Building and Environment 141 (2018) 143–154

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

# Surface heat assessment for developed environments: Optimizing urban temperature monitoring

Carl Malings[a,*], Matteo Pozzi[a], Kelly Klima[b], Mario Bergés[a], Elie Bou-Zeid[c], Prathap Ramamurthy[d]

[a] Civil and Environmental Engineering Department, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA
[b] RAND Corporation, 1776 Main Street, Santa Monica, CA, 90401, USA
[c] Department of Civil and Environmental Engineering, Princeton University, 59 Olden Street, Princeton, NJ, 08544, USA
[d] Department of Mechanical Engineering, City College of New York, 160 Convent Avenue, New York, NY, 10031, USA

ARTICLE INFO

*Keywords:*
Entropy
Value of information
Sensor placement
Optimization methods
Temperature
Urban areas

ABSTRACT

The urban heat island effect, exacerbated by rising average surface temperatures due to climate change, can lead to adverse impacts on city populations. Fine resolution modeling of the spatial and temporal distribution of extreme heat risk within a city can improve the strategies used to mitigate this risk, such as the issuance of targeted heat advisories to city residents. In this paper, we combine a recently developed method for probabilistic modeling of urban temperatures with previously developed vulnerability assessments, and then implement sensor placement optimization techniques to guide temperature monitoring in urban areas. A variety of metrics are used to optimize the placement of temperature measures to best support decision-making for monitoring and responding to extreme heat risk. This optimal sensor placement methodology is demonstrated for the city of Pittsburgh, PA, resulting in several proposed temperature monitoring schemes based on the various sensor performance metrics investigated. We quantitatively and qualitatively compare these schemes to identify the relative merits of each proposed metric.

## 1. Introduction

It is well established that increasing temperatures exacerbate both heat-related and air-quality-related illnesses and deaths [1], so it is important to understand how interactions between hazard, exposure, and vulnerability to high temperature create risk. These interactions are typically quantified as a product of terms relating to the hazard posed by a given factor, and the vulnerability and exposure of the population to this hazard.

Vulnerability indicates a population's sensitivity to extreme heat, and is usually quantified in terms of mortality [2], which is positively correlated with temperature beyond about 27 °C [3,4] (the exact threshold depends on the background climate to which the population has adapted). Many factors impacting vulnerability have been identified, e.g. age, socio-economic status, and whether individuals live alone [5–9]. Factor analysis has been used to identify which factors contribute the most to vulnerability [10–12]. Exact exposure to extreme heat is difficult to quantify due to daily population movements, and so vulnerability and exposure are typically analyzed together, with all individuals considered equally exposed [13–15].

Hazard indicates the severity of the heat to which a population is exposed. Hazard can be assessed in a variety of ways, including average air temperature [13], surface temperature at a specific time [11,14], or the number of heat wave days per year [15]. In previous work [16], we developed a probabilistic surrogate Gaussian process model of spatio-temporal heat hazard in an urban area that captures the properties of the urban heat island (UHI) and quantifies the variability in surface temperatures. Measured temperature data can update this prior model to an improved posterior (i.e. post-measurement) model via Bayesian techniques.

In this paper, we investigate the placement of sensors to monitor surface temperature in urban areas. This monitoring is done in order to reduce the physical and monetary consequences of extreme heat through better-informed response and mitigation activities. There have been numerous efforts to optimize sensing in indoor environments to support decision-making about temperature [17,18], air quality [19,20], and energy usage [21] within buildings in an effort to improve human health and comfort while reducing operational costs. The relationship between the outdoor and indoor environments, their effects on one another, and the movement of people between these

* Corresponding author.
*E-mail address:* cmalings@andrew.cmu.edu (C. Malings).

https://doi.org/10.1016/j.buildenv.2018.05.059
Received 7 April 2018; Received in revised form 24 May 2018; Accepted 28 May 2018
Available online 29 May 2018
0360-1323/ 

environments throughout the day make the monitoring of outdoor temperatures an important consideration as well [22]. The practical aim of this monitoring is to support strategies to address urban heat risk. These include short-term responses such as the issuance of heat advisories [13] or the opening of heat shelters [23] and long-term strategies including cool or green roofs [24,25] or green infrastructure [26]. These actions have intrinsic costs (e.g. economic disruption, increased energy consumption) and benefits (reduced mortality and morbidity), and can have varying effectiveness [27].

To address the issue of outdoor urban temperature monitoring, we make use of a calibrated spatio-temporal heat hazard model for the city of Pittsburgh, Pennsylvania, created using high-resolution WRF-PUCM urban microclimate simulations [28,29] of the region, as described in a recent publication [16]. Using this model, we apply a set of sensor placement metrics for the case study area to optimize the collection of information [30,31]. Specifically, we investigate the use of conditional entropy [32], weighted prediction error [33,34], and value of information [35–37] metrics to improve the predictive capabilities of the heat hazard model. We compare their performance to draw general conclusions on their applicability to urban temperature monitoring.

## 2. Heat hazard, vulnerability, and risk

This section describes the hazard model of [16] and the vulnerability assessment of [12], which we combine into an urban heat risk model.

### 2.1. Heat hazard modeling

We begin by summarizing the probabilistic urban temperature model developed in Ref. [16]; the reader is referred to that paper for details. We denote the spatio-temporal field of surface temperature at location $\mathbf{x}$ and time $t$ as $T(\mathbf{x}, t)$, and decompose it as:

$$T(\mathbf{x}, t) = T_0(t) + T_1(\mathbf{x}, t) + T''(\mathbf{x}, t) \tag{1}$$

where $T_0(t)$ is the average temperature of the region at time $t$ (note that in this paper time is always discretized hourly), $T_1(\mathbf{x}, t)$ is the temperature pattern (which captures the expected or typical daily cyclic UHI pattern) and $T''(\mathbf{x}, t)$ is the residual, i.e. the variability in temperature that is not captured in the previous terms. Fig. 1 demonstrates the decomposition of the temperature field for a typical time in and around the city of Pittsburgh, PA.

For each term in Eq. (1), we calibrate a Gaussian process model. Similar models have been proposed for a wide variety of spatio-temporal phenomena, such as rainfall intensity, corrosion initiation, and seismic risk [36,38,39]. For a general overview of Gaussian process modeling, the reader is referred to [40]. To perform model calibration, we use fine resolution (1 km grid scale, 1 h temporal resolution) numerical weather simulation data from an historical extreme temperature event in the region. The resulting calibrated spatio-temporal temperature field model is:

$$T(\mathbf{x}, t) \sim \mathcal{GP}\left(\mu_{T_0}, \mathrm{K}_{T_0}(t_i, t_j) + \mathrm{K}_{T_1}(\mathbf{x}_i, t_i, \mathbf{x}_j, t_j) + \mathrm{K}_{T''}(\mathbf{x}_i, t_i, \mathbf{x}_j, t_j)\right) \tag{2}$$

where $\mu_{T_0}$ denotes the mean temperature for the region and the covariance functions $\mathrm{K}_{T_0}(t_i, t_j)$, $\mathrm{K}_{T_1}(\mathbf{x}_i, t_i, \mathbf{x}_j, t_j)$, and $\mathrm{K}_{T''}(\mathbf{x}_i, t_i, \mathbf{x}_j, t_j)$ describe the covariance between the regional average temperature, cyclic temperature pattern, and residual temperatures, respectively, at different spatio-temporal coordinates $\{\mathbf{x}_i, t_i\}$ and $\{\mathbf{x}_j, t_j\}$. This Gaussian process model defines a multivariate Gaussian distribution for $\mathbf{T}$, the vector of temperatures at all space-time coordinates in the spatial domain $X = \{\mathbf{x}_1, ..., \mathbf{x}_{n_X}\}$ (i.e. the set of points in a 1 km by 1 km grid covering the region of interest) and time duration $\tau = \{t_1, ..., t_{n_\tau}\}$ (i.e. the set of time steps spaced at 1 h intervals over the analysis period of interest):

$$\mathbf{T} \sim \mathcal{N}(\boldsymbol{\mu}_T, \mathbf{K}_T) \tag{3}$$

where $\boldsymbol{\mu}_T$ is the mean vector and $\mathbf{K}_T$ is the covariance matrix obtained by evaluating the Gaussian process model at all pairs of space-time coordinates.

Observations of the temperature at specific locations and times (or observations of related quantities such as the regional average temperature) can update this prior model. Let $Y$ denote a set of such observations (i.e. a set of locations and times where observations are taken) and $\mathbf{y}$ be the specific data collected under measurement scheme $Y$. We model these data as a linear function of $\mathbf{T}$ with zero-mean Gaussian noise:

$$\mathbf{y} = \boldsymbol{\beta}_Y \mathbf{T} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0, \mathbf{K}_\varepsilon) \tag{4}$$

where matrix $\boldsymbol{\beta}_Y$ specifies which elements (or linear combinations of elements) in $\mathbf{T}$ are observed under scheme $Y$, and $\boldsymbol{\varepsilon}$ denotes the observation noise, which is potentially correlated between measures as described by noise covariance matrix $\mathbf{K}_\varepsilon$. For problems with measurements taken over time, matrix $\boldsymbol{\beta}_Y$ should be defined differently at different times, such that observations which are planned for the future but have not yet been made are not included when defining the matrix.

