Extremum Tracking in Sensor Fields with Spatio-temporal Correlation

Abhishek Nadamani[1], Prithwish Basu[2] and Lang Tong[1]
[1]School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853
[2]BBN Technologies, Cambridge, MA 02138

Abstract

Physical phenomena such as temperature, humidity, and wind velocity often exhibit both spatial and temporal correlation. We consider the problem of tracking the extremum value of a spatio-temporally correlated field using a wireless sensor network. If the extremum is determined at a particular instant of time, all sensor nodes could transmit their measurements to the fusion center which, then evaluates the extremum. This is not energy-efficient because the spatio-temporal correlation of the field is not exploited. We present an optimal centralized algorithm that utilizes the aforementioned correlation to not only minimize the number of transmitting sensors but also ensure low tracking error with respect to the actual extremum. We use recent order statistics bounds in the formulation of the cost function. Since the centralized algorithm has high time complexity, we propose a suboptimal distributed algorithm based on a similar cost function. Our simulations indicate that a small fraction of sensors is often sufficient to track the extremum. Simulations indicate that the centralized algorithm can achieve about 70% energy savings with almost perfect tracking. Furthermore, the performance of the distributed algorithm is comparable to that of the centralized algorithm with up to 25% more energy expenditure.

1 Introduction

Recent advances in embedded sensing and wireless communications and networking technologies has resulted in the proliferation of low power sensor devices that are capable of sharing sensed information with each other over a wireless medium and thus forming a wireless sensor network. A common application of wireless sensor networks is in the domain of environmental monitoring; in particular, they are useful for sensing and tracking variations in physical phenomena such as temperature, pressure, humidity, wind velocity etc. over a geographical area. Wireless sensor nodes usually have low power RF transceivers that regularly transmit sensed data either directly or over multiple hops to the fusion center which supports user queries on gathered data samples.

A major concern that plagues sensor networks researchers and users alike is the limited battery life on sensor nodes since RF communication and also idle listening have a significant drain on battery. Since sensors can rarely be recharged once deployed in a remote environment, extending battery life is of paramount importance. Therefore, it is prudent to develop energy efficient mechanisms for using these sensors and thus increase the lifetime of the sensor network. In this paper we consider the problem of accurately monitoring a spatio-temporally correlated physical phenomenon in an energy-efficient manner. We assume that there are a finite number of static sensor nodes that can sense the phenomenon at their locations – we refer to this discretized sampling of a continuous phenomenon as a *sensor field*. Specifically, we are interested in tracking the *maximum* value in the sensor field over time.

In this paper, we focus on the single hop network scenario where all sensors and the fusion center are within transmission range of each other. This is interesting in monitoring applications that involve sensors spread across a geographical area such as inside or on the rooftop of a building. Commercial sensor networks such as Zigbee that target building, home or industrial automation often organize nodes into single hop "star" topologies [7]. In those scenario, most sensors are not multi-hop routers; instead they listen to the basestation (called the PAN coordinator) for data and transmission schedule information, and transmit sensed data to the latter. Typically, the coordinator node (in our case, the fusion center) has much greater energy reserves than the sensor nodes. Applications that require sensors to be spread across a large geographical area and require multi-hopping are not a focus of this paper. Extremum tracking in multi-hop sensor networks is a topic of active research for us.

If the maximum is to be determined at a particular instant of time, then all sensors could transmit their measurements to the fusion center which would then determine the maximum value of the sensor field. However, if there are spatio-temporal correlations in the sensed data, we show that it is possible to exploit such correlations intelligently for the purpose of saving energy while still accurately tracking the maximum value.

1.1 Main Contributions

We first propose a centralized algorithm that is executed at the fusion center. It determines *which* sensor nodes should be transmitting their sensed measurements in the next epoch. In this algorithm, the fusion center maintains recent history of readings from various sensors and attempts to minimize an objective function that captures both energy consumption and the expected value of *deviation from maximum*, if only a subset of nodes were transmitting. Minimization of this objective function yields low tracking error even if only a small fraction of sensor nodes transmit their measurements.

We also propose a distributed algorithm in which the transceivers of sensor nodes are in either ON or OFF state. When in ON state, they overhear measurements reported by a subset of other transmitting nodes and then make a decision about whether to transmit their measurements in the next epoch or to sleep. In OFF state, the nodes do not sense, transmit, or receive. Every node independently attempts to minimize an objective function that is similar to the one used in the centralized algorithm.

To summarize, the main contributions of this paper are:

1. We give a novel analytical formulation of the problem of energy-efficient extremum tracking while exploiting the inherent spatio-temporal correlation in the sensor field. The spatio-temporal correlation structure is computed for the data collected from a weather sensor testbed in our institution.
2. In the formulation of the cost function, we use recent order statistics bounds to compute the expected value of the maximum of a given number of random variables.
3. We derive an optimal policy for the selection of sensors that should transmit in every epoch while minimizing energy consumption as well as tracking error. We use this policy in a centralized algorithm.
4. We propose a suboptimal distributed algorithm where each sensor determines whether to transmit or sleep by optimizing a local cost function.

1.2 Related Work

Boulis et al. propose a heuristic mechanism for performing energy-efficient aggregation in sensor networks [4]. One of the applications of their algorithm is to determine the maximum in a sensor field. Unlike our proposed technique, their scheme does not take into account the spatio-temporal correlations of the phenomenon.

Fuhrmann and Widmer address the problem of determining the maximum (or minimum) in a network with large multicast groups [14, 6]. Their objective is to minimize the number of transmissions for determining the maximum using multicast feedback. The data measured at nodes is assumed to be independent, and the algorithm terminates after it has computed the maximum value. On the other hand, our proposed approach exploits the spatio-temporal correlation in the sensor field and continuously tracks the extremum.

Many researchers have investigated exploiting spatial correlation for wireless sensor network applications [13, 10]. Borrowing terminology of [4], these applications correspond to "snapshot aggregation" and do not take into account the temporal characteristics of the sensed data.

Akyildiz et al. emphasize the importance of spatial and temporal correlation in designing MAC protocols for wireless sensor networks [1]. But they consider the spatial and temporal correlation independently and not together. Moreover, their analysis on how spatial and temporal correlation can be used for MAC protocol design is mostly qualitative.

Like our proposed schemes, Mergen et. al. [9] have also considered the case where the sensors communicate directly to the fusion center. Their analysis shows that the cost of listening can be a dominant factor in a dense network. They have proposed wake up schemes and have determined the capacity of such systems from an information theoretic perspective.

The rest of the paper is organized as follows. Section 2 describes the data collection experiments that we conducted on our weather sensor testbed. The data that was collected exhibits different degrees of correlation for a variety of phenomena. We use these correlation values to drive our simulations that are presented later in this paper. In Section 3 we introduce the optimization problem. An optimal centralized algorithm and a distributed algorithm are presented in Sections 4 and 5, respectively. Section 6 concludes the paper and presents some directions of future research.

