University of Reading

WRF-TEB: implementation and evaluation of the coupled Weather Research and Forecasting (WRF) and Town Energy Balance (TEB) model

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Meyer, D., Schoetter, R., Riechert, M., Verrelle, A., Tewari, M., Dudhia, J., Masson, V., Reeuwijk, M. and Grimmond, S. ORCID: https://orcid.org/0000-0002-3166-9415 (2020) WRF-TEB: implementation and evaluation of the coupled Weather Research and Forecasting (WRF) and Town Energy Balance (TEB) model. Journal of Advances in Modeling Earth Systems, 12 (8). e2019MS001961. ISSN 1942-2466 doi: 10.1029/2019MS001961 Available at https://centaur.reading.ac.uk/92181/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1029/2019MS001961

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other

University of Reading

copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2019MS001961

Key Points:
- WRF-TEB accounts for urban canopy energy exchanges, building energy demand, and neighborhood vegetation
- Coupled weather model development and scientific reproducibility are demonstrated
- The coupling is verified through an integration test and evaluated with meteorological observations

Correspondence to:
D. Meyer,
d.meyer@pgr.reading.ac.uk

Citation:
Meyer, D., Schoetter, R., Riechert, M., Verrelle, A., Tewari, M., Dudhia, J., et al. (2020). WRF-TEB: Implementation and evaluation of the coupled Weather Research and Forecasting (WRF) and Town Energy Balance (TEB) model. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001961. https://doi.org/10.1029/2019MS001961

Received 4 DEC 2019
Accepted 8 JUN 2020
Accepted article online 9 JUN 2020

©2020. The Authors.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

WRF-TEB: Implementation and Evaluation of the Coupled Weather Research and Forecasting (WRF) and Town Energy Balance (TEB) Model

D. Meyer[1,2], R. Schoetter[3], M. Riechert[4], A. Verrelle[3], M. Tewari[5], J. Dudhia[6], V. Masson[3], M. van Reeuwijk[2], and S. Grimmond[1]

[1]Department of Meteorology, University of Reading, Reading, UK, [2]Department of Civil and Environmental Engineering, Imperial College London, London, UK, [3]CNRM UMR 3589, Université Fédérale de Toulouse, Météo-France/CNRS, Toulouse, France, [4]Microsoft Research, Cambridge, UK, [5]IBM Thomas J Watson Research Center, Yorktown Heights, NY, USA, [6]Mesoscale and Microscale Meteorology Laboratory, NCAR, Boulder, CO, USA

Abstract Urban land surface processes need to be represented to inform future urban climate and building energy projections. Here, the single layer urban canopy model Town Energy Balance (TEB) is coupled to the Weather Research and Forecasting (WRF) model to create WRF-TEB. The coupling method is described generically, implemented into software, and the code and data are released with a Singularity image to address issues of scientific reproducibility. The coupling is implemented modularly and verified by an integration test. Results show no detectable errors in the coupling. Separately, a meteorological evaluation is undertaken using observations from Toulouse, France. The latter evaluation, during an urban canopy layer heat island episode, shows reasonable ability to estimate turbulent heat flux densities and other meteorological quantities. We conclude that new model couplings should make use of integration tests as meteorological evaluations by themselves are insufficient, given that errors are difficult to attribute because of the interplay between observational errors and multiple parameterization schemes (e.g., radiation, microphysics, and boundary layer).

Plain Language Summary With increasing urbanization and climate change, estimates of possible future urban climate and building energy scenarios are needed. Weather models (e.g., Weather Research and Forecasting; WRF) provide the state of the atmosphere, and urban land surface models (e.g., Town Energy Balance; TEB) allow the interactions and feedbacks of people—buildings—atmosphere to be investigated. WRF-TEB is a free and open-source model aimed at the urban climate and energy community to investigate applications such as the energy consumption associated with air conditioners, or the generation of solar energy in cities.

1. Introduction

With increasing urbanization (United Nations, 2019) and climate change (Collins et al., 2013), the study of urban atmospheric phenomena such as the spatial variation of temperature (Arnfield, 2003), or the impact of the urban environment on moisture (Unger, 1999), precipitation (Liu & Niyogi, 2019; Shepherd, 2005), wind fields (Martilli, 2002; Moonen et al., 2012), boundary layer (Chen et al., 2009; Lin et al., 2008; Masson, 2006), air conditioning (Salamanca et al., 2013, 2014; Takane et al., 2017), and heating energy demand (Santamouris et al., 2001), are increasingly relevant.

Coupled numerical weather prediction (NWP) and urban land surface models allow a diverse range of urban climate phenomena to be studied (e.g., Best, 2005; Chen et al., 2011; Hamdi et al., 2012). While NWP models simulate the prevailing meteorological conditions at kilometer resolution, land surface models (LSM) parameterize subgrid surface processes that are too small-scale, or (currently) too complex, to be explicitly modeled. Urban land surface models (ULSM) or urban canopy models (UCM), aim to capture the urban form (and sometimes function) created by buildings, roads, and vegetation.

UCMs, applicable to horizontal scales of the order 1–10 km, provide surface radiative and turbulent fluxes to NWPs. Parameters for the different aerodynamic (e.g., roughness length, drag force), radiative (e.g., albedo, emissivity), and thermal (e.g. heat capacity, conductivity) processes are required to capture radiative shading and trapping, large storage heat fluxes, and strongly varying turbulent heat fluxes. The energy balance may

Table 1
Models and Software Used in the Coupling, Where "Model" Refers to the Science (i.e., as Outlined in the Literature), "Software" Refers to the Actual Software and "Version" the Exact Software-Version Used in Running a Simulation

Model	Reference	Software	Reference	Version	Reference
WRF	Skamarock et al. (2019)	WRF-CMake	Riechert and Meyer (2019a)	4.1.5	Riechert and Meyer (2020)
TEB	Masson (2000)	TEB	Meyer et al. (2020)	4.0.1	Masson et al. (2020)

be solved for individual facets (e.g., roof, walls, roads) with different levels of complexity (Grimmond et al., 2010, 2011).

A coupled NWP-UCM may treat the surface as a single vertical layer (single layer UCM), where the entire urban canopy layer is collapsed into a single point, or as multiple layers (multilayer UCM), where the UCM is "immersed" in the NWP to account for the interaction between bluff-bodies (e.g., buildings) and the atmosphere (Chen et al., 2011). The assumptions and simplifications can also vary from treating buildings as being arranged to create infinitely long street canyons (e.g., Kondo et al., 2005, Kusaka et al., 2001; Martilli et al., 2002; Masson, 2000), or as cuboids (e.g., Mills, 1997).

Complex models may not perform systematically better than simpler ones (Grimmond et al., 2010, 2011). However, simpler models tend to lack features, thus limiting the study of specific urban climate processes (e.g., estimation of building energy consumption or details of it such as air conditioners (AC) energy demand, energy production from solar photovoltaic panels (PV)), which may be of interest to the broad urban climate community.

Previous NWP-UCM coupling work implemented and evaluated the linkage between the single layer Town Energy Balance (TEB Masson, 2000, and subsequent papers) and NWPs (e.g., Freitas et al., 2007; Hamdi et al., 2012; Lemonsu & Masson, 2002; Lemonsu et al., 2009; Rozoff et al., 2003) or the Weather Research and Forecasting (WRF Skamarock et al., 2019) model and UCMs (e.g., Kikegawa et al., 2014; Martilli et al., 2002). However, to our knowledge, none have employed tests to verify the effects of coupling alone.

Here we both outline a technical approach to couple and verify model components and link TEB with WRF to add to other UCM options already available in WRF (UCAR, 2020). To date, these are: bulk urban parameterization within the Noah-LSM (Chen & Dudhia, 2001; Ek et al., 2003), single layer Urban Canopy Model (SLUCM Chen et al., 2011), and the multi-layer Building Effect Parameterization (BEP Martilli et al., 2002) with optional Building Energy Model (BEP+BEM Salamanca et al., 2010). Given the conclusions from (Grimmond et al., 2010, 2011) model comparison, we do not expect WRF-TEB to perform systematically better than other models currently available in WRF but we have undertaken this work to (a) offer researchers and practitioners a greater number of features currently unavailable in other models (section 2.2), (b) simplify the evaluation of offline and online TEB in future research and comparison projects (section 4.1), and (c) simplify the integration with future TEB developments (section 4.1).

We describe the coupling between WRF and TEB (section 2) both conceptually (section 3) and technically (section 4) in a way that may be generalizable beyond the scope of WRF (or WRF-TEB). We release the complete source code, data, and tools to make our results reproducible (section 5) and evaluate the model (section 6) with a technical integration test (section 6.3) and meteorological observations (section 6.4).

2. Models and Software

2.1. Weather Research and Forecasting (WRF)

WRF is a popular atmospheric model used in research and NWP applications (Powers et al., 2017). It has been developed under two variants: the Advanced Research WRF (ARW Skamarock et al., 2008, 2019), and the Nonhydrostatic Mesoscale Model (NMM Janjic et al., 2001; Janjic, 2003). The support for the latter recently ended (see Developmental Testbed Center, 2018). Here, we exclusively refer to the WRF-ARW variant (as WRF).

WRF-TEB is developed using WRF-CMake version 4.1.5 (Riechert & Meyer, 2019a, 2020; Table 1) as it adds CMake (Kitware Inc., 2019a) support to the latest versions of WRF to simplify the configuration and build process of WRF and WPS (WRF Preprocessing System). Although, WRF-CMake version 4.1.5 does not

include support for WRF-Chem (Grell et al., 2005), WRF-DA (Huang et al., 2009), WRFPLUS (Guerrette & Henze, 2015), or WRF-Hydro (Gochis et al., 2018), its benefits may outweigh these limitations to model developers, code maintainers, and end-users wishing to build WRF, as it includes: robust incremental rebuilds, dependency analysis of Fortran code, flexible library dependency discovery, integrated support for shared (Open Multi-Processing; OpenMP) and distributed (Message Passing Interface; MPI) memory, support for automated testing using continuous integration (CI), and availability of experimental prebuilt binary releases for Linux, macOS, and Windows from the project's GitHub page or through the integration with GIS4WRF (Meyer & Riechert, 2019a), a QGIS (QGIS Development Team, 2019) toolkit for preprocessing and postprocessing, visualizing, and running simulations in WRF. Here we refer to both the physical model and the software (i.e., WRF-CMake) as WRF, unless highlighting specific software features.

2.2. Town Energy Balance (TEB)

The physically based single-layer UCM TEB (Masson, 2000) characterizes cities by their surface area of building roofs, walls, roads, and integrated vegetation using a simplified infinite street canyon geometry. The energy balance of impervious and pervious (vegetation) surfaces are calculated independently before being aggregated. To characterize the urban area, TEB requires a surface fraction of vegetation/garden, building, and road area, building height and vertical to horizontal surface ratio. For the calculation of shadowing effects and radiative trapping, the street orientation is assumed isotropic.