Since any linear combination of Gaussian variables is itself described by a multivariate Gaussian, the distribution of measurement values is:

$$\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_Y, \mathbf{K}_Y) \tag{5}$$

where the expected value of the observations is $\boldsymbol{\mu}_Y = \boldsymbol{\beta}_Y \boldsymbol{\mu}_T$ and the covariance of the observations is $\mathbf{K}_Y = \boldsymbol{\beta}_Y \mathbf{K}_T \boldsymbol{\beta}_Y^{\mathrm{T}} + \mathbf{K}_\varepsilon$. Given measurement data $\mathbf{y}$, we update the prior model in Eq. (3) to a posterior model:

$$\mathbf{T}|\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_{T|\mathbf{y}}, \mathbf{K}_{T|\mathbf{y}}) \tag{6}$$

where the parameters are evaluated as:

$$\begin{aligned} \boldsymbol{\mu}_{T|\mathbf{y}} &= \boldsymbol{\mu}_T + \mathbf{K}_{TY}\mathbf{K}_Y^{-1}(\mathbf{y} - \boldsymbol{\mu}_Y) \\ \mathbf{K}_{T|Y} &= \mathbf{K}_T - \mathbf{K}_{TY}\mathbf{K}_Y^{-1}\mathbf{K}_{TY}^T \end{aligned} \tag{7}$$

with $\mathbf{K}_{TY} = \mathbf{K}_T \boldsymbol{\beta}_Y^{\mathrm{T}}$. Note that $\mathbf{K}_{T|Y}$ is a function of measurement scheme $Y$ only, while $\boldsymbol{\mu}_{T|\mathbf{y}}$ is a function of specific measurements $\mathbf{y}$. Following this procedure, we can process any measurement of the temperature field described by Eq. (4) to update the prior temperature model of Eq. (3) to a posterior model as in Eq. (6). We then use the posterior model for temperature prediction, taking any remaining uncertainty into account. Note that measurements processed as above may come from a variety of sources, e.g. measurements of local temperature $T(\mathbf{x}, t)$, predictions of regional temperature $T_0(t)$ from coarse-resolution weather models, and assessments of temperature pattern $T_1(\mathbf{x}, t)$ from prior fine resolution simulations of the region or remote sensing data (e.g. Ref. [41]).

### 2.2. Heat vulnerability

We quantify heat vulnerability using an index previously developed for Pittsburgh [12]. The index was created after analyzing 13 socio-economic metrics (listed in Table 1), such as age, income, and education. These metrics are identified in the literature as contributing to heat-related vulnerability [10,11,15]. A factor analysis was conducted to group these metrics into factors that explain at least 70% of the variance in all metrics. These factors were then averaged and binned into 6 increments for each city census block. The result was an index from 1 to 6, with 6 corresponding to a higher vulnerability (with the factors 2 or more standard deviations above the mean) and 1 to a lower vulnerability (2 or more standard deviations below the mean) [12]. In this paper, we define the vulnerability index $V(\mathbf{x})$ for location $\mathbf{x}$ (which applies to the one-square-kilometer box centered at that coordinate) as the area-weighted average of vulnerability indices associated with the census blocks located within that one-square-kilometer box. This vulnerability index field is depicted in Fig. 2.

The vulnerability index used here already considers exposure-

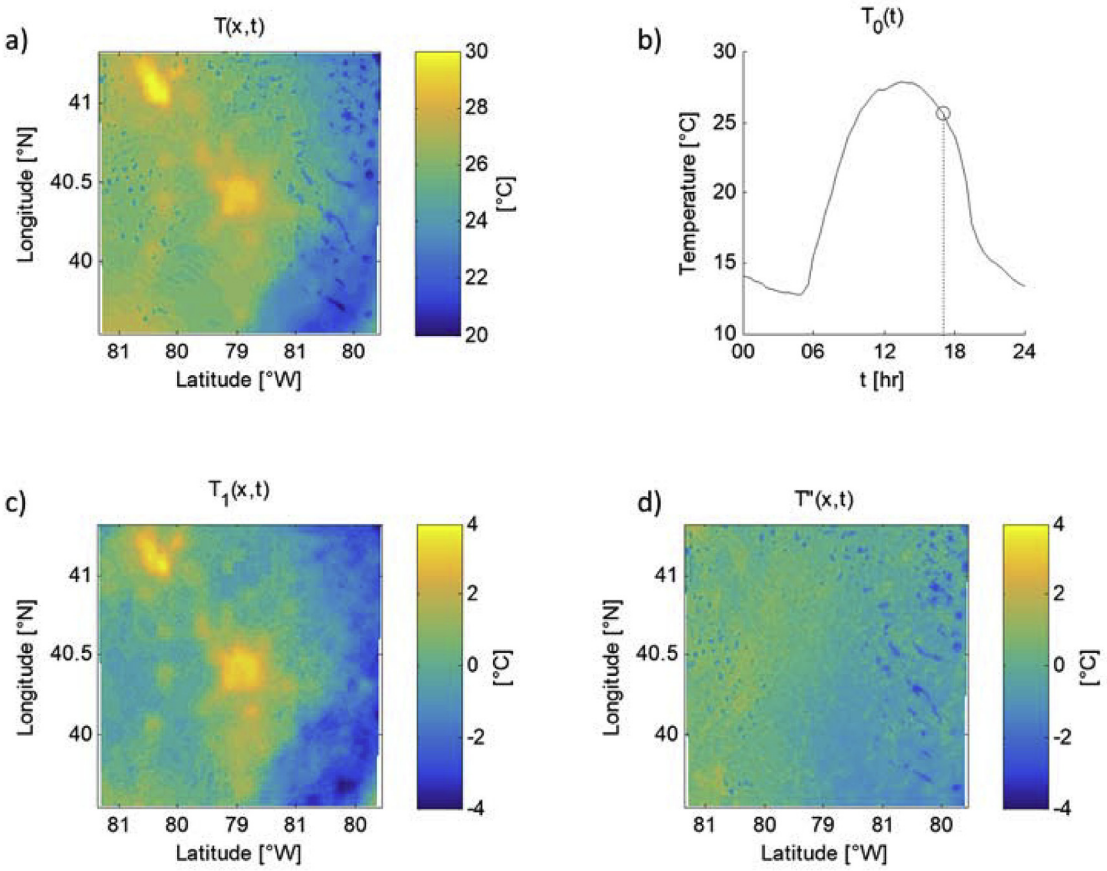


**Fig. 1.** Example of decomposition of the temperature field for Pittsburgh, PA, for $t$ = 17: 00 local time: a) total temperature field across the region at this time; b) regional average temperature for the day, with the current time indicated; c) temperature pattern for the region at this time of day; d) residual between the temperature pattern and the true temperature for this time.

related factors, such as access to air conditioning and to green spaces. Therefore, following the approach of [13], we consider exposure and vulnerability together in this work. We assume that all areas except those which are mostly made up by rivers have populations exposed to extreme heat. While a spatio-temporal model of urban population movements might be created to describe the varying distribution of people in a city throughout the day, no such models were identified in the literature and the development of a separate model is beyond the scope of this work. Therefore, we assume vulnerability and exposure are constant with time, i.e. variations in population distributions throughout the day are not considered, although the framework developed here could easily be extended to consider temporal variations in these properties when such data become available.

### 2.3. Heat risk

We combine our hazard value, i.e. the surface temperature $T(\mathbf{x}, t)$, with our spatial vulnerability index $V(\mathbf{x})$ to define a spatio-temporal risk index:

$$R(\mathbf{x}, t) = T(\mathbf{x}, t)\, V(\mathbf{x}) \tag{8}$$

The spatio-temporal risk is described by a Gaussian process model, since it is the product of a Gaussian process model with a deterministic scaling field. While we must note that actual health hazard is related to air temperature, humidity, wind speed, and ambient radiation as well as the surface temperature, here we seek to demonstrate a proposed methodology for optimizing sensor placement, and therefore, we restrict our analysis to surface temperature. The methodology presented here can be readily replicated for all other hazard factors, or for a

**Table 1**
Factors considered in the vulnerability index. The list is reproduced from Ref. [12].

| | |
|---|---|
| Persons over 65 years old | (by percent of census block population) |
| Persons living alone | (by percent of census block population) |
| Persons over 65 years old and living alone | (by percent of census block population) |
| Persons below the poverty line | (by percent of census block population) |
| Persons with poor English language skills | (by percent of census block population) |
| Persons without high school diplomas | (by percent of census block population) |
| Ethnic minorities | (by percent of census block population) |
| Immigrants from Latin America | (by percent of census block population) |
| Persons with Diabetes | (by percent of census block population) |
| Persons in homes lacking central air conditioning | (by percent of census block population) |
| Persons in homes lacking any air conditioning | (by percent of census block population) |
| Access to green space (average) | (by percent of census block area) |
| Access to nearby green space (variance) | (by percent of census block area) |

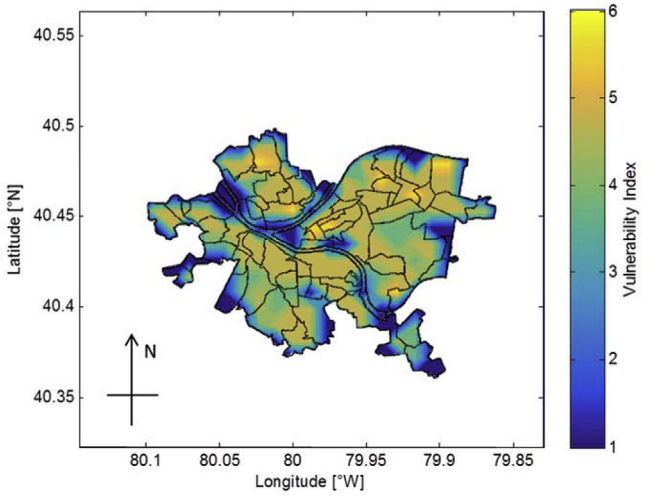


**Fig. 2.** Vulnerability index map for Pittsburgh, PA. Blue areas correspond to sparsely populated zones, usually due to the prevalence of parks, water, and/or industrial areas. Based on results from Ref. [12]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

combined thermal comfort indicator (e.g. Ref. [42]). Furthermore, note that this formulation ignores any potential synergistic effects or non-linear interactions between the temperature hazard and vulnerability index.

