Supporting Information for "The Urban-PLUMBER land surface model evaluation project: the water balance representation"

H.J. Jongen[1,2], M. Lipson[3], A.J. Teuling[1], C.S.B Grimmond[4], J.-J. Baik[5], M. Best[6], M. Demuzere[7,8], K. Fortuniak[9], Y. Huang[10], M.G. De Kauwe[11], R. Li[12,13], J. McNorton[14], N. Meili[15,16], K. Oleson[17], S.-B. Park[18], T. Sun[12], A. Tsiringakis[2,19], M. Varentsov[20], C. Wang[10,21], Z.-H. Wang[22]and G.J. Steeneveld[2]

[1]Hydrology and Environmental Hydraulics, Wageningen University, Wageningen, The Netherlands.

[2]Meteorology and Air Quality, Wageningen University, Wageningen, The Netherlands.

[3]Bureau of Meteorology, Canberra, Australia.

[4]Department of Meteorology, University of Reading, Reading, United Kingdom.

[5]School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea.

[6]Met Office, Exeter, United Kingdom.

[7]Urban Climatology Group, Department of Geography, Ruhr-University Bochum, Bochum, Germany.

[8]B-Kode, Ghent, Belgium.

[9]Department of Meteorology and Climatology, Faculty of Geographical Sciences, University of Łódź, Łódź, Poland.

[10]School of Meteorology, University of Oklahoma, Norman, Oklahoma, United States of America.

[11]School of Biological Sciences, University of Bristol, Bristol, United Kingdom.

[12]Institute for Risk and Disaster Reduction, University College London, London, United Kingdom.

[13]Department of Hydraulic Engineering, Tsinghua University, Beijing, China.

[14]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, United Kingdom.

[15]Department of Civil and Environmental Engineering, National University of Singapore, Singapore, Singapore.

January 22, 2024, 3:39pm

[16]Future Cities Laboratory Global, Singapore-ETH Centre, Singapore, Singapore.

[17]U.S. National Science Foundation National Center for Atmospheric Research (NSF NCAR), Boulder, Colorado, United States of America.

[18]School of Environmental Engineering, University of Seoul, Seoul, South Korea.

[19]European Centre for Medium-Range Weather Forecasts (ECMWF), Bonn, Germany.

[20]Faculty of Geography/Research Computing Center, Lomonosov Moscow State University, Moscow, Russia.

[21]Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, Oklahoma, United States of America.

[22]School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, Arizona, United States of America

Contents of this file

1. Figures S1 to S6

Introduction This supplementary information contains seven additional figures further visualizing the analyses that we present in the paper. We show the auto-correlations for precipitation for all sites to determine the duration of no rain before a new event starts (Figure S1). Additionally, we include the relation between the impervious fraction of a site and the runoff ratio per model (Figure S2). We added the overviews of both the

---

Corresponding author: H.J. Jongen, Hydrology and Environmental Hydraulics Group, Wageningen University, Droevendaalsesteeg 3, Wageningen, 6708PB, The Netherlands. (harro.jongen@wur.nl)

curve number (Figure S3) and lag time (Figure S4) for all model runs. Next, we give an overview of the urban water balance representation scores averaged per site (Figure S5). We also present the evaluation of the latent heat flux compared to the UWBR score (Figure S6).