The outer surface of each facet is assumed to be sufficiently thin that the layer-averaged temperature can be used to determine the radiative and turbulent surface flux densities (i.e., the impervious skin temperature equals that of first layer-averaged temperatures). Thermal diffusion into materials is calculated using the thermal properties and thickness of the specified layers. The momentum flux is calculated for the whole canopy using a representative roughness length of the city (at model grid point scale), whereas thermal and hydrological fluxes for impervious areas are computed using an aerodynamic resistances network that considers local energy exchange within and above the canyon. Turbulent exchanges inside the urban canyon, and those between the canyon and the atmosphere above, depend on an aerodynamic resistances network with exchange coefficients that depend on wind speed and stability conditions (see Figure 1 in Lemonsu et al., 2004). Other TEB original features include the following: a water reservoir on roofs and road, and a snow mantel on roofs and roads (Masson, 2000), but more recent TEB developments now also include the following:

1. Building Energy Model (BEM Bueno et al., 2012): internal building energy balance (indoor air, floor, and internal mass), windows, heat-ventilation-air-conditioning (HVAC), infiltration, shading devices, and natural ventilation (opening of windows).
2. Road orientation (Lemonsu et al., 2012): specified road orientation, and separate energy balance for adjacent walls.
3. Gardens (Lemonsu et al., 2012): vegetation inside canyons.
4. Green roofs (de Munck et al., 2013).
5. Human behavior related to building energy consumption (Schoetter et al., 2017).
6. Calculation of urban carbon dioxide fluxes (Goret et al., 2019).
7. Irrigation (de Munck et al., 2013): irrigation of green roofs, gardens, and watering of roads.
8. Solar panels (Masson et al., 2014) for hot water and/or photovoltaic (PV).

To implement WRF-TEB, TEB version 4.0.1 (Masson et al., 2020; Meyer et al., 2020; Table 1) is used as it includes MinimalDX (Meyer & Raustad, 2020) to improve the modeling of air conditioners (AC), and Psychrolib (Meyer & Thevenard, 2019) to calculate psychrometric functions. Furthermore, support for Linux, macOS, and Windows with CMake allows direct integration in WRF-CMake. In TEB 4.0.1, Features 3 and 4 use a simplified vegetation scheme with a fixed albedo and Bowen ratio, whereas Lemonsu et al. (2012) treats the vegetation by coupling to ISBA (Interaction Soil Biosphere Atmosphere; Noilhan and Planton (1989); see simplifications in section 3). Furthermore, feature 5 and 6 are not available in TEB 4.0.1.

3. Coupling Approach

TEB is coupled to WRF following the generalized coupling methodology described in Best et al. (2004), where atmospheric quantities from the NWP's lowest model level are passed to the LSM to improve the

Table 2
Variables Used in the Coupling With Outputs (O) and Inputs (I) and Units

Symbol	Name	Unit	WRF	TEB
(a) WRF outputs, TEB inputs				
T	Dry-bulb air temperature	K	O	I
p_0	Atmospheric pressure at the surface[a]	Pa	O	I
p	Atmospheric pressure at forcing level	Pa	O	I
r	Mass mixing ratio of water vapor	kg kg^{-1}	O	I
u	Zonal component of wind velocity	m s^{-1}	O	I
v	Meridional component of wind velocity	m s^{-1}	O	I
$S^{\downarrow\downarrow}$	Downwelling direct shortwave radiation flux density	W m^{-2}	O	I
$S^{\downarrow\Downarrow}$	Downwelling diffuse shortwave radiation flux density	W m^{-2}	O	I
$L^{\downarrow}$	Downwelling longwave radiation flux density	W m^{-2}	O	I
RR	Rainfall rate	kg m^{-2} s^{-1}	O	I
SR	Snowfall rate	kg m^{-2} s^{-1}	O	I
Z	Solar zenith angle	rad	O	I
Δt	Time step	s	O	I
t_{local}	Current (local) time	s	O	I
(b) WRF inputs, TEB outputs				
Q_H	Turbulent sensible heat flux density	W m^{-2}	I	O
Q_E	Turbulent latent heat flux density	W m^{-2}	I	O
E	Evaporation mass flux density	kg m^{-2} s^{-1}	I	O
Q_G	Ground heat flux density	W m^{-2}	I	O
α	Surface albedo	1	I	O
ε	Surface emissivity	1	I	O
T_s	Surface (skin) temperature	K	I	O
w_s	Mass mixing ratio of water vapor	kg kg^{-1}	I	O
u_*	Shear (friction) velocity	m s^{-1}	I	O
(c) Diagnostic outputs				
Q^*	Net all-wave radiation flux density[b]	W m^{-2}	O	O
T_{canyon}	Dry-bulb air temperature at half building height	K	O	O
r_{canyon}	Mass mixing ratio of water vapor at half building height	kg kg^{-1}	O	O
u_{canyon}	Zonal component of wind velocity at half building height	m s^{-1}	O	O
v_{canyon}	Meridional component of wind velocity at half building height	m s^{-1}	O	O
Q_{cooling}	Buildings' power demand for cooling	W m^{-2}	O	O
Q_{heating}	Buildings' power demand for heating	W m^{-2}	O	O
Q_{thermal}	Thermal power production of solar panels on roofs	W m^{-2}	O	O
Q_{electric}	Electric power production of solar panels on roofs	W m^{-2}	O	O

Note. Full list of parameters are reported in `URBAPARM.TBL`. [a]WRF atmospheric surface pressure is assumed to be the same as the atmospheric pressure at the roof level in TEB due to the coupling assumption (see $h_{0,\text{TEB}}$ and $h_{0,\text{WRF}}$; Figure 1). WRF-TEB assumes the atmospheric pressure to be the uniform throughout the canyon. [b]$Q^* = S^{\downarrow} - S^{\uparrow} + L^{\downarrow} - L^{\uparrow}$, where $S^{\downarrow}$ is the (total) downwelling shortwave radiation flux density calculated as $S^{\downarrow\downarrow} + S^{\Downarrow\downarrow}$, and $S^{\uparrow}$ and $L^{\uparrow}$ is the upwelling shortwave and longwave radiation flux density.

calculation of surface fluxes (Best et al., 2004). The current implementation of WRF-TEB is designed to work in both WRF (UCAR, 2020) and WRF-CMake (Riechert & Meyer, 2019a) alongside the Yonsei University (YSU) planetary boundary layer (PBL) scheme (Hong et al., 2006; Noh et al., 2003), the revised MM5 Monin-Obukhov surface layer scheme (Jimnez et al., 2012), and the Noah-LSM as they have been shown to perform reasonably in recent comparison studies for several types of environments (Greve et al., 2013; Hari Prasad et al., 2016; Hu et al., 2010, 2013; Shin & Dudhia, 2016; Xie et al., 2012).

The general workflow used in WRF-TEB is as follows: (1) for each grid cell defined as *urban*, and each time step, WRF provides TEB with atmospheric and model-specific quantities from the lowest model level (Table 2a), and site-specific characteristics such as grid cell spatial coordinates and surface parameters (e.g., building height and roof albedo) from a lookup table in WRF (`URBPARM.TBL`). (2) From these, TEB computes area-averaged surface quantities (Table 2b) and passes them to WRF (Figure 1). (3) Surface diagnostics (Table 2c) are calculated in TEB and passed directly to WRF as outputs without affecting calculations in the WRF dynamical core.

As with other TEB features (section 2.2), anthropogenic heat flux options in WRF-TEB are specified in the `URBPARM` file and modeled as follows: (i) traffic and industrial heat fluxes are user specified, (ii) heating

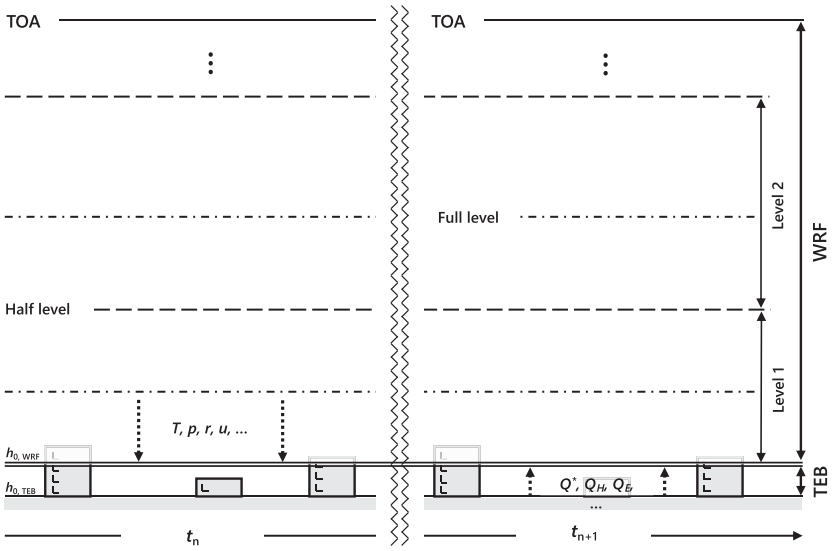

Figure 1. Coupling approach. For each urban grid cells, and at each time step (t_n), atmospheric quantities from the first model level such as dry-bulb air temperature (T), atmospheric pressure (p), mass mixing ratio of water vapor (r), and horizontal wind components (u, v; Table 2a) are passed to TEB. In turn, TEB computes area-averaged surface quantities such as net all-wave radiation flux density (Q^*), and turbulent sensible and latent heat flux density (Q_H and Q_E; Table 2b) and passes them to WRF at the surface (t_{n+1}). Under the single-layer UCM assumption, the vertical extent of the atmospheric model extend from $h_{0,\mathrm{WRF}}$ to the top of the atmosphere (TOA) whereas the UCM is assumed *below* the ground. This therefore creates a mismatch between the real ground level $h_{0,\mathrm{TEB}}$ and the ground level as seen by the atmospheric model $h_{0,\mathrm{WRF}}$.

and air conditioning from buildings are a function of meteorological conditions, physical characteristics of the building envelope, internal heat release from electrical appliances, and heating/air conditioning set points. All options available in TEB (and therefore WRF-TEB) are given in the TEB section of the Surfex scientific documentation (Le Moigne et al., 2018).

Given the different nature of the two models, with atmospheric equations solved explicitly (WRF) and processes parameterized (TEB), several assumptions and simplifications are made:

In TEB, the urban canopy layer (UCL) is represented as a point. This results in a mismatch in elevation between TEB and WRF (i.e., $h_{0,\mathrm{TEB}}$ and $h_{0,\mathrm{WRF}}$; Figure 1). As with other single-layer UCMs, $h_{0,\mathrm{WRF}}$ is located at the mean building height (given in `URBPARM.TBL`). This means that the urban canopy layer is located below the surface of the NWP model. This is an important assumption of any single-layer UCM and rarely explicitly stated. This assumption may be acceptable if buildings are low to midrise (e.g., <20 m) and of uniform height, such as typical of extensive suburban areas. However, areas with tall and/or variable height roughness elements (e.g., skyscrapers), as in central business districts in many cities worldwide, may not suit the use of a single-layer UCM. Depending on the user's needs and model configuration, outputs from WRF-TEB may require postprocessing to account for this assumption.

