The surface temperatures of Earth: steps towards integrated understanding of variability and change

C. J. Merchant$^{1,2}$, S. Matthiesen$^2$, N. A. Rayner$^3$, J. J. Remedios$^4$, P. D. Jones$^{5,6}$, F. Olesen$^7$, B. Trewin$^8$, P. W. Thorne$^9$, R. Auchmann$^{10}$, G. K. Corlett$^4$, P. C. Guillevic$^{11,12}$, and G. C. Hulley$^{13}$

$^1$Department of Meteorology, University of Reading, Reading, UK
$^2$School of GeoSciences, The University of Edinburgh, Edinburgh, UK
$^3$Hadley Centre, MetOfﬁce, Exeter, UK
$^4$University of Leicester, Physics and Astronomy, Leicester, UK
$^5$Climatic Research Unit, University of East Anglia, Norwich, UK
$^6$Center of Excellence for Climate Change Research/Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia
$^7$Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
$^8$National Climate Centre, Australian Bureau of Meteorology, Melbourne, Australia
$^9$Nansen Environmental and Remote Sensing Center, Bergen, Norway
$^{10}$Institute of Geography and Oeschger Centre for Climate Change Research, University of Bern, Switzerland
$^{11}$Cooperative Institute for Climate and Satellites (CICS), North Carolina State University, Asheville, NC, USA
$^{12}$National Oceanic and Atmospheric Administration (NOAA), National Climatic Data Center, Asheville, NC, USA
$^{13}$Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

Correspondence to: C. J. Merchant (c.j.merchant@reading.ac.uk)

Received: 23 May 2013 – Published in Geosci. Instrum. Method. Data Syst. Discuss.: 10 June 2013
Revised: 18 November 2013 – Accepted: 5 December 2013 – Published: 18 December 2013

Abstract. Surface temperature is a key aspect of weather and climate, but the term may refer to different quantities that play interconnected roles and are observed by different means. In a community-based activity in June 2012, the EarthTemp Network brought together 55 researchers from five continents to improve the interaction between scientific communities who focus on surface temperature in particular domains, to exploit the strengths of different observing systems and to better meet the needs of different communities. The workshop identified key needs for progress towards meeting scientific and societal requirements for surface temperature understanding and information, which are presented in this community paper. A “whole-Earth” perspective is required with more integrated, collaborative approaches to observing and understanding Earth’s various surface temperatures. It is necessary to build understanding of the relationships between different surface temperatures, where presently inadequate, and undertake large-scale systematic intercomparisons. Datasets need to be easier to obtain and exploit for a wide constituency of users, with the differences and complementarities communicated in readily understood terms, and realistic and consistent uncertainty information provided. Steps were also recommended to curate and make available data that are presently inaccessible, develop new observing systems and build capacities to accelerate progress in the accuracy and usability of surface temperature datasets.

1 Introduction

Surface temperature is a key aspect of weather and climate, relevant to human health, agriculture and leisure, ecosystem services, infrastructure development and economic activity. The term “surface temperature” encompasses several distinct temperatures that differently characterise even a single place and time on Earth’s surface, as well as encompassing different domains of Earth’s surface (surface air, sea, land, lakes and ice; see Fig. 1).
Different surface temperatures play interconnected yet distinct roles in Earth’s surface system, and are observed with different complementary techniques. To better meet the needs of various applications and users communities, creative exploitation of the strengths of different observing system components is needed. Cooperation between scientific communities who focus on particular domains of Earth’s surface and on different components of the observing system is essential to accelerate scientific understanding and multiply the benefits of this understanding for society. A “whole-Earth” perspective on surface temperature is required. With this in mind, the EarthTemp Network held its inaugural meeting in June 2012 (Edinburgh, UK). The 55 participants convened from five continents with expertise on all of Earth’s surfaces and a full range of relevant techniques. The workshop identified the following needs for progress towards meeting societal needs for surface temperature understanding and information:

- develop more integrated, collaborative approaches to observing and understanding Earth’s various surface temperatures;
- build understanding of the relationships between different surface temperatures, where presently inadequate;
- demonstrate novel underpinning applications of various surface temperature datasets in meteorology and climate;
- make surface temperature datasets easier to obtain and exploit for a wider constituency of users;
- consistently provide realistic uncertainty information with surface temperature datasets;
- undertake large-scale systematic intercomparisons of surface temperature data and their uncertainties;
- communicate differences and complementarities of different types of surface temperature datasets in readily understood terms;
- rescue, curate and make available valuable surface temperature data that are presently inaccessible;
- maintain and/or develop observing systems for surface temperature data;
- build capacities to accelerate progress in the accuracy and usability of surface temperature datasets.

The needs are broadly expressed above. Twenty-eight specific ambitious steps, relevant to these objectives, are recommended in the remainder of this community position paper and, for easy reference, also summarised in Table 1 and Fig. 2. Our recommendations can also be seen as a concrete application of many of the Global Climate Observing System (GCOS) climate monitoring principles (Global Climate Observing System, 2003).
Table 1. The steps recommended in this paper for improving our understanding of Earth’s surface temperatures.