2 Spatio-temporal Correlation in Sensor Data

In order to verify our intuition about spatio-temporal correlation in physical phenomena, we gathered weather data from a sensor network testbed recently deployed on the rooftops of various buildings in our institution[1]. Each node consists of an embedded PC, a 802.11a/b/g interface, and various sensors for monitoring weather conditions as well as air pollutants. The sensors can detect weather measurements, such as wind speed and direction, temperature, air pressure, relative humidity, and rainfall.

We collected pressure, temperature, humidity, wind speed and direction data for two sensors in the testbed over a period of 12 hours. The pressure, temperature and humidity measurements are spaced a minute apart. The wind speed and direction measurements are spaced 30 seconds apart. The two sensor nodes are located approximately 85 meters apart. These sensors currently transmit the monitored data to the fusion center periodically. Since they are connected to AC power outlets, these sensors have abundant energy resources, and energy efficiency is currently not an issue in the testbed.

[1]The details about the location of the testbed have been suppressed to meet the double-blind review requirement.

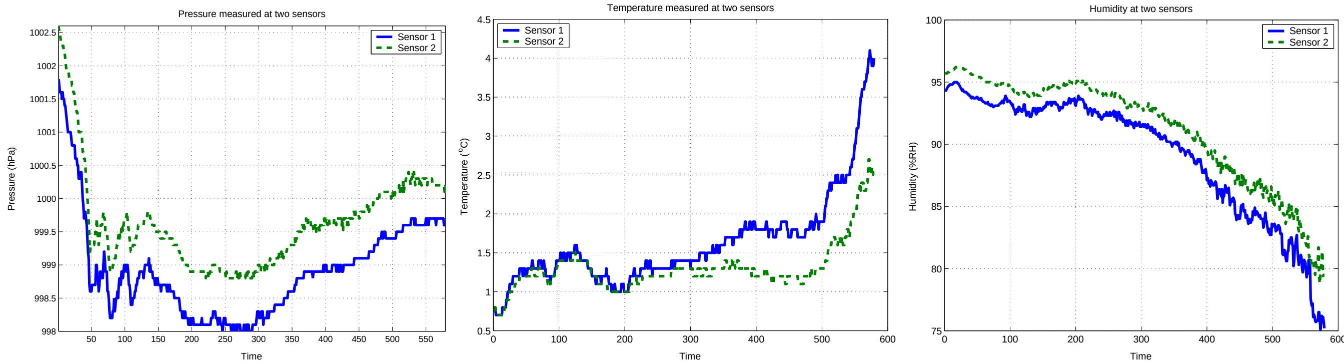

Figure 1. Variation of pressure, temperature, and humidity at two sensors over 12 hours

However, in the foreseeable future, some of these nodes may be operating on battery while monitoring physical phenomena such as maximum temperature or maximum wind velocity. Therefore, if all sensors transmit their sensed data it would result in unnecessary energy expenditure and would reduce the lifetime of the sensor network significantly. If however, the sensors could learn the spatio-temporal statistics of the sensed data and adapt themselves such that only certain sensors transmit, then we can determine and track the extremum values of the underlying fields with high accuracy.

In Figures 1 and 2, we plot all measured data from the two sensors over the entire time period of 12 hours. We observe from Figure 1(a) that, to determine the maximum value of pressure over this period of time, sensor 1 need not have transmitted at all, and the maximum could have been determined if only sensor 2 were transmitting. Similarly consider Figure 2(a) and Figure 2(b); the maximum values could be determined with sensor 2 transmitting most of the time and sensor 1 transmitting at certain times. Therefore, it makes sense to use the spatio-temporal correlation characteristics of the sensor field in order to make decisions about which sensors should transmit with the objective of saving energy while not sacrificing accuracy significantly.

Stochastic modeling of weather data is a reasonably mature discipline. In particular, modelling wind velocity is of great importance in civil engineering from a structural engineering point of view [12, 2]. It is used to forecast maximum wind speed to determine the worst case wind load for a structure. It is also required in wind energy production systems for forecasts of power which are generally derived from forecasts of speed. Researchers have explored the possiblities of modelling wind speed data with a first order Markov chain model [12]. Others have proposed techniques for forecasting wind speed, based on cross correlation at neighboring sites [2]. Accurate modeling of wind direction is important in coastal applications for determining worst case directional load on the structure.

We modeled the data collected from our testbed as a first order Spatio Temporal Auto Regressive (STAR) model. In particular, the first order spatio temporal process can be written as:

$$X_1[n] = \phi_{10} X_1[n-1] + \phi_{11} X_2[n-1] + \varepsilon[n]$$

where ϕ_{10} and ϕ_{11} are called the space-time partial autocorrelation functions. ϕ_{10} measures the extent of temporal correlation whereas ϕ_{11} measures spatial correlation. $X_i[n]$ is the measured value at sensor i at time n. More details about the STAR model have been presented in Appendix A.

We used techniques in [11] for estimating the parameters for the spatio-temporal process measured from the testbed. Specifically for the wind speed data,

$$\begin{bmatrix} \phi_{10} \\ \phi_{11} \end{bmatrix} = \begin{bmatrix} 0.6305 \\ 0.2568 \end{bmatrix}$$

and for the wind direction data

$$\begin{bmatrix} \phi_{10} \\ \phi_{11} \end{bmatrix} = \begin{bmatrix} 0.9535 \\ 0.0260 \end{bmatrix}$$

From the above correlation coefficients, we observe that wind speed has both spatial and temporal correlation but wind direction is mostly temporally correlated with little spatial correlation.

The basic tradeoff here that should be exploited is the following: if more sensors transmit, the tracking error would be minimized but the energy consumption would be high. On the contrary if too few sensors transmit, the energy consumed would be low, but we would run the risk of erroneous tracking. Therefore, we want to develop an algorithm that minimizes both the tracking error and the energy using the measured spatio-temporal characteristics of the sensed field. If we know the spatio-temporal correlation structure, current measurements can be used to predict future values and then a decision can be made by the algorithm about which sensors should transmit. We use this observation in Sections 4.5 and 5.1 for driving the simulations of our algorithms.