Surface diagnostics in WRF are given at standard World Meteorological Organization weather station heights (i.e., 10 m for wind, and 2 m for air temperature and humidity), thus representing quantities below the mean building height in the case of single layer UCMs. In WRF-TEB, we do not calculate these explicitly but simply rely on TEB's surface diagnostics (i.e., representative at half building height; see Masson, 2000). Given the single-layer assumption, WRF-TEB's surface diagnostics are, in effect, representative at half the building height *below* the first model level.

The treatment of vegetation in WRF-TEB can be done in three ways. Users have the option to define the fraction of vegetation as (a) *integrated* (i.e., vegetation inside canyons Lemonsu et al., 2012 or as part of green roofs de Munck et al., 2013), (b) *nonintegrated*, or (c) *both* (i.e., integrated *and* nonintegrated). In (a), WRF-TEB uses a simple urban vegetation scheme with time-constant Bowen ratio (default 0.25) and albedo (default 0.15). These values can be modified (e.g., higher for drought conditions), but are constant during a

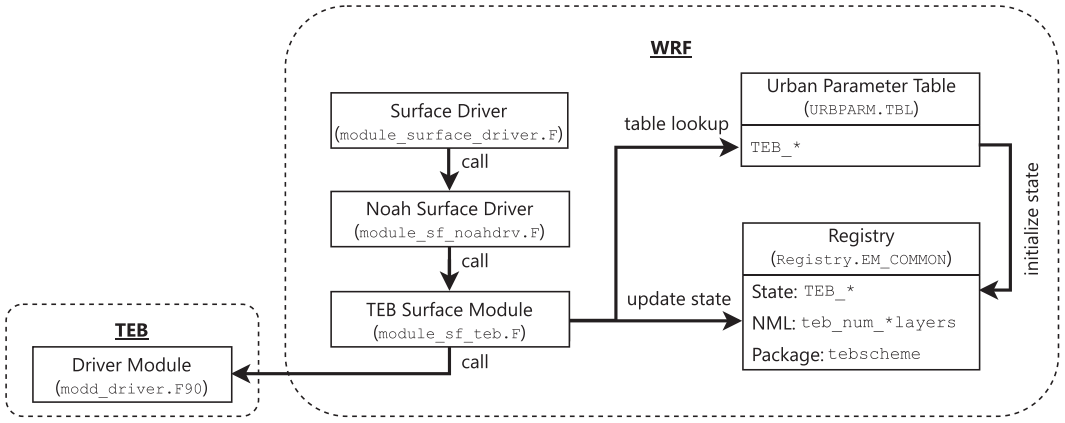

Figure 2. Software coupling. When TEB is enabled (`sf_urban_physics = 4` in `namelist.input`), the Noah Surface Driver calls the TEB subroutine in the TEB Surface Module. This calls the `TEB_DRIVER` subroutine in the Driver Module in the TEB library. The TEB driver is invoked for each urban grid cell at every time step. TEB parameters and initial conditions (IC; e.g., initial wall temperature) are read from a lookup table (`URBPARM.TBL`) and state variables are initialized with those IC at the start of the simulation. Conversions between units and dimensions are done in TEB Surface Module.

model simulation. The temporal evolution of soil temperature, surface water storage or vegetation is not represented. The approach of Lemonsu et al. (2012) to simulate in-canyon vegetation with the ISBA model is expected to be more universally applicable than the approach implemented in the current WRF-TEB. The default option in WRF-TEB (b) is to calculate the surface energy balance separately for urban and vegetated areas and to aggregate the fluxes. For the vegetated areas, the Noah-LSM cropland (MODIS class 12 in `VEGPARM.TBL`) class is assumed, whereas for built areas, TEB is used assuming no vegetation. In (c), users define the fraction of vegetation to model as integrated and nonintegrated respectively.

4. Software Implementation

The implementation of model to software and its testing are critical aspects of any model development. In WRF-TEB (Figure 2), the data flow and sequence of Fortran subroutine invocations is similar to existing couplings (e.g., SLUCM, BEM) and can be coupled serially as parallelization will be inherited from WRF over the number of grid cells in the domain(s). The main differences are in its code structure, implementation, and location. Although a generalizable framework (Common Community Physics Package; CCPP) is currently being developed (see Developmental Testbed Center, 2019), WRF does not yet provide a clear extension or plugin mechanism for integrating external models. Here we summarize the main techniques and issues encountered during its development that may aid other future meteorological model developments:

1. *Modularity*: As existing coupled models are tightly interwoven into WRF and partly modified, deviations from the original code can be hard to detect. By treating models as libraries we increase modularity.
2. *Clarity*: As different couplings are typically in one large Fortran file or subroutine, understanding existing coupling is difficult. This is complicated further by reuse of some variables and parameters. By separating the coupling code, using consistent naming conventions, and introducing a small amount of duplication, we aim to increase clarity and reduce the time to undertake the coupling.
3. *Reliability*: WRF allows simulations to run even when required inputs are omitted, resulting in uninitialized values that may change results. By using a stricter input validation, we increase reliability of model output.

4.1. Modularity

Copying the coupled model's source code (e.g., TEB) directly into the WRF codebase is likely to cause code fragmentation, ultimately hindering model development, collaboration, and (possibly) the formation of a strong coherent community around that model.

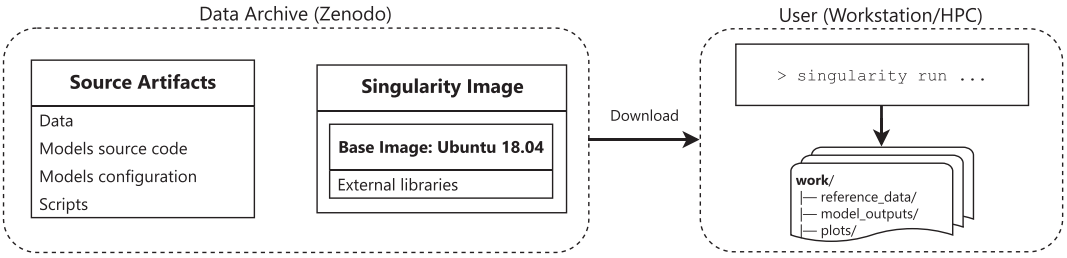

Figure 3. The data archive (Meyer, 2020) contains all data and code used to evaluate the models. Users wishing to reproduce our results can download the data archive and run Singularity on their local or remote systems. As small differences in outputs may be possible because of the different hardware used, model outputs are also provided as reference.

Often, coupled models are not readily available in both offline and online versions (e.g., SLUCM, BEP, BEP +BEM, PBL), or are modified from their original (i.e., offline) version (e.g., Noah-LSM). For this reason it can be challenging to evaluate both versions.

To prevent this, we (a) keep TEB in its original repository on GitHub (Meyer et al., 2020), (b) include TEB as a library dependency in WRF (similar to what is done with other external libraries such as `libnetcdf`, Unidata, 2019 or `libjasper`; Adams, 2019), and (c) link the TEB library to WRF through a thin coupling interface (`module_sf_teb.F`; Figure 2). Then, to minimize complexity and overheads involved in publishing the TEB library in package managers (e.g., Ubuntu package repositories; Canonical Ltd, 2019), TEB is downloaded and compiled from source during the general WRF build process by relying on CMake's ExternalProject module (Kitware Inc., 2019b).

Once coupled in this way, any improvements, bug fixes, or other changes to the ULSM are inherited from the model source code repository (e.g., on GitHub) and included in WRF with a new commit (or version) identifier that is downloaded during the build process. This provides a central location for "issues," maintains the community around the model, and facilitates the creation of offline tests (e.g., on units or individual components, or end-to-end integration tests of the whole model).

From this, a natural step is to make use of freely available Continuous Integration (CI) services to automate the execution of such tests on new commits on different operating systems and using multiple compiler versions and options, thus providing a stronger sense of reproducibility and trust. Where an appropriate testing methodology exists, CI can also enable model developers to accept code contributions from the community with more confidence.

4.2. Clarity

WRF/urban models (e.g., SLUCM, BEM/BEP) often reuse a subset of variables, parameters, and parts of the coupling "glue" code. Problems can arise as WRF allows each model to declare which state variables should be allocated in memory (via "schemes" in its registry). As the glue code for all models of one type is typically in a single large Fortran file (e.g., `module_sf_noahdrv.F`) with many conditional branches (e.g., `IF(SF_URBAN_PHYSICS == 1) THEN ...`) that is run irrespective of which model is activated, a model code contributor must understand most of the glue code to not introduce unintended memory issues. In the best case, using a variable that is not allocated leads to program crashes. In the worst case, it may lead to reading from, or writing to, other variables that are nearby in memory. In such case, the program may not fail but simply change results, possibly without the user being aware.

Here, we separate the configuration and coupling of TEB as much as possible from other models to improve clarity and reduce the time needed to understand the coupling. The prefix `TEB_` is added to state variables, array dimension names, and parameters used by TEB, while as much glue code as possible is moved into the TEB coupling module (`module_sf_teb.F`; Figure 2). Without shared variables and parameters both the coupling and model can be understood in isolation and evolve independently from other models, hopefully encouraging community contributions. Despite duplicating state variables and parameters between urban models, we believe the benefits outweigh the disadvantages.

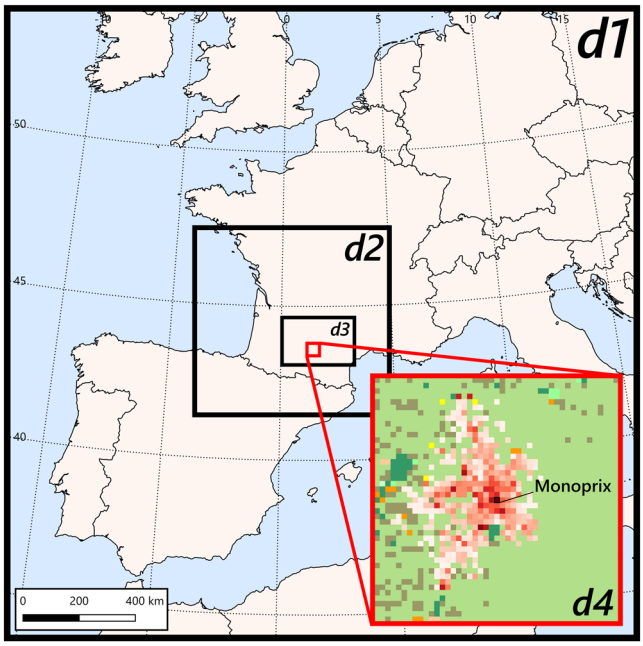

Figure 4. Four nested domains (d1–d4) shown on a base map from Natural Earth (2019). Innermost domain (d4, 1 km horizontal grid spacing) has WPS MODIS 30 arc-sec land cover/use (UCAR, 2019) with River Garonne replaced with urban land use and 1 km interpolated MApUCE urban fraction (Bocher et al., 2018) overlaid (see section 6.2). Higher urban fraction shown with darker shades of red. Manually assigned urban fraction (0.15) shown in yellow. See namelists in `configs/wrf/capitoul` in Meyer (2020) for the complete list of options used.

