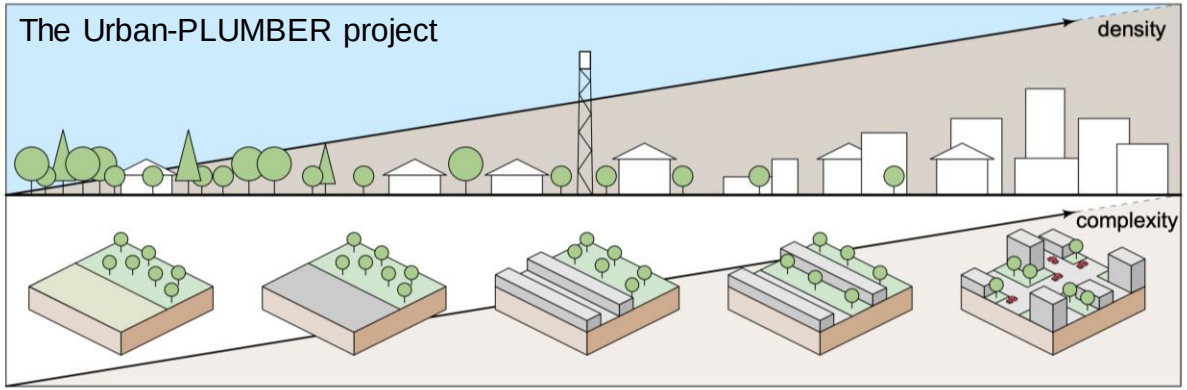
The Urban-PLUMBER land surface model evaluation project: Phase 1 results

mathew.lipson@bom.gov.au



Mathew Lipson (Bureau of Meteorology)

with Sue Grimmond (Reading University, UK), Martin Best (UK Met Office)

Alberto Martilli, Andreas Christen, Andres Simon, Andrew Coutts, Andy Pitman, Aristofanis Tsiringakis, Aude Lemonsu, Ben Crawford, Beom-Soon Han, Bert Heusinkveld, Chenghao Wang, Cécile De Munck, David Meyer, Doo-Il Lee, Erik Velasco, Gab Abramowitz, Gabriele Manoli, Gert-Jan Steeneveld, Helen Claire Ward, Hiroaki Kondo, Hirofumi Sugawara, Je-Woo Hong, Jinkyu Hong, Joseph McFadden, Joey McNorton, Jonathan Evans, Jong-Jin Baik, Keith Oleson, Kerry Nice, Keunmin Lee, Krzysztof Fortuniak, Leena Järvi, Maggie Hendry, Marcus Thatcher, Martin De Kauwe, Matthias Demuzere, Matthias Roth, Meiring Beyers, Michael Roth, Mikhail Varentsov, Naika Meili, Nektarios Chrysoulakis, Nigel Tapper, Oliver Michels, Robert Schoetter, Sang-Hyun Lee, Seung-Bu Park, Simone Fatichi, Simone Kotthaus, Souhail Boussetta, Stevan Earl, Sungsoo Jo, Tiago Machado, Ting Sun, Valéry Masson, Winston Chow, Wlodzimierz Pawlak, Yeon-Hee Kim, Yukihiro Kikegawa, Yuya Takane, Zhihua Wang

State of knowledge

PILPS-Urban (2011) key conclusions

Grimmond et al., (2011): https://doi.org/10.1002/joc.2227
Best and Grimmond (2015): https://doi.org/10.1175/BAMS-D-14-00122.1

Important for energy fluxes:

- land cover information
- vegetation/soil processes
- bulk albedo in day
- longwave trapping at night