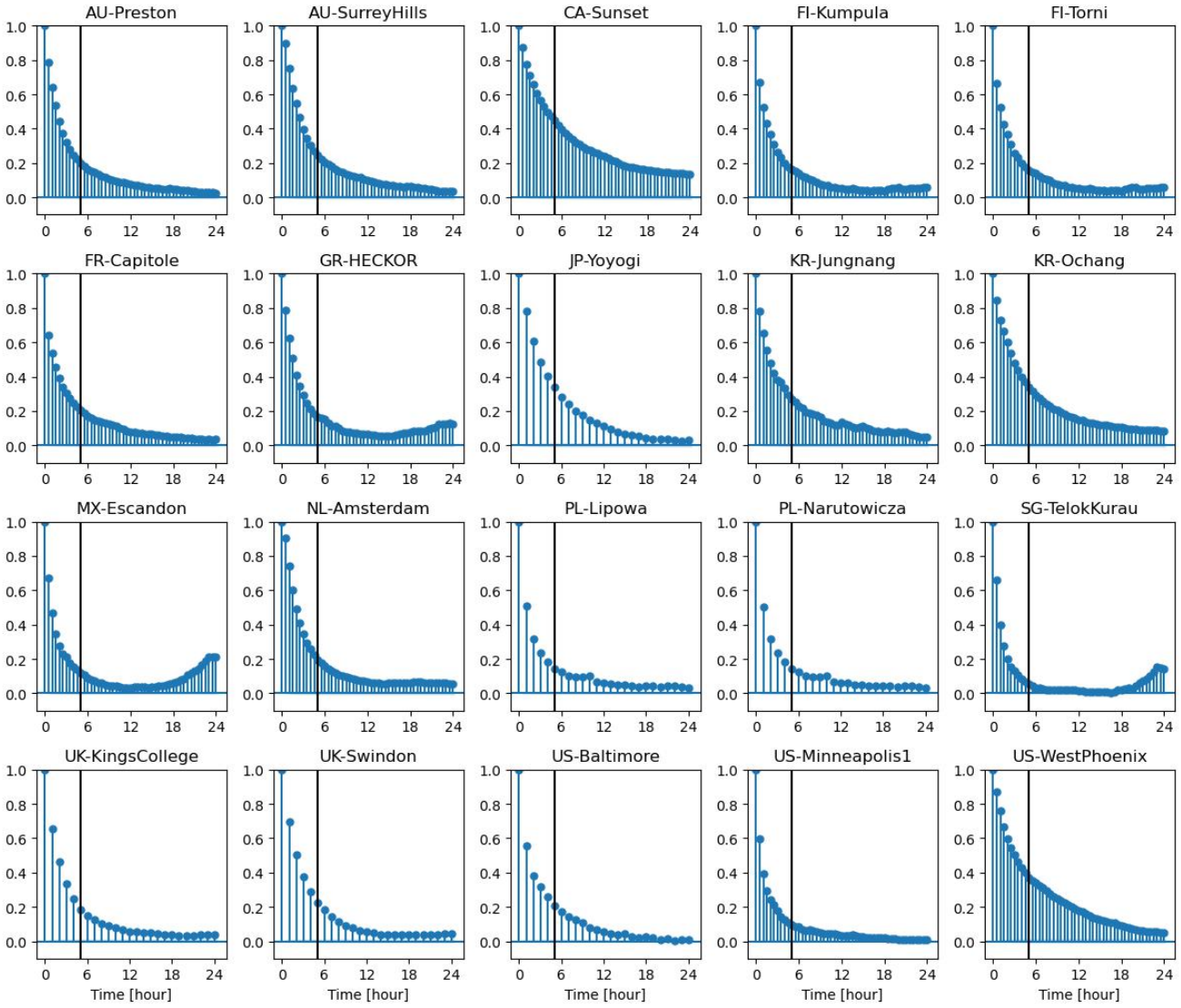

Figure S1. Auto-correlation of precipitation for all sites to determine separate events with the threshold of 5 hours (vertical line).

