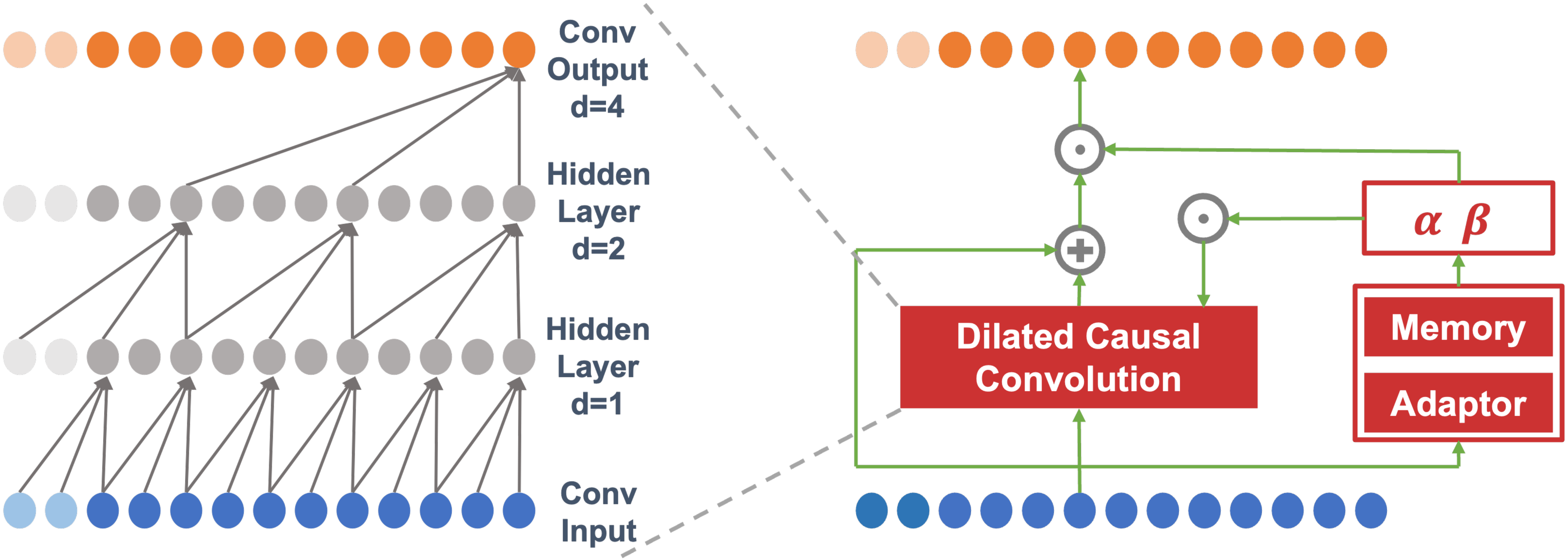
(Temporal Convolutional Network) TCN



(a) Dilated Causal Convolution

- TCN is a popular convolutional architecture in timeseries. (sometimes also known as WaveNet)
- The **dilation** operation exponentially increase receptive field of subsequent layers.
- The **causal** part means that the receptive field is only over preceding nodes, so they are not exposed to data from the future.