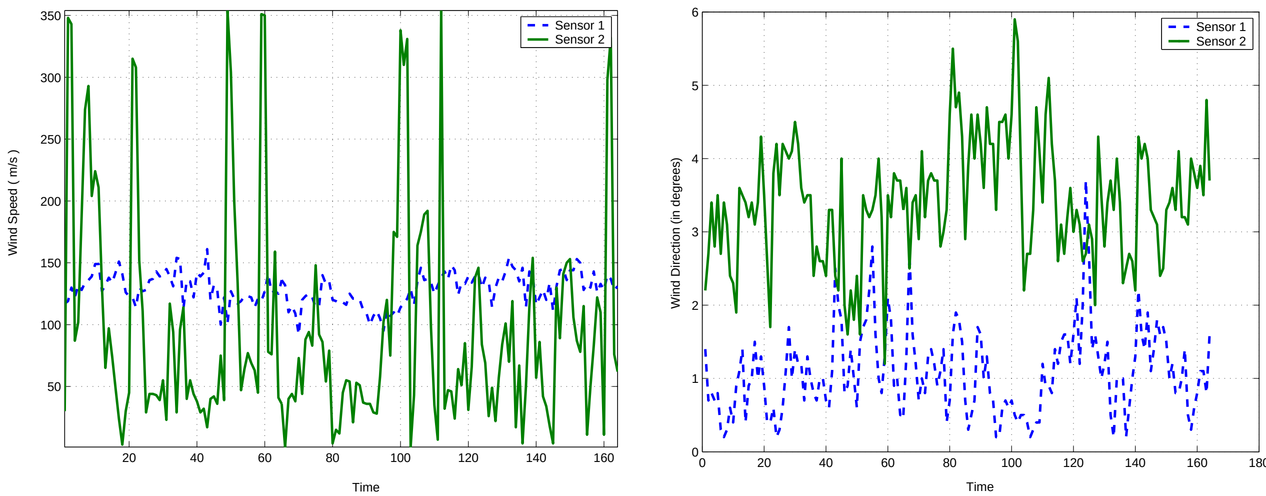

Figure 2. Wind velocity measurements at two sensors for one hour: (a) Wind Speed; (b) Wind Direction

3 Extremum Tracking Problem under Spatio-temporal Correlation

We consider the scenario in which N sensors are distributed randomly over an area $\mathcal{A}$. The sensors are attempting to sense a spatio-temporal field denoted by $V(x,y,n)$[2]. The field is sensed at discrete time instants $1,2,3,\ldots,n,\ldots$ by each sensor and the results are transmitted to the fusion center. The fusion center then determines the extremum value $\max(V(x_1,y_1,n),V(x_2,y_2,n),\ldots,V(x_N,y_N,n))$ for each n.

In this paper we consider both centralized and distributed versions of the problem. In the centralized version of the problem, the fusion center keeps history of all measurements from all sensors that it has received so far. At time instant n, the fusion center decides which of the N sensors need to transmit after examining the previously received data, and tasks those selected sensors to transmit in the next epoch. Upon receiving the measurements from the chosen sensors in the next epoch, the fusion center determines the maximum and repeats the process. In the distributed version of the problem, the fusion center does not make a decision about which sensor should transmit; instead the sensor nodes themselves make a decision about whether to transmit after overhearing transmissions from other nodes and exploiting the spatio-temporal correlations in the sensed field.

3.1 Mathematical Notation

We introduce some useful notation for the description of our algorithms. Uppercase characters indicate random variables and lowercase characters indicate a realization of the corresponding random variable, e.g., x is a realization of random variable X. Vectors of random variables are denoted by bold characters.

- n represents a discrete instant (or epoch) of time. We use this convention instead of the usual t because the latter is typically used for modeling continuous time.
- $X_i(n) = V(x_i,y_i,n)$ is the random variable that indicates the value of the field being measured at node i at time instant n. (x_i,y_i) is the location of sensor i.
- $\mathbf{X}[n] = [X_1(n),X_2(n),\ldots X_N(n)]$ is the vector of measurements at the sensors at time n.
- U_i is a random variable associated with sensor i that indicates sensor selection; $U_i = 1$ if sensor i transmits; otherwise, $U_i = 0$.
- Define $\mathbf{U} = [U_1,U_2,\ldots U_N]$ to be a vector of random variables. $\mathbf{u} = [u_1,u_2,\ldots u_N]$ is a specific realization of $\mathbf{U}$ that represents a transmission policy.
- $\mathbf{H}[n]$. This is the historical information available at the fusion center F at time instant n. This includes information received by F *before* time n. $\mathbf{H}[n] = [\mathbf{H}[n-1],X_i[n]$ s.t. $u_i = 1]$.
- $\delta_{\mathbf{u}} = \Pr(\mathbf{U} = \mathbf{u}|\mathbf{H}[n])$ is the conditional probability of a particular transmission policy $\mathbf{u}$ being selected by the algorithm given the historical measurements. The optimal policy could be randomized, which is given by $\delta_{\mathbf{u}}$. Note that $\sum_{\mathbf{u}} \delta_{\mathbf{u}} = 1$.
- $M[n] = \max(X_1[n],X_2[n],\ldots X_N[n])$ is the maximum value of the measurements from all sensors. $M[n]$ is a function of $\mathbf{X}[n]$ and is the best possible tracking.
- $F[n] = \max(X_i[n]$ s.t. $U_i = 1)$. $F[n]$ denotes the maximum value of the measurements received from the sensors who were tasked to transmit by the algorithm and is a function of $\mathbf{X}[n]$ and $\mathbf{U}$. $F_{\mathbf{u}}[n]$ denotes the measured maximum value if the sensors were following transmission policy $\mathbf{u}$.
- $\mathcal{E}$ is the energy consumed by a sensor when it transmits. For the sake of simplicity we assume equal transmission

[2]Note that we follow the convention of using n to denote discrete time instants.

power. It would be straightforward to extend it to a case where the transmission power depends on the distance to the fusion center.

- Define $k(\mathbf{u})$ as the number of ones in the vector $\mathbf{u}$ and denotes the number of sensors transmitting.

4 Centralized Algorithm for Tracking the Maximum

We assume that each sensor can transmit once in an information gathering epoch. The epochs are assumed to be equal in time duration. At the beginning of each information gathering epoch n, the fusion center uses the previously received measurements (in the $n-1^{st}$ epoch) to make a decision about which sensors should transmit their measurements in the current epoch. Upon receiving this information, the requisite sensors transmit to the fusion center.

We assume a slotted TDMA style MAC protocol in operation. The fusion center uses a broadcast slot to communicate the information about "which sensors transmit next" to all sensors listening in that slot. Along with this information, the fusion center also specifies a slot schedule for "which sensors transmits when" in the current epoch. Such a contention-free scheme allows sensors to put their transceivers to sleep at all times except when it is their turn to transmit. In addition to mitigation of packet losses due to collisions, this generally results in tremendous energy savings because of the reduction in idle listening. The fusion center, however, listens to all ongoing transmissions in the network.

For this analysis we assume a perfect channel with no transmission errors or data loss due to fading or time asynchrony due to clock skews. The fusion center chooses the maximum among the received values and denotes it as the maximum of the entire field. The fusion center also updates its history $\mathbf{H}[n]$ with the received values.

