STATISTICALLY-BASED EVENT DETECTION USING WIRELESS SENSOR NETWORKS

TIBLEZ OGBAGABRIEL
February, 2012

SUPERVISORS:
Dr. N.A.S. Hamm
Dr. N. Meratnia
Prof.Dr.Ir. A. Stein

STATISTICALLY-BASED EVENT DETECTION USING WIRELESS SENSOR NETWORKS

TIBLEZ OGBAGABRIEL
Enschede, The Netherlands, February, 2012

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:
Dr. N.A.S. Hamm
Dr. N. Meratnia
Prof.Dr.Ir. A. Stein

THESIS ASSESSMENT BOARD:
Prof.Dr.Ir. A. Stein (Chair)
Dr. Yang Zhang (External Examiner, University of Twente)

ABSTRACT

Widespread spatial and temporal monitoring of the environment is required for drawing accurate scientific conclusions and for providing reliable forecasting. Traditional environmental measurement systems use expensive sensing stations which makes it difficult to deploy a large number of sensing units. Wireless Network Sensors (WSNs) emerged as a solution to such issues. WSNs can be deployed in great numbers to observe the environment with high temporal and spatial resolution. On the other hand, WSNs are constrained by computational, storage and communication limitations.

Event detection is an important application in the practical deployment of wireless sensor networks (WSNs). One of the key challenges in detecting events using WSNs is how to detect it accurately and fast (real time) in an online and decentralized way. In order to assure reliable detection of interesting events using WSNs, we need to develop an accurate outlier detection technique while paying attention to the computational, storage and communication limitations in WSNs. As a result there is a need of developing analytical techniques for outlier and event detection which can fit to the resource constraint nature of WSNs. This research takes advantage of the spatial and temporal correlations that exist between sensor data in order to ensure reliable detection of events using WSNs, while maintaining the resource consumption of the WSN to a minimum.

The analysis was done based on data from a real world deployment on a high mountain pass (the Grand St. Bernard pass) in Switzerland. The spatial and temporal correlation in sensor data is exploited by using statistical approaches. This research provides domain specific definition for outliers based on the guidelines of World Meteorological Organization (WMO). Strategies for defining events and distinguishing them from errors are provided based on temporal and spatial correlations.

This research proposed detecting obvious outliers using plausible value and minimum variability checks and investigated the use of model independent and computationally simple outlier detection techniques for detecting unclear outliers. The outlier detection accuracy of the proposed techniques was evaluated using detection rate (DR) and false positive rate (FPR) based on the results of three labelling techniques (running average-based, Mahalanobis distance-based and density-based) by Zhang (2010). Event detection accuracy was evaluated by relabeling these labelled datasets so that they can be used to evaluate the event detection.

This research examined the use of geostatistical analysis for modelling spatial correlation using the variogram averaging method of Sterk & Stein (1997) which is based on the assumption that a constant correlation exists over time. The prediction accuracy of the spatial correlation model was evaluated using leave-one-out cross validation technique and the assumption of constant spatial correlation over time in which the variogram averaging method of Sterk & Stein (1997) is based in is verified.

Keywords: Outlier detection, Event detection, Wireless sensor networks, Time series analysis, Geostatistics, Temporal correlation, Spatial correlation.

ACKNOWLEDGEMENTS

To God be the glory for giving me the strength to finish this thesis.

I owe my deepest gratitude to those, in one way or another, contributed their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude goes to my supervisors who have made available their support in a number of ways. I would like to thank my first supervisor Dr. N.A.S. Hamm for his helpful comments and suggestions throughout this thesis work. He has been very supportive and has given me great supervision and advice during all those months. Deep gratitude goes to my second supervisors Dr. N. Meratnia and Prof.Dr.Ir. A. Stein. I am heartily thankful to Dr. N. Meratnia whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. I would like to show my greatest appreciation to Prof.Dr.Ir. A. Stein for his comments and suggestion.

Special thanks to Dr. Yang Zhang for his quick response whenever I needed help and for providing his labelled dataset, without which I couldn't evaluate the outlier and event detection techniques.

My acknowledgement goes to the Joint Japan World Bank Graduate Scholarship Program (JJ/WBGSP) for funding my studies and this research.

Last but not the least, I offer my regards and blessings to my parents and all my family for their moral support and prayers. I express my deepest appreciation to Kin for his encouragement and help. I am thankful to all my friends, especially Fevi and Mike, for their many hours of discourse in academic and general discussion which made life in ITC enjoyable. I want to thank all of those who supported me in any respect during the completion of this thesis. God bless you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CI	confidence level
DR	detection rate
EMA	exponential moving average
FPR	false positive rate
GPRS	general packet radio service
GSM	global system for mobile communication
ME	mean error
MM	moving median
n	smoothing window size
RMSE	root mean squared error
SMA	Simple moving average
SSErr	Sum of squared error
TOD	temporal outlier detection
WMA	weighted moving average
WMO	World Meteorological Organization
WSN(s)	wireless Sensor network(s)

1. INTRODUCTION

Wireless Sensor Networks (WSNs) are new source of geodata that measure various environmental variables such as temperature, rainfall, atmospheric pressure, precipitation, humidity etc. They consist of a number of low cost sensor nodes with sensing, processing and communication capabilities. These sensors can be deployed in harsh and unattended regions to monitor the natural environmental conditions by detecting interesting events that are associated with the natural environment such as fire and storm (Zhang et al., in press). Figure 1 shows an example of a Wireless Sensor Network (WSN) setup where data from the sensing stations are transmitted through the wireless network to a base station (GPRS master station). The base station then forwards the data by means of long range connection (GSM network) to a central server and makes it available to the outside world in real time (Barrenetxea et al., 2008a; SensorScope, 2007b).

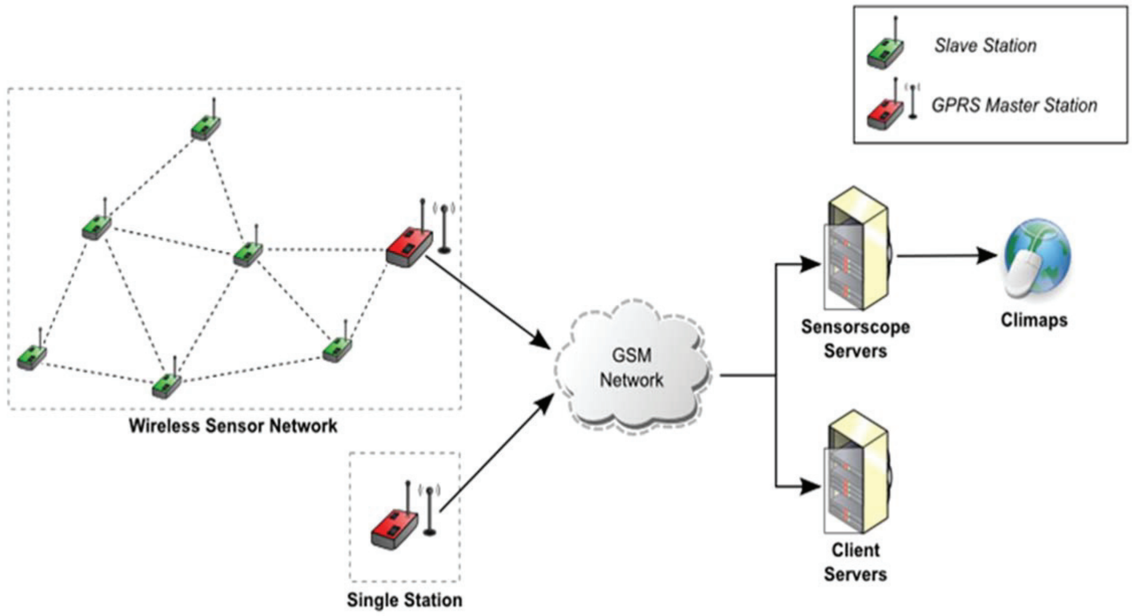


Figure 1: An example of wireless sensor network architecture (Source: http://www.sensorscope.ch/solution)

Widespread spatial and temporal monitoring of the environment is required for drawing accurate scientific conclusions and for providing reliable forecasting. Conventionally, in-situ measurement stations, which use data loggers to store the data, are utilized for environmental observations. This method is not only time consuming and tedious but also expensive to use as it requires setting up of different instruments with the data loggers and then physical accusation of data from the log (SensorScope, 2007b). Moreover, high cost of in-situ measurement systems makes it difficult to deploy a large number of sensing stations and this result in a limited spatial coverage. WSNs emerged as a solution to such issues.

Real-time communication capability between sensor nodes is another asset of WSNs. Depending on the application, there is a need to respond rapidly to what the sensor values represent. For example, disastrous events need to be detected fast to create awareness and to generate timely alarms (Bahrepour et al., 2010). Furthermore, the fact that WSNs can work under harsh conditions, such as extremely high and low temperature, very high or low humidity, wet environments, high and low pressure, and in an extremely noisy environment (Akyildiz et al., 2002) is a valuable property of sensor nodes. Therefore, the real time communication capability of WSNs and their ability to work in harsh environments should be exploited and used in environmental monitoring.

Wireless sensor nodes are smaller, lighter and cheaper compared to previous generation of sensors. However, they may have low data quality and are constrained by limited battery life and the network cost of communication. The quality of data measured by low cost and low quality sensor nodes is affected by noise or error, missing values, duplicated data or inconsistent data (Zhang, 2010). As a result, WSNs are more likely to generate outliers.

Outliers are measurements that significantly deviate from the normal pattern of the phenomenon (Chandola et al., 2007; Zhang et al., 2008). These deviations do not always signify some kind of error (either environmental noise or faulty sensor); they could also correspond to some interesting events (changes in the state of the environment) to which we should be alerted. Hence, having identified an outlier, further analysis is required to differentiate it as error or event.

Event detection using WSNs can be of great help for real time decision making and situation awareness (Bahrepour, et al., 2010) along with long-term monitoring of natural events in areas of varying size. It also contributes to remote environmental monitoring, which is crucial in harsh conditions. Remote surveillance prevents the need to have onsite staff at remote locations and can save people's time, effort and travelling cost. Therefore, event detection functionality of WSNs helps in remote environmental monitoring and it also supports real time decision making and awareness on environmental situations.

1.1. Problem statement

Traditional environmental measurement systems use expensive sensing stations which makes it difficult to deploy a large number of sensing units. This results in limited spatial coverage and lack of appropriate observations (Barrenetxea et al., 2008b). To overcome the limitations in spatial coverage by traditional environmental measurement systems, low cost WSNs can be deployed in great numbers to observe the environment with high spatial and temporal resolution (Elson & Estrin, 2004). Furthermore, traditional outlier detection techniques are computationally expensive, require much memory for data analysis and storage, use centralized approach for data analysis, analyse data in an offline manner, and do not distinguish between errors and events (Zhang, et al., 2008). These techniques are not suitable for the context of WSNs. Therefore, in order to optimize the traditional outlier detection techniques so that they can be suitable for WSNs; (near) real-time, decentralized outlier and event detection functionality of WSNs should be exploited.

To detect an event in the environment, analysing the WSN data is important. However, WSNs produce large quantities of spatio-temporal data, causing challenge for online data analysis (Zhang, et al., in press). One of the key challenges in detecting an event in a WSN is how to detect it accurately and fast (real time) in an online and decentralized way (Ma et al., 2004) while maintaining the resource consumption to a minimum and providing sufficient details about the event (Banerjee et al., 2008; Martinez et al., 2004; Zhang et al., 2010). As temporal and areal coverage of an event are important factors in environmental monitoring, tracking the temporal and spatial development of an event is also important once it is detected. Hence, it is essential to develop analytical techniques of event detection and event interpolation which are able to operate under the limited WSN resources.

1.2. Research identification

1.2.1. Research objective

The general objective of this research is to detect and monitor the temporal evolution and spatial extent of events, accurately and fast (real time), using analytical techniques that are suitable for WSNs.

1.2.2. Specific objectives and research questions

In order to support the accomplishment of the research objective, five specific objectives followed by one or more research questions are formulated in this research. The specific objectives with their corresponding research questions are presented in Table 1.

Table 1: Specific objectives and research questions

Specific Objectives	Research questions
1. To define an outlier and an event	1.1. How can we define an outlier? 1.2. How can we define an event? 1.3. How can we distinguish outliers as events or errors? 1.4. When does an event become usual behaviour of an environmental variable?
2. To model the data behaviour (i.e. to model ambient temperature)	2.1. How can we use time series analysis for modelling temporal correlation to identify outliers? 2.2. How can we use geostatistical analysis for modelling spatial correlation to identify outliers? 2.3. Is it realistic to assume a constant spatial correlation over time?
3. To detect an event	3.1. How can we use spatial and temporal correlations existing in sensor data and exploit them to detect events in WSNs?
4. To characterize an event	4.1. What is the temporal evolution of an event? 4.2. What is the spatial extent of an event?
5. To evaluate the event detection	5.1. How accurately can we detect events?

1.3. Structure of the Thesis

This thesis is structured in to seven chapters, including the introductory chapter. The rest of the thesis is organized as follows. Related work in WSNs for outlier and event detection, guideline for defining outliers and events as well as fundamentals of time series and geostatistics are presented in chapter 2. In chapter 3, description of the study area and the dataset used for the experiment is presented. In chapter 4, the methodology used to undertake the analysis is explained. Results are presented in chapter 5 and discussed in chapter 6. Finally, I conclude this thesis in chapter 7 by summarizing the key results and discussing future directions.