- simpler models did well
  - more easily able to use provided information

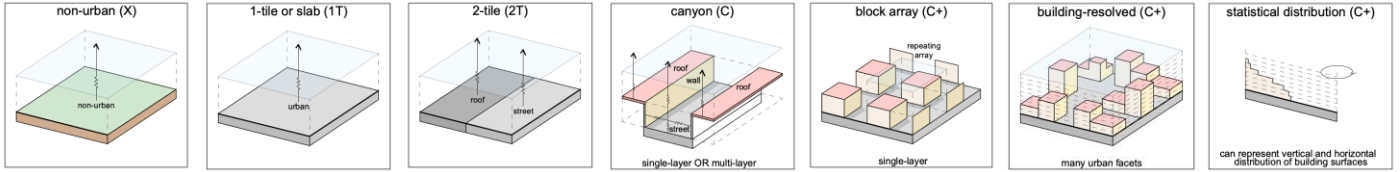
Site: Melbourne, Australia

Urban-PLUMBER: Phase 1

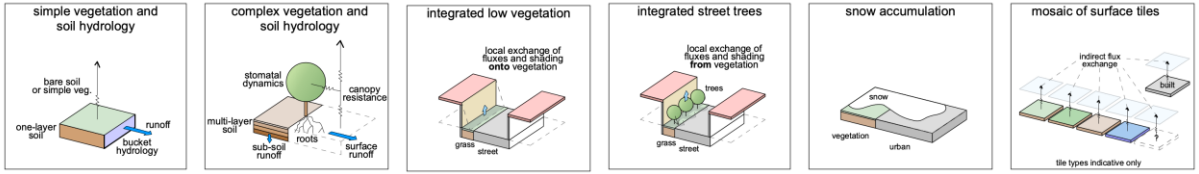
30 land surface models at same site (Preston, Melbourne)

ID	Submission name	Participating author(s)
01	ASLUMv2.0	Wang, Wang
02	ASLUMv3.1	Wang, Wang
03	BEPCOL	Simón-Moral, Martilli
04	CABLE	De Kauwe
05	CHTESSEL	McNorton, Boussetta
06	CHTESSEL_U	McNorton, Boussetta
07	CLMU5	Oleson
08	CM	Takane, Kondo
09	CM-BEM	Takane, Kikegawa
10	JULES_1T	Best
11	JULES_2T	Best
12	JULES_MORUSES	Hendry, Best
13	K-UCMv1	Beyers, Roth
14	Lodz-SUEB	Fortuniak
15	Manabe_1T	Best
16	Manabe_2T	Best
17	MUSE	Lee, Lee
18	NOAH-SLAB	Steeneveld, Tsiringakis
19	NOAH-SLUCM	Tsiringakis, Steeneveld
20	SNUUCM	Park, Baik
21	SUEWS	Sun, Blunn
22	TARGET	Nice
23	TEB-CNRM	Machado, de Munck, Schoetter, Masson, Lemonsu
24	TEB-READING	Meyer
25	TEB-SPARTCS	Machado, de Munck, Schoetter, Masson, Lemonsu
26	TERRA_4.11	Demuzere, Varentsov
27	UCLEM	Thatcher, Lipson
28	UT&C	Meili, Fatichi, Manoli, Bou-Zeid
29	VTUF-3D	Nice
30	VUCM	Lee, Han