4.1 Objective Function

At the beginning of every epoch, the fusion center, depending on the history gathered so far, chooses the sensors that ensure minimum deviation from an estimate of the actual maximum in the epoch. If all sensors transmit, the tracking error would be zero but we should choose the minimum number of sensors possible for reducing energy consumption. We capture the competing goals of minimizing both energy consumption and the tracking accuracy by using the following linear cost function as the objective function for the optimization problem (λ is a multiplier that weighs the relative importance of energy efficiency and tracking error):

$$\lambda \times \text{Energy consumed} + (1-\lambda) \times \text{Tracking Error}$$

Since the fusion center may not have access to the measurements ($\mathbf{X}$) before making a decision, we attempt to minimize the expected value of the deviation to account for all possible realizations of $\mathbf{X}$. Also the optimal policy $\mathbf{U}$ may be probabilistic. So we take the expectation over $\mathbf{U}$ as well. Our optimization problem then becomes the following:

$$\min_{\delta} \{\lambda \times \mathbb{E}_{\mathbf{U},\mathbf{X}}(E_{total}|\mathbf{H}[n]) + (1-\lambda) \times \mathbb{E}_{\mathbf{U},\mathbf{X}}(M[n]-F[n]|\mathbf{H}[n])\} \quad (1)$$

subject to constraint:

$$\sum_{\mathbf{u}} \delta_{\mathbf{u}} = 1$$

where E_{total} denotes the total energy consumed by all sensors and $\mathbb{E}_{\mathbf{U},\mathbf{X}}$ denotes the expectation over all possible realizations of $\mathbf{U}$ and $\mathbf{X}$ given the history $\mathbf{H}[n]$.

4.2 Optimal Policy

THEOREM 1. The cost function given in Equation 1 is minimized by a deterministic policy given by

$$\begin{aligned} \delta_{\mathbf{u}} &= 1, \text{ if } \mathbf{u} = \mathbf{u}_{opt} \\ &= 0, \text{ otherwise} \end{aligned} \quad (2)$$

where

$$\mathbf{u}_{opt} = \arg\min_{\mathbf{u}} \{\lambda \times k(\mathbf{u}) \times \mathcal{E} + (1-\lambda) \times \mathbb{E}_{\mathbf{U},\mathbf{X}}(M[n]-F_{\mathbf{u}}[n]|\mathbf{H}[n])\} \quad (3)$$

$F_{\mathbf{u}}[n]$ is the maximum among the readings of the sensors which transmit as indicated by $\mathbf{u}$.

PROOF. Consider the first term

$$\lambda \times \mathbb{E}_{\mathbf{U},\mathbf{X}}(E_{total}|\mathbf{H}[n]) = \lambda \times \sum_{\mathbf{u}} \delta_{\mathbf{u}} \times k(\mathbf{u}) \times \mathcal{E}$$

$\mathbb{E}_{\mathbf{U},\mathbf{X}}(M[n]|\mathbf{H}[n])$ is independent of $\mathbf{U}$ and hence does not depend on $\delta_{\mathbf{u}}$.

$$\begin{aligned} \mathbb{E}_{\mathbf{U},\mathbf{X}}(F[n]|\mathbf{H}[n]) &= \sum_{\mathbf{u}} \int_{\mathbf{x}} F[n]\, p(\mathbf{u},\mathbf{x}|\mathbf{H}[n])\, d\mathbf{x} \\ &= \sum_{\mathbf{u}} \int_{\mathbf{x}} F[n]\, p(\mathbf{x}|\mathbf{H}[n])\, \delta(\mathbf{u}|\mathbf{H}[n])\, d\mathbf{x} \\ &= \sum_{\mathbf{u}} \delta_{\mathbf{u}}\, \mathbb{E}_{\mathbf{X}}(F_{\mathbf{u}}[n]|\mathbf{H}[n]) \end{aligned}$$

We can now formulate this optimization as a linear programming problem of the form

$$\text{Minimize } \sum_{\mathbf{u}} \delta_{\mathbf{u}} C_{\mathbf{u}}$$

$$\text{subject to } \sum_{\mathbf{u}} \delta_{\mathbf{u}} = 1$$

where $C_{\mathbf{u}} = \lambda \times k(\mathbf{u}) \times \mathcal{E} + (1-\lambda) \times \mathbb{E}(M[n]-F_{\mathbf{u}}[n]|\mathbf{H}[n])$.

By the Fundamental Theorem of Linear Programming, the solution to this optimization problem is attained at the end points of the $2^N - 1$ simplex formed by the $\delta_{\mathbf{u}}$'s [8]. Hence the optimal solution is given by Equation 2. □

4.3 The Algorithm

In this section, we propose the following algorithm to track the maximum of a spatio-temporally correlated sensor field. Algorithm can broken down into the start-up phase ($n = 0$) and subsequent phases ($n > 0$).

Initialization phase ($n = 0$):

1. All the sensors transmit their measurements to the fusion center
2. $F(0) = \max(X_1[n], X_2[n], \ldots X_N[n])$
3. $\mathbf{H}[0] = [X_1[0], X_2[0], \ldots X_N[0]]$

Subsequent phases ($n > 0$):

1. The fusion center determines which sensors have to transmit based on the algorithm
2. $F[n] = \max(X_i[n] \text{ s.t. } u_i = 1)$
3. $\mathbf{H}[n+1] = [\mathbf{H}[n], X_i[n] \text{ s.t. } u_i = 1]$

4.4 Exploiting the Correlation Model

We assume that the measured data are jointly zero-mean gaussian random variables with the following correlation structure:

$$\mathbb{E}(X_i[n_1]X_j[n_2]) = \sigma^2 e^{-B|n_1-n_2|} e^{-Ad_{ij}} \quad (4)$$

where n_1 and n_2 denote discrete instants of time, d_{ij} is the distance between sensor i and j, and σ^2 is the variance. A and B control the degree of correlation between the measured data samples – higher A implies low spatial correlation and vice-versa. Similarly, B controls temporal correlation. We now explain how we use historical information to affect the expected value of the highest order statistic.

Given two jointly gaussian random variables $\mathbf{X}$ and $\mathbf{Y}$ we have,

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} \begin{pmatrix} \Sigma_{\mathbf{XX}} & \Sigma_{\mathbf{XY}} \\ \Sigma_{\mathbf{YX}} & \Sigma_{\mathbf{YY}} \end{pmatrix} \right)$$

where $\Sigma_{\mathbf{XX}}, \Sigma_{\mathbf{XY}},$ and $\Sigma_{\mathbf{YX}}$ are covariance matrices. This implies:

$$\mathbb{E}(\mathbf{X}|\mathbf{Y}) = \Sigma_{\mathbf{XY}} \Sigma_{\mathbf{YY}}{}^{-1} \mathbf{Y} \quad (5)$$

$$\mathrm{Var}(\mathbf{X}|\mathbf{Y}) = \Sigma_{\mathbf{XX}} - \Sigma_{\mathbf{XY}} \Sigma_{\mathbf{YY}}{}^{-1} \Sigma_{\mathbf{YX}} \quad (6)$$

The determination of the optimal policy requires the evaluation of the expected value of the highest order statistic conditioned on the history. The bound on the expected order statistic can then be evaluated by calculating the conditional mean and conditional variance of the random variables, conditioned on the history. We can determine this by using Equations 5 and 6, where $\mathbf{Y} = \mathbf{H}[n]$ and $\mathbf{X} = X_i[n]$ for all i.

We use order statistics bounds given by Equations 16 and 17 in Appendix B to compute the expected value of the maximum.