4.3. Reliability

WRF checks the validity of some, but not all, user inputs. Unfortunately, when an urban parameter is unspecified, unexpected simulation results can occur as uninitialized memory values are used instead of an error being raised. These errors may be hard to trace or can go undetected.

To solve this we separate parameter reading (in `module_sf_urban.F`) into three phases: (a) initialize all parameters with a known out-of-bounds value, (b) read the user-supplied parameters (from `URBPARM.TBL`), and (c) check that parameter values are not equal to the out-of-bounds value. Missing parameters cause an error message with the parameter name provided and stop WRF, thus saving time by early detection. Additional improvements could include checking value ranges for each parameter.

5. Scientific Reproducibility

Issues of lack of scientific reproducibility have been noted by several authors in various disciplines (e.g., Atmanspacher & Maasen, 2016; Chen et al., 2019; Cohen-Boulakia et al., 2017; Grning et al., 2018; Redish et al., 2018; Sochat et al., 2017; Van Bavel et al., 2016). To achieve a reasonable level of scientific reproducibility, several aspects must be considered. In the current context we identify: CPU (micro) architecture, operating system (OS), compiler vendor, compiler version, compiler options, external library versions, source and version of data sets, preprocessing and postprocessing steps of data, model version and configuration, and plotting routines.

One way to achieve a reasonable level of scientific reproducibility is to automate the generation of results included in this article by using a combination of Shell and Python scripts run through a Singularity container (Kurtzer et al., 2017). This approch remains architecture dependent but provides full control over OS (i.e., Ubuntu 18.04 in our case), compiler vendor (i.e., GNU in our case), source of data sets, processing steps for data, model version and configuration, and plotting routines. Although this level of reproducibility may be deemed sufficient, the container would still rely on the Ubuntu package repository to install the compiler and external libraries without defining exact versions, or to download the "latest" (un-versioned) WPS high-resolution geographical data set (from the UCAR website). This can therefore lead to downloading newer compiler, library, or data set versions when re-building the Singularity image and ultimately alter results included in this paper.

Here, we achieve a reasonable level of scientific reproducibility by archiving all tools, data (including our results for reference, as equal results can only be guranteed on the same hardware), and software together with a Singularity image containing OS and external libraries to Zenodo (Meyer, 2020; Figure 3). By doing this we remove the need for duplicating configuration settings in tables or appendices, thus reducing accidental errors and allowing reproducibility on local or high-performance computing (HPC) systems. Users wishing to reproduce the results described in this paper can download the data archive (Meyer, 2020) and run Singularity on their local or remote systems (Figure 3).

6. Model Evaluation

A fundamental aspect of any software development is testing. Although neither WRF nor TEB have been developed with testing in mind, in this section we outline tests for: the coupling (integration test, section 6.3) and meteorological evaluation (section 6.4). The former assesses the coupling technically, while the latter is used to explore the scientific benefit of the coupling. Both tests use similar model configurations (section 6.2) with meteorological and geographical data for Toulouse, France (section 6.1).

Figure 5. View of Toulouse downtown roofs from the terrace of the central site. Source: Masson et al. (2008).

6.1. Site and Observational Data

Toulouse (Local Climate Zone 2; Hidalgo et al., 2019), the fouth largest city in France (475,438 inhabitants INSEE, 2016), is located on a plain, 80 km north of the Pyrenees mountains (Figure 4). It has mild wet winters and dry hot summers (Joly et al., 2010). Central Toulouse (1°26′4″E, 43°36′15″N Monoprix; Figure 5) has homogeneous dense midrise buildings (see Figure 2 in Pigeon et al., 2007). Buildings (average height 15 m) cover 55% of the plan area, whereas vegetation covers 10% (see Table 2 in Pigeon et al., 2007).

During the CAPITOUL (Canopy and Aerosol Particles Interactions in TOulouse Urban Layer Masson et al., 2008) campaign, a Gill HS 50 sonic anemometer for eddy covariance (EC), and other meteorological sensors, were mounted on a tower at 48 m to 27 m a.g.l. (depending on local wind conditions). EC sensors should be above the roughness sublayer to observe local-scale rather than microscale (e.g., individual urban obstacles) fluxes (Grimmond et al., 2004; Roth, 2000). The Goret et al. (2019) analysis of observed momentum fluxes confirms that they are located in the inertial sublayer (constant flux layer).

EC measurements (sampled at 50 Hz) postprocessing includes double-rotation (azimuth and pitch correction), recursive filtering according to McMillen (1988) with filter parameter set to 200 s, prior to 30 min covariance flux calculation (Pigeon et al., 2007). The EC flux footprint, calculated using the Kljun et al. (2015) model for each 30 min interval, identifies that the probable mean 80% fetch extends to around 500 m in all wind directions, except for southerlies where it extends to 1 km (Goret et al., 2019). Given the homogeneous characteristics within 500 m radius of the tower and areas further south, we assume that the observed turbulent fluxes are comparable to the modeled turbulent fluxes (horizontal grid spacing 1 km; Figure 4, d4).

Radiation fluxes observed with a Kipp and Zonen CNR1 radiometer (sample rate 0.1 Hz) mounted at the tower top are averaged to 1 min (used to force the offline TEB) and to 30 min (used in the meteorologicale-valuation). Air temperature and relative humidity measured with a Vaisala HMP233 Thermo-hygrometer (sample rate 0.1 Hz) at 43.3, 34.2, and 25 m a.g.l. when the mast was in the high, medium, and lowest position (respectively) are used as 1 and 30 min averages. The atmospheric pressure measured with a Vaisala PTB220 class A barometer (sample rate 0.1 Hz) at 20 m a.g.l., is used as 1 min average. Missing data are gap-filled every 1 min using observations from the routine observation station at Toulouse-Blagnac airport (7 km northwest of the tower) or a station operating at the site of Météo-France (flat grassland 6.5 km west southwest of the tower). The temperature and relative humidity values measured at these stations (2 m a.g.l.) are corrected by the average daily cycle of the differences between the values measured at the mast and at these sites. Wind speed measured at these stations (10 m a.g.l.) is corrected to the height of the mast assuming a logarithmic wind profile and neutral stratification. The values of aerodynamical roughness length and displacement height are 1.5 and 10.5 m, respectively.

The evaluation is undertaken between 2 and 5 July 2004 when the tower was at 48 m a.g.l. (28 m above roof height). A strong urban canopy layer heat island was present before sunrise on 4 July 2004 (Hidalgo et al., 2008). An offline TEB simulation is forced with the required meteorological data and the surface morphological parameters averaged for the area within in a 500 m radius from the tower (see Figure 1 in Goret et al., 2019 and data in Meyer, 2020).

6.2. Model Setup

Four nested domains (Figure 4) are set up using GIS4WRF version 0.14.2 (Meyer & Riechert, 2019b) with the innermost domain (Figure 4, d4) centered on the EC tower. The grid spacing is set to 1 km for the horizontal and to 66 m (increasing with height and in pressure (η) level equivalent) for the vertical, thus allowing equal comparison with observations (i.e., under single layer UCM assumptions (section 3). The 48 m a.g.l. EC tower is represented at 33 m a.g.l., because the vertical extent of the buildings (mean building height 15 m) is not represented in WRF (see single-layer UCM assumptions in section 3). The WPS MODIS land use (UCAR, 2019) and WRF urban fraction for the innermost domain are modified using GIS4WRF: (a) *lake*

Table 3
General WPS/WRF Configuration Settings Used in Integration Test and Meteorological Evaluation

Option	Value	TS/Unit
(a) General		
Start Time	1 July 2004 00:00	UTC
End Time	5 July 2004 08:18	UTC
Spin-up	(0)1	day
Timestep	(180)108, (60)36, 12, 4	s
(b) Grid		
Map Projection	LCC	—
Horizontal Spacing	27, 9, 3, 1	km
Vertical Spacing	$f(h)$ with L1 = 66[†]	m
Vertical Levels	61	—
Nests and Grid Ratio	(2)4 and 1:3	—
Nesting Approach	1-way	—
Urban Classes	1	—
(c) Initial and Boundary Condition		
Data Set Provider	ECMWF	—
Data Set Name	Cycle 28r2 analysis[a]	—
Horizontal Spacing	TL511 (≈ 40 km)	—
Vertical Levels	61	—
Time Interval	6	h
(d) Physical parameterization		
Shortwave Radiation	Dudhia[b]	—
Longwave Radiation	RRTM[c]	—
Microphysics	Single-moment 3-class[d]	—
PBL	YSU[e]	—
Surface layer	Revised MM5[f]	—
LSM	Noah-LSM[g]	—
UCM	TEB[h] version 4.0.1[h]	—

Note. Integration test values deviating from the common shared configuration are enclosed in parenthesis. The model time step is indicated for each domain (d1–d4). Time Standard (TS), Coordinated Universal Time (UTC), Lambert Conformal Conic (LCC). European Centre for Medium-Range Weather Forecasts (ECMWF). [†]Vertical grid spacing increasing with height (h) and first level (L1) set to 66 m a.g.l. See namelists in `configs/wrf/capitoul` in Meyer (2020) for the complete list of options used.
[a]ECMWF (2015), [b]Dudhia (1989), [c]Mlawer et al. (1997) [d]Hong et al. (2004), [e]Hong et al. (2006) [f]Jimnez et al. (2012), [g]Chen and Dudhia (2001) [h]Meyer et al. (2020), [i]Masson et al. (2020)

(MODIS class 21) used to indicate the River Garonne within central Toulouse is reassigned to *cropland* (MODIS class 12) as the river would otherwise be over represented in the 1 km grid (the river is ≈ 200 m wide), and (b) the urban fraction (i.e., constant for the whole grid for pixels defined *urban*; MODIS class 13, i.e., built) is replaced with that from the MApUCE (100 m resolution) data (Bocher et al., 2018) linearly interpolated to 1 km to provide spatial variability within the domain. Grid cells classified as urban in the MODIS 30 arcsec data set but not present in MApUCE (Figure 4, d4 yellow) are given a default fraction of 0.15 to be representative of the area. For the initial and boundary conditions we used ECMWF Cycle 28r2 analysis (ECMWF, 2015) with native horizontal grid spacing TL511 (≈ 40 km), gridded to 0.36 arc-degree.

Domains are generated using WPS-CMake version 4.1.0 (Riechert & Meyer, 2019b). WPS/WRF are configured with parameters from Table 3 (see namelists in `configs/wrf/capitoul` in Meyer (2020) for the complete configuration). TEB (offline and online) surface characteristics are derived and adapted from Lemonsu et al. (2004) and Schoetter et al. (2017). To reduce the computational and storage cost during the integration test, we only run the outermost two domains with a longer simulation timestep (Table 3). Other differences in configuration settings between integration test and meteorological evaluation are reported in Table 4.

