International Journal of Forecasting 19 (2003) 623–634

international journal of forecasting

www.elsevier.com/locate/ijforecast

Predicting the geo-temporal variations of crime and disorder

Jonathan J. Corcoran, Ian D. Wilson, J. Andrew Ware*

School of Computing, University of Glamorgan, Pontypridd, Mid Glamorgan, CF37 1DL, UK

Abstract

Traditional police boundaries—precincts, patrol districts, etc.—often fail to reflect the true distribution of criminal activity and thus do little to assist in the optimal allocation of police resources. This paper introduces methods for crime incident forecasting by focusing upon geographical areas of concern that transcend traditional policing boundaries. The computerised procedure utilises a geographical crime incidence-scanning algorithm to identify clusters with relatively high levels of crime (hot spots). These clusters provide sufficient data for training artificial neural networks (ANNs) capable of modelling trends within them. The approach to ANN specification and estimation is enhanced by application of a novel and noteworthy approach, the Gamma test (GT).

Keywords: Crime forecasting; Cluster analysis; Geographic information system; Artificial neural networks; Gamma test; Autoregressive model

1. Introduction

A system that intelligently interrogates a constantly updated database of crime incidence and provides accurate indicators of where and when crime is likely to be highest would be of great utility in real-time police resource allocation. A limiting factor, however, is that crime incidence counts are generally low in relation to crime type, time and space, and subject to randomness. Gorr, Olligschlaeger, and Thompson (2003) have shown that crime forecast error measures vary inversely with increasing incidence count utilised in estimating time series forecast models. Average crime counts per unit time period and geographic area of at least 25–35 are needed before forecast errors become acceptable.

This paper details a forecasting framework for short-term, tactical deployment of police resources in which the objective is the identification of areas where the levels of crime are high enough to enable accurate predictive models to be produced. This work differs from other recent studies dealing with hot-spot methods (e.g., Ratcliffe & McCullagh, 1999) and their statistical significance (for example, Chainey & Reid, 2002). Whereas these researchers employ hot-spot methods as a means of visualising and comprehending crime distributions, here their utility is extended to use identified hot-spot regions as the foundation for predictive models.

The methodology presented in this paper follows three key stages (summarised in Fig. 1). The first (spatial analysis) identifies geographical clusters; the second (cluster modelling) determines the data quality of each cluster; the third (prediction) develops a corresponding artificial neural network (ANN) model based on an autoregressive predictive specification.

*Corresponding author. Tel.: +44-1443-482650; fax: +44-1443-482711.

E-mail address: jaware@glam.ac.uk (A.J. Ware).

0169-2070/03/$ – see front matter
doi:10.1016/S0169-2070(03)00095-5

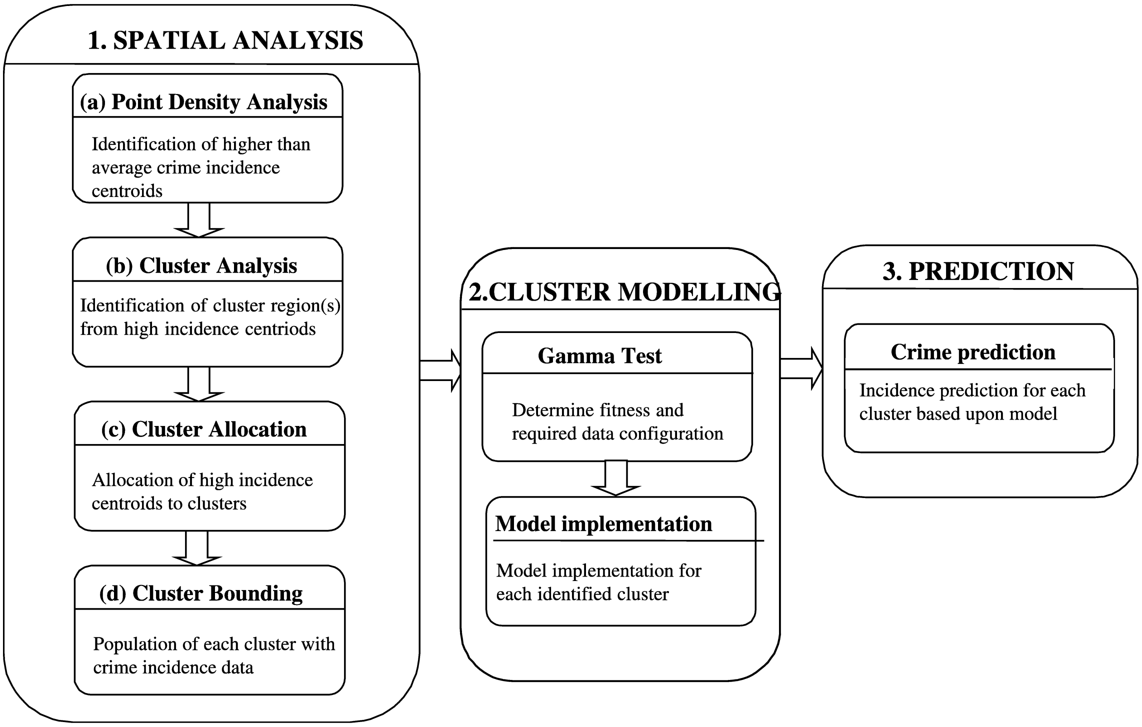

Fig. 1. The CLAP/Gamma test/predictive process model.

