

# Urban Water Security Research Alliance



## An Application of Pattern Recognition for the Location and Sizing of Leaks in Pipe Networks

**John Mashford**

**System Losses**

**17 August 2009**

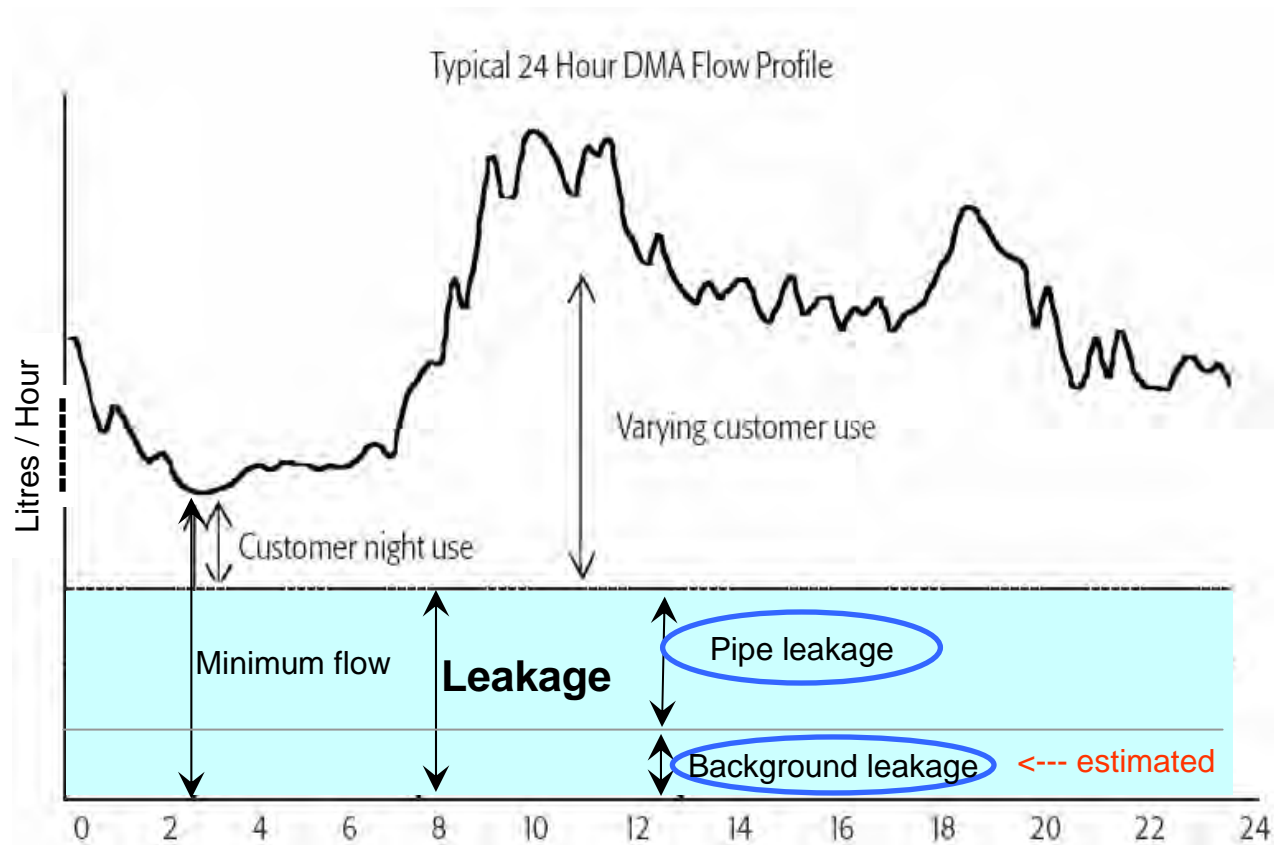


# Background to the leakage problem

- Large leaks such as burst water mains are not an issue
  - These are 'visible' and get reported within a short time
  - They get fixed fast, usually within 72 hours
- The problem is with small leaks -
  - Leaks of 1 -10 L/Hr (eg. fully open tap is about 30-50 L/minute)
  - These leaks run for months before they get fixed
  - Very small leaks are never fixed as its 'not economical'
- Average losses, incl. leakage, for a city is 6% - 25% of total supplied
  - In some European cities, its as high as 40%
  - Parts of London is supposed to be about 35%
  - Australian capital city average is about 12%
  - Melbourne average is about 9%
  - Total leakage is about 171 GL/Year for Australia