6.3. Integration Test

The model coupling is validated using an integration (i.e., end-to-end) test. Although neither WRF nor TEB carry out unit tests on their components, we assume that each on their own is working correctly. In this test, variables passed in the coupling (Table 2a) and TEB specific output variables (Tables 2b and 2c) are compared between offline (i.e., TEB) and online (i.e., WRF-TEB) models (Figure 6). As the inputs required to force TEB are not provided as standard WRF outputs, we introduce new variables in WRF's registry. To avoid permanently allocating memory for the additional 11 variables (Table 2) a new registry package (`teb_test`) is enabled through the namelist configuration option `teb_integration_test=1`. Thus, only the necessary variables are allocated without performance overheads when tests are not conducted, and further testing can be performed when new software releases are available. Any difference larger than machine precision is attributed to coupling implementation errors. Results are evaluated graphically and statistically (metrics Appendix A).

The implemented test detects errors that cause incorrect

1. loading of parameters from the urban parameter table file;
2. passing of parameters to TEB;
3. conversion of date/time from WRF to TEB conventions;
4. passing of geographical and date/time coordinates to TEB;
5. passing of TEB-internal state variables;
6. storing of TEB-specific output diagnostics;
7. updating of WRF state variables from TEB outputs;
8. grid cell looping; and
9. activation of TEB based on grid cell vegetation type and global scheme selector.

These errors can lead to software crashes or nonidentical results between TEB offline and WRF-TEB. With offline TEB forced with data from the WRF dynamical core quantities converted prior to use in TEB are not assessed. For example, as WRF uses mixing ratio whereas TEB uses specific humidity, it must be converted

Table 4
Differences in TEB Configuration Between the Integration Test and the Meteorological Evaluation

Option	Integration test	Meteorological evaluation
Vegetation Model	B	B/N
Air Conditioning	On	Off
Heating	On	Off
Green Roofs	On	Off
Solar Panels	On	Off

Note. Bowen ratio (B), Noah-LSM (N). See namelists in `configs/teb` in Meyer (2020) for the complete list of options used.

for TEB. Conversion errors would propagate in TEB and be evident in this comparison. Meteorological evaluation is undertaken separately (section 6.4).

The regular WRF output are directly comparable to offline TEB outputs, as the relevant WRF quantities are not modified further. Modifications are prevented by setting the test grid urban fraction to 100%. As this may be inappropriate for other submodels, the same method as used to obtain TEB forcing could be used. However, it may not be easy to test if the output quantities are correctly passed without using other techniques.

This testing approach (Figure 6) requires a common calling method of the offline model (TEB). As TEB can be compiled both as a library for online use and as an executable for offline use (Figure 6), the integration test tool (Figure 6) can run both TEB and WRF-TEB with the same source code. Thus, there is a strict testing of the coupling. The CCPP effort (Developmental Testbed Center, 2019) aims to organize models (/schemes) in a central location independent of a target framework. This may solve similar issues in the future. Unfortunately, it is not (yet) ready for use within WRF, or as an offline model and coupling testing tool.

Although a single configuration cannot represent all the degrees of freedom defined by the different model options and input parameters, we activate as many TEB options as possible (Table 4). Results show no visible differences between TEB and WRF-TEB (Figure 7). Similarly, no errors are detected using statistical metrics (NRMSE = 0%, Figure 7).

6.4. Meteorological Evaluation

With the coupling code verified (section 6.3), a meteorological evaluation allows the scientific benefit of the coupling to be explored. Given the wide range of WRF options (Powers et al., 2017) the individual choices (e.g., radiation, microphysics, and boundary layer) may have a larger impact than the UCM selected. Here, the evaluation is focused on TEB.

For the evaluation period (2–5 July 2004) the net all-wave radiation flux density (Q^*) is simulated well by TEB but only moderately well by WRF-TEB (Figure 8a; Table 5). Unsurprisingly, TEB forced with observations has a lower mean absolute error (MAE) for Q^* (MAE $\approx$ 7.7 W m^{-2}; Figure 8a) than when forced with

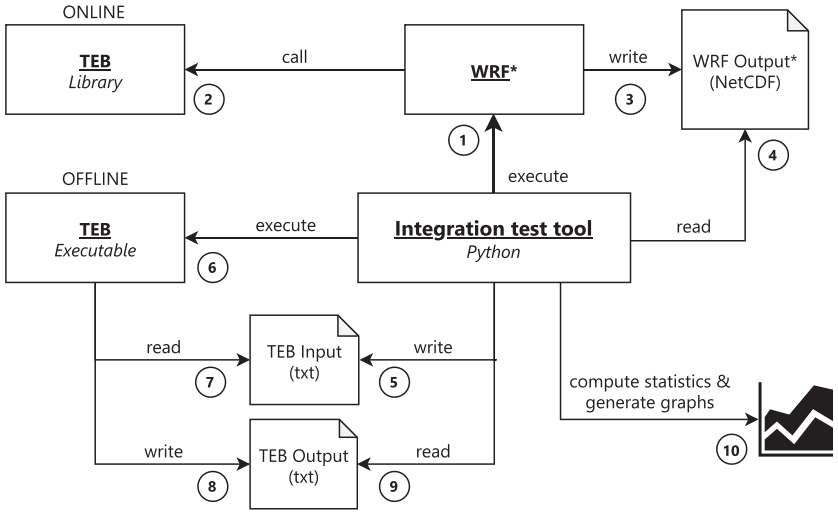

Figure 6. Integration test workflow: (1) The integration test tool runs the integration test case in WRF, (2) the TEB library is called from WRF, and (3) WRF writes outputs (including TEB inputs) to NetCDF files. At the end of the WRF simulation, (4) TEB inputs are read and (5) TEB inputs are generated (6) for offline execution. (7) TEB reads the inputs, (8) simulates the same case, and stores outputs as text files. (9) Outputs are read and (10) compared to the WRF outputs using statistics and graphs. Asterisk (*) indicates that WRF is run with `teb_integration_test=1` (see section 6.3).

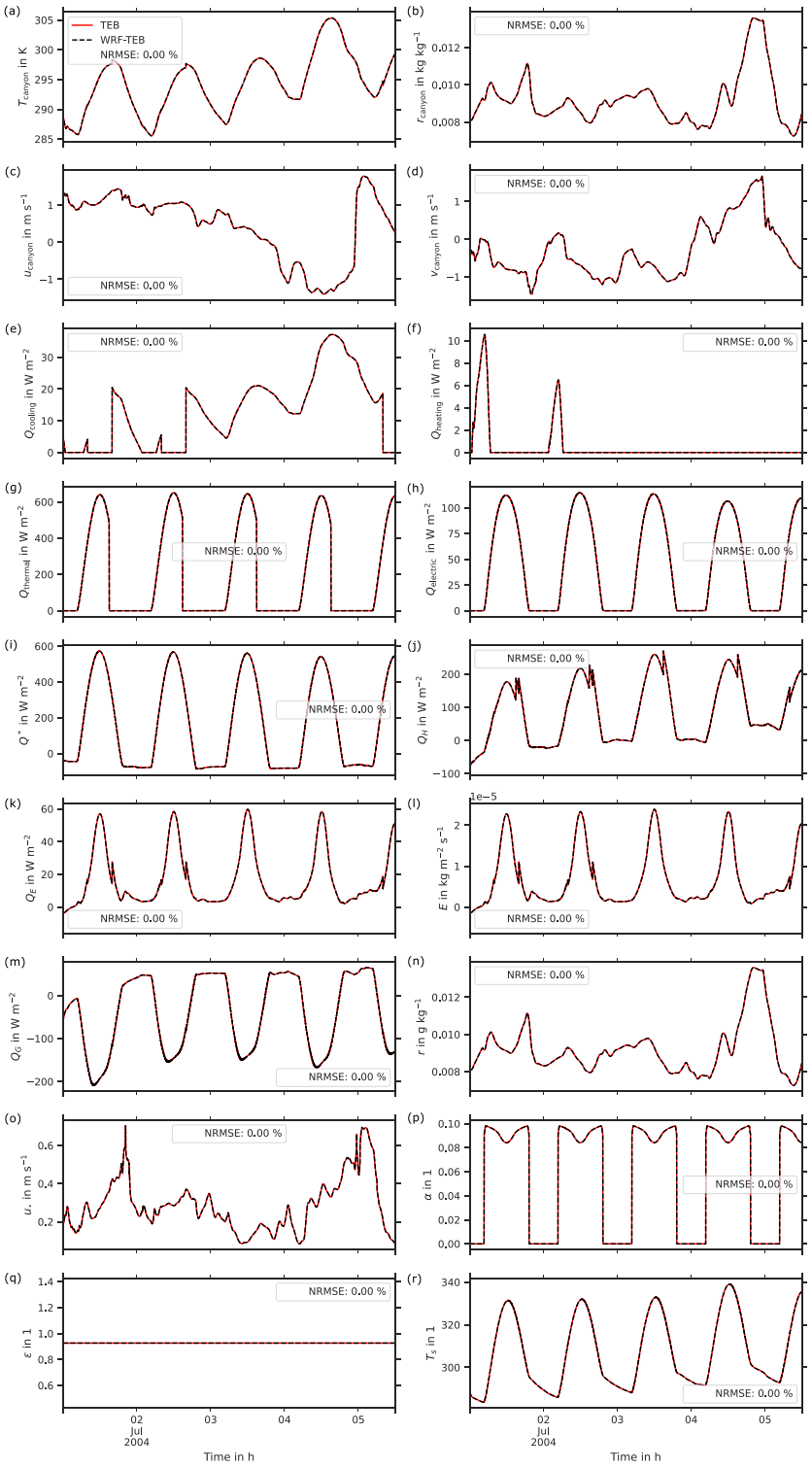

Figure 7. Results from the integration test used to verify the coupling. Line graph and normalized root-mean-square error (NRMSE) calculated from offline TEB and coupled WRF-TEB outputs. (a) Dry-bulb air temperature, (b) mass mixing ratio of water vapor, and (c) zonal and (d) meridional component of wind velocity at half building height. Buildings' power demand for (e) cooling and (f) heating. (g) Thermal and (h) electric power production of solar panels on roofs. (i) Net all-wave radiation, (j) turbulent sensible, and latent (k) flux density. (l) Evaporation mass and (m) ground heat flux density. (n) Mass mixing ratio of water vapor. Surface (o) shear (friction) velocity, (p) albedo, (q) emissivity, and (r) skin temperature.