The paper demonstrates how artificial neural networks can be trained using geographical clusters of crime data to facilitate predictive modelling. The extent to which each cluster has the potential to facilitate prediction is estimated using a novel technique known as the Gamma test (GT) (Durrant, 2001; Evans & Jones, 2002; Jones, Evans, Margetts, & Durrant, 2002; Stefánsson, Koncar, & Jones, 1997). The paper also explains the details of the spatial analysis undertaken to geographically identify crime clusters. The paper concludes with a discussion of the results and focus for future research.

2. Artificial neural networks models for prediction

Building forecasting models with neural networks is not a new phenomenon (e.g., Balkin & Ord, 2000; Tkacz, 2001; Zhang, Patuwo, & Hu, 1998). In the case of crime level forecasting, the models tend to be autoregressive with input and output vectors being counts of crime: multiple inputs $y_{t-1} \ldots y_{t-n}$ and a single output y_t. The models built in this paper are of this type.

Testing the corresponding ANN involves presenting the network with a series of input vectors for which the tester, but not the network, knows the corresponding crime levels. The answers given by the network as to what it determines to be the level of crime, given the presented input vector, can then be used by the tester to determine the robustness of the training process. If the robustness of the network is deemed sufficient, the network can be used in a truly predictive capacity. Here, the network is presented with an input vector for which the output is not known and its answer is assumed reliable.

3. The Gamma test

The GT estimates the best mean square error (MSE) that can be achieved when modelling the data using any continuous model fitting method, such as least squares regression or an ANN, for an unknown function. Suppose that we have

$$y = f(\mathbf{x}) + \boldsymbol{\varepsilon} \quad (1)$$

where y is the given output of an unknown, smooth but unknown function f, $\mathbf{x}$ is a vector of inputs, and ε is the error term. Consider an input/output data set

$$\{(x_i,y_i)|1 \le i \le M\}. \quad (2)$$

The GT is based on $N[i,k]$, which are the kth ($1 \le k \le p$) nearest neighbours $x_{N[i,k]}$ ($1 \le i \le M$) for each vector x_i ($1 \le i \le M$). Specifically, the GT is derived from the Delta function of input vectors:

$$\delta_M(k) = \frac{1}{M}\sum_{i=1}^{M}|x_{N[i,k]} - x_i|^2 \qquad (1 \le k \le p) \quad (3)$$

where $|\ldots|$ denotes Euclidean distance, and corresponding Gamma function of output values:

$$\gamma_M(k) = \frac{1}{2M}\sum_{i=1}^{M}(y_{N[i,k]} - y_i)^2 \qquad (1 \le k \le p) \quad (4)$$

where $y_{N[i,k]}$ is the corresponding y-value for the kth nearest neighbour of x_i in Eq. (3). Next, we fit the regression line:

$$\gamma = A\delta + \Gamma \quad (5)$$

of the points $(\delta_M(k), \gamma_M(k))$ $(1 \le k \le p)$.

This algorithm produces $(\delta(k), \gamma(k))$ $(1 \le k \le p)$ coordinates or simply (δ, γ) pairs, which can be displayed using a two-dimensional scatter graph. Fig. 2 shows an example, where also shown is the estimated regression line $\gamma = A\delta + \Gamma$ for the p nearest neighbours. Here, similar inputs plotted along the horizontal, Delta axis will have similar output values plotted along the vertical, Gamma axis if a good $f(x)$ exists (Evans & Jones, 2002). However, noise within the data will result in different Gamma points for a given Delta value (illustrated using the frequency histogram show in Fig. 3). Ideally, the histogram should show a preponderance of points close to the origin, indicating that similar inputs are generally producing similar outputs. This provides a useful means for visualising patterns within numerical input/output data.

The graphical output, specifically the regression line (shown in Fig. 2), provides two indicators. First, it is remarkable that the vertical intercept Γ of the y (or Gamma) axis offers an estimate of the best MSE achievable utilising a modelling technique for unknown smooth functions of continuous variables (Evans & Jones, 2002). Second, the gradient A offers an indication of model complexity (where a steeper gradient indicates a model of greater complexity). Results can indicate variations in the two variables (e.g., estimates of low MSE being associated with a high level of complexity), with the preferred scenario being a low MSE and shallow gradient.

