

Drink. Water Eng. Sci. Discuss., 3, 149–175, 2010  
www.drink-water-eng-sci-discuss.net/3/149/2010/  
© Author(s) 2010. This work is distributed under  
the Creative Commons Attribution 3.0 License.

This discussion paper is/has been under review for the journal Drinking Water Engineering and Science (DWES). Please refer to the corresponding final paper in DWES if available.

# Understanding and managing large sensor networks

**D. D. Ediriweera and I. W. Marshall**

Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK

Received: 5 February 2010 – Accepted: 12 March 2010 – Published: 12 April 2010

Correspondence to: D. D. Ediriweera (damjee@gmail.com)

Published by Copernicus Publications on behalf of the Delft University of Technology.

**DWESD**

3, 149–175, 2010

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Abstract

The water supply industry is trialing a range of sensor network designs for monitoring distributed infrastructure. The paper investigates the performance of such a sensor system deployed to monitor a water distribution network. The study reveals up to one fifth of the data intended to be collected either to be missing or erroneous. Findings reinforce the importance of in-depth design consideration of all aspects of large scale sensor systems, and the necessity for expertise on every detail of the system, or access to a rule set which embeds this knowledge allowing non-specialists to make near optimal choices. First steps towards defining such a rule set is presented here with supporting evidence.

## 1 Introduction

### 1.1 Sensor network design

Sensor networks perform an important role in facilitating efficient management of distributed industrial infrastructure. Sensors are commonly used for monitoring operational status of critical assets with the objective of identifying potential issues early on. Benefits of such monitoring include the possibility of performing proactive maintenance leading to significant financial savings, and providing regulated industries with efficient means of maintaining and managing distributed infrastructure within regulatory requirements.

Sensors have been previously used by the water distribution and supply industry for offline monitoring of assets. Since recently they are also being trailed for online monitoring of water distribution infrastructure (Stoianov et al., 2007). Within this new context sensors are used for collecting information in near real-time to facilitate improved and proactive management of pipe systems. Successful achievement of this objective is highly dependent upon the operational performance of the sensor systems.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and  
I. W. Marshall

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Proper design, deployment, and management of sensor systems are critical for achieving efficient and reliable operation. Acquiring skills and expertise to achieve this can sometimes be expensive and difficult leading to compromised system quality. Non-optimal design choices not only lead to poor performance of the sensor system itself, but also compromise any system which relies on it for information.

Deploying sensors for monitoring distributed infrastructure unavoidably introduces a second distributed system to the mix, this being the sensor system itself. Being a distributed system, sensor networks themselves commonly carry complexities similar to the systems which they are designed to monitor. Therefore it is essential to identify that even well designed sensor systems require suitable tools embedded in them to support close monitoring to ensure sustainable performance (Tierney et al., 2001).

**1.2 Work outline**

A large scale industrial sensor system deployed for online monitoring of pressure and flow measurements of a wide area water distribution and supply system is investigated. System performance, system design, system deployment, and system management are looked at. The work aims to identify operational limitations which the system encountered during its initial four year deployment period and attempts to understand which of such limitations were related to poor design-time choices.

A series of correlations tests are presented between identified system failures and possible contributing factors. The correlation analysis aims to identify failure mechanisms affecting the system and reasons behind the failures. Findings and results are presented as set of basic design guidelines aimed at helping non-specialist engineers make high quality design choices with an opportunity to avoid certain elusive runtime limitations experienced by large scale sensor networks.

## 2 Related work

The discovery of new design practices in development of sensor systems is an ongoing interest. A brief overview of relevant work within this area is presented below.

Common sensor network architecture for equipment failure prediction through vibration monitoring within industrial environments has been proposed (Krishnamurthy et al., 2005). The architecture is trialed within two distinct industrial surroundings: A semiconductor manufacturing plant; and a North Sea Oil Tanker. The two implementations use different hardware processors and communication capabilities in-order to test the architecture against different platforms. Lessons learned from the trials are presented as design insights for future deployments. Authors identify a relationship between sensor hardware configuration and power efficiency. It is claimed that more capable hardware with sufficient RAM and I/O bandwidth were found to be more power efficient as they require less software intelligence for resource management. The paper also reveals certain wireless communication limitations which were experienced under varying radio conditions. The use of retry mechanisms, and oversampling to overcome such issues are discussed.

Work of Cerpa et al. (2001) identifies design challenges of a different sensor network application. Authors present a sensor network platform for habitat monitoring. The paper addresses sensor network design issues such as miniaturisation, energy efficiency, localization algorithms, and time synchronisation. The platform has been trialed in a test environment. A similar architecture has been trialed outdoors by Szewczyk et al. (2004). Though the architecture used here is relatively simple the trials reveal interesting and unexpected behavior throughout the four month deployment period. 150 nodes have been deployed with both single hop and multi hop communication capability. The paper provides an analysis of the trial period and looks at various performance characteristics in relation to changing conditions such as weather, battery performance, deployment depth, and hardware robustness. Useful design insights are presented.

