Assessment of different kernel-driven models for daytime urban thermal radiation directionality simulation

Lu Jiang [a], Wenfeng Zhan [a,b,*], Leiqiu Hu [c], Fan Huang [a], Falu Hong [a], Zihan Liu [a], Jiameng Lai [a], Chenguang Wang [a]

[a] Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, International Institute for Earth System Science, Nanjing University, Nanjing, Jiangsu 210023, China
[b] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, Jiangsu 210023, China
[c] Department of Atmospheric Science, University of Alabama in Huntsville, AL 35899, USA

ARTICLE INFO

ABSTRACT

Parametric kernel-driven models are crucial for operationally adjusting satellite-derived urban land surface temperatures (LSTs) obtained at slant angles to hemispherically-representative values. Various parametric models have been proposed to simulate urban thermal radiation directionality, but a comprehensive comparison of the performances of the published parametric models, especially over a variety of urban surfaces under different solar radiation conditions, remains lacking. It is also unknown whether the combination of the available hotspot and base shape kernels can be used to derive new parametric models with even better performances compared with existing models. Based on both forward-modelling and satellite datasets, here we systematically evaluate three single-kernel and eight dual-kernel parametric models. The main findings are as follows: (1) Amongst the three single-kernel models, the VIN model has the best overall performance, with an average root-mean-square error (RMSE) of 0.79 and 1.35 K, based on forward-modelling and satellite data, respectively. However, the ROU and RL models outperform the VIN model when the solar zenith angle is less than 30°, and in particular it has a higher accuracy for hotspot description. (2) The dual-kernel models usually perform better than the single-kernel models. Amongst the eight dual-kernel models, those with the hotspot kernel $K_{\text{Hotspot\_rou}}$ (used by the ROU model) are more competent than those using $K_{\text{Hotspot\_vin}}$ (obtained from the Vinnikov model) as the hotspot kernel. The RVI model, in general, has the highest accuracy, with average RMSEs of 0.49 and 0.77 K based on forward-modelling and satellite data, respectively. (3) Compared with the single- and dual-kernel models, the multi-kernel models sometimes have better accuracies but the performance improvements are relatively limited. We also provide recommendations for model selection under various scenarios. Our systematic assessment improves our understanding of urban thermal radiation directionality regimes and potentially enables the improved correction of remotely-sensed urban LSTs, thus helping to advance thermal remote sensing of the urban environment.

1. Introduction

Land surface temperature (LST) is a key parameter resulting from but also modulating energy exchanges between land surfaces and the atmosphere. Thus, it has been widely used in various studies related to climate change, vegetation monitoring, and the urban environment (Voogt and Oke, 2003; Sobrino et al., 2005; Hansen et al., 2010; Li et al., 2020). Satellite thermal remote sensing is an important technique for obtaining the LST at a large scale (Li et al., 2013). However, most satellite thermal sensors (e.g., MODIS, AVHRR, and SEVIRI) can only sample the surface thermal status from a specific viewing angle each

Symbols

$T(\theta_s, \theta_v, \varphi)$ directional radiometric temperature
T_0 radiometric temperature at nadir
f_{iso} isotropy coefficient
$f_{Hotspot}$ hotspot kernel coefficient
$f_{BaseShape}$ base shape kernel coefficient
$K_{Hotspot}$ hotspot kernel
$K_{BaseShape}$ base shape kernel
θ_s solar zenith angle
θ_v viewing zenith angle
φ_s solar azimuth angle
φ_v viewing azimuth angle
φ relative azimuth angle between the sun and sensor
D angular distance between the sun and sensor
ξ phase angle related to the sun-surface-sensor position
k scale factor

time, yet there are large differences amongst the directional radiometric land surface temperatures (DRTs) from different observation angles, especially during daytime (Hu et al., 2016a; Cao et al., 2019b). Such complex thermal anisotropic variations make it challenging to estimate hemispherically-representative LSTs unbiased by observation angles (Li et al., 2013), and they greatly constrain the application of satellite-derived LST products (Cao et al., 2019b).

An in-depth understanding of the hemispherical characteristics of surface thermal radiation directionality (STRD) is critical for eliminating the angular impact on satellite-derived LSTs. Comprehensive investigations of thermal anisotropy have been conducted over natural surfaces such as croplands grassland and forest (Cao et al., 2019b). The effective thermal anisotropy intensity (i.e., the temperature difference between nadir and off-nadir observations) over typical vegetation surfaces usually ranges from 2.0 to 4.0 K (or slightly greater), as reported based on purely mechanistic forward 3D modelling (hereafter termed forward-modelling), as well as on ground-based and remotely- sensed observations (Lagouarde et al., 2000; Gastellu-Etchegorry et al., 2004; Pinheiro et al., 2004, 2006; Du et al., 2007; Liu et al., 2007; Rasmussen et al., 2011; Huang et al., 2012; Guillevic et al., 2013; Duffour et al., 2015). Forward-modelling strategies often require elaborate surface geometry and physical information to capture the thermal anisotropy regime. Such strategies are inappropriate for the angular normalization of satellite-derived DRTs into the nadir because complex surface information is usually unknown. It is also cumbersome and even unfeasible for ground-based and remotely-sensed observations to adequately characterize the hemispherical characteristics of STRD. For the purposes of angular normalization and inversion, simple but effective modelling strategies are therefore required. Parametric kernel-driven models can meet such needs when only a few DRTs are available (Cao et al., 2019b). The current kernel-driven STRD models can be generally grouped into two categories (Cao et al., 2021). The first category is adapted from the kernel-driven modelling framework developed for the visible and near-infrared (VNIR) spectra, which usually includes a geometric optic kernel, a radiative transfer kernel, and an isotropic kernel, i.e., the regular kernel-driven bidirectional reflectance distribution function (BRDF) models yet revised by replacing the reflectance with DRT (e.g., the Ross-Li series model) (Peng et al., 2011; Ren et al., 2014; Hu et al., 2016b, 2017; Liu et al., 2018; Cao et al., 2019a, 2021). The second category is initiated from the general thermal infrared (TIR) kernel-driven modelling framework (Cao et al., 2021), which often contains a base shape kernel, a hotspot kernel, and an isotropic kernel. Compared with the first category, the second category is specifically designed for STRD simulation; and it includes newly designed kernels, e.g., the Vinnikov model (Vinnikov et al., 2012), the Roujean-Lagouarde (RL) model

(Lagouarde and Irvine, 2008; Duffour et al., 2016), and the kernel-hotspot model that combines the Vinnikov and RL models (Ermida et al., 2018a, 2018b).

Compared with relatively homogeneous natural land surfaces that are usually covered by vegetation, soil and snow (or ice), urban surfaces are highly heterogeneous with a greater variety of surface thermal properties and a more prominent 3D structure that results in a complicated shading pattern amongst urban surface objects (e.g., buildings) (Oke et al., 2017; Mao and Li, 2020). The intensity of urban thermal radiation directionality (UTRD) is greater than that of natural surfaces. The UTRD intensity can even reach up to 10.0 K or more during daytime (Voogt and Oke, 1998; Voogt, 2008; Lagouarde et al., 2004, 2010, 2012; Yu et al., 2006; Zhan et al., 2012; Zhao et al., 2012; Hu et al., 2016a). Likewise, there are generally two groups of kernel-driven models used for UTRD simulation. The first group is a straightforward application of the available kernel-driven models used for natural surfaces. For example, the Vinnikov and RL models have also been shown to be effective for simulating UTRD (Lagouarde and Irvine, 2008; Duffour et al., 2016; Jiang et al., 2018). Similarly, the urban surface emissivity anisotropy (USEA) model, which was developed based on the Vinnikov model but contains a revised semi-empirical base shape kernel, has been illustrated applicable for the UTRD simulation (Sun et al., 2015). The second group conceives new kernels based on a particular consideration of urban geometry. For example, Wang et al. (2018a, 2018b, 2020) and Wang and Chen (2019) designed the GUTA series models (i.e., the GUTA -sparse, GUTA -osg, GUTA -dense, and GUTA -T models) for urban surfaces with different building densities.

For vegetation surfaces, the available models for simulating thermal anisotropy (e.g., the BRDF-based, RL, and Vinnikov models) have been comprehensively compared under numerous scenarios (Ermida et al., 2018b; Liu et al., 2018, 2020; Cao et al., 2019a). For urban surfaces, however, only a small number of parametric models have been evaluated, and the comparisons have only been conducted under a few scenarios with limited representation of urban surfaces and observation times (Sun et al., 2015). Previous investigations have shown that (1) new parametric kernel-driven models can be derived directly by combining the available hotspot and base shape kernels (Cao et al., 2021); and (2) the UTRD regime depends substantially on urban surface morphology and observation time (i.e., solar radiation conditions) (Voogt, 2008; Krayenhoff and Voogt, 2016; Hu and Wendel, 2019). We therefore identify the following two major issues for UTRD simulation with parametric models: First, it remains unclear whether there exist new kernel-driven models for UTRD simulation with even better performances based on combining all the available hotspot and base shape kernels. Second, we lack a comprehensive assessment of the performances of all of the available parametric models under scenarios that incorporate an adequate range of urban morphology and solar radiation conditions.

To fill these knowledge gaps, in this study we screen all of the published kernels suitable for UTRD simulation, from which six new parametric models are obtained. With both forward-modelling and satellite data, we then conduct a systematic performance assessment of eleven parametric models (including five existing models and six new models) for daytime UTRD simulation. This assessment can assist practitioners in choosing the most appropriate models under different scenarios, as well as in designing parametric models with even better performance in the future.

2. Datasets

Both forward-modelling and satellite datasets are used to assist model assessments; both provide adequate DRTs at different angles and are therefore suitable for model comparison. Note that the airborne dataset is not included in this study, and that further clarification is provided in Section 4.4.2.

2.1. Forward-modelling datasets

Forward-modelling datasets have been widely used for investigating UTRD, as well as for assessing parametric models (Duffour et al., 2015; Krayenhoff and Voogt, 2016; Dyce and Voogt, 2018; Wang et al., 2018a, 2018b, 2020; Wang and Chen, 2019; Hu and Wendel, 2019; Cao et al., 2019a). The forward-modelling datasets were produced using three major steps:

2.1.1. Step (1): Design of urban surface models (USMs)

The UTRD regime is closely related to urban morphology (Hu and Wendel, 2019). To represent the characteristics of typical urban surfaces as comprehensively as possible, nine USMs (termed USM01 to USM09, see Fig. A1 in Appendix A) were designed according to the local climate zone (LCZ) concept (Stewart and Oke, 2012). These nine USMs are exemplified with different building densities (densely or sparsely distributed) and building heights (high, low, or mixed heights), which were created using 3ds-Max software.

2.1.2. Step (2): Simulation of component temperatures

The Envi-met software extensively used in urban environment studies (Chow et al., 2011) was used to simulate the component temperatures of these nine USMs. The RMSE of the component temperature (including surface and wall temperatures) simulated by the Envi-met software has been demonstrated to range from 1.0 to 2.0 K (Yang et al., 2013; Huang et al., 2015). Two typical dates during summer and winter in Nanjing (118.54°E, 31.56°N), July 29, 2015 (summer) and January 15, 2016 (winter), were selected for the simulation. The simulation was initialized with input parameters (including meteorological, soil, and building material data) identical to those of Jiang et al. (2018), and the simulation was run for 12 h from 06:00 to 18:00 local time (UTC + 8). The urban surfaces are divided into 16 components: trees, grass, bare ground, roofs, and east-, west-, south-, and north-facing walls under both sunlit and shaded situations (Zhan et al., 2012). The component surface temperatures were then estimated as the averages of the same component.

2.1.3. Step (3): Generation of all-direction DRTs

By combining the component temperatures simulated in Step (2), CoMSTIR (Computer Model to Simulate the Thermal Infrared Radiation of 3-D urban targets) was used to obtain the all-direction DRTs. CoMSTIR is a forward sensor view model for UTRD simulation, with typical urban surfaces being divided into 10 components (Wang et al., 2020); it has been demonstrated to have acceptable accuracy (represented by RMSE) of ~1.0 K based on in-situ data (Ma et al., 2013). For this study, three typical observation times (including the typical overpass times of most sun-synchronous polar orbiters) were selected to simulate the daytime DRTs of all nine USMs: 11:00, 13:00, and 15:00 on Jul. 29, 2015 (summer) and Jan. 15, 2016 (winter). The viewing zenith angles (VZAs) were set from nadir (0°) to 60° with an increment of 10°, covering the VZAs of most polar orbiters. The viewing azimuth angles (VAAs) (0–360°) were set with an increment of 30°. The sensor observation distance and the sensor field of view (FOV) were set as 500 m and 20°, respectively, where the FOV effect can be ignored (Zhan et al., 2010).

Following these steps, forward-modelling datasets include 54 scenarios (nine USMs and three cases for summer and winter respectively), each with 73 DRTs. To assess the impact of building height on the comparison of model performance, we divided the nine USMs into four categories, including high-rise (USM01 and USM04, 12 scenarios), mid-rise (USM02 and USM05, 12 scenarios), low-rise (USM03, USM06, USM08, and USM09, 24 scenarios), and hybrid-rise (USM07, 6 scenarios). To assess the influence of observation season on model comparison, these 54 scenarios were divided into summer (27 scenarios) and winter (27 scenarios).

In addition to building height and observation season, the UTRD can also be impacted by controls such as street orientation (related to urban surface morphology) and city latitude (related to solar radiation) (Voogt, 2008; Lagouarde et al., 2010; Krayenhoff and Voogt, 2016; Wang et al., 2018a; Hu and Wendel, 2019). Similar to *Steps (1)* to *(3)*, we therefore also generated two supplementary datasets to assess the impacts of street orientation and city latitude on model comparison. To assess the impact of street orientation on model comparison, the original USM02 (with south-to-north streets) was rotated clockwise by 15°, 30°, 45°, 60°, and 75° to obtain five new sub-USMs. They are termed USM02-15°, USM02-30°, USM02-45°, USM02-60°, and USM02-75°, respectively, with the original urban surface model termed USM02-0°. A total of 36 scenarios were generated to assess the impact of street orientation. To assess the impact of city latitude (i.e., solar zenith angle variation), we took USM01 as an example and set the forward-modelling at another four latitudes corresponding to four cities, i.e., 45.93°N (Harbin), 40.08°N (Beijing), 26.08°N (Fuzhou), and 20°N (Haikou), in addition to 31.56°N (Nanjing). A total of 30 scenarios were produced to assess the impact of city latitude.

2.2. Satellite datasets

In addition to forward-modelling datasets, satellite-derived anisotropy datasets can provide observational evidence to support the evaluation of parametric model performances. However, when compared with forward-modelling, it is more difficult to simultaneously obtain multi-angle LST data from direct satellite observations. The combination of two or more geostationary sensors (e.g., GOES-EAST and GOES-WEST) with fixed observation angles is applicable for investigating the thermal anisotropy over natural surfaces (Vinnikov et al., 2012). Nevertheless, the spatial resolutions of most geostationary thermal sensors (usually ~5 km) are too coarse relative to the size of a city. Along-track polar-orbiters with observation geometry similar to the Along Track Scanning Radiometer −2 can acquire finer (with the resolution of ~1 km) multi-angle LST observations quasi-simultaneously, but the angle number remains limited for model performance assessments. Cross-track polar-orbiters such as MODIS also sample the surface with a spatial resolution (also ~1 km) which is sufficiently fine to observe cities, and with wider fields of view that enable the acquisition of sufficient multi-angle thermal data with various VZAs from nadir to ±65°. However, MODIS is unable to acquire multi-angle LSTs simultaneously – its multi-angle thermal data are obtained from temporally-adjacent days (Wan, 2006) and consequently the original MODIS LSTs cannot be directly used for model evaluations (Cao et al., 2019a).