Using the estimated MSE, a useful test is to establish the minimum quantity of data points needed

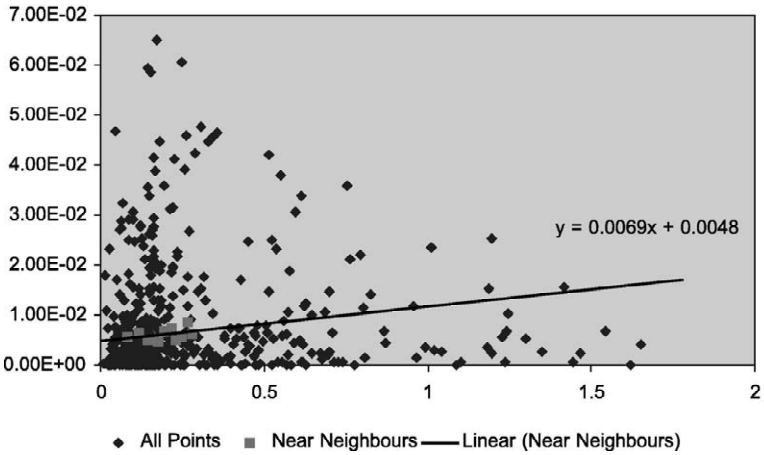

Fig. 2. Gamma test 2D graphical analysis.

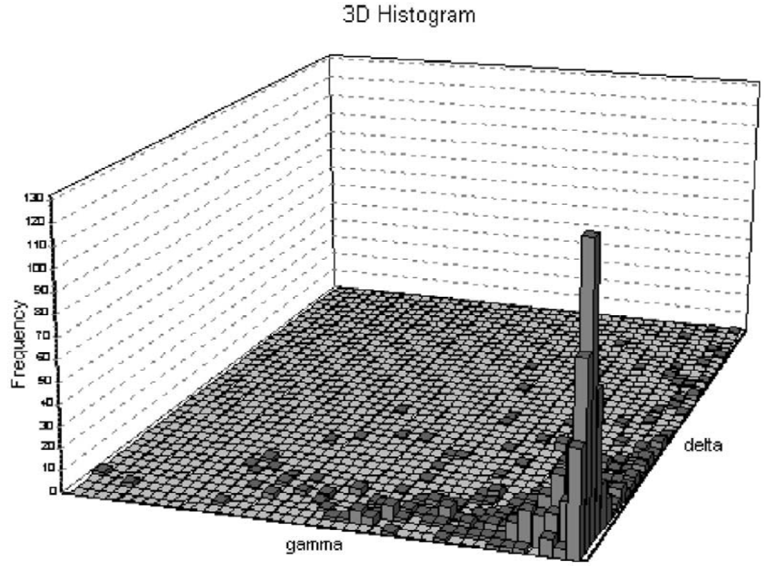

Fig. 3. Histogram of Delta, Gamma plot frequencies.

to model an underlying function. This is accomplished using the M-test, where the GT is applied to an increasing sample size $(M_1 \ldots M_m)$ and the Gamma value plotted against that of M. In an ideal model, the output may exhibit large variation in Gamma at small values of M but ultimately stabilise at a higher value of M, which is indicative of the true noise variance inherent within the data. The region of asymptotic values of Gamma against M identifies the minimum data required to establish best possible accuracy in prediction.

The two indicators from the GT offer a basis from which an ANN model can be assembled and trained, given that an estimate of the best possible MSE has been provided by the GT. Hence, model training can be terminated once the estimated MSE is reached, thus avoiding overfitting.

4. The crime incident data set

The data used in this study are 18,498 violent incidents (violence against the person, criminal damage, and disorder), spanning 1 year in an urban area measuring approximately 242,700,000 m^2. Given this, it can be said that on average, one crime took place during the year per 13,120 m^2, or approximately one crime per 65-m radius. Included in the database of crime incidents are a number of variables relating to time, day, month, weather, and location (represented as geographical coordinates).

4.1. Training sets from crime clusters

For this analysis, a similar technique to the GAM/1 geographical analysis machine developed by Openshaw (1987, 1988) was used, augmented to allow the clustering of centroids of high incidence. This four-stage process consists of

- point density analysis,
- geographic representation and cluster analysis,
- allocation of centroids to clusters, and
- relation of incidents to cluster boundaries.

4.1.1. Stage one: point density analysis

A simple search algorithm, provided in Appendix A2, identifies small areas of greater than average crime incidence. The ensuing results for the test data set are illustrated in Fig. 4.