| Short description | Description |
|-------------------|-------------|
| 1 A whole-Earth perspective | We recommend the scientific communities, agencies and programmes involved in surface temperature research and applications develop more integrated, collaborative approaches to observing and understanding Earth’s various surface temperatures. |
| 2 Build understanding of the relationships of different STs | We recommend work to build understanding of the relationships of different surface temperatures, where presently inadequate. |
| 2.1 Reconcile discrepancies of satellite IST and in situ measurements | Satellite IST and field measurements over ice sheets and sea ice show discrepancies that are not fully explained. We recommend continuation of intensive efforts to reconcile these. |
| 2.2 Dialogue between ST and NWP re-analysis communities | We recommend closer dialogue between the surface temperature and NWP re-analysis communities, to clarify the correspondence between model and observed surface temperatures and to maximise their mutual exploitation. |
| 2.3 Global analysis of LSAT vs. LST | We recommend global systematic analysis of LSAT vs. LST relationships. |
| 2.4 Elucidate STs along “edge-lands” (marginal ice zone, coastal zones, suburbs) | We recommend research to elucidate the inter-relationships of surface temperature along “edge-lands”: the marginal ice zone (SST, IST, MAT), coastal zones (LSAT, LST, MAT) and suburbs (heat island fringes). There are complex issues around representation of surface temperature in the vicinity of boundaries and transition zones between domains. |
| 3 Demonstrate new underpinning applications of various ST datasets | We recommend demonstration of new underpinning applications of various surface temperature datasets in meteorology and climate. |
| 3.1 STs in NWP and re-analysis | The exploitation of improved LST, LSWT and IST within numerical weather prediction and re-analysis should be further demonstrated. |
| 3.2 Development of climate-quality time series from satellite obs. | Climate quality, > 10 yr long time series of LST, LSWT, SST and IST should be systematically developed from satellite observations (some exist), assessed against in situ-based trends and exploited in climate model evaluation. |
| 3.3 Advance use of LST for (sub)urban temperatures | The use of LST in understanding urban and suburban temperature distributions (heat island effects) should be advanced. |
| 3.4 Trial use of LSTs to validate adjustments of LSAT time series | There should be large-scale trials of the use of LSTs to help validate step-change detection and adjustments applied to LSAT time series from weather stations. |
| 3.5 Develop use of LSTs for interpolating LSAT | The use of LSTs in informing interpolation of LSAT across areas without meteorological stations should be developed. This includes historical reconstruction using spatially complete modes of variability. |
| 4 Make ST datasets easier to obtain and exploit for a wide constituency of users | We recommend that surface temperature datasets of all types be made easier to obtain and exploit for a wide constituency of users. Specific steps towards this need to be undertaken with extensive consultation of potential users. |
| 4.1 Create and sustain GDAC and LTSF | Regarding satellite datasets, we recommend creating and sustaining a global data assembly centre (GDAC) and long-term stewardship facility (LTSF) that collect, curate and disseminate datasets in common, self-describing formats, with free and open data access. |
| 4.2 ST providers should participate in Obs4MIPS | We recommend that surface temperature data providers with datasets relevant to climate modelling applications should participate in Obs4MIPS. |
| 4.3 Expand and simplify access to in situ ST data | We recommend expanding and simplifying access to the fundamental data holdings for in situ surface temperature records of all types. |
| 5 Consistently provide realistic uncertainty information | We recommend that all surface temperature measurements or estimates be provided with a realistic estimate of surface temperature uncertainty. |
| 5.1 Validate uncertainty information | We recommend that uncertainty information associated with surface temperatures measurements or estimates is itself subject to validation. |
| 5.2 Develop common uncertainty vocabulary | We recommend that a common uncertainty vocabulary be developed and adopted by the surface temperature community, building where possible on agreed usage from the metrological community. |
| 5.3 Improve interactions across community and users | We identify the need for improved interactions on the topic of uncertainty characterisation, across the surface temperature science community and with users. |
|编号 | 简要描述 | 详细描述 |
|---|---|---|
| 6 | 实施大规模系统的ST数据集和它们不确定性的互相比对 | 我们建议，所有开发和扩展ST数据集的项目都应投入资源进行大规模系统的互相比对（包括表面温度及其不确定性）。 |
| 6.1 | 系统的互相比对不同的观测和记录实践 | 我们建议，进行大规模系统的互相比对不同观测和记录实践，如气象服务对表面温度的观测。 |
| 6.2 | 开发多产品组合 | 我们建议开发一个直接可比的多产品组合，代表不同的ST数据集。 |
| 7 | 通信差异和ST数据集在容易理解的术语中互补性 | 我们建议改进ST数据集差异和互补性的通信。 |
| 7.1 | 为一般科学用户撰写的评估论文 | 我们认为需要撰写一篇评估论文，详细介绍ST数据集的差异和互补性。 |
| 7.2 | 采用通用方法提供用户简介 | 我们建议ST数据集生产商采用通用方法提供用户简介（约5页）。 |
| 8 | 救援、编目和可获取有价值的ST数据 | 我们建议协调救援、编目和可获取有价值的ST数据。 |
| 8.1 | 支持数据的救援、编目和访问历史气象观测 | 我们支持数据救援和编目，以及历史气象观测数据的免费访问。 |
| 8.2 | EO机构应积极进行数据救援和管理 | 我们建议，负责ST数据的EO机构应积极进行数据救援和管理。 |
| 8.3 | 国际协调数据救援和编目 | 我们建议协调国际数据救援和编目。 |
| 9 | 维护或开发观测系统 | 观测系统需要维护或开发。 |
| 9.1 | 维持卫星星座以观测所有STs | 我们建议维持卫星星座，包括GHRSST推荐的SST。 |
| 9.2 | 开发和维护全球多波段高分辨率热成像 | 除了9.1的建议外，还需要维护全球多波段高分辨率热成像。 |
| 9.3 | 部署更多用于长期气候参考的LSAT站点 | 我们建议部署更多用于长期气候参考的LSAT站点。 |
| 9.4 | 部署新长期站点用于卫星ST数据的验证 | 我们建议部署新长期站点用于卫星ST数据的验证。 |
| 10 | 增强加速ST数据的容量 | 我们建议增强加速ST数据的容量。 |
| 10.1 | 增强放射传输模拟的容量 | 我们建议增强放射传输模拟的容量。 |
| 10.2 | 增强多传感器匹配的共享能力 | 我们建议增强多传感器匹配的共享能力。 |


2. **Recommended steps towards integrated understanding**

The temperature at a location on Earth’s surface is profoundly important. Surface temperature is a basic environmental/meteorological parameter that directly affects human life and well-being; influences the function and viability of ecosystems, including agriculture; exercises controls on surface–atmosphere exchanges of energy, water, gases and aerosols; and is a primary variable of climatology and one indicator of climate change. For these reasons and more, the scientific and societal importance of surface temperature has long been obvious, and surface air temperature has been observed and investigated quantitatively for several hundred...
years (Middleton, 1966; Peterson and Vose, 1997; Strange-ways, 2009).