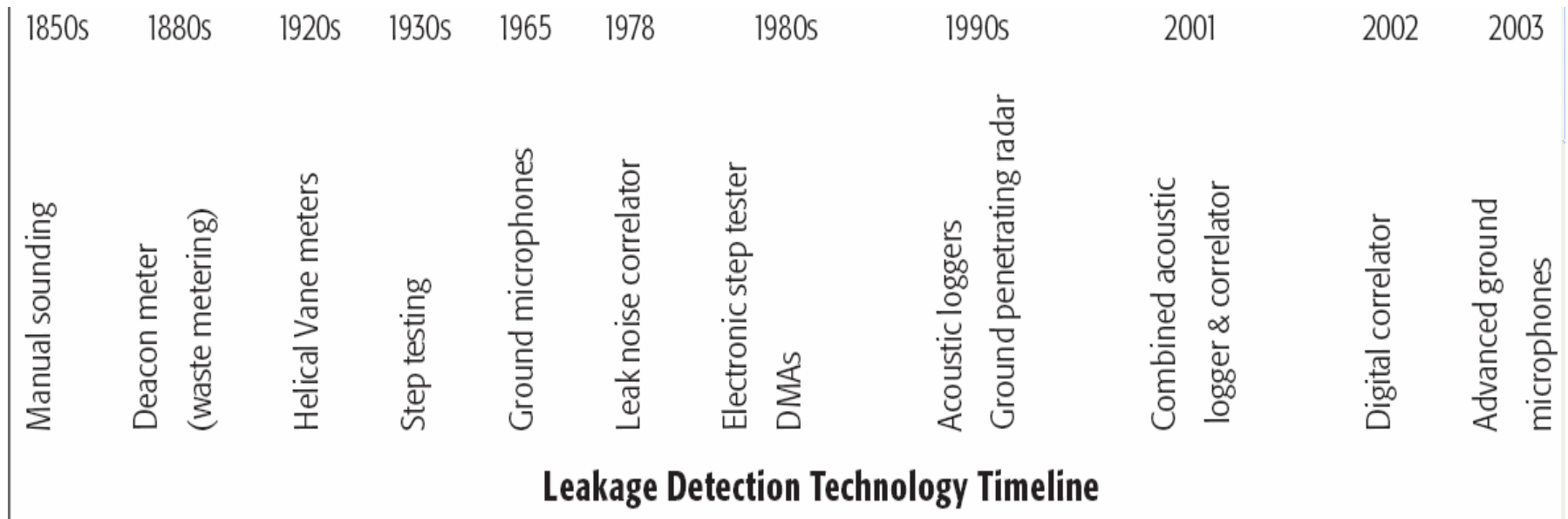
# Leakage quantification

- Typical 24 hour flow profile of components of DMA leakage



# Current Leakage Detection Methods

- As water escapes through a hole in the pipe, a hissing noise is generated
- Detection methods, from 100 years ago to present, are based on identifying the source of this noise
- 100 years ago listening sticks were used. Today electronic microphones combined with digital correlators are used



# New techniques for leak detection

- By sending radio frequency signal into the ground and detecting the reflection.
  - Flowing water from a leak causes phase and amplitude modulation of the reflected signal
  - The reflected signal is detected by an interferometer
- By injecting a fluid transient into the pipeline
  - Leaks in pipeline contribute to damping of transient events
  - By analyzing the transient trace it is possible to locate a leak
  - Transient pressure waves used for attenuation studies
- Using Artificial Neural Network (ANN) based models
  - ANN based on pressure and flow rate information
  - Measurements from pressure and flow rate sensors distributed on the network branches and nodes
  - The ANN recognizes patterns among such data, and classifies them to identify and locate the leak