The non-synchronization of multi-angle MODIS LSTs can be partly overcome by the strategy proposed by Hu et al. (2016a). To eliminate LST variations induced by weather changes and differences in atmospheric path radiance on adjacent days, this strategy uses homogeneous water bodies adjacent to cities, which exhibit minimal thermal anisotropy, as a reference. With this strategy, the LSTs that are only related to the observation angle (i.e., the quasi-simultaneous multi-angle LSTs) can be obtained. This strategy consists of two steps (Hu et al., 2016a): (1) Calculation of the LST difference (ΔLST) and air temperature difference (ΔT_{air}) between land surfaces and water bodies; and (2) generation of the quasi-simultaneous multi-angle MODIS LSTs ($\Delta LST_{\Delta T_{air}}$) by subtracting the anomaly of ΔT_{air} from ΔLST with a magnitude adjustment in a time series.

In this study, the quasi-simultaneous multi-angle MODIS LSTs (i.e., the $\Delta LST_{\Delta T_{air}}$ dataset) produced by Hu et al. (2016a) over Chicago (87.6°W, 41.9°N) and New York City (74°W, 40.7°N) were directly employed for model comparisons. For these two cities, the LST products used are the daytime L2 products (i.e., the MOD11_L2 from Terra and MYD11_L2 from Aqua) acquired at ~11:00 (Terra day) and 13:00 (Aqua day) local time, each day. The MODIS LSTs were collected within the growing season (May to September) from a 10-year period (2003−2012), mostly considering that (1) urban vegetation is relatively stable in the growing season, and (2) that observation and retrieval noise can be significantly reduced by using the average of data for a 10-year

period.

For these MODIS DRTs, the VAAs are relatively fixed while the VZAs vary from −65° (westerly) to +65° (easterly) (see Fig. A2 in Appendix A for the Chicago case). The MODIS DRTs were further averaged according to the VZA in each 5°. The corresponding solar positions are also the averages during the corresponding observational periods for each 2-h overpass window (see Fig. A2 in Appendix A for the Chicago case). To further evaluate the impact of urban morphology on model performance, the MODIS DRTs were categorized into nine groups by urban fraction (urban%), including (0, 10%), [10%, 20%), [20%, 30%), …, [70%, 80%), [80%, 100%]. Altogether, the MODIS DRT dataset consists of 36 scenarios (9 groups by urban fraction, 2 observation times during the day, and 2 cities), each with DRTs at 27 different VZAs. For this dataset, the maximum daytime thermal anisotropy intensity can be up to 9.0 K (Hu et al., 2016a). Readers can refer to Hu et al. (2016a) for more information about the 36 DRT lines.

3. Summary of kernel-driven models and schemes for model comparison

This section summarizes the selected parametric kernel-driven models for comparison, as well as the associated comparison schemes. According to the general TIR kernel-driven modelling framework proposed by Cao et al. (2021), we mainly discuss four base shape kernels and three hotspot kernels suitable for UTRD simulation (see Section 3.1). By integrating these available hotspot and base shape kernels, eleven parametric models with either single or dual kernels were obtained, including six new models and five existing models (see Section 3.2). We finally provide three schemes to evaluate these parametric models (see Section 3.3).

3.1. Available parametric kernels for UTRD simulation

The general TIR kernel-driven modelling framework can be written as follows (Cao et al., 2021):

$$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{BaseShape} \cdot K_{BaseShape} + f_{Hotspot} \cdot K_{Hotspot} \quad (1)$$

where $T(\theta_s, \theta_v, \varphi)$ represents the DRT, wherein θ_s, θ_v, and φ are the solar zenith angle (SZA), viewing zenith angle (VZA), and sun-sensor relative azimuth angle, respectively; f_{iso}, $f_{BaseShape}$, and $f_{Hotspot}$ are the isotropy coefficient, base shape kernel coefficient, and hotspot kernel coefficient, respectively; and $K_{BaseShape}$ and $K_{Hotspot}$ are the base shape and hotspot kernels used to simulate the base shape and shadowing effect anisotropy, respectively.

According to the general TIR kernel-driven modelling framework, we discuss four base shape kernels (see Section 3.1.1) and three hotspot kernels (see Section 3.1.1.3) suitable for UTRD simulation. Noting that we disregarded the Li-related kernels (i.e., the Li-sparse (Wanner et al., 1995), Li-dense (Wanner et al., 1995), upgraded Li-sparseR (Lucht et al., 2000), and Li-transit kernels (Yang et al., 2002)) because they were designed for vegetation canopy, contain vegetation structure parameters (e.g., the crown size), and are consequently inappropriate for urban surfaces characterized primarily by cuboid-like buildings.

3.1.1. Four base shape kernels suitable for UTRD simulation

3.1.1.1. Ross-thick kernel ($K_{BaseShape\_thick}$). The Ross-thick kernel ($K_{BaseShape\_thick}$) was proposed to describe the volume scattering of dense vegetation (Roujean et al., 1992; Jupp, 2000). $K_{BaseShape\_thick}$ is applicable as a base shape kernel for simulating the thermal anisotropy of urban surfaces with dense vegetation. The formula of $K_{BaseShape\_thick}$ is

given as:

$$\begin{cases} K_{BaseShape\_thick}(\theta_s, \theta_v, \varphi) = \dfrac{(\pi/2 - \xi)cos\xi + sin\xi}{cos\theta_s + cos\theta_v} - \dfrac{\pi}{4} \\ cos\xi = cos\theta_s cos\theta_v + sin\theta_s sin\theta_v cos\varphi \end{cases} \quad (2)$$

where ξ is the phase angle related to the sun-surface-sensor position (Roujean et al., 1992). The meanings of the angle-related parameters in Eq. (2) are identical to those given in Eq. (1). The $K_{BaseShape\_thick}$ is determined by VZA, VAA, and sun-sensor relative azimuth angle. Its kernel values along the solar principle plane (SPP) with SZA = 30° and SAA = 180° are illustrated in Fig. 1a.

3.1.1.2. Ross-thin kernel ($K_{BaseShape\_thin}$). The Ross-thin kernel ($K_{BaseShape\_thin}$) is similar to $K_{BaseShape\_thick}$ but was designed for a vegetation canopy with a relatively low leaf area index (Wanner et al., 1995; Jupp, 2000). $K_{BaseShape\_thin}$ can be used as a base shape kernel for simulating the thermal anisotropy of urban surfaces with sparse vegetation. The formula of $K_{BaseShape\_thin}$ is given as:

$$K_{BaseShape\_thin}(\theta_s, \theta_v, \varphi) = \dfrac{(\pi/2 - \xi)cos\xi + sin\xi}{cos\theta_s cos\theta_v} - \dfrac{\pi}{2} \quad (3)$$

The kernel values of $K_{BaseShape\_thin}$ along the SPP with SZA = 30° and SAA = 180° are given in Fig. 1b.

3.1.1.3. Base shape kernel of the Vinnikov model ($K_{BaseShape\_vin}$). The semi-empirical Vinnikov model was proposed by Vinnikov et al. (2012) for simulating the thermal anisotropy of natural surfaces. The capability of the Vinnikov model for simulating vegetation thermal anisotropy was further confirmed (Duffour et al., 2016; Ermida et al., 2018b; Liu et al., 2018; Cao et al., 2019a). It has been recently shown applicable for UTRD simulation (Jiang et al., 2018). The base shape kernel (denoted by $K_{BaseShape\_vin}$) of the Vinnikov model can be written as follows:

$$K_{BaseShape\_vin}(\theta_v) = 1 - cos\theta_v \quad (4)$$

The $K_{BaseShape\_vin}$ is only VZA-dependent. Its kernel values along the SPP with SZA = 30° and SAA = 180° are displayed in Fig. 1c.

3.1.1.4. Base shape kernel of the USEA model ($K_{BaseShape\_uea}$). Based on the Vinnikov model, Sun et al. (2015) proposed the USEA model to better simulate UTRD by replacing the original base shape kernel with a revised kernel $K_{BaseShape\_uea}$, given as follows:

$$K_{BaseShape\_uea}(\theta_v) = sin\theta_v \quad (5)$$

Similar to $K_{BaseShape\_vin}$, the $K_{BaseShape\_uea}$ is also only VZA-dependent. Its kernel values along the SPP with SZA = 30° and SAA = 180° are demonstrated in Fig. 1c.

3.1.2. Three hotspot kernels suitable for UTRD simulation

3.1.2.1. Roujean kernel ($K_{Hotspot\_rou}$). The Roujean model (hereafter termed the ROU model) was originally developed by Roujean et al. (1992) for BRDF simulation and was later proved capable of serving as a hotspot kernel for simulating vegetation thermal anisotropy. The Roujean kernel (given as $K_{Hotspot\_rou}$) is obtained based on the existence of identical long wall-like protrusions (similar to buildings) within a scene. The ROU model has not previously been used for UTRD simulation but we incorporate this model for comparison because its assumption of wall-like protrusions within a scene makes it suitable for UTRD simulation. The $K_{Hotspot\_rou}$ can be written as follows:

$$\begin{cases} K_{\text{Hotspot\_rou}}(\theta_s, \theta_v, \varphi) = \dfrac{1}{2\pi}[(\pi - \varphi)cos\varphi + sin\varphi\,]tan\theta_s tan\theta_v - \dfrac{1}{\pi}(tan\theta_s + tan\theta_v + D) \\ D = \sqrt{\tan^2\theta_s + \tan^2\theta_v - 2tan\theta_s tan\theta_v cos\varphi} \end{cases} \tag{6}$$

where D is the angular distance between the sun-sensor. The hotspot position simulated by $K_{\text{Hotspot\_rou}}$ is consistent with the solar location; and the $K_{\text{Hotspot\_rou}}$ values along the SPP with SZA = 30° and SAA = 180° are given in Fig. 1d.

3.1.2.2. Hotspot kernel of the Vinnikov model ($K_{\text{Hotspot\_vin}}$). The hotspot kernel (denoted by $K_{\text{Hotspot\_vin}}$) of the Vinnikov model can be expressed as follows:

$$K_{\text{Hotspot\_vin}}(\theta_s, \theta_v, \varphi) = sin\theta_v cos\theta_s sin\theta_s cos(\theta_s - \theta_v)cos\varphi \tag{7}$$

The $K_{\text{Hotspot\_vin}}$ values along the SPP are demonstrated in Fig. 1e. As a pure semi-empirical kernel, the hotspot position (i.e., the maximum kernel value) simulated with $K_{\text{Hotspot\_vin}}$ (i.e., VZA = 60° and VAA = 180°) can be inconsistent with the solar position (i.e., SZA = 30° and SAA = 180°).

3.1.2.3. Hotspot kernel of the RL model ($K_{\text{Hotspot\_rl}}$). The parametric hotspot model initially used for BRDF simulation (Roujean, 2000) was demonstrated to be capable of simulating UTRD (Lagouarde and Irvine, 2008; Duffour et al., 2016). Under the condition that the DRT at nadir (T_0) is known, the RL model contains two parameters: the DRT difference between nadir and hotspot (ΔT_{HS}) and the scale factor (k). It is defined by the following formula:

$$\begin{cases} T(\theta_s, \theta_v, \varphi) = f_{\text{iso}} + f_{\text{Hotspot}} \cdot K_{\text{Hotspot\_rl}} \\ K_{\text{Hotspot\_rl}} = \dfrac{e^{-kD} - e^{-ktan\theta_s}}{1 - e^{-ktan\theta_s}} \end{cases} \tag{8}$$

The kernel values of $K_{\text{Hotspot\_rl}}$ along the SPP with different k values (corresponding to different hotspot widths) are demonstrated in Fig. 1f. Similar to $K_{\text{Hotspot\_rou}}$, the hotspot position simulated with $K_{\text{Hotspot\_rl}}$ is consistent with the solar position.

3.2. Illustration of 11 parametric kernel-driven models

After screening the usable kernels as introduced in Section 3.1, we obtained six candidates, including two hotspot kernels ($K_{\text{Hotspot\_rou}}$ and $K_{\text{Hotspot\_vin}}$) and four base shape kernels ($K_{\text{BaseShape\_thin}}$, $K_{\text{BaseShape\_thick}}$, $K_{\text{BaseShape\_vin}}$, and $K_{\text{BaseShape\_uea}}$). Here, $K_{\text{Hotspot\_rl}}$ was treated individually and did not participate in the kernel combination. This is mainly because $K_{\text{Hotspot\_rl}}$ possesses a specific unknown parameter (i.e., the scale factor k), while the other six candidates are only determined by the sun-surface-sensor geometry (i.e., θ_s, θ_v, and φ).

According to the general TIR kernel-driven modelling framework as given by Eq. (1), we obtained 11 parametric kernel-driven models, including three single-kernel models and eight dual-kernel models (Table 1). To ensure consistency in the abbreviations of the parametric models, we used 3-letter abbreviations for all 10 models except the RL model. The abbreviation of a single-kernel model is given as the first three letters of the last name of the associated first author, e.g., the
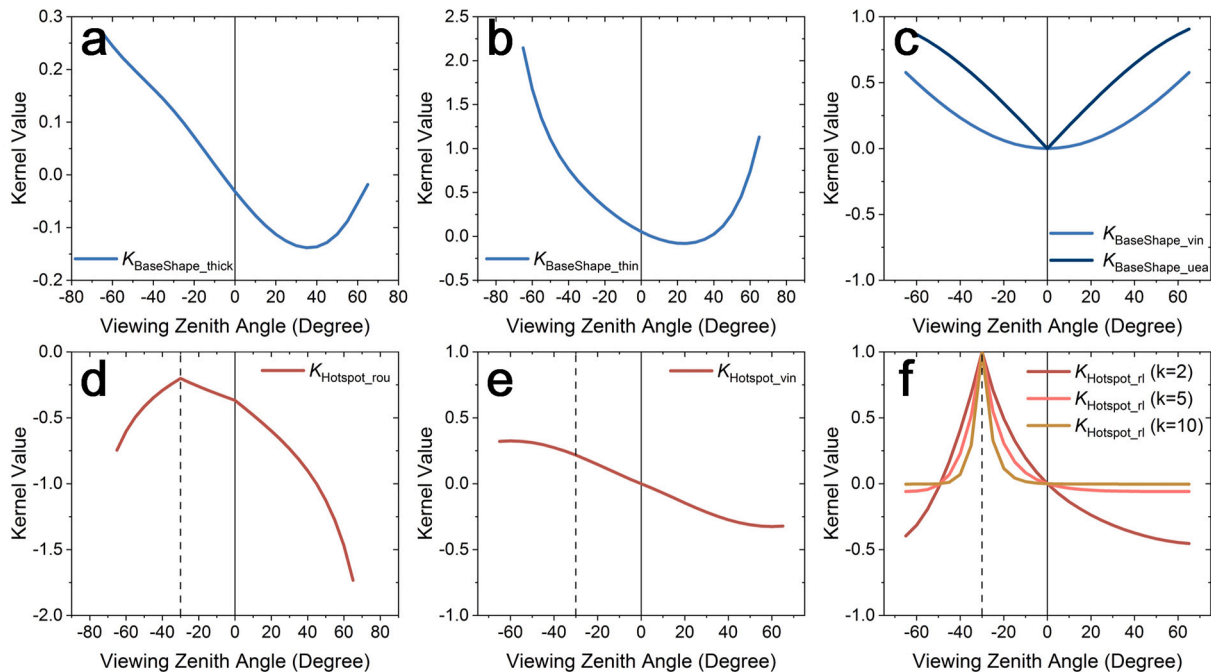

Fig. 1. Kernel values along the solar principle plane (SPP), with the SZA and SAA designated respectively as 30° and 180° for demonstrating a typical scenario. The negative and positive VZAs correspond to the VAA being equivalent to 180° and 0°, respectively; and the solid straight and dashed lines represent the VZA being equal to 0° (i.e., nadir) and − 30° (i.e., the designated sun position), respectively. (a) and (b) show the kernel values for $K_{\text{BaseShape\_thick}}$ and $K_{\text{BaseShape\_thin}}$, respectively; (c) illustrates the kernel values for $K_{\text{BaseShape\_vin}}$ and $K_{\text{BaseShape\_uea}}$, respectively; (d) and (e) display the kernel values for $K_{\text{Hotspot\_rou}}$ and $K_{\text{Hotspot\_vin}}$, respectively; and (f) demonstrates the kernel values for $K_{\text{Hotspot\_rl}}$ with different scale factors (k = 2, 5, and 10).