2. LITERATURE REVIEW

2.1. Related work

Due to resource constraints of WSNs, different algorithms for event detection have been proposed considering specific application and scenario characteristics. These algorithms can be divided in to three categories: time-driven, query-driven and event-driven (Barrenetxea, et al., 2008b). In WSN applications, the time-driven and event –driven approaches are commonly used (Grenon & Smith, 2004). Time-driven algorithm initiates sensor nodes to periodically forward gathered data to the base station (e.g., (Lian et al., 2007; Wark et al., 2007)) while query-driven algorithm initiates the nodes to send gathered data only upon request from the base station. In event-driven algorithm, an alert is forwarded to the base station when a particular event occurs. Figueiredo et al. (2009) proposed the event driven model for applications which are related to infrequent events such as fire. King & Nittel (2010) present a strategy for developing and running small models called 'tiny models' on individual sensor nodes. These models are developed from much larger detailed parent models and they can capture the predictable behaviour of a phenomenon. In circumstances where the event-driven algorithm is used, the individual nodes communicate only when a particular event is observed (Martinez, et al., 2004). These methods reduce the overall data communication and focus in handling outlying values. Shi & Winter (2010) developed a model that incorporates both time-driven and event-driven algorithms and proposed that their model can be further developed to support query-drive approach.

Most of the studies of outlier and event detection using WSNs have been undertaken outside of a practical situation and have a theoretical nature. This is challenging for real world environmental monitoring in which the performance of outdoor sensor networks can be influenced by factors like harsh weather condition (Barrenetxea, et al., 2008b; Wark, et al., 2007). Martinez et al. (2006) designed a real world WSN for extreme environment and investigated that there was a difference between the actual and predicted behaviour of the sensor networks. Many fundamental environmental processes have hardly been studied due to remoteness or unattainability (Martinez, et al., 2004). Moreover, (Martinez, et al., 2004) pointed out that, in addition to the sensing, communication and computing technologies, domain knowledge is an important component in the field of environmental sensor networks and this is desirable to realistically monitor the natural environment.

Detection accuracy and energy efficiency are two important aspects in event detection using WSNs. Zheng & Gang (2011) suggested a trade-off between these two aspects in order to design an effective detection method. They pointed out that high detection accuracy requires more data collection, processing and transmission, which results in high energy consumption. The highest energy consumption in WSNs is communication or data transmission (Qu et al., 2010; Wang & Sodini, 2004). Liang & Wang (2005) proposed that executing more local processing using a single sensor node and less data transmission between sensor nodes can save energy and increase the life time of a WSN. An accurate but resource hungry outlier detection method is not suitable for WSNs.

Some research have been undertaken to detect outliers in WSNs by using spatial and temporal correlations. Shi & Winter (2010) developed a spatio-temporal data model for dynamic areal objects which support the analysis of qualitative spatial change in sequences of time. A geostatistical technique to identify outliers at single sensor nodes based on spatial and temporal analysis has been developed by Zhang (2010). He proposed a distributed and online outlier detection techniques based on quantification of spatial and temporal correlations.

2.2. Basis for defining outliers and events

The perception of outliers differs in terms of data type, application domains and detection techniques. Even though it seems that no universally accepted definition exists, some general definition that can cope with different data types and application domains are provided. For example, a classical definition that describes an outlier as "an observation, which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism", was provided by Hawkins (1980: p.1). Chandola et al. (2007) and Zhang et al. (2008), defined outliers as measurements that do not conform to the normal behavioural pattern of the phenomenon. Outliers might represent a change in the state of the environment, or errors of measurement or transcription mistakes (Webster & Oliver, 2008). Zhang (2010) identified occurrence of various types of outliers in WSNs and classifies them in to four: incidental absolute error (very short sequence of extreme high or low values), clustered absolute error (continuous sequence of incidental absolute errors), random error (observations not falling within a threshold of normal behaviour of data) and long-term error (continuous sequence of random errors). These outliers must be investigated in order to decide what to do about them which in turn helps in proper data management (Onoz & Oguz, 2003). Obvious errors can be replaced or eliminated and those representing a change in the state of the environment should be investigated further.

Domain-specific definitions for outliers are given by Záhumenský (2004) based on the standards given by World Meteorological Organization (WMO). These definitions are based on plausible value and plausible rate of change of measured values and are explained in sections 2.2.1 and 2.2.2 respectively.

Once an outlier is detected, it needs to be investigated whether it represents a change in the state of the environment or some kind of error. In order to do this, the length of sequence of detected outliers should be considered depending on the sampling rates (Yoneki & Bacon, 2005; Zhang, et al., in press). In cases where there is no a priori knowledge on incoming event, it is difficult to assign a specific sequence of detected outliers to represent an event. This is because some events may last for a short time while others may last longer. Therefore the length of outlier sequence should be determined based on the application requirements. Zhang (2010) recommended that the length of the outlier sequence be small, so that there is no delay in classifying outliers as events or errors.

To make event detection in WSNs more reliable, local collaboration between neighbouring nodes is helpful (Chongming et al., 2009). For example, Zhang (2010) defined an event as long-term error plus specific minimum number of nodes displaying the same behaviour. Therefore, if a number of consecutive outliers from one node, which signify the possibility of an event, correspond with another such observations from at least half of the neighbouring nodes this signifies the occurrence of an event (Krishnamachari & Iyengar, 2003). But it should be noted that if neighbouring nodes experience similar noisy data, which has both temporal and spatial extent, it can be misinterpreted as an event (King & Nittel, 2010). Some events can have spatial extent less than the distance between neighbouring nodes, meaning that an event can be detected by a single node only. But since single node is liable to failures, distributed collaboration among neighbouring nodes is required. Environmental conditions are spatially correlated but errors due to sensor faults are likely to be white noise (Krishnamachari & Iyengar, 2003). Furthermore, a single node cannot perform the task like event boundary detection (Ding et al., 2005).

2.2.1. Check on plausible value

The aim of the plausible value check, which is also called the gross error check, is to verify if the observations are within the acceptable range limits Záhumenský (2004). On the basis of the nature of the variable under consideration and the geographical area and time of year in which the data is taken, a credible range limit can be set.

The nature of the normal state may not necessarily be constant. For example, the nature of the normal state of ambient temperature is not constant value, but regular condition. Night time temperatures are expected to be lower than day time temperatures. Moreover, we may not know in advance the event type that we are looking for. In such conditions, where the level of deviation from the normal pattern is unknown and can change when the environment changes or if unwanted interferers go on and off, it is quite difficult to set a fixed threshold (Liang & Wang, 2005). Observations that fall within the threshold but are inconsistent with the successive observations cannot be detected as outliers. Therefore, we cannot define an outlier based on a fixed predefined threshold when we do not have much priori knowledge on incoming event (Chongming, et al., 2009).

Therefore, for applications where outliers cannot be defined based on a specific threshold, the plausible value check can only be used to detect specific types of outliers, absolute errors. For example, Zhang (2010) used this technique, which he called fault detection (or absolute error check), to identify errors using a pre-defined threshold.

2.2.2. Check on a plausible rate of change

The aim of the check on plausible rate of change, which is also called the time consistency check, is to verify the rate of change of instantaneous data Záhumenský (2004). The check is best applicable to data of high temporal resolution (a high sampling rate) as the correlation between the adjacent samples increases with the sampling rate. It checks if there are sudden jumps in values or 'dead band' caused by faulty or blocked sensors over a specific period of time. This check is done based on minimum and maximum variability of instantaneous values.

2.2.2.1. Minimum variability

The check on minimum required variability of instantaneous values is called a persistence test. In order to undertake the minimum variability check, measurement of a parameter has to be done for a specific period (Záhumenský, 2004). If the values do not vary over the specified period by more than a specific limit, all observations from the period are considered as outliers. The minimum required variability and the specific time limit are dependent on some factors as variable type, geographic location and time. Based on the WMO guidelines, Záhumenský (2004) provided possible limits of minimum required variability for some environmental variables. For example, 0.1ºC for ambient temperature, 1% for relative humidity, 10 degrees for wind direction, 0.5m/s for wind speed and so on. These limits are used to check for minimum variability of one-minute values over a period of at least 60 minutes.

2.2.2.2. Maximum variability

The check on maximum allowed variability of an instantaneous value is also called a step test. It checks if a current instantaneous value varies from the prior one by more than a specific limit (Pawlowski et al., 2009). If the specified limit is exceeded, the observed value is considered an outlier. For example, Záhumenský (2004) proposed an algorithm for calculating the maximum variability. The formula which compares the current instantaneous observation with the previous and the next ones is shown in equation 1. If the condition in equation 1 is fulfilled the current observation is considered an outlier. In cases where either the previous or the next observations are missing, or when we want to detect an outlier based on previous observations only; Záhumenský (2004) proposed omitting the corresponding part of the formula and changing the comparison term to $2.\sigma_v$. Similarly, (Basu & Meckesheimer, 2007) proposed two-sided and one-sided median methods for detecting outlier and replacing them with credible values.

$$|V_i - V_{i-1}| + |V_i - V_{i+1}| > 4.\sigma_v \tag{1}$$

Where: V_i is the current observation,
V_{i-1} is the previous observation,
V_{i+1} is the next observation, and
σ_v is the standard deviation calculated at least from the last 10 minute period.

2.3. Fundamentals of time series and geostatistics

2.3.1. Time series

A time series is a sequence of data points that follow non-random order. Time series data have a natural temporal ordering in which successive observations are dependent. The analysis of time series, unlike the analysis of random samples of observations, takes the time order of observations in to account (Chatfield, 2004). In order to reduce irregularities in time series data, smoothing techniques are used. A moving average is the most common type of smoothing technique (Easton & McColl, 1997), which is used to analyse a set of data points by creating a series of averages of different subsets of the full data set. The moving average can be obtained by taking the average of the first subset. Then, the subset is shifted forward creating a new subset which is averaged. This process is repeated over the full data set and the plot line which connects all the averages is the moving average. Simple moving average (SMA) approach uses equal weights for each data value in the subset and calculates the mean. Equation 2 shows the formula for calculating the average of the first subset using SMA. For calculating the successive averages, there is no need of computing the full summation each time. An alternative formula for calculating the successive averages based on the average of the previous subset is presented in equation 3. This is done by including the new value and excluding the last value of the previous subset. A moving average may also use unequal weights for each data value in the subset to emphasize particular values. Weighted moving average (WMA) and exponential moving average (EMA) are examples of moving average which give unequal weights to data values in the subset.

$$SMA = (P_m + P_{m-1} + \cdots + P_{m-(n-1)})/n \tag{2}$$

Where: P_m is the last data value in a subset, and
n is the size of the smoothing window.

$$SMA_{current} = SMA_{previous} - P_{m-n}/n + P_m/n \tag{3}$$

In estimating the underlying trend using SMA, the mean value of a given set of distribution can greatly be affected by the extreme values. Moving median (MM) approach is believed to give a more robust estimate of the trend. The formula for using MM over n time points is shown in equation 4.

$$MM = Median\ (P_m,\ P_{m-1}, \dots,\ P_{m-(n-1)}) \tag{4}$$

Where: P_m is the last data value in a subset, and
n is the size of the smoothing window.

2.3.2. Geostatistics

Geostatistics deals with spatially autocorrelated data. That is, each data value is associated with a location in space and there is an inferred relationship between the location and the data value. A function that describes the degree of spatial dependence of a spatial random field is called a variogram. Webster & Oliver (2008: p 55) described the variogram as "the cornerstone of practical geostatistics". As a result, in order to model spatial correlation of data using geostatistics, we need to estimate and model the variogram. Geostatistical analysis for modelling spatial correlation include calculating a sample variogram from the data, fitting a model to the sample variogram and predicting at unsampled locations (Webster & Oliver, 2008).

The sample variogram represents the relationship between the semivariances and the corresponding separation distances of two observations. It is calculated by averaging half the squared difference of the values over all pairs of observations with the specified separation distance. The usual computing formula for the sample variogram is shown in equation 5.

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{s=1}^{n(h)} (y_s - y_{s+h})^2 \qquad 5$$

Where: s is vector of spatial coordinates,

h is the lag distance representing separation between two spatial locations,

y_s is variable under consideration as a function of spatial location,

y_{s+h} is lagged version of variable under consideration,

s is the different spatial locations of the data, and

$n(h)$ is the number of point pairs(y_s, y_{s+h}) separated by h.

There are three parameters that characterise the variogram: sill, range and nugget. The sill is the semivariance value at which the variogram levels off. The range is the distance up to which the spatial dependence extends. In other words, the range is the lag distance at which the variogram reaches the sill value. Nugget represents variability at distances smaller than the typical sample spacing, including measurement error. The difference between the sill and nugget is referred as partial sill.