We live in an era of operational numerical weather prediction (NWP), Earth Observation and rapid data communications. Measurements, indirect estimates and information that constrains surface temperatures are available. The availability of different types of surface temperature observation differs enormously in frequency, spatial density, spatial completeness, and length and consistency of record. In some ways, we are simultaneously data-rich and data-poor with regards to surface temperature observations.

Surface temperature is not only profoundly important: it is complex. There are in fact several “surface temperatures” that can characterise a given place at a given time (see below). These distinct surface temperatures interact, they partially co-vary (albeit with distinct time constants), they play distinct geophysical and ecological roles, and often vary rapidly with time and distance.

2.1 Recommendation 1: a whole-Earth perspective

(R 1) We recommend the scientific communities, agencies and programmes involved in surface temperature research and applications develop more integrated, collaborative approaches to observing and understanding Earth’s various surface temperatures in order to accelerate progress in this area and multiply the benefits to society. This whole-Earth perspective aims to understand and exploit all forms of surface temperature observation across all domains, to develop clearer, more integrated and more informative knowledge of the surface temperatures of Earth, how they vary and how they may be changing. To multiply benefits and services to science and society, this activity needs to be supplemented by knowledge exchange, both to communicate comprehensive data and insight conveniently to users and to draw in improved understanding and refined requirements from users.

The surface temperature observations included in this comprehensive perspective are as follows (see also Fig. 1):

- land surface air temperature (LSAT) measured at approximately 2 m height at meteorological stations;

- land surface temperature (LST) estimated from satellite thermal and passive microwave sensors measurements;

- marine air temperatures (MAT) measured from ships and buoys;

- sea surface temperatures measured at depth from ships, buoys, etc. (SST-depth);

- sea surface temperature estimated from satellite thermal sensors (SST-skin) and passive microwave sensors (SST-subskin) measurements;

- lake surface water temperature (LSWT, skin and depth), both measured in situ and estimated from satellites, and including inland seas, reservoirs, etc.;

- ice surface temperatures (IST), both measured in situ and estimated from satellites (IST is sometimes also called LST in the literature when referring to land-based ice);

- various more specialist surface temperature measurements (in situ thermal radiometry ice-buoy thermistor chains, micrometeorological measurements, etc.).

2.2 Recommendation 2: build understanding of the relationships of different surface temperatures, where presently inadequate

While there is understanding of the relationships of different surface temperatures, research is required in several areas to reach a maturity of understanding where we can effectively exploit all forms of surface temperature observation across all domains, to develop clearer, more integrated and more informative knowledge of the surface temperatures of Earth, how they vary and how they may be changing.

(R 2) We recommend work to build understanding of the relationships of different surface temperatures, where presently inadequate.

(R 2.1) Satellite IST and field measurements over ice sheets and sea ice show discrepancies that are not fully explained. We recommend continuation of intensive efforts to reconcile these. For example, Hall et al. (2008) compared satellite-derived IST products with in situ observations over Greenland and found large apparent uncertainties in the in situ data, possibly related to unrepresentative local surface topography and other local factors, while the satellite-derived IST was shown to be of low relative bias but unknown precision.

Strategic efforts are required to investigate these discrepancies, and Recommendation 9.4 on dedicated reference sites is relevant.

(R 2.2) We recommend closer dialogue between the surface temperature and NWP re-analysis communities, to clarify the correspondence between model and observed surface temperatures and to maximise their mutual exploitation (see also Recommendation 3.1).

For example, satellite LST products provide useful information about surface energy and water cycles, and can be used in land data assimilation systems to monitor the climate and climate change (Reichle et al., 2009, 2010; Ghent et al., 2010, 2011). Data from meteorological stations have been shown to be useful for assessing re-analysis products (Simmons et al., 2004, 2010).

However, significant challenges remain for the use of ST in NWP, particularly over land. Typical issues are discrepancies between the spatial and/or temporal coverage, insufficient knowledge of surface emissivities and the geophysical
interpretation of the surface layers in NWP models in relation to observed ST.

(R 2.3) We recommend global systematic analysis of LSAT vs. LST relationships. The programme of research should encompass statistical relationships and how these vary with meteorological, micrometeorological, geographical and landcover context, taking into account that different types of ST may show distinct trends under transient climate change; model/process studies and experiments designed to account for observed relationships; and assessment of observed relationships in comparison to those present in major re-analysis products. Such a programme would support developments such as merged LSAT–LST datasets and use of LST in validating interpolated/gridded LSAT datasets.

(R 2.4) We recommend research to elucidate the inter-relationships of surface temperature along “edge-lands”: the marginal ice zone (SST, IST, MAT), coastal zones (SST, LSAT, LST, MAT) and suburbs (heat island fringes). There are complex issues around representation of surface temperature in the vicinity of boundaries and transition zones between domains. Datasets straddling such boundaries can disagree markedly depending on the assumptions made about how to combine data. Basic observational challenges for remote sensing are often even more complex because of the heterogeneity of “edge-lands”, where LST, SST and/or IST may be less accurate. For more detailed discussions of the issues in different types of “edge-lands”, see e.g. Høyer et al. (2012) for marginal ice zones, Castro et al. (2012) for small-scale coastal variability, and Armfield (2003) and Mirzaei and Haghhighat (2010) for reviews of urban/suburban inhomogeneities. Moreover, the true surface temperatures we seek to quantify and understand may be spatially variable (e.g. the land–sea temperature contrast) and interact (e.g. the land–sea breeze).