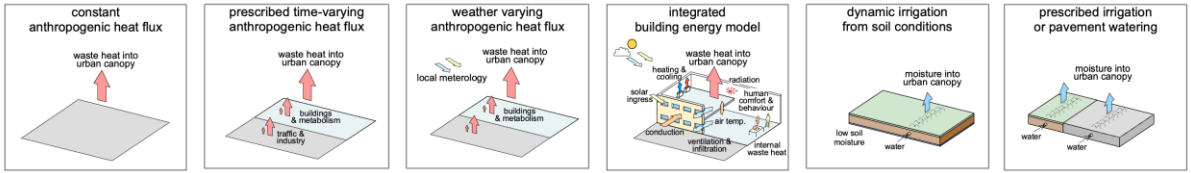
built representation



hydrological attributes

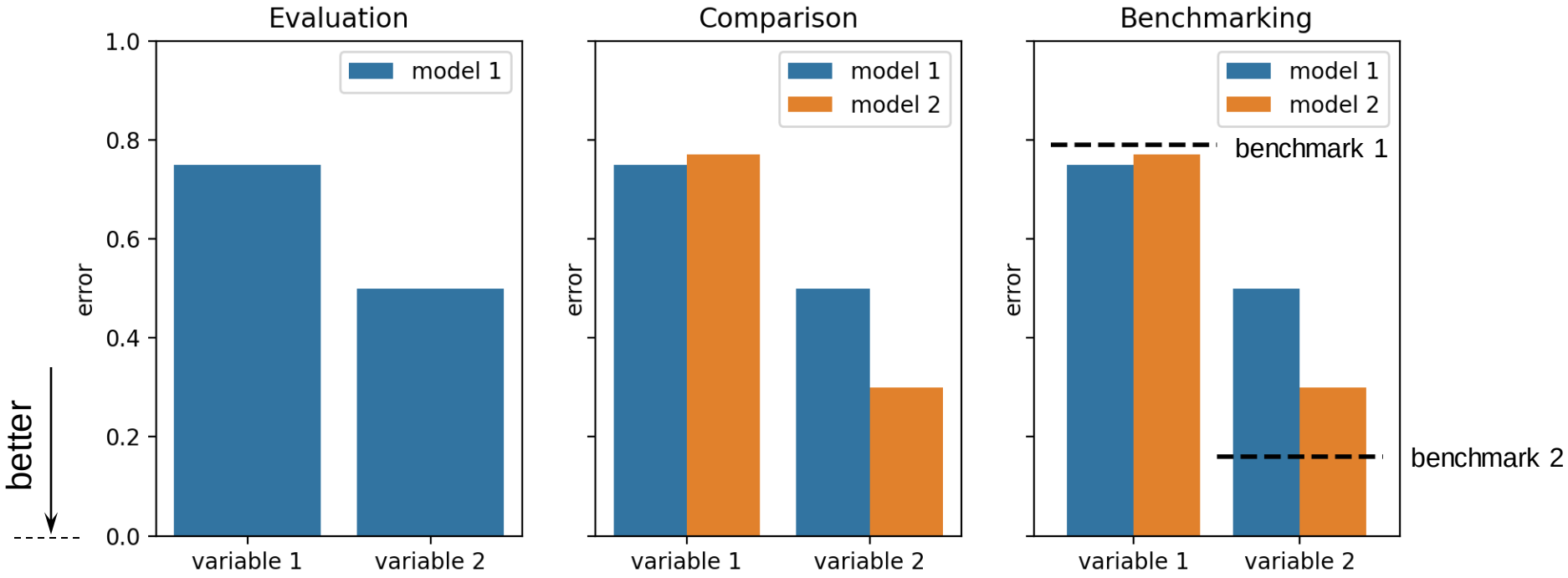


behavioural attributes



Evaluation

The benefits of benchmarking



Adapted from the PLUMBER project (Best et al., 2015: https://doi.org/10.1175/JHM-D-14-0158.1)

Evaluation

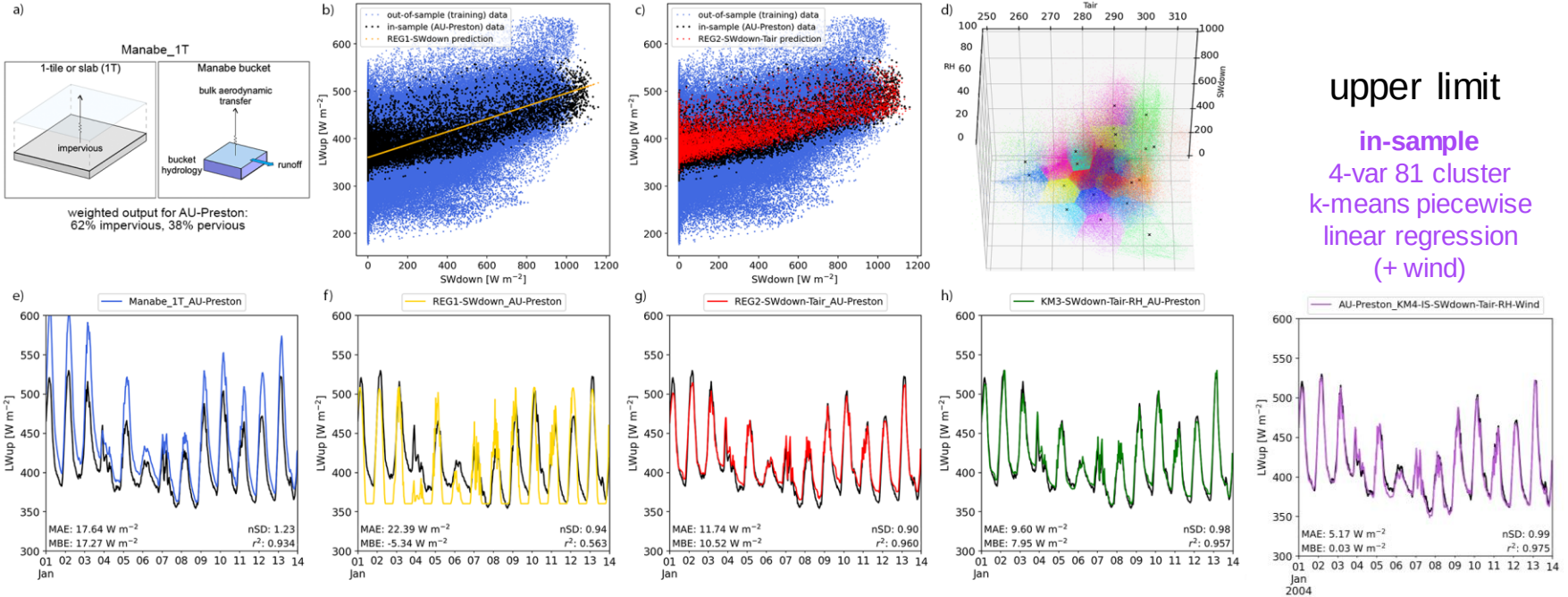
Urban-PLUMBER benchmarks

simple physically-based (slab and bucket)