The probabilistic character of the heat hazard model allows for different heat risk indices to be derived from it. For instance, previous studies, e.g. Refs. [11,14], quantify hazard using the temperatures observed at a specific time during a heat wave. Such information can be derived in the Gaussian process hazard model via appropriate applications of Eqs. (6)–(8) to define a posterior temperature field conditioned on available observations. Alternatively [15], defines hazard using the number of heat wave days (i.e. days with peak temperatures exceeding a set threshold) per year. This measure can also be obtained from the Gaussian process hazard model by conducting a forward propagation analysis (e.g. using Monte Carlo simulations) and counting the number of days the threshold is exceeded.

As an example of how the probabilistic model can be used to generate a static heat risk index map (such as has been used in previous work), we consider the 95th percentile of the local temperature throughout the day as a metric to quantify the heat hazard. Let $T_{95\%}(\mathbf{x}, t)$ be the 95th percentile temperature at location $\mathbf{x}$ and time $t$, as predicted by the probabilistic model. We take the maximum of these temperatures over a 24 h period as the hazard, or $T_{95\%}(\mathbf{x})$. This quantity is the peak of a time series for the 95th percentile of the probabilistic temperature distribution at a given location, and represents a conservative estimate for the peak daily temperature. The model $T(\mathbf{x}, t)$ used to define these percentiles is calibrated for the region of Pittsburgh and conditioned on measurements of $T_i(\mathbf{x}, t)$ obtained from the training data set as discussed by Ref. [16]. The resulting $T_{95\%}(\mathbf{x})$ temperature field is depicted in Fig. 3. Note that this is only one potential method for defining hazard and is used here for illustrative purposes only. To allow comparison between hazard and vulnerability, these hazards are binned into 6 increments based on the range of values across the city. This is denoted by bin[·], which describes a mapping from the continuous temperature values to a integer between 1 and 6 (this is the same binning function used to develop the vulnerability index, as described in 2.2). Following this transformation, the hazard and vulnerability measures are both indices ranging from 1 to 6, and are directly multiplied to define a risk metric:

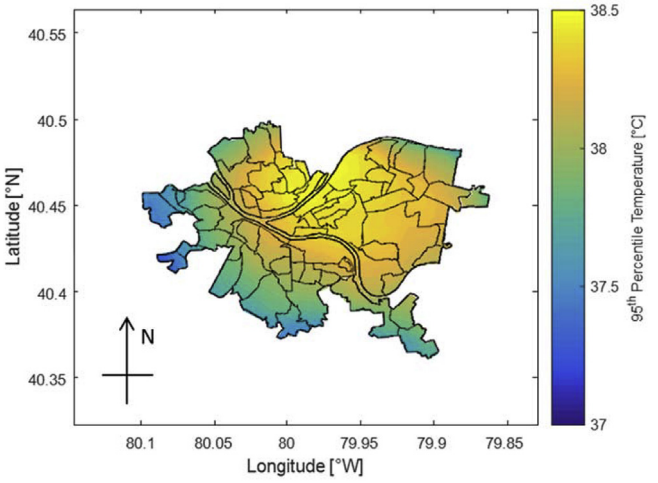


**Fig. 3.** Map for the 95th percentile peak daily temperature, $T_{95\%}(\mathbf{x})$, in Pittsburgh, PA.

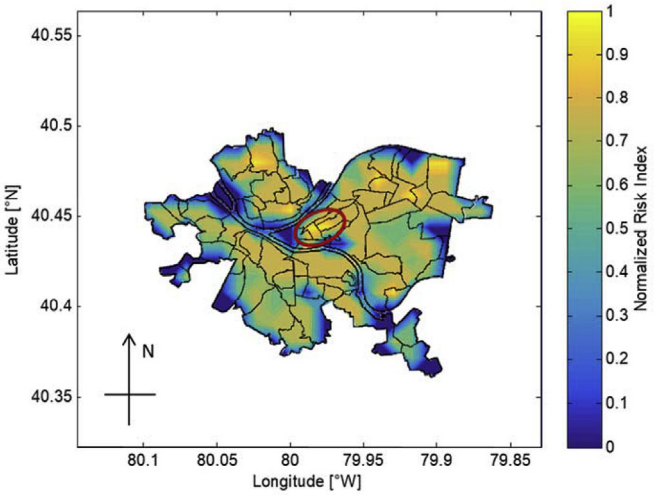


**Fig. 4.** An example risk index based on 95th percentile peak daily temperature in Pittsburgh, PA, and the vulnerability map of Fig. 2. This index is normalized between 0 and 1. Downtown Pittsburgh is indicated with a red circle. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

$$R_{95\%}(\mathbf{x}) = \text{bin}\left[\max_{t\in[0,24]} T_{95\%}(\mathbf{x}, t)\right] V(\mathbf{x}) \tag{9}$$

The resulting risk index map is shown in Fig. 4, normalized to vary from 0 to 1. High-risk areas include downtown Pittsburgh (indicated with a circle in the figure) as well as the northeast and northwest.

Many heat risk indices, including that of Eq. (9), involve non-linear functions of the temperature field, and so can no longer be described as Gaussian. Therefore, it is computationally more difficult to analyze these mappings directly in order to determine the utility of sensor placements in reducing uncertainty and supporting risk mitigation actions. Furthermore, static hazard measures (e.g. using the 95th percentile temperature, as demonstrated above) ignore the changing nature of hazard and risk over time, while the full spatio-temporal hazard model captures these changes. For these reasons, the following sections apply sensor placement metrics directly to the underlying Gaussian temperature and risk fields, as in Eq. (8). While not used directly, the risk index of Eq. (9) and Fig. 4 provides context on the nature of the urban heat risk in Pittsburgh.

## 3. Sensor placement optimization theory

Many factors should be considered when selecting optimal locations to measure temperature in an urban area. These include the uncertainty or variability in the local temperature field, the vulnerability of people in different areas to extreme heat, and the available actions (and their consequences) which can be undertaken to mitigate extreme heat risk. A map of the spatially distributed risk, such as Fig. 4, may not be sufficient. For example, monitoring temperature in places where risk is high is an intuitive solution, but may neglect areas where additional information could significantly reduce uncertainties in the risk; conversely, focusing on areas of high uncertainty but very low overall risk is also inefficient.

We investigate three sensor placement metrics for the urban temperature monitoring problem: a conditional-entropy-based metric, a metric based on the predictive performance of the temperature model through its weighted error, and a value of information (VoI) metric. Each of these metrics corresponds to a particular viewpoint on what information is important when monitoring temperature. The conditional entropy metric's goal is to reduce the uncertainty in the temperature field. The weighted prediction error metric is also focused on reducing uncertainty, but puts higher weight on uncertainty in areas which are assessed to be more vulnerable to extreme heat. VoI models a decision-making problem for issuing heat advisories and identifies additional information which might best support this decision-making. Additional details on these metrics are provided below as well as in Appendix C. To optimize sensor placement via these metrics, we make use of an efficient but approximate forward greedy optimization algorithm.

### 3.1. Conditional entropy

Entropy is a way of measuring the uncertainty in a random field [32]. Observations of this field can reduce its uncertainty, with this uncertainty after observation (termed the “posterior uncertainty”) quantified as the conditional entropy. An optimal sensor placement scheme will minimize uncertainty after observation, i.e., it will minimize the conditional entropy:

$$Y^* = \text{argmin}_Y \text{H}(T|Y) \tag{10}$$

where $\text{H}(T|Y)$ denotes the conditional entropy of temperature field $T$ conditioned on measurement set $Y$.

### 3.2. Weighted prediction error

Prediction error measures the sum of square differences (i.e. the square of the L-2 norm) between temperature $T(\mathbf{x}, t)$ and predicted temperature $\hat{T}(\mathbf{x}, t)$. The expected prediction error is:

$$\text{Err}(\hat{T}) = \sum_{t\in\tau} \omega_t \sum_{\mathbf{x}\in X} \omega_{\mathbf{x}} \mathbb{E}_T (T(\mathbf{x}, t) - \hat{T}(\mathbf{x}, t))^2 \tag{11}$$

where $\mathbb{E}_T$ denotes the expected value with respect to the uncertain temperature field. To account for the effects of spatially and (potentially) temporally-varying vulnerability, weighting coefficients $\omega_t$ and $\omega_{\mathbf{x}}$ are used. In this paper, we use the vulnerability index $V(\mathbf{x})$ of 2.2 to define the weight, with $\omega_{\mathbf{x}} = V(\mathbf{x}) \ \forall \ \mathbf{x} \in X$. In this way, prediction errors in more vulnerable areas are penalized more heavily, with penalties proportional to the product of the local vulnerability and the temperature variance. Note that, under this choice of weighting, this metric measures the uncertainty in the risk field $R(\mathbf{x}, t)$ of Eq. (8). Also note that we assign $\omega_t$ to be constant (i.e. $\omega_t = 1 \ \forall \ t \in \tau$) since we assume (as discussed in 2.2) that vulnerability and exposure do not vary with time. The optimal sensor placement by this weighted prediction error metric is that which minimizes the prediction error:

$$Y^* = \text{argmin}_Y \text{Err}(\hat{T}|Y\ ) \tag{12}$$

where $\text{Err}(\hat{T}|Y\ )$ denotes the expected weighted prediction error using the best possible prediction $\hat{T}$ of $T$, informed by the information collected under measurement scheme $Y$.