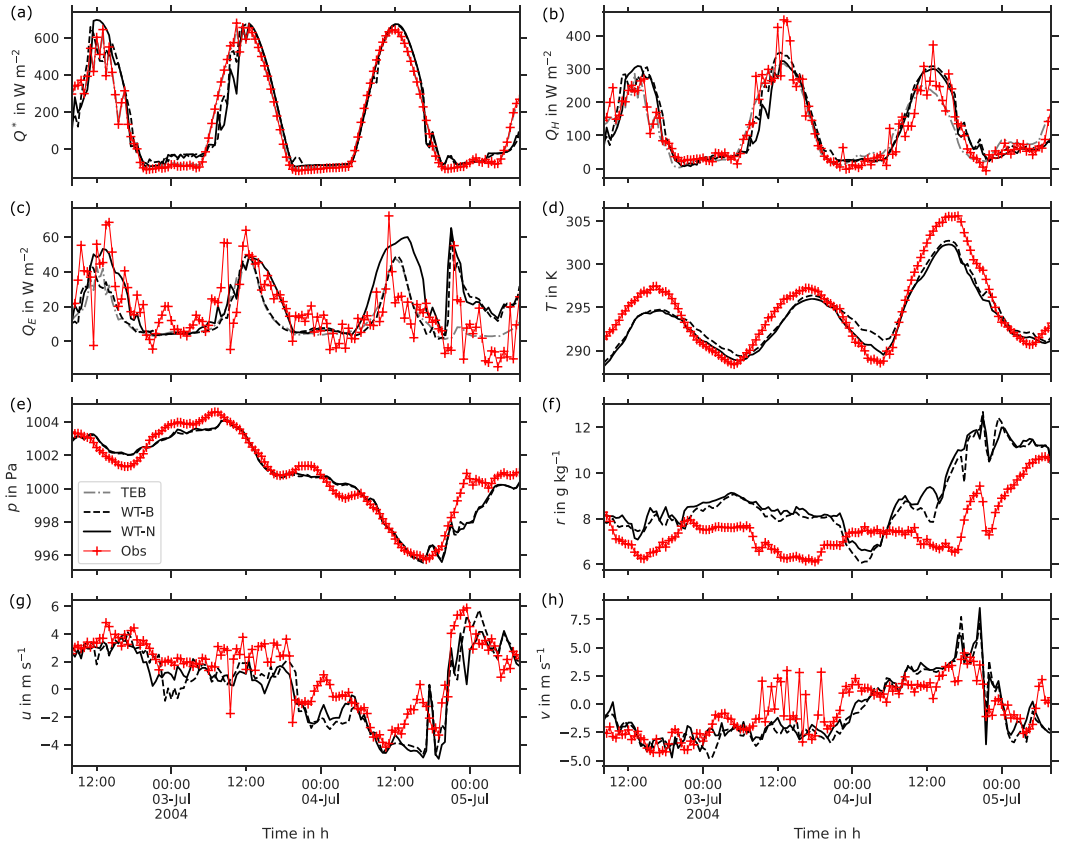

Figure 8. Observations (Obs) and model simulations (WRF-TEB with Noah-LSM (WT-N), WRF-TEB with Bowen ratio (WT-B), offline TEB (TEB)) for (a) net all-wave radiation flux density, (b) turbulent sensible heat flux density, and (c) turbulent latent heat flux density. Results between WT-N, WT-B, and Obs are (d) dry-bulb air temperature, (e) atmospheric pressure, (f) mass mixing ratio of water vapor, (g) zonal component of wind velocity, and (h) meridional component of wind velocity. Observations from the Monoprix tower measured at 48 m. g.l. (a–c, g, and h), 43.3 m a.g.l.; (d, f), 20 m a.g.l. (e). Atmospheric pressure corrected for height by linearly interpolating between surface pressure and first model level (center) to measurement height at roof level because of the assumption used in single-layer UCM. All other quantities are uncorrected as changes would be minimal.

quantities simulated by WRF (MAE ≈ 63 W m^{-2}; Figure 8a; Table 5). This difference is most likely caused by the cloud microphysics scheme, which simulates too much cloud overnight (2 and 3 July) and the next morning (3 July). This leads to an overestimation of Q^* during the night and an underestimation during

Table 5
Meteorological Evaluation of Half-Hourly Values for 2–5 July 2004

		MAE			MBE			$\overline{x}$
Quantity	Unit	WT-N	WT-B	TEB	WT-N	WT-B	TEB	OBS
Q^*	W m^{-2}	62.6	62.5	7.7	7.9	11.6	1.0	154.5
Q_H	W m^{-2}	34.9	34.1	27.7	−2.7	7.9	−4.9	120.0
Q_E	W m^{-2}	13.4	13.1	11.3	5.7	−1.1	−4.6	18.0
T	K	1.5	1.6	—	−1.2	−0.7	—	294.7
p	hPa	0.5	0.5	—	−0.2	−0.3	—	1000.9
r	g kg^{-1}	1.5	1.4	—	1.4	1.2	—	7.5
u	m s^{-1}	1.3	1.4	—	−1.0	−0.9	—	1.3
v	m s^{-1}	1.4	1.5	—	−0.2	−0.4	—	−0.5

Note. WRF-TEB with Noah-LSM (WT-N), WRF-TEB with Bowen ratio (WT-B), offline TEB (TEB), and arithmetic mean ($\overline{x}$) of observation (OBS) values. Appendix A defines statistics. N = 145.

the morning by WRF-TEB (Figure 8a). Turbulent sensible and latent heat flux densities (Q_H and Q_E) are captured reasonably well by both TEB and WRF-TEB (MAE ≈ 34 and 27 W m^{-2} for Q_H and MAE ≈ 13 and 11 W m^{-2} for Q_E respectively; Figures 8b and 8c; Table 5).

Dry-bulb air temperature (T) at 48 m a.g.l. (28 m above roof level) is generally underestimated by WRF-TEB (mean bias error (MBE) ≈ −1.2 K; Figure 8d; Table 5) for the whole period and during the day, but slightly overestimated at night. Such underestimation requires further investigation but may be caused by other WRF processes (e.g., too low advected temperature). Mass mixing ratio of water vapor (r) is overestimated (MBE ≈ 1.4 g kg^{-1}; Figure 8f). Pressure (p) and wind components (u, v) are simulated reasonably well (MAE ≈ 0.5 Pa and 1.4 m s^{-1}, respectively; Figures 8e, 8g, and 8h; Table 5), indicating that WRF has captured the general atmospheric dynamics.

Overall, the choice of vegetation scheme used (i.e., WRF-TEB with Noah-LSM (WT-N) or with Bowen ratio (WT-B)) results in similar simulation performance (Table 5, Figure 8).

7. Concluding Remarks

The coupled WRF-TEB model enables a wide range of urban climate processes to be analyzed. In this paper, we describe techniques to help with the coupling approach, implementation, verification, and scientific reproducibility.

In implementing the coupling interface, we do not alter the current WRF framework but, instead, implement techniques to help with software modularity, clarity, and reliability, for example, treating TEB as an external library. We assess the software linkage with an integration test to ensure that the coupling is technically correct. The results of the integration test show no detectable differences with the offline TEB. The meteorological evaluation is used to confirm that the results are physically reasonable; these generally show reasonable agreement with net all-wave radiation and turbulent heat flux densities and other near-surface observations. Although improvements of surface fluxes and near-surface meteorological quantities may possibly be gained from using alternative parameters or parameterization schemes (e.g., microphysics, radiation) when configuring WRF, the interplay of these make attribution difficult. Furthermore, errors may arise from differences between observational source area (e.g., eddy covariance) and model grid length; parameter specification and uncertainties (e.g., from lack of availability, difficulty of "measurement" and theoretical understanding. This highlights the importance of undertaking both integration tests and meteorological evaluations.

Scientific reproducibility is addressed by providing model source code, configurations, data, and scripts with a Singularity image deposited on Zenodo. The coupled WRF-TEB model has been integrated into WRF and WRF-CMake and released as a free, open-source software on GitHub at https://github.com/teb-model/wrf-teb.

We encourage future versions of WRF to include the implementation of a flexible number (i.e., beyond three) of urban classes to allow for a greater heterogeneity of urban form and function to be represented.

Appendix A: Evaluation Metrics

The vector of differences $\boldsymbol{d} = (d_1,...,d_i)$ between two vectors $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ of paired quantities $x_{a,i}$ and $x_{b,i}$ is defined as

$$\boldsymbol{d} = \boldsymbol{x}_a - \boldsymbol{x}_b. \tag{A1}$$

The mean bias error (MBE), mean-absolute error (MAE), and normalized root-mean-square error (NRMSE) for a time series of times 1,...,N are defined as

$$\mathrm{MBE} = \frac{1}{N}\sum_{i=1}^{N} d_i, \tag{A2}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |d_i|, \tag{A3}$$

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\bar{x}_a} \times 100\ \%, \tag{A4}$$

where $\bar{x}_a$ is the arithmetic mean of $\boldsymbol{x}_a$ and the RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} d_i^2}{N}}. \tag{A5}$$

Data Availability Statement

Data for this paper are available in Meyer (2020).

Acknowledgments
We thank the European Centre for Medium-Range Weather Forecasts (ECMWF) for relicensing the data used in this study for redistribution. We thank Dev Niyogi (Purdue University) and one anonymous reviwer for their time and useful comments.

References

Adams, M. (2019). mdadams/jasper. https://github.com/mdadams/jasper

Arnfield, A. J. (2003). Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island. *International Journal of Climatology*, *23*(1), 1–26. https://doi.org/10.1002/joc.859

Atmanspacher, H., & Maasen, S. (2016). *Reproducibility: Principles, problems, practices, and prospects*. Hoboken, NJ: John Wiley & Sons.

Best, M. J. (2005). Representing urban areas within operational numerical weather prediction models. *Boundary-Layer Meteorology*, *114*(1), 91–109. https://doi.org/10.1007/s10546-004-4834-5

Best, M. J., Beljaars, A., Polcher, J., & Viterbo, P. (2004). A proposed structure for coupling tiled surfaces with the planetary boundary layer. *Journal of Hydrometeorology*, *5*(6), 1271–1278. https://doi.org/10.1175/JHM-382.1

Bocher, E., Petit, G., Bernard, J., & Palominos, S. (2018). A geoprocessing framework to compute urban indicators: The MApUCE tools chain. *Urban Climate*, *24*, 153–174. https://doi.org/10.1016/j.uclim.2018.01.008

Bueno, B., Pigeon, G., Norford, L. K., Zibouche, K., & Marchadier, C. (2012). Development and evaluation of a building energy model integrated in the TEB scheme. *Geoscientific Model Development*, *5*(2), 433–448. https://doi.org/10.5194/gmd-5-433-2012

Canonical Ltd (2019). Ubuntu—Ubuntu packages search. https://packages.ubuntu.com/

Chen, X., Dallmeier-Tiessen, S., Dasler, R., Feger, S., Fokianos, P., Gonzalez, J. B., et al. (2019). Open is not enough. *Nature Physics*, *15*(2), 113–119. https://doi.org/10.1038/s41567-018-0342-2

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 Modeling System. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4), 569. https://doi.org/10.1175/1520-0493(2001)129<0569:caalsh>2.0.co;2

Chen, Y., Jiang, W. M., Zhang, N., He, X. F., & Zhou, R. W. (2009). Numerical simulation of the anthropogenic heat effect on urban boundary layer structure. *Theoretical and Applied Climatology*, *97*(1-2), 123–134. https://doi.org/10.1007/s00704-008-0054-0