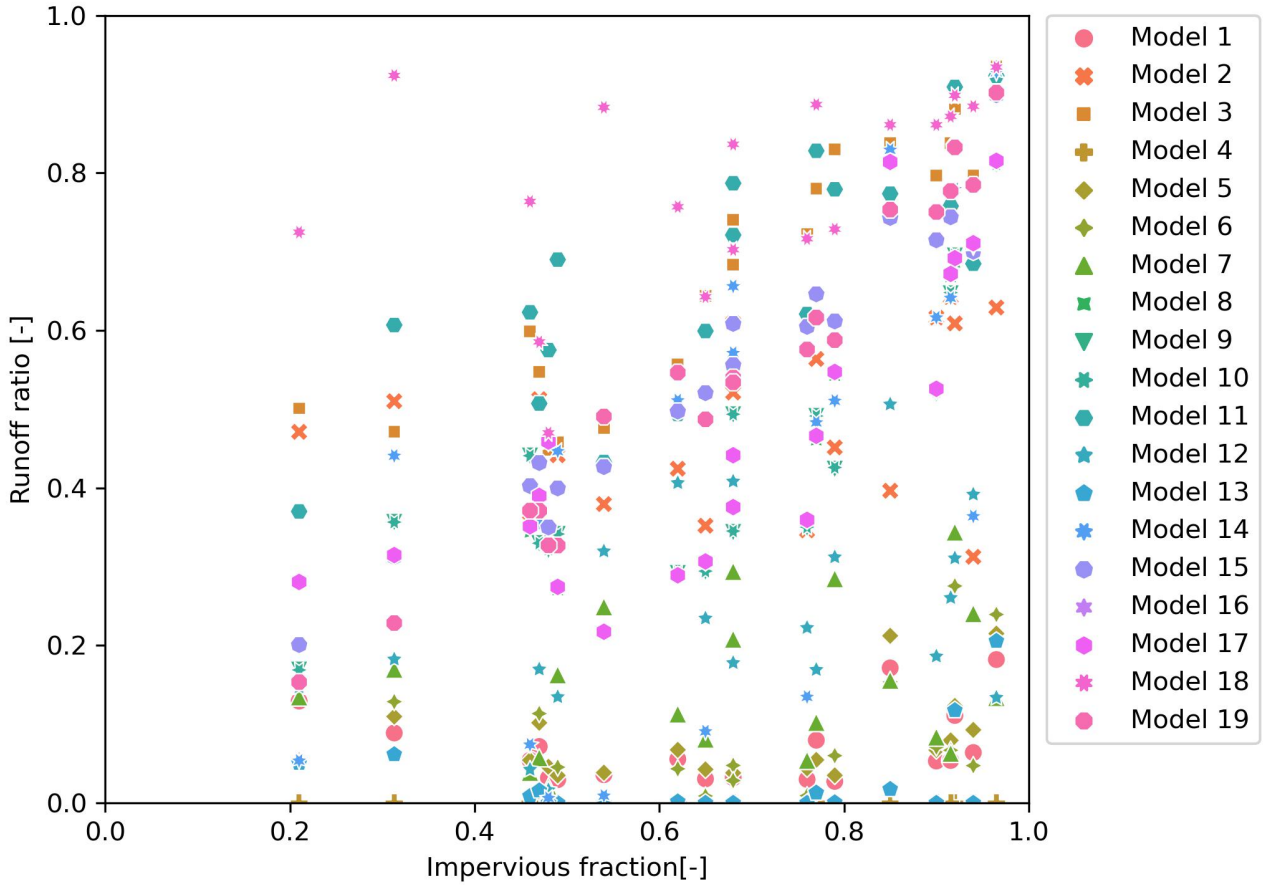

Figure S2. Relation between the site impervious fraction and runoff ratio per model run (R/P) for each model (marker).