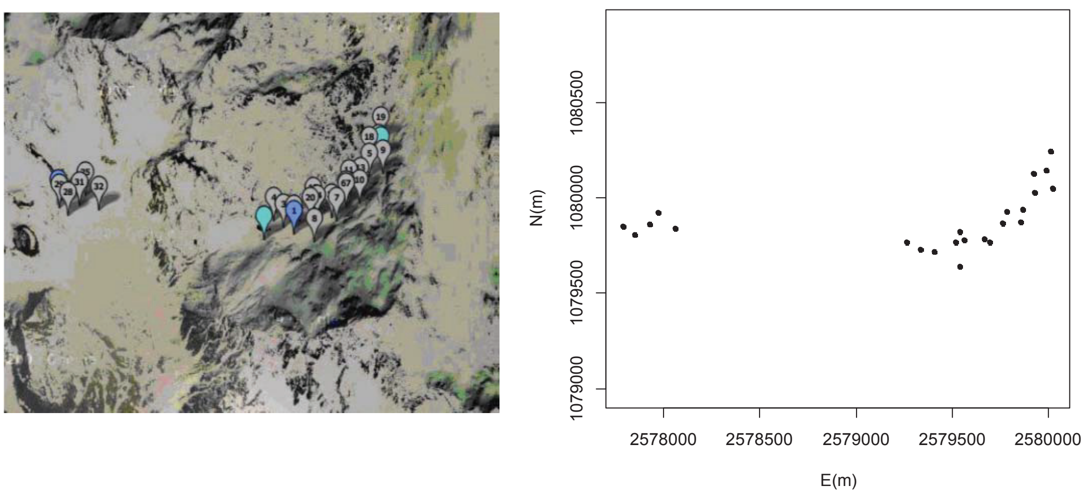
The sample variogram summarizes the spatial relations in the data by calculating semivariances at particular lags. However, the true variogram that represent the variance of a region is continuous. Additionally, in order to estimate or predict values at each location (or to undertake kriging), semivariances at lag distances other than those used in the empirical variogram are required. Therefore, in order to describe the variance of the region and predict at unsampled locations, we need to fit a continuous function to the sample variogram. In that case, we have to choose from a palette of authorized variogram models which are monotonically increasing (this may include fluctuation or periodicity) and have a constant or asymptotic sill as well as non-negative y-intercept (nugget). Variance may continue to increase as the region is expanded. The intrinsic hypothesis (intrinsic stationarity) therefore, assumes stationarity of the increment in cases where the variogram increases without bound. In general, authorized variogram models are those that ensure non-negative variances and represent either bounded or unbounded variation (Webster & Oliver, 2008). Variogram models which represent bounded variation may reach the sill at a specific range (example, bounded linear and spherical models) or approach the sill asymptotically (example, exponential and Gaussian models) and are called transitive models. Variogram models which represent unbounded variation have no sill and range (example, power model) and are called non-transitive models. Webster & Oliver (2008) recommended using weighted least-squares approximation for fitting plausible models. Additionally they suggested using statistical criteria for comparing the deviations between the observed semivariances and the ones expected from the model.

3. STUDY SITE AND DATA DESCRIPTION

The SensorScope dataset of the Grand St. Bernard pass deployment is investigated in this research. SensorScope is an interdisciplinary project developed by environmental and networking researchers (Ingelrest et al., 2010). It is based on a wireless sensor network with built-in capacity to produce high temporal and spatial density measures. It has already been successfully deployed multiple times in various environments (e.g., mountainous, urban) one of which is the Grand St. Bernard pass deployment. The Grand St. Bernard pass, which is the most ancient pass through the western Alps, is situated between Switzerland and Italy at an elevation of 2469 m with coordinates 45º52'08" N, 7º10'14" E.

A WSN consisting of 23 sensor nodes was deployed in a 600m by 2200m area at the Grand St-Bernard pass. The sensing capability of these sensors usually ranges from -20ºC to 60ºC (SensorScope, 2007a) and the precision is ±0.3ºC (Ingelrest, et al., 2010). The nodes communicate in a multi-hop fashion and communication range goes up to 1.5 km (SensorScope, 2007a). The nodes were deployed in two clusters: a big cluster consisting of 18 nodes and a small cluster consisting of 5 nodes. The two clusters were separated by approximately 1200 m.

Figure 2 illustrates distribution of the nodes in the study area and their corresponding coordinates according to the Swiss grid. These nodes measured different environmental attributes from September 13, 2007 to October 26, 2007 with sampling frequency of 2 minutes (SensorScope, 2007a). One of the different environmental attributes, the ambient temperature was analysed in this research. Data collected on 29th and 30th September 2007 from 06:00 to 14:00 was used for the experiment. According to the Federal Office of Meteorology and Climatology Switzerland, the historical data range of temperature in the Swiss Alps for the months of September and October of the last 10 years ranges from -5ºC to 25ºC (Meteosuisse, 2005 - 2011).



(a) (b)

Figure 2: (a) The distribution of the nodes in the study area(SensorScope, 2007a) (b) the corresponding coordinates according to the Swiss Grid

4. METHODOLOGY

4.1. Data preprocessing

In order to reduce potential corruption of processed data, pre-processing was done to detect and eliminate obvious outliers. This was done based on the plausible value check and the minimum variability check. The values that fail these tests were not used for further data analysis.

In order to do the plausible value and minimum variability checks, domain specific definitions were given to outliers based on the standards discussed in section 2.2.1 and section 2.2.2.1 respectively. These standards in relation to the variable of interest (i.e., ambient temperature) and the study site (i.e. the Grand St. Bernard pass) were used to define obvious outliers.

4.1.1. Plausible value check

In section 2.2.1, a maximum and minimum limit for gross error check in ambient temperature is provided based on the WMO guidelines. This is a broad and general range and was narrowed by considering the sensing capability of the sensors that were deployed in the Grand St. Bernard pass. This was further narrowed by considering the climatological conditions of the site and the season. As a result, the acceptable range limit was set to be a minimum of -5ºC and a maximum of 25ºC. Observations that exceed this limit were detected as outliers and were not used for further analysis.

4.1.2. Minimum variability check

In order to check if there was a 'dead band' caused by sensor failure, a check on minimum variability was done. As specified in section 2.2.2.1, the threshold value for minimum variability check of ambient temperature, adopted from the conformance criteria of WMO, is 0.1ºC over a period of 60 minutes. However, the precision of the sensors, which were deployed at the Grand St. Bernard pass, is ±0.3ºC (Ingelrest, et al., 2010). As a result, instead of checking for less than 0.1ºC variability, it was checked if there was no variability in the sensor data for a period of 60 minutes. Observations that do not vary over the specified time were considered as outliers. These kinds of outliers are most likely errors due to sensor failure. Sensor nodes that fail the minimum variability test were not used for further analysis.

After eliminating obvious outliers, the temperature measurements on the specified dates (29th and 30th September 2007) range from -1ºC to 11ºC. The next step was detecting outliers that fall within this range but are inconsistent with the successive observations. This was done based on the maximum variability check and the moving average. Limit of maximum variability for ambient temperature according to WMO is 3ºC in 10 minutes (Záhumenský, 2004). Observations that vary by more than this maximum variability limit were detected as outliers. This method is discussed in detail in section 4.2.1.1. Similarly, outliers were also defined based on a specific confidence level by creating a series of averages of different subsets of the full data set. When an observation in the subset lay outside the specified confidence level, then it was detected as an outlier. This method is discussed in detail in section 4.2.1.2.

4.2. Temporal and spatial correlations in WSNs

In this section, we describe how temporal and spatial correlations existing in WSN data were used for defining the normal behaviour of temperature data. Then the measurements that significantly deviate from the normal pattern were identifying as outliers. Temporal correlations were based on time series analysis and spatial correlations were based on geostatistics.

4.2.1. Time series analysis to identify temporal outliers

Uninterrupted series is required for time series analysis (Zhang, et al., in press). As a result, some missing values were generated using median smoothing (Basu & Meckesheimer, 2007). Temporal outliers were detected using two different techniques: maximum variability check and moving average.

The analysis was done based on the data from the nodes in the small cluster (i.e. nodes 25, 28, 29, 31 and 32). Figure 3 illustrates distribution of these nodes and their corresponding coordinates according to the Swiss grid. Data from these five nodes on September 30, 2007 from 6:00 to 14:00 was used. With a sampling frequency of 2 minutes, 240 observations from each node were used for the analysis.

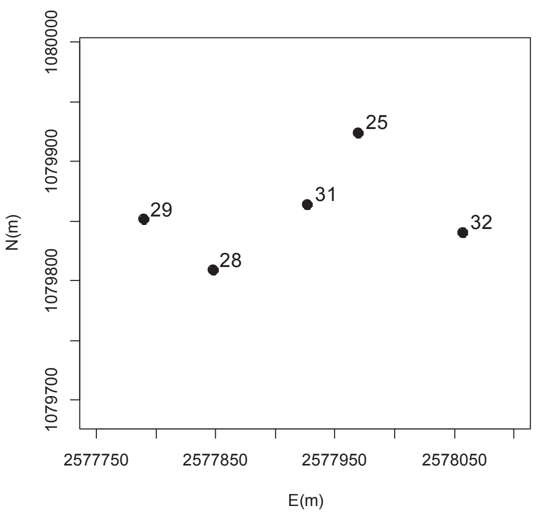


Figure 3: Distribution of the nodes in the small cluster

R (R Development Core Team, 2011) version 2.14.0, with some packages, was used for the analysis. R is an open-source language and environment for statistical computing and graphics. It is an integrated suite of software facilities for data manipulation, calculation and visualization. Packages are collections of R functions, data, and compiled code in a well-defined format. Besides the 'stats', which is a standard package which comes with R, other packages as 'zoo' (Zeileis & Grothendieck, 2005), 'xts' (Ryan & Ulrich, 2011) and 'TTR' (Ulrich, 2011) were loaded in to R and used for the analysis.

4.2.1.1. Temporal outlier detection using maximum variability check

Temporal outliers were detected by checking for unrealistic jumps in values over a specific period. The rate of change of instantaneous data was examined based on the specified maximum variability limit (i.e., 3ºC in 10 minutes). In our dataset, where the sampling frequency is 2 minutes, 6 consecutive observations represent 10 minutes data. A current reading was checked against the previous 5 observations. When the difference to the previously recorded readings is significant (i.e. if it varies by more than 3ºC), then it is labelled as an outlier.

4.2.1.2. Temporal outlier detection using moving average

Time series data have a natural temporal ordering. This makes time series analysis distinct from other common data analysis problems, in which there is no natural ordering of the observations. A time series model will generally reflect the fact that observations close together in time will be more closely related than observations further apart. In cases where a time series model is calibrated based on data, the model will no longer apply if the weather conditions change and this result in an inaccurate future value prediction. Hence, if there is consistent change in the weather conditions, the time series model needs to be re-calibrated and updating the model consumes high resources in terms of memory and processor (Zhang, 2010). To solve these problems, an automated approach for outlier detection, which is computationally inexpensive for real time analysis is required (Basu & Meckesheimer, 2007). A moving average is commonly used with time series data to smooth out short term fluctuations and estimate trends. Using moving average approach for outlier detection is computationally simple, can detect temporal outliers upon arrival of a new observation and no need to train a model. Furthermore, depending on the time frame (or window width), the system can automatically detect when an event becomes the usual behaviour of the environment (Basu & Meckesheimer, 2007) and no need of model updating as in case of model dependent approaches. Model updating is memory and processor intensive (Zhang, et al., in press) and do not comply with the limited WSN resources.

To model the temporal correlation of the data, the moving average approach was used. Figure 4 shows an example of a moving average with a smoothing window size, n=5. The averaging was done using simple moving average (SMA) technique, which gives equal weights for each data value in the subset. The SMA approach sets a dynamic threshold with maximum and minimum bound of a given confidence level of a specific subset. An observation that lay outside the maximum and minimum bound of a given confidence level of a specific subset was detected as an outlier.

Temporal outliers were detected using SMA of different sizes of smoothing window, n (5, 15, 30, 45, 60 and 90) for different confidence levels, CI (90%, 95% and 99%). The effect of the size of smoothing window and the value of confidence level was examined.

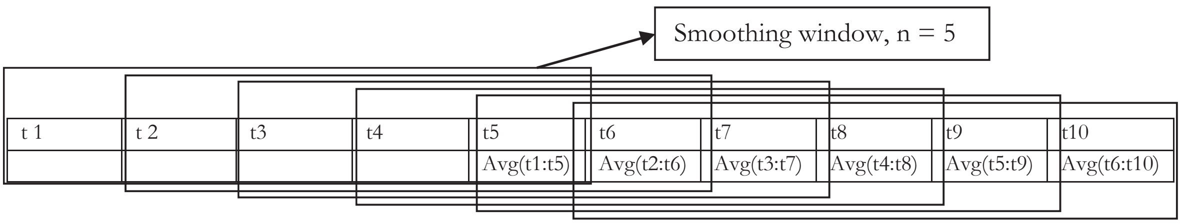


Figure 4: An example of a moving window, for n = 5 (right aligned index of results)

As shown in Figure 4, the index of the result was set to be right aligned in relation to the moving window of observations, meaning that the averaging was done by using only past data. By doing so, there is no need to wait for some more observations to classify an observation as an outlier or normal. In this case an observation is classified as an outlier or normal based on the mean of the previous n observations and the confidence level. As a result, the classification starts from the n^{th} observation and the first n-1 observations are not classified as outliers or normal. For example, if the size of smoothing window for the moving average is set to 5 as in Figure 4, the first 4 observations are not classified as outliers or normal and the classification starts from the 5^{th} observation. Since the number of unclassified data differs with different n, this makes it difficult to assess the effect of the size of the moving window. In order to solve this problem, n-1 observations from the previous hour/s were taken and all the 240 observations (from 06:00 to 14:00) from each node were classified as outliers or normal.

In estimating the underlying trend using SMA, the mean value of a given set of distribution will greatly be affected by the extreme values. As a result, moving median (MM) approach, which is believed to give a more robust estimate of the trend, was also used.

4.2.2. Accuracy of temporal outlier detection

In order to assess the accuracy of the outlier detection techniques discussed in section 4.2.1; there is no available reference dataset. Therefore, a labelled dataset from Zhang (2010), where every observation had been labelled as outlier or normal, was used as a reference dataset. This labelled dataset is obtained by a posteriori labelling using 3 techniques: running average-based, Mahalanobis distance-based and density based. Zhang (2010) provides a detailed description of the labelling techniques.

The outlier detection techniques were assessed using detection rate ($DR_{outlier}$) and false positive rate ($FPR_{outlier}$) based on the three labelled datasets. $DR_{outlier}$ represent the percentage of outliers that are correctly detected. $FPR_{outlier}$ represent the percentage of normal data that are incorrectly detected as outliers. The formulas for $DR_{outlier}$ and $FPR_{outlier}$ are presented in Equation 6 and Equation 7 respectively.

$$DR_{outlier} = (CO/TO) \times 100 \quad 6$$

$$FPR_{outlier} = (IO/TN) \times 100 \quad 7$$

Where: CO is the number of correctly detected outliers,
TO is the total number of true outliers,
IO is the number of normal data incorrectly detected as outliers, and
TN is the total number of normal data.