# New Techniques for leak detection

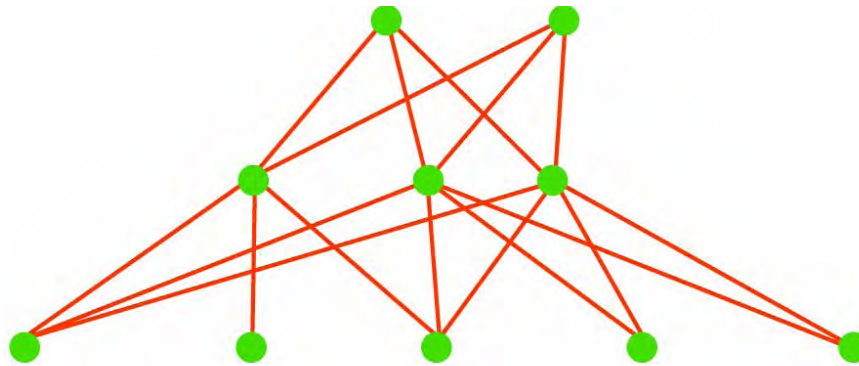
- Current technologies for detection are based on dependant and non-dependant process parameters
- Technologies based on Non-dependant parameters include:
  - analysis of acoustic emissions from leaks
  - pressure wave transmission
  - identification of radar or radio frequencies emitted from transmitters located inside the pipes, etc.
- These are expensive, often intrusive, and labour intensive
- Dependant parameters are Pressure ( $p$ ) and Flow Rate ( $q$ )
- However, computation of boundary conditions (i.e. leaks) starting from measured  $p$  and  $q$  values at the moment presents mathematical difficulties
- Therefore this is a classical inverse engineering problem:
  - determination of causes (leaks) starting from knowledge of their effects ( $p$ - $q$  maps)
  - suited for pattern recognition based models.

# Method using Artificial Neural Networks

- Deterministic models often used for simulation of various processes such as
  - hydrodynamic, morphological, ecological, water quality, groundwater flow, etc
- These require detailed information on underlying processes
- ANNs do not require the explicit knowledge of physical processes and the relations can be fitted on the basis of measured data
- If significant variables are known, without knowing the exact relationships, ANN is suitable to perform a kind of function fitting

# Method using Artificial Neural Networks

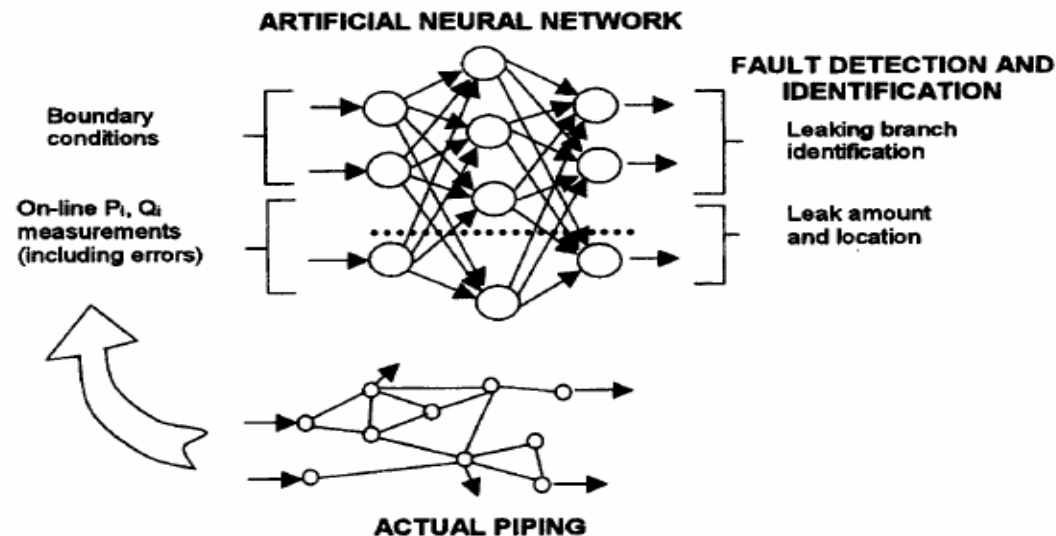
- ANNs are made up of a collection of nodes connected together by a number of arcs