1-var linear regression (SWdown)

2-var regression (SWdown-Tair)

3-var 27 cluster k-means piecewise linear regression (+ RH)



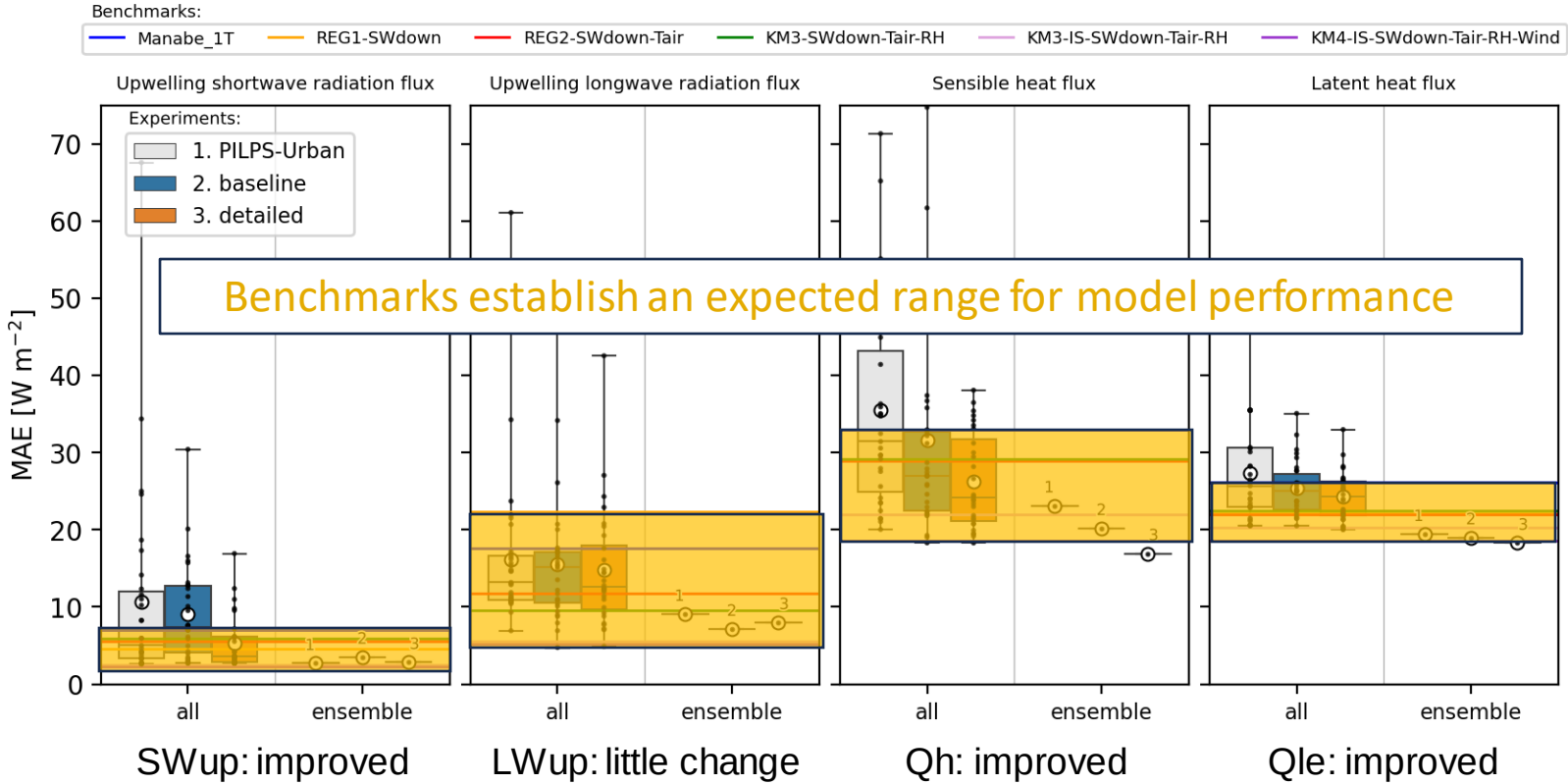
upper limit

in-sample
4-var 81 cluster k-means piecewise linear regression (+ wind)

subset of upward longwave radiation observations shown, error metrics are for full analysis period (474 days)