4.2.3. Geostatistical analysis for modelling spatial correlation to identify outliers

Geostatistical analysis for modelling spatial correlations was used to identify spatial outliers. This includes calculating a sample variogram from the data, fitting a model to the sample variogram and predicting at unsampled locations (Webster & Oliver, 2008).

The analysis was done based on the data from 06:00 to 14:00 on September 29, 2007. All the nodes from both the small and big cluster were considered. However, 6 nodes from the big cluster were detected to be faulty in the pre-processing and are excluded from the analysis. Therefore, measurements from the remaining 17 nodes were considered for the analysis.

In order to compute sample variogram, generally from 100 to 200 data point is required and those calculated with less than 50 are liable to sudden unpredictable change (Webster & Oliver, 2008). Consequently, the spatial data from these 17 locations were too few to represent spatial variation of sensor data. As a result, the usual method for calculating sample variogram, which is presented in equation 5, could not be used. To mitigate this constraint regarding required sample size, observations at different time periods from the 17 locations were used (Sterk & Stein, 1997). Accordingly the formula for variogram in equation 5 was modified to equation 8.

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{t=1}^{m} \sum_{s=1}^{n_t(h)} (y_{s,t} - y_{s+h,t})^2 \qquad 8$$

Where: m is the number of different time periods, and
$n_t(h)$ is the number of point pairs $(y_{s,t} - y_{s+h,t})$ separated by h at time period t.

The variogram averaging method of Sterk & Stein (1997), which is presented in equation 8, is based on the assumption that a constant spatial correlation exists over time. To check if this assumption is realistic, spatial variations for different hours and for the whole morning were modelled.

R (R Development Core Team, 2011) version 2.14.0, with 'gstat' (Pebesma, 2004) and 'sp' (Bivand et al., 2008a; Pebesma & Bivand, 2005) packages, was used for the analysis.

4.2.4. Accuracy of spatial correlation model

To evaluate the prediction accuracy of the spatial correlation model, leave-one-out cross validation technique (Webster & Oliver, 2008) was used. Temperature values at 6:30, 7:30… 13:30 were left out from the data and the hourly variograms and the average variogram were estimated. The left out values were then predicted using the respective hourly variogram models and the average variogram model. For example, the left out values at 06:30 were predicted using the hourly variogram model for 6:00 ~ 7:00 and the average variogram model for 06:00 ~ 14:00. Subsequently, in order to compare prediction accuracy between the hourly and the average variogram models, the RMSE was determined. In order to determine the RMSE, the residuals (i.e., the differences between the left out values and the corresponding predicted values) were calculated as in Equation 9. Then ME and RMSE are calculated as in Equation 10 and Equation 11 respectively.

$$e_i = x_i - \hat{x}_i \qquad 9$$

$$ME = \sum_{i=1}^{N} e_i / N \qquad 10$$

$$RMSE = \sqrt{\sum_{i=1}^{N} {e_i}^2 / N} \qquad 11$$

Where: i denote a specific observation,
e denote residual,
x denote a left out value,
$\hat{x}$ denote a predicted value, and
N denote the number of residuals.

4.3. Event detection

Once an outlier is detected by identifying deviation from the normal behaviour, it was further investigated to distinguish between events and errors. This was done based on the length of the outlier sequence. Choosing a specific length of sequence of outliers for distinguishing between events and errors is not easy when we do not have a priori information about the event. Zhang (2010) recommended that the length of the outlier sequence be small so that outliers be classified as events or errors without delay. In this research, the length of sequence of outliers for defining an event was set to be four consecutive outliers (i.e., 8 minutes).

The occurrence of four consecutive outliers at a single node could also be a long term error. Therefore, when four consecutive outliers from one node, which signify the possibility of an event, correspond with another such observations from at least two of the neighbouring nodes this signify the occurrence of an event and such consecutive outliers were labelled as an event otherwise they were considered as a long term error.

An event pertains to a place of limited geographic extent and a limited time frame of the observation period. If such an extreme reading pertains for a time frame more than a specific limit and its geographic extent covers almost the entire region under consideration, it is no more an event (i.e. it becomes the usual behaviour of the environment). However, in cases where event detection is used for environmental monitoring, the focus is not in detecting a specific kind of event. As a result, it is difficult to specify a limit on time frame and geographic extent for considering an event as the normal environmental variable. For convenience the minimum limits for the time frame and geographic extent for classifying an event as normal environmental variable were set to be one fourth of the total observation period and two third of the total number of nodes respectively. It is obvious that a disastrous event, such as fire, can never be considered as normal behaviour of the environment. The specified limits in the time frame and geographic extent assume that, if such a disastrous event happens, immediate action will be taken before the specified time limit and if it continue more than the specified limit, it is less likely that the sensor nodes will survive.

The analysis was done based on the data from the five nodes in the small cluster using the outliers detected by the temporal outlier detection (TOD) method of Zhang, et al. (in press). The time sequence of the temporal outliers was first checked and when it signify the possibility of an event, it was checked if at least two of the neighbouring nodes also detect similar sequence of temporal outliers.

4.3.1. Accuracy of event detection

In order to assess accuracy of the event detection method, there was no reference dataset available. As a result, the observations which were labelled as outliers using the three labelling techniques by Zhang (2010) were relabelled in to events and errors based on the definition provided in section 4.3. Then these labelled events were used as a reference to assess the accuracy of the event detection technique.

The event detection technique was assessed using detection rate ($\mathrm{DR_{event}}$) and false positive rate ($\mathrm{FPR_{event}}$). $\mathrm{DR_{event}}$ represent the percentage of events that are correctly detected. $\mathrm{FPR_{event}}$ represent the percentage of outliers that are incorrectly detected as events. The formulas for $\mathrm{DR_{event}}$ and $\mathrm{FPR_{event}}$ are presented in Equation 12 and Equation 13 respectively.

$$DR_{event} = (CE/TE) \times 100 \qquad 12$$

$$FPR_{event} = (IE/TO) \times 100 \qquad 13$$

Where: CE is the number of outliers that are correctly classified as events,
TE is the total number of outliers that represent a true event,
IE is the number of outliers that are incorrectly classified as events, and
TO is the total number of outliers.

5. RESULTS

5.1. Temporal correlation-based outliers

The results of temporal outliers detected using the techniques discussed in sections 4.2.1.1 and 4.2.1.2 are presented in sections 5.1.1 and 5.1.2 respectively.

5.1.1. Outliers detected using maximum variability check

An observation that varies by more than 3ºC from five preceding observations was labelled as an outlier. Figure 5 shows the results of the detected outliers at each node of the small cluster where the outliers are marked as (red) solid circle. In the figure the time series of the data in each node is shown by black line. Likewise, the number of outliers and the points at which they occurred are presented in Table 2.

Table 2: Number of outliers that exceed the maximum variability limit

Nodes	Number of outliers	Occurred points
Node 25	8	220, 221, 222, 223, 224, 225, 226, 228
Node 28	9	196, 204, 205, 207, 208, 222, 223, 224, 227
Node 29	5	187, 223, 224, 228, 229
Node 31	5	203, 204, 208, 209, 226
Node 32	2	223, 224

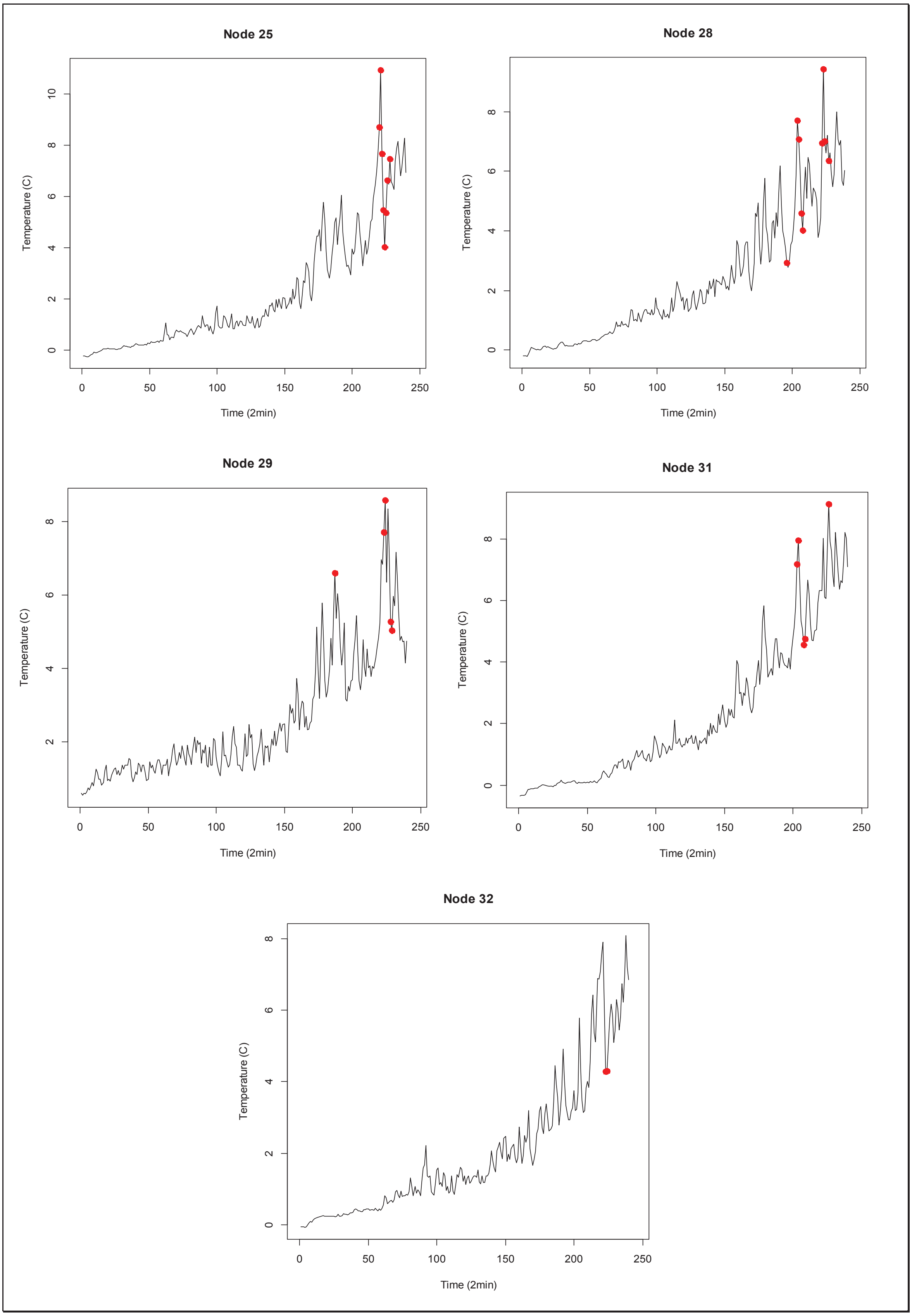


Figure 5: Outliers that exceed the maximum variability limit

5.1.2. Outliers detected using moving average

This section presents the results of outliers detected using the simple moving average (SMA) and moving median (MM) techniques described in section 4.2.1.2. In the experiments the size of the smoothing window, n was assigned different values (5, 15, 30, 45, 60 and 90) for different confidence level, CI (90%, 95% and 99%). In order to examine the effect of n and CI in the outlier detection method, all the assigned values of n for the different CI were tried. For n=5, there were no outliers detected at CI 95% and 99% using the SMA approach and at CI 99% using the MM approach. Figure 6 and Figure 7 are presented as examples to show the effect of using different n and CI. The temporal outliers detected at node 25 for n=45 at different CI by using SMA are presented in Figure 6. Likewise, the temporal outliers detected at node 28 for CI=95% at different n by using MM are presented in Figure 7. In the figures, the original time series of the data in each node is shown by black line. The dotted lines show the upper bound and lower bound of standard deviation of the corresponding smoothed time series. The data points which lie outside these bounds are labelled as outliers and are marked as (red) solid circle.