4.1.2. Stage two: geographic representation and cluster analysis

At this stage, a heuristic approach is taken to determine the level of crime incidence required for a cluster to be considered salient. The heuristic rules

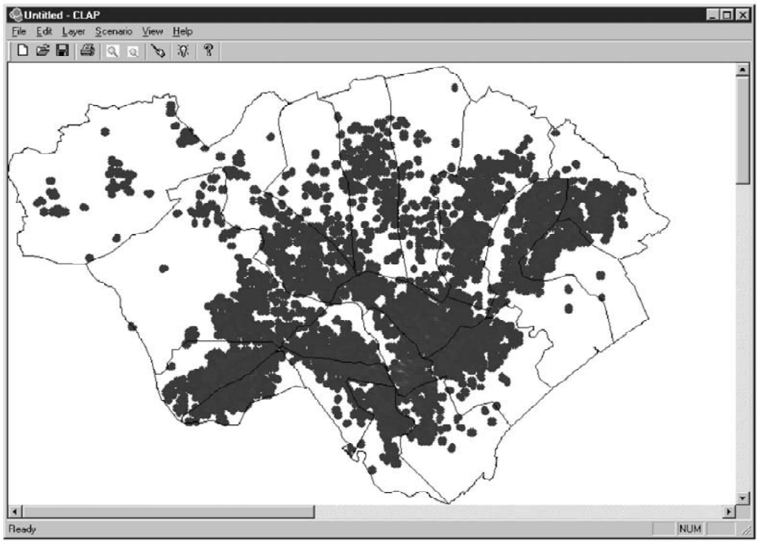

Fig. 4. Centroids of higher than average crime incidence.

utilised to make this determination are based on the assumption that most incidence of crime tends to be concentrated within relatively small geographic areas (i.e., ''hot spots'').

Given this, a scatter graph representation of the geographical data was utilised to heuristically increase both the radius of the area associated with that centroid. As the density of the centroids displayed increases, so does the radius of influence associated with that centroid. Experimentation and user interaction resulted in the radius of influence, or *gravity*, being set to *density**40, where *density* is the count of crimes associated with the centroid during stage one of the analysis process. This results in the identification of seven clusters of interest (illustrated in Fig. 5 below).

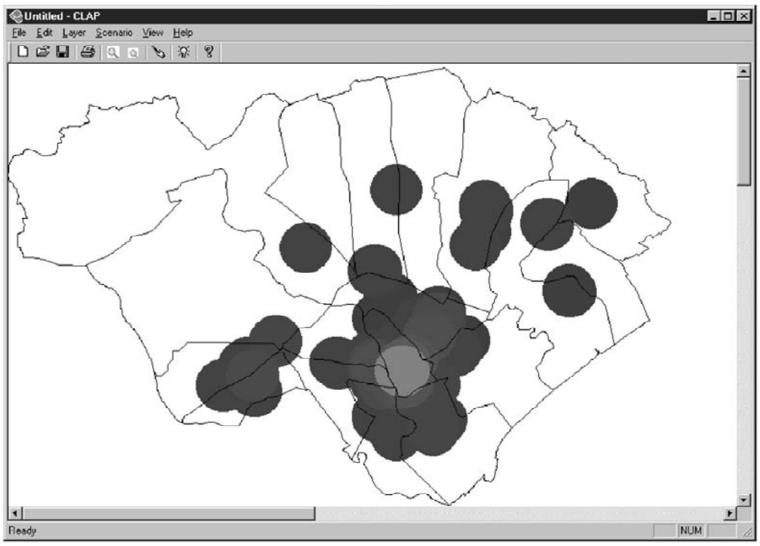

Fig. 5. Clusters of high crime incidence.

Table 1
Crime incidence by cluster

Cluster	Centroids	Crime count
1	0, 4, 5, 6	1254
2	1, 2, 3	50
3	7, 8, 9, 10, 11, 12	954
4	13–109, 111–148, 165, 166	4097
5	110	161
6	149	228
7	150–164	2094
Total clustered:		8838
Total violent crime:		18,498

4.1.3. Stage three: allocation of centroids to clusters

Next, the centroids that should be grouped together to form clusters are identified. The density and gravity parameters, together with a centroid list generated in stage two, forms the basis for this iterative procedure.

4.1.4. Stage four: relating incidents to cluster boundaries

Finally, each of the clusters is populated with data ready for training a series of ANNs (one per cluster). Each crime record contains a unique identifier, the cluster it belongs to and the weekday during which the crime was committed. In addition, each cluster record has a unique identifier, a list of its member centroids and a total crime count (see Fig. 5 and Table 1).

5. Application of Gamma test to the cluster data

Taking the results from the cluster analysis, two techniques were used to model the clustered data. The first sought to model day of week against crime volume, the second used an autoregressive formulation. Initial attempts using the first technique failed to achieve the accuracy accomplished by the second. The autoregressive model was selected as preferred for the short-term forecasting problem. The following discusses the associated methodology.

6. Forecasting using artificial neural networks

Implementation of an ANN model requires careful consideration of model parameters impacting model stability and efficiency. These included decisions that concern architecture type (number of input/output nodes and hidden layers), selection of training algorithm, and volume of data to be used for training and testing.