(Recommendations 2.3 and 2.4 are fundamental also to Recommendation 6 below.)

2.3 Recommendation 3: demonstrate new underpinning applications of various surface temperature datasets in meteorology and climate

The recommendations made in this paper are justifiable only because of the significant benefits foreseen from better, more accessible, more consistent surface temperature datasets. Applications in areas as diverse as crop monitoring and modelling (e.g. Moulin et al., 1998; Moran, 2003; Ramírez-Villegas and Challinor, 2012; Ericsson et al., 2012), volcano monitoring (e.g. Ramsey and Harris, 2013), water management (e.g. Agam et al., 2008; Kalma et al., 2008; Anderson et al., 2011, 2012), emergency responses to heat waves (e.g. Endlicher et al., 2008; Tong et al., 2012; Vanecкова et al., 2011; Zanobetti et al., 2012), infectious disease control (e.g. Reisen et al., 2006; Midekisa et al., 2012), lake ecology (e.g. MacCallum and Merchant, 2012), sea-ice forecasting for shipping, and many more will benefit downstream.

Weather forecasting and climate services exploit understanding of various forms of surface temperature, and progress here underpins a wide range of benefits to society.

(R 3) We recommend demonstration of new underpinning applications of various surface temperature datasets in meteorology and climate.

(R 3.1) The exploitation of improved LST, LSWT and IST within numerical weather prediction and re-analysis should be further demonstrated.

SST is already used widely for NWP, and the UK MetOffice’s Operational SST and Sea Ice Analysis System (OSSTIA) (Stark et al., 2007; Donlon et al., 2012) has been designed to meet the needs of the NWP community. On the other hand, the exploitation of LST, LSWT and IST for NWP is much less well developed. Recommendation 2.2. is also relevant here.

(R 3.2) Climate quality, > 10 yr time series of LST, LSWT, SST and IST should be systematically developed from satellite observations (some exist), assessed against in situ based trends and exploited in climate model evaluation.

Examples for such existing datasets for SST are ARC SST, a 20 yr SST record from along-track scanning radiometers (ATSRs) produced in the ATSR Reprocessing for Climate (ARC) project (Merchant et al., 2012), and the NOAA Optimum Interpolation (OI) SST (Reynolds et al., 2002). For LSWT, there is the ARC-Lake database (MacCallum and Merchant, 2012, 2013) and the JPL Large Inland Waterbody Database, which comprises AVHRR, MODIS and ATSR-series datasets (Schneider et al., 2009; Schneider and Hook, 2010). Development of LSWT in particular is rendered difficult by lack of accessible in situ validation data for many major lakes outside of North America and Europe (MacCallum and Merchant, 2012).

As far as we are aware, there are no long-term global LST datasets.

(R 3.3) The use of LST in understanding urban and suburban temperature distributions (heat island effects) should be advanced. LST influences our understanding of radiation, heat fluxes, evapotranspiration and other climatic factors in urban environments, and thermal remote sensing is valuable for assessing urban temperature effects, e.g. because of its geographically complete coverage (Stefanov et al., 2001; Carlson, 2003; Voogt and Oke, 2003). However, low temporal coverage and viewing angles that do not cover the three-dimensionality of the urban canyon create limitations (Mirzaei and Haghhighat, 2010), and the downscaling of satellite thermal imagery for uses in urban climatology remains a challenge (e.g. Stathopoulou and Cartalis, 2009; Essa et al., 2013).

(R 3.4) There should be large-scale trials of the use of LSTs to help validate step-change detection and adjustments applied to LSAT time series from weather stations.

Many LSAT time series from weather stations have inhomogeneities, e.g. due to site moves, changes in local site environment or instrument, or observing practice changes...
Users of surface temperature information are varied, and no single type of surface temperature dataset or spatio-temporal resolution will meet all their requirements. Users vary in their capacity to identify and obtain suitable environmental datasets and in their capacity to handle varied data formats. There are probably many potential non-specialist surface temperature data users in areas of health, planning, agriculture, etc. Measures such as the adoption of common surface temperature file contents and a common format may greatly expand the user base.

At the same time, the current diverse formats and contents are often driven by the interests of specific target communities with their own norms. For example, some communities expect datasets to be readily understood by geographical information systems, whereas the climate community would expect the same data in netCDF files compliant with the climate and forecasting (CF) convention. A solution is a common standard across all domains, formats tailored to specific communities and tools to convert data between standards.

(R 4.2) We recommend that surface temperature data providers with datasets relevant to climate modelling applications should participate in Obs4MIPs where this is not already the case.

(R 4.3) We recommend expanding and simplifying access to the fundamental data holdings for in situ surface temperature records of all types. This recommendation seeks to build on the progress of the International Surface Temperatures Initiative (ISTI) (Thorne et al., 2011; Lawrimore et al., 2013) in rescuing, standardising and serving free and open meteorological station data from a single portal. In principle, it is attractive to integrate records from ISTI, the International Comprehensive Ocean-Atmosphere Dataset (ICOADS) (Worley et al., 2005; Woodruff et al., 2011), and the International Arctic Buoy Programme (IABP) (http://iabp.apl.washington.edu) into a shared access point. This would not supersede the need for ISTI, ICOADS, IABP and similar programmes but would rather depend on such programmes, and arguably would augment their reach into wider surface temperature user communities.

A useful interim step is a single location where links are maintained to freely available datasets covering all ST types and domains (e.g. GHCNDaily, ECA&D and many national datasets).

2.5 Recommendation 5: consistently provide realistic uncertainty information with surface temperature datasets

Uncertainty information provided with surface temperature datasets needs to be consistently provided in two senses. First, uncertainty information should always be...
provided. Second, uncertainty information provided in different datasets needs to be comparable, certainly for different instances of the same sort of dataset, and ideally across different domains and types of observation.