Roujean model is abbreviated as the ROU model. The abbreviation of a dual-kernel model consists of the first letter of the name of the used hotspot kernel (i.e., 'R' stands for $K_{Hotspot\_rou}$, 'V' stands for $K_{Hotspot\_vin}$) plus the first two letters of the name of the used base shape kernel (i.e., 'TH' stands for $K_{BaseShape\_thin}$, 'TK' stands for $K_{BaseShape\_thick}$, 'VI' stands for $K_{BaseShape\_vin}$, and 'US' stands for $K_{BaseShape\_uea}$). According to this abbreviation rule, the Vinnikov model, for example, can be abbreviated as the VVI model; and the USEA model can be abbreviated as the VUS model.

Amongst these 11 parametric models, only three have been tested for UTRD simulation over typical urban surfaces with buildings: the RL model (Duffour et al., 2016), the VVI model (tested in Jiang et al. (2018)), and the VUS model (tested in Sun et al. (2015)); two (i.e., the RTH and RTK models) will be tested for the UTRD simulation for the first time; the remaining six are newly derived models that will be tested for UTRD simulation, including two single-kernel and four dual-kernel models. The two new single-kernel models use hold one hotspot kernel with the base shape kernel disregarded (Vinnikov et al., 2012), i.e., $K_{Hotspot\_rou}$ for the ROU model and $K_{Hotspot\_vin}$ for the VIN model (Table 1). The four new dual-kernel models incorporate one hotspot kernel and one base shape kernel. They include the VTH, VTK, RVI, and RUS models (Table 1). Possible issues related to this empirical kernel combination will be discussed in Section 4.4.2.

Compared with the single- and dual-kernel models, the multi-kernel models characterized by more parameters are expected to have higher accuracies; however, they are also expected to have a less stable performance when the number of input DRT is limited, which is the case for satellite thermal observations. Therefore, in this study we focused mainly on the assessments of single- and dual-kernel models. Nevertheless, we provided a quantitative analysis and discussion of the performances of multi-kernel models. Additional details of the formulae and performances of multi-kernel models are therefore not provided in Section 3 but in Section 4.4.1 and Appendix B.

3.3. Schemes for model comparison

The performances of the 11 kernel-driven models were assessed by the following three schemes. The first scheme (termed Scheme #1 hereafter) employed all the DRTs simulated by CoMSTIR to test the performances of parametric models. The model performances were evaluated by comparing the simulated DRTs provided by the parametric models and the forward-modelling dataset (more details are given in

Table 1
Summary of the 11 kernel-driven models.

Model type	Model name	Model expression
Single-kernel models	ROU[b]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rou}$
	VIN[b]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_vin}$
	RL[a]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rl}$
Dual-kernel models	RTH[a]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rou} + f_{BaseShape} \cdot K_{BaseShape\_thin}$
	VTH[b]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_vin} + f_{BaseShape} \cdot K_{BaseShape\_thin}$
	RTK[a]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rou} + f_{BaseShape} \cdot K_{BaseShape\_thick}$
	VTK[b]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_vin} + f_{BaseShape} \cdot K_{BaseShape\_thick}$
	RVI[b]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rou} + f_{BaseShape} \cdot K_{BaseShape\_vin}$
	VVI[a]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_vin} + f_{BaseShape} \cdot K_{BaseShape\_vin}$
	RUS[b]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rou} + f_{BaseShape} \cdot K_{BaseShape\_uea}$
	VUS[a]	$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_vin} + f_{BaseShape} \cdot K_{BaseShape\_uea}$

Superscripts 'a' and 'b' respectively represent the five published and six newly-derived parametric models.

Section 4.1). This is the major scheme used in this study as the forward-modelling dataset provides abundant observation angles and scenarios that ensure a relatively comprehensive assessment. Note that quantitative evaluation of the performances of the multi-kernel models is also conducted under Scheme #1.

The second scheme (Scheme #2) is similar to Scheme #1 except that all the MODIS DRTs along the scanning line from east to west were used (refer to Section 4.2). The evaluations under Scheme #2 can provide a reference for the application of parametric models to actual satellite LSTs: e.g., the angular normalization of urban satellite LSTs. Notably, the number of observation angles and scenarios for Scheme #2 are less than that for Scheme #1, so this scheme is used to supplement Scheme #1.

The third scheme (Scheme #3) tested model sensitivity (or stability) with only a limited number of DRTs from the forward-modelling dataset (see Section 4.3). Scheme #3 is based on the forward-modelling rather than the satellite datasets because the former provides more data for model sensitivity assessments. The limited DRT inputs used to fit parametric models are determined using the following two steps. The first step is to determine the number of DRT inputs each time. The single-kernel ROU and VIN models require at least two DRTs to solve the unknown parameters, while the RL model and the dual-kernel models require at least three DRTs; in other words, the number of DRT inputs should be equal to or greater than three. We added one more DRT input each time (i.e., therefore there are four DRT inputs) to increase the model solvability. The second step is to determine the angular positions of the four DRT inputs within the upper hemisphere. In general, the angular positions of these four DRTs were generated randomly. However, we further complied with the following two rules to refine the random selection, in order to avoid systematic errors in model inversion caused by the possible aggregation of the four DRTs once the entirely random selection was used. (i) The VZA for of the four DRTs should be different; and (ii) the VAA difference between two adjacent DRTs should be 60–120°; e.g., if the VAA of the first DRT is between 0 and 30°, those of the other three DRTs are 90–120°, 180–210°, and 270–300°. In total, we generated 8100 groups of four-DRTs based on the forward-modelling dataset (i.e., 54 scenarios, 150 four-DRT groups for each scenario).

Three indicators were used for assessing model performance, including the mean absolute error (MAE), root mean square error (RMSE), and distance of hotspots (DHS). Amongst these, the DHS is a newly-defined indicator, which can be calculated by the spherical distance formula (i.e., the Great-circle distance formula) in geometry (Bullock, 2007). The DHS is used to evaluate the model capability for simulating hotspot location. The DHS is calculated as the angular distance between the 'reference' and fitted hotspots, with the following formula:

$$DHS = arccos\left[cos(\pi/2 \text{-} \theta_r)cos(\pi/2 \text{-} \theta_f)cos\varphi_{r\_f} + sin(\pi/2 \text{-} \theta_r)sin(\pi/2 \text{-} \theta_f)\right]$$

(9)

where θ_r and θ_f are the VZAs of the reference and fitted hotspots, respectively; and φ_{r_f} is the relative VAA between these two hotspots.

4. Results and discussion

4.1. Comparison with forward-modelling data

4.1.1. Comparison of the overall performance of parametric models
The overall performances of the 11 parametric kernel-driven models under Scheme #1 (with all DRTs used as input) are shown in Fig. 2. The results show that the dual-kernel models generally have a higher accuracy and are more robust than the single-kernel models. The average RMSE of the dual-kernel models is ~0.30 K less than that of the single-kernel models, and the error distributions of the dual-kernel models are more concentrated (Fig. 2). The better performance of the dual-kernel models for UTRD simulation can be attributed to their combination of
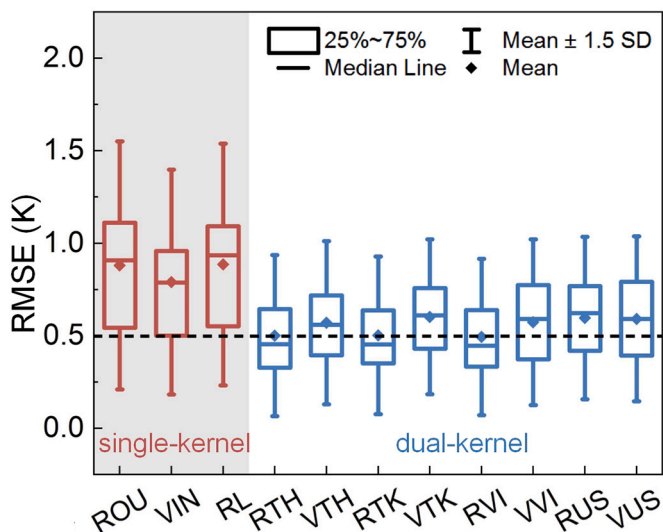
Fig. 2. RMSEs of the 11 parametric kernel-driven models under Scheme #1. The top and bottom whiskers indicate the highest and lowest value within mean ± 1.5 times standard deviation (SD), respectively; the top and bottom of a box are the 25th and 75th percentile of the RMSEs, respectively; and the middle line and point within a box denote the median and mean RMSE, respectively. The black horizontal dashed line corresponds to the lowest mean RMSE. Note that the RMSEs are averaged based on all nine USMs at six typical times.

a hotspot kernel (which is also used by the single-kernel models) and a base shape kernel. Deviations in performance are also observed within the dual- or single-kernel model categories.

Amongst the single-kernel models, the VIN model has the highest accuracy (RMSE: 0.79 K), followed by the ROU and RL model (RMSE of 0.88 and 0.89 K, respectively). The superiority of the VIN model over the two other single-kernel models lies in its better performance in modelling the DRTs over the angular ranges of 55–60°/120–150° and 20–40°/0–60° (VZA/VAA) (see Fig. 3).

Amongst the dual-kernel models, the RVI model has the lowest RMSE (0.49 K), followed by the RTH and RTK models (the RMSEs are both 0.50 K). In addition, once given a determined base shape kernel, the dual-kernel models with the hotspot kernel $K_{Hotspot\_rou}$ are generally better than those with the hotspot kernel $K_{Hotspot\_vin}$. For example, when

$K_{BaseShape\_thick}$ is taken as the base shape kernel, the RTK model with $K_{Hotspot\_rou}$ (the RMSE is 0.50 K) performs better than the VTK model with $K_{Hotspot\_vin}$ (the RMSE is 0.60 K). The better performance of the $K_{Hotspot\_rou}$-based dual-kernel models is the most significant within the angular range of 50–60°/255–285° (VZA/VAA) (Fig. 4). The superiority of $K_{Hotspot\_rou}$ may be due to its assumption that surface objects are long-walled protrusions, which coincides well with the geometry of typical urban surfaces dominated by rectangular buildings. In contrast, as a pure empirical kernel, $K_{Hotspot\_vin}$ was proposed based on a rough statistical regression of DRTs over natural surfaces (Cao et al., 2019a). It therefore becomes less capable for thermal directionality simulation over typical urban surfaces dominated by rectangular buildings. The RMSE differences between models with different base shape kernels are generally less than those with different hotspot kernels (except for the VUS model). This indicates that, in terms of the performance of kernel-driven models, the influence of the base shape kernel is less than that of the hotspot kernel. For example, the RMSEs of the RTH (by combining $K_{Hotspot\_rou}$ and $K_{BaseShape\_thin}$) and RTK (by combining $K_{Hotspot\_rou}$ and $K_{BaseShape\_thick}$) models are both 0.50 K, while the RMSE of the VTH model (by combining $K_{Hotspot\_vin}$ and $K_{BaseShape\_thin}$) is 0.57 K. The higher dependence of model performance on the hotspot rather than the base shape kernel during daytime is reasonable because daytime UTRD is primarily determined by the differences in component temperatures caused by the sunlit and shaded conditions of urban surfaces under strong solar radiation.

We now further analyze the comparison of model performance under scenarios with different building heights and observation seasons.

Building height directly affects the relative proportions of both horizontal and vertical surface components, which in turn affect the UTRD (Krayenhoff and Voogt, 2016; Hu and Wendel, 2019). For example, the difference between the maximum thermal anisotropy intensities (i.e., the difference in the maximum minus minimum DRTs) of USM04 and USM06 at the same observation time (i.e., 11:00 local time in summer) can be up to 2.0 K (6.0 K vs. 4.0 K). The evaluations in terms of building height illustrate that the model fitting error generally increases with building height (Fig. 5), mostly due to the more significant sunlit-shaded conditions with a greater building height. In spite of the large influence of building height on model performance, the results show that this would rarely affect the relative performance ranking of these models.

Observational season largely determines the level of solar radiation and therefore affects the UTRD regime (Lagouarde et al., 2010). For example, the maximum difference in DRTs of USM01 at 13:00 local time
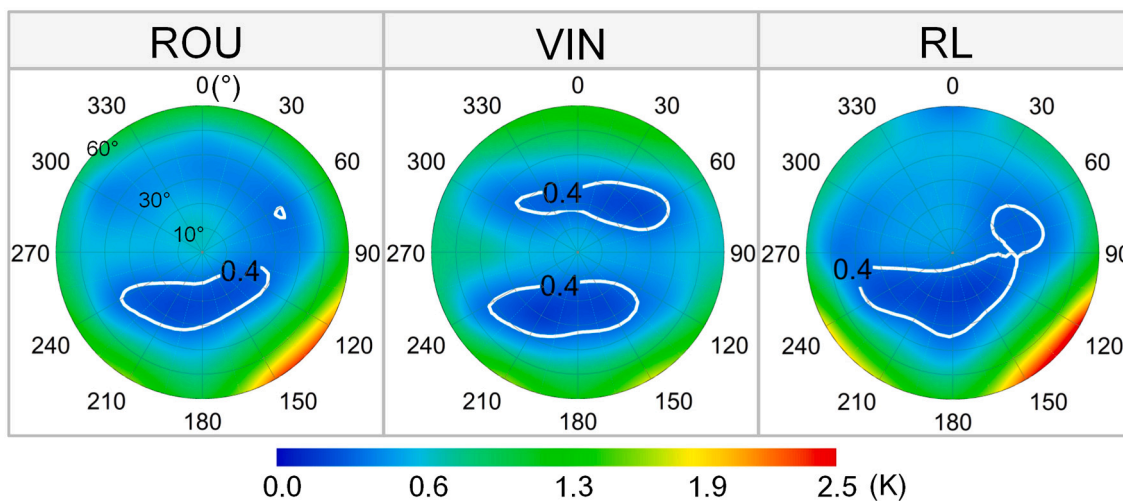