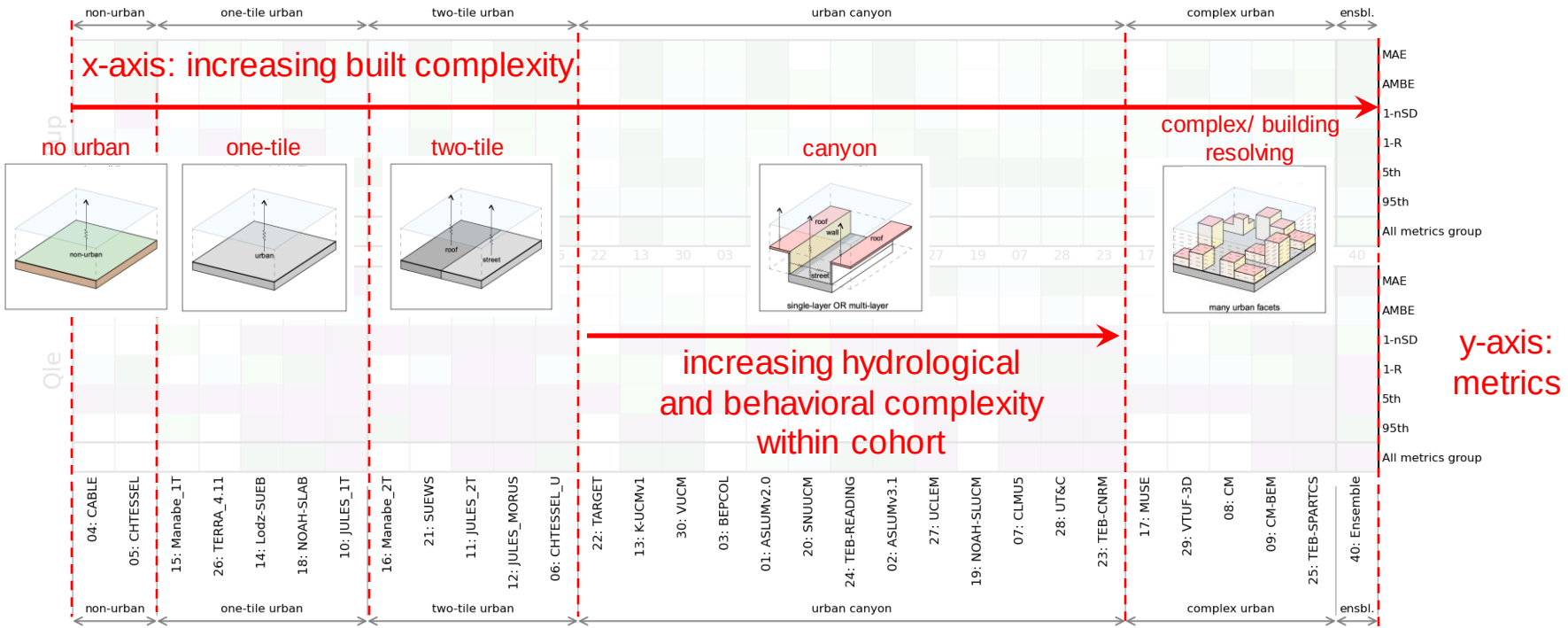
Results

Analysis: single metric comparison (mean absolute error)