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C. S. B., Grossman-Clarke, S., et al. (2011). The integrated WRF/urban modellingsystem: Development, evaluation, and applications to urban environmental problems. *International Journal of Climatology*, *31*(2), 273–288. https://doi.org/10.1002/joc.2158

Cohen-Boulakia, S., Belhajjame, K., Collin, O., Chopard, J., Froidevaux, C., Gaignard, A., et al. (2017). Scientific workflows for computational reproducibility in the life sciences: Status, challenges and opportunities. *Future Generation Computer Systems*, *75*, 284–298. https://doi.org/10.1016/j.future.2017.01.012

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., et al. (2013). Long-term Climate Change: Projections, Commitments and Irreversibility. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

de Munck, C., Pigeon, G., Masson, V., Meunier, F., Bousquet, P., Tréméac, B., et al. (2013). How much can air conditioning increase air temperatures for a city like Paris, France? *International Journal of Climatology*, *33*(1), 210–227. https://doi.org/10.1002/joc.3415

Developmental Testbed Center (2018). DTC || WRF-NMM users page. https://dtcenter.org/wrf-nmm/users/

Developmental Testbed Center (2019). Common Community physics package (CCPP). https://dtcenter.org/community-code/common-community-physics-package-ccpp

Dudhia, J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences*, *46*(20), 3077–3107. https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2

ECMWF (2015). Cycle 28r2 summary of changes. https://www.ecmwf.int/en/forecasts/documentation-and-support/evolution-ifs/cycle-archived/cycle-28r2-summary-changes

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., et al. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*, *108*(D22), 8851. https://doi.org/10.1029/2002JD003296

Freitas, E. D., Rozoff, C. M., Cotton, W. R., & Dias, P. L. S. (2007). Interactions of an urban heat island and sea-breeze circulations during winter over the metropolitan area of São Paulo, Brazil. *Boundary-Layer Meteorology*, *122*(1), 43–65. https://doi.org/10.1007/s10546-006-9091-3

Gochis, D. J., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., et al. (2018). The WRF-Hydro modelingsystem technical description, (Version 5.0) (*Tech. Rep.*): NCAR Technical Note. https://ral.ucar.edu/sites/default/files/public/WRF-HydroV5TechnicalDescription.pdf

Goret, M., Masson, V., Schoetter, R., & Moine, M.-P. (2019). Inclusion of CO_2 flux modelling in an urban canopy layer model and an evaluation over an old European city centre. *Atmospheric Environment: X*, *3*, 100,042. https://doi.org/10.1016/j.aeaoa.2019.100042

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled online chemistry within the WRF model. *Atmospheric Environment*, *39*(37), 6957–6975. https://doi.org/10.1016/j.atmosenv.2005.04.027

Greve, P., Warrach-Sagi, K., & Wulfmeyer, V. (2013). Evaluating soil water content in a WRF-Noah downscaling experiment. *Journal of Applied Meteorology and Climatology*, *52*(10), 2312–2327. https://doi.org/10.1175/JAMC-D-12-0239.1

Grimmond, C. S. B., Blackett, M., Best, M. J., Baik, J.-J., Belcher, S. E., Beringer, J., et al. (2011). Initial results from Phase 2 of the international urban energy balance model comparison. *International Journal of Climatology*, *31*(2), 244–272. https://doi.org/10.1002/joc.2227

Grimmond, C. S. B., Blackett, M., Best, M. J., Barlow, J., Baik, J.-J., Belcher, S. E., et al. (2010). The international urban energy balance models comparison project: First results from phase 1. *Journal of Applied Meteorology and Climatology*, *49*(6), 1268–1292. https://doi.org/10.1175/2010JAMC2354.1

Grimmond, C. S. B., Salmond, A. J., Oke, T. R., Offerle, B., & Lemonsu, A. (2004). Flux and turbulence measurements at a densely built-up site in Marseille: Heat, mass (water and carbon dioxide), and momentum. *Journal of Geophysical Research*, *109*, D24101. https://doi.org/10.1029/2004JD004936

Grning, B., Chilton, J., Kster, J., Dale, R., Soranzo, N., van den Beek, M., et al. (2018). Practical computational reproducibility in the life sciences. *Cell Systems*, *6*(6), 631–635. https://doi.org/10.1016/j.cels.2018.03.014

Guerrette, J. J., & Henze, D. K. (2015). Development and application of the WRFPLUS-Chem online chemistry adjoint and WRFDA-Chem assimilation system. *Geoscientific Model Development*, *8*(6), 1857–1876. https://doi.org/10.5194/gmd-8-1857-2015

Hamdi, R., Degrauwe, D., & Termonia, P. (2012). Coupling the town energy balance (TEB) scheme to an operational limited-area NWP model: Evaluation for a highly urbanized area in Belgium. *Weather and Forecasting*, *27*(2), 323–344. https://doi.org/10.1175/WAF-D-11-00064.1

Hari Prasad, K. B. R. R., Venkata Srinivas, C., Venkateswara Naidu, C., Baskaran, R., & Venkatraman, B. (2016). Assessment of surface layer parameterizations in ARW using micro-meteorological observations from a tropical station: Surface layer parameterization. *Meteorological Applications*, *23*(2), 191–208. https://doi.org/10.1002/met.1545

Hidalgo, J., Dumas, G., Masson, V., Petit, G., Bechtel, B., Bocher, E., et al. (2019). Comparison between local climate zones maps derived from administrative datasets and satellite observations. *Urban Climate*, *27*, 64–89. https://doi.org/10.1016/j.uclim.2018.10.004

Hidalgo, J., Masson, V., & Pigeon, G. (2008). Urban-breeze circulation during the CAPITOUL experiment: Numerical simulations. *Meteorology and Atmospheric Physics*, *102*(3-4), 243–262. https://doi.org/10.1007/s00703-008-0345-0

Hong, S.-Y., Dudhia, J., & Chen, S.-H. (2004). A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly Weather Review*, *132*(1), 103–120. https://doi.org/10.1175/1520-0493(2004)132<0103:aratim>2.0.co;2

Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341. https://doi.org/10.1175/MWR3199.1

Hu, X.-M., Klein, P. M., & Xue, M. (2013). Evaluation of the updated YSU planetary boundary layer scheme within WRF for wind resource and air quality assessments. *Journal of Geophysical Research: Atmospheres*, *118*, 10,490–10,505. https://doi.org/10.1002/jgrd.50823

Hu, X.-M., Nielsen-Gammon, J. W., & Zhang, F. (2010). Evaluation of three planetary boundary layer schemes in the WRF model. *Journal of Applied Meteorology and Climatology*, *49*(9), 1831–1844. https://doi.org/10.1175/2010JAMC2432.1

Huang, X.-Y., Xiao, Q., Barker, D. M., Zhang, X., Michalakes, J., Huang, W., et al. (2009). Four-dimensional variational data assimilation for WRF: Formulation and preliminary results. *Monthly Weather Review*, *137*(1), 299–314. https://doi.org/10.1175/2008MWR2577.1

INSEE (2016). Populations lgales 2016—Commune de Toulouse (31555). https://www.insee.fr/fr/statistiques/3681328?geo=COM-31555

Janjic, Z. I. (2003). A nonhydrostatic model based on a new approach. *Meteorology and Atmospheric Physics*, *82*, 271–285. https://doi.org/10.1007/s00703-001-0587-6

Janjic, Z. I., Gerrity, J. P., Nickovic, S. (2001). An alternative approach to nonhydrostatic modeling. *Monthly Weather Review*, *129*(5), 1164–1178. https://doi.org/10.1175/1520-0493(2001)129<1164:aaatnm>2.0.co;2

Jimnez, P. A., Dudhia, J., Gonzlez-Rouco, J. F., Navarro, J., Montvez, J. P., & Garca-Bustamante, E. (2012). A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, *140*(3), 898–918. https://doi.org/10.1175/MWR-D-11-00056.1

Joly, D., Brossard, T., Cardot, H., Cavailhes, J., Hilal, M., & Wavresky, P. (2010). Les types de climats en France, une construction spatiale. *Cybergeo*. http://dx.doi.org/10.4000/cybergeo.23155

Kikegawa, Y., Tanaka, A., Ohashi, Y., Ihara, T., & Shigeta, Y. (2014). Observed and simulated sensitivities of summertime urban surface air temperatures to anthropogenic heat in downtown areas of two Japanese Major Cities, Tokyo and Osaka. *Theoretical and Applied Climatology*, *117*(1-2), 175–193. https://doi.org/10.1007/s00704-013-0996-8

Kitware Inc. (2019a). CMake. https://cmake.org/

Kitware Inc. (2019b). ExternalProject CMake 3.15.3 documentation. https://cmake.org/cmake/help/latest/module/ExternalProject.html

Kljun, N., Calanca, P., Rotach, M. W., & Schmid, H. P. (2015). A simple two-dimensional parameterisation for flux footprint prediction (FFP). *Geoscientific Model Development*, *8*(11), 3695–3713. https://doi.org/10.5194/gmd-8-3695-2015

Kondo, H., Genchi, Y., Kikegawa, Y., Ohashi, Y., Yoshikado, H., & Komiyama, H. (2005). Development of a multi-layer urban canopy model for the analysis of energy consumption in a big city: Structure of the urban canopy model and its basic performance. *Boundary-Layer Meteorology*, *116*(3), 395–421. https://doi.org/10.1007/s10546-005-0905-5

Kurtzer, G. M., Sochat, V., & Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLOS ONE*, *12*(5), e0177459. https://doi.org/10.1371/journal.pone.0177459

Kusaka, H., Kondo, H., Kikegawa, Y., & Kimura, F. (2001). A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models. *Boundary-Layer Meteorology*, *101*(3), 329–358. https://doi.org/10.1023/A:1019207923078

Le Moigne, P., Albergel, C., Boone, A., Belamari, S., Brun, E., Calvet, J.-C., Decharme, B., et al. (2018). *Surfex scientific documentation*. In P. Le Moigne (Ed.). https://www.umr-cnrm.fr/surfex/IMG/pdf/surfex_scidoc_v8.1.pdf

Lemonsu, A., Belair, S., & Mailhot, J. (2009). The new Canadian urban modelling system: Evaluation for two cases from the joint urban 2003 Oklahoma City experiment. *Boundary-Layer Meteorology*, *133*(1), 47–70. https://doi.org/10.1007/s10546-009-9414-2

Lemonsu, A., Grimmond, C. S. B., & Masson, V. (2004). Modeling the surface energy balance of the core of an old Mediterranean City: Marseille. *Journal of Applied Meteorology*, *43*(2), 312–327. https://doi.org/10.1175/1520-0450(2004)043<0312:mtsebo>2.0.co;2

Lemonsu, A., & Masson, V. (2002). Simulation of a summer urban breeze over Paris. *Boundary-Layer Meteorology*, *104*(3), 463–490. https://doi.org/10.1023/A:1016509614936