Fig. 3. Polar diagrams of the MAEs of the single-kernel models. The concentric circles correspond to the off-nadir viewing zenith angles (φ_v, 0–360°) whereas the radial lines indicate the viewing azimuth angles (θ_v, 0–60°).
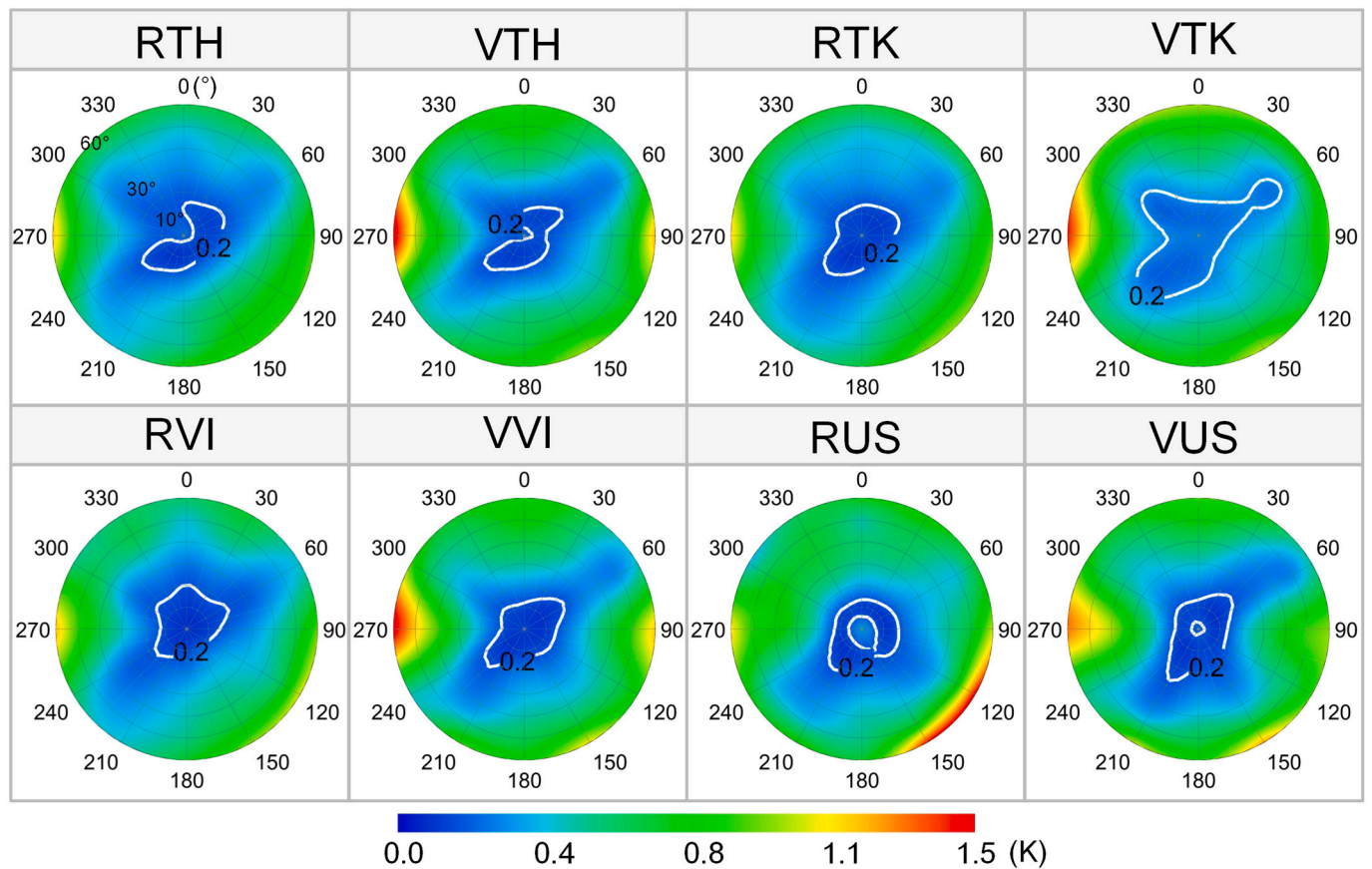
Fig. 4. Polar diagrams of the MAEs of the dual-kernel models.

in winter (13.3 K) can be twice that in summer (6.3 K). To assess the influence of observation season on the model performances, we compared the simulation errors of models in both summer and winter (Fig. 6). The results show that observation season generally exerts a smaller impact on model performance when compared with building height (Fig. 6). For the same model, the fitting RMSE in summer is only slightly larger than that in winter (with ΔRMSE of 0.03–0.06 K on average), whereas ΔRMSE between the high- and low-rise buildings can reach 0.33–0.49 K (Figs. 5 and 6). Nevertheless, observation season would not substantially affect the relative ranking of model performance, which is similar to that for building height.

In addition to building height and observation season, additional assessments of street orientations and city latitude also confirm that urban surface morphology and solar radiation affect the fitting error of the models but they rarely affect the performance ranking (detailed assessments are given in Appendix C).

4.1.2. Comparison based on observation angle and observation time

The overall assessments in Section 4.1.1 are unable to provide a detailed view of the performance differences amongst the models. We now further examine the fitting errors at each observation time (Table 2) and the accuracy of the simulated DTRs depending on the observation angle. The assessments based on the analysis of 594 DRT polar diagrams under 54 scenarios (see Section 3.3) indicate that the performance differences between the models can largely be explained by their ability to simulate DRTs and, in particular, their ability to describe hotspots. We take the UTRD simulation of USM01 at 13:00 as an example to examine the ability of the 11 parametric models to simulate hotspots (Figs. 7 and 8).

For the single-kernel models, the results show that the VIN model is

superior to the other two single-kernel models (i.e., the ROU and RL models) in winter (Table 2). In summer, the VIN model remains superior to the ROU and RL models at 11:00 and 15:00, but it becomes less competent than the other two at 13:00 (Fig. 7 and Table 2). The complexity of model performance comparison in summer may be related to their associated capabilities in simulating hotspot location for different observation times (different SZAs). When the SZA is relatively higher (e.g., SZA > 30°; in winter or at 11:00 and 15:00 in summer), all three models describe the hotspot locations more accurately. In contrast, when the SZA is small (e.g., SZA < 30° at 13:00 in summer), the hotspot location simulated by the VIN model is less accurate than those obtained by the RL and ROU models (Fig. 7). The differences in model design likely contribute to this contrast in model performance in different seasons and observation times (i.e., different SZAs). Unlike the ROU ($K_{Hotspot\_rou}$) and RL models that fully consider the sensor-surface-sun geometry (including the zenith angles of the sun and sensor as well as the angular distance and the relative azimuthal angle between the sun and sensor) and set the hotspot location strictly based on the sun position during UTRD simulation, the VIN model ($K_{Hotspot\_vin}$) disregards the angular distance between the sun and sensor (i.e., the parameter 'D' in Eq. (6) for the ROU model and Eq. (8) for the RL model) (Liu et al., 2018; Cao et al., 2019a). The disregard of the angular distance between the sun and sensor for the VIN model can lead to an inaccurate simulation of hotspot location, especially when the SZA is small.

For the dual-kernel models, the performance differences between the models with the hotspot kernel $K_{Hotspot\_rou}$ and those with $K_{Hotspot\_vin}$ are more significant in summer than in winter. In summer, the models with $K_{Hotspot\_rou}$ (the RTH, RTK, and RVI models) depict more accurately both the hotspots and DRTs than those with $K_{Hotspot\_vin}$ (Table 3 and Fig. 7). The differences in the DRT distributions simulated by these different
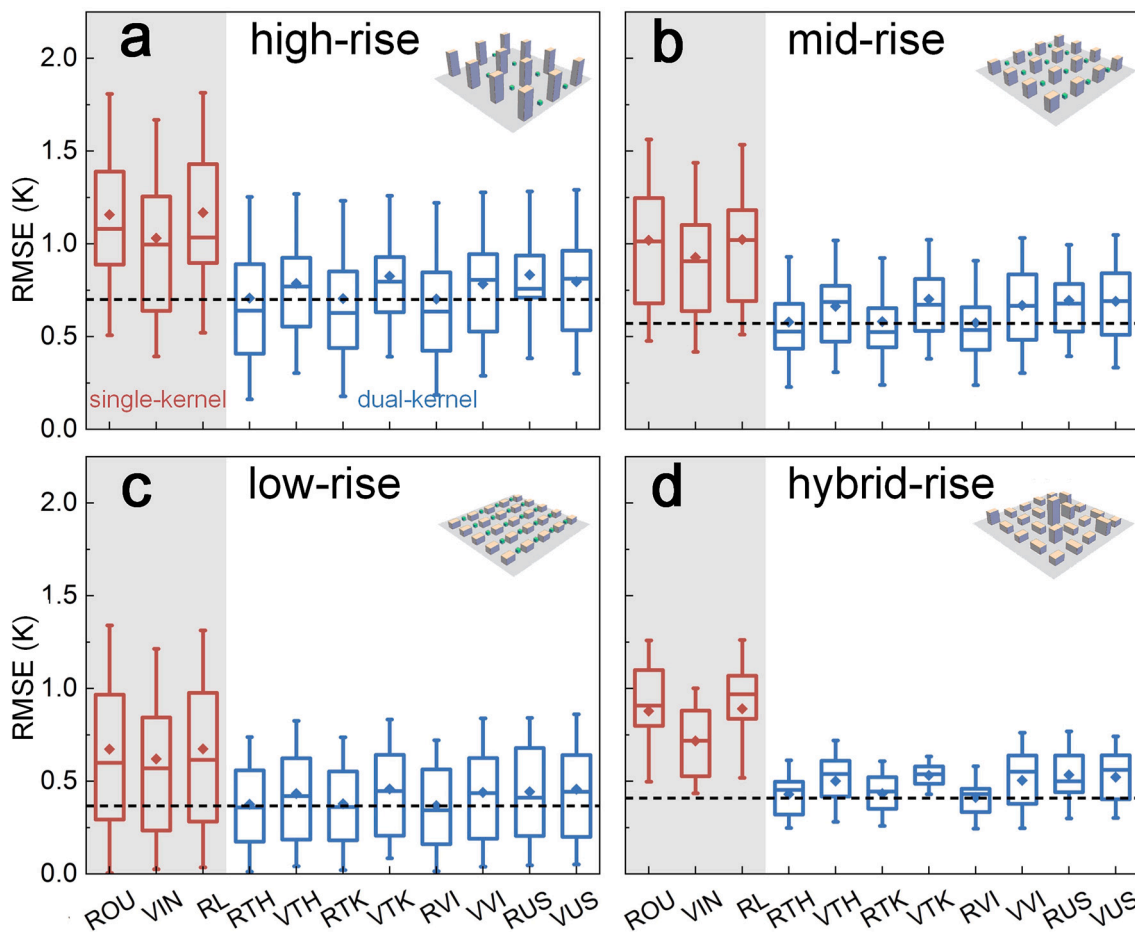
Fig. 5. Boxplots of the RMSEs. Same as Fig. 2, but for four types of USMs: (a) high-rise, (b) mid-rise, (c) low-rise, and (d) hybrid-rise, respectively.
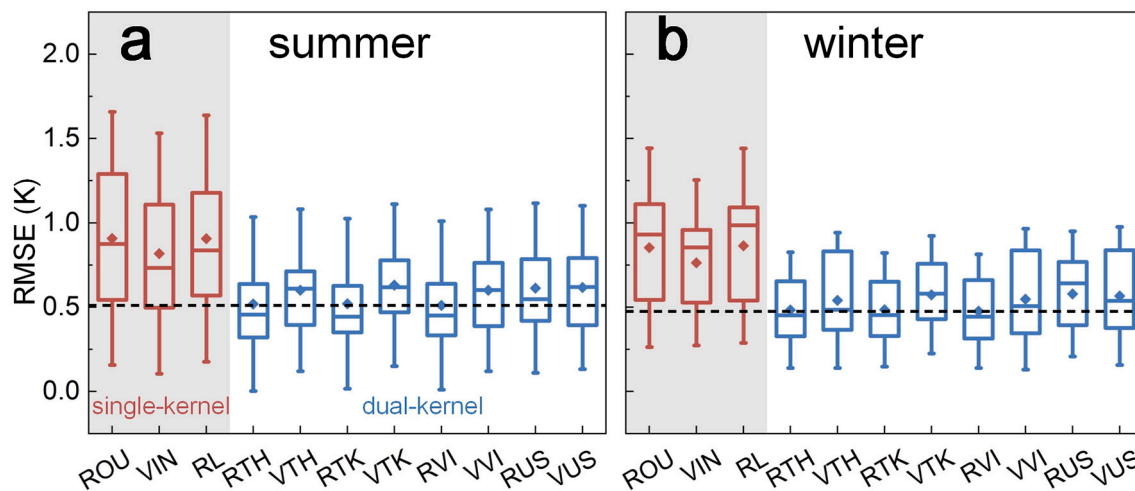
Fig. 6. Boxplots of the RMSEs. Same as Fig. 2, but for (a) summer and (b) winter, respectively.

models may be the result of the synergy of different combinations of hotspot and base shape kernels, rather than relying on only one of them (Peng et al., 2011). In winter, by comparison, each model describes the hotspot and DRT with comparable accuracy (Fig. 8). The comparable performances amongst models with $K_{Hotspot\_rou}$ and $K_{Hotspot\_vin}$ in winter is probably due to the relatively monotonous UTRD regime with a large SZA for this season (Fig. 8a).

4.2. Comparison based on satellite data

Reprocessed MODIS DRTs were used to evaluate model performances (Scheme #2). Figs. 9 and 10 show that the performances of the 11 parametric models under different scenarios over Chicago and New York City. The RMSEs of the parametric models based on MODIS DRTs are generally greater than those based on forward-modelling DRTs, and

Table 2
RMSEs (K) of the 11 parametric kernel-driven models at different observation times.

Model name	Summer			Winter		
	11:00	13:00	15:00	11:00	13:00	15:00
ROU	0.77	0.51	0.70	1.12	1.04	1.16
VIN	0.44	0.59	0.66	0.99	0.92	1.14
RL	0.76	0.50	0.75	1.09	1.01	1.19
RTH	0.34	0.36	0.42	0.53	0.54	0.82
VTH	0.37	0.38	0.65	0.56	0.67	0.79
RTK	0.36	0.36	0.44	0.51	0.52	0.82
VTK	0.42	0.59	0.65	0.53	0.62	0.80
RVI	0.34	0.36	0.43	0.51	0.53	0.79
VVI	0.58	0.45	0.50	0.61	0.62	0.82
RUS	0.37	0.38	0.64	0.63	0.74	0.80
VUS	0.37	0.38	0.64	0.63	0.74	0.80

accordingly the performance differences amongst models are relatively smaller, although the results based on MODIS LSTs remain mostly consistent with those with the simulation dataset. This may be due to the following characteristics of the MODIS LST dataset: (1) the retrieval error of the MODIS LSTs, especially over complex urban areas (1.0 K or higher) (Li et al., 2013), can contribute to the decreased model performance. (2) The MODIS LSTs are not simultaneous observation from multiple angles but are values averaged at each observation angle throughout a 10-year period. (3) The model performances, if evaluated with the MODIS DRTs, can only be reflected along a single profile rather than within the entire upper hemisphere (refer to the profiles in Fig. A2 given in Appendix A).