DWESD

3, 149–175, 2010

### Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

The efficient design of sensor networks for monitoring continuously moving objects has been explored (Nikoletseas et al., 2008). The paper introduces a novel combi-  
national model designed to estimate the number of sensors required, and where they  
should be deployed, for solving a given tracking problem. Work is focused on build-  
ing a flexible model capable of designing optimum sensor layout for different problem  
specifications.

Based on considered literature, it is evident that existing design guidelines for sensor  
network development are typically geared towards addressing issues which are appar-  
ent at design time, and in most cases such guidelines are local to specific problem  
domains. In contrary guidelines presented here are aimed at assisting designers avoid  
issues which are illusive and subtle at design time, but impactful during live operation.  
These guidelines are derived by analyzing the performance of a live network in ret-  
rospect, and hence should allow the analysis to effectively capture runtime limitations  
and their relationship to prior design time choices.

### 3 Survey details

#### 3.1 Information used

A multi-faceted approach has been chosen for the analysis. A variety of variables and  
potential contributing factors have been analyzed with the aim of understanding how  
each effects system performance. Primary analysis has been performed based on  
the dataset collected by the sensor network over the studied period. Following data  
sources and information have been used in addition.

- Discussions with the network operator, and analysis of other data pertaining to  
the water distribution system being monitored.
- Analysis of the sensor locations and their likely impact on communications perfor-  
mance.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Analysis of local weather conditions during the study period.
- Analysis of other sensor networks and knowledge of monitoring systems in the telecommunications industry.

### 3.2 System configuration

5 The investigated system is a wide area sensor network deployed for monitoring of a water distribution and supply system. The sensor net consists of 520 plus sensor nodes monitoring flow and pressure of the pipe system. The nodes are distributed within an area roughly around the size of 50 km<sup>2</sup>. Majority of the nodes are deployed underground inside manholes located on public roads. All nodes including their sensors,  
10 radios and loggers are commercially manufactured and are designed to be water proof. Electronics of the nodes are handmade with the possibility of some minor manufacturing imperfections. Each node is equipped with a GSM modem capable of GPRS data connectivity. Data collected by the nodes are relayed through a public GSM/GPRS network. Nodes are powered using a single battery pack with an estimated lifetime of  
15 approximately twenty four months.

Each node is designed to support up to a maximum of 4 data channels. Each channel can be connected to an individual sensor. 68% of the nodes deployed were using only one connected channel while the remaining 32% nodes were using 2 channels. Nodes with two active channels were monitoring both flow and pressure, where as  
20 nodes with a single active channel were measuring either flow or pressure.

The data path configuration of the system is highlighted in Fig. 1. The nodes are configured to record readings every 15 min starting at the top of the hour. Pressure readings are recorded as snapshots whereas flow is averaged over several readings within the 15 min period. The loggers report back to the central data collector every  
25 30 min starting at the top of the hour. The dataset used for the analysis consists of a total of over 76 million records over a four year period from 2006 to 2009. Around

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



49 million were readings of pressure, and around 27 million were flow. The data is reported from 529 distinct nodes recording 697 distinct channels.

## 4 Analysis

### 4.1 Missing data analysis – overview

5 Table 1 provides a summary of yearly performance of the system. Failures in the table all refer to events of missing data within the dataset. A missing event in this instance can be identified as occurring when two consecutive measurements for a specific channel occur with a time difference of greater than 15 min. This can be viewed as a conservative measurement as it does not take in to account any records which  
10 continued to be missing at the end of the studied period. Considering Table 1, key performance indicators are Logger Failure Percentage (LFP), Missing Data Percentage (MDP), and average Records Lost per Logger per Day (RL/L/D).

LFP identifies the percentage of loggers which have failed at least a single instance during a specific year. Over the four years LFP is found to be fairly stable. This is  
15 a rough indication that the manufacturing quality of original hardware and any newly installed or replacement loggers are fairly stable over the years. MDP identifies the amount of data missing for a year as a percentage of the total data recorded during the same period. MDP is a strong indicator of the overall performance of the system. MDP for 2006 and 2007 are stable around 6–8%, but dramatically increases to 32% in  
20 2008. This indicates a much higher level of failures in 2008. MDP for 2009 indicates a big improvement at 0.8%, but may not be fully indicative as the figures are only for a short period of the year. RL/L/D identifies on average the number of records missing per logger per day. RL/L/D figures are consistent with MDP. RL/L/D peaks in 2008 and then falls off in 2009. In summary these findings indicate that apart from 2009 (were  
25 only a short period of the year is analysed) system performance of all other years show significant loss of data, with 2008 being unusually high.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



## 4.2 Missing data by date

Figure 2 illustrates the distribution of the short term (less than 6 h) missing data events over four years. The plot indicates failures to be somewhat clustered with some periods illustrating higher failure rates than others.

5 Periods of high error rates can be identified as Jan-07 to May-07, Mar-08 to Aug-08, and Oct-08 to Dec-08. The calmer low failure periods occur between Jan-06 to Nov-06, and May-07 to Jan-08. No low failure periods occur after Jan-08.