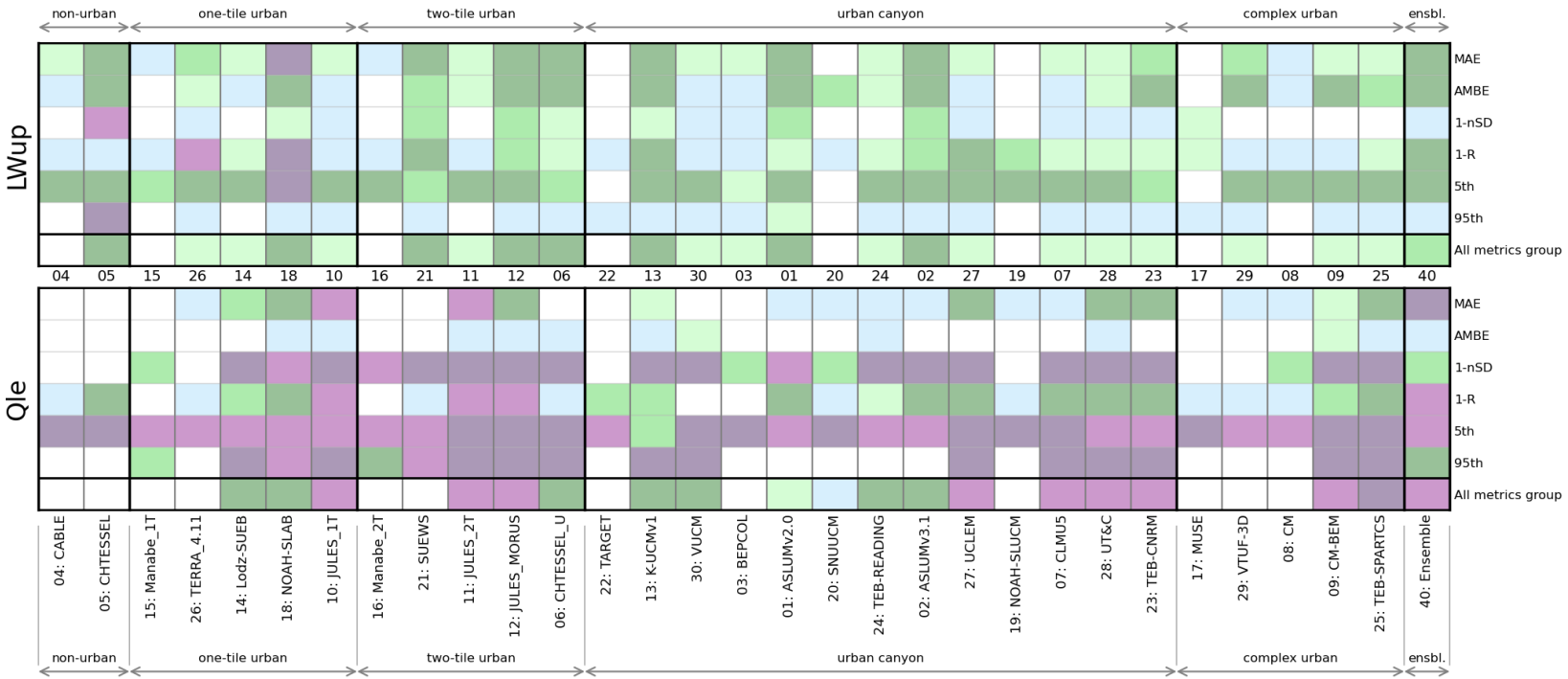
Results

Analysis: Multi-metric benchmarking (metrics: MAE, MBE, nSD, R, 5th, 95th)



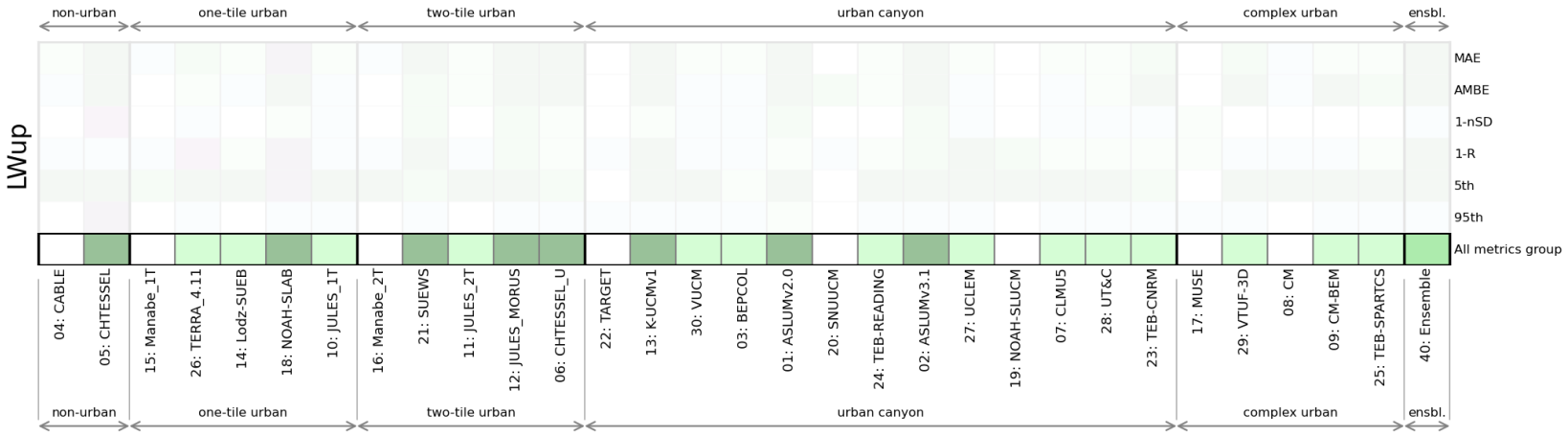
Results

Analysis: Multi-metric benchmarking (metrics: MAE, MBE, nSD, R, 5th, 95th)