The results in Fig. 9 further show that the model accuracies decrease with a higher urban fraction. This is probably because the UTRD intensity for surfaces with a smaller urban fraction is usually less than that
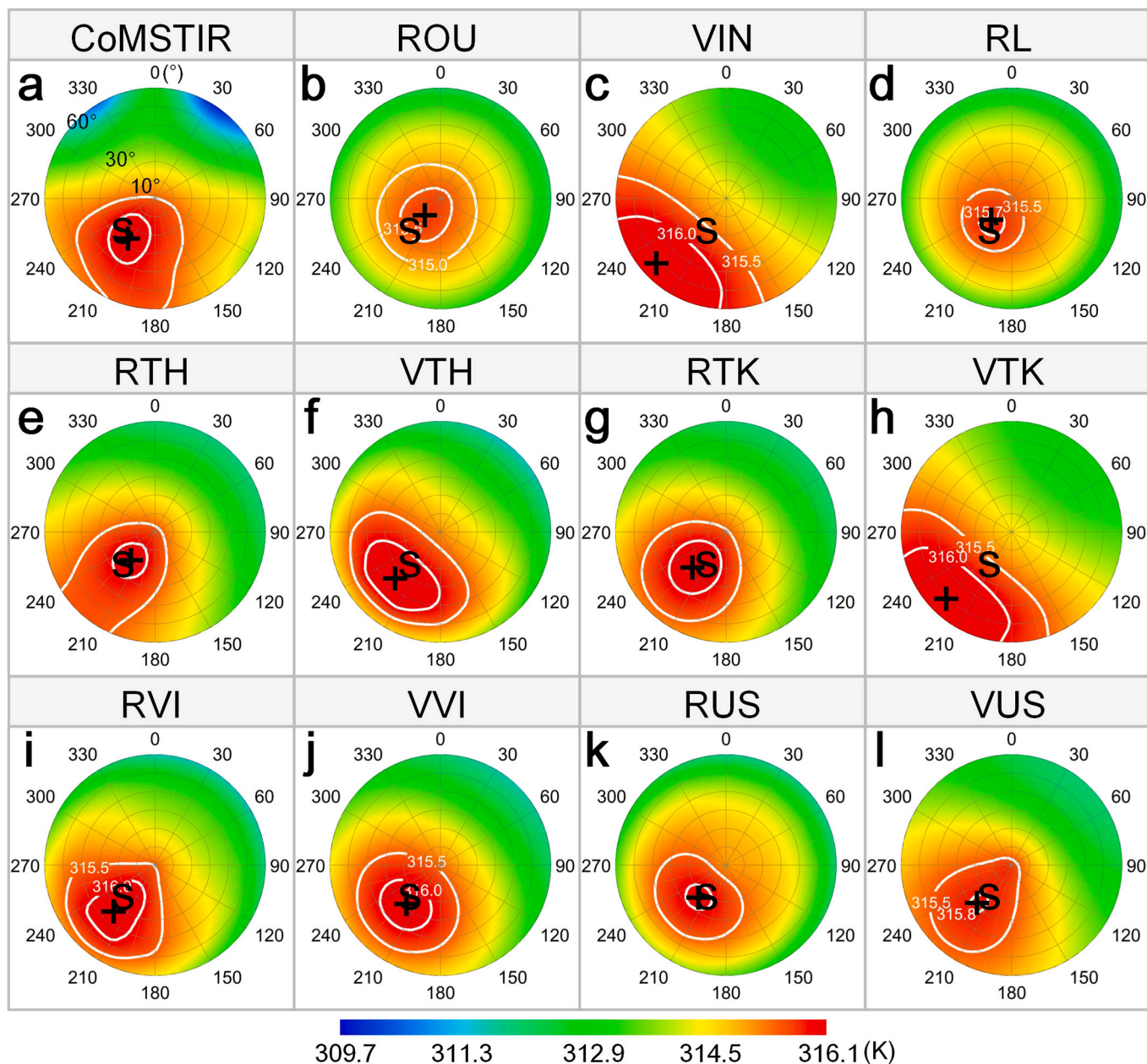

Fig. 7. Polar-DRTs for USM01 simulated by the CoMSTIR (a) and 11 parametric models at 13:00 in summer (from (b) to (l)). The concentric circles correspond to the sun/sensor off-nadir zenith angles (θ_v/θ_s, 0–60°) whereas the radial lines indicate the sun/sensor azimuth angles (φ_v/φ_s, 0–360°). '+' and 'S' indicate the hotspot and sun positions, respectively.
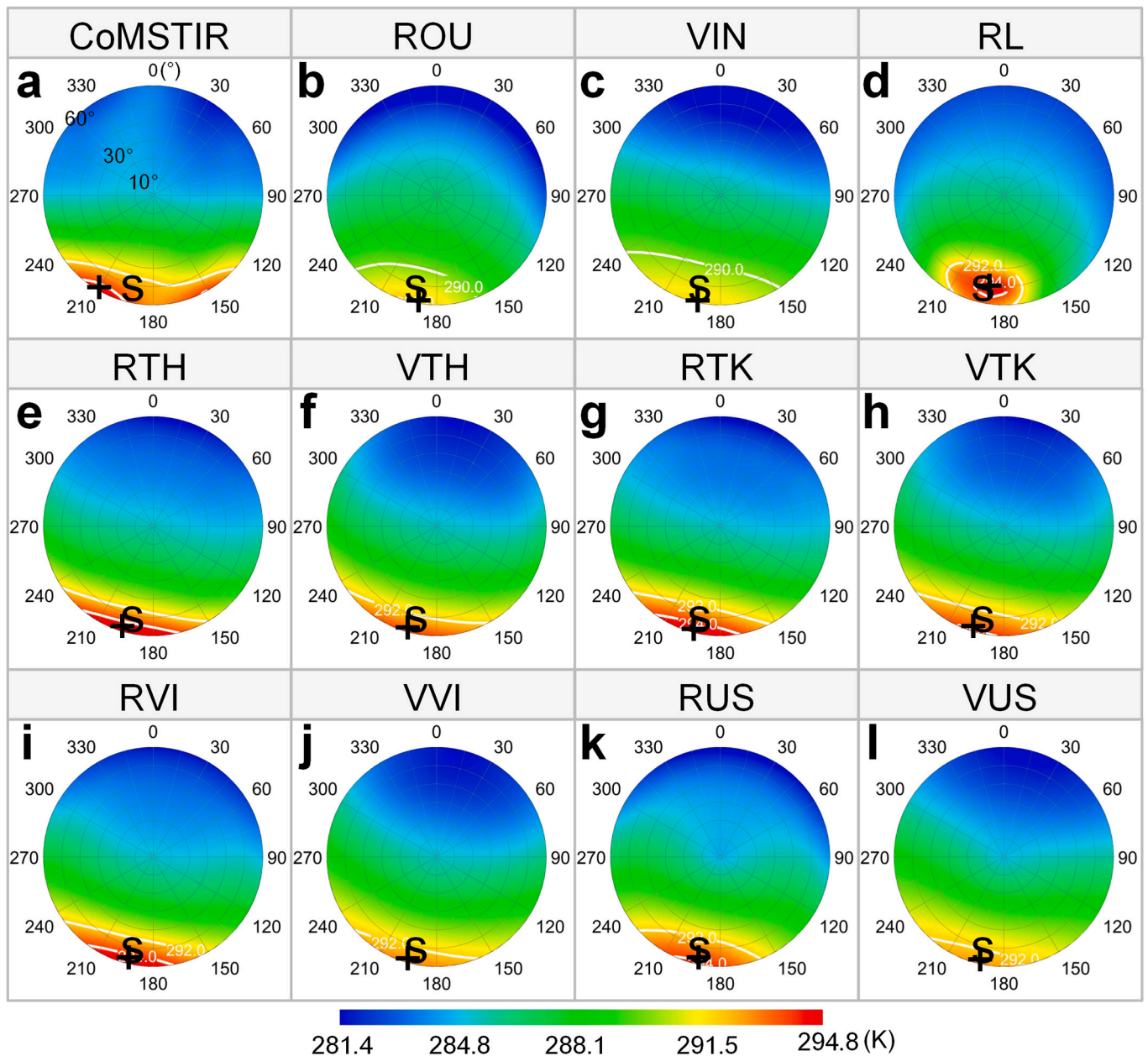
Fig. 8. Polar-DRTs for USM01 simulated by the CoMSTIR (a) and 11 parametric models at 13:00 in winter (from (b) to (l)).

Table 3
DHSs ($\times 10^{-2}$) of the 11 parametric models at different observation times in summer.

Model name	11:00	13:00	15:00	Average
ROU	52.91	20.61	16.68	30.07
VIN	19.65	50.48	4.96	25.03
RL	52.89	16.25	21.33	30.16
RTH	20.36	13.65	9.76	14.59
VTH	22.48	22.51	4.57	16.52
RTK	22.84	12.30	4.07	13.07
VTK	21.00	51.75	4.96	25.90
RVI	16.66	12.67	4.24	11.19
VVI	22.38	13.71	4.76	13.62
RUS	34.73	11.85	6.99	17.86
VUS	16.26	31.29	4.46	17.33

Note that the DHSs in winter are not listed because their values are very small: i. e., the simulated hotspots are sufficiently accurate during winter when the SZA is large (i.e., the hotspots are all concentrated at ~60° of VZA, see Fig. 8a).

with a larger urban fraction, at least for the Chicago and New York cases (Hu et al., 2016a). For this dataset, we find no significant relationship between simulation accuracy and observation time. The model accuracies are very slightly higher at 11:00 than at 13:00 (Figs. 9 and 10), likely because the UTRD intensity at 13:00 is also relatively greater than that of 11:00 (Hu et al., 2016a). Nevertheless, the ranking of model performance remains mostly unchanged with different urban fractions and observation times. As indicated previously, the overall performance of the dual-kernel models is, by and large, better than that of the single-kernel models – the difference in RMSE between these two types of models is 0.16 K. However, some of the single-kernel models are more capable than some dual-kernel models; for example, the ROU model (a single-kernel model; RMSE = 0.82 K) is even slightly better than the dual-kernel VTK model (a dual-kernel model; RMSE = 1.04 K).

For the single-kernel models, the overall performance ranking is as follows: ROU model (RMSE = 0.82 K) = RL model (RMSE = 0.82 K) >
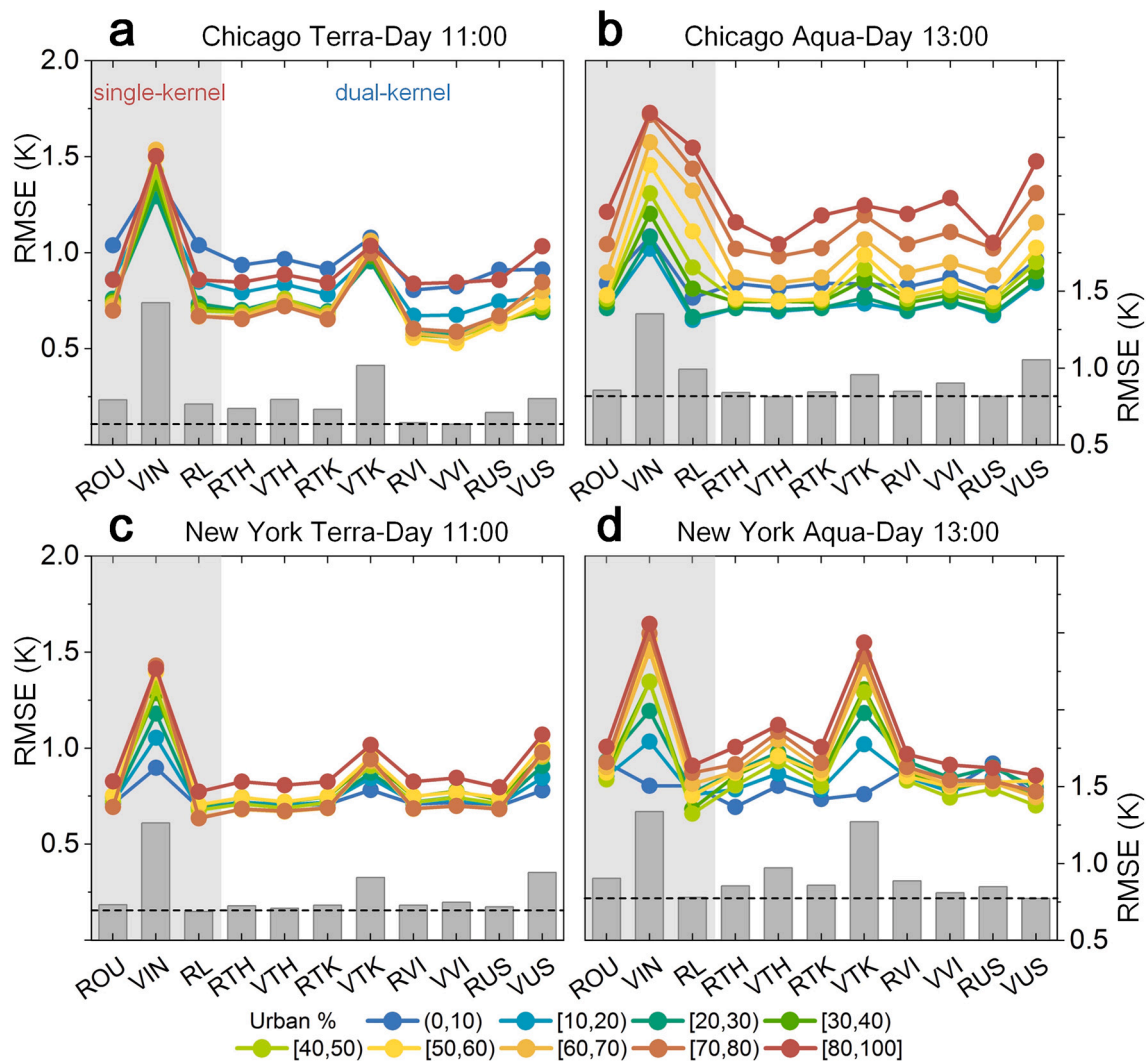
Fig. 9. RMSEs of the 11 kernel-driven models with daytime MODIS DRTs for Chicago and New York City. The grey shading encloses the three single-kernel models. The colored solid lines (left-y-axis) represent the RMSEs grouped into nine urban fraction intervals. The histograms (right-y-axis) indicate the mean RMSEs averaged from the nine urban fraction intervals. The black horizontal dashed line corresponds to the lowest mean RMSE amongst the eleven models.

VIN model (RMSE = 1.35 K). It is interesting that, contrary to the overall assessments with forward-modelling hemispherical DRTs (see Section 4.1.1), the ROU and RL models perform significantly better than the VIN model when evaluated with MODIS DRTs (Figs. 9 and 10). One possible explanation may lie in the sun-surface-sensor (i.e., MODIS) geometry, for which the SZAs of the daytime MODIS overpasses (i.e., around 11:00 and 13:00) are mostly lower than 30° over these two cities (Hu et al., 2016a). In comparison, it has been shown that the VIN model is better than the other two single-kernel models when the SZA is large (see Section 4.1.2). For the dual-kernel models, the results again illustrate that the average RMSE of the models with $K_{Hotspot\_rou}$ is smaller than those with $K_{Hotspot\_vin}$ by 0.1 K, but the model performance ranking amongst the dual-kernel models varies slightly under different scenarios (Figs. 9 and 10). For example, the RVI (RMSE = 0.64 K) and VVI (RMSE = 0.63 K) models perform better than the other dual-kernel models at 11:00 in Chicago (Fig. 9a), while the VTH (RMSE = 0.81 K) and RUS (RMSE = 0.82 K) models perform better at 13:00 in Chicago (Fig. 9b). The VTK model performs worse than the other dual-kernel models in almost all cases (Figs. 9 and 10). By excluding the VTK model, the dual-kernel models with $K_{Hotspot\_rou}$ (the RTH, RTK, RVI, and RUS models) and those with $K_{Hotspot\_vin}$ (the VTH, VVI, and VUS models) have similar performance, with mean RMSEs of 0.78 and 0.83 K, respectively.

Amongst them, the overall performances of the RVI, VVI and RUS models are slightly better (the RMSEs are both 0.77 K).

We however should note that the model performance comparison with the 10-year averaged MODIS multi-angle LST dataset is still very preliminary because this dataset possesses uncertainties. More in-depth investigations based on better satellite multi-angle DRTs are therefore necessary for achieving a finer comparison between these parametric models.