None of the peak failure periods correlate significantly with rainfall during the same period plotted in Fig. 3. The main rainfall event in Jul-07 which had even caused heavy flooding in the area does not seem to cause any interference affecting the sensor network. Figure 2 illustrates minimum failures during this time. Based on this it is fair to rule out any rain related interference as a main contributing factor of the failures. The rainfall has been calculated as the average of 6 rain gauges in the general area of the sensor network. Figure 5 plots all recorded maintenance work carried out on the sensor network. It is interesting to note that short term failures clustered in Jul-08 and the stand alone peak in Sep-08 closely correlate with the maintenance work. This strongly suggests the possibility of these missing data events being caused by engineers performing maintenance work on the devices.

10  
15  
20  
25 Figure 4 plots long-term errors with duration of greater than 6 h. Two long-term error clusters appear from Jan-07 to May-07 and from Mar-08 to Jul-08. Based on operator feedback these periods strongly correlate with the battery replacement schedule and could possibly indicate devices switching off due to battery depletion. Specific battery replacement records were not available to verify this. Long-term errors which do not fall within these periods potentially indicate out-of-synch battery failures, and other types of hardware failures.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



### 4.3 Missing data by duration and group

Figure 6 plots missing data events grouped by duration, and color-coded by the number of simultaneously effected nodes. Two distinct distributions can be identified in the plot; short-term and long-term errors. Short-term errors peak at “<1 h” and tail off towards “24 to 36 h”. Long-term errors peak at “>30 days” and tail off around “2 to 3 days”. 99.5% of long-term errors were affecting individual nodes. This strongly supports the idea of these being hardware or battery related failures.

Considering short-term errors, 84% of them are affecting individual loggers. These errors could be caused due to a host of issues such as node maintenance, temporary signal-loss, ad-hoc occurrences such vehicular traffic, and radio interference etcetera. The specific cause cannot be confirmed as error logs from the nodes are not available. 13% of the short-terms errors have simultaneously affected groups of 2 to 5 loggers. Failure mechanisms here are potentially different to the individual logger failures. Short-term small group failures could potentially be temporary radio conditions affecting a small locality, GSM/GPRS network traffic, and even node maintenance work in a specific area. Short-term errors effecting larger groups of nodes also exist. One such event is identified in the 50 to 100 nodes category. Such a wide outage is possibly caused due to a network wide event. Failure points for such events include GSM/GPRS network, back-end data collector failure, and/or data mishandling issues. Figure 7 illustrates the relationship between nodes which correlate with each others failures versus the distance between them. Based on Fig. 7 it is clear that strongly correlating nodes are almost always located closer to each other. This is strong evidence that the short-term group failures are mostly locality related.

### 4.4 Missing data by hour of day

Figures 8 and 9 plot missing data events against time of day which they occur at. Figure 8 plots the events during weekdays, and Fig. 9 for the weekends. A distinctly different pattern emerges between weekdays and weekends. During weekdays a strong

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



correlation of the missing data events can be seen with the GSM/GPRS network peak hours. This suggests some relationship to exist between the missing data events and network traffic. This hypothesis is further reinforced by the fact that the high error rate during network peak hours disappearing during weekends where no peak traffic occurs. Although likely the said hypothesis cannot be confirmed as access to the error logs from the nodes are not available for analysis.

The cause for the bell shaped distribution of Fig. 8 could also be due to an entirely different mechanism. Ironically the GPRS peak time also correlates with working hours of sensor network maintenance engineers. The missing data event distribution in Fig. 8 could easily be caused through maintenance work. This could also explain the dip within the hours 12 and 13 as this is generally the break for lunch. However following the plot of the short term errors in Fig. 2, and recorded maintenance in Fig. 5 it is clear that only a subset of the short-term errors correlate with maintenance. This suggests the possibility of the distribution in Fig. 8 being caused by a combination of factors including both maintenance work, and network traffic effects.

The classification of failures by category in Table 2 is only a rough estimate based on the analysed data. These estimates cannot be confirmed as no error logs are available from the nodes. The breakdown in Table 2 nevertheless provides a rough idea of the causes of the failures. A majority of the failure events are short-term. However short-term failures only contribute to around 2.5% of the total records lost. The bulk of the records are missing due to the long-term failures. This does not however indicate which missing data is most damaging as even a short-term failure during a critical time, at a critical point can be more damaging than a long-term failure within a remote area of the network. A portion of the unclassified failures are possibly due to maintenance work not recorded. Remaining unclassified failures might be due to data handling errors at the collector.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

## 4.5 Data quality analysis

Figure 10 plots the distribution of flow values within the dataset. The distribution appears free of any obvious errors with a fairly smooth tail and a clear peak at zero value. Zero flow is an acceptable value as this could occur on certain pipes during off-peak usage periods. However it was found that approximately 14% of the flow records were inaccurately valued zero. This includes any zero flow readings which occur continuously for periods of over 24 h. The pressure value distribution of Fig. 11 illustrates a lower level of data quality in comparison to flow. Observing Fig. 11 three peaks in the distribution are visible. The first and the highest occur at zero. However, zero pressure is not feasible in normal operation of the pipe system. The second peak occurring at 14 is again a mystery. Given the magnitude of the peak it is clear that the readings are non-genuine indicating an issue with the dataset. The distribution also indicates some pressure readings to be negative which again is not feasible within normal operation of the pipe system. Overall pressure readings of the system illustrate poor data quality.