Results

Analysis: Multi-metric benchmarking (metrics: MAE, MBE, nSD, R, 5th, 95th)



For upward longwave radiation flux (LWup):

- a few models do well (dark green) in each built cohort; except the most complex
- no clear pattern within cohorts

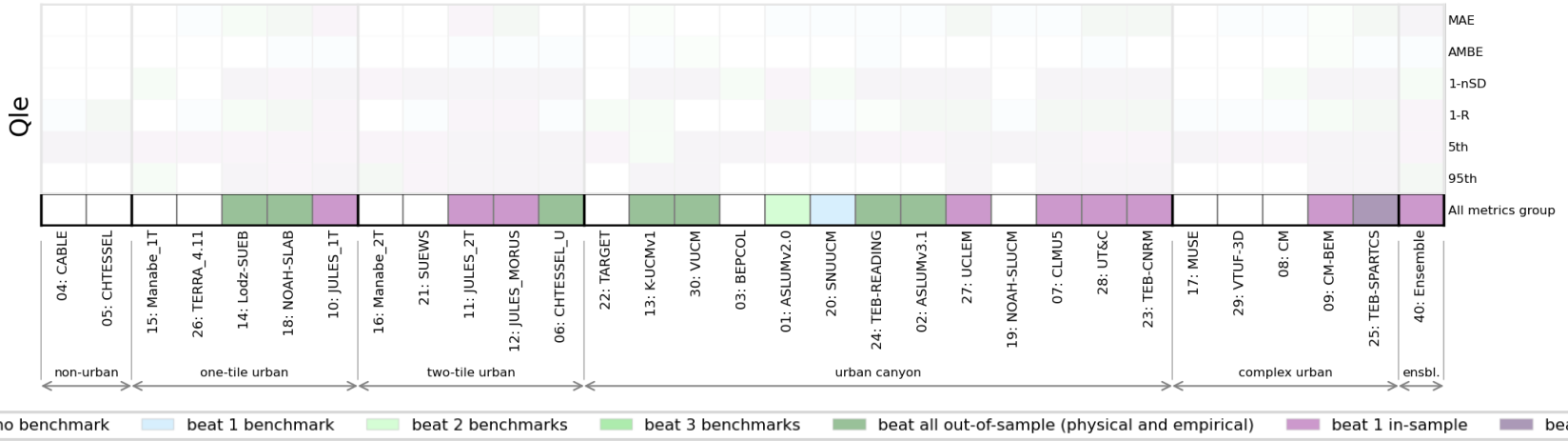
beat no benchmark | beat 1 benchmark | beat 2 benchmarks | beat 3 benchmarks | beat all out-of-sample (physical and empirical) | beat 1 in-sample | beat 2 in-sample

Results

Analysis: Multi-metric benchmarking (metrics: MAE, MBE, nSD, R, 5th, 95th)

For latent heat flux (Qle):

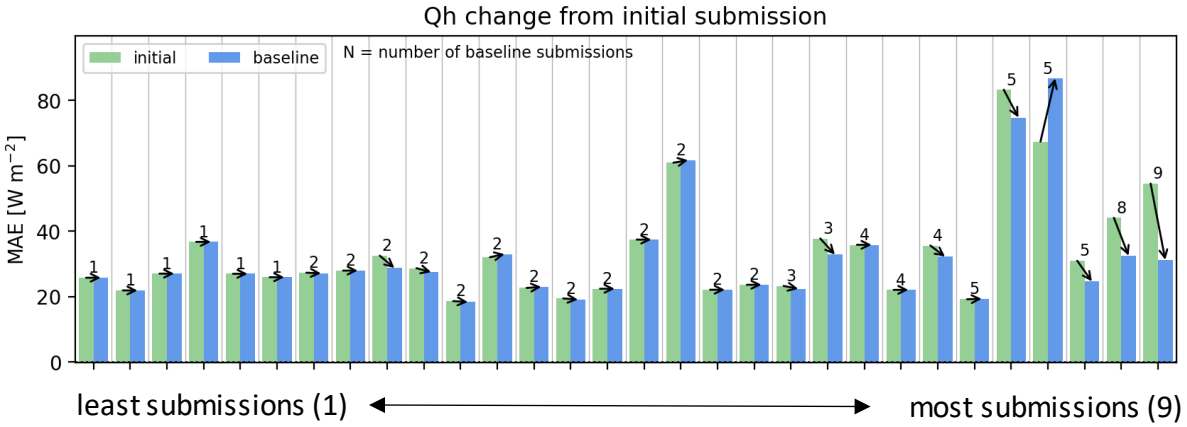
- many models do well (dark green or purple) in each built cohort; except non-urban
- within cohorts, more complex hydrological and behavioural attributes tend to help



Human factors

Model intercomparisons do not just test models!