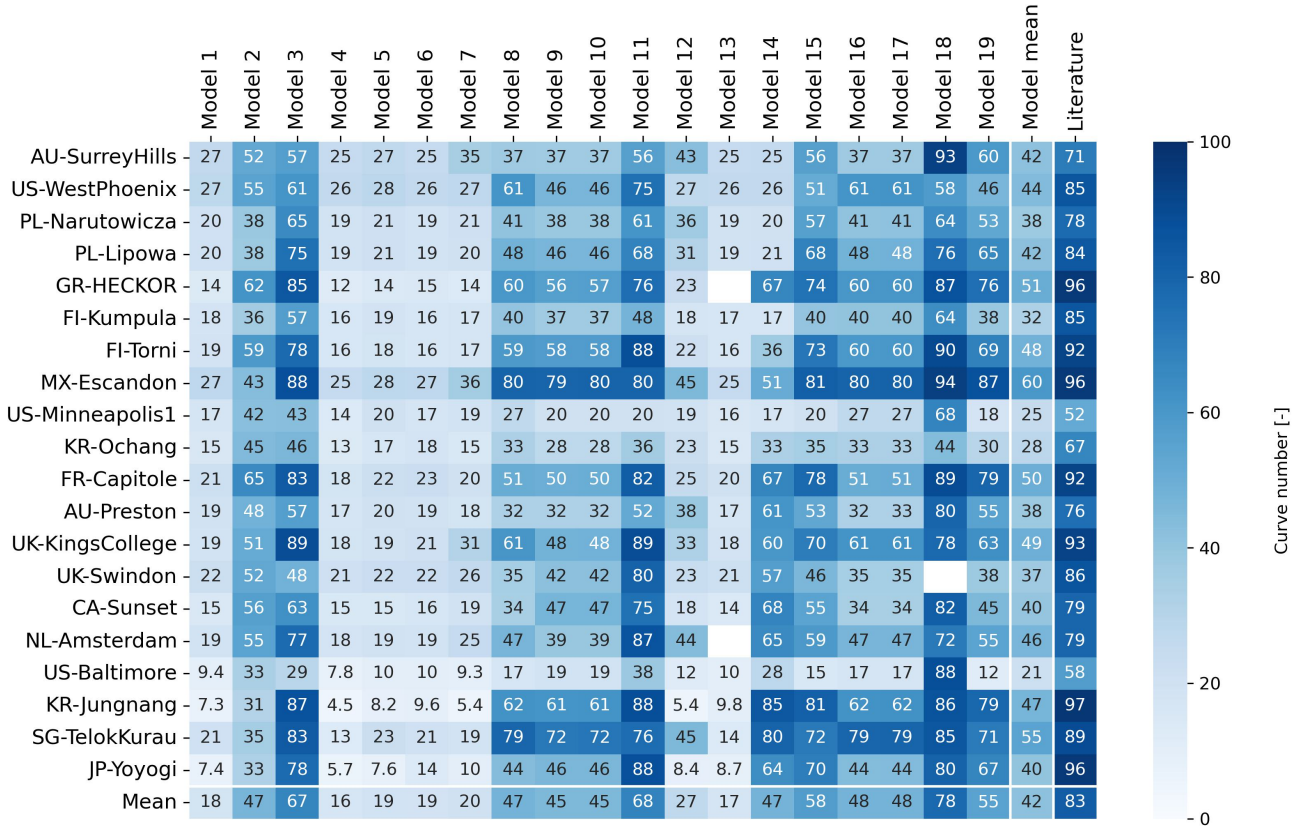

Figure S3. Curve numbers estimated from modeled runoff for each model run and from curve number reference tables in Croshney et al. (1985), see methods section.