Uncertainty is easily underestimated, and it is also easily misunderstood, both semantically (what do we mean by uncertainty?) and practically (how is it aggregated and propagated during processing of the data?). Good practice needs to be developed and adopted to make uncertainty information realistic. This will make it usable in contexts where relative uncertainty in different datasets is crucial, such as statistical and assimilation-based applications of surface temperature data. Following Einstein’s famous dictum, uncertainty information needs to be as simple as possible – but not simpler. Uncertainties need to be appropriately propagated when data are aggregated into higher-level products in order to ascribe realistic, consistent uncertainties to these higher-level datasets. This implies at least some representation of uncertainty components with differing degrees of spatial and temporal correlation. Equiprobable ensemble approaches are also attractive for capturing the complexities of uncertainties, where practicable. Users’ needs regarding uncertainty information will need to be surveyed. Perhaps less obviously, users’ exploitation of improved uncertainty information will need to be actively facilitated.

(R 5) We recommend that all surface temperature measurements or estimates be provided with a realistic estimate of surface temperature uncertainty. Uncertainty varies within products, from location to location (e.g. Jiménez-Muñoz and Sobrino, 2006; Freitas et al., 2010; Kennedy et al., 2011; Hulley et al., 2012; Guillevic et al., 2012). Much of this variation is usually amenable to quantification. Uncertainty information specific to each surface temperature measurement or estimate is preferable to generic estimates, but this is not universal practice. The surface temperature community should develop shared vocabulary and objectives about what forms of uncertainty information to provide. Probably, it is a necessary minimum to distinguish and quantify random, partially correlated and systematic components of uncertainty. Where components cannot be estimated and are missing from uncertainty estimates, this needs to be clear to give a fair picture to users. Providing uncertainty estimates does not supersede the need for quality- and/or confidence flags in datasets.

We note that this is a challenging area. Measurement uncertainty (which may have spatio-temporal correlation), parametric uncertainty and structural uncertainty may all be present in a dataset. The propagation of uncertainty from individual measurements through to end products can be complex. Interactions with metrologists and statistical experts can help define appropriate approaches to these challenges.

(R 5.1) We recommend that uncertainty information associated with surface temperature measurements or estimates is itself subject to validation. For confidence in the realism and comparability of surface temperature uncertainty estimates, the surface temperature community should develop shared approaches and good practice for validation of uncertainty information. It is important to minimise underestimation of uncertainty, avoiding situations where surface temperature products disagree by more than their supposed uncertainties plausibly explain. Relevant techniques will include inference from distributions of discrepancy between measurements from different components of the observing system, as well as triple collocation (multi-sensor) approaches (Diamond et al., 2013). It will sometimes be necessary to develop a better understanding of the true geophysical discrepancies between the different measurements. For structural components of uncertainty in creating datasets, the benchmarking approach can be informative (Venema et al., 2012; Williams et al., 2012; Thorne et al., 2011). (Benchmarking quantifies the impacts of different choices and methods of dataset generation using test cases that are synthetic, and thus perfectly known, and realistic.)

(R 5.2) We recommend that a common uncertainty vocabulary be developed and adopted by the surface temperature community, building where possible on agreed usage from the metrological community. This will facilitate communication on uncertainty and quality issues within the surface temperature community, with metrologists, and informed users. The vocabulary needs to be intellectually rigorous, and consistent with metrological usage where applicable. The vocabulary also needs to address all the types of uncertainty inherent in measuring or estimating surface temperature using satellite or in situ sensors and in creating and using such datasets, including those related to spatio-temporal sampling and correlation.

(R 5.3) We identify the need for improved interactions on the topic of uncertainty characterisation, across the surface temperature science community and with users. We recommend workshops involving producers and users, testing of different approaches to uncertainty information with use cases, and other interactions intended to improve the provision and exploitation of uncertainty information. We consider that appropriate uncertainties have to be conveyed clearly in terms recognisable by the users and answering their needs, whilst maintaining scientific detail behind the process for generation of the error characterisation. Workshops should explore the needs and formats for uncertainty estimates, the methods for calculating and conveying complex, correlated error estimates and representivity (sampling) errors in an accessible way, and the confidence in the error estimation process. Unification with appropriate vocabulary will be essential, as noted above. Such workshops should facilitate dialogue in both directions, informing users about data products and their uncertainties as well as informing data providers about the requirements of user communities.
2.6 Recommendation 6: undertake large-scale systematic intercomparisons of surface temperature datasets and their uncertainties

Users require guidance about the suitability of different surface temperature datasets for different applications, as well as information about how and why surface temperature datasets differ (see also Recommendations 2.3 and 2.4). Surface temperature data providers need to be able to summarise and interpret differences for users, and in part this depends on doing systematic intercomparisons of diverse surface temperature datasets.

Systematic intercomparison between different types of surface temperature datasets will be fruitful both in communicating the differences between different temperatures and in challenging and developing our understanding of the physics underlying differences. Systematic intercomparison between datasets of nominally the same sort of surface temperature is crucial in communicating the full degree of discrepancies across the choice of products that a user faces in selecting datasets for their application. For example, in the case of satellite LST, surface heterogeneity, geolocation uncertainty, spectral dependencies and view-angle dependencies produce relatively large, complex, localised differences in LSTs from different sensors (e.g. Jiménez-Muñoz and Sobrino, 2006; Freitas et al., 2010; Hulley and Hook, 2011; Guillevic et al., 2013). For a given sensor, LSTs generated using different methods and assumptions can likewise differ significantly (e.g. Hulley and Hook, 2009; Niclós et al., 2011; Göttzsche and Hulley, 2012).