6.1. The network architecture and estimation

The ANN presented in this paper comprises of an input layer (corresponding to the length of the input vector), two layers of hidden nodes, and an output layer providing the forecast value. Modelling a time series involves generating a set of input vectors and corresponding output values. The optimal number of lags for the autoregressive model is established by generating a Gamma statistic for incremental lag lengths. The lag length is that with Gamma statistic closest to zero (for example, 13 for the city centre cluster as shown in Fig. 6).

The best topology for nodes in the hidden layers was also determined empirically. Previous research has indicated that use of a single hidden layer is sufficient to learn any complex nonlinear function (Hornik, 1991). However, Chester (1990), Srinivasan, Liew, and Chang (1994), and Zhang (1994) suggest that two hidden layers can produce more efficient architectures.

Hence, we use two hidden layers. A one-step-ahead forecasting horizon was used, represented by a single node in the output layer.

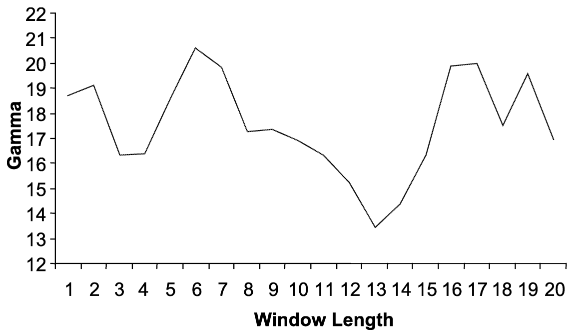

Fig. 6. Example of increasing embedding (city centre, cluster four).

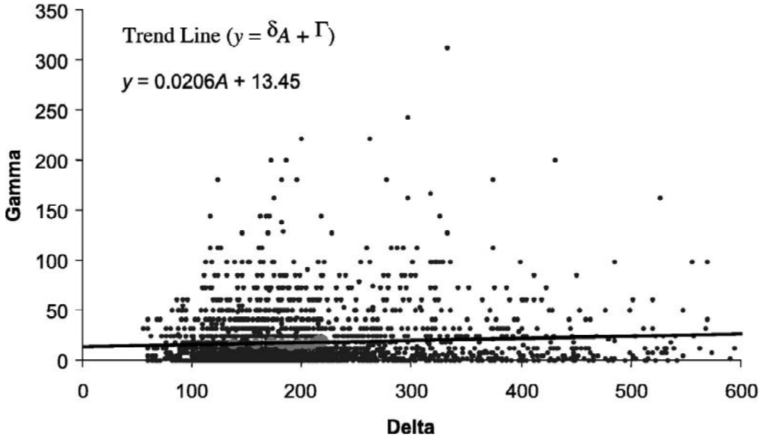

Fig. 7. Gamma scatter plot showing the trend line.

Initial large numbers of nodes ($2N+1$ split evenly between both hidden layers, where N is the number of inputs) in the hidden layers were incrementally reduced to a minimum whilst maintaining acceptable forecasting capabilities. The shallow gradient (shown in Fig. 7 for the city centre cluster) suggested that a relatively few number of hidden nodes in proportion to the $2N+1$ rule would be sufficient to model the underlying function, and this proved to be the case.

The standard gradient descent method for adjusting weights is replaced with conjugate gradient descent (Bishop, 1996), which uses past gradient measures to improve the error minimisation process.

6.2. Terminating the training procedure

As overfitting is a widely accepted problem associated with modelling utilising ANNs, the GT's ability to accurately measure the noise within a data set, and consequently, the point at which training should stop provides a significant utility for practitioners. Overfitting occurs because the ANN will eventually attempt to fit all data encountered, including any noise present. Providing a measure of noise present in the data set allows training to be terminated at a near optimal point. This is because an ANN will tend to fit useful data before any noise. Therefore, the GT statistic Γ provides an MSE value at which training can be stopped (for example, an approximate target value of 13.45, shown in Fig. 7, for the city centre cluster).

6.3. Partitioning the vectors into training and test sets

Once the number of inputs required to model the output is known, the data can be transformed to fit the optimal set-up. Using this set-up, an M-test is performed to establish whether the available number of vectors is sufficient to model any underlying function. An asymptotic level for the Gamma statistic (which approximates to the inherent noise of the output) indicates that there is sufficient data and provides a point where the data can be split into

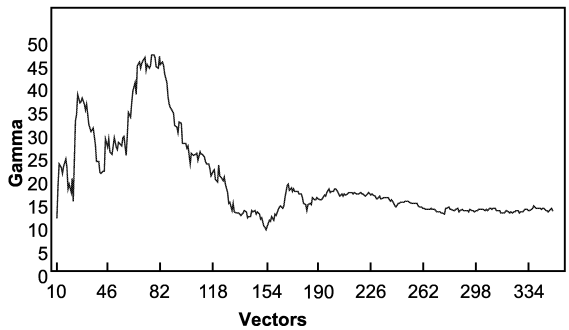

Fig. 8. M-test for the city centre.