FSNet architecture with TCN backbone

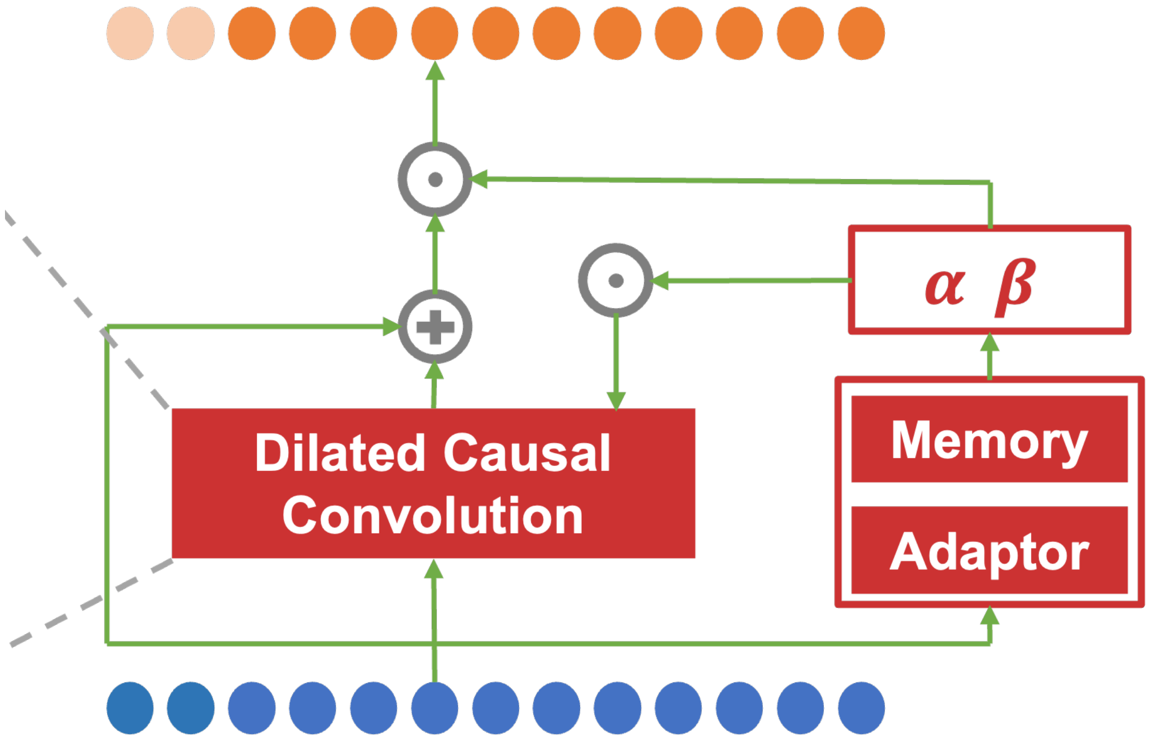


(a) Dilated Causal Convolution

(b) Convolution Layer with Adaptor and Memory

FSNet architecture with TCN backbone

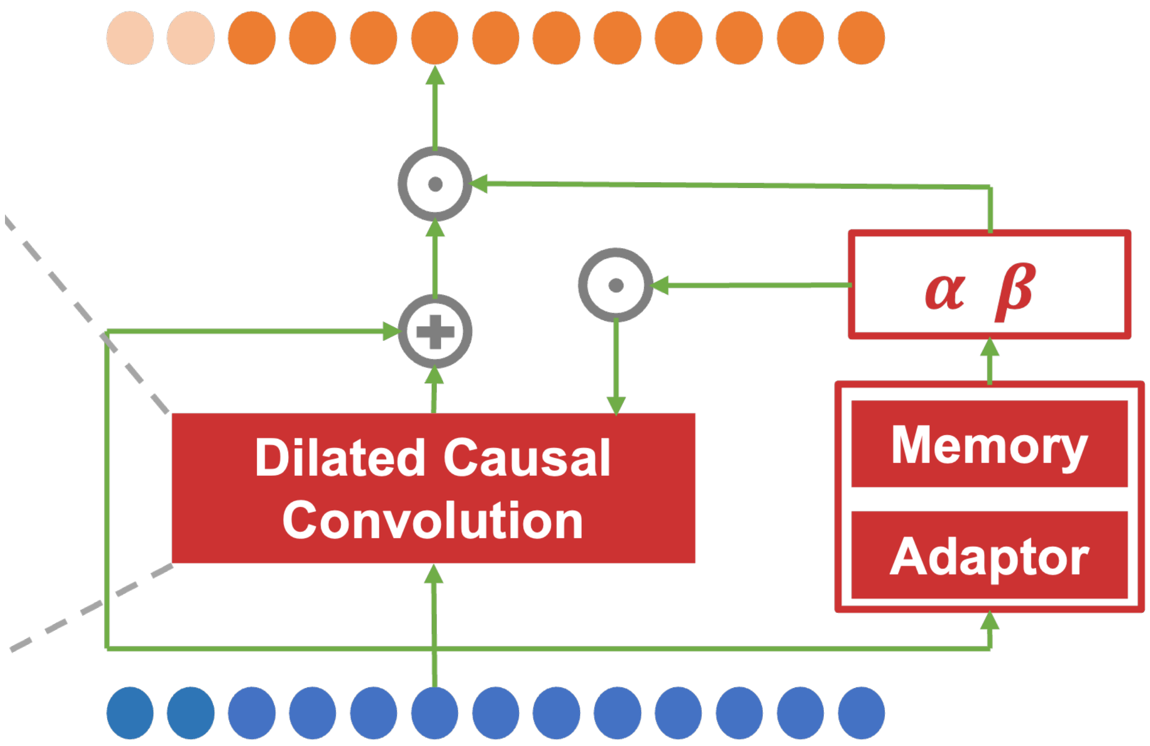
- Updating the entire model (OGD) whenever there's a new data is <u>**inefficient**</u>, and risk <u>**catastrophic forgetting**</u> where the model forget how to perform well on older task / distribution.



(b) Convolution Layer with Adaptor and Memory

FSNet architecture with TCN backbone

- Updating the entire model whenever there's a new data is inefficient, and risk catastrophic forgetting.
- So we only update the **memory module**.
- And the **adapter module** decides when to update the memory module.



(b) Convolution Layer with Adaptor and Memory