## 4.6 Data analysis overview

The overall analysis reveals 12% of the data which the system should have collected to be missing.

A further 7% of the data are zero valued with no explanation of being genuine, and therefore considered erroneous. 2% of the data are negative values, and 1% impossibly large values. In total this amounts to a figure high as 22% of the data expected to be collected either to be unavailable, or erroneous.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

## 5 Rules for future design

A fundamental oversight in the design of the current sensor system is its lack of support for effective monitoring and management. The sensor system being a distributed entity requires constant management. Necessary tools and mechanism for achieving this should therefore be designed in to the system. Within the current system it is clear that designers were unaware of the complexities of managing a large sensor network, and as a result failed to provide even fundamental management information such as error logs. Resulting from this lack of information the operators of the network were unaware of the scale of the performance lapses within the system. To overcome such limitations it is vital that system designers envisage how a system can be managed when it is deployed, and use this insight to integrated necessary tools to support such management. At a minimum such support should include the possibility of monitoring sensor nodes continuously and red-flagging possible issues as they occur. Error logs should be available for further investigation of such issues. Support should also be introduced for automated monitoring of battery performance and scheduling of battery replacements. This would help minimise downtime of nodes due to battery failure. Considering the above, it was surprising to discover that batteries were hardwired to sensor nodes within the current system, making their replacement a complex process. Issues such as hardwired batteries and non-availability of error logs strongly indicated lapses in the requirement specifications provided by the operators to the equipment vendors.

Another important consideration of a well designed sensor systems is the careful evaluation and selection of suitable communication mediums, and fitting communication strategies. Within the current system operators have opted for GSM/GPRS solution while the industry norm has largely remained towards GSM/SMS. Although such forward thinking is valuable, extreme caution needs to be taken to fully understand their implications. As an example, the effects of peak GSM/GPRS network traffic discussed earlier is potentially a hidden issue with significantly adverse impacts. The existence

DWESD

3, 149–175, 2010

### Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

of such an elusive limitation strongly suggests the importance of undertaking performance evaluations such traffic profiling, and quality of service checks on any third party infrastructure which is to be used.

Selecting a proper communication strategy to match the selected communication platform can be as vital as selecting the communication platform itself. Within the current system all nodes are polled every 30 min starting from the top of the hour. This approach of simultaneously polling all nodes may itself be sufficient to cause lower capacity GSM/GPRS network cells to be overloaded, creating traffic errors. More importantly however, this approach creates a bottle neck at the data collector which would need to handle incoming data from all nodes simultaneously, leading to potentially data mishandling and other communication errors within the backend system. The simplest approach to overcome this would be to stagger the polling times such that the nodes report back at different individual times during each 30 min reporting block. To complement such an approach it is also useful to avoid the use of standard times for reporting back data, especially when using third party infrastructure. The rationale being, the probability of standard times such as the top of the hour, fifteen past, or midnight being used by other systems on same network is much higher compared to non-standard times such as for example three minutes past the hour. Using such non-standard times increases the probability of avoiding traffic being generated by other systems.

## 6 Conclusions

The study revealed up to one fifth of the data intended to be collected by the evaluated system to be either missing or erroneous. Contributing factors included lack of operator knowledge on communication systems issues and sensor management, weaknesses in the specification given to the equipment vendor, back end IT system failures, and failing transducer hardware. Findings reveal several important considerations which should be made when designing large sensor systems. These include the design of integrated management functions within the systems, the careful selection of suitable

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



communication platforms and identifying matching communication strategies. Further work is planned on developing these observations further to formulate a more comprehensive rule base aimed at supporting more effective design of large scale industrial sensor systems in the future.

- 5 *Acknowledgements.* This work is has been funded by project Neptune (EPSRC Reference: EP/E003192/1).

## References

- 10 Cerpa, A., Elson, J., Estrin, D., Girod, L., Hamilton, M., and Zhao, J.: Habitat monitoring: application driver for wireless communication technology, Proceedings of the ACM Sigcomm Workshop on Data Communication, 20–41, 2001.
- Krishnamurthy, L., Adler, R., Buonadonna, P., Chhabra, J., Flanigan, M., Kushalnagar, N., Nachman, L., and Yarvis, M.: Design and deployment of industrial sensor networks: experiences from a semiconductor plant and the north sea, Proceedings of the 3rd International Conference on Embedded networked sensor systems, 64–75, 2005.
- 15 Nikolettseas, S. and Spirakis, P.: Efficient sensor network design for continuous monitoring of moving objects, Theor. Comput. Sci., 402, 1, 56–66, 2008.
- Stoianov, I., Nachman, L., Madden, S., and Tokmouline, T.: PIPENETa wireless sensor network for pipeline monitoring, Proceedings of the 6th International Conference on Information processing in sensor networks, 264–273, 2007.
- 20 Szewczyk, R., Mainwaring, A., Polastre, J., Anderson, J., and Culler, D.: An analysis of a large scale habitat monitoring application, Proceedings of the 2nd International Conference on Embedded networked sensor systems, 214–226, 2004.
- Tierney, B., Crowley, B., Gunter, D., Lee, J., and Thompson, M.: A Monitoring Sensor Management System for Grid Environments, Cluster Comput., 4, 1, 19–28, 2001.