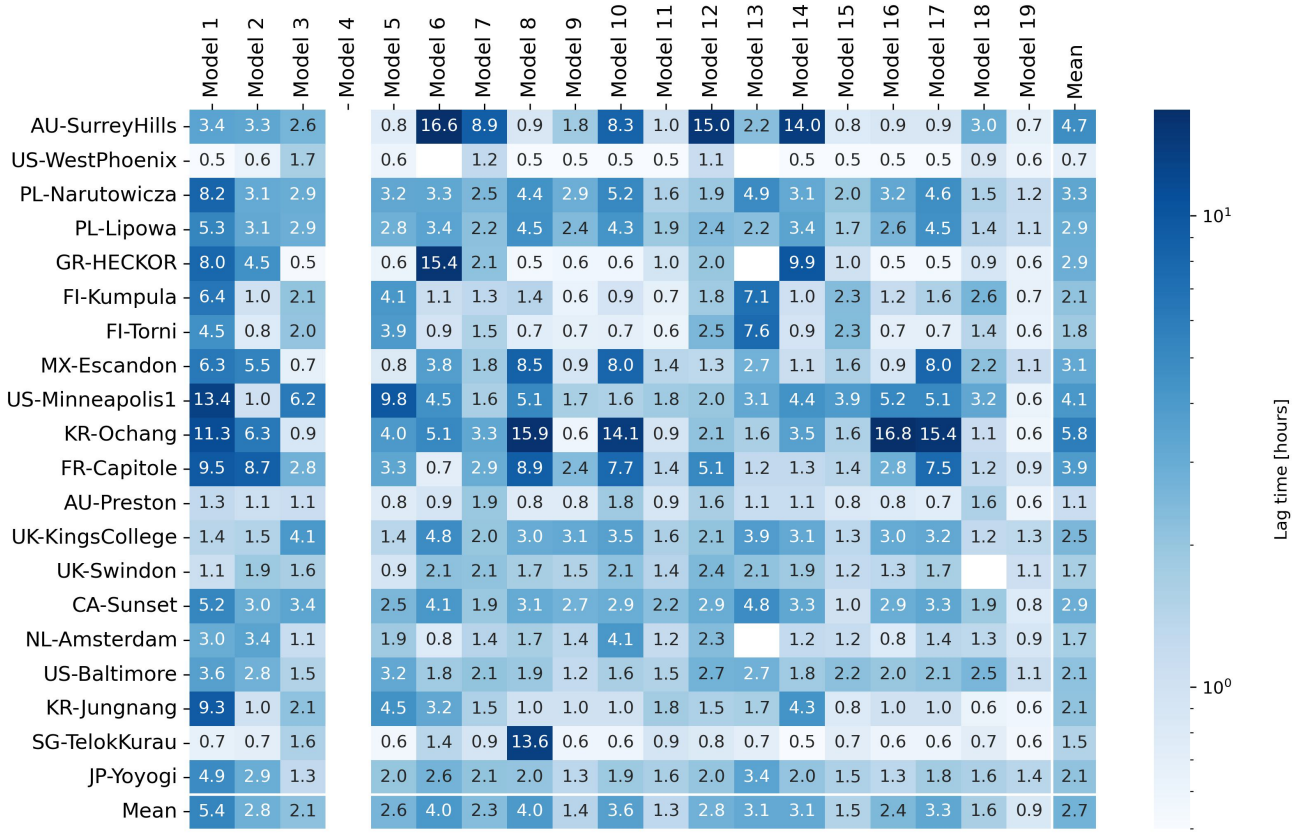

Figure S4. Average lag times (hours) for all model runs at each site. Note that the color scale is logarithmic.

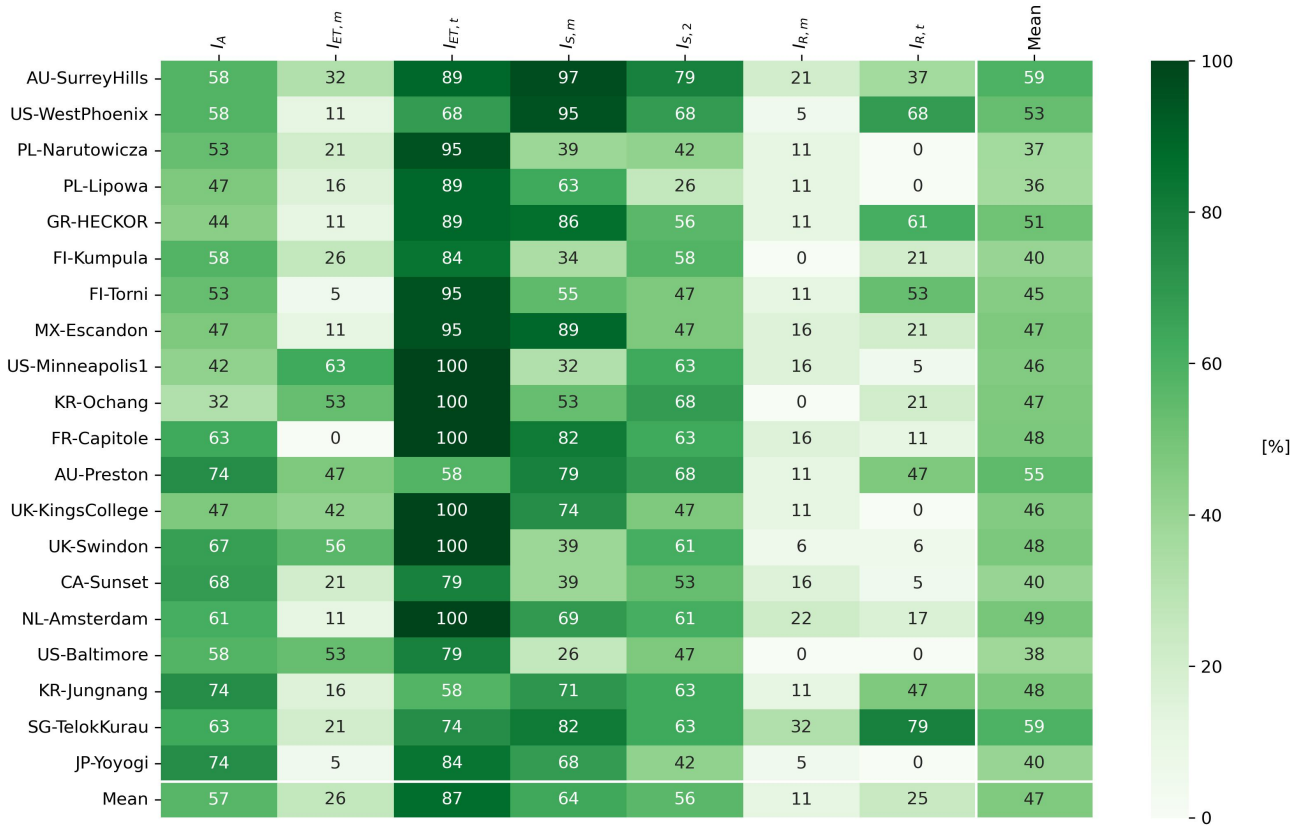

Figure S5. As Figure 4, but are averaged per site.