Systematic comparison of differences between datasets with the corresponding estimated uncertainties may well reveal the need to uncover and estimate additional components of uncertainty that are sometimes neglected, but are nonetheless relevant to potential users. Satellite, gridded and blended surface temperature datasets are created by a complex sequence of steps in relation to data screening, aggregation and/or interpolation. These steps often involve detailed choices that, while based on reasoning and testing, are not fully objective. For example, when aggregating data, the weights of different inputs may depend on assumptions used to model various sources of uncertainty. Intercomparison of the consequences of such decisions can be fruitful at these intermediate stages, in addition to intercomparison of surface temperature and surface temperature uncertainty information. Benchmarking approaches are useful here (see also Recommendation 5.1).

(R 6) We recommend that all projects to develop and extend surface temperature datasets include resources dedicated to large-scale systematic intercomparisons (of both surface temperatures and their uncertainties). The intercomparisons need to include but extend well beyond comparison to proximate datasets (such as previous versions or alternative products derived from the same raw observations) in order to expose the full range of dataset differences relevant to potential users.

(R 6.1) We recommend more systematic intercomparison of the effects of different observing and recording practices between meteorological services on the surface air temperature record. Measurement and recording practices for LSAT from weather stations are only partly standardised across meteorological services by the World Meteorological Organization (Aguilar et al., 2003). Different practices significantly affect the absolute time series and climatology obtained (Parker, 1994; Brunet et al., 2008; van der Meulen and Brandsma, 2008; Brandsma and van der Meulen, 2008; Trewin, 2010), with more subtle impacts on anomaly time series (Peterson et al., 1998; Jones and Wigley, 2010). It is important to understand properly the effects of variations in practice on important surface temperature time series and on the relationships between different types of surface temperature (Recommendation 2). There is also a need for more systematic assessment of the impacts of different methods of homogenising LSAT records.

(R 6.2) We recommend development of a multi-product ensemble of directly comparable representations of different surface temperature datasets. The GHRSSST multi-product ensemble (GMPE) provides a useful precedent here. Surface temperature datasets may have, for good reasons, a range of spatial resolutions and binning/averaging in time. Large-scale intercomparison can be addressed by creating consistent representations of different datasets on a common time and space grid. These representations can then be readily manipulated to explore commonalities and differences. A web service providing on-the-fly visualisations of the ensemble and differences between members over time is a powerful way of allowing users to explore differences and build their understanding of surface temperature datasets.

2.7 Recommendation 7: communicate differences and complementarities of surface temperature datasets in readily understood terms

(R 7) We recommend improved communication of the differences and complementarities of surface temperature datasets in readily understood terms. In some cases, this needs to be underpinned by a firmer understanding of physical relationships between measurands (Recommendation 2). Recommendation 6.2 (for a surface temperature multi-product ensemble) is also relevant and needs to be augmented by a range of written information.

(R 7.1) We identify the need for a review paper, adopting a whole-Earth surface temperature perspective, explaining to general scientific users the range of surface temperature measurands, their physical significance, their interrelationships and the status of their corresponding measurements. Useful precursors for this exist covering certain domains. For example, Kerr et al. (2004) focus on a comparative overview of existing split window methods. LST
standard products from MODIS, SEVIRI, VIIRS and future GOES-R ABI sensors are based on these methods (with the effect of view angle explicitly represented by an additional term in the retrieval algorithms used for VIIRS and ABI). The work of Jacob et al. (2008) discusses the different types of temperature measurands in vegetated regions and their interrelation in the context of thermal infrared (TIR) remote sensing, and that of Li et al. (2013) reviews the current state of different algorithms for estimating LST from satellite TIR data.

(R 7.2) We recommend adoption by surface temperature dataset producers of a common approach to providing briefing notes (of approximately 5 pages) for users. Models for this exist (e.g. Obs4MIPs; Gleckler et al., 2011), and ideally an existing approach already in use within the surface temperature community should be adopted. It may be necessary to supplement an existing approach – for example, with a structured discussion of how the surface temperature of a particular dataset is different to, as well as complements, other types of surface temperature.

2.8 Recommendation 8: rescue, curate and make available valuable surface temperature data that are presently inaccessible

Many potentially valuable datasets are at present not (easily) accessible to the research community. Besides historical records, there is a wealth of data from individual research campaigns which often have not been used beyond the aims of the original project for which they were collected. Older records may not exist in digital form, and data (both in non-digital and digital form) are often kept by individual research groups or institutions, where they are difficult to discover and obtain. Their format and metadata are not always documented adequately, and datasets are in danger of disappearing due to organisational changes. Curating such data is necessary to obtain best value and realise the benefits from previous investments.

(R 8) We recommend coordinated efforts to rescue, curate and make available valuable surface temperature data that are presently inaccessible.

(R 8.1) We support data rescue and curation initiatives related to historical meteorological observations, and recommend these include free and open access to digitised data. Data rescue and curation is scientifically critical and an issue of intergenerational responsibility. We strongly support initiatives such as the All-Russia Research Institute of Hydrometeorological Information World Data Centre Baseline Datasets (http://meteo.ru/english/climate/), Old Weather (http://www.oldweather.org/), Atmospheric Circulation Reconstructions over the Earth (ACRE) (http://www.metr-acre.org/), ISTI (Thorne et al., 2011), ICOADS (including the “value added” initiative) (Woodruff et al., 2011) and Mediterranean Data Recovery (MEDARE) (http://www.omm.urn.cat/MEDARE/). (There may be many more of which we are not aware.) Necessary elements of the most useful initiatives are digitisation; free and open access online; convenient integration of new data within already-available datasets; and maintenance of datasets, including migration to new storage media. “Citizen science” approaches (Hand, 2010) to digitisation can be scientifically effective and cost effective, and can have benefits relating to public engagement with science. Other data recorded with surface temperature records (e.g. precipitation, pressure) should be digitised as part of a single effort, for reasons of cost effectiveness and because they help to interpret the temperature records. Recent significant progress made in meteorological data rescue is welcomed, yet there is much more that can be done. We also note that the World Meteorological Organization (WMO) “commits itself to broadening and enhancing the free and unrestricted international exchange of meteorological and related data and products” in its Resolution 40 (World Meteorological Organization, 1995).