### 3.3. Value of information

VoI measures the utility of sensor measurements, quantifying how costs can be reduced in a decision-making problem by incorporating additional information to make better decisions [35]. In such a problem, an agent must select a set of actions $A$ for managing an uncertain system. For this paper, we consider the actions to be choices of whether or not to issue heat advisories to certain areas. Based on the selected actions and the true temperature field $T$, the agent suffers a loss or penalty $\text{L}(T, A)$. This penalty can be a monetary value, or any other mapping from the actions and temperatures to a descriptive quantitative value. A higher value of $\text{L}(T, A)$ represents a less desirable outcome. In this paper, the vulnerability index $V(\mathbf{x})$ of Section 2.2 is incorporated into the penalty function, such that failing to issue an appropriate heat advisory in a more vulnerable area will incur a higher penalty (see Appendix C for details). VoI is then defined as the difference between the prior and posterior expected losses:

$$\text{VoI}(Y) = \min_A \mathbb{E}_T \text{L}(T, A) - \mathbb{E}_Y \min_A \mathbb{E}_{T|\mathbf{y}} \text{L}(T, A) \tag{13}$$

That is, VoI is the amount by which the penalty is expected to be reduced by having access to information from $Y$ prior to making decisions $A$, and thereby being able to make more appropriate decisions based on more complete information. The optimal sensor placement scheme under the VoI metric is the set of measurements $Y$ that maximizes the net VoI:

$$Y^* = \text{argmax}_Y \text{VoI}(Y) - \text{C}(Y) \tag{14}$$

where $\text{C}(Y)$ denotes the cost to the managing agent of carrying out measurement scheme $Y$ (expressed in the same units as VoI). Appendix C provides further information about VoI and how it is evaluated, and introduces the concept of regret, the complement to the VoI.

### 3.4. Greedy optimization

Sensor placement is a problem of combinatorial optimization, i.e. selecting a few measurements from the set of potential measurements being considered. In general, the only guaranteed way to find the true optimal solution in this type of problem is to enumerate every possible set of measurements, compute the sensor placement metric value for each set, and finally select the best set from among all possibilities. In most practical problems, this exhaustive search method is infeasible, and so a fast but approximate approach is used [31].

For this paper, we use a forward greedy algorithm: this approach builds up the set of measurements one at a time. It does so by selecting the single measurement that most improves the sensing metric at each step and adding this measurement to the chosen set. Pseudo-code for forward greedy maximization is presented in Algorithm 1. This algorithm can be directly used for the VoI metric with Eq. (14) as the objective. For the minimization of conditional entropy and weighted prediction error, the objectives of Eq. (10) and Eq. (12) respectively should be multiplied by negative one for use as the objective.

In general, because this algorithm is an approximate solution method, there is no guarantee that it will produce an optimal result. However, both the conditional entropy and weighted prediction error metrics exhibit a set functional property known as submodularity, and theoretical guarantees on the performance of the greedy optimization approach on functions with this property are available [33,43]. For the VoI metric, there is unfortunately no such guarantee, but empirical evidence suggests it performs sufficiently well to justify its use in most problems [36].

**Algorithm 1.** Pseudo-code for forward greedy maximization

**Inputs:** Set of candidate measurements: $Y_{\text{all}}$
Objective function: $\text{Objective}(\cdot)$
Constraint function: $\text{Constraint}(\cdot)$
Upper limit for constraint: $b$

**Initialize:** $Y_{\text{selected}} \leftarrow \emptyset$
$Y_{\text{possible}} \leftarrow Y_{\text{all}}$

**while** $|Y_{\text{possible}}| > 0$
  $y_{\text{new}} \leftarrow \text{argmax}_{y \in Y_{\text{possible}}} \text{Objective}(Y_{\text{selected}} \cup y)$
  **if** $\text{Constraint}(Y_{\text{selected}} \cup y_{\text{new}}) \leq b$
    $Y_{\text{selected}} \leftarrow Y_{\text{selected}} \cup y_{\text{new}}$
  **end**
  $Y_{\text{possible}} \leftarrow Y_{\text{possible}} \backslash y_{\text{new}}$
**end**

**Output:** Set of selected measurements: $Y_{\text{selected}}$

In the above code, $\leftarrow$ denotes assignment, $\emptyset$ denotes the empty set, $|\cdot|$ denotes the size of a set, $\in$ denotes an element of a set, $\cup$ denotes the union operator, and $\backslash$ denotes set subtraction.

## 4. Case study: Pittsburgh, PA

We conduct sensor placement using a previously calibrated probabilistic model for urban temperatures in Pittsburgh [16] together with the three metrics discussed in 3, using the objective functions of Eqs. (10), (12) and (14), optimized by using the greedy algorithm discussed in 3.4. We assume that a) accurate forecasts of the average temperature in the city ($T_0$), b) prior simulations of the cyclic temperature pattern ($T_1$), and c) local temperature measurements are available to update the probabilistic model (note that this corresponds with use case number 4 in Ref. [16]). It is the locations of these latter local measurements which are being optimized. We consider this the most realistic use case because accurate regional temperature forecasts are commonly available (e.g., from the National Weather Service) and historical simulations of temperature patterns can be generated by WRF-PUCM (as in Ref. [16]).

We further assume that, once placed, a temperature sensor will continuously gather data over time at a negligible additional cost. Thus, we need only choose the locations within the domain ($X$) to place these monitors. By adjusting the method, simultaneous placement and scheduling of measurements could also be considered at the cost of increased computational time [44]. The potential measurement sites considered correspond to a 1 km grid over the city. This set of possible measurement sites, which we denote $Y_{\text{all}}$, consists of 122 locations, one for each grid point.

### 4.1. Optimal sensor placement via the conditional entropy metric

First, sensor placements are optimized to reduce the conditional entropy of the posterior temperature field within the city of Pittsburgh. For reference, the prior standard deviation (which is related to the prior entropy) is shown in Fig. 5. The results in Fig. 6 plot the locations of 9 greedily optimized temperature measurement locations throughout the city. These sensor locations are distributed evenly over the domain of interest to reduce overall uncertainty in temperature values throughout the city. The posterior standard deviation of the temperature field, i.e. the remaining uncertainty in the field after measurements are taken at the prescribed locations, is also indicated. Note that conditional entropy decreases as more sensors are considered, and therefore the set $Y_{\text{all}}$ would have the smallest conditional entropy. However, the use of this set is assumed to be impractical due to the high cost of installing so many sensors. The number of sensors to use is chosen such that the sets optimized under each metric have the same number (nine) of sensors, allowing for direct comparison between them, as discussed in Appendix B.

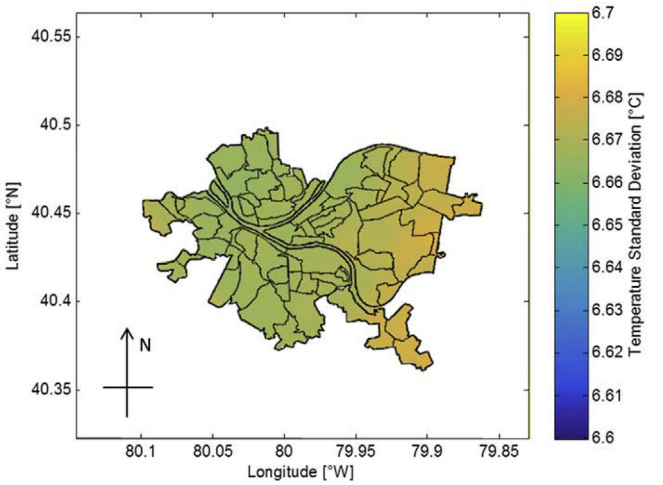


**Fig. 5.** Prior standard deviation of the temperature field $T''(\mathbf{x}, t)$ for Pittsburgh, PA, before the placement of sensors.

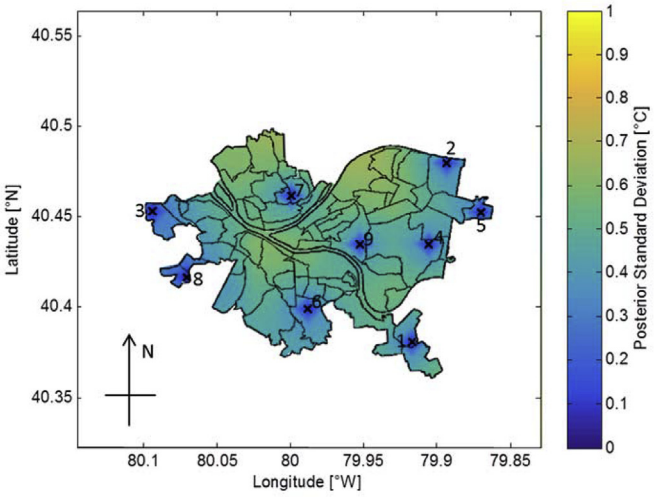


**Fig. 6.** Optimized measurement locations for reducing the conditional entropy of the temperature field in Pittsburgh, PA. Sensing locations are indicated by x's, while the pseudocolors indicate the posterior standard deviation of the temperature field conditional to these measurements. Numbers next to the sensing locations indicate the order they were selected by the greedy optimization algorithm. The city's outline is also indicated.