- The arcs have weights associated with them
- Neural networks can be trained on a collection of input pattern and output pattern pairs
- During training, the weights are adjusted to minimise the least mean squares deviation between the predicted output and the actual output pattern

# Applying ANN method for leaks

- In the methodology, an ANN is fed data representative of the system status (leak size) and measurements from pressure sensors distributed on the network branches and nodes
- The ANN recognizes patterns among such data, and classifies them to identify and locate the leak
- In this way, the ANN is used to correlate effects (pressure conditions) with causes (network's boundary conditions such as leaks)



# Using Support Vector Machine

- Support Vector Machines (SVMs) are related to ANNs since:

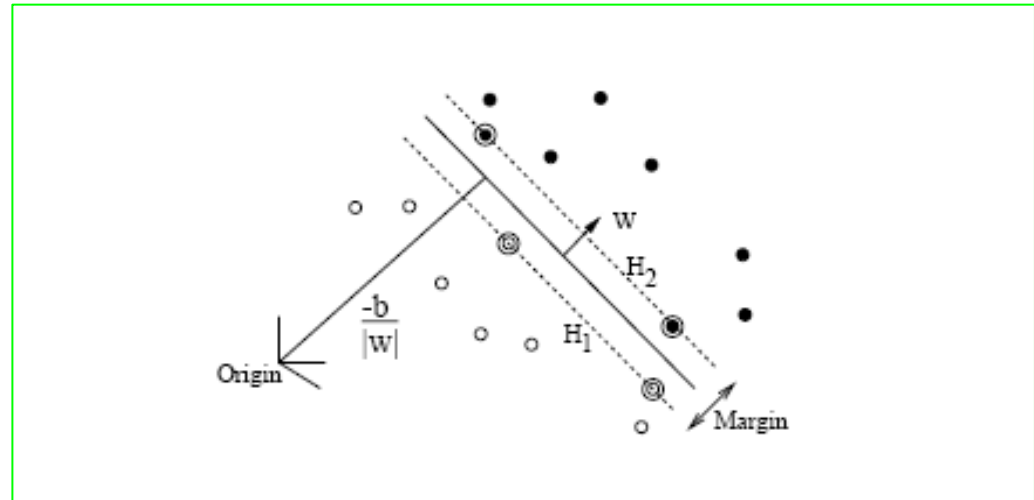
- SVMs are statistical pattern recognisers that can be used for classification or regression
- The SVM output is a predicted class (classification) or a real number (regression)

- SVMs have some properties that make them preferable to ANNs:

- SVMs can act effectively on very high dimensionality input pattern spaces
- They can be trained with smaller training sets than ANNs

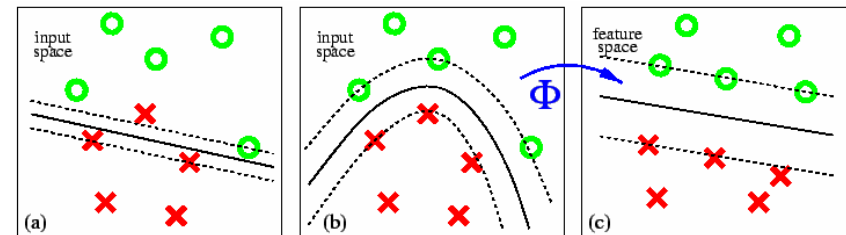
# Principles of support vector machines

- Structural risk minimisation
- Separating hyperplanes
  - for binary classification problems
- Margin maximisation
- Quadratic optimisation problem
- Sequential minimal optimisation



# Non-linear problems

- First pre-process the data by some non-linear mapping  $\Phi$
- Then apply the linear algorithm just described to the image space of  $\Phi$



# Method using SVM

- Leak detection using Support Vector Machine
  - The method is to monitor the pressure at a number of nodes in a pipe network and to feed these values to train the SVMs to predict leak size and leak location
  - The SVMs require 100s or possibly 1000s of cases in their training sets
  - Therefore it is not possible to generate the training sets by introducing actual leaks into a real pipe network
    - The training sets can be generated by simulation of the pipe network under consideration
    - This can be achieved by using the simulation tool EPANET