If the measured data like wind direction is spatially uncorrelated, i.e.,

$$\mathbb{E}(X_i(n_1)X_j(n_2)) = \sigma^2 e^{-A|n_1-n_2|} \delta_{ij} \quad (7)$$

where δ_{ij} is the kronecker delta, then it is sufficient to keep the most recent received values in the history.

CLAIM 1. The set of most recently received values is sufficient as history for the correlation structure of Equation 7.

PROOF. For the correlation structure of Equation 7, the measured data at each sensor forms a one-dimensional Gauss-Markov process. We use history information to estimate the conditional mean and the variance of the next measurement. Without loss of generality, consider sensor 1. Since the measurements are spatially uncorrelated, any measurements from other sensors do not affect the conditional mean or the variance at this sensor. Let us assume that we are using all received readings of sensor 1 as the history i.e. $X_1[n_1], X_1[n_2], \ldots X_1[n_N]$ and $n_1 > n_2 > \ldots n_N$.

$$\begin{aligned} &\mathbb{E}(X_1[n]|\text{all history from all sensors}) \\ &= \mathbb{E}(X_1[n]|X_1[n_1], X_1[n_2], \ldots X_1[n_N]) \quad (8) \\ &= \mathbb{E}(X_1[n]|X_1[n_1]) \quad (9) \end{aligned}$$

Equation 8 follows because the measurements are spatially uncorrelated. Equation 9 follows due to the fact that for each sensor, the temporal measurements form a 1-D Gauss Markov Random Process. □

Note that for the case where the measured data is spatio-temporally correlated, from a theoretical point of view, only the most recently received data is not sufficient for the history. However, from a practical point of view, we observed from our simulation experiments (presented in Section 4.5) that using most recent history alone yields good tracking performance.

4.5 Simulation Results

In the simulations of the centralized algorithms, 15 sensors are distributed randomly in a $80m \times 80m$ square. They sense a zero-mean jointly Gaussian field with the correlation structure of Equation 4. We assume that each sensor requires unit energy to transmit to the fusion center. The energy consumed in every epoch is therefore the number of sensors transmitting in that epoch.

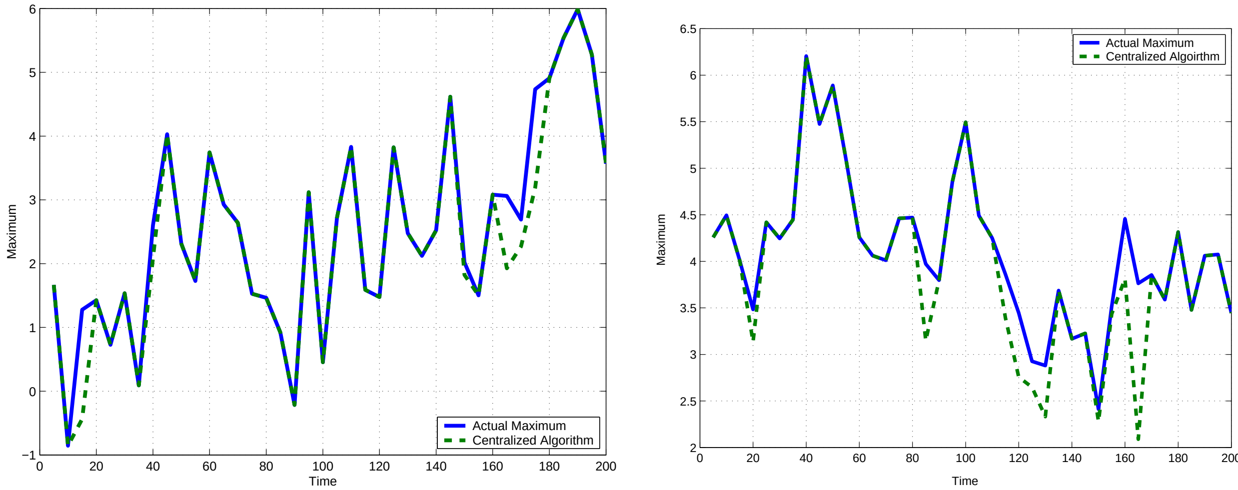

Figure 3. Tracking Maximum Wind Characteristics using Centralized Algorithm: (a) Speed; (b) Direction

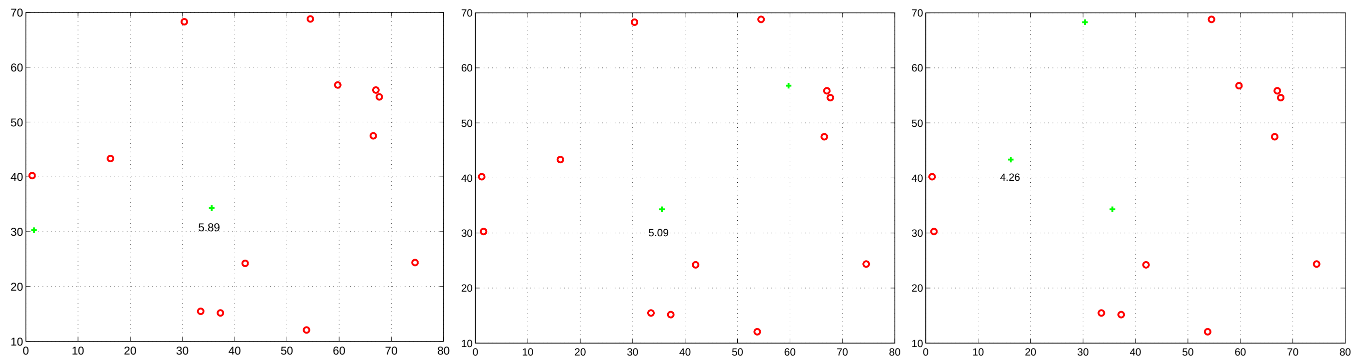

Figure 4. Status of sensors: (a) at t=50s; (b) t=55s; (c) t=60s

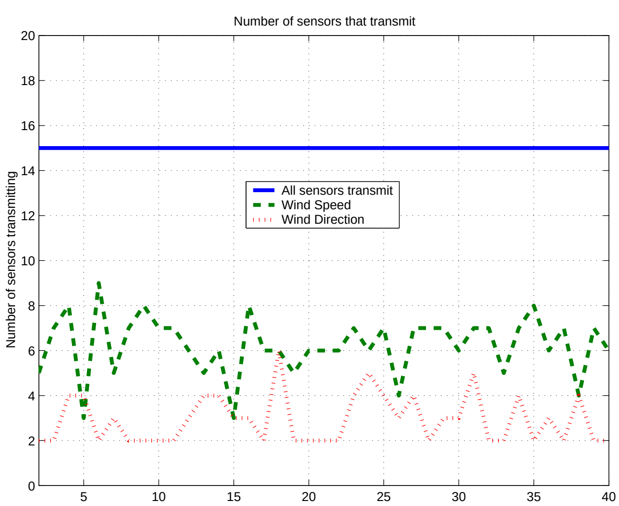

Figure 5. Centralized Algorithm: Energy Consumption Over Time (Number of Transmissions)

In our simulations, we used the partial space time autocorrelation functions determined by the experimental data presented in Section 2. ϕ_{10} is the correlation coefficient for a temporal lag of one unit and no spatial lag. ϕ_{11} is the correlation coefficient for a temporal lag of one unit and a spatial lag of one. Hence, from Equation 7, we have:

$$\begin{aligned} e^{-B\Delta n} &= \phi_{10} \\ e^{-Ad_{12}} e^{-B\Delta n} &= \phi_{11} \end{aligned}$$

where Δn is the time difference between the samples and d_{12} is the distance between the two sensors.