training and test vector sets. This is an important consideration as it allows the data set to be split into two rather than the commonly practised three parts (the training, validation, and test set). Consequently, Wilson, Paris, Ware, and Jenkins (2002) has demonstrated that there is no need to set aside part of the data as a validation set, which is used to determine when continuing to train would result in overfitting allowing a higher proportion of data to be utilised during training. Thus, selection of the appropriate amount of data for modelling is confirmed at a point in advance of where the *M*-test reaches a stable level indicative of inherent noise (e.g., an approximate volume of 300, shown in Fig. 8, might be taken as the partitioning point).

7. Experimental work

This section compares forecast performance of the ANN model just discussed with a linear regression model and a modification of the random walk. The modified random walk forecasts the change from t_t to t_{t+1} based upon the average change from one period to the next. For example, taking the known number of crimes for a Thursday, the forecast for Friday is based upon the average observed change (over the entire time series) between Thursday and Friday.

7.1. Comparison between ANN, linear regression and random walk

As an example of the results obtained, the ANN models discussed here focus on two clusters analysed according to daily incident count.

7.2. Cluster seven (residential)

Cluster Seven, which is a residential area with very few owner–occupiers, showed almost no general correlation between incident rate and weekday (illustrated in Fig. 9). However, an increased tendency for violent crime towards weekends was noted, warranting a closer examination of other causal factors.

The GT procedures were utilised to determine the model input length (30 lagged observations of the dependent variable), the partitioning of the data set into training (308 observations) and test (28 observations) sets, and to provide an estimate of the inherent noise within the data (8.01 robberies, or 32% of the range, a high value indicating a very chaotic data series). In addition, the gradient statistic estimate (0.0159) suggested that relatively few hidden nodes (10 in each of the 2 hidden layers) would be required to reach the estimated Gamma statistic. The ANN, regression, and RW models resulted in forecast MAPEs of 31.3%, 30.5%, and 30.9% of the observed

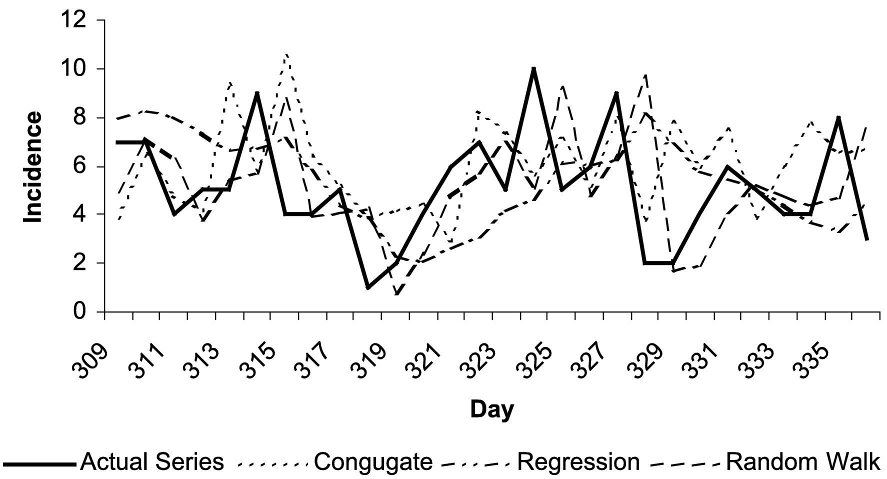

Fig. 9. Incidence and forecast of violent crime (cluster 7).

data range of 9, respectively. Forecast MSEs were 7.95, 7.61, and 7.77, respectively. With such low incidence rates, this cluster is difficult to forecast accurately and none of the three methods excels.

7.3. Cluster four (city centre)

Cluster four is the city centre and includes a concentrated collection of night clubs, public houses, public transport centre, and a sporting stadium. It showed a higher incidence of crime on weekends, with peeks during times of known sporting events.

The GT procedures were utilised to determine the model input length (13 lagged observations of the dependant variable), data set partitioning into training (330 vectors) and test (23 vectors) sets, and to estimate the inherent noise within the data (13.17, or 27% of the range). In addition, the gradient statistic estimate (0.0159) suggested that relatively few hidden nodes (five in each of the two hidden layers) would be required to reach the estimated Gamma statistic. The results generated by the resultant ANN, regression, and RW models (shown in Fig. 10) produced forecast MAPEs of 24.2%, 33.5%, and 36.4% of the observed data range of 13, respectively. The forecast MSEs were 9.94, 18.96, and 22.500, respectively. In this case, the ANN forecasts much more accurately than the regression or modified random walk methods.