# Method using SVM

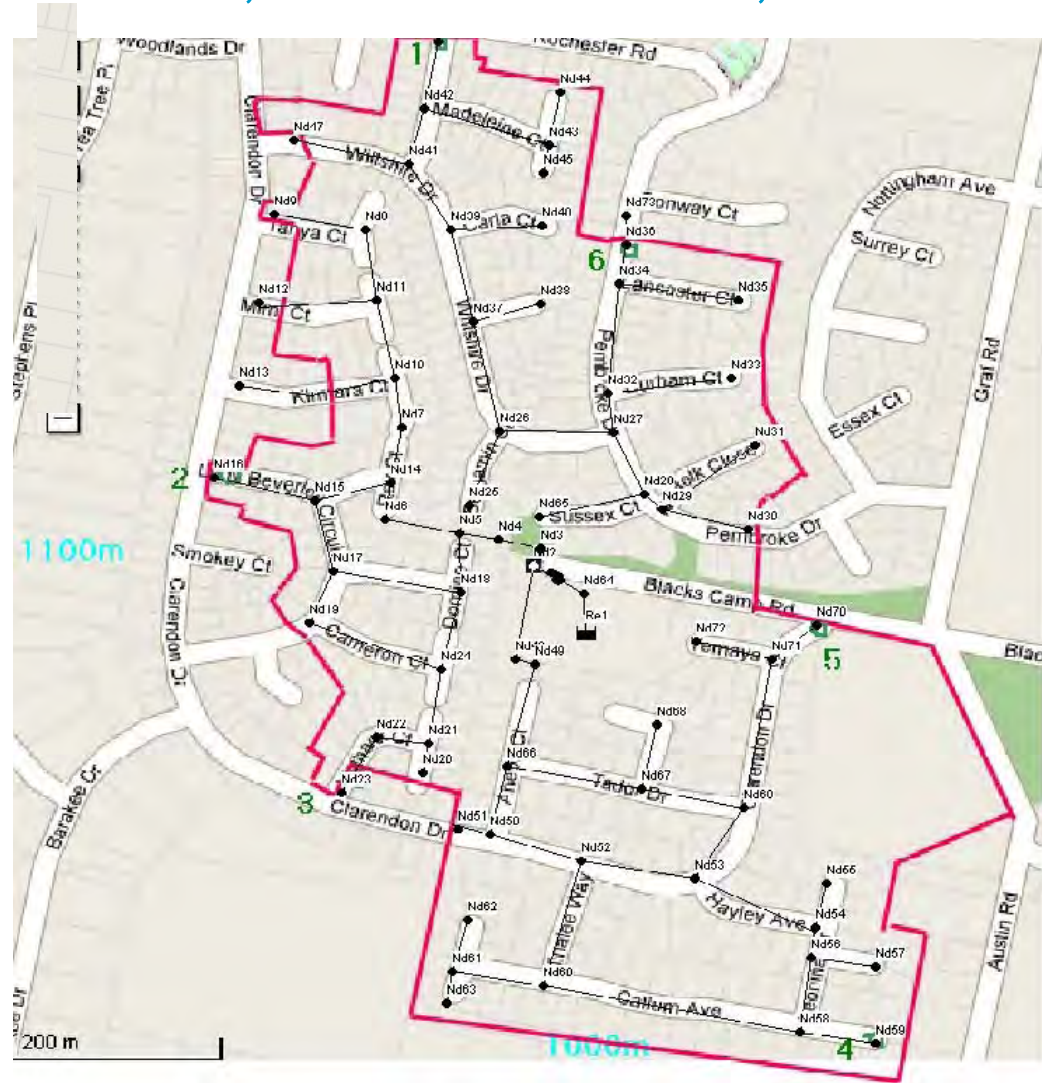
- EPANET

- EPANET is a computerised hydraulic simulation model produced by the Environmental Protection Agency of the USA
- It predicts the dynamic hydraulic and water quality behaviour of a drinking water distribution system
- Leaks of various sizes can be simulated in EPANET using the 'EMITTER' function, and the resulting pressures and flows in the network can be calculated
- For SVM training set generation, the running of EPANET was automated by means of a program which calls EPANET many times with various values of the leak size