4.3. Sensitivity analysis of model comparison with limited DRTs

In contrast to the model comparisons with hemispherical DRTs (Scheme #1), this section assesses the stability and sensitivity of the parametric models with limited DRTs used as model inputs (Scheme #3). The error increments of models between Schemes #1 and #3 can reflect their stability (Table 4). The results show that the single-kernel VIN and ROU models are generally more stable than the dual-kernel models when only limited DRTs are available. The RMSE increments of the VIN and ROU models (i.e., 0.20 and 0.28 K, respectively) are considerably smaller than those of the dual-kernel models (the lowest RMSE increment for the dual-kernel models is 0.53 K, obtained by the RTH model).
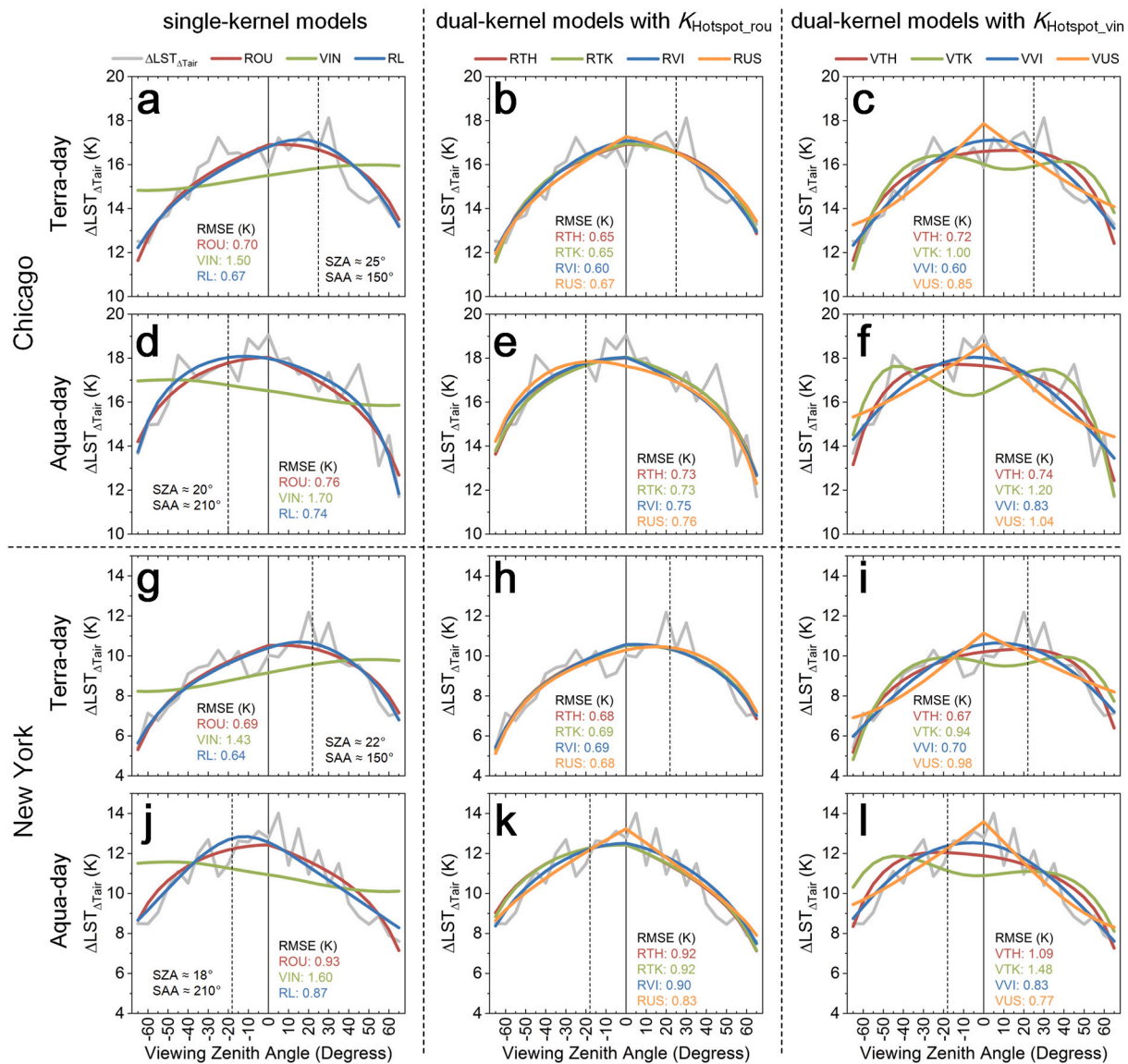
Fig. 10. MODIS multi-angle DRTs as well as the associated simulations by the 11 kernel-driven models at 11:00 (Terra-day) and 13:00 (Aqua-day) for Chicago and New York City. Here only the pixels with the urban fraction of 70% ∼ 80% are demonstrated. The negative and positive VZAs represent the viewing azimuth angles that are larger (negative) and smaller (positive) than 180°. The solid straight and dashed lines represent the VZA being equal to 0° (i.e., nadir) and corresponding sun positions, respectively. (a), (d), (g), and (j) are the simulations by the single-kernel models; (b), (e), (h), and (k) are the simulations by the dual-kernel models with $K_{Hotspot\_rou}$; and (c), (f), (i), and (l) are the simulations by the dual-kernel models with $K_{Hotspot\_vin}$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
RMSEs (K) of the 11 parametric models under Schemes #1 and #3.

Model type	Model name	Scheme #3 (K)	Scheme #1 (K)	Scheme #3 - Scheme #1 (K)
Single-kernel	ROU	1.16	0.88	0.28
	VIN	0.99	0.79	0.20
	RL	1.89	0.89	1.00
Dual-kernel	RTH	1.03	0.50	0.53
	VTH	1.22	0.57	0.65
	RTK	1.06	0.50	0.56
	VTK	1.33	0.60	0.73
	RVI	1.05	0.49	0.56
	VVI	1.16	0.57	0.59
	RUS	1.20	0.60	0.60
	VUS	1.18	0.59	0.59

For the three single-kernel models, the VIN model is the most stable,

followed by the ROU model, while the RL model is the most sensitive to the number of DRTs, with an RMSE increment of 1.0 K between Schemes #1 and #3. For the eight dual-kernel models, those with the hotspot kernel $K_{Hotspot\_rou}$ (the RTH, RTK, RVI, and RUS models) show better stability than those with the hotspot kernel $K_{Hotspot\_vin}$ (the VTH, VTK, VVI, and VUS models). Amongst the four dual-kernel models with $K_{Hotspot\_vin}$, the stabilities of the RTH, RTK, and RVI models are very similar, with RMSE increments of 0.53, 0.56, and 0.56 K, respectively. Considering both model accuracy and stability (Table 4), the VIN model is recommended for UTRD simulation when the available DRTs are very limited and relatively uniformly distributed within the upper hemisphere, followed by the dual-kernel models with $K_{Hotspot\_rou}$ (e.g., the RTH, RTK, and RVI models).

Table 5
RMSEs (K) of four kernel-driven models (i.e., the RVI, RVIc, KRL, GUTA-sparse models).

Urban surface model	RVI (K)	RVIc (K)	KRL (K)	GUTA-sparse (K)
USM01	0.84	0.81	0.90	0.85
USM02	0.68	0.67	1.02	0.68
USM03	0.64	0.62	0.69	0.63
USM04	0.57	0.56	0.63	0.58
USM05	0.46	0.45	0.52	0.45
USM06	0.41	0.40	0.53	0.39
USM07	0.41	0.40	0.66	0.43
USM08	0.34	0.33	0.45	0.33
USM09	0.09	0.09	0.11	0.09
Mean $\pm$ SD	0.49 $\pm$ 0.21	0.48 $\pm$ 0.20	0.61 $\pm$ 0.25	0.49 $\pm$ 0.21

4.4. Discussion

4.4.1. Comparison with several multi-kernel models

The performances of several multi-kernel models were also compared and evaluated under Scheme #1, benchmarked by the RVI model (dual-kernel), which has the best overall performance amongst the aforementioned 11 single- and dual-kernel models. These multi-kernel models include four or more parameters: the hotspot-modified RVI model (five parameters, termed the RVIc model), the Kernel-Hotspot model (four parameters, termed the KRL model), and the GUTA-sparse model (four parameters). Detailed descriptions of three models are given in Appendix B. Note that the GUTA-osg model was not

included because it requires urban morphology information as inputs, which is difficult to obtain at coarse resolutions.

In general, the multi-kernel models perform better than the associated single- and dual-kernel models under some specific scenarios, but the performance improvement is relatively small (refer to Table 5 and Fig. 11). For different urban surface types, the mean RMSE with standard deviation (SD) of the dual-kernel RVI model is 0.49 $\pm$ 0.21 K; and those of the RVIc, KRL, and GUTA-sparse models are 0.48 $\pm$ 0.20, 0.61 $\pm$ 0.26, and 0.49 $\pm$ 0.21 K, respectively (Table 5). Specifically, the RVIc model performs slightly better than the other three multi-kernel models because its hotspot-modified kernel simulates the hotspot more accurately. The KRL model derived from the combination of the RL and VVI models has the lowest accuracy, although its performance is significantly improved compared with the original RL model (the RMSE decreased from 0.89 K for the RL model to 0.61 K for the KRL model). The GUTA-sparse model was designed for sparse urban surfaces and consequently it performs well over sparse urban surfaces (e.g., USM05, USM06, USM08, and USM09) but slightly worse over the other urban surfaces with a high building density. For example, the hotspot simulated by the GUTA-sparse model is relatively less accurate over USM01 (a neighborhood with densely-distributed buildings) for the summer case (Fig. 11).

4.4.2. Limitations and prospects

We have summarized the available parametric models and kernels for UTRD simulation and have then comprehensively compared the performances of three single- and eight dual-kernel models and discussed those of three multi-kernel models. We have also provided recommendations for choosing appropriate models under different scenarios. Arguably, the accuracy of the forward-modelling component
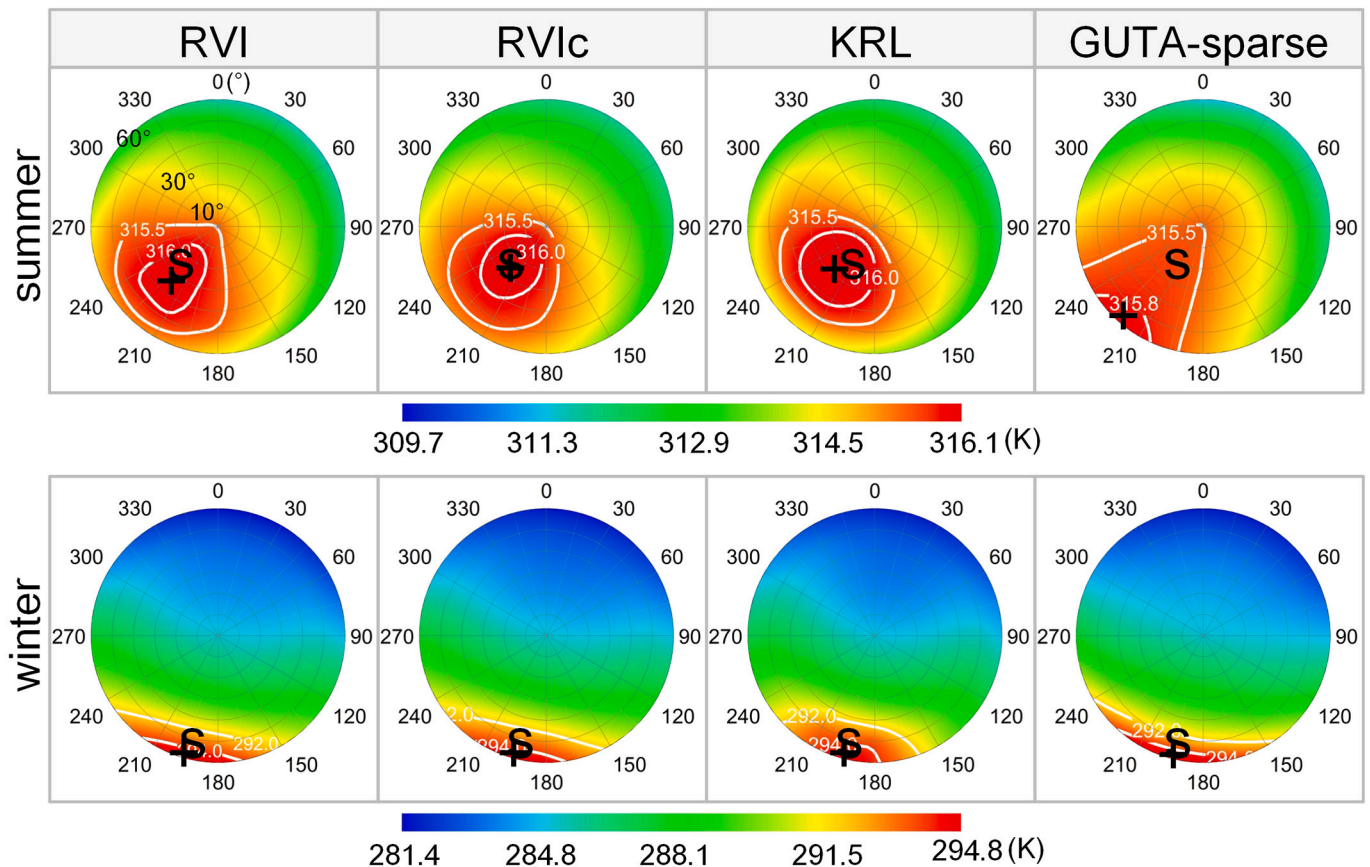

Fig. 11. Polar DRTs simulated by the RVI, RVIc, KRL, and GUTA-sparse models over USM01 at 13:00 during summer and winter. The original polar DRTs (simulated by the CoMSTIR) for validation are provided in Figs. 7 and 8.

temperatures simulated by Envi-met (with the MAE of ~1.0–2.0 K) and accordingly the forward-modelling DRTs simulated by CoMSTIR (with the RMSE of ~1.0 K) is probably inadequate for differentiating the performances amongst the parametric models. Nevertheless, here we are not focused on the absolute accuracy of the forward-modelling DRTs but more on the inversion abilities of parametric models in approximating the complex forward 3D mechanistic models. In addition, we have identified the following aspects that require further consideration:

First, the construction of a large database of all-angle DRTs for a great variety of urban surface types and environmental conditions (including illumination and climatic conditions). Both forward-modelling and satellite datasets (MODIS LSTs) have been used for assessment of model performances in this study. Though various urban surface types and environmental conditions were considered for the forward-modelling datasets (designed based on different local climate zones, see Fig. A1), as well as for the MODIS LST dataset (data of two megacities with different fractions of impervious surfaces were incorporated, see Fig. 9), these two types of datasets may have several shortcomings. The forward-modelling dataset is still based on conceptual rather than real urban surfaces (Jiang et al., 2018). Differences between the conceptual and real urban surface types do exist and there are additional urban surface types that are not considered. The MODIS LST dataset does not comprise the original instantaneous observations but rather reprocessed ones averaged over a long period; in addition, its spatial resolution is relatively low (Hu et al., 2016a). Airborne measurements can be used to obtain DRTs from real urban surfaces at fine resolutions, but it is very costly to obtain adequate amounts of such data for different urban surface types from different cities and under arbitrary environmental conditions. In recent years there has been the rapid development of unmanned aerial vehicle (UAV) technology, and UAV thermal observations can obtain quasi-real-time DRTs at very high spatial resolution. UAV technology can complement forward-modelling and satellite and airborne DRTs (Feng et al., 2020), allowing for the establishment of a large DRT database over sufficiently diverse conditions.

Second, the design of parametric kernel-driven models that better suit UTRD simulation. The models or kernels compared in this study were mostly designed originally for vegetation surfaces, without strictly considering urban morphology. In other words, the newly-derived parametric models based on the kernel combination strategy are primarily statistical (or empirical) rather than grounded strictly in urban morphology. Although this empirical kernel combination is well accepted for kernel-driven modelling (Wanner et al., 1995), it is still necessary to design kernel-driven models that are formulated specifically for urban surfaces. We suggest the following strategies. The first strategy lies in the incorporation of the urban thermal inertia effect, which is mostly caused by the heat capacity differences amongst urban elements and which scales with impervious surface fraction, building structure, and building material. For surfaces where there are more elements with higher heat capacity, there is likely to be more lag and correspondingly a larger azimuthal gap between the sun and hotspot (Lagouarde et al., 2010; Liu et al., 2019). Parametric models that consider this azimuthal gap can be designed to improve model performance. The other strategy is to consider urban structure parameters (e. g., the sky view factor) that affect the illumination conditions of urban surface components and consequently the component temperatures. Great progress has been made recently in the parametric modelling of UTRD for urban surfaces. By considering urban morphology, the GUTA series models are able to simulate hotspot shapes distinctively occurring over urban areas, e.g., for either sparsely (i.e., the GUTA-sparse model) or densely (i.e., the GUTA-osg and GUTA-dense models) distributed buildings (Wang et al., 2018a, 2018b; Wang and Chen, 2019).