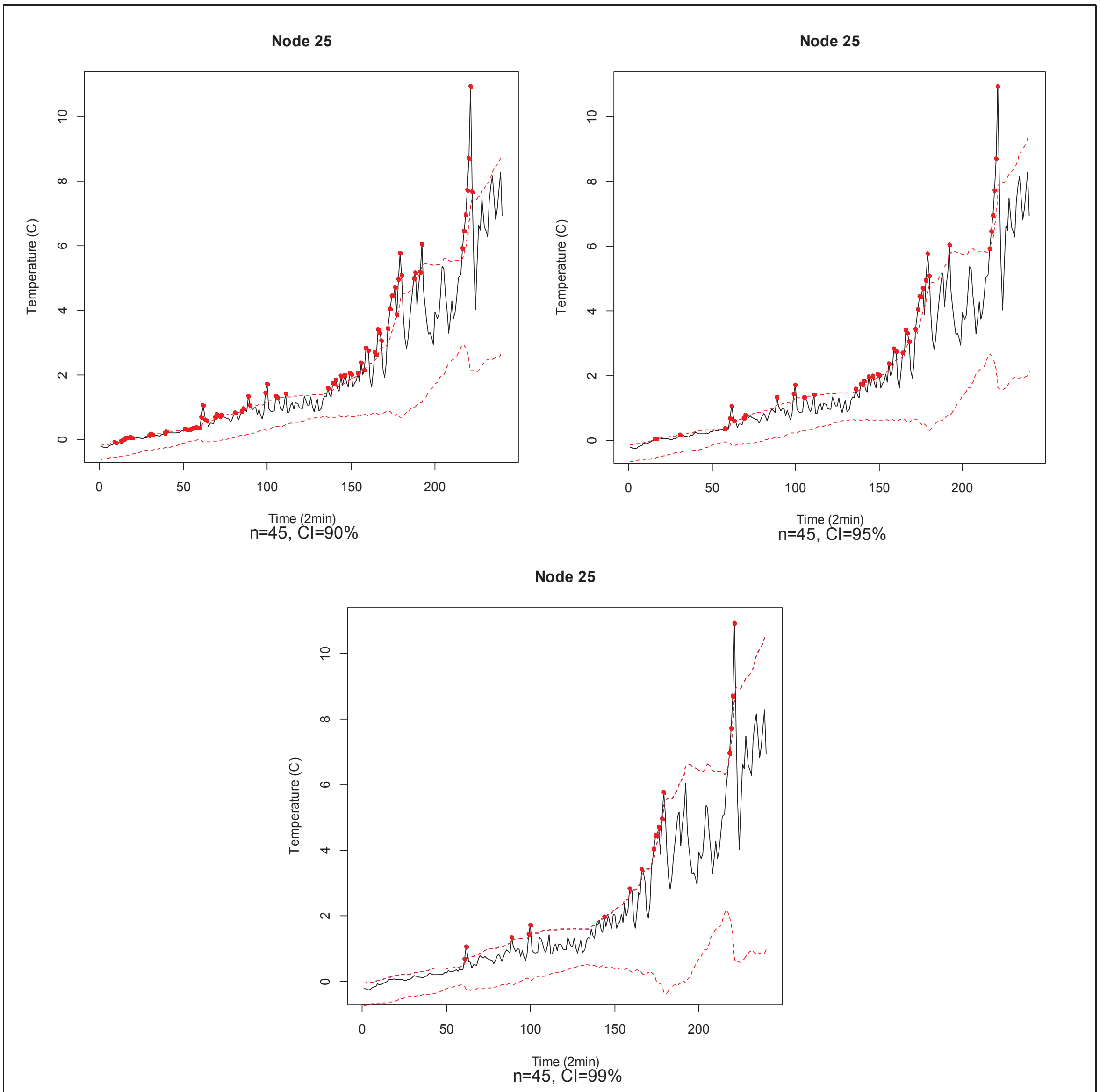


Figure 6: Temporal outliers detected at node 25 using SMA for n=45 at different CI

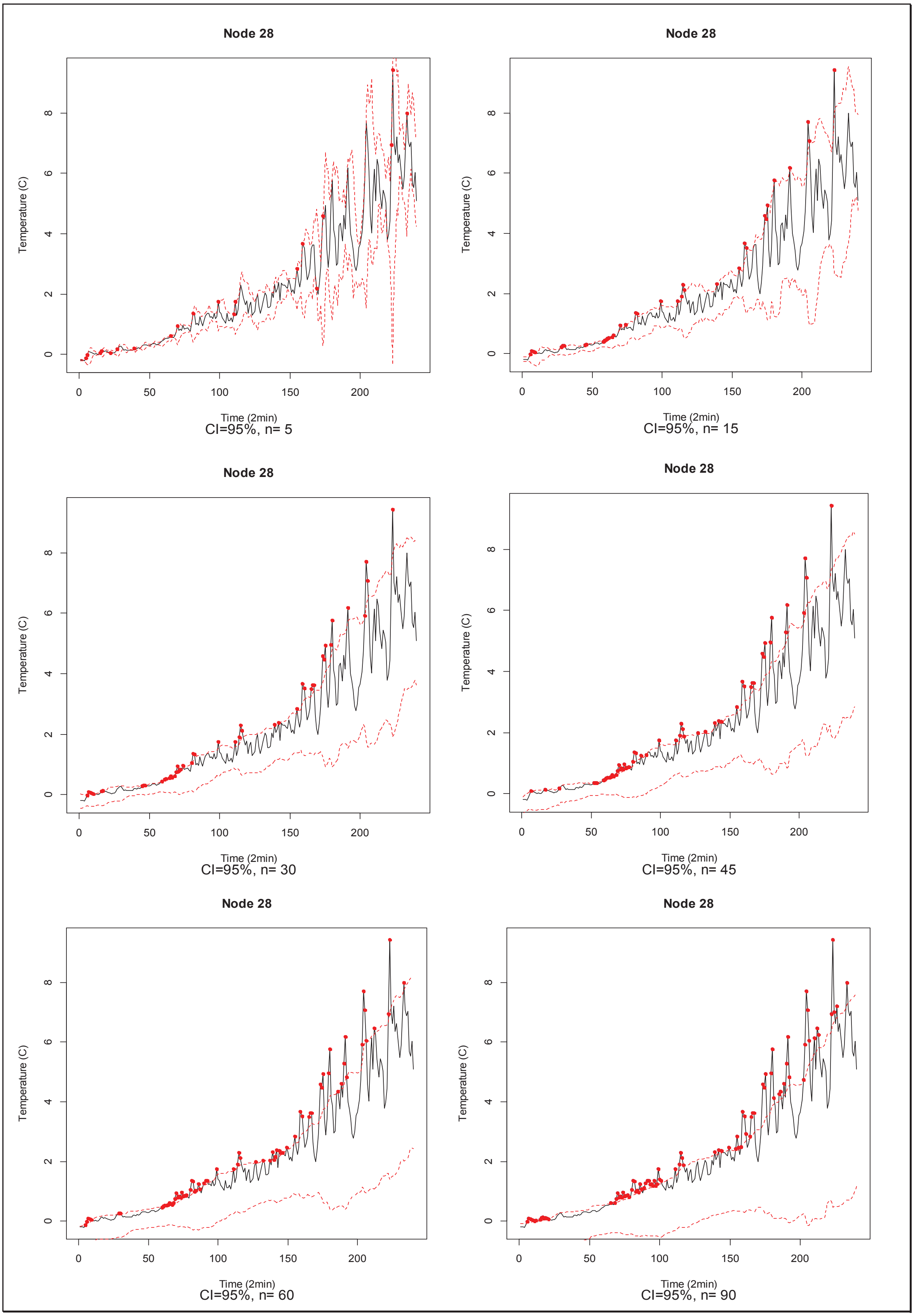


Figure 7: Temporal outliers detected at node 28 using MM for CI=95% at different n

5.1.3. Outlier detection accuracy

The outlier detection accuracy was calculated based on DR and FPR using Equation 6 and Equation 7 respectively. This was based on the results of three labelling techniques (running average-based, Mahalanobis distance-based and density-based techniques). For comparison purpose the accuracy of temporal outlier detection approach used by Zhang, et al. (in press) is presented at the last rows of Table 3, Table 4, and Table 5.

5.1.3.1. Outlier detection accuracy of the maximum variability check

The outlier detection accuracy of the maximum variability check is presented in Table 3.

Table 3: Outlier detection accuracy of the maximum variability check

n	Running average-based		Mahalanobis distance-based		Density-based	
	DR%	FPR%	DR%	FPR%	DR%	FPR%
6	19	1	44	2	73	2
TOD	72	11	100	15	100	15

5.1.3.2. Outlier detection accuracy of the moving average

The outlier detection accuracy of the SMA and the MM are presented in Table 4 and Table 5 respectively. This was done for different sizes of smoothing window and different confidence levels.

Table 4: Outlier detection accuracy of the SMA

n	CI	Running average-based		Mahalanobis distance-based		Density-based	
		DR%	FPR%	DR%	FPR%	DR%	FPR%
5	90%	6	4	5	4	20	4
15	90%	32	15	59	16	73	16
	95%	24	9	55	10	73	10
	99%	7	2	11	3	27	3
30	90%	39	22	69	22	93	23
	95%	35	15	65	15	93	16
	99%	24	5	53	5	73	6
45	90%	47	29	82	29	100	30
	95%	39	17	69	17	93	18
	99%	32	6	62	6	93	7
60	90%	57	34	88	35	100	36
	95%	46	20	79	20	100	21
	99%	34	7	62	8	93	9
90	90%	74	46	98	46	100	47
	95%	55	26	86	26	100	27
	99%	35	9	62	9	93	10
TOD		72	11	100	15	100	15

Table 5: Outlier detection accuracy of the MM

n	CI	Running average-based		Mahalanobis distance-based		Density-based	
		DR%	FPR%	DR%	FPR%	DR%	FPR%
5	90%	23	16	44	16	50	17
	95%	8	6	20	6	33	6
15	90%	35	19	63	19	73	20
	95%	31	12	60	12	73	13
	99%	15	3	45	3	53	4
30	90%	42	27	72	27	93	28
	95%	37	18	65	18	93	19
	99%	29	8	60	8	93	9
45	90%	55	33	86	33	100	34
	95%	44	20	72	20	93	21
	99%	35	7	65	8	93	9
60	90%	68	40	97	41	100	42
	95%	52	25	86	25	100	26
	99%	37	10	65	10	93	11
90	90%	83	53	100	54	100	55
	95%	66	33	97	34	100	35
	99%	39	11	67	12	93	13
TOD		72	11	100	15	100	15

5.2. Geostatistical analysis for modelling spatial correlation to identify outliers

5.2.1. Empirical variogram

The sample variogram for every one hour and for the whole morning was calculated according to equation 8. In calculating the sample variogram, the ‘gstat’ makes a number of decisions, some of which are the cut-off and width, by default. Cut-off refers to the maximum distance up to which point pairs are considered and width refers to the distance interval over which point pairs are averaged. The default value ‘gstat’ uses for cut-off is one third of the maximum possible lag (or the largest diagonal of the bounding box of the data) and cut-off value divided by 15 is used as default width by ‘gstat’ (Bivand et al., 2008b). The default cut-off and width for our data, which were provided by ‘gstat’, were 745 and 49 respectively.

Figure 8 illustrates average empirical variograms for the whole morning (06:00 ~ 14:00) at (a) the default cut-off and width, as well as at different cut-offs and widths (b) cut-off = 1000, width = 100 and (c) cut-off = 1100 and width = 150.

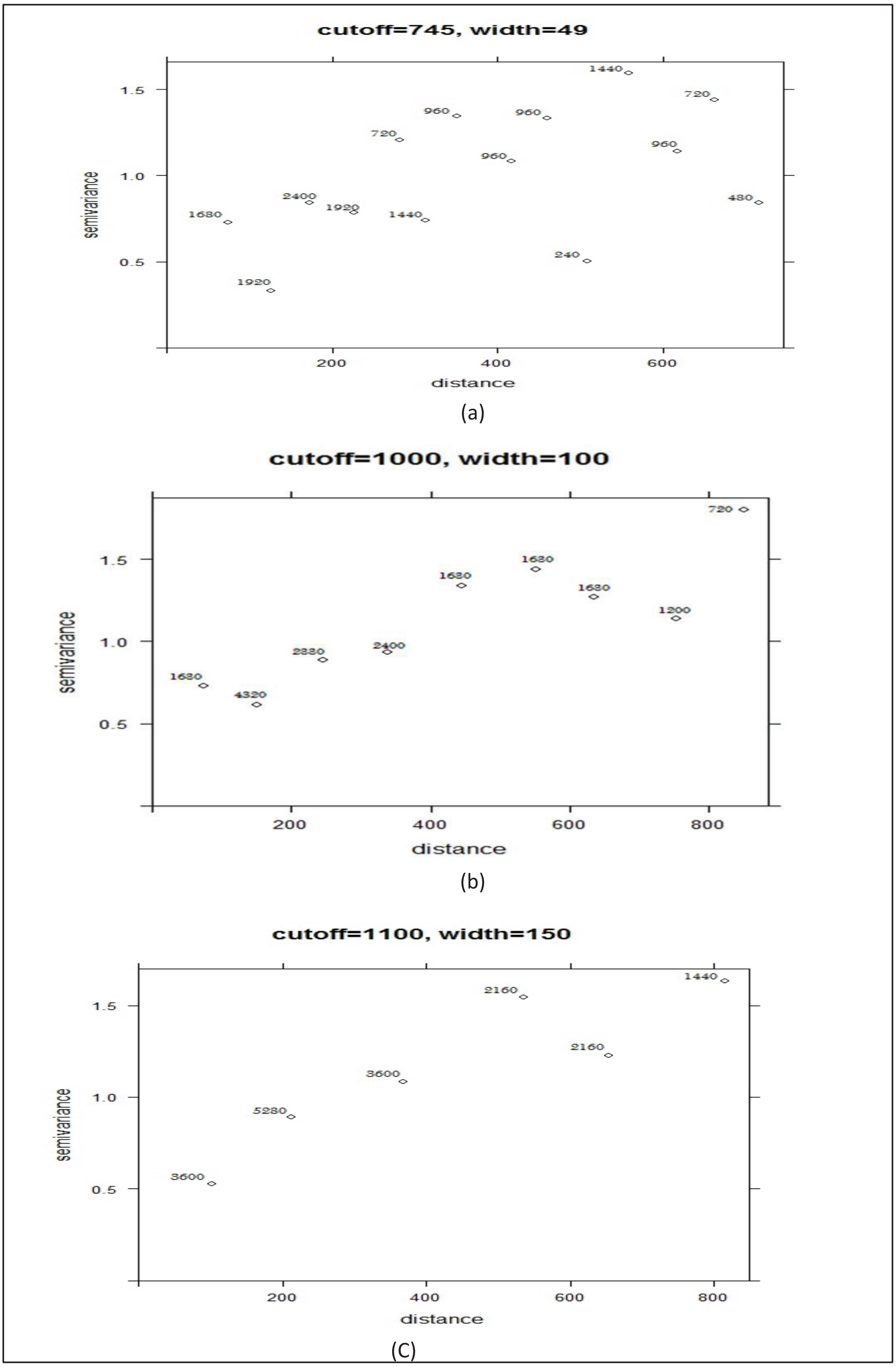


Figure 8: Average empirical variograms at different cut-offs and widths (time 6:00 ~ 14:00)

5.2.2. Fitting variogram model

After investigating the nature of the empirical variograms of every hour and the average variogram of the whole morning, appropriate variogram models were fitted. To improve the fitting to the sample variogram, 'gstat' default weighting method was used and use was made of the bounded linear and the power models which are defined in Equation 14 and Equation 15 respectively.