### 4.2. Optimal sensor placement via the weighted prediction error metric

Next, sensing locations are optimized based on the residual prediction error weighted by the local vulnerability to extreme temperatures, as discussed in 3.2. The results are shown in Fig. 7, plotting the nine locations where sensors should be placed to reduce the overall weighted temperature prediction error throughout the city. Again, even though adding more sensors will always decrease the prediction error, we use an optimized subset of $Y_{\text{all}}$ with the same number of sensors as selected under other metrics. Because the vulnerability index is used as a weighting factor, the sensing scheme shown here is different from that indicated in Fig. 6, although one sensor placement is common to both schemes. Here, sensor placements generally correspond with areas where temperature variability and vulnerability are both relatively

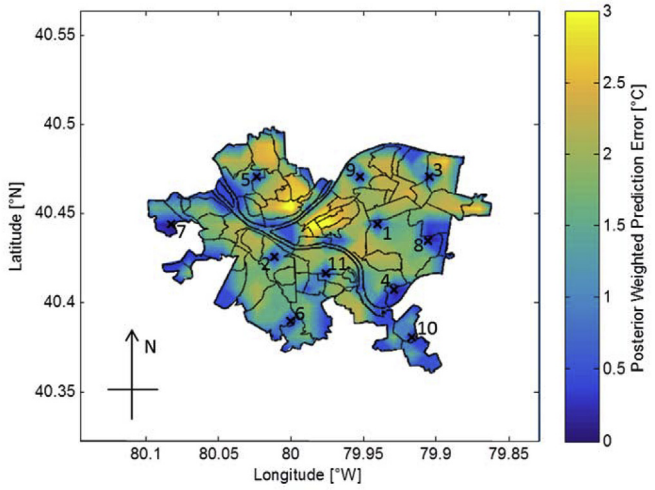


**Fig. 7.** Optimized measurement locations for reducing prediction error of the temperature field weighted by the vulnerability field in Pittsburgh, PA. Sensing locations are indicated by x's, while the pseudocolors indicate the posterior weighted prediction error conditional to these measurements. Numbers next to the sensing locations indicate the order the placements were selected by the greedy optimization algorithm. The city's outline is also indicated.

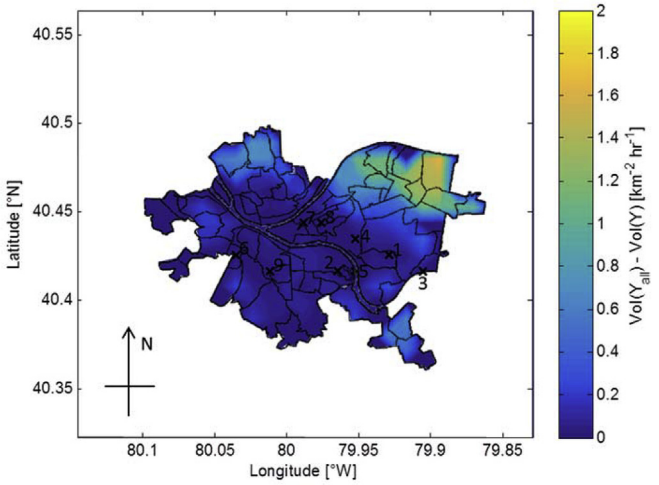


**Fig. 8.** Optimized measurement locations supporting decision-making for heat advisory issuance via the VoI metric in Pittsburgh, PA. Sensing locations are indicated by x's, while the pseudocolors indicate the expected regret (the difference between the VoI of all possible measures and the VoI provided by measuring at the selected locations) corresponding with the indicated sensor placements, distributed over space, for June 27, 2012 at 1630 h. Numbers next to the sensing locations indicate the order the placements were selected by the greedy optimization algorithm. The city's outline is also indicated.

high. Using this metric, the error in the temperature prediction across the city is higher than when using the conditional entropy metric, although the error in the risk prediction is lower, as we will illustrate in 4.4.

### 4.3. Optimal sensor placement via the value of information metric

Finally, we optimize sensor placements using the VoI metric as described in 3.3 to support the issuance of heat advisories. In this case, we assume that the cost for measuring temperature continuously during a summer is 0.025 per location (in the same units as the penalty function $L(T, A)$), representing about 1% of the economic penalty for issuing a heat advisory for that location for 1 h. Using this cost, the optimal number of sensors to use for this problem is evaluated to be 9 (the sensitivity of this result to the sensing cost is examined in Appendix B). The results of this optimization are shown in Fig. 8. The chosen measurement locations are superimposed on the expected regret (i.e. the difference between the VoI provided by all potential sensors, or $\text{VoI}(Y_{all})$, and that provided by the selected set of sensors, $\text{VoI}(Y)$) for a particular date and time, spatially distributed over the city. The regret plotted in Fig. 8 illustrates how information collected at just a few locations within the city (less than one tenth of the possible measurement locations) allows the predicted temperature field to be updated throughout the area, and for these updated predictions to be used to improve decision-making to reduce losses even in places where no direct measurements are made.

### 4.4. Comparison of metrics

Table 2 compares the performance of the sensor sets selected using the three metrics considered. The set optimized under the conditional entropy metric, depicted in Fig. 6, is denoted $Y^*_{\text{Ent}}$. The set optimized under the weighted prediction error metric, depicted in Fig. 7, is denoted $Y^*_{\text{Err}}$. The set optimized under the VoI metric, depicted in Fig. 8, is denoted $Y^*_{\text{VoI}}$. The performance of each set is assessed using the three different metrics. Performances are listed as percentages of the value for each metric which would be achieved by measuring $Y_{\text{all}}$. A higher percentage corresponds to a better sensor placement under all metrics in this table. Note that sets $Y^*_{\text{Ent}}$, $Y^*_{\text{Err}}$, and $Y^*_{\text{VoI}}$ all contain 9 sensing locations, while $Y_{\text{all}}$ contains 122.

**Table 2**
Comparative results for sensor placement metrics.

| Sensor Set | Percent of Metric Value for $Y_{\text{all}}$ | | |
|---|---|---|---|
| | $\text{H}(T\|Y)$ | $\text{Err}(\hat{T}\|Y)$ | $\text{VoI}(Y)$ |
| $Y^*_{\text{Ent}}$ | 11.3% | 68% | 23% |
| $Y^*_{\text{Err}}$ | 10.8% | 72% | 35% |
| $Y^*_{\text{VoI}}$ | 9.7% | 57% | 42% |

The set optimized under each metric is the best set when evaluated by that metric (e.g. $Y^*_{\text{Ent}}$ is the best set as evaluated by the conditional entropy metric); this result is to be expected, and indicates how the greedy optimization approach used, although approximate, is still effective. Comparing across metrics, the conditional entropy metric does not vary greatly across sets, with $Y^*_{\text{VoI}}$ performing only slightly worse than $Y^*_{\text{Ent}}$. By contrast, under the VoI metric, $Y^*_{\text{Ent}}$ achieves only about half the value of $Y^*_{\text{VoI}}$. The set $Y^*_{\text{Err}}$ always performs either best or second-best, while the set $Y^*_{\text{VoI}}$ performs worst except under the VoI metric. This suggests that the weighted prediction error may be a good proxy for the other metrics for performing optimization, but that the VoI metric tends to select sensor placements that are not favored by the other metrics.

Comparing sensor placements qualitatively, those selected under the conditional entropy metric seem to be evenly dispersed across the region, as well as along its perimeter. Placements selected under the weighted predication error metric are similarly distributed, but with fewer measurements along the perimeter. Placements selected under the VoI metric are concentrated in the downtown area, as well as south and southeast of there. Comparing these results with the risk index map of Fig. 4, there is no clear indication that high-risk or low-risk areas are particularly focused on by any metric. For example, while the VoI places some measures in the high-risk downtown area, similarly high-risk areas to the north are ignored, while many sensors are allocated to more moderate-risk areas in the south.

This comparison supports the assertion in 3 that heat risk alone is not sufficient to identify areas of interest for temperature monitoring. Some combination of the temperature pattern, prior uncertainty, and

vulnerability should be considered. Of the metrics considered here, the VoI takes all of these factors into account through a decision-making problem taking into account temperature thresholds and vulnerability. The conditional entropy metric does not take local vulnerability into account, and the weighted prediction error metric seeks to minimize uncertainty in risk, even in areas where the magnitude of the risk may be low. Therefore, we assert that the VoI metric should be preferred, as it accounts for relevant decision-making factors that the other metrics do not. However, these other metrics may still be of use in cases when the decision-making problem being supported is not well-defined, when the inputs needed for evaluating the losses are too uncertain, or when the goal of the sensing effort is to reduce temperature uncertainty or prediction errors only, without consideration for heat response decision-making implications. Furthermore, the computational costs associated with evaluating conditional entropy and weighted prediction error are lower than those related to evaluating VoI.