Wind speed data has both spatial and temporal correlation; we have determined that $A = 0.0107$ and 0.0922 respectively. We set $\lambda = 0.5$ and $\sigma^2 = 5$ for these simulations since we wanted to put equal imporance on energy conservation and tracking accuracy. In Figure 3(a), the actual maximum, $M[n]$, and the maximum determined by the centralized algorithm, $F[n]$, is plotted against time. The centralized algorithm tracks the maximum with very little deviation. Our simulations showed energy savings of 62.67% over the time period of 200s in comparison with the policy of all sensors tranmsitting in every time epoch.

For less spatially correlated data like wind direction, $A = 0.044$ and $B = 0.0095$ were respectively derived from the space-time partial autocorrelation functions. In this simulation we chose $\lambda = 0.45$ and $\sigma^2 = 5$. We plotted the actual maximum and the maximum as determined by the centralized algorithm in Figure 3(b). For this 200s simulation run, we achieved an energy savings of nearly 79%.

Higher energy savings can be expected when sensing highly temporally correlated data – if a sensor has sensed the maximum at a particular instant of time, it is likely that in the next epoch the same sensor would continue to report the maximum value. This is because the measurement of this sensor is unlikely to be affected by that of the neighbor due to low spatial correlation.

We now analyze the ability of the algorithm to select the appropriate sensors to transmit. We have used the simulation data of the parameters obtained for the wind direction data. In Figures 4(a-c), the circles represent sensors in the area that are not transmitting and the diamonds represent sensors that are transmitting. Nodes with juxtaposed real numbers denote the sensor that had measured the maximum value in that time epoch. At time=50s, the status of the sensors is illustrated by Figure 4(a).

Sensor status at times 55s and 60s has been shown in Figures 4(b) and 4(c), respectively. The sensor reporting the maximum value at time 60s was not transmitting at times 50 and 55s. This demonstrates that the algorithm is intelligent enough to adapt to this situation and query the particular sensor at time=60s which did not transmit in the previous three epochs.

Figure 5 illustrates how the energy consumption of the centralized algorithm varies over time for tracking both wind speed and direction. We observe that, on average, only 6 sensors (out of 15) need to transmit for tracking the maximum wind speed and even fewer number of sensors are needed to track maximum wind direction. This is because the latter phenomenon has a very high degree of temporal correlation (ϕ_{10}) whereas the former phenomenon has much less temporal correlation although it does have some spatial correlation (ϕ_{11}).

5 Distributed Algorithm for Tracking the Maximum

The centralized algorithm presented in Section 4 suffers from the problem of exponential time complexity since it uses a brute force approach to optimization over 2^N transmission policies. Hence, we propose a suboptimal distributed algorithm which is a simple extension of our centralized algorithm. In the distributed algorithm, the fusion center does not make a decision about the sensors that should be transmitting. Instead, each sensor node makes that decision locally.

We define:

$$M[n] = \max\{X_1[n], X_2[n], \ldots, X_N[n]\} \quad (10)$$

The sensors overhear transmissions and update their local history. We assume that the time is slotted as in Section 4 and the fusion center distributes transmission slot-schedules *a priori* to all nodes. Each node knows which slot (in each epoch) to transmit its measurements in.

At every time step, each sensor tries to locally minimize the cost function according to the policy defined by Equation 3. Each sensor S_i follows the following rule: if last-received values from certain other sensors $S_{i_1}, \ldots, S_{i_k}$ in S_i's history are greater than S_i's value, then S_i assumes that $S_{i_1}, \ldots, S_{i_k}$ will transmit. If the decision is to transmit, S_i senses and transmits its measurement to the fusion center.

In this distributed algorithm, each sensor node keeps awake in other slots as well in order to overhear the transmissions from other sensors. If the decision is not to transmit, that sensor sleeps for that time epoch. However, if the sensor sleeps for a long period, it may lose track of the state of the field. Hence, we propose that all sensors wake up, listen, and transmit at *a priori* fixed periodic intervals of duration T.

More formally, when $n = 0$ or $n \bmod T = 0$:

1. All sensors transmit to the fusion center
2. Each sensor updates its history about other sensors after hearing their transmissions to the fusion center
3. The fusion center determines the maximum according to Equation 10

In subsequent phases ($n > 0$ and $n \bmod T \neq 0$), for every sensor i, the following steps are taken:

1. $\mathbf{H}_i[n]$ is the vector denoting history available at sensor i at time n. In formal terms, $\mathbf{H}_i[n] = [H_i^1, H_i^2, \ldots, H_i^N]$, where H_i^j is the last received value from sensor j at sensor i.
2. Let $u_j = 1$ if $H_i^j \geq H_i^i$
3. Define
$$\begin{aligned} \mathbf{u}_i^1 &= [u_1, \ldots, u_{i-1}, 1, u_{i+1}, \ldots, u_N] \\ \mathbf{u}_i^0 &= [u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_N] \end{aligned}$$
4. Every sensor i calculates:
$$\lambda \times k(\mathbf{u}_i^0) \times \mathcal{E} + (1-\lambda) \times \mathbb{E}(M[n] - F_{\mathbf{u}_i^0} | \mathbf{H}_i[n]) \quad (11)$$
$$\lambda \times k(\mathbf{u}_i^1) \times \mathcal{E} + (1-\lambda) \times \mathbb{E}(M[n] - F_{\mathbf{u}_i^1} | \mathbf{H}_i[n]) \quad (12)$$
5. If Equation 11 > Equation 12, then sensor i transmits.
6. If it transmits, it also listens to the transmissions of the other sensors and updates $\mathbf{H}_i[n+1]$, else it sleeps until the beginning of the next time epoch.