# Case Study

A 400 property area in Somerville, SE of Melbourne, selected by SEWL

- The zone has one input; -supply from a single pump station
- It is a new network with almost no leaks
- This allows leak simulation by opening fire hydrants
- Also allows pressure monitoring through fire hydrants





# Leak size Prediction

- The EPANET driving program was used to generate 300 cases corresponding to node Nd73 leaking with emitter values ranging from 0 to 0.3 in steps of 0.001
- 200 cases were randomly selected as a training set and the rest were used as a testing set
- On testing, the mean squared error (MSE) and squared correlation coefficient ( $r^2$ ) were:  
MSE = 4.476e-005,  
 $r^2 = 0.994$   
which indicate good prediction accuracy

# Leak Location Prediction

The next experiment was to determine the effectiveness of using SVM for leak location prediction when there are 10 candidate leak nodes.

- The EPANET driving program was used to generate data representing the pressures at the 6 monitoring nodes
  - when one of the candidate leak nodes is leaking with emitter values going from 0 to 0.3 in steps of 0.002
  - then repeated for the other 9 leak nodes
- 10 data sets of 150 cases each were generated (1500 cases)
- 1000 cases were randomly selected as a training set and the rest were used as a testing set
- After optimisation of the training parameters the best testing accuracy obtained was 76.8%
  - This means that the predicted leaking node equals the actual leaking node in 76.8% of the cases in the testing set

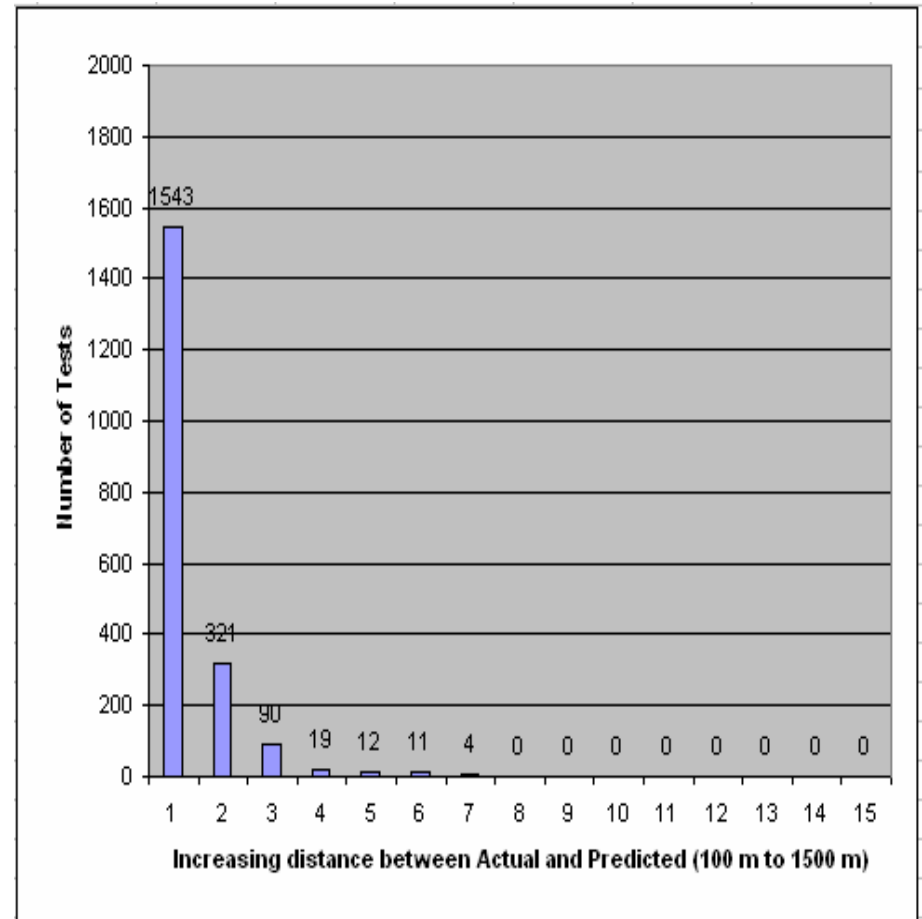
# Leak location prediction

- Similar experiments were carried out for the cases of 20 and 40 candidate leak nodes with resulting accuracies of 77.2% and 57.25%
- Accuracy of 57.25% for the case of 40 candidate leak nodes not so bad because
  - if the system were guessing randomly it would get the right answer 1 time in 40 or 2.5% of the time

# Leak location prediction

How far the actual leaking node is from the predicted leaking node?