For temperature monitoring in the city of Pittsburgh, this case study suggests two alternative approaches to temperature monitoring. First, sensing efforts can be distributed across the city; this will best support accurate prediction of temperature throughout, even when taking into account differing vulnerabilities between neighborhoods. Second, monitoring efforts can be focused in central and southern Pittsburgh, as these will best support heat wave advisory issuance under the current decision-making model, depending on the priorities of city or local government, and/or the National Weather Service.

## 5. Conclusion

In this paper, we optimize the locations of temperature measurements for updating the Gaussian temperature field model of [16]. This allows sensing resources to be used optimally to support decision-making for the response to and mitigation of extreme heat events in urban areas. The Gaussian temperature model decomposes the temperature field into consistent time- and space-varying components ($T_0$ and $T_1$) and an uncertain perturbation component ($T''$), and provides a statistical framework for optimizing sensor placement that is not feasible using temperature fields obtained by numerical models alone.

We present a case study application for Pittsburgh, Pennsylvania, comparing three potential sensing metrics. Of these, only VoI captures information about uncertainty, vulnerability, and decision-making in guiding sensor placement. Although the relationship between extreme temperatures and their impacts is still incompletely understood, the VoI has the potential to become the key metric for supporting decision-making for urban heat risk response.

The methods presented in this paper can support targeted heat wave advisory issuance. Location- and time-specific warnings (e.g. “any outdoor activities in Schenley Park should be rescheduled until after 4 P M due to extreme heat”) might overcome the lack of understanding regarding response to city-wide heat advisories in Pittsburgh identified by Ref. [27]. More generally, these methods support an improved understanding of urban heat risk, including its uncertainty, which is important for robust urban decision-making regarding climate change [45].

In order to extend this methodology to other urban areas, two main components are needed. First, the probabilistic temperature model must be created and calibrated, which requires fine-resolution temperature data obtained through numerical simulation using WRF-PUCM or a comparable modeling approach. Second, an assessment of the vulnerability is needed. This can be conducted using the same vulnerability index used here, or a different methodology can be used to assess local vulnerability to extreme temperature based on the specific characteristics of the urban area in question. With these two inputs available, the methodologies presented in this paper should be directly applicable to any urban area.

Several future improvements to the methods presented here are possible. First, temporally varying vulnerability models, based on the movements of populations throughout the day and their activities in response to extreme temperatures, might be incorporated. These models, together with the spatio-temporal hazard model, would create a full spatio-temporal model for the heat risk, allowing for more robust analysis of the problem of urban heat monitoring and mitigation across time as well as space. Second, improved models of sensor noise can be used. These models would account not only for the intrinsic uncertainty in temperature measurement but also for the fact the measurements at specific locations may be more or less representative of the average temperature across the surrounding local area (e.g. Ref. [46]). Third, the proliferation of low-cost temperature sensors may allow for the incorporation of data obtained from mobile sensing platforms and/or from “crowd-sourced” temperature data contributed by citizen volunteers (e.g. Ref. [44]). Finally, the potential of the modeling and sensing optimization approaches discussed in this paper to support long-term extreme temperature mitigation decisions, such as investments in cool and/or green roofs or green infrastructure investments, remains to be investigated.

## Acknowledgments

Funding: This work was supported by the National Science Foundation [Grant No. ICER 1664091, Sustainability Research Network Cooperative Agreement No. 1444758]; and the 2016 Dowd Fellowship from the College of Engineering at Carnegie Mellon University. The authors would like to thank Philip and Marsha Dowd for their financial support and encouragement. The authors would like to thank S. Harris for contribution of the project title and acronym (SHADE).

## Appendix A. Nomenclature

The following table lists some of the common notations and symbols used in this paper.

Table A.1
Standard nomenclature used in this paper

| Term | Definition | Units |
|---|---|---|
| $t$ | time | [h] |
| $\mathbf{x}$ | spatial location or coordinate | [km] |
| $T$ | surface temperature | [°C] |
| $T_0$ | global average temperature | [°C] |
| $T_1$ | cyclic temperature pattern | [°C] |
| $T''$ | residual temperature | [°C] |
| $\hat{T}$ | temperature predictions | [°C] |
| $\mathbf{T}$ | vector of temperature values | [°C] |

(continued on next page)

Table A.1 (*continued*)

| Term | Definition | Units |
|---|---|---|
| $\mathbf{y}$ | vector of measurements | [°C] |
| $\mathcal{N}$ | normal or Gaussian distribution | |
| M | mean function | |
| K | covariance function | |
| $\mathbb{E}$ | statistical expectation (with respect to subscripted variable) | |
| $\omega$ | weight parameter | [dimensionless] |
| $V$ | vulnerability index | [dimensionless] |
| $R$ | risk index | [dimensionless] |
| $\tau$ | temporal domain | |
| $X$ | spatial domain | |
| H(·) | entropy | [dimensionless] |
| Err(·) | expected prediction error | [°C] |
| VoI(·) | value of information | [dimensionless] |
| L(·) | loss function | [dimensionless] |
| C(·) | cost function | [dimensionless] |
| $Y$ | set of measurements, with subscript denoting the metric used for optimization | |
| $\mathcal{GP}$ | Gaussian process model | |
| $\boldsymbol{\mu}$ | mean vector | |
| $\mathbf{K}$ | covariance matrix | |
| $\mathbb{I}$ | indicator function (takes value 1 if argument is true, 0 otherwise) | |
| $\boldsymbol{\Omega}$ | matrix of weight parameters | |
| $T_{95\%}$ | 95th percentile temperature | [°C] |
| $R_{95\%}$ | 95th percentile risk index | [dimensionless] |
| $\tau$ | temporal domain | |

## Appendix B. Sensor set size selection

The plots below show the benefit derived from greedily chosen sensor sets of various sizes under the metrics in question. For the VoI metric, this benefit is directly traded off against sensing cost to determine the optimal number of sensors. The same number of sensors is then used across all metrics to allow for comparison.

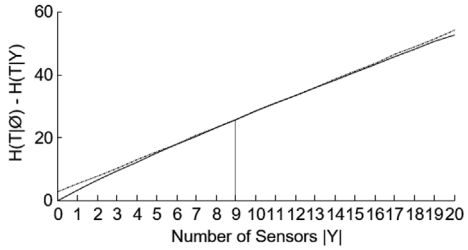


**Fig. B.1.** Entropy reduction resulting from greedy sensor placement (solid line). The tangent line to this curve at 9 sensors is also indicated (dashed line).

Conditional entropy reduction is the difference between the marginal entropy $\mathrm{H}(T|\varnothing)$ and conditional entropy $\mathrm{H}(T|Y)$, where $\varnothing$ denotes the empty set of measurements. Note that this reduction is nearly linear in the number of sensors, as shown in Fig. B.1. Note that this metric is a logarithmic function of the probabilistic volume spanned by the posterior covariance, quantified via the determinant of $\mathbf{K}_{T|Y}$. Because this volume is reduced at an almost exponential rate as more sensors are added, the metric reduction is nearly linear. The tangent line indicates that the ninth sensor placed yields 82% of the entropy reduction of the first.

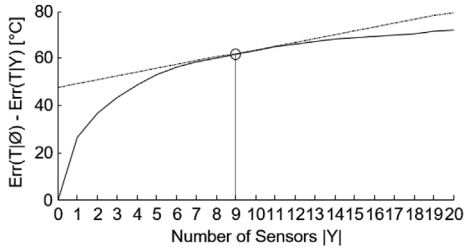


**Fig. B.2.** Weighted prediction error reduction resulting from greedy sensor placement (solid line). The tangent line at 9 sensors is also indicated (dashed line).

Weighted prediction error reduction is measured as the difference between the marginal weighted prediction error $\mathrm{Err}(\hat{T}|\varnothing)$ and conditional weighted prediction error $\mathrm{Err}(\hat{T}|Y)$. Fig. B.2 shows a strong diminishing returns property, i.e. the slope decreases quickly as the number of sensors increases, with the ninth sensor providing only 6% of the reduction of the first. Note the contrast with the weak property exhibited in Fig. B.1 due to differences in the definitions of these metrics.

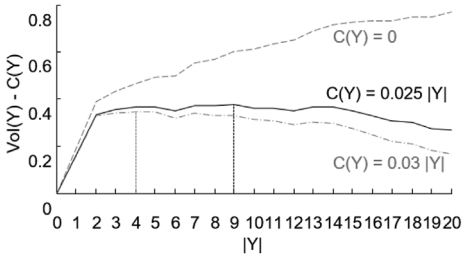


**Fig. B.3.** Net VoI resulting from greedy sensor placements, which is maximized when $|Y| = 9$ for a cost of 0.025 per sensor. This is the optimal number of sensors indicated in Fig. 8. The net VoI for other costs are also shown.

Finally, the net VoI, $\mathrm{VoI}(Y) - \mathrm{C}(Y)$, is plotted directly in Fig. B.3, and the peak of this curve identifies the optimal number of sensors as nine. Note that for the given cost per sensor (0.025) any number of sensors between 2 and 14 would provide roughly the same net VoI, and so the optimal number of sensors is sensitive to cost in this case. For example, a cost of 0.03 per sensor indicates 4 sensors are optimal. However, VoI remains strictly increasing with the number of sensors.

## Appendix C. Additional details of sensor placement metric computation

*Conditional entropy*

For Gaussian process models, conditional entropy can be efficiently computed from the posterior covariance $\mathbf{K}_{T|Y}$ of Eq. (7). Efficient and provably near-optimal sensor placement in Gaussian process models using the conditional entropy metric and greedy optimization is discussed in detail by Ref. [43]; the same general method is used in this paper.