The above distributed algorithm involves computation of only two values (Equations 11 and 12) at each sensor; therefore, it is much more computationally efficient in compari-

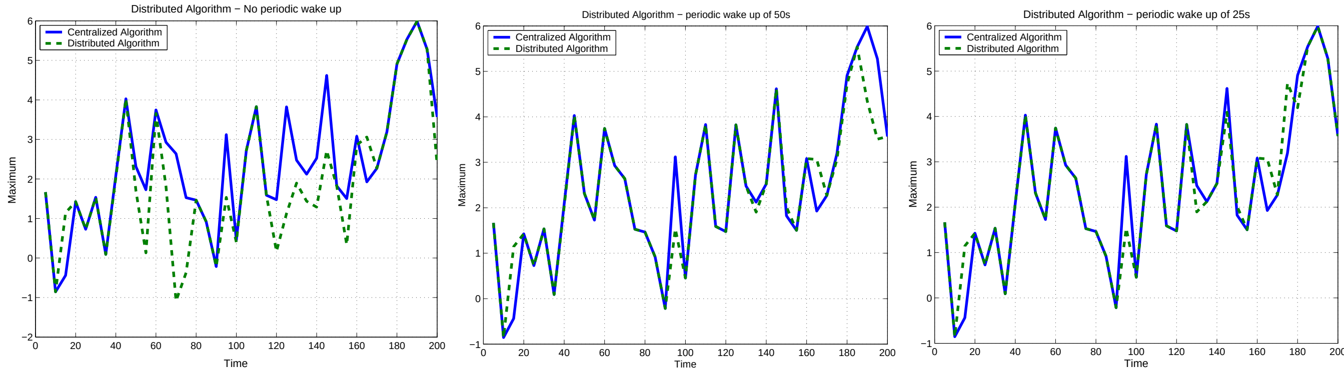

Figure 6. Tracking the Maximum with the Distributed Algorithm

sion with its centralized counterpart which has to make 2^N computations at each time epoch. In Section 5.1 we observe that the tracking and energy-efficiency of the distributed algorithm is comparable to that of the centralized algorithm, hence it is a good candidate for energy-efficient tracking of extrema in sensor networks.

5.1 Simulation Results

We simulated the distributed algorithm with the same data that was used in the simulation of the centralized algorithm, and compared their performance. The simulation was carried out on the wind direction data. From Figure 6, we can observe that the distributed algorithm performs almost as well as the centralized algorithm although the total energy consumed is $\approx 25\%$ more than the latter (for periodic wakeup of $T = 25s$). The mean squared error is 0.1026 which is lower in comparison with 0.1773 for the centralized algorithm because more sensors transmit in the case of the former, thus resulting in the reduction of the error.

The distributed algorithm is efficient in its operation since it just involves the computation of two values before making a decision about whether to transmit. With minimal extra energy expenditure, we obtain very good tracking performance.

Figure 7 illustrates how the energy consumption of the distributed algorithm varies over time for tracking wind direction. We observe that the distributed algorithm is able to conserve energy almost as significantly as its centralized counterpart. However, the exact energy efficiency depends on the value of the periodic wakeup interval T. For the case where there is no wakeup at all, the energy consumption is quite low (lower than centralized) at the cost of moderate tracking error. For the case where sensors wake up periodically with a certain frequency, ($T = 50s$) we observe that energy consumption rate occasionally shoots up but it is low overall. From Figures 6 and 7 we conclude that a distributed scheme with a low frequency periodic wakeup has decent energy efficiency and good extremum tracking performance.

6 Conclusion

Many natural weather phenomena such as temperature, humidity, pressure, and wind velocity exhibit spatio-

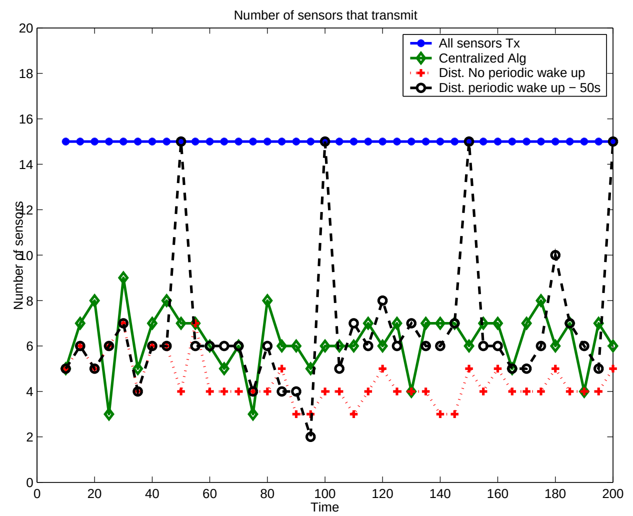

Figure 7. Distributed Algorithms: Energy Consumption Over Time (Number of Transmissions)

temporal correlation of varying degrees. In this paper we show that a single hop wireless sensor network that is tracking the extrema of such fields can intelligently exploit the underlying spatio-temporal correlations of the field in order to save precious battery energy and thus extend the lifetime of the sensor network. In particular, we proposed both centralized and distributed versions of an optimization algorithm that simultaneously attempt to minimize the energy consumption as well as the tracking error. We showed that both the centralized and the distributed algorithms could track the maximum of a spatio-temporally correlated field remarkably well over varying degrees of correlation while spending little energy. The algorithms were also adaptive to detect sudden changes in the phenomena.

Future work: In this paper we focused on single hop sensor networks where all sensor nodes can directly communicate with the fusion center and can also hear each other's transmissions. While we were able to perform systematic analysis for this case, the more general case of multihop sensor networks is more interesting. We believe that a similar optimization framework is a good starting point for the multihop scenario. In particular the cost function should capture both the energy consumption and the tracking error components. The biggest challenge in this case is that we will need to worry about how the data is *routed* from any given sensor node to the fusion center. The choice of routes is highly likely to impact the tracking error as well as the energy efficiency. Some proposals in the literature address this topic heuristically. We believe that the objective function proposed in this paper can be adapted to incorporate the routing metric and is promising for future analysis.

Also, our centralized algorithm is brute force in nature with exponential time complexity. It is worthwhile to investigate approximation algorithms that run faster and produce nearly optimal results.

7 References

[1] I. F. Akyildiz, M. C. Vuran, and O. B. Akan, "On Exploiting Spatial and Temporal Correlation in Wireless Sensor Networks," *Proc. WiOpt*, 2004.

[2] M. C. Alexiadis, P. S. Dokopoulos and H. S. Sahsamanoglou, "Wind speed and power forecasting based on spatial correlation models," *IEEE Transactions on Energy Conversion*, September 1999.

[3] D. Berstimas, K. Natarajan and C. Teo, "Tight bounds on Expected Order Statistics," *Probability in the Engineering and Informational Sciences*, 2006.

[4] A. Boulis, S. Ganeriwal and M. Srivatsava, "Aggregation in Sensor Networks: An Energy Accuracy Tradeoff," *Proc. Sensor Network Protocols and Applications*, 2003.

[5] H. A. David and H. N. Nagaraja, "Order Statistics," 3rd Edition, Wiley, 2003.

[6] T. Fuhrmann and J. Widmer, "Extremum Feedback with Partial Knowledge," *Lecture Notes in Computer Science*, Springer, 2004.

[7] P. Kinney, "ZigBee Technology: Wireless Control that Simply Works," *Communication Design Conference*, October 2003. `http://www.zigbee.org/imwp/idms/popups/pop_download.asp?contentID=5162`.

[8] D. G. Luenberger, "Linear and Non Linear Programming," 2nd Edition, Addison-Wesley, 1989.

[9] G. Mergen, Q. Zhao, and L. Tong, "Sensor Networks with Mobile Access: Energy and Capacity Considerations," *IEEE Transactions on Communications*, vol. 54, no. 11, pp 2033-2044, Nov., 2006.