- The distance between the predicted and the actual in the case of 40 candidate leaking nodes and a testing set of 2000 cases was analysed
- The distance interval between 0 m and 1500 m was divided up into 15 bins

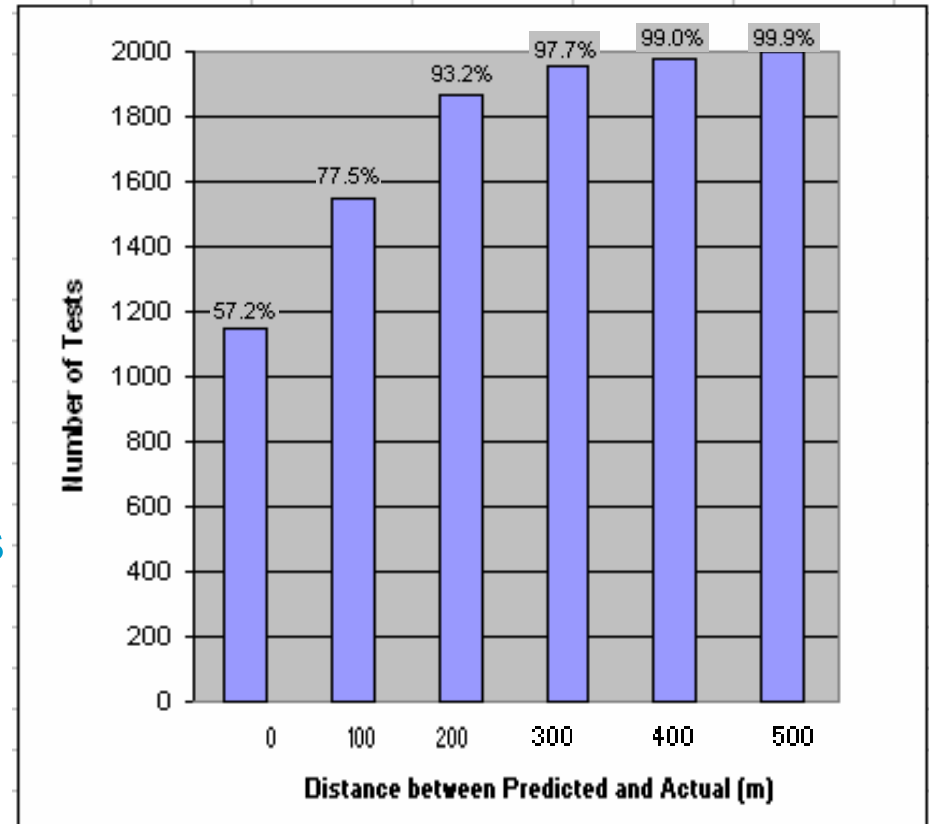


# Leak location prediction

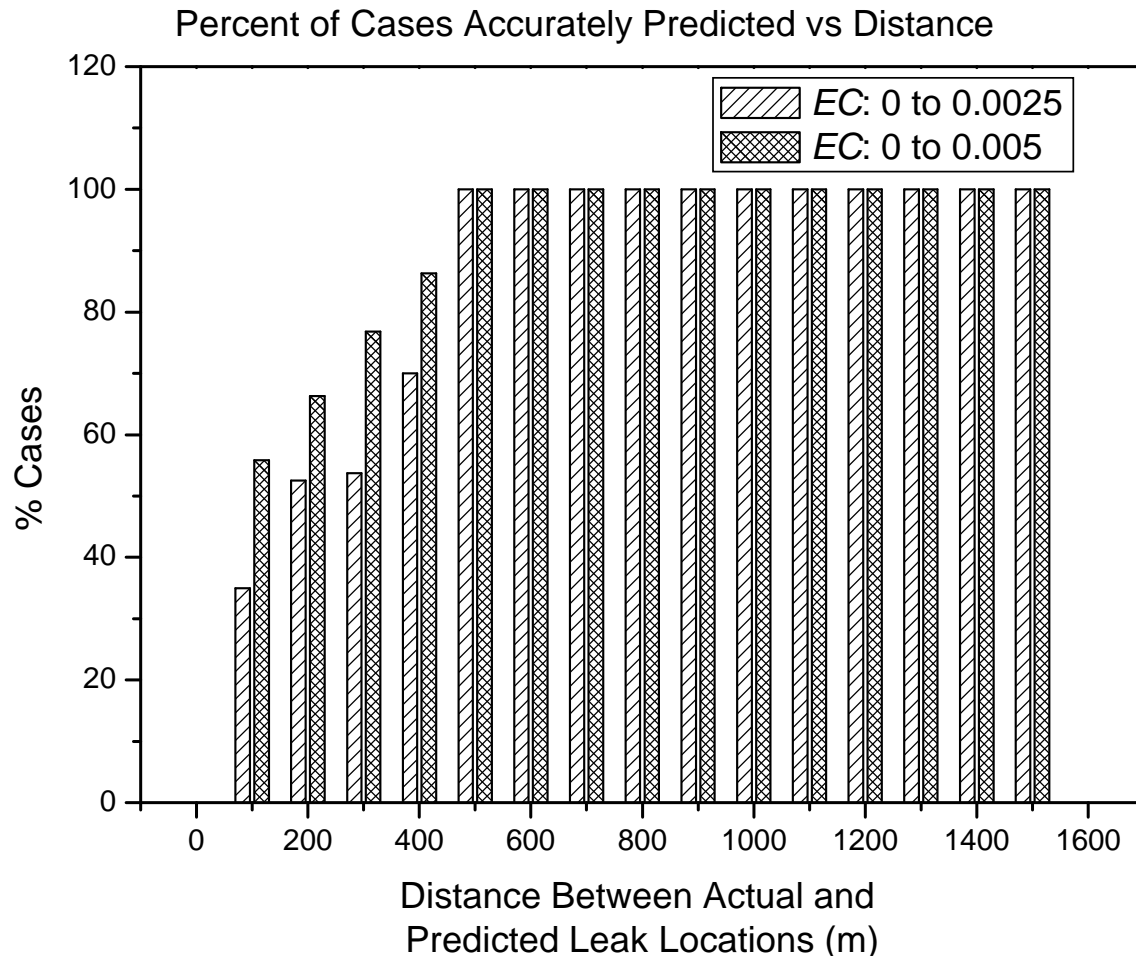
The predicted node equals the actual node in 57.25% of the cases, however:

- predicted node is within 100 m of the actual node in 77.5% of the cases
- predicted node is within 200 m of the actual node in 93.2% of the cases and within 300 m of the actual node in 97.7% of the cases

Thus the prediction would provide useful information to the water authority to direct their search for the leaking node



# SVM Performance for lower leak sizes



# Conclusion

- Desktop study has shown that using EPANET simulation for SVM training it may be possible to detect leaks of size down to 90 litres per hour
- For a leak of size 90  $\text{Lhr}^{-1}$  the pressure difference between the leak and no leak situation resulting from EPANET simulation is of the order of 0.1 Pa
- This is too small to be detected by available pressure measuring devices
- Another group within CSIRO has been working towards the development of a sensitive flow measuring device
- We plan to carry out the analyses on the basis of flows rather than pressures

# Further work

- The next stage of the work is validation of the predictions against field data at the study area for South East Water
- Leaks will be simulated by opening fire hydrant(s) to a predetermined flow level(s)
- Comparison of the measured pressure against the predicted pressure will be used to validate the procedure
- It is anticipated that the validated system will be field tested in Ipswich before it is rolled out to other parts of SE Queensland

# Urban Water Security Research Alliance



THANK YOU

[www.urbanwateralliance.org.au](http://www.urbanwateralliance.org.au)

Email: [John.Mashford@csiro.au](mailto:John.Mashford@csiro.au)