## Understanding and managing large sensor networks

D. D. Ediriweera and  
I. W. Marshall

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and  
I. W. Marshall

**Table 1.** Sensor net performance summarised by year.

Year	Total Days Operational	Total Failings	Loggers Operational	Loggers Failed	Logger Failure %	Missing Data %	Total Records Lost	Records Lost per Logger per Day
2006	364	822	479	321	0.670	8.03	1 196 956	6.86
2007	364	1242	462	401	0.867	7.45	1 389 133	8.26
2008	365	2527	430	349	0.811	32.8	4 465 913	28.45
2009	71	137	387	289	0.746	0.32	21 987	0.80

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall

**Table 2.** Overall breakdown of failures by estimated type.

Type	# Failures	# Lost Records	% Failures	% Lost Records
Battery (LT)	455	839 676	9.74%	11.75%
Hardware (LT)	1170	6 121 813	25.05%	85.68%
Network (ST)	538	32 250	11.52%	0.45%
Maintenance (ST)	659	37 980	14.11%	0.53%
Unclassified (ST)	1848	112 515	39.57%	1.57%

(LT) – Long-Term, (ST) – Short-Term

[Title Page](#)

[Abstract](#) [Introduction](#)

[Conclusions](#) [References](#)

[Tables](#) [Figures](#)

[⏪](#) [⏩](#)

[◀](#) [▶](#)

[Back](#) [Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

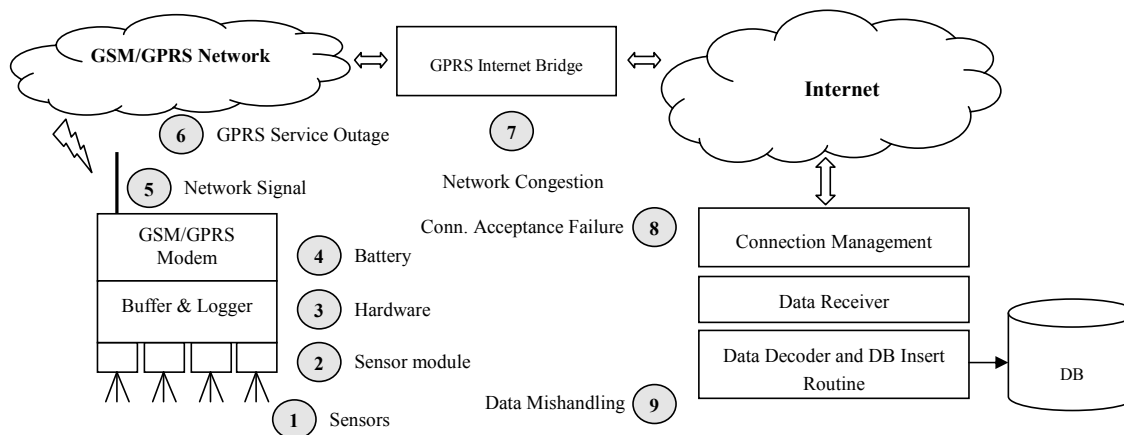
[Interactive Discussion](#)





**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 1.** The diagram provides an overall view of the configuration of the sensor network being investigated. An abstracted view of the process is captured to improve clarity. Circled numbers indicate possible failure points within the infrastructure which are relevant to the analysis.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

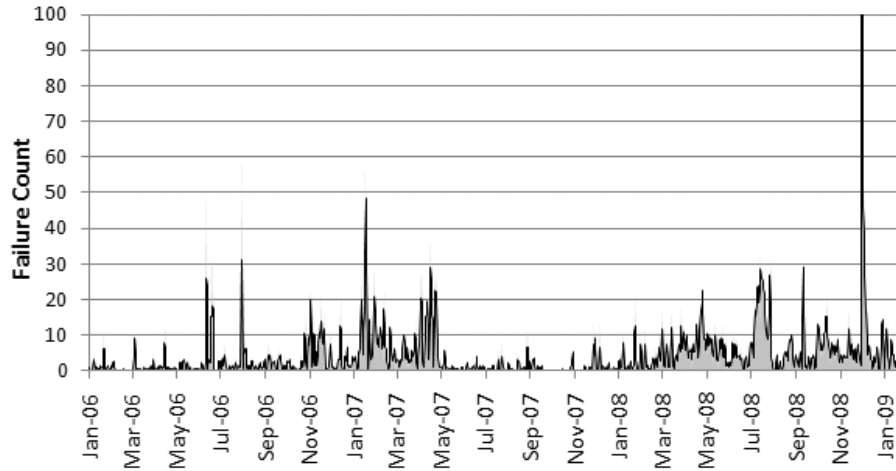
[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 2.** Short-term (<6 h) failures vs. date.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