Specifically, a priori knowledge of building morphology information is required for the GUTA-osg and GUTA-dense models, while such information is unnecessary for the linear GUTA-sparse model. We believe that the GUTA series models are anticipated to be better options in the near future when increasingly more urban morphology information becomes available. We suggest that much more effort is needed to develop models or kernels suitable for a greater variety of urban surface types and solar illumination conditions.

Third, the UTRD can be greatly impacted by canopy morphology and vegetation abundance (Dyce and Voogt, 2018). The variations of morphological and thermal properties within a heterogeneous urban landscape can influence the UTRD – the UTRD intensity over urban surfaces with vegetation can be either increased or decreased, depending on the vegetation abundance as well as the vegetation and building morphology (Dyce and Voogt, 2018). The USMs used in this study have considered the importance of the vegetation cover – four amongst the nine USMs have incorporated vegetation. Nevertheless, the realistic representation of vegetation remains inadequate in current USMs. The spatial pattern of vegetation and buildings is also relatively simple, e.g., surfaces with vegetation fraction higher than that of building were not considered. In other words, these USMs were still dominated by buildings, based on which models have been compared. For the scenarios where vegetation dominates, the model performances are anticipated to be slightly different; and for these urban surfaces, the optimal parametric models may be close to those that were entirely based on pure vegetation (i.e., without impervious surface) (Liu et al., 2018; Cao et al., 2019a, 2019b, 2021). To this end, we provide the following suggestions for model selection: Results of the current study are recommended when the urban landscape is dominated by impervious surfaces such as buildings; results from Cao et al. (2019a, 2019b, 2021) are recommended when the urban landscape is dominated by vegetation; for urban landscape with approximately equal fractions between impervious surface and vegetation, practitioners should be more careful and should refer to both types of studies.

5. Conclusions

Parametric kernel-driven models used for UTRD simulation are crucial for the angular normalization of remotely-sensed LSTs obtained at slant angles. However, it remains largely unknown which of the parametric kernel-driven models has the best performance for simulating UTRD under various types of urban surface. It is also unclear whether models with even better performance can be developed by combining the available hotspot and base shape kernels. Based on both forward-modelling and satellite datasets, we have summarized and then compared 14 usable parametric models (including 3 single-kernel, 8 dual-kernel, and 3 multi-kernel models) over a large variety of urban surfaces and solar radiation conditions.

Our major findings are as follows: (1) The dual-kernel models that combine one hotspot kernel and one base shape kernel perform better than the single-kernel models with only a hotspot kernel. The mean RMSE of the dual-kernel models is 0.30 K lower than that of the single-kernel models. (2) For the three single-kernel models, the overall performance of the VIN model is superior to that of the ROU and RL models. However, the hotspot simulated by the VIN model is less accurate than the other two models, especially when the SZA is small (SZA $\leq 30°$). For this reason, the VIN model becomes less competent than the ROU and RL models with the MODIS dataset as model inputs, mainly because the SZA is often small for daytime MODIS overpasses (i.e., close to solar noon). (3) For the eight dual-kernel models, those with $K_{\mathrm{Hotspot\_rou}}$ as the hotspot kernel are mostly more capable than those with $K_{\mathrm{Hotspot\_vin}}$, probably because the hotspot kernel $K_{\mathrm{Hotspot\_rou}}$ are parametrized based on

rectangular objects that are more comparable to urban buildings while $K_{Hotspot\_vin}$ is less capable for the simulation of hotspot location. Amongst the three dual-kernel models with $K_{Hotspot\_rou}$ (i.e., the RVI, RTH, and RTK models), the RVI model has the best performance (the RMSE is ~0.49 and ~0.77 K based on forward-modelling and satellite data, respectively), although the differences between these three models are small. (4) The multi-kernel models have higher accuracies compared with the dual-kernel models, but their performance improvements are relatively limited.

Based on these findings, if the available DRTs are very limited (e.g., three to five observed DRTs), we recommend the use of the single-kernel models: the VIN model is preferred when the SZA is larger than 30°, otherwise the ROU model is recommended. When there are adequate DRTs, we recommend practitioners use the dual-kernel models, and the RVI model is generally the optimal option. Our findings are potentially helpful for the angular normalization of satellite-derived urban LST products, as well as for the remote estimation of complete urban LSTs.

Overall, they contribute significantly to the enhancement of thermal remote sensing of urban climates and environment.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A. Urban surface models and distribution of viewing angles of the used MODIS data
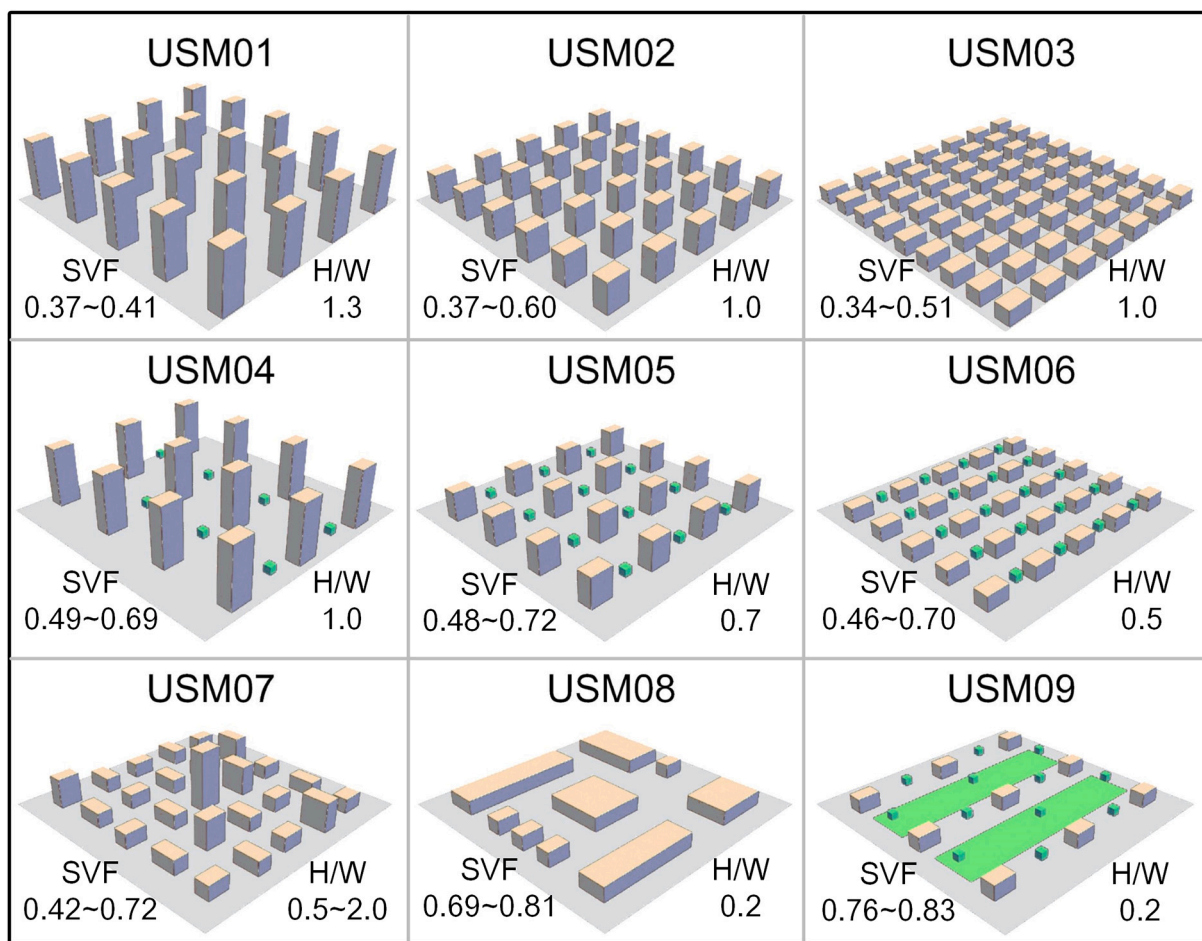

Fig. A1. Nine USMs with different the sky view factors (SVFs) and aspect ratios (*H/W*, where *H* and *W* are the wall height and street width, respectively), adapted from Jiang et al. (2018).
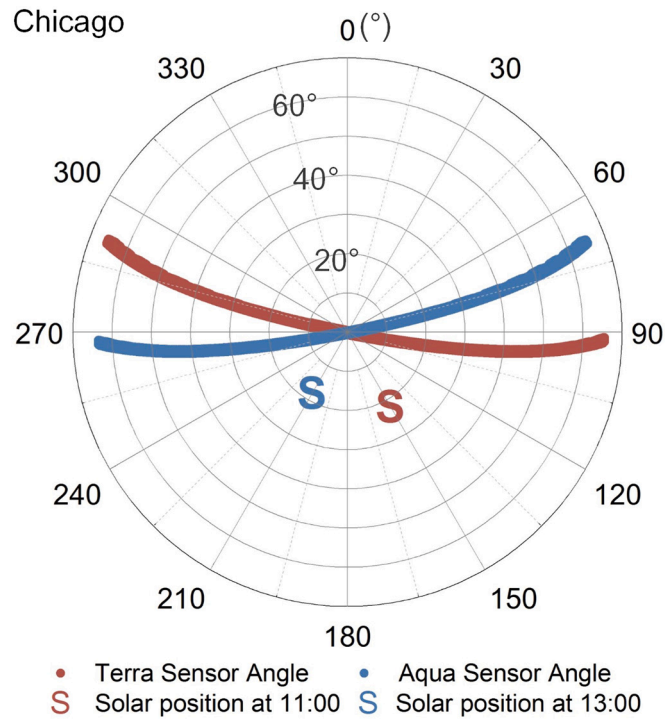
Fig. A2. Polar diagram of the distribution of the viewing angles of MODIS in Chicago during daytime, adapted from Hu et al. (2016a). The concentric circles (0–65°) and radial lines (0–360°) of the polar diagram represent the sensor/sun zenith and azimuth angles, respectively.

Appendix B. Formulae of several multi-kernel models

This appendix presents the formulae for several parametric models mentioned in Section 4.4.1, which can be used for UTRD simulation. They include the hotspot-modified RVI model (termed the RVIc model), the Kernel-Hotspot model (termed the KRL model), and the GUTA-sparse model.

Chen and Cihlar (1997) proposed a modification function (F_m) to improve the performance of kernel-driven BRDF models by improving hotspot simulation. The function F_m is also applicable to modifying the hotspot description in UTRD simulation. The hotspot sharpening function F_m was therefore incorporated into the RVI model, which has the overall best performance amongst 11 parametric models (see Section 4), resulting in a new hotspot-modified RVI model (termed the RVIc model). The RVIc model can be written as follows:

$$\begin{cases} T(\theta_s, \theta_v, \varphi) = \left(f_{iso} + f_{Hotspot} \cdot K_{Hotspot\_rou} + f_{BaseShape} \cdot K_{BaseShape\_vin}\right) \cdot F_m \\ F_m = 1 + C_1 e^{-C_2 \frac{\xi}{\pi}} \end{cases} \tag{B1}$$

where F_m is the hotspot sharpening function, with the detailed formula given in Chen and Cihlar (1997); C_1 and C_2 are two coefficients; and ξ is the phase angle related to the sun-surface-sensor position.

The Kernel-Hotspot model (termed the KRL model) integrates the Vinnikov and RL models and is more capable of simulating the shadowing effect during the day and the base shape kernel anisotropy at night (Ermida et al., 2018b). The validity of the KRL model has been tested over vegetation surfaces (Ermida et al., 2018b). This study further tested the performance of the KRL model for UTRD simulation. To facilitate model comparison, we converted the formula of the original KRL model according to the general expression Eq. (8). Specifically, (1) the 'T_0' on the left side of the equation was moved to the right side as f_{iso}; (2) the coefficient of hotspot kernel '$B \cdot Rad^*_{TOA}$' was integrated into a single coefficient $f_{Hotspot}$. From the modelling perspective, the KRL model is exactly equivalent before and after the conversion. The converted KRL model can be expressed as follows:

$$T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{Hotspot} \cdot sin(2\theta_s) \cdot K_{Hotspot\_rl} + f_{BaseShape} \cdot K_{BaseShape\_vin} \tag{B2}$$

where $sin(2\theta_s) \cdot K_{Hotspot\_rl}$ is the hotspot kernel, and D is the angular distance between the sun and sensor (refer to Eq. (2)).

The GUTA-sparse model was designed for UTRD simulation over a sparse urban canopy without mutual shadowing (Wang et al., 2018b). It can be expressed by the following formula:

$$\begin{cases} T(\theta_s, \theta_v, \varphi) = f_{iso} + f_{bgd} K_{bgd}(\theta_s, \theta_v, \varphi) + f_{ori} K_{ori}(\theta_s, \theta_v, \varphi) + f_{shw} K_{shw}(\theta_s, \theta_v, \varphi) \\ K_{bgd}(\theta_s, \theta_v, \varphi) = \frac{2}{\pi} tan\theta_v \\ K_{ori}(\theta_s, \theta_v, \varphi) = \frac{1}{2\pi}((\pi - \varphi)cos\varphi + sin\varphi)tan\theta_v \\ K_{shw}(\theta_s, \theta_v, \varphi) = \frac{1}{\pi} tan\theta_s \left(\frac{\sqrt{tan^2\theta_s + tan^2\theta_v - 2tan\theta_s tan\theta_v cos\varphi}}{tan\theta_s + tan\theta_v} - 1\right)(cos\varphi + 1) \end{cases} \tag{B3}$$

where f_{iso}, f_{bgd}, f_{ori}, and f_{shw} are the coefficients; and K_{bgd}, K_{ori}, and K_{shw} are three kernels weighting the contributions from the surface background, orientation effect, and shadow, respectively.

Appendix C. Analysis of the impacts of scenarios with different street orientations and city latitudes on model performances

The performances of the 11 parametric models for the six USMs with different street orientations (see Section 2.1) are listed in Table C1. The accuracy for USM02-0° is slightly lower than those USMs with a rotation angle (with ΔRMSE of 0.03–0.13 K on average). Specially, when the rotation angle is less than 45°, the RMSE remains mostly unchanged; when the rotation angle exceeds 45°, the RMSE can be decreased by about 0.1 K. The slight increase of the model accuracy when the rotation angle is >45° may be related to the smaller UTRD intensity under this scenario, which results from the significantly reduced fractions of shadow components for USMs with building lengths greater than widths (e.g., USM02). However, the relative performance ranks amongst models remains relatively highly consistent amongst these scenarios (Table C1).

Table C1
RMSEs (K) of the 11 parametric kernel-driven models for USM02 with different street orientations.

Model name	0°	15°	30°	45°	60°	75°
ROU	1.21	1.24	1.23	1.11	1.17	1.12
VIN	1.08	1.01	1.04	0.90	0.90	0.91
RL	1.21	1.24	1.21	1.09	1.17	1.14
RTH	0.69	0.66	0.66	0.59	0.60	0.59
VTH	0.78	0.72	0.72	0.63	0.65	0.64
RTK	0.69	0.66	0.66	0.58	0.60	0.59
VTK	0.82	0.75	0.78	0.68	0.67	0.67
RVI	0.69	0.64	0.64	0.58	0.58	0.57
VVI	0.78	0.71	0.71	0.62	0.65	0.63
RUS	0.82	0.81	0.80	0.72	0.75	0.71
VUS	0.80	0.73	0.74	0.64	0.67	0.65
Mean	0.87	0.83	0.84	0.74	0.76	0.75

The performances of the 11 parametric models for the five latitudes (including Nanjing) (see Section 2.1) for USM01 are listed in Table C2. For the same model, the difference in RMSE amongst cities with different latitudes ranges from 0.22 to 0.61 K, but this difference is less related to city latitude. As before, the relative performance ranking of the models remains relatively highly consistent for scenarios with different latitudes.