7.4. Discussion of results

The results concurred with expectations given the GT's output. Importantly, exceptional incidence levels that occur infrequently appear as noise and are excluded from the underlying model accounting for a large portion of the error margin. The utility of the GT was demonstrated as a premodel evaluation technique. The city centre (cluster four) offered the best predictive model using the ANN, cluster seven (residential area) generated relatively poor models for ANN, regression, and RW. Further experiments are now needed with an initial data set covering a longer period.

8. Conclusions and future work

This paper introduces a forecasting framework (summarised in Fig. 1) focusing upon geographical areas of concern that may well transcend traditional policing boundaries. The paper focuses upon the development of a practical solution for use in an operational policing environment, which ameliorates the deficiencies of rigid boundaries and moves towards a more dynamic methodology. The computerised procedure utilises a geographical crime incidence-scanning algorithm to identify clusters with relatively high levels of crime (hot spots).

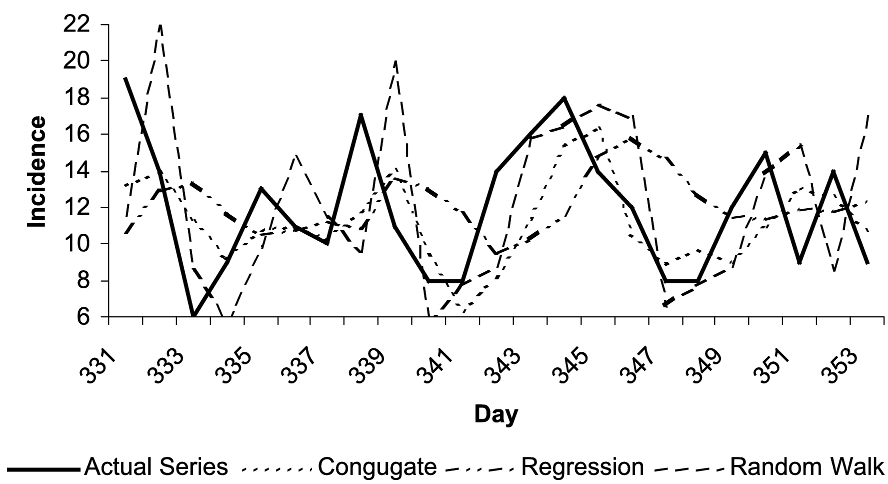

Fig. 10. Incidence and forecast of violent crime (city centre).

These clusters provide sufficient data that can be analysed using the GT procedures assessing fitness and required configuration (for example, quantity of data required and number of inputs) for predictive modelling. Using the outputs from the GT, two techniques were implemented (ANN and regression). The ANN generally exhibits a superior capacity to model the trends within each cluster. A modified random walk was used a naïve forecasting method, the results demonstrating a comparable forecasting accuracy to the other techniques for cluster seven (residential area) where the GT indicated a chaotic data series.

Future developments will include the modelling of more detailed scenarios to facilitate prediction based upon selected input criteria. Thus, for example, the impact upon the region of say a forthcoming public holiday, where the weather is predicted to be warm, could be evaluated. The objective here was to extract an underlying generalised model of crime incidence. However, specific localities might best be modelled independently of the other data at specific times of the year (for example, sporting events that generate exceptionally high crime *spikes* that fall out of a generalised model). These spikes could be extracted and treated as a separate modelling exercise, given sufficient high quality data. Alternatively, a statistical analysis of exceptional events would provide an estimate of the change against normal levels that could subsequently be encoded as a set of rules that modifies the incidence count accordingly. These two differing approaches will form the basis for continued experimentation.

Appendix A. Computational algorithms

A.1. The Gamma test algorithm

```
Load input/output vectors {(x_i, y_i) : 1 ≤ i ≤ M} where x_i ∈ P^m and y_i ∈ P.
For i = 1 to M
    For j = 1 to M
            InputDistance[i, j] = |x_i - x_j| (n-dimension Euclidean distance)
            OutputDistance[i, j] = |y_i - y_j|
        Next j
Next i
Sort corresponding (InputDistance[i, j], OutputDistance[i, j]) into ascending order.
For i = 1 to M-1
    For j = 1 to M
        If (InputDistance[i, j] = InputDistance[i+1, j]) then
            OutputDistance[i, j] = (OutputDistance[i, j] + OutputDistance[i+1, j]) / 2
            InputDistance[i+1, j] = InputDistance[i+2, j]
        Endif
    Next j
Next i
For i = 1 to M
        For k = 1 to p
            δ(k) = δ(k) + InputDistance[i, k]^2 (Sum squares of input distances.)
            γ(k) = γ(k) + OutputDistance[i, k]^2 (Sum squares of output distances.)
        Next k
Next i
For k = 1 to p
        δ(k) = δ(k) / M (Calculate the average.)
        γ(k) = γ(k) / 2M (Calculate the half of the average.)
Next k
Compute linear regression on (δ(k), γ(k)) where (1 ≤ k ≤ p) as in (4).
```