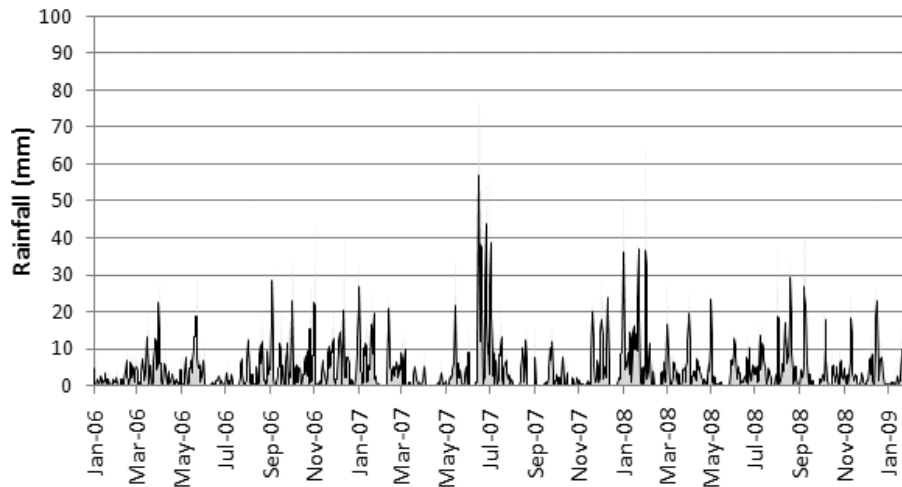
[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 3.** Rainfall vs. date.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

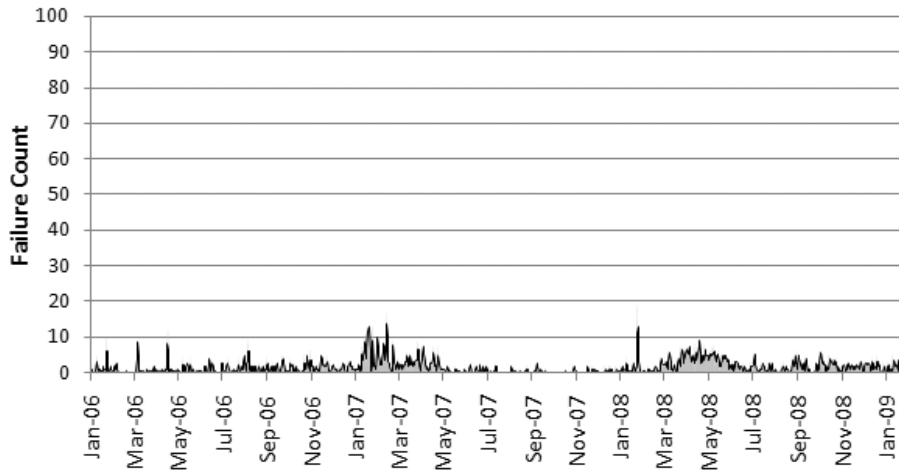
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 4.** Long-term (>6 h) failures vs. date.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

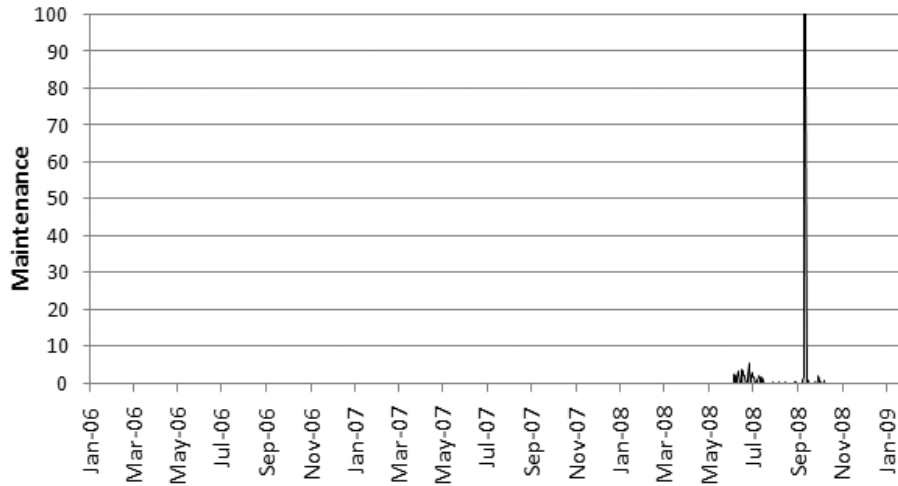
[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 5.** Sensor network maintenance vs. date.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