Table C2
RMSEs (K) of the 11 parametric kernel-driven models with different city latitudes: 45.93°N (Harbin), 40.08°N (Beijing), 31.56°N (Nanjing), 26.08°N (Fuzhou), and 20°N (Haikou).

Model name	45.93°N	40.08°N	31.56°N	26.08°N	20°N	Mean ± SD
ROU	1.39	1.52	1.34	1.41	1.07	1.35 ± 0.15
VIN	1.36	1.33	1.19	1.26	1.08	1.24 ± 0.10
RL	1.72	1.61	1.36	1.41	1.11	1.44 ± 0.21
RTH	1.01	1.06	0.85	0.99	0.69	0.92 ± 0.13
VTH	1.12	1.10	0.91	1.06	0.81	1.00 ± 0.12
RTK	1.03	1.07	0.84	0.98	0.69	0.92 ± 0.14
VTK	1.12	1.11	0.96	1.11	0.90	1.04 ± 0.09
RVI	0.97	1.03	0.84	0.97	0.69	0.90 ± 0.12
VVI	0.98	1.08	0.90	1.04	0.75	0.95 ± 0.12
RUS	1.07	1.18	0.97	1.10	0.79	1.02 ± 0.13
VUS	0.99	1.09	0.92	1.06	0.77	0.97 ± 0.11

References

Bullock, R., 2007. Great Circle Distances and Bearings Between Two Locations. Available from: https://dtcenter.org/met/users/docs/write_ups/gc_simple.pdf.

Cao, B., Gastellu-Etchegorry, J.-P., Du, Y., Li, H., Bian, Z., Hu, T., Fan, W., Xiao, Q., Liu, Q., 2019a. Evaluation of four kernel-driven models in the thermal infrared band. IEEE Trans. Geosci. Remote Sens. 57, 5456–5475.

Cao, B., Liu, Q., Du, Y., Roujean, J.L., Xiao, Q., 2019b. A review of earth surface thermal radiation directionality observing and modeling: historical development, current status and perspectives. Remote Sens. Environ. 232, 111304.

Cao, B., Roujean, J.-L., Gastellu-Etchegorry, J.-P., Liu, Q., Du, Y., Lagouarde, J.P., Huang, H., Li, H., Bian, Z., Hu, T., 2021. A general framework of kernel-driven modeling in the thermal infrared domain. Remote Sens. Environ. 252, 112157.

Chen, J., Cihlar, J., 1997. A hotspot function in a simple bidirectional reflectance model for satellite applications. J. Geophys. Res. Atmos. 102, 25907–25913.

Chow, W.T.L., Pope, R.L., Martin, C.A., Brazel, A.J., 2011. Observing and modeling the nocturnal park cool island of an arid city: horizontal and vertical impacts. Theor. Appl. Climatol. 103, 197–211.

Du, Y., Liu, Q., Chen, L., Liu, Q., Yu, T., 2007. Modeling directional brightness temperature of the winter wheat canopy at the Ear stage. IEEE Trans. Geosci. Remote Sens. 45, 3721–3739.

Duffour, C., Olioso, A., Demarty, J., Van der Tol, C., Lagouarde, J.P., 2015. An evaluation of SCOPE: a tool to simulate the directional anisotropy of satellite-measured surface temperatures. Remote Sens. Environ. 158, 362–375.

Duffour, C., Lagouarde, J.P., Roujean, J.L., 2016. A two parameter model to simulate thermal infrared directional effects for remote sensing applications. Remote Sens. Environ. 186, 250–261.

Dyce, D.R., Voogt, J.A., 2018. The influence of tree crowns on urban thermal effective anisotropy. Urban Clim. 23, 91–113.

Ermida, S.L., Trigo, I.F., DaCamara, C.C., Pires, A., 2018a. A methodology to simulate LST directional effects based on parametric models and landscape properties. Remote Sens. 10, 1114.

Ermida, S.L., Trigo, I.F., Dacamara, C.C., Roujean, J.L., 2018b. Assessing the potential of parametric models to correct directional effects on local to global remotely sensed LST. Remote Sens. Environ. 209, 410–422.

Feng, L., Tian, H., Qiao, Z., Zhao, M., Liu, Y., 2020. Detailed variations in urban surface temperatures exploration based on unmanned aerial vehicle thermography. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 204–216.

Gastellu-Etchegorry, J.-P., Martin, E., Gascon, F., 2004. DART: a 3D model for simulating satellite images and studying surface radiation budget. Int. J. Remote Sens. 25, 73–96.

Guillevic, P.C., Bork-Unkelbach, A., Gottsche, F.M., Hulley, G., Gastellu-Etchegorry, J.-P., Olesen, F.S., Privette, J.L., 2013. Directional viewing effects on satellite land surface temperature products over sparse vegetation canopies—a multisensor analysis. IEEE Geosci. Remote Sens. Lett. 10, 1464–1468.

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. Rev. Geophys. 48, 7362–7388.

Hu, L., Wendel, J., 2019. Analysis of urban surface morphologic effects on diurnal thermal directional anisotropy. ISPRS J. Photogramm. Remote Sens. 148, 1–12.

Hu, L., Monaghan, A., Voogt, J.A., Barlage, M., 2016a. A first satellite-based observational assessment of urban thermal anisotropy. Remote Sens. Environ. 181, 111–121.

Hu, T., Du, Y., Cao, B., Li, H., Bian, Z., Sun, D., Liu, Q., 2016b. Estimation of upward longwave radiation from vegetated surfaces considering thermal directionality. IEEE Trans. Geosci. Remote Sens. 54, 6644–6658.

Hu, T., Cao, B., Du, Y., Li, H., Wang, C., Bian, Z., Sun, D., Liu, Q., 2017. Estimation of surface upward longwave radiation using a direct physical algorithm. IEEE Trans. Geosci. Remote Sens. 55, 4412–4426.

Huang, H., Liu, Q., Liu, Q., Qin, W., 2012. Validating theoretical simulations of thermal emission hot spot effects on maize canopies. Int. J. Remote Sens. 33, 746–761.

Huang, H., Xie, W., Sun, H., 2015. Simulating 3D urban surface temperature distribution using Envi-met model: case study on a forest park. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1642–1645. https://doi.org/10.1109/IGARSS.2015.7326100.

Jiang, L., Zhan, W., Voogt, J.A., Zhao, L., Gao, L., Huang, F., Cai, Z., Ju, W., 2018. Remote estimation of complete urban surface temperature using only directional radiometric temperatures. Build. Environ. 135, 224–236.

Jupp, D., 2000. A Compendium of Kernel & Other (Semi-) Empirical BRDF Models. Office of Space Science Applications-Earth Observation Centre available only as online document (last accessed November 2008). www.cossa.csiro.au/tasks/brdf/k_summ.pdf.

Krayenhoff, E., Voogt, J.A., 2016. Daytime thermal anisotropy of urban neighbourhoods: morphological causation. Remote Sens. 8, 108.

Lagouarde, J.P., Irvine, M., 2008. Directional anisotropy in thermal infrared measurements over Toulouse city centre during the CAPITOUL measurement campaigns: first results. Meteorol. Atmos. Sci. 102, 173–185.

Lagouarde, J.P., Ballans, H., Moreau, P., Guyon, D., Coraboeuf, D., 2000. Experimental study of brightness surface temperature angular variations of maritime pine (Pinus pinaster) stands. Remote Sens. Environ. 72, 17–34.

Lagouarde, J.P., Moreau, P., Irvine, M., Bonnefond, J.M., Voogt, J.A., Solliec, F., 2004. Airborne experimental measurements of the angular variations in surface temperature over urban areas: case study of Marseille (France). Remote Sens. Environ. 93, 443–462.

Lagouarde, J.P., Hénon, A., Kurz, B., Moreau, P., Irvine, M., Voogt, J.A., Mestayer, P., 2010. Modelling daytime thermal infrared directional anisotropy over Toulouse city Centre. Remote Sens. Environ. 114, 87–105.

Lagouarde, J.P., Hénon, A., Irvine, M., Voogt, J.A., Pigeon, G., Moreau, P., Masson, V., Mestayer, P., 2012. Experimental characterization and modelling of the nighttime directional anisotropy of thermal infrared measurements over an urban area: case study of Toulouse (France). Remote Sens. Environ. 117, 19–33.

Li, Z., Tang, B., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F., Sobrino, J.A., 2013. Satellite-derived land surface temperature: current status and perspectives. Remote Sens. Environ. 131, 14–37.

Li, H., Li, R., Yang, Y., Cao, B., Liu, Q., 2020. Temperature-based and radiance-based validation of the collection 6 MYD11 and MYD21 land surface temperature products over barren surfaces in northwestern China. IEEE Trans. Geosci. Remote Sens. 1–14. https://doi.org/10.1109/TGRS.2020.2998945.

Liu, Q., Huang, H., Qin, W., Fu, K., Li, X., 2007. An extended 3-d radiosity–graphics combined model for studying thermal-emission directionality of crop canopy. IEEE Trans. Geosci. Remote Sens. 45, 2900–2918.

Liu, X., Tang, B., Li, Z., 2018. Evaluation of three parametric models for estimating directional thermal radiation from simulation, airborne, and satellite data. Remote Sens. 10, 420.

Liu, X., Tang, B., Li, Z., Rasmussen, M.O., 2019. Influence of temperature inertia on thermal radiation directionality modeling based on geometric optical model. IEEE Trans. Geosci. Remote Sens. 58, 3445–3457.

Liu, X., Tang, B., Li, Z., Zhang, X., Shang, G., 2020. On the derivation of geometric optical kernels for directional thermal radiation. Earth Space. Sci. 7 e2019EA000895. https://doi.org/000810.001029/002019EA000895.

Lucht, W., Schaaf, C.B., Strahler, A.H., 2000. An algorithm for the retrieval of albedo from space using semiempirical BRDF models. IEEE Trans. Geosci. Remote Sens. 38, 977–998.

Ma, W., Chen, Y., Zhan, W., 2013. Thermal anisotropy model for simulated three dimensional urban targets. J. Remote Sens. 17, 62–76.

Mao, B., Li, B., 2020. Graph-based 3D building semantic segmentation for sustainability analysis. J. Geovis. Spat. Anal. 4, 4. https://doi.org/10.1007/s41651-019-0045-y.

Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017. Urban Climates. Cambridge University Press.

Peng, J., Liu, Q., Liu, Q., Li, H., Ma, H., Fang, L., 2011. Kernel-driven model fitting of multi-angle thermal infrared brightness temperature and its application. J. Infrared Millimeter Waves 30, 361–365.

Pinheiro, A.C.T., Privette, J.L., Mahoney, R., Tucker, C.J., 2004. Directional effects in a daily AVHRR land surface temperature dataset over Africa. IEEE Trans. Geosci. Remote Sens. 42, 1941–1954.

Pinheiro, A.C.T., Privette, J.L., Guillevic, P.C., 2006. Modeling the observed angular anisotropy of land surface temperature in a Savanna. IEEE Trans. Geosci. Remote Sens. 44, 1036–1047.

Rasmussen, M.O., Gottsche, F.M., Olesen, F.S., Sandholt, I., 2011. Directional effects on land surface temperature estimation from meteosat second generation for savanna landscapes. IEEE Trans. Geosci. Remote Sens. 49, 4458–4468.

Ren, H., Liu, R., Yan, G., Mu, X., Li, Z.L., Nerry, F., Liu, Q., 2014. Angular normalization of land surface temperature and emissivity using multiangular middle and thermal infrared data. IEEE Trans. Geosci. Remote Sens. 52, 4913–4931.

Roujean, J.L., 2000. A parametric hot spot model for optical remote sensing applications. Remote Sens. Environ. 71, 197–206.

Roujean, J.L., Leroy, M., Deschamps, P.Y., 1992. A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data. J. Geophys. Res. Atmos. 97, 20455–20468.

Sobrino, J.A., Jiménez-Muñoz, J.C., Verhoef, W., 2005. Canopy directional emissivity: comparison between models. Remote Sens. Environ. 99, 304–314.

Stewart, I.D., Oke, T.R., 2012. Local climate zones for urban temperature studies. Bull. Am. Meteorol. Soc. 93, 1879–1900.

Sun, H., Chen, Y., Zhan, W., Wang, M., Ma, W., 2015. A kernel model for urban surface thermal emissivity anisotropy and its uncertainties. J. Infrared Millimeter Waves 34, 66–73.

Vinnikov, K.Y., Yu, Y., Goldberg, M.D., Tarpley, D., Romanov, P., Laszlo, I., Chen, M., 2012. Angular anisotropy of satellite observations of land surface temperature. Geophys. Res. Lett. 39, 23802.

Voogt, J.A., 2008. Assessment of an urban sensor view model for thermal anisotropy. Remote Sens. Environ. 112, 482–495.

Voogt, J.A., Oke, T.R., 1998. Effects of urban surface geometry on remotely-sensed surface temperature. Int. J. Remote Sens. 19, 895–920.

Voogt, J.A., Oke, T.R., 2003. Thermal remote sensing of urban climates. Remote Sens. Environ. 86, 370–384.

Wan, Z., 2006. MODIS Land Surface Temperature Products Users' Guide. Institute for Com-putational Earth System Science. University of California, Santa Barbara, CA. Available from: http://www.icess.ucsb.edu/modis/LstUsrGuide_v5/MODIS_LST_products_Users_guide.pdf.

Wang, D., Chen, Y., 2019. A geometric model to simulate urban thermal anisotropy in simplified dense neighborhoods (GUTA-Dense). IEEE Trans. Geosci. Remote Sens. 57, 6226–6239.

Wang, D., Chen, Y., Cui, Y., Sun, H., 2018a. A geometric model to simulate urban thermal anisotropy for simplified neighborhoods. IEEE Trans. Geosci. Remote Sens. 56, 4930–4944.

Wang, D., Chen, Y., Zhan, W., 2018b. A geometric model to simulate thermal anisotropy over a sparse urban surface (GUTA-sparse). Remote Sens. Environ. 209, 263–274.

Wang, D., Chen, Y., Voogt, J.A., Krayenhoff, E.S., Wang, J., Wang, L., 2020. An advanced geometric model to simulate thermal anisotropy time-series for simplified urban neighborhoods (GUTA-T). Remote Sens. Environ. 237, 111547.

Wanner, W., Li, X., Strahler, A.H., 1995. On the derivation of kernels for kernel-driven models of bidirectional reflectance. J. Geophys. Res. Atmos. 100, 21077–21089.

Yang, H., Li, X., Gao, F., 2002. An algorithm for the retrieval of albedo from space using new go kernel-driven BRDF model. J. Remote Sens. 6, 246–251.

Yang, X., Zhao, L., Bruse, M., Meng, Q., 2013. Evaluation of a microclimate model for predicting the thermal behavior of different ground surfaces. Build. Environ. 60, 93–104.

Yu, T., Tian, Q., Gu, X., Wang, J., Liu, Q., Yan, G., 2006. Modelling directional brightness temperature over a simple typical structure of urban areas. J. Remote Sens. 10, 661–669.

Zhan, W., Chen, Y., Ma, W., Zhou, J., 2010. FOV effect analysis in directional brightness temperature observations for urban targets. J. Remote Sens. 14, 372–386.

Zhan, W., Chen, Y., Voogt, J.A., Zhou, J., Wang, J., Ma, W., Liu, W., 2012. Assessment of thermal anisotropy on remote estimation of urban thermal inertia. Remote Sens. Environ. 123, 12–24.

Zhao, L., Gu, X., Yu, T., Wan, W., Zhang, L., Xie, Y., 2012. A directional thermal radiance model for multiple scattering over surfaces. J. Infrared Millimeter Waves 31, 528–535.