(R 8.2) We recommend that space- and other agencies with responsibility for Earth-observation data relevant to surface temperature (and climate in general) are proactive in data rescue and stewardship. This includes recovery and curation of the satellite observations (at all data processing levels) and of all pre-flight and in-flight calibration information. Full calibration information is critical to future reprocessing of satellite observations and should be readily accessible and curated along with mission data. This will support satellite reprocessing initiatives using best techniques (arising from improved radiative transfer modelling including advances in understanding surface emissivity, improved input data from satellite data rescue and recalibration efforts, and theoretical advances in image classification and retrieval). The GHR SST community includes an initiative to rescue full-resolution (locally downlinked) NOAA meteorological satellite data (http://earthdata.nasa.gov/our-community/community-data-system-programs/ghrsst-avhrr-gac-hrpt), and cooperation with this initiative is recommended to all relevant holders of such data. There is relevant effort within the ERA-CLIM project (Dee et al., 2011) and the NOAA Climate Data Record programme (http://www.ncdc.noaa.gov/cdr).

(R 8.3) We recommend international coordination of a programme of data rescue and curation related to research campaign data that include meteorological observations, including surface temperature. Research campaign data with surface temperature information are often not publicly accessible, and can provide especially valuable data from sparsely observed regions and epochs. Such data, which are usually very high quality and taken at fine spatio-temporal resolution, can independently test satellite retrievals, merged datasets, re-analysis fields and/or historical reconstructions. We recommend a systematic effort to collect these data with all necessary metadata, engaging research councils and institutes internationally. The rescued data should be transformed to a standard form and made freely and openly accessible.
Good precedents exist that can be followed and extended, such as the open data access to observations of the Shipboard Automated Meteorological and Oceanographic System programme.

2.9 Recommendation 9: maintain and/or develop observing systems for surface temperature data

(R 9) Observing systems for surface temperature need to be maintained and/or developed. As regards satellite-based observations of surface temperature, the requirements for observations across all domains can be based on those articulated for the operational SST satellite constellation (Donlon et al., 2009), with some additional requirements.

(R 9.1) We recommend maintenance of a satellite constellation in line with GHRSST recommendations for SST as the baseline for a constellation for observing all surface temperatures. The SST constellation comprises complementary observations: ≈1 km resolution polar-orbiting visible and thermal imagery; frequent geostationary imagery for diurnal cycle observation; high-accuracy, low-noise, two-point-calibrated, dual-view thermal imagery; passive microwave for low-resolution, all-weather capacity with a channel suitable for high-latitude SST estimation. A GHRSST position paper from 2009 (Donlon et al., 2009) foresaw the risk of lack of continuity and overlap of the passive microwave and dual-view thermal components of the system, which unfortunately came to pass in 2011 and 2012 with the failure of the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) onboard the Aqua satellite (4 October 2011) and the loss of the Advanced Along Track Scanning Radiometer (AATSR) onboard Envisat, when contact to the satellite was lost unexpectedly (8 April 2012). The importance of appropriate redundancy of observation to maintain continuity, attaining both high spatial resolution and all-weather data, is re-affirmed here. From the whole-Earth perspective, additional elements in a satellite constellation, particularly diurnal observations, are required to maximise the utility of LST.

(R 9.2) In addition to the baseline from Recommendation 9.1, the whole-Earth surface temperature constellation requires development and maintenance of global multi-band thermal imagery with high spatial resolution (objective approximately 10 m). Progress can be made by improving fundamental knowledge of spatial heterogeneity in surface temperature and surface emissivity at scales smaller than resolved by the meteorological-style sensors (as demonstrated in some circumstances with ASTER). The high resolution is particularly required for understanding urban areas, which are extremely heterogeneous and of great societal relevance. We note that maintenance of a stable local observation time (by maintaining satellites in stable orbits) is more crucial for LST than for SST because of the larger diurnal cycle of the former.

(R 9.3) We recommend more LSAT sites specifically designed for long-term climate reference purposes at strategic locations globally, with access to specifications and metadata. One possible starting point for this is the US Climate Reference Network (http://www.ncdc.noaa.gov/crn) (Diamond et al., 2013).

(R 9.4) We recommend new long-term sites suitable for radiometric validation of satellite surface temperature and traceability to SI standards. Some should be co-located with LSAT climate reference observations. Long-term radiometric sites are required for land and ice surfaces spanning a wide range of climate regimes globally. No radiometric reference fixed sites currently exist for SST (in addition to regular cruise routes), and this should be rectified. All these sites should be maintained with quantified, high levels of stability.

2.10 Recommendation 10: build capacities to accelerate progress in the accuracy and usability of surface temperature datasets

(R 10) We recommend broad-impact building of capacity to accelerate progress in the accuracy and usability of surface temperature datasets. Specific examples follow below.

Significant errors in satellite LSWT (MacCallum and Merchant, 2012), LST (Jin and Dickinson, 2010), IST (Hall et al., 2008) and (to a lesser degree) SST arise, still, from image classification errors. Accurate surface temperature retrieval using thermal sensors depends on cloud-free conditions. Clear-sky conditions over water, land, ice and snow need to be distinguished from cloud-affected conditions, which is particularly challenging when the surface is highly reflective. Classification and cloud detection problems are also more acute along boundaries; for example, SSTs are routinely absent for the coastal pixels in many satellite datasets.

A key to rapid and consistent progress here is the capacity to estimate a priori, with known error covariance, the plausible radiiances for each possible class given the observational situation (atmospheric state and surface characterisation). The necessary radiative transfer knowledge and models exist, but are not integrated and easy-to-use across domains.

Therefore: (R 10.1) We recommend building integrated capacity for radiative transfer simulation across all surface-temperature-relevant sensors (all wavelengths/channels, surface domains, view and illumination conditions), in support of mitigating cloud detection errors in satellite surface temperature datasets.