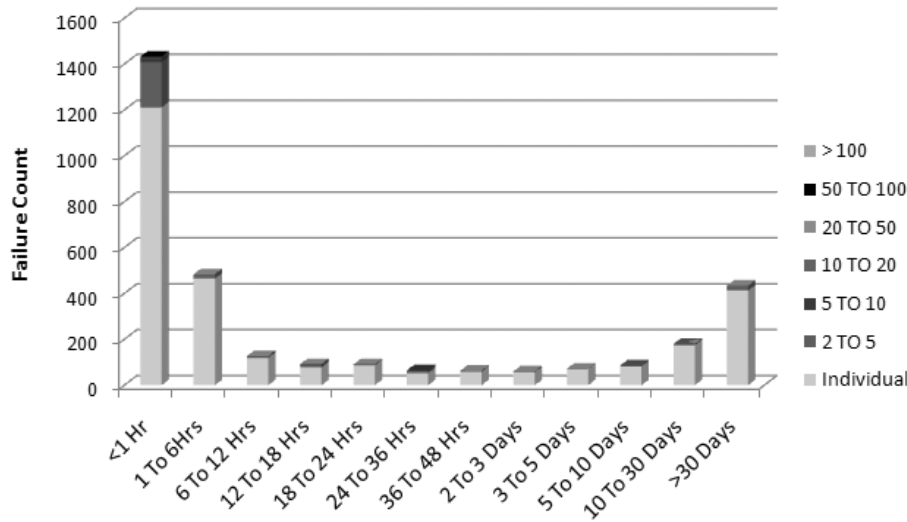
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall

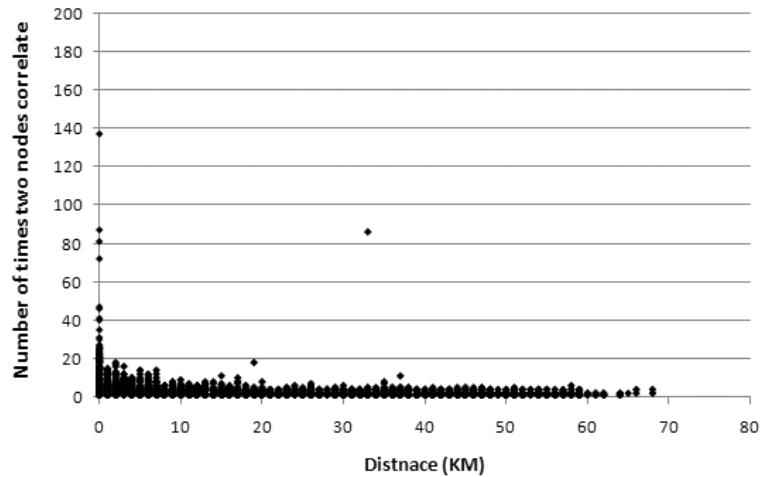


**Fig. 6.** Missing data events grouped by duration and number of affected nodes.

- [Title Page](#)
- [Abstract](#)   [Introduction](#)
- [Conclusions](#)   [References](#)
- [Tables](#)   [Figures](#)
- [⏪](#)   [⏩](#)
- [◀](#)   [▶](#)
- [Back](#)   [Close](#)
- [Full Screen / Esc](#)
- [Printer-friendly Version](#)
- [Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 7.** Relationship between individual nodes which correlate in their short-term failures versus the distance between them. A 1 h time window is used for correlating missing data events between two individual nodes.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

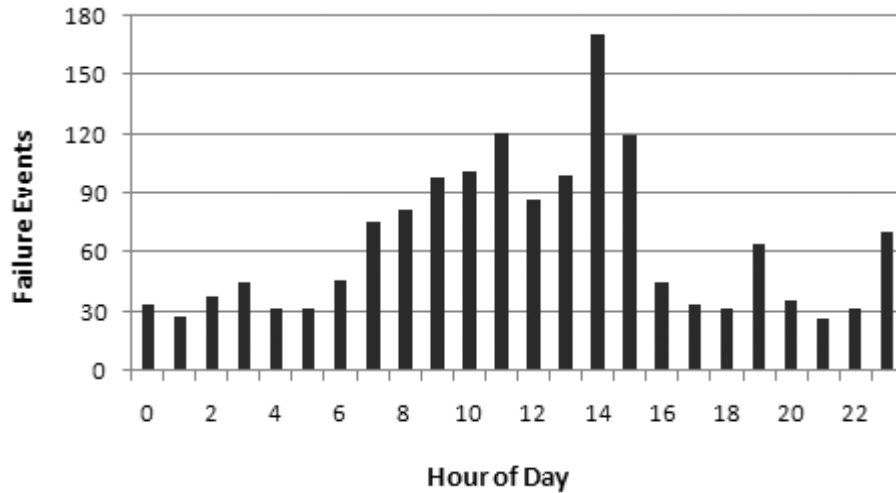
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