A.2. Incidence density algorithm

```
ScanRadius is the radius of circle within which points will be counted.
StepSize is the distance that the centroid will be moved at each iteration.
StartX is the bottom left longitudinal start co-ordinate.
StartY is the bottom left latitudinal start co-ordinate.
EndX is the top right longitudinal end co-ordinate.
EndY is the top right latitudinal end co-ordinate.
Average is the number of crimes per area of circle.
For X = StartX to EndX step StepSize (moves longitudinally across the map)
   For Y = StartY to EndY step StepSize (moves latitudinally up the map)
      Count the number of points within ScanRadius
      If Count is greater than Average then
        Add co-ordinates to centroid list.
      Endif

   Next Y

Next X
```

Count is a function that iterates through the crime data, incrementing a counter each time one is found to be within *ScanRadius* of the centroid (*X*, *Y* coordinates), making use of

$$\text{if } (ScanRadius < \sqrt{(CrimeX - X)^2 + (CrimeY - Y)^2}) \text{ then add 1 to count} \qquad (6)$$

where *CrimeX* and *CrimeY* are projected coordinates of the crime incident.

References

Balkin, S. D., & Ord, J. K. (2000). Automatic neural network modelling for univariate time series. *International Journal of Forecasting*, *16*, 509–515.

Bishop, C. M. (1996). *Neural networks for pattern recognition*. Oxford: Oxford University Press.

Chainey, S., & Reid, S. (2002). When is a hotspot a hotspot? A procedure for creating statistically robust hotspot maps of crime. In Kidner, D. B. et al. (Ed.), *Socio-economic applications in geographical information science*. London: Taylor and Francis, pp. 21–36.

Chester, D. L. (1990). Why two hidden layers are better than one? *Proceedings of the International Joint Conference on Neural Networks*, *1*, 265–268.

Durrant, P. J. (2001). *winGamma*™: A non-linear data analysis and modelling tool with applications to flood prediction. Ph.D. thesis, Department of Computer Science, University of Wales, Cardiff, Wales UK.

Evans, D., & Jones, A. J. (2002). A proof of the Gamma test. *Proceedings of the Royal Society of London*. *A*, *458*, 1–41.

Gorr, W., Olligschlaeger, A., & Thompson, Y. (2003). Short-term forecasting of Crime. *International Journal of Forecasting, 19*, 579–594.

Hornik, K. (1991). Approximation capabilities of multi-layer feed-forward networks. *Neural Networks*, *4*, 251–257.

Jones, J. A., Evans, D., Margetts, S., & Durrant, P. J. (2002). The gamma test. In Sarker, R., Abbass, H., & Newton, C. (Eds.), *Heuristic and optimization for knowledge discovery*. Hershey: Idea Group Publishing, pp. 142–168.

Openshaw, S. (1987). An automated geographical analysis system. *Environment and Planning*. *A*, *19*, 431–436.

Openshaw, S. (1988). Leukaemia patterns in Northern England: A new method of finding cancer clusters. *Northern Economic Review*, *16*, 52–69.

Ratcliffe, J. H., & McCullagh, M. J. (1999). Hotbeds of crime and the search for spatial accuracy. *Journal of Geographical Systems*, *1*, 385–398.

Srinivasan, D., Liew, A. C., & Chang, C. S. (1994). A neural network short-term load forecaster. *Electric Power Systems Research*, *28*, 227–234.

Stefánsson, A., Koncar, N., & Jones, A. J. (1997). A note on the Gamma test. *Neural Computing and Applications*, *5*, 131–133.

Tkacz, G. (2001). Neural network forecasting of Canadian GDP growth. *International Journal of Forecasting*, *17*, 57–69.

Wilson, I. D., Paris, S. D., Ware, J. A., & Jenkins, D. H. (2002). Residential property price time series forecasting with neural networks. *Knowledge-Based Systems*, *15*, 335–341.

Zhang, G. B., Patuwo, E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35–62.

Zhang, X. (1994). Time series analysis and prediction by neural networks. *Optimization Methods and Software*, *4*, 151–170.

Biographies: Jonathan CORCORAN is a researcher based in the GIS Research Group, at the University of Glamorgan. His research interests include the exploitation of Geographical Information Systems for crime mapping and the use of neural networks to predict the development of emerging crime hot spots.

Ian D. WILSON is a Research Fellow within the Artificial Intelligence Technologies Modelling Group, at the University of Glamorgan. Ian has broad research interests including knowledge engineering, system modelling and complex combinatorial problem optimisation.

Andrew WARE is the leader of the Artificial Intelligence Modelling Group, at the University of Glamorgan. He has acted as consultant to industry in the deployment of Artificial Intelligence techniques to help solve real world problems. His interests include housing in the third world and he has worked with Habitat for Humanity on projects in India, Nepal, Uganda, South Africa and Mexico.