[10] S. Pattem, B. Krishnamachari, R. Govindan," The impact of spatial correlation on routing with compression in wireless sensor networks", *Symposium on Information Processing in Sensor Networks (IPSN)*, 2004.

[11] P. E. Pfeifer, S. J. Deutsch, "A Three Stage Iterative Procedure for Space- Time Modeling," *Technometrics*, Vol. 22, No. 1, Feb 1980.

[12] A. D. Sahin and Z. Sen, "First-order Markov chain approach to wind speed modelling," *Journal of Wind Engineering and Industrial Aerodynamics*, Volume 89, Issues 3-4, March 2001.

[13] Y. Sung, S. Misra, L. Tong and A. Ephremides, " Cooperative Routing for Signal Detection in Large Sensor Networks," *IEEE JSAC Special Issue on Cooperative Communications and Networking*, vol. 25, no. 2, pp. 471–483, Feb. 2007.

[14] J. Widmer and T. Fuhrmann, "Extremum Feedback for Very Large Multicast Groups," *Lecture Notes in Computer Science*, Springer, 2001.

A Spatio Temporal Processes

A flexible class of empirical models for space time models are the ones that are characterized by the autoregressive and moving average forms of univariate time series lagged in both space and time. These models have been referred to as space time moving average (STARMA) models. The STARMA model is characterized by linear dependance in both space and time. Assuming that observation $z_i(t)$ of the random variable $Z_i(t)$ are available at each of N fixed locations in space $(i = 1, 2, \ldots, N)$ over T time periods of $(t = 1, 2, \ldots, T)$.

The dependance of $z_i(t)$ on past observations and errors at site i can be expressed by the following equation.

$$\begin{aligned} z_i(t) &= \sum_{k=1}^{p}\sum_{l=0}^{\lambda_k} \phi_{kl} L^{(l)} z_i(t-k) \\ &\quad - \sum_{k=1}^{q}\sum_{l=0}^{m_k} \theta_{kl} L^{(l)} \varepsilon_i(t-k) + \varepsilon_i(t) \end{aligned} \tag{13}$$

where p is the autoregressive order, q is the moving average order, λ_k is the spatial order of the k^{th} autoregressive term, m_k is the spatial order of the k^{th} moving average term, ϕ_{kl} and θ_{kl} are parameters, and the $\varepsilon_i(t)$ are random normal errors with

$$\mathbb{E}[\varepsilon_i(t)] = 0$$

$$\mathbb{E}[\varepsilon_i(t)\varepsilon_j(t+s)] = \begin{cases} \sigma^2 & i = j, s = 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$L^{(l)} z_i(t) = \sum_{j=1}^{N} w_{ij}^{(l)} z_j(t)$$

where w_{ij}^{l} are a set of weights with

$$\sum_{j=1}^{N} w_{ij}^{(l)} = 1$$

for all i and $w_{ij}^{(l)}$ nonzero only if sites i and j are l^{th} order neighbors.

We have used the first order spatio temporal autoregressive (STAR) model. It is obtained by substituting $p = 1, q = 0, \lambda_1 = 1$ in Equation 13, which then can be written as

$$z_i(t) = \phi_{10} z_i(t-1) + \phi_{11} \sum_j w_{ij}^{(1)} z_i(t-1) \tag{14}$$

ϕ_{10} and ϕ_{11} are called the space time partial autocorrelation function. ϕ_{10} represents correlation for no spatial lag and unit time lag. ϕ_{11} represents correlation for a spatial lag of 1 (immediate neighbors) and a time lag of 1. These parameters can be solved by solving the space time analog of Yule - Walker equations

$$\begin{bmatrix} \gamma_{00}(1) \\ \gamma_{10}(1) \end{bmatrix} = \begin{bmatrix} \gamma_{00}(0) & \gamma_{01}(0) \\ \gamma_{10}(0) & \gamma_{11}(0) \end{bmatrix} \begin{bmatrix} \phi_{10} \\ \phi_{11} \end{bmatrix}$$

where $\gamma_{lk}(s)$ denotes the covariance between the weighted l^{th} order neighbors of any site and the weighted k^{th} order neighbors of the same site at a time lag of s.

$$\gamma_{lk}(s) = \mathbb{E}\left\{ \sum_{i=1}^{N} \frac{L^{(l)} z_i(t) L^{(k)} z_i(t+s)}{N} \right\} \tag{15}$$

B Using Expected Order Statistics

If random variables $X_1, X_2, \ldots X_n$ are arranged in order of magnitude and written as

$$X_{(1)} \leq \ldots \leq X_{(n)}$$

We call $X_{(i)}$ as the ith order statistic ($i = 1, 2, \ldots, N$). $X_{(n)}$ denotes the maximum among the random variables. It is a central problem in statistics to evaluate, bound or approximate the moments of the highest order statistic under varying assumptions on the distribution of the underlying random variables. For a more detailed review on this subject see [5].

We use bounds developed in [3] for evaluating the expected highest order statistic, i.e. the maximum. The authors of [3] derive bounds under moment information of the random variables, in particular the first and second moments. No assumptions are made about the independence or distribution of random variables.

DEFINITION 1. Z_k^* is a tight upper bound on the kth-order statistic if

$$Z_k^* = \sup_{\mathbf{X} \sim \theta^{\mathbf{m}}} \mathbb{E}_\theta[X_{(k)}]$$

that is, there exists a feasible distribution or a limit of a sequence of feasible distributions that achieves the upper bound.

Let $\mathbf{X} \sim \theta^{\mathbf{m}}$ denote the set of feasible distributions θ that satisfies the given moments $\mathbf{m}$ for the random variables.

THEOREM 2. [3] The tight upper bound on the expected value of the highest-order statistic Z_n^* given $\mathbf{X} \sim \theta(\mu, \sigma^2)$is obtained by solving the strictly convex univariate minimization problem.

$$Z_n^* = \min_z f_n(z) = \min_z \left(z + \sum_{i=1}^{n} \frac{1}{2} [\mu_i - z + \sqrt{(\mu_i - z)^2 + \sigma_i^2}] \right)$$

[3] also proposes the following closed form bounds:

$$Z_n^* \leq \frac{1}{2} \left(\sum_{i=1}^{n} \left[\mu_i + \sqrt{\left(\mu_i - \max_{1 \leq i \leq n} \{Y_i\} \right)^2 + \sigma_i^2} \right] + (2-n) \left[\max_{1 \leq i \leq n} \{Y_i\} \right] \right) \tag{16}$$

and

$$Z_n^* \leq \frac{1}{2} \left(\sum_{i=1}^{n} \left[\mu_i + \sqrt{\left(\mu_i - \min_{1 \leq i \leq n} \{Y_i\} \right)^2 + \sigma_i^2} \right] + (2-n) \left[\min_{1 \leq i \leq n} \{Y_i\} \right] \right) \tag{17}$$

where $Y_i = \mu_i + \frac{n-2}{2\sqrt{n-1}} \sigma_i$.