- results are highly dependent on user configuration
- participants with more experience generally did better
- initial feedback and encouraging resubmissions helps level playing field
- try to avoid:
  - non-physical model behaviour
  - date/ time-of-day errors
  - i/o processing errors
  - variable labelling errors
  - forcing interpolation errors



Urban PLUMBER: Phase 2

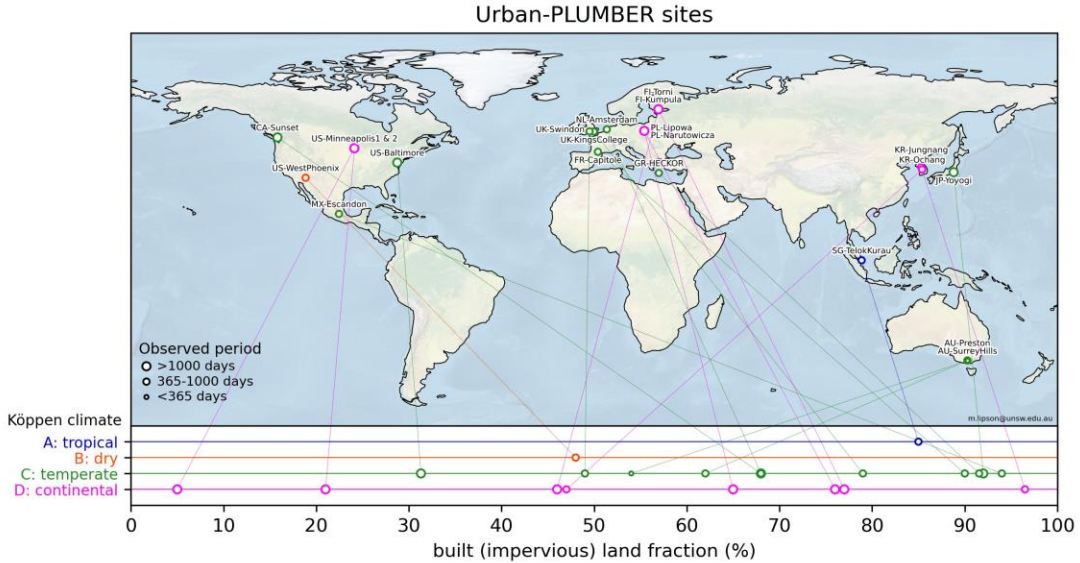
20 urban sites; 50 years of data

https://urban-plumber.github.io/sites

Earth Syst. Sci. Data, 14, 5157–5178, 2022
https://doi.org/10.5194/essd-14-5157-2022

Open Access Earth System Science Data

Harmonized gap-filled datasets from 20 urban flux tower sites

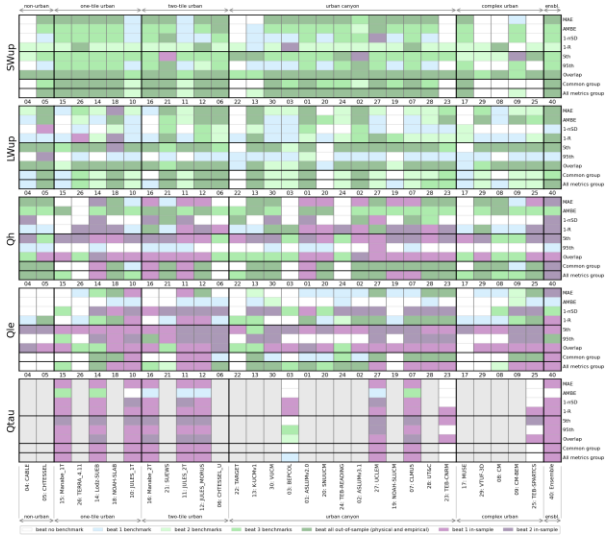


Conclusions

mathew.lipson@bom.gov.au

Urban-PLUMBER: Phase 1

- compared with PILPS-Urban:
  - **improved:** shortwave, sensible and latent heat fluxes
  - **little change:** longwave and momentum fluxes
  - **benchmarks:** show potential for improvement in LWup
- more complete hydrological and behavioural attributes help
  - **developments:** efforts in last decade appear beneficial
  - **human factors:** impactful, not just model physics
- observation and benchmark timeseries are openly available
- Phase 1 paper soon to be published in QJRMS



Urban-PLUMBER: Phase 2 (underway)

- 20 urban sites, from highly vegetated to highly urbanised

Thank you!