*Weighted prediction error*

In the specific case of a Gaussian process model, the optimal prediction $\hat{T}(\mathbf{x}, t)$ with respect to the L-2 norm is the mean of the model, i.e. $\boldsymbol{\mu}_{T|\mathbf{y}}$ in the posterior case. The expected square error of a prediction at a given space-time coordinate is the posterior variance of the temperature field at that coordinate. Therefore, the prediction error can be expressed as:

$$\mathrm{Err}(\hat{T}|Y) = \min_{\hat{T}} \mathrm{Err}(\hat{T}) = \mathrm{tr}(\boldsymbol{\Omega}\mathbf{K}_{T|Y}\boldsymbol{\Omega}^{\mathrm{T}}) \tag{C.1}$$

where $\mathrm{tr}(\cdot)$ denotes the matrix trace and $\boldsymbol{\Omega}$ is a diagonal weight matrix, incorporating square roots of the coefficients $\omega_t$ and $\omega_{\mathbf{x}}$. Thus, as with conditional entropy, this metric is a function of the posterior covariance of the Gaussian model only, and therefore can also be efficiently evaluated. In this paper, optimal sensor placement via this metric is performed similarly to that for optimal placement via conditional entropy.

*Value of information*

Taking the expectation over (potentially multivariate) sensor observations and field states and minimizing over a potentially large set of possible actions is computationally intensive, making VoI an inefficient metric for sensor placement in general. However, under certain assumptions on the problem structure, VoI can be efficiently computed. One such assumption is the cumulative assumption, which states that the loss function can be expressed as the sum of local loss functions that depend only on subsets of the inputs. In the case of urban temperatures, we assume that the city-wide loss function can be expressed as the sum of local loss functions:

$$\mathrm{L}(T, A) = \sum_{t\in\tau}\sum_{\mathbf{x}\in X} \mathrm{L}_{\mathbf{x},t}(T(\mathbf{x}, t), A(\mathbf{x}, t)) \tag{C.2}$$

where each local loss is a function only of the temperature and the selected action at the corresponding location and time. The VoI can then be expressed as:

$$\mathrm{VoI}(Y) = \sum_{t\in\tau}\sum_{\mathbf{x}\in X} \mathrm{VoI}_{\mathbf{x},t}(Y) \tag{C.3}$$

where $\mathrm{VoI}_{\mathbf{x},t}(Y)$ is computed as in Eq. (13), using the corresponding local loss function. Efficient computation of the VoI under this cumulative assumption in Gaussian models is discussed by Ref. [36] for spatial random fields and by Ref. [44] for spatio-temporal random fields. In general, losses for future events may be discounted; a discounting factor can easily be included in the above formulation for this purpose.

In this paper, we use a decision-making problem motivated by heat advisory issuance, and we use the vulnerability to define the loss function in this problem. Consider that a cost $C_f(\mathbf{x})$ is incurred in an area if the temperature in that area exceeds a threshold $T_{limit}$ (with $T_{limit} = 30°C$ in this paper). This cost captures the consequences of the area's population being exposed to extreme temperatures. A heat advisory may or may not be issued for this area at this time. This decision is indicated by choosing $A(\mathbf{x}, t) \in \{0,1\}$, where a choice of "1" corresponds with the issuance of a heat advisory, and a choice of "0" corresponds with no warnings being issued. By issuing a warning, people in the region will be encouraged to seek shelter, and so we assume that the consequences of exposure will be avoided. The cost for issuing this advisory, $C_r(\mathbf{x})$, quantifies the effort necessary to issue the warning, as well as the loss of economic productivity in the area where the advisory is issued due to restrictions on outdoor activity. Under this decision-making problem, the local loss function of Eq. (C.2) is expressed as:

$$\mathrm{L}_{\mathbf{x},t}(T(\mathbf{x}, t), A(\mathbf{x}, t)) = C_f(\mathbf{x})\,\mathbb{I}[T(\mathbf{x}, t)(1 - A(\mathbf{x}, t)) > T_{limit}] + C_r(\mathbf{x})\,A(\mathbf{x}, t) \tag{C.4}$$

where $\mathbb{I}[\cdot]$ is the indicator function, taking on value 1 when its argument is true and 0 otherwise. Under this definition, the VoI represents the average

hourly reduction in expected loss in the decision-making problem over the city.

Ideally, both $C_f(\mathbf{x})$ and $C_r(\mathbf{x})$ can be quantified in monetary values, e.g. by converting mortality, morbidity, and productivity loss to common monetary quantities using accepted practices. However, such a detailed quantification is beyond the scope of this work. As discussed by Ref. [12], the heat vulnerability index is not a measure of the consequences of a heat wave event. It is, rather, one input to a utility function (or loss function) for decision-making in response to extreme temperatures. The relationship between the societal factors captured in this index and the impacts of extreme temperatures, e.g. in terms of hospitalizations due to heat stress, is an open area of research [4,12,15]. Instead, we assume for illustrative purposes that consequences are equal to the vulnerability index of an area, and so $C_f(\mathbf{x}) = V(\mathbf{x})$. On this vulnerability scale, the consequence of issuing a heat advisory is chosen as $C_r(\mathbf{x}) = 3 \ \forall \ \mathbf{x} \in X$: under these assumptions, for some sparsely populated or low-vulnerability areas of the city, the consequence of being exposed to extreme heat is lower than the consequence of issuing an advisory, and so it will never be beneficial to do so. However, in areas with larger and/or more vulnerable populations, an advisory may be issued to avoid the higher consequences of extreme heat exposure.

In examining Eq. (13), note that $Y$ represents a set of observations (e.g. a plan to measure temperature at specific locations and times), while $\mathbf{y}$ indicates a specific set of measurements related to these observations (e.g. the temperature measurement data gathered as a result of carrying out plan $Y$). In the posterior case, because additional information is available to improve temperature estimates, a better set of actions can be chosen, reducing the losses incurred by the managing agent. Since the possible values $\mathbf{y}$ of the observations $Y$ are themselves random quantities, an expectation must also be taken over these values. VoI thus quantifies the utility of measurements based on how much they can be expected to improve decision-making outcomes by reducing losses. Note also that the prior expected loss is fixed with respect to the set $Y$, and so maximizing $\mathrm{VoI}(Y)$ is equivalent to minimizing the posterior expected loss.

It is sometimes useful, as in the case of Fig. 8, to utilize regret rather than VoI. Regret is a complementary concept to the VoI, and assesses how much additional value might still be gained by including more measurements on the system. Specifically, it is defined as:

$$\mathrm{Regret}(Y) = \mathrm{VoI}(Y_{\mathrm{all}}) - \mathrm{VoI}(Y) \tag{C.5}$$

where $Y_{\mathrm{all}}$ denotes the set of all measurements which might be made.

In the example of Section 4.3, the VoI and regret are computed using the temperature data generated to calibrate this probabilistic model [16]. Some of these data, corresponding to the locations selected for temperature measurement, are used to update the temperature field prediction at this time. Therefore, note that these measurements correspond to simulation data rather than actual field temperature measurements. Thus, while for the previous optimization metrics demonstrated, only the locations of measurements (and not their values) were needed to compute the metrics, the computation of VoI or regret requires measurement data (or a range of potential measurement data acquired from simulations).

Using the updated field prediction, choices are made on whether or not to issue heat advisories for different areas based on the predicted temperatures, the remaining uncertainty in these predictions, and the relative consequences of issuing versus failing to use the heat advisory. The resulting penalties are then calculated based on the decisions made and the city-wide temperature field data at this time. VoI represents the difference between penalties incurred without and with the updated predictions made using a subset of these data corresponding to the selected measurement locations. The regret, on the other hand, represents the difference between the VoI and minimum possible penalties which would have been incurred if data from all measurement locations in $Y_{all}$ were used to update the predictions (i.e., the amount of additional penalty reductions which would be possible if more sensors were included).

## References


[1] J.E. Clougherty, L.D. Kubzansky, A framework for examining social stress and susceptibility to air pollution in respiratory health, Environ. Health Perspect. 117 (2009) 1351–1358, http://dx.doi.org/10.1289/ehp.0900612.

[2] B. Doyon, D. Bélanger, P. Gosselin, The potential impact of climate change on annual and seasonal mortality for three cities in Québec, Canada, Int. J. Health Geogr. 7 (2008) 23, http://dx.doi.org/10.1186/1476-072X-7-23.

[3] F.C. Curriero, K.S. Heiner, J.M. Samet, S.L. Zeger, L. Strug, J.A. Patz, Temperature and mortality in 11 cities of the Eastern United States, Am. J. Epidemiol. 155 (2002) 80–87, http://dx.doi.org/10.1093/aje/155.1.80.

[4] D.M. Hondula, A.G. Barnett, Heat-related morbidity in Brisbane, Australia: spatial variation and area-level predictors, Environ. Health Perspect. (2014), http://dx.doi.org/10.1289/ehp.1307496.

[5] E. Brunner, Commentary: education, education, education, Int. J. Epidemiol. 30 (2001) 1126–1128, http://dx.doi.org/10.1093/ije/30.5.1126.

[6] S. Hajat, R.S. Kovats, K. Lachowycz, Heat-related and cold-related deaths in England and Wales: who is at risk? Occup. Environ. Med. 64 (2006) 93–100, http://dx.doi.org/10.1136/oem.2006.029017.