Bounded linear model, which is the simplest function for describing bounded variation (Webster & Oliver, 2008), was fitted to the empirical variograms of every hour and to the average variogram of the whole morning at cut-off 1000 and width 100. This is shown in Figure 9 and the estimated variogram parameter values and the SSErr are presented in Table 6.

Bounded linear model: $$\gamma(h) = \begin{cases} c\left(\frac{h}{a}\right), & h \leq a \\ c, & h > a \end{cases} \qquad 14$$

Where h is lag distance,
c is the sill, and
a is the range.

Power model, which is the simplest function for describing unbounded variation (Webster & Oliver, 2008), was fitted to the empirical variograms of every hour and to the average variogram of the whole morning at cut-off 1100 and width 150. This is shown in Figure 10 and the estimated variogram parameter values and the SSErr are presented in Table 7.

Power model: $$\gamma(h) = ch^{\propto}, \quad 0 < \propto < 2 \qquad 15$$

The power model has no sill and range; instead, it quantifies the variation by using a positive slope parameter (c) which has dimensions of the variance and a dimensionless quantity ($\propto$) which indicate how fast the variance increases. Note that the value $\propto = 1$ yields a straight line. If $\propto < 1$, the variogram is convex upward and if $\propto > 1$, the variogram is concave upward.

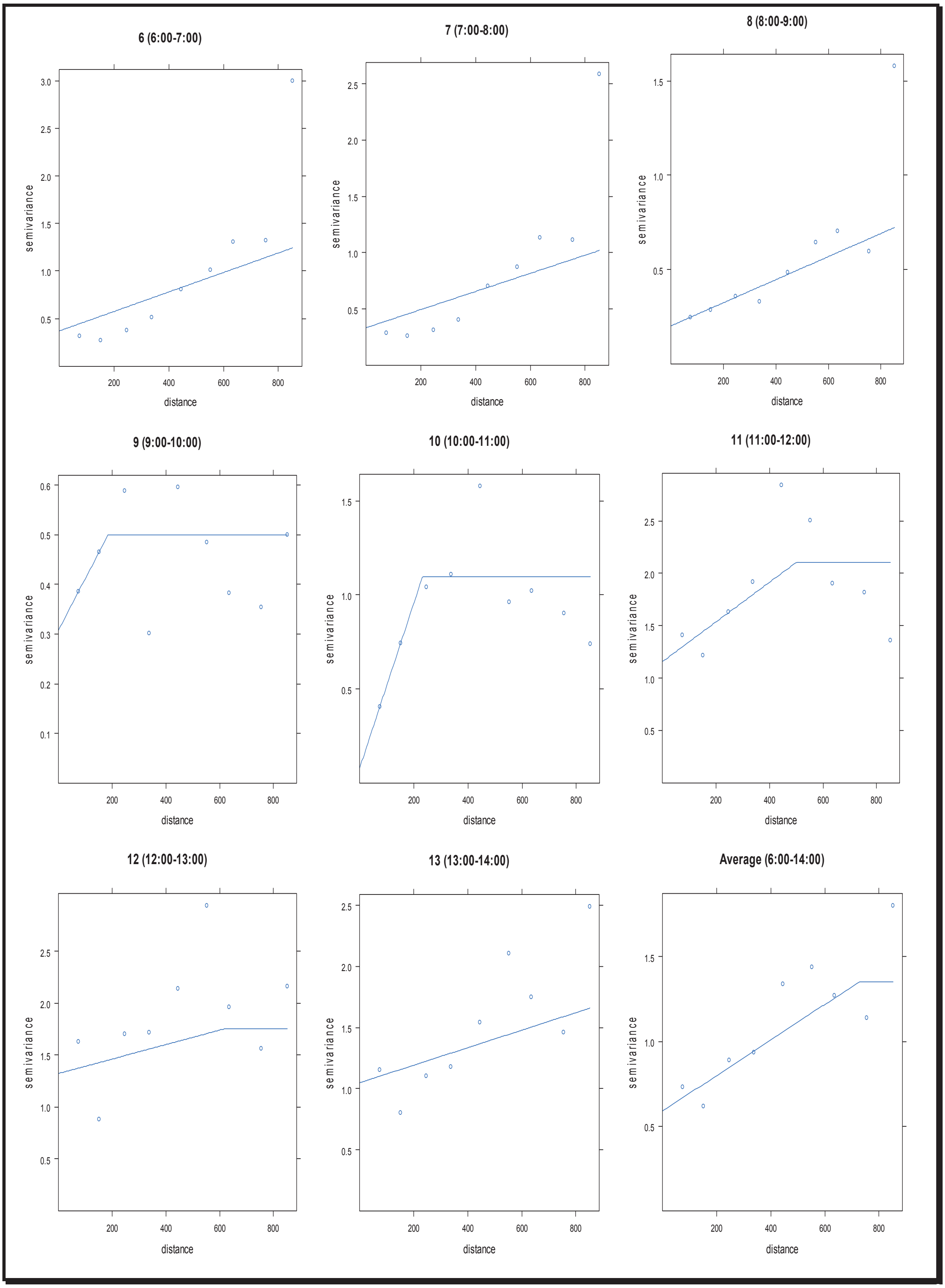


Figure 9: Bounded linear model fitted to the hourly and average sample variograms at cut off=1000 & width=10

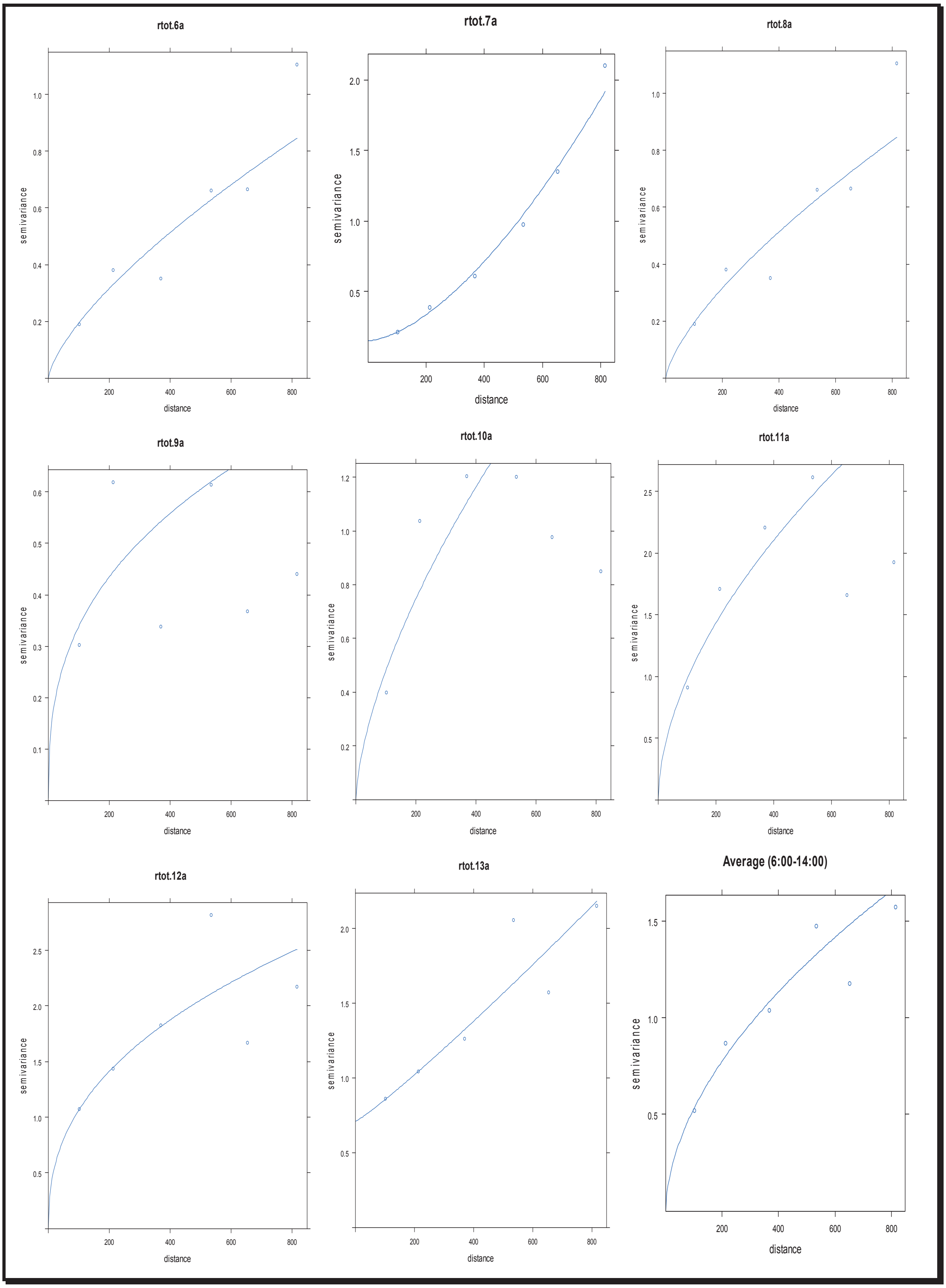


Figure 10: Power model fitted to the hourly and average sample variograms at cut-off=1100 & width=150

Table 6: Estimated variogram parameter values and SSErr for the linear model

Linear model, Cut-off=1000 and Width =100				
Hour	Partial sill	Range	Nugget	SSErr
6 (6:00-7:00)	0.881	860	0.370	3.07×10^{-3}
7 (7:00-8:00)	0.693	860	0.333	2.07×10^{-3}
8 (8:00-9:00)	0.604	988	0.201	1.23×10^{-4}
9 (9:00-10:00)	0.192	185	0.308	1.72×10^{-4}
10 (10:00-11:00)	1.015	232	0.080	3.06×10^{-4}
11 (11:00-12:00)	0.949	501	1.157	2.73×10^{-3}
12 (12:00-13:00)	0.433	621	1.322	1.12×10^{-2}
13 (13:00-14:00)	0.617	860	1.048	3.64×10^{-3}
Average(6:00-14:00)	0.764	727	0.588	5.95×10^{-3}

Table 7: Estimated variogram parameter values and SSErr for the power model

Power model, Cut-off= 1100 and Width=150				
Hour	Partial sill	Range	Nugget	SSErr
rtot.6a (6:00-7:00)	7.77×10^{-3}	0.70	0	1.21×10^{-4}
rtot.7a (7:00-8:00)	1.30×10^{-5}	1.74	0.166	2.84×10^{-5}
rtot.8a (8:00-9:00)	7.77×10^{-3}	0.70	0	1.21×10^{-4}
rtot.9a (9:00-10:00)	6.43×10^{-2}	0.36	0	7.21×10^{-4}
rtot.10a (10:00-11:00)	2.56×10^{-2}	0.64	0	1.88×10^{-3}
rtot.11a (11:00-12:00)	7.74×10^{-2}	0.55	0	2.27×10^{-3}
rtot.12a (12:00-13:00)	1.63×10^{-1}	0.41	0	7.50×10^{-4}
rtot.13a (13:00-14:00)	9.11×10^{-4}	1.10	0.709	2.34×10^{-4}
Average(6:00-14:00)	4.72×10^{-2}	0.53	0	1.71×10^{-3}

5.2.3. Prediction accuracy

In order to predict at the left-out observations, the fitted power model with the estimated parameters was applied to the data. The left-out observations were predicted using the respective hourly variogram models and the average variogram model. Table 8 shows comparison of prediction accuracy between the hourly and average models using ME and RMSE.

Table 8: Comparison of the prediction accuracy between the hourly and average models

Prediction point	Variogram model	ME	RMSE
6:30	6:00-7:00(rtot.6a)	0.0128	0.4365
	6:00-14:00(Average)	0.0092	0.4552
7:30	7:00-8:00(rtot.7a)	-0.0065	0.6389
	6:00-14:00(Average)	-0.0040	0.6424
8:30	8:00-9:00(rtot.8a)	0.0181	0.6181
	6:00-14:00(Average)	0.0142	0.6377
9:30	9:00-10:00(rtot.9a)	0.0077	0.6396
	6:00-14:00(Average)	0.0131	0.6078
10:30	10:00-11:00(rtot.10a)	0.0227	0.7488
	6:00-14:00(Average)	0.0186	0.7488
11:30	11:00-12:00(rtot.11a)	0.0409	1.4271
	6:00-14:00(Average)	0.0395	1.4216
12:30	12:00-13:00(rtot.12a)	0.0238	0.8621
	6:00-14:00(Average)	0.0354	0.8630
13:30	13:00-14:00 (rtot.13a)	-0.0030	0.8196
	6:00-14:00(Average)	0.0150	0.8884

5.3. Event detection

The event detection process was done based on the outliers detected using the temporal outlier detection (TOD) method of Zhang, et al. (in press). Based on the definition provided in section 4.3, sub section of the observed points, where occurrence of an event was identified, is presented in Table 9. First the occurrence of four or more consecutive outliers was labelled as a possible event. Then the possible events were further classified as events or long-term errors by checking if at least two out of the four neighbouring nodes also detect the same deviations.