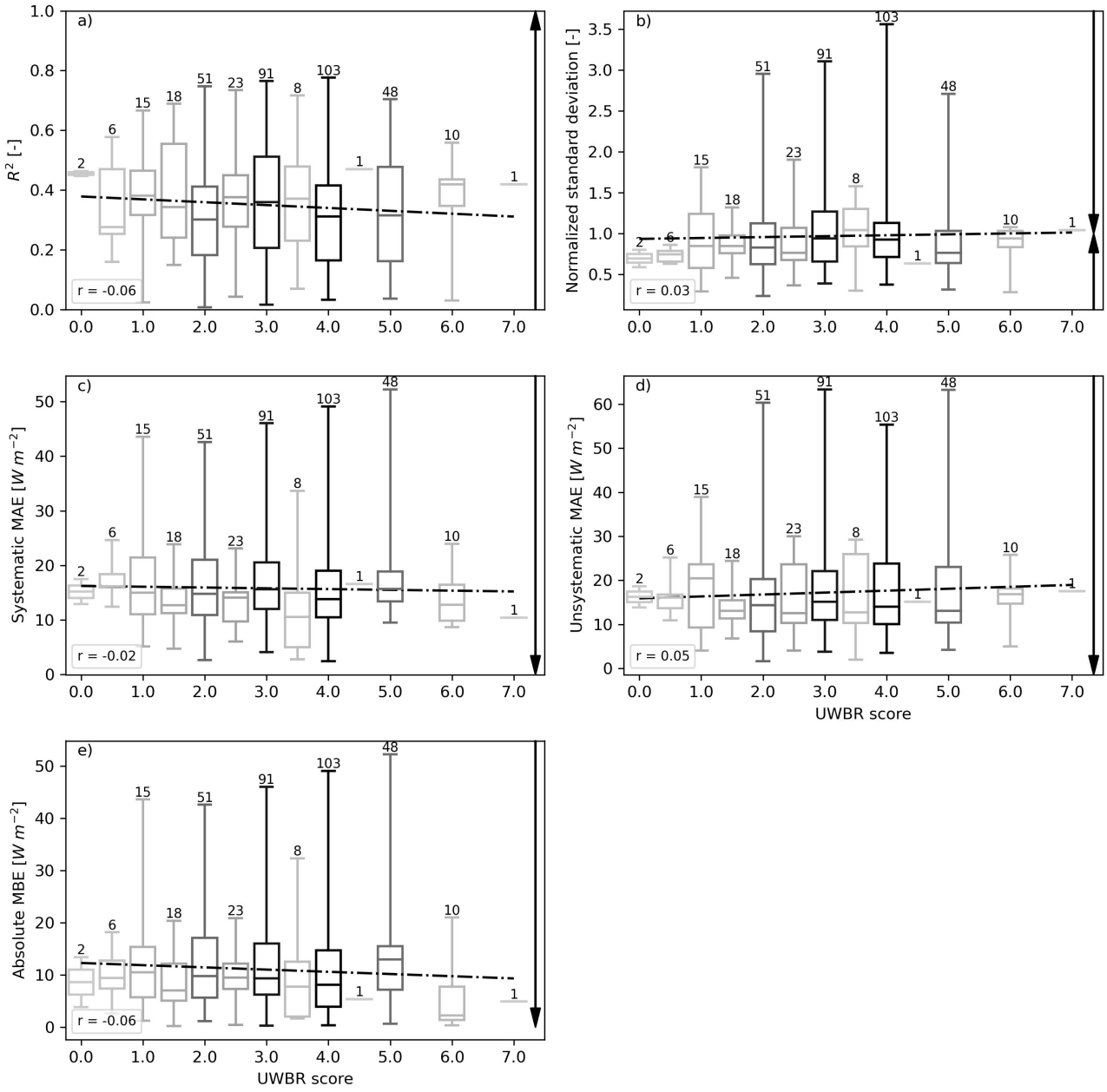

Figure S6. Relation between the urban water balance representation (UWBR) score (0.5) classes and (half-)hourly MBE for the modeled Q_E separated between (a) R^2, (b) normalized standard deviation, (c) systematic, (d) unsystematic MAE, and (e) absolute MBE. The number of model runs per class is given above the box. Whiskers indicate the minimum and maximum values. A Wald (1943) test indicates a significant correlation (*). Arrows indicate the direction of better model performance.