Lemonsu, A., Masson, V., Shashua-Bar, L., Erell, E., & Pearlmutter, D. (2012). Inclusion of vegetation in the town energy balance model for modelling urban green areas. *Geoscientific Model Development*, *5*(6), 1377–1393. https://doi.org/10.5194/gmd-5-1377-2012

Lin, C.-Y., Chen, F., Huang, J. C., Chen, W.-C., Liou, Y.-A., Chen, W.-N., & Liu, S.-C. (2008). Urban heat island effect and its impact on boundary layer development and landsea circulation over northern Taiwan. *Atmospheric Environment*, *42*(22), 5635–5649. https://doi.org/10.1016/j.atmosenv.2008.03.015

Liu, J., & Niyogi, D. (2019). Meta-analysis of urbanization impact on rainfall modification. *Scientific Reports*, *9*(1), 7301. https://doi.org/10.1038/s41598-019-42494-2

Martilli, A. (2002). Numerical study of urban impact on boundary layer structure: Sensitivity to wind speed, urban morphology, and rural soil moisture. *Journal of Applied Meteorology*, *41*(12), 1247–1266. https://doi.org/10.1175/1520-0450(2002)041<1247:nsouio>2.0.co;2

Martilli, A., Clappier, A., & Rotach, M. W. (2002). An urban surface exchange parameterisation for mesoscale models. *Boundary-Layer Meteorology*, *104*(2), 261–304. https://doi.org/10.1023/A:1016099921195

Masson, V. (2000). A physically-based scheme for the urban energy budget in atmospheric models. *Boundary-Layer Meteorology*, *94*(3), 357–397. https://doi.org/10.1023/A:1002463829265

Masson, V. (2006). Urban surface modeling and the meso-scale impact of cities. *Theoretical and Applied Climatology*, *84*(1-3), 35–45. https://doi.org/10.1007/s00704-005-0142-3

Masson, V., Bonhomme, M., Salagnac, J.-L., Briottet, X., & Lemonsu, A. (2014). Solar panels reduce both global warming and urban heat island. *Frontiers in Environmental Science*, *2*, 14. https://doi.org/10.3389/fenvs.2014.00014

Masson, V., Gomes, L., Pigeon, G., Liousse, C., Pont, V., Lagouarde, J.-P., et al. (2008). The Canopy and Aerosol Particles Interactions in TOulouse Urban Layer (CAPITOUL) experiment. *Meteorology and Atmospheric Physics*, *102*(3-4), 135–157. https://doi.org/10.1007/s00703-008-0289-4

Masson, V., Lemonsu, A., Pigeon, G., Schoetter, R., de Munck, C., Bueno, B., et al. (2020). The Town Energy Balance (TEB) model (Version 4.0.1). Zenodo. https://doi.org/10.5281/zenodo.3894784

McMillen, R. T. (1988). An eddy correlation technique with extended applicability to non-simple terrain. *Boundary-Layer Meteorology*, *43*(3), 231–245. https://doi.org/10.1007/BF00128405

Meyer, D. (2020). WRF-TEB data archive MinimalDX(Version 1.0.0) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.3554517

Meyer, D., & Raustad, R. (2020). MinimalDX (Version 0.1.4). Zenodo. https://doi.org/10.5281/zenodo.3892452

Meyer, D., & Riechert, M. (2019a). The gis4wrf plugin (Version 0.14.2). Zenodo. https://doi.org/10.5281/zenodo.3593935

Meyer, D., & Riechert, M. (2019b). Open source QGIS toolkit for the Advanced Research WRF modelling system. *Environmental Modelling & Software*, *112*, 166–178. https://doi.org/10.1016/j.envsoft.2018.10.018

Meyer, D., Schoetter, R., & Masson, V., Grimmond, S. (2020). Enhanced software and platform for the Town Energy Balance (TEB) model. *Journal of Open Source Software*, *5*(50), 2008. https://doi.org/10.21105/joss.02008

Meyer, D., & Thevenard, D. (2019). PsychroLib: A library of psychrometric functions to calculate thermodynamic properties of air. *Journal of Open Source Software*, *4*(33), 1137. https://doi.org/10.21105/joss.01137

Mills, G. (1997). An urban canopy-layer climate model. *Theoretical and Applied Climatology*, *57*(3-4), 229–244. https://doi.org/10.1007/BF00863615

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, *102*(D14), 16,663–16,682. https://doi.org/10.1029/97JD00237

Moonen, P., Defraeye, T., Dorer, V., Blocken, B., & Carmeliet, J. (2012). Urban physics: Effect of the micro-climate on comfort, health and energy demand. *Frontiers of Architectural Research*, *1*(3), 197–228. https://doi.org/10.1016/j.foar.2012.05.002

Natural Earth (2019). Natural Earth. https://www.naturalearthdata.com/

Noh, Y., Cheon, W. G., Hong, S. Y., & Raasch, S. (2003). Improvement of the K-profile model for the planetary boundary layer based on large eddy simulation data. *Boundary-Layer Meteorology*, *107*(2), 401–427. https://doi.org/10.1023/A:1022146015946

Noilhan, J., & Planton, S. (1989). A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review*, *117*(3), 536–549. https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2

Pigeon, G., Legain, D., Durand, P., & Masson, V. (2007). Anthropogenic heat release in an old European agglomeration (Toulouse, France). *International Journal of Climatology*, *27*(14), 1969–1981. https://doi.org/10.1002/joc.1530

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., et al. (2017). The weather research and forecasting model: Overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, *98*(8), 1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1

QGIS Development Team (2019). QGIS geographic information system. Open source geospatial foundation project.https://qgis.osgeo.org/

Redish, A. D., Kummerfeld, E., Morris, R. L., & Love, A. C. (2018). Opinion: Reproducibility failures are essential to scientific inquiry. *Proceedings of the National Academy of Sciences*, *115*(20), 5042–5046. https://doi.org/10.1073/pnas.1806370115

Riechert, M., & Meyer, D. (2019a). WRF-CMake: Integrating CMake support into the advanced research WRF (ARW) modelling system. *Journal of Open Source Software*, *4*(41), 1468. https://doi.org/10.21105/joss.01468

Riechert, M., & Meyer, D. (2019b). WPS-CMake. Zenodo. Version WPS-CMake-4.1.0. https://doi.org/10.5281/zenodo.3407075

Riechert, M., & Meyer, D. (2020). WRF-CMake: Integrating CMake support into the advanced research WRF (ARW) modelling system. Zenodo Version WRF-CMake-4.1.5 https://doi.org/10.5281/zenodo.3733593

Roth, M. (2000). Review of atmospheric turbulence over cities. *Quarterly Journal of the Royal Meteorological Society*, *126*(564), 941–990. https://doi.org/10.1002/qj.49712656409

Rozoff, C. M., Cotton, W. R., & Adegoke, J. O. (2003). Simulation of St. Louis, Missouri, land use impacts on thunderstorms. *Journal of Applied Meteorology*, *42*(6), 716–738. https://doi.org/10.1175/1520-0450(2003)042<0716:SOSLML>2.0.CO;2

Salamanca, F., Georgescu, M., Mahalov, A., Moustaoui, M., & Wang, M. (2014). Anthropogenic heating of the urban environment due to air conditioning: Anthropogenic heating due to AC. *Journal of Geophysical Research: Atmospheres*, *119*, 5949–5965. https://doi.org/10.1002/2013JD021225

Salamanca, F., Georgescu, M., Mahalov, A., Moustaoui, M., Wang, M., & Svoma, B. M. (2013). Assessing summertime urban air conditioning consumption in a semiarid environment. *Environmental Research Letters*, *8*(3), 034022. https://doi.org/10.1088/1748-9326/8/3/034022

Salamanca, F., Krpo, A., Martilli, A., & Clappier, A. (2010). A new building energy model coupled with an urban canopy parameterization for urban climate simulations part I. formulation, verification, and sensitivity analysis of the model. *Theoretical and Applied Climatology*, *99*(3-4), 331–344. https://doi.org/10.1007/s00704-009-0142-9

Santamouris, M., Papanikolaou, N., Livada, I., Koronakis, I., Georgakis, C., Argiriou, A., & Assimakopoulos, D. N. (2001). On the impact of urban climate on the energy consumption of buildings. *Solar Energy*, *70*(3), 201–216. https://doi.org/10.1016/S0038-092X(00)00095-5

Schoetter, R., Masson, V., Bourgeois, A., Pellegrino, M., & Lvy, J.-P. (2017). Parametrisation of the variety of human behaviour related to building energy consumption in the Town Energy Balance (SURFEX-TEB v. 8.2). *Geoscientific Model Development*, *10*(7), 2801–2831. https://doi.org/10.5194/gmd-10-2801-2017

Shepherd, J. M. (2005). A Review of current investigations of urban-induced rainfall and recommendations for the future. *Earth Interactions*, *9*(12), 1–27. https://doi.org/10.1175/EI156.1

Shin, H. H., & Dudhia, J. (2016). Evaluation of PBL parameterizations in WRF at subkilometer grid spacings: Turbulence statistics in the dry convective boundary layer. *Monthly Weather Review*, *144*(3), 1161–1177. https://doi.org/10.1175/MWR-D-15-0208.1

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). A description of the advanced research WRF version 3 (NCAR Technical Note Nos. NCAR/TN475+STR). University Corporation for Atmospheric Research. https://doi.org/10.5065/D68S4MVH

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., et al. (2019). A description of the advanced research WRF model Version 4 (No. NCAR/TN-556+STR). https://doi.org/10.5065/1dfh-6p97

Sochat, V. V., Prybol, C. J., & Kurtzer, G. M. (2017). Enhancing reproducibility in scientific computing: Metrics and registry for singularity containers. *PLOS ONE*, *12*(11), e0188511. https://doi.org/10.1371/journal.pone.0188511

Takane, Y., Kikegawa, Y., Hara, M., Ihara, T., Ohashi, Y., Adachi, S. A., et al. (2017). A climatological validation of urban air temperature and electricity demand simulated by a regional climate model coupled with an urban canopy model and a building energy model in an Asian megacity. *International Journal of Climatology*, *37*, 1035–1052. https://doi.org/10.1002/joc.5056

UCAR (2019). WPS V4 geographical static data. http://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html

UCAR (2020). The official repository for the Weather Research and Forecasting (WRF) model. https://github.com/wrf-model/WRF

Unger, J. (1999). Comparisons of urban and rural bioclimatological conditions in the case of a Central-European city. *International Journal of Biometeorology*, *43*(3), 139–144. https://doi.org/10.1007/s004840050129

Unidata (2019). Unidata/netcdf-c. Unidata https://github.com/Unidata/netcdf-c

United Nations (2019). World urbanization prospects: The 2018 revision (*Tech. Rep. No. ST/ESA/SER.A/420*): United Nations: New York.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). *Proceedings of the National Academy of Sciences*, *113*(23), 6454–6459. https://doi.org/10.1073/pnas.1521897113

Xie, B., Fung, J. C. H., Chan, A., & Lau, A. (2012). Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model. *Journal of Geophysical Research*, *117*, D12103. https://doi.org/10.1029/2011JD017080