Table 9: Detected events based on TOD method of (Zhang, et al., in press)

Occurred points	Node 25	Node 28	Node 29	Node 31	Node 32
213	Normal	Normal	Normal	Normal	Error
214	Normal	Event	Normal	Normal	Error
215	Normal	Event	Normal	Normal	Error
216	Event	Event	Normal	Normal	Normal
217	Event	Event	Normal	Normal	Event
218	Event	Normal	Normal	Event	Event
219	Event	Error	Normal	Event	Event
220	Event	Error	Normal	Event	Event
221	Event	Normal	Event	Event	Normal
222	Normal	Event	Event	Event	Event
223	Event	Event	Event	Normal	Event
224	Event	Event	Event	Normal	Event
225	Event	Event	Event	Event	Event
226	Event	Normal	Normal	Event	Event
227	Normal	Event	Event	Event	Event
228	Error	Event	Event	Event	Event
229	Normal	Event	Event	Event	Normal
230	Normal	Event	Event	Event	Normal
231	Normal	Normal	Normal	Normal	Error
232	Long-term error	Error	Error	Normal	Normal
233	Long-term error	Error	Normal	Long-term error	Normal
234	Long-term error	Error	Normal	Long-term error	Normal
235	Long-term error	Normal	Error	Long-term error	Error
236	Long-term error	Normal	Normal	Long-term error	Normal
237	Normal	Error	Normal	Normal	Error
238	Normal	Error	Normal	Error	Error
239	Error	Normal	Normal	Error	Error
240	Normal	Error	Normal	Normal	Normal

5.3.1. Event detection accuracy

The accuracy of the event detection technique was calculated based on DR (Equation 12) and FPR (Equation 13) using results of the classified outliers of the three labelling techniques. The classified outliers of the Mahalanobis distance-based and density-based techniques detect no event. Therefore, the accuracy assessment was based on the classified outliers of the running average-based technique. The results for DR and FPR are 49% and 16% respectively.

6. DISCUSSION

6.1. Temporal correlation based outliers

The maximum variability check was first computed by calculating the difference between the maximum and minimum observations in a subset of six consecutive observations. But this caused multiple observations to be labelled as outliers even when there is only one outlier in the subset. In order to mitigate this constraint, the last observation was checked against the maximum and minimum of the previous five observations. This reduces the effect of a single outlying value in labelling other normal observations as outliers. However, this cannot be fully eliminated. Extra investigation in this approach is required.

As it can be seen from Table 3 the maximum variability check resulted in a very low FPR. The DR was relatively higher when the Density-based labelling technique was uses to assess the accuracy of the maximum variability check. However, when running average-based and Mahalanobis distance-based labelling techniques were used, the DR was low. According to WMO guidelines, the maximum allowed variability for air temperature over 10 minutes period is 3ºC (Záhumenský, 2004). In the maximum variability check approach, this limit was used to detect unrealistic jumps in values. However this limit is a general one and could be too big for detecting outliers in areas where the temperature variability is very low. And it is possible that this could be the reason for getting low DR. Using a lower limit could improve the DR while still keeping the FPR to a minimum.

The effect of using medians (MM) instead of means (SMA) was examined. When MM approach was used instead of SMA approach, the confidence intervals were based on a symmetric distribution and were the same as in the case of moving mean. However, the standard deviations for the MM were from the median not from the mean and a different result from that of the moving mean was obtained. As it can be seen from Table 4 and Table 5, taking the same size of smoothing window (n), the percentage of outliers that were correctly detected was higher when medians were used instead of means. This is because in case of SMA approaches, the mean value of a given set of distribution was greatly affected by the extreme values. As it can be seen from Table 4 and Table 5, this resulted to a better DR in MM than in SMA. However, the FPR in the MM was also higher than in the SMA. When emphasis is given to DR, it can be concluded that using the MM can give more accurate results than the SMA.

The accuracy of the outlier detection techniques in relation to the effect of the different labelling techniques is examined. As it can be seen from Table 4 and Table 5, the FPR for a specific outlier detection approach (with a specific size of smoothing window, n at specific confidence level, CI) were similar for all the three labelling techniques. The DR was higher when density-based labelling technique was used to assess the accuracy of the outlier detection approaches. However, a bit higher false positive rate was found when using the density-based labelling technique compared to using the running average based and Mahalanobis distance based. According to the running average based labelling, the moving average approaches had low DR for the different sizes of smoothing window, n and different confidence level, CI. The running average based labelling did not use the full range of values in the dataset. It rather used the surrounding values on the time axis (Zhang, 2010). This is a possible reason for the low accuracy of outlier detection when the running average-based labelled data was used for assessing the accuracy.

The effect of the value of confidence level in detecting temporal outliers was examined. As it can be seen from Figure 6, relatively lower CI resulted in the detection of large number of outliers. Because of the low confidence interval, some normal observations were considered as outliers. This resulted in high DR and high FPR as it can be seen from Table 4 and Table 5. On the other hand, higher CI resulted in the detection of small number of outliers. Because of the high confidence interval, some outliers were considered as normal. This resulted in lower DR and lower FPR.

The effect of the size of smoothing window in detecting temporal outliers was examined. As it can be seen from Figure 7, relatively small size of smoothing window can keep the original data structure and this results to the detection of relatively small number of outliers. Even no outlier was detected in all the nodes when small size of smoothing window was used with higher confidence level. As it can be seen from Table 4 and Table 5, small size of smoothing window results in low detection rate based on all the three labelled datasets. From this we can conclude that using small size of smoothing window, which keeps the original data structure, has a negative impact in detecting outliers. In the contrary, large size of smoothing window changes the original data structure resulting to the detection of large number of outliers. This in turn resulted in higher DR based on all the three labelled datasets. But the FPR was also higher for large size of smoothing window.

There is a trade-off in choosing the size of smoothing window and the value of confidence level. If our main emphasis is in acquiring high DR, we can use large size of smoothing window (n) while keeping the CI level low. On the other hand if we give emphasis to acquiring low FPR, we can use small size of smoothing window (n) at higher CI.

The analysed dataset has an upward trend and the used moving average approaches do not take trend in to consideration. Furthermore, the index of the results was set to be right aligned in comparison to the moving window of observations. An example of a moving window with right aligned index of results is provided in Figure 4. When the index of results is right aligned, there is no need to wait for more observations in order to classify an observation as an outlier or normal. On the other hand, as it can be seen from Figure 6 and Figure 7, this approach has a drawback that it can only detect outliers above the upper bound in a dataset with an upward trend. By taking small size of sliding window at a lower CI, we can detect outliers below the lower bound of confidence level. However, as discussed earlier small size of smoothing window has a negative impact in detecting outliers. Setting the alignment of the index of the results to centre in comparison to the moving window of observations could alleviate the problem with the right aligned index of the results, but of course there is some time delay (depending on the size of the smoothing window) in classifying an observation as an outlier or normal. The accuracy of the moving average approaches when the index of the results was aligned to centre is not investigated in this research.

The accuracy of the moving average approaches was compared to Zhang's (2010) temporal outlier detection (TOD) technique. The DR of the moving average approaches gave comparable results to Zhang's (2010) TOD technique when Mahalanobis distance-based and density-based labeling techniques were used. However, when running average based labeling technique was used, the detection rate of the moving average approaches was much lower. FPR of both the SMA and MM approaches using all labeling techniques were much higher compared to Zhang's (2010) TOD technique. Hence, for the further analysis on event detection, the outliers detected using Zhang's (2010) TOD technique were used.

6.2. Geostatistical analysis for modelling spatial correlation to identify outliers

The sample variograms calculated by using the default cut-off and width were quite noisy. As a result it was difficult to fit any model to them. Bivand, et al. (2008b) suggests that the default cut-off and width values, when not appropriate, can be overridden and replaced with different values. To find a sample variogram with a clear trend and less noise, different computations were tried by changing the cut-off and width values. First, it was tried by increasing the default cut-off and width values to 1000 and 100 respectively. This result to a less noisy variograms compared to those calculated at the default cut-off and width values and as shown in Figure 9 a bounded linear model was fitted to them. Bounded linear model, which consists of two straight lines as described in equation 14, is the simplest function for describing bounded variation (Webster & Oliver, 2008). The bounded linear model was chosen because it is simple and can fit to the computational limitations of WSNs. In order to improve the fitting to the sample variogram, gstat default weighting method, which uses weights, N_h/h^2 with N_h the number of point pairs and h the distance, was used. The fitted variogram model had a logical and numeric attributes. The logical attribute indicates whether the model converged or ended in a singularity and the numeric attribute indicates the sum of squared error (SSErr) of the fitted model (Pebesma, 2004).The variogram parameters (partial sill, range, and nugget), which are discussed in section 2.3.2 were estimated using the fitted bounded linear model and these parameters together with the SSErr are presented in Table 6. The SSErr indicates the squared deviations between the observed semivariances and the ones expected from the model (Webster & Oliver, 2008). Fitting exponential model was also tried. However, the model did not converge and ended in a singularity when fitted to some of the hourly variograms.

The cut-off and width values were further increased to 1100 and 150 respectively. As illustrated in Figure 8, this result to variograms with much less noises compared to the previous ones, but were unbounded. As a result, we need to fit non-transitive model and the power model was fitted as shown in Figure 10, and the estimated parameters and SSErr are presented in
Table 7.

.

Succeeding step was to evaluate the prediction accuracy between the hourly and the average variogram models. The power model was used for prediction and the prediction accuracy of the hourly and average models were compared using ME and RMSE based on the leave-one –out cross-validation strategy. A model with lower RMSE is more accurate and ideally ME should be zero indicating that prediction is unbiased.

There are a number of factors (such as latitude, altitude, season, etc.) that influence the temperature at any particular place. Some other factors like wind direction, time of day and present weather condition can control variations in temperature over short periods. It could be the effect of these factors that caused the difference in the empirical variogram of every hour and the average empirical variogram. However, from Table 8, we can see that there is no big difference in the RMSE for the prediction using the average variogram and using the hourly variogram. This indicates that comparable prediction can be done using the average variogram. Predictions at 9:30 and 11:30 were even more accurate when we use the average variogram model of the whole morning than the hourly variogram models. From these results, we can see that the observations collected at different time periods can be characterized by the same spatial correlation structure. Hence it can be concluded that the assumption in which the variogram averaging method of Sterk & Stein (1997) is based on is realistic.

6.3. Event detection

Once outliers are detected, it is useful to distinguish between events and errors. Events were detected by exploiting spatial and temporal correlations existing in the sensor data. Outliers that represent a change in the state of the environment are more likely to be spatially and temporally correlated. However, outliers that occur due to measurement error or environmental noise are less likely to be spatially or temporally correlated.

From Table 9 we can see how the event at the 5 nodes developed temporally. For all the nodes, the event disappeared and appeared again at least once. The event did not start at the same time in all the nodes. It started at node 28; then node 25, node 32, node 31 and finally node 29. The event did not develop at the same time in all the nodes. It took it about 14 minutes to reach node 29 once it started at node 28. The event lasted for about 34 minutes. The event finally stopped at the same time for nodes 28, 29 and 31. At nodes 32 and 25, it stopped 4 minutes and 8 minutes earlier respectively.

In order to assess the accuracy of the event detection process, there was no reference dataset. As a result, the observations which were labelled as outliers using the three labelling techniques by Zhang (2010) were relabelled as events or errors based on the definition of errors provided in section 4.3. This relabelled dataset was then used to assess the accuracy of the event detection technique. However, the number of outliers labelled using Mahalanobis distance-based and density-based techniques are too small. Consequently no outlier was labelled as an event when data from the Mahalanobis distance-based and density-based techniques was used. Hence, the accuracy assessment was based on the classified outliers of the running average-based technique only. It can be seen from Table 4 or Table 5, Zhang's (2010)TOD technique achieved the lowest DR when the running average based labelling technique was used to assess the accuracy. Therefore, it is difficult to conclude on the accuracy of the event detection process based on the running average bases relabelled dataset.

7. CONCLUSION

The objective of this research was to develop analytical techniques which are suitable for event detection using wireless sensor networks (WSNs). To achieve this objective, five specific objectives, followed by one or more research questions for each objective, were formulated (Table 1). From the reviewed literature, obtained results and discussion the specific objectives are achieved to conclude the following findings.

Defining an outlier and an event

No universally accepted definition for outliers exists. The perception of outliers differs in terms of data type, application domains and detection techniques. But generally, measurements that significantly deviate from the normal pattern of a phenomenon are called outliers. Domain specific definitions were given to outliers based on the WMO guidelines in relation to the variable of interest (i.e., ambient temperature) and the study site (i.e. the Grand St. Bernard pass). These definitions are based on plausible value and plausible rate of change of measured values. Similarly, outliers were also defined based on a specific confidence level by creating a series of averages of different subsets of the full data set.

Outliers are classified in to errors and events. Errors signify either environmental noise or faulty sensor. Events are changes in the state of the environment. The distinction between events and errors was done based on the length of outlier sequence depending on the sampling rates. In cases where there is no a priori knowledge on incoming event, it is difficult to assign a specific sequence of detected outliers to represent an event. In such cases the length of outlier sequence should be determined based on the application requirements. Nevertheless, since single node is liable to failures, we cannot make reliable distinction between events and errors based on the length of outlier sequence only. Therefore, distributed collaboration among neighbouring nodes is required. Zhang (2010) recommended that the length of the outlier sequence be small, so that there is no delay in classifying outliers as events or errors. Moreover, in order outliers to be distinguished as events, Krishnamachari & Iyengar (2003) recommended that a number of consecutive outliers from one node should correspond with another such observations from at least half of the neighbouring nodes. Based on these recommendations, an event is defined as four consecutive outliers that correspond with another such observations from at least half of the neighbouring nodes.