[7] F. Bosello, R. Roson, R.S.J. Tol, Economy-wide estimates of the implications of climate change: sea level rise, Environ. Resour. Econ. 37 (2007) 549–571, http://dx.doi.org/10.1007/s10640-006-9048-5.

[8] M. Stafoggia, F. Forastiere, D. Agostini, N. Caranci, F. de'Donato, M. Demaria, P. Michelozzi, R. Miglio, M. Rognoni, A. Russo, C.A. Perucci, Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis, J. Epidemiol. Community Health 62 (2008) 209–215, http://dx.doi.org/10.1136/jech.2007.060715.

[9] J.K. Rosenthal, P.L. Kinney, K.B. Metzger, Intra-urban vulnerability to heat-related mortality in New York City, 1997–2006, Health Place 30 (2014) 45–60, http://dx.doi.org/10.1016/j.healthplace.2014.07.014.

[10] C. Reid, M. O'Neill, C. Gronlund, S. Brines, D. Brown, A. Diez-Roux, J. Schwartz, Mapping community determinants of heat vulnerability, Environ. Health Perspect. (2009), http://dx.doi.org/10.1289/ehp.0900683.

[11] S.L. Harlan, J.H. Declet-Barreto, W.L. Stefanov, D.B. Petitti, Neighborhood effects on heat deaths: social and environmental predictors of vulnerability in Maricopa County, Arizona, Environ. Health Perspect. 121 (2012) 197–204, http://dx.doi.org/10.1289/ehp.1104625.

[12] K. Bradford, L. Abrahams, M. Hegglin, K. Klima, A heat vulnerability index and adaptation solutions for Pittsburgh, Pennsylvania, Environ. Sci. Technol. 49 (2015) 11303–11311, http://dx.doi.org/10.1021/acs.est.5b03127.

[13] S. Grower, C. Mee, M. Campbell, Protecting Vulnerable People from Health Impacts of Extreme Heat, Toronto Public Health, Toronto, Ontario, Canada, 2011http://www.climateontario.ca/doc/ORAC_Products/TPH/ProtectingVulnerablePeoplefromHealthImpactsofExtremeHeat.pdf.

[14] C. Buscail, E. Upegui, J.-F. Viel, Mapping heatwave health risk at the community level for public health action, Int. J. Health Geogr. 11 (2012) 38, http://dx.doi.org/10.1186/1476-072X-11-38.

[15] C. Aubrecht, D. Özceylan, Identification of heat risk patterns in the U.S. National Capital Region by integrating heat stress and related vulnerability, Environ. Int. 56 (2013) 65–77, http://dx.doi.org/10.1016/j.envint.2013.03.005.

[16] C. Malings, M. Pozzi, K. Klima, E. Bou-Zeid, P. Ramamurthy, M. Bergés, Surface heat assessment for developed environments: probabilistic urban temperature modeling, Comput. Environ. Urban Syst. 66 (2017) 53–64.

[17] Z. Du, P. Xu, X. Jin, Q. Liu, Temperature sensor placement optimization for VAV control using CFD–BES co-simulation strategy, Build. Environ. 85 (2015) 104–113, http://dx.doi.org/10.1016/j.buildenv.2014.11.033.

[18] A. Barbaresi, D. Torreggiani, S. Benni, P. Tassinari, Indoor air temperature monitoring: a method lending support to management and design tested on a wine-aging room, Build. Environ. 86 (2015) 203–210, http://dx.doi.org/10.1016/j.buildenv.2015.01.005.

[19] A.D. Fontanini, U. Vaidya, B. Ganapathysubramanian, A methodology for optimal placement of sensors in enclosed environments: a dynamical systems approach, Build. Environ. 100 (2016) 145–161, http://dx.doi.org/10.1016/j.buildenv.2016.02.003.

[20] J. Li, H. Li, Y. Ma, Y. Wang, A.A. Abokifa, C. Lu, P. Biswas, Spatiotemporal distribution of indoor particulate matter concentration with a low-cost sensor network, Build. Environ. 127 (2018) 138–147, http://dx.doi.org/10.1016/j.buildenv.2017.11.001.

[21] T. Lovett, J. Lee, E. Gabe-Thomas, S. Natarajan, M. Brown, J. Padget, D. Coley, Designing sensor sets for capturing energy events in buildings, Build. Environ. 110 (2016) 11–22, http://dx.doi.org/10.1016/j.buildenv.2016.09.004.

[22] P.A. Mirzaei, F. Haghighat, Approaches to study Urban Heat Island – abilities and limitations, Build. Environ. 45 (2010) 2192–2201, http://dx.doi.org/10.1016/j.buildenv.2010.04.001.

[23] C. Kisner, K. Mulder, B. VanGessel, Assessing Heat Vulnerability and Access to Cooling Centers in Detroit, Michigan, University of Michigan, Ann Arbor, MI, 2012http://www.imagin.org/awards/sppc/2013/2013_sppc_paper_kisner_mulder_

vangessel.pdf.
[24] D. Li, E. Bou-Zeid, M. Oppenheimer, The effectiveness of cool and green roofs as urban heat island mitigation strategies, Environ. Res. Lett. 9 (2014) 055002, http://dx.doi.org/10.1088/1748-9326/9/5/055002.
[25] Y.K. Yang, I.S. Kang, M.H. Chung, S. Kim, J.C. Park, Effect of PCM cool roof system on the reduction in urban heat island phenomenon, Build. Environ. 122 (2017) 411–421, http://dx.doi.org/10.1016/j.buildenv.2017.06.015.
[26] T.E. Morakinyo, K.K.-L. Lau, C. Ren, E. Ng, Performance of Hong Kong's common trees species for outdoor temperature regulation, thermal comfort and energy saving, Build. Environ. 137 (2018) 157–170, http://dx.doi.org/10.1016/j.buildenv.2018.04.012.
[27] E. Wells, K. Klima, Excessive heat event risk & response: how the public and public health experts prepare, respond, and interpret extreme heat, Cities Environ. (2016) (in review).
[28] W.C. Skamarock, J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, W. Wang, J.G. Powers, A Description of the Advanced Research WRF Version 2, National Center for Atmospheric Research, 2005.
[29] Z.-H. Wang, E. Bou-Zeid, J.A. Smith, A coupled energy transport and hydrological model for urban canopies evaluated using a wireless sensor network, Q. J. R. Meteorol. Soc. 139 (2013) 1643–1657, http://dx.doi.org/10.1002/qj.2032.
[30] C. Currin, T. Mitchell, M. Morris, D. Ylvisaker, Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments, J. Am. Stat. Assoc. 86 (1991) 953–963.
[31] A. Krause, Optimizing Sensing: Theory and Applications, Carnegie Mellon University, 2008.
[32] T. Cover, J. Thomas, Elements of Information Theory, second ed., John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2006.
[33] A. Krause, B. McMahan, C. Guestrin, A. Gupta, Robust submodular observation selection, J. Mach. Learn. Res. 9 (2008) 2761–2801.
[34] C. Malings, M. Pozzi, Value of information for spatially distributed systems: application to sensor placement, Reliab. Eng. Syst. Saf. 154 (2016) 219–233, http://dx.doi.org/10.1016/j.ress.2016.05.010.
[35] H. Raiffa, R. Schlaifer, Applied Statistical Decision Theory, Harvard University Press, Cambridge, Massachusetts, USA, 1961.
[36] C. Malings, M. Pozzi, Conditional entropy and value of information metrics for optimal sensing in infrastructure systems, Struct. Saf. 60 (2016) 77–90.
[37] C. Malings, M. Pozzi, Value-of-information in spatio-temporal systems: sensor placement and scheduling, Reliab. Eng. Syst. Saf. 172 (2018) 45–57, http://dx.doi.org/10.1016/j.ress.2017.11.019.
[38] N.A.C. Cressie, C.K. Wikle, Statistics for Spatio-temporal Data, Wiley, Hoboken, N.J, 2011.
[39] D. Straub, Reliability updating with inspection and monitoring data in deteriorating reinforced concrete slabs, Proc. ICASP11, 2011.
[40] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning, MIT Press, Cambridge, Mass, 2006.
[41] L. Jiang, W. Zhan, J. Voogt, L. Zhao, L. Gao, F. Huang, Z. Cai, W. Ju, Remote estimation of complete urban surface temperature using only directional radiometric temperatures, Build. Environ. 135 (2018) 224–236, http://dx.doi.org/10.1016/j.buildenv.2018.03.005.
[42] P. Höppe, The physiological equivalent temperature - a universal index for the biometeorological assessment of the thermal environment, Int. J. Biometeorol. 43 (1999) 71–75.
[43] A. Krause, A. Singh, C. Guestrin, Near-optimal sensor placements in Gaussian processes: theory, efficient algorithms and empirical studies, J. Mach. Learn. Res. 9 (2008) 235–284.
[44] C. Malings, M. Pozzi, Optimal Sensing Using Value of Information in Spatio-temporal Random fields, Pittsburgh, PA, USA (2016).
[45] F. Hoss, K. Klima, P. Fischbeck, Ten strategies to systematically exploit all options to cope with anthropogenic climate change, Environ. Syst. Decis 34 (2014) 578–590, http://dx.doi.org/10.1007/s10669-014-9517-2.
[46] C.L. Muller, L. Chapman, C.S.B. Grimmond, D.T. Young, X. Cai, Sensors and the city: a review of urban meteorological networks, Int. J. Climatol. 33 (2013) 1585–1600, http://dx.doi.org/10.1002/joc.3678.