(R 10.2) We recommend building shared capability for multi-sensor matched-data techniques across all domains of Earth’s surface. Relationships between varied in situ and satellite surface temperatures can be powerfully elucidated using matched multi-sensor data augmented by auxiliary information. Multi-sensor matching also supports improvement in surface temperature estimates, development of uncertainty information, validation of uncertainty information, interpretation of differences in surface temperatures, inter-comparison of sensors and algorithms, and design of quality control. Systems to provide multi-sensor match-up datasets...
are difficult to design and create. Some precedent exists with the SST community developed under ESA funding, e.g. the Sea Surface Temperature Climate Change Initiative (ESA SST CCI, http://www.esa-sst-cci.org/). Reference sites (Recommendations 9.3 and 9.4) would be an appropriate focus for initial developments of a multi-sensor match-up system. Such a system should also seek to enable Recommendation 5.

3 Conclusions

Significant benefits are foreseen to arise from better, more accessible, more consistent surface temperature datasets, and these justify the considerable effort that our recommendations require. The whole-Earth perspective adopted here will, we consider, accelerate progress and multiply benefits to society from investments in meteorological and Earth observation. This will happen because of the efficiency of shared capacity building, the willingness of the surface temperature community to share ideas and agree on common approaches, and because of increased quality, accessibility and usability of surface temperature datasets. Improved dialogue with users will be required and is necessary to ensure the most effective translation of the improved surface temperature data into applications.

Appendix A

About the EarthTemp Network and the writing of this position paper

The EarthTemp Network (http://www.EarthTemp.net) is a funded research network, sponsored by the UK Natural Environment Research Council, with the aim of increasing international cooperation and progress in quantifying and understanding variability and change in surface temperature across all domains of Earth’s surface. The initiative does not aim to replace or supersede any existing programmes or activities but rather to build collaboration.

The EarthTemp Network hosted its first workshop in Edinburgh in June 2012. Fifty-five participants gathered from five continents attended, with almost all of the desired range of expertise represented: scientists working on every domain of Earth’s surface, making or using in situ measurements, satellite products and re-analysis.

The meeting included networking activities to build relationships across the new community, overviews of the state of the art in the field, and a series of 20 intensive small-group discussions on current gaps in our knowledge and scientific priorities on 5 to 10 yr timescales across a number of themes. This position paper captures, as concretely as possible, the community conclusions of these discussion groups. The chairs of each discussion group, aided by notetakers, presented the outcomes of each group in plenary at the end of the workshop, with further opportunity to discuss and refine the points captured.

The principal investigator of the network took these presentations as the starting point to draft a discussion paper. The next draft captured the comments and amendments of the project’s co-investigators and international steering group, as well as those of the chairs and notetakers of the discussion sessions. Finally, the draft was sent to all participants for their comment and final approval. This process was intended to ensure that the final version is truly a consensus white paper from the EarthTemp Network membership. This version for peer-review publication was then developed from that.

The following researchers participated in the workshop and endorsed the draft of the discussion paper (Names in italics are co-authors of this paper): Lisa Alexander, Renate Auchmann, David Berry, Manola Brunet, Claire Bulgin, Gary K. Corlett, Katarzyna Dąbrowska-Zielińska, Emma Dodd, Claude Duguay, Owen Embury, Eirik Forland, Darren Ghent, Frank Göttsche, Elizabeth Good, Pierre Guillevic, Yatian Guo, Dorothy Hall, Jacob L. Høyer, Glynn Hulley, Maria A. Jiménez, Juan Carlos Jimenez-Muñoz, Philip D. Jones, Alexey Kaplan, John Kennedy, Elizabeth Kent, Albert Klein Tank, Jean-Pierre Lagouarde, Aisling Layden, Michele Lazzerini, David Llewellyn-Jones, Giuseppina Lopardo, Stuart MacCallum, Cristian Mattar, Stephan Matthiesen, Mark McCarthy, Matt Menne, Christopher J. Merchant, Colin Morice, Folke Olesen, Simon Pinnock, Fred Prata, Nick A. Rayner, Viatcheslav Razuvaev, John J. Remedios, Ignatius Rigor, Hervé Roquet, Matilde Rusticucci, Prashant Sardeshmukh, Philipp Schneider, Kay B. Smith, Jose Sobrino, Blair Trewin, Peter Thorne, Karen Veal, Yunyue Yu.

Acknowledgements. This work was supported by NERC grant NE/I030127/1. Thorne and Guillevic acknowledge support by NOAA through the Cooperative Institute for Climate and Satellites – North Carolina under cooperative agreement NA09NES4400006. Auchmann received travel support through COST Action “HOME” (ES0601).

Edited by: L. Eppelbaum

References

Agam, N., Kustas, W. P., Anderson, M. C., Li, F., and Coiazzzi, P. D.: Utility of thermal image sharpening for monitoring field-scale evapotranspiration over rainfed and irrigated agricultural regions, Geophys. Res. Lett., 35, L02402, doi:10.1029/2007GL032195, 2008.

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., and Wieringa, J.: Guidelines on climate metadata and homogenization: WCDMP-No. 53, WMO-TD No. 1186, WCDMP 53, World Meteorological Organization, Geneva, 2003.
C. J. Merchant et al.: The surface temperatures of Earth

ical activities, https://www.wmo.int/pages/about/Resolution40_en.html, 1995.

Worley, S. J., Woodruff, S. D., Reynolds, R. W., Lubker, S. J., and Lott, N.: ICOADS release 2.1 data and products, Int. J. Climatol., 25, 823–842, doi:10.1002/joc.1166, 2005.

Zanobetti, A., O’Neill, M. S., Gronlund, C. J., and Schwartz, J. D.: Summer temperature variability and long-term survival among elderly people with chronic disease, P. Natl. Acad. Sci. USA, 109, 6608–6613, doi:10.1073/pnas.1113070109, 2012.