An event pertains to a place of limited geographic extent and a limited time frame of the observation period. If such an extreme reading pertains for a time frame more than a specific limit and its geographic extent covers majority of the entire region under consideration, it is no more an event (i.e. it becomes the usual behaviour of the environment). It is difficult to specify a limit on time frame and geographic extent for considering an event as the normal environmental variable. For convenience the minimum limits for the time frame and geographic extent for classifying an event as normal environmental variable were set to be one fourth of the total observation period and two third of the total number of nodes respectively.

Modelling the data behaviour

Temporal and spatial correlations existing in WSN data, based on time series analysis and geostatistics respectively, were used for defining the normal behaviour of temperature data. Then the measurements that significantly deviate from the normal pattern were identifying as outliers.

A time series model will generally reflect the fact that observations close together in time will be more closely related than observations further apart. Time series analysis using two different techniques (maximum variability check and moving average) was used to detect temporal outliers. Maximum variability check detected outliers by checking for unrealistic jumps in values over a specific period. In the moving average approach averaging was done using simple moving average (SMA) and moving median (MM) techniques. These techniques set a dynamic threshold with maximum and minimum bound of a given confidence level of a specific subset. An observation that lay outside the maximum and minimum bound of a given confidence level of a specific subset was detected as an outlier.

Geostatistics deals with spatially autocorrelated data. Geostatistical analysis for modelling spatial correlations was used to identify spatial outliers. This include calculating a sample variogram from the data, fitting a model to the sample variogram and predicting at unsampled locations (Webster & Oliver, 2008).

When using geostatistical analysis for modelling spatial correlation to identify outliers variogram averaging method of Sterk & Stein (1997) was used . This method assumes that a constant spatial correlation exists over time. To check if this assumption is realistic, spatial variations for different hours and for the whole morning were modelled and tested using the leave-one-out cross validation technique (Webster & Oliver, 2008) . The results proved that the average models can do comparable predictions to the hourly models. Hence, the assumption in which the variogram averaging method of Sterk & Stein (1997) is based on is realistic.

Detecting an event

In order to assure reliable detection of interesting events using WSNs, we need to develop or use an accurate outlier detection technique while paying attention to the computational, storage and communication limitations in WSNs. The maximum variability check, the SMA and the MM approaches of detecting outliers are model independent and computationally simple. Using these methods for detecting outliers in a single node can save the communication overhead. However, the results indicate that the accuracy of these approaches is not as good as the model based TOD technique by Zhang, et al (in press). As a result, the TOD technique was used for event detection. Event detection process was done by exploiting temporal and spatial correlations existing in sensor data. Based on the previously provided definition, four consecutive outliers that correspond with another such observations from at least half of the neighbouring nodes were detected as events.

Characterizing an event

Once an event is detected, its temporal evolution and spatial extent can be described. From Table 9, we can see that the event did not start at the same time in all the nodes. It started at node 28; then goes to node 25, node 32, node 31 and finally to node 29. Besides, the event did not develop at the same time in all the nodes. It took it about 14 minutes to reach node 29 once it started at node 28. The event lasted for about 34 minutes. The event finally stopped at the same time for nodes 28, 29 and 31. At nodes 32 and 25, it stopped 4 minutes and 8 minutes earlier respectively.

Evaluation of event detection

Commonly, no pre-labelled data is available for sensor data (Zhang, 2010). This made the evaluation of outlier detection technique challenging. This in turn made the event detection technique even more challenging. Due to the lack of availability of reference dataset, evaluation of the event detection technique was made by relabeling a labelled dataset. Zhang (2010) labelled the observations of the Grand St. Bernard pass data as outliers or normal data using three labelling techniques. The observations which were labelled as an outlier by Zhang (2010) were relabelled as events or errors, based on the previously given definition of events, and were used for evaluating the event detection technique. The event detection approach can detect events with DR of 49% and FPR of 16%. Due to the limitation in selecting the length of sequence of outliers the event detection approach may not be reliable.

Recommendations

The accuracy of the maximum variability check could be improved by carefully setting the maximum variability limit in relation to the behaviour of the variable of interest in the study area. Similarly, the accuracy of the moving average technique could be improved by giving emphasis to the most recent values using moving average techniques that use unequal weights for each data value in the subset (for example, weighted least square). These cannot be analysed in this research due to the time constraint and are proposed for future work.

LIST OF REFERENCES

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey. *Computer Networks, 38*(4), 393-422.

Bahrepour, M., Meratnia, N., Poel, M., Taghikhaki, Z., & Havinga, P. J. M. (2010). *Distributed Event Detection in Wireless Sensor Networks for Disaster Management.* Paper presented at the 2nd International Conference on Intelligent Networking and Collaborative Systems (INCOS).

Banerjee, T., Xie, B., & Agrawal, D. P. (2008). Fault tolerant multiple event detection in a wireless sensor network. *Journal of Parallel and Distributed Computing, 68*(9), 1222-1234.

Barrenetxea, G., Ingelrest, F., Schaefer, G., & Vetterli, M. (2008a). *Wireless Sensor Networks for Environmental Monitoring: The SensorScope Experience.* Paper presented at the Int. Zurich Seminar on Communications (IZS).

Barrenetxea, G., Ingelrest, F., Schaefer, G., Vetterli, M., Couach, O., & Parlange, M. (2008b). *SensorScope: Out-of the-box environmental monitoring.* Los Alamitos: IEEE Computer Soc.

Basu, S., & Meckesheimer, M. (2007). Automatic outlier detection for time series: an application to sensor data. *Knowledge and Information Systems, 11*(2), 137-154.

Bivand, R. S., Pebesma, E. J., & Gomez-Rubio, V. (2008a). Applied spatial data analysis with R. Springer, NY. *http://www.asdar-book.org/*

Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2008b). Interpolation and Geostatistics *Applied Spatial Data Analysis with R* (pp. 191-235): Springer New York.

Chandola, V., Banerjee, A., & Kumar, V. (2007). *Outlier detection: a survey*: Technical Report, University of Minnesota, USA.

Chatfield, C. (2004). *The Analysis of Time Series : An Introduction*: Chapman & Hall/ CRC, London.

Chongming, Z., Chunmei, W., Dazhi, L., Xi, Z., & Chuanshan, G. (2009). *Unspecific Event Detection in Wireless Sensor Networks.* Paper presented at the ICCSN '09. International Conference on Communication Software and Networks.

Ding, M., Chen, D., Xing, K., & Cheng, X. (2005). *Localized fault-tolerant event boundary detection in sensor networks.* Paper presented at the INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE.

Easton, V. J., & McColl, J. H. (1997, September 1997). Statistics Glossary. Retrieved October 5, 2011, from http://www.stats.gla.ac.uk/steps/glossary/index.html

Elson, J., & Estrin, D. (2004). Sensor Networks: A Bridge to the Physical World *Wireless Sensor Networks* (pp. 3-20): Springer US.

Figueiredo, C. M. S., Nakamura, E. F., & Loureiro, A. A. F. (2009). A Hybrid Adaptive Routing Algorithm for Event-Driven Wireless Sensor Networks. *Sensors, 9*(9), 7287-7307.

Grenon, P., & Smith, B. (2004). SNAP and SPAN: Towards dynamic spatial ontology. *Spatial cognition and computation, 4*(1), 69-103.

Hawkins, D. M. (1980). *Identification of Outliers.* London: Chapman and Hall.

Ingelrest, F., Barrenetxea, G., Schaefer, G., Vetterli, M., Couach, O., & Parlange, M. (2010). SensorScope: Application-Specific Sensor Network for Environmental Monitoring. *ACM Transactions on Sensor Networks, 6*(2).

King, K., & Nittel, S. (2010). Efficient Data Collection and Event Boundary Detection in Wireless Sensor Networks Using Tiny Models *Geographic Information Science* (Vol. 6292, pp. 100-114): Springer Berlin / Heidelberg.

Krishnamachari, B., & Iyengar, S. (2003). Efficient and Fault-Tolerant Feature Extraction in Wireless Sensor Networks (Vol. 2634, pp. 551-551): Springer Berlin / Heidelberg.

Lian, J., Chen, L., Naik, K., Liu, Y., & Agnew, G. B. (2007). Gradient boundary detection for time series snapshot construction in sensor networks. *IEEE Transactions on Parallel and Distributed Systems, 18*(10), 1462-1475.

Liang, Q., & Wang, L. (2005, March 31 2005-April 1 2005). *Event detection in wireless sensor networks using fuzzy logic system.* Paper presented at the CIHSPS 2005. Proceedings of the 2005 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety.

Ma, X. L., Yang, D. Q., Tang, S. W., Luo, Q., Zhang, D. H., & Li, S. F. (2004). Online mining in sensor networks. In H. Jin, G. R. Gao, Z. W. Xu & H. Chen (Eds.), *Network and Parallel Computing, Proceedings* (Vol. 3222, pp. 544-550). Berlin: Springer-Verlag Berlin.

Martinez, K., Hart, J. K., & Ong, R. (2004). Environmental sensor networks. *IEEE Computer Society, 37*(8), 50-56.

Martinez, K., Padhy, P., Elsaify, A., Zou, G., Riddoch, A., Hart, J. K., et al. (2006). *Deploying a sensor network in an extreme environment.* Paper presented at the IEEE International Conference on Sensor Networks.

Meteosuisse. (2005 - 2011). Federal Office of Meteorology and Climatologt Switzerland. Retrieved Oct. 15, 2011, from http://www.meteosuisse.admin.ch/web/en/weather.html

Onoz, B., & Oguz, B. (2003). Assessment of outliers in statistical data analysis. In N. B. Harmancioglu, S. D. Ozkul & O. Fistikoglu (Eds.), *Integrated Technologies for Environmental Monitoring and Information Production* (Vol. 23, pp. 173-180). Dordrecht: Springer.

Pawlowski, A., Guzman, J. L., Rodriguez, F., Berenguel, M., Sanchez, J., & Dormido, S. (2009). Simulation of Greenhouse Climate Monitoring and Control with Wireless Sensor Network and Event-Based Control. [Article]. *Sensors, 9*(1), 232-252.

Pebesma, E. J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences, 30*, 683-691.

Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News, 5*(2), http://cran.r-project.org/doc/Rnews/

Qu, Y., Fang, J., & Shuai, Z. (2010). *Nearest neighbor nodes and connectivity of wireless sensor networks with Poisson point process.* Paper presented at the Control Conference (CCC), 2010 29th Chinese.

R Development Core Team. (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Ryan, J. A., & Ulrich, J. M. (2011). xts: eXtensible Time Series. R package version 0.8-2. http://CRAN.R-project.org/package=xts

SensorScope. (2007a). Grand-St-Bernard pass deployment. Retrieved 20 August 2011, from http://sensorscope.epfl.ch/index.php/Grand-St-Bernard_Deployment

SensorScope. (2007b). SensorScope System. Retrieved 20 August, 2011, from http://sensorscope.epfl.ch/index.php/MainPage

Shi, M. Z., & Winter, S. (2010). Detecting Change in Snapshot Sequences. In S. I. Fabrikant, T. Reichenbacher, M. VanKreveld & C. Schlieder (Eds.), *Geographic Information Science* (Vol. 6292, pp. 219-233).

Sterk, G., & Stein, A. (1997). Mapping wind-blown mass transport by modeling variability in space and time. *Soil Science Society of America Journal, 61*(1), 232-239.

Ulrich, J. (2011). TTR: Technical Trading Rules. *R package version 0.21-0. http://CRAN.R-project.org/package=TTR*

Wang, A. Y., & Sodini, C. G. (2004). *A simple energy model for wireless microsensor transceivers.* Paper presented at the Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE.

Wark, T., Corke, P., Sikka, P., Klingbeil, L., Ying, G., Crossman, C., et al. (2007). Transforming Agriculture through Pervasive Wireless Sensor Networks. *Pervasive Computing, IEEE, 6*(2), 50-57.

Webster, R., & Oliver, M. A. (2008). *Geostatistics for Environmental Scientists* (2nd ed.): John Wiley & Sons Ltd.

Yoneki, E., & Bacon, J. (2005). Unified semantics for event correlation over time and space in hybrid network environments *On the Move to Meaningful Internet Systems 2005: Coopis, Doa, and Odbase, Pt 1, Proceedings* (Vol. 3760, pp. 366-384). Berlin: Springer-Verlag Berlin.

Záhumenský, I. (2004). Guidelines on quality control procedures for data from automatic weather stations. *CIMO/OPAG-SURFACE/ET ST&MT-1/Doc.6.1(2). World Meteorological Organization.*

Zeileis, A., & Grothendieck, G. (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software, 14*(6), 1-27. URL http://www.jstatsoft.org/v14/i06/

Zhang, Y. (2010). *Observing the unobservable : distributed online outlier detection in wireless sensor networks. PhD thesis.* University of Twente, The Netherlands.

Zhang, Y., Hamm, N. A. S., Meratnia, N., & Stein, A. (in press). Statistically-Based Outlier Detection Techniques for Wireless Sensor Networks. *International Journal of Geographical Information Science.*

Zhang, Y., Meratnia, N., & Havinga, P. J. M. (2008). Outlier detection techniques for wireless sensor networks: a survey. *IEEE, Communications Surveys and Tutorials 12*(2), 159-170.

Zhang, Y., Meratnia, N., & Havinga, P. J. M. (2010). Ensuring high sensor data quality through use of online outlier detection techniques. *Int. J. Sensor Networks, 7*(3), 141-151.

Zheng, Z., & Gang, Q. (2011, 11-14 Sept. 2011). *An energy efficient adaptive event detection scheme for wireless sensor network*. Paper presented at the Architectures and Processors (ASAP), 2011 IEEE International Conference on Application-Specific Systems.