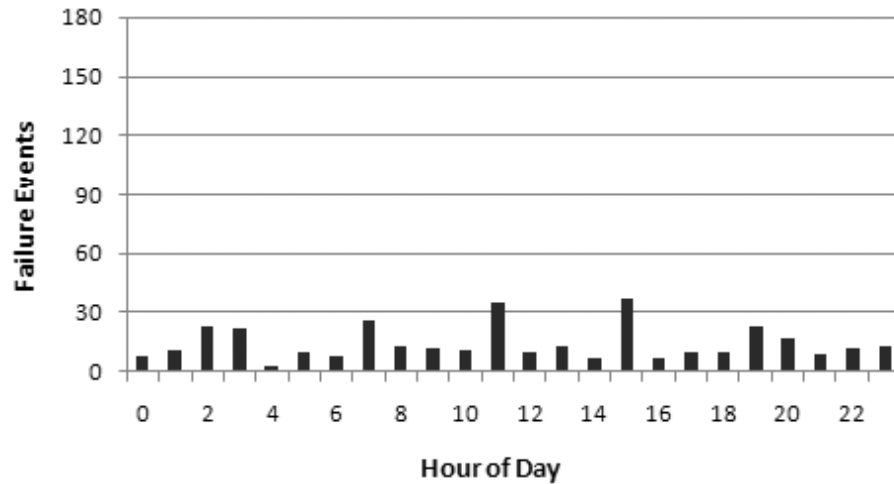
**Fig. 8.** Weekday missing data events against hour of day.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
⏪	⏩
◀	▶
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	



**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall

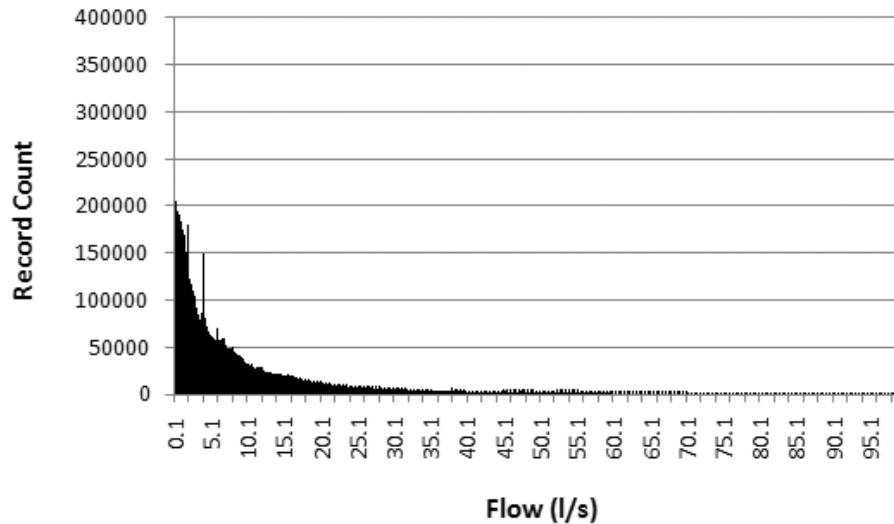


**Fig. 9.** Weekend missing data events against hour of day.

- [Title Page](#)
- [Abstract](#)
- [Introduction](#)
- [Conclusions](#)
- [References](#)
- [Tables](#)
- [Figures](#)
- [⏪](#)
- [⏩](#)
- [◀](#)
- [▶](#)
- [Back](#)
- [Close](#)
- [Full Screen / Esc](#)
- [Printer-friendly Version](#)
- [Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall



**Fig. 10.** Distribution of flow values.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

**Understanding and managing large sensor networks**

D. D. Ediriweera and I. W. Marshall

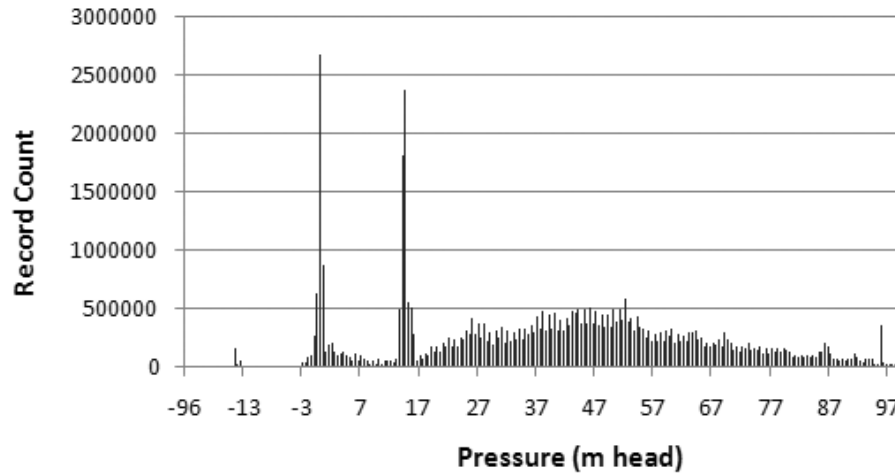


Fig. 11. Distribution of pressure values.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion