Jurnal
Teknologi

Full Paper

MODELING OF VIOLENT CRIME RATES WITH ECONOMIC INDICATORS USING HYBRIDIZATION OF GREY RELATIONAL ANALYSIS AND SUPPORT VECTOR REGRESSION

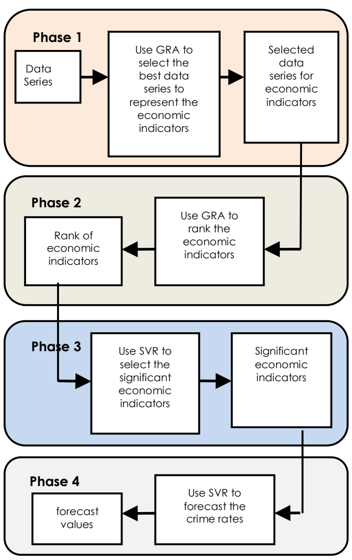
Article history
Received
7 January 2015
Received in revised form
24 March 2015
Accepted
1 August 2015

Razana Alwee[a*], Siti Mariyam Shamsuddin[b], Roselina Sallehuddin[a]

[a]Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]UTM Big Data Centre, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

*Corresponding author
razana@utm.my

Graphical abstract



Abstract

Regression and econometric models are commonly applied in modeling of violent crime rates. However, these models are mainly linear and only capable in modeling linear relationships. Moreover, the econometric models are quite complex to develop. Although time series model is a promising alternative tool, limited historical data of crime rates makes the standard time series models less suitable for modeling the violent crime rates. Thus, in this study, a hybrid model that can handle limited historical data is proposed for modeling the violent crime rates. The proposed hybrid model combines grey relational analysis and support vector regression. Since inaccurate parameters setting leads to inaccuracy of support vector regression model, particle swarm optimization is used to increase the accuracy of the model. The proposed hybrid model is used to model the violent crime rates of United State based on economic indicators. The proposed model also has additional features such as able to choose the data series for economic indicators and significant economic indicators for the violent crime rates. The experimental results showed that the proposed model produces more accurate forecast as compared to multiple linear regression in forecasting the violent crime rates.

Keywords: Support vector regression, grey relational analysis, particle swarm optimization, violent crime rates, economic indicators

Abstrak

Model regresi dan ekonometrik sering digunakan dalam pemodelan kadar jenayah keganasan. Walau bagaimanapun, model regresi dan ekonometrik adalah model linear yang hanya mampu memodelkan hubungan linear. Selain itu, model ekonometrik agak kompleks untuk dibangunkan. Walaupun, kaedah siri masa dianggap sebagai kaedah alternatif untuk pemodelan kadar jenayah keganasan, namun data sejarah yang terhad menjadikan model siri masa piawai kurang sesuai untuk pemodelan kadar jenayah keganasan. Oleh itu, kajian ini mencadangkan model hibrid yang mampu mengendalikan data sejarah yang terhad bagi memodelkan kadar jenayah keganasan. Model hibrid yang dicadangkan menggabungkan analisis hubungan kelabu dan regresi vektor sokongan. Oleh kerana, ketepatan model regresi vektor sokongan bergantung pada tetapan nilai parameternya, pengoptimuman pengelompok zarah digunakan bagi mengatasi masalah. Model hibrid cadangan digunakan bagi memodelkan kadar jenayah keganasan Amerika Syarikat bersama dengan penunjuk ekonomi. Model cadangan dititip dengan ciri-ciri tambahan seperti dapat memilih siri data bagi penunjuk ekonomi dan penunjuk ekonomi yang signifikan kepada kadar jenayah kekerasan. Keputusan eksperimen menunjukkan bahawa model cadangan menghasilkan nilai peramalan yang lebih tepat berbanding model regresi linear berganda dalam meramalkan kadar jenayah kekerasan.

Kata kunci: regresi vektor sokongan, analisis hubungan kelabu, pengoptimuman pengelompok zarah, kadar jenayah keganasan, penunjuk ekonomi

1.0 INTRODUCTION

Crime rate represents social stability and is closely related to macro economy of the country. Forecast on crime rate can provide additional information for government as input in planning and developing a sustainable and healthy economy [1]. There are two types of crime rates, namely property and violent crime rates. This study focuses on the violent crime rates. The violent crime rate can be forecasted through the development of crime rate models. Econometric models are often used by researchers to study violent crime rate [2-4]. However, the development of the econometric models are quite complex and requires theoretical assumptions about the relationships between the explanatory variables. Therefore, time series model has been considered as a promising alternative tool for forecasting the violent crime rate. However, standard time series models usually require a substantial number of observations. Due to limited historical data, the application of time series models for forecasting the violent crime rate is still rare.

The application of time series models for forecasting the violent crime rate is a new field of research and rarely studied currently. Standard time series model, such as ARIMA requires sufficient historical data to produce an accurate model. Several applications of ARIMA model are only for short-term forecasting, which involve the use of weekly [5] and monthly data [6-7]. However, time series forecasting is one of the ongoing active research area. Continuous studies on time series methods, especially on soft computing techniques have improved the time series models. Soft computing technique such as support vector regression (SVR) was successfully applied in time series forecasting. Unlike statistical models, SVR is capable of modeling the nonlinear components of time series data and is robust to limited historical data. Thus, there is possibility for time series models to be used in forecasting the violent crime rates.

There are several factors that influence crimes including economic indicators, social indicators and demographic. The economic indicator is the factor that is extensively studied by researchers, where unemployment is frequently selected [8]. In previous studies, the unemployment has been found to influence violent crimes [9-12]. In addition to unemployment, other economic indicators such as consumer price index [13], gross domestic product [14] and consumer sentiment index [8] were also studied by previous researchers and were found to have influence on violent crimes. Since there are relationships between violent crimes and economic indicators, several researchers used economic indicators in modeling the violent crime rate [12-13].

Violent crime rates model can be developed using multivariate model that involves the use of one or more explanatory variables. By having other variables in the model, multivariate models have the ability to predict new patterns that have never been observed in the past. The accuracy of the multivariate model is highly dependent on explanatory variables or input features used in the model. Only significant features need to be used in the model. Hence, features selection process is very important to produce an accurate model [15]. One of the commonly use methods is statistical correlation analysis. However, most of the statistical correlation analysis requires sufficient data to determine the distribution of data and to ensure statistical significance. In addition, statistical methods also require assumption that data distribution is linear, exponential or logarithmic, and errors are normally distributed with zero means [16]. Grey relational analysis (GRA) is a mathematical method of analysis based on geometry that was proposed by Prof. Deng Julong in 1982 [17]. Unlike statistical methods, GRA can be applied in any condition without the requirements such as a particular sample size or statistical law [16, 18-19]. Thus, GRA is an alternative method that can be used to identify relationships between input features and violent crime rates.

Since GRA can only produce rank of input features according to its importance or priority, another method is required to determine the significant input features. Artificial neural network (ANN) is one of the methods that have been used together with GRA for features selection [20]. ANN is a popular and commonly used method in time series forecasting [21]. However, ANN suffers several problems such as the need for controlling numerous parameters, uncertainty in solution (network weights), and the danger of over fitting. ANN also requires sufficient data to produce a good model [22]. The drawbacks of ANN can be overcome by using SVR. SVR is a nonlinear model to solve regression problems. There are four factors that contributed to the success of SVR, which are good generalization, global optimal solution, the ability to handle nonlinear problems, and the sparseness of the solution. This has made SVR very robust with small training data, nonlinear, and high-dimensional problems [23]. However, SVR model parameters must be set correctly as inappropriate parameters may lead to over-fitting or under-fitting that can affect the regression accuracy [24]. In order to overcome this problem, researchers used particle swarm optimization (PSO) to estimate the parameters of SVR model [25-26]. PSO has the ability to escape from local optima, easy to implement and has fewer parameters to be adjusted [27].

In this study, the hybridization of grey relational analysis (GRA) and support vector regression (SVR) is proposed in modeling the violent crime rates. PSO is used to estimate the parameters of SVR model. GRA and SVR are suitable for limited historical data. GRA is used to select the most suitable data series to represent each input feature. After that, GRA will rank each input feature according to its importance or priority. Meanwhile, SVR will choose the optimal features and forecast the violent crime rates. Economic indicators are used as input features for the proposed hybrid model. The violent crime rates of the United States (US) and four economic indicators are used in this study. The economic indicators are unemployment rate (UR), gross domestic product (GDP), consumer price index (CPI) and consumer sentiment index (CS).

The remainder of this study is organized as follows. In Section 2, the proposed method, GRA_SVR, particle swarm optimization, support vector regression and grey relational analysis are described. Section 3 describes the data set and model evaluation. The experimental results are presented in Section 4. Finally, a brief conclusion is drawn in Section 5.

2.0 METHODOLOGY

This section discusses the methodology used in this study including the proposed method, GRA_SVR, particle swarm optimization, support vector regression and grey relational analysis.

2.1 The Proposed Method (GRA_SVR)

Figure 1 shows the components and the implementations of the proposed GRA_SVR. In the proposed method, GRA is used to determine the most suitable data series for each economic indicator and rank the economic indicators. Meanwhile, SVR is used to select the significant economic indicators and forecast the crime rates. There are four phases in GRA_SVR:

Phase 1: Select the best data series to represent the economic indicators
Phase 2: Rank the economic indicators
Phase 3: Select the significant economic indicators
Phase 4: Forecast the crime rates

Each phase is sequentially implemented starting with phase 1 followed by phase 2, 3 and 4, in which, output from each phase will be the input for the next phase.

Phase 1: *Select the most suitable data series to represent the economic indicators*. For each economic indicator, there are several candidate data series that can be used to represent the economic indicators in modeling the violent crime rate. The selection of the data series for economic indicators is implemented according to the crime rates. For each crime rate, a data series that is most related to the crime rate (the reference series) is selected to represent the economic indicator. For each economic indicator, GRA is used to rank all the candidate data series (compared series). The rank is based on the grey relational grade (GRG) value. The data series with the greatest GRG value is placed in the first position, the second largest is placed in the second position, and so on. The data series which is in the first position is chosen as the best data series for the economic indicator. The best data series is then used to represent the economic indicator in phase 2.

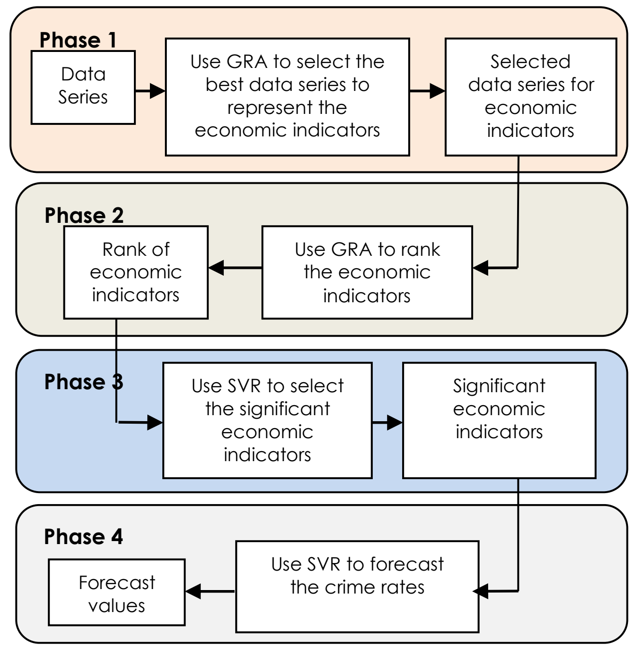


Figure 1 The proposed GRA_SVR

Phase 2: *Rank the economic indicators*. In this step, the economic indicators are arranged according to their importance to crime rate. The selected data series from phase 1 is used to represent the economic indicator. The GRA is used to examine the relationship between each economic indicator and the reference series (crime rate). The economic indicators (compared series) are arranged according to their importance to the reference series based on the GRG value. The economic indicator with the greatest GRG value is placed in the first position, the second largest is placed in the second position, and so on.

Phase 3: *Select the significant economic indicators*. This step is intended to select the significant economic indicators for crime rate using SVR model. Figure 2 shows the flowchart in phase 3. SVR is used to develop the crime rate models. A different set of economic indicators is used as input to the SVR model. Starting with all available economic indicators, the least important economic indicator is removed one at a time from the input set. The data is divided into two, training and testing data sets. SVR model is developed using the training data set. After that, the model is applied to the testing data set to forecast the crime rate. The model accuracy is calculated using four types of quantitative error measurements, namely root mean square error (RMSE), mean square error (MSE), mean absolute percentage error (MAPE), and mean absolute deviation (MAD). After all the required SVR models have been developed, the best model is determined based on error measurements, RMSE, MSE, MAPE, and MAD. If the results given by the four error measurements are inconsistent, then MAPE is chosen as a benchmark [28]. The economic indicators used as the input set for the best model is selected as the significant economic indicators.

Phase 4: *Forecast the crime rates.* The significant economic indicators are used as input features in SVR model to forecast the crime rate.

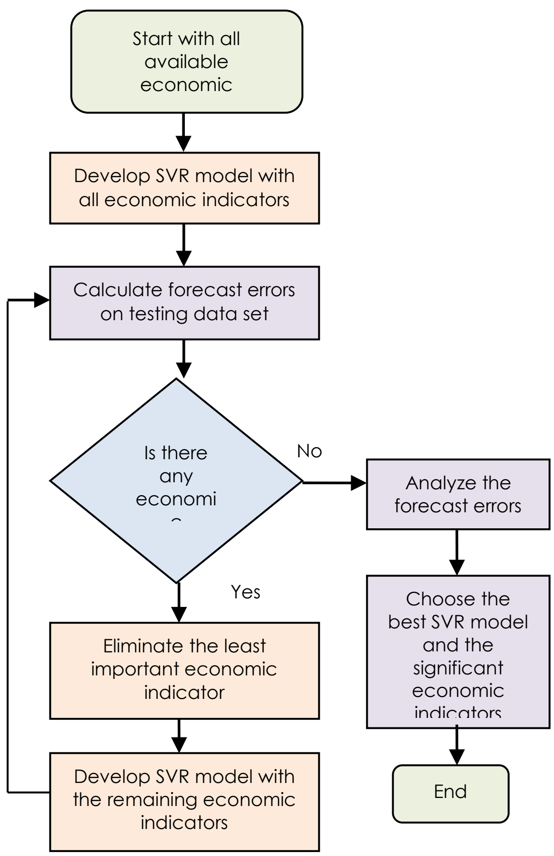


Figure 2 Flowchart to select the significant economic indicators

2.2 Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) is one of the stochastic optimization methods introduced by Kennedy and Eberhart [29]. This method is based on the natural evolution process which uses swarming strategies in bird flocking and fish schooling. PSO is population-based which consists of particles. Initially, the particles are randomly generated. Each particle has a position and velocity, which represents a potential solution to a problem in *D*-dimensional space. The position and velocity of *i*th particle are denoted by $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$ and $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$, respectively. While solving the search problem, each particle explores the search space by moving in previous direction, its previous best particle (*pbest*) and the best solution for the entire population (*gbest*). The velocity and position of each particle are updated by using formula (1) and (2) respectively [30].

$$v_{ij}(t+1) = w.v_{ij}(t) + c_1.rand1_{ij}.\left(pbest_{ij}(t) - x_{ij}(t)\right)$$
$$+c_2.rand2_{ij}.\left(gbest_j(t) - x_{ij}(t)\right) \quad (1)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (2)$$

where $v_{ij}(t)$ is the velocity of *i*th particle at iteration *t*, $x_{ij}(t)$ is the position of *i*th particle at iteration *t*, *j*=1, 2,..., *D*, is the dimension of the search space, *w* is the inertia weight to balance the global and local search abilities of particles, $rand1_{ij}$ and $rand2_{ij}$ are two uniform random numbers generated independently within the range of [0, 1], c_1 and c_2 are two learning factors which control the influence of the social and cognitive components, $pbest_{ij}(t)$ is the best previous position yielding the best fitness value for *i*th particle at iteration *t*, and $gbest_j$ is the global best particle by all particles at iteration *t*. After changing the position of the particle, the particle's fitness value is evaluated. The *pbest* and *gbest* are updated based on the current position of the particles. As this process is repeated, the whole population evolves towards the optimum solution.

The following are the steps in PSO implementation [31]:

Step 1: Initialize the positions and velocities of all the particles randomly in the *D*-dimensional search space by using the uniform probability distribution function.
Step 2: Evaluate the fitness values of the particles.
Step 3: Update *pbest* for each particle: if the current fitness value of the particle is better than its *pbest* value, set the *pbest* equal to the current position of the particle.
Step 4: Update *gbest*: if the current fitness value of the particle is better than the *gbest* value, then set *gbest* equal to the current position of the particle.
Step 5: Update the velocity and position of each particle using equation (1) and (2), respectively.
Step 6: Repeat steps 2 to 5, until a stopping criteria is met, such as a sufficient good fitness value or a maximum number of iterations.

2.3 Support Vector Regression (SVR)

Support vector regression (SVR) is a nonlinear model to solve regression problems. SVR training process is equivalent to the process of solving the linearly constrained quadratic programming problems that provide a unique optimal value where there is no local minimum problem. The solution has sparseness, as only the essential data is used to solve the regression function. Lagrangian multipliers are introduced to solve this problem. The SVR model is given by formula (3) [32],

$$f(\mathbf{x}) = (\mathbf{z}.\phi(\mathbf{x})) + b$$
(3)

(5)

where

$\mathbf{z}$ = weight vector
b = bias value
$\phi(\mathbf{x})$ = kernel function

SVR used ε-insensitivity loss function which can be expressed as formula (4),

$$L_\varepsilon(f(\mathbf{x}) - y) = \begin{cases} |f(\mathbf{x}) - y| - \varepsilon, & \text{if } |f(\mathbf{x}) - y| \geq \varepsilon \\ 0, & \text{otherwise} \end{cases}$$
(4)

(6)

where ε is the region for ε-insensitivity. Loss is accounted only if the predicted value falls out of the band area. The SVR model can be constructed to minimize the quadratic programming problem as given in formula (5),

$$\min: \tfrac{1}{2}\mathrm{z}^T\mathrm{z} + C\sum_i(\xi_i + \xi_i^*)$$
(5)

(7)

subject to

$$\begin{cases} y_i - \mathrm{z}^T\mathrm{x}_i - b \leq \varepsilon + \xi_i \\ z^T\mathrm{x}_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

where

i = 1, 2, ..., n is the number of training data
$(\xi_i + \xi_i^*)$ = the empirical risk
$\frac{1}{2}\mathbf{z}^T\mathbf{z}$ = the structure risk preventing over-learning and lack of applied universality
C = the regularization parameter

After selecting proper regularization parameter (C), width of band area (ε) and kernel function (K), the optimum of each parameter can be resolved though Lagrange function (6) [26]

$$\min: \tfrac{1}{2}\sum_{i,j}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j)$$
(6)

(8)

$$-\varepsilon\sum_i(\alpha_i + \alpha_i^*) + \sum_i y_i(\alpha_i - \alpha_i^*)$$

subject to

$$\begin{cases} 0 \leq \alpha_i \\ \alpha_i^* \leq C \\ \sum_i \alpha_i - \sum_i \alpha_i^* = 0, \ i = 1, .., n \end{cases}$$

where α_i and α_i^* are Lagrange multipliers. The kernel function (K) enables the original nonlinear input space mapped to the higher-dimensional feature space inside in which the original data are linear. Thus, the optimization problem can be solved. Finally, the SVR model can be expressed as (7) [23].

$$f(x) = \sum_i(\alpha_i - \alpha_i^*)K(x_i, x) + b \tag{7}$$

The commonly used kernels are linear kernel, polynomial kernel, radial basis function (RBF) or Gaussian kernel and sigmoid kernel. Formula (8), (9), (10), and (11) are the formula for linear kernel, polynomial kernel, RBF kernel [33] and sigmoid kernel [34], respectively.

$$K(x_i, x_j) = x_i^T x_j \tag{8}$$

$$K(x_i, x_j) = (1 + x_i.x_j)^d \tag{9}$$

$$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2) \tag{10}$$

$$K(x_i, x_j) = \tanh[v(x_i, x_j) + \alpha] \tag{11}$$

The type of kernel function influences the parameters of SVR kernel. The kernel function and parameters of SVR kernel function should be set properly because it can affect the regression accuracy. Inappropriate parameters may lead to over-fitting or under-fitting [24]. This study uses the RBF kernel function because it is suitable for solving most forecasting problems [25]. The RBF kernel is also effective and has a fast training process [35]. For the RBF kernel function, there are three important parameters to be determined, which are regularization parameter (C), kernel parameter (γ), and tube size of ε-insensitive loss function (ε) [26].

These parameters must be set correctly, in order to produce accurate estimation model. Experience and prior knowledge can be used to set these parameters, however, this is not an efficient way [36]. Meanwhile, traditional method, namely grid search requires a long searching time to get the optimal parameters. Therefore, optimization algorithms are applied by several researchers in setting the SVR parameters. Among the optimization algorithms are PSO [26-27, 37-39], simulated annealing [40], GA [35], ant colony optimization [41], differential evolution [36] and ant colony-immune clone PSO [42].

In this study, PSO is used to estimate the SVR parameters as shown in Figure 3. The position and velocity of ith particle are represented by the three-dimensional vectors, $X_i = (x_{i1}, x_{i2}, x_{i3})$ and $V_i = (v_{i1}, v_{i2}, v_{i3})$, where the first, second and third dimensions of the vectors are referred to C, γ and ε, respectively. PSO is used in conjunction with k-fold cross-validation to reduce the adverse effects of over fitting phenomenon. In k-fold cross-validation, the training data set is randomly divided into k subsets of equal size. The regression function is built with a given set of parameters (C, γ, ε) using the k-1 subsets and the model performance is measured by the one remaining subset (validation data set). Each subset is used once for validation and the process is repeated k times. In this study, root mean squared error (RMSE) is selected to be the performance criterion, and the 5-fold cross-validation is utilized to evaluate the performance of the model [43-44]. The average of RMSE on the validation set from k trials is used as a measure of fitness. The RMSE is defined as equation (12).

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - p_t)^2} \quad (12)$$

where n is the number of validation data; y_t is the actual value and p_t is the forecast value.

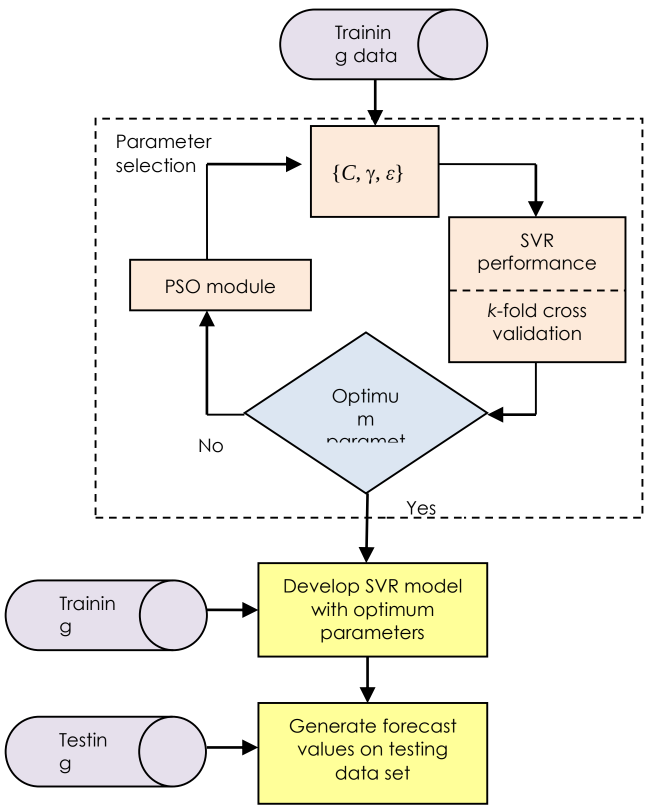


Figure 3 PSO in estimating the SVR parameters

2.4 Grey Relational Analysis (GRA)

Grey relational analysis (GRA) is a method of analysis proposed in the Grey system theory and founded by Professor Deng Julong [17, 45]. This method is based on geometrical mathematics, which is in compliance with the principles of normality, symmetry, entirety and proximity [46-47]. GRA is a distinct similarity measurement that uses data series to obtain grey relational order to describe the relationship between the related series [48]. Thus, GRA can be used to measure the correlations between the reference series and other compared series.

The relative distance between a compared series on the reference series, which is referred to as grey relational grade (GRG) represents the degree of influence between these series. A small distance indicates a significant influence [16]. GRG is a numeric value ranging between 0 and 1. GRG value close to 1 indicates that there is a strong relationship between the two series. The following are the basic steps in the GRA [49]:

i. Determine compared series and reference series

$$X_0 = \{x_0(1), x_0(2), \ldots, x_0(n)\} \quad (13)$$

$$X_i = \{x_i(1), x_i(2), \ldots, x_i(n)\}\,,\ i = 1,2,\ldots,m \quad (14)$$

ii. Use formula (15) to transform the series to dimensionless form.

$$x_i^*(k) = \frac{x_i(k) - \min\limits_k x_i(k)}{\max\limits_k x_i(k) - \min\limits_k x_i(k)},\quad i = 1,..,m, k = 1,..,n \quad (15)$$

iii. Use formula (16) to calculate the grey relational coefficient (ξ_{0i}) between X_0 and X_i, which reflects the degree of influence of two compared series at one time.

$$\xi_{0i} = \frac{\min\limits_i \min\limits_k |x_0^*(k) - x_i^*(k)| + \rho \max\limits_i \max\limits_k |x_0^*(k) - x_i^*(k)|}{|x_0^*(k) - x_i^*(k)| + \rho \max\limits_i \max\limits_k |x_0^*(k) - x_i^*(k)|} \quad (16)$$

where $\rho \in (0,1)$, is the distinguishing coefficient. Generally, $\rho = 0.5$. According to the mathematic proof, the value of ρ does not affect the rank of the grey relational grade [50].

iv. Calculate the grey relational grade (GRG), which is the average value of grey relational coefficient using formula (17).

$$r_{0i} = \frac{1}{n}\sum_{k=1}^{n} \xi_{0i}(k) \quad (17)$$

where r_{0i} represents the level of correlation between the reference series and the compared series.

The grey relational order is constructed based on the calculated value of GRG, r_{0i}. The order obtained is a list of priorities in selecting a series closely related to the reference series, X_0. Generally, $r > 0.9$ indicates a marked influence, $r > 0.8$ a relatively marked influence, $r > 0.7$ a noticeable influence, and $r < 0.6$ a negligible influence [46].

3.0 DATA SET AND MODEL EVALUATION

This section describes the data set used and the model evaluation carried out in this study.

3.1 Data Set

This study includes four violent crime rates, which are robbery, aggravated assault, forcible rape and murder and non-negligent manslaughter rates. There are four economic indicators studied, namely consumer price index, gross domestic product, consumer sentiment index and unemployment rate. The data used are annual data from 1961 until 2009. The violent crime rates data are obtained from the Uniform Crime Reporting Statistics web site (http://www.ucrdatatool.gov), while economic indicators data are available on Economic Research Federal Reserve Bank of St. Louis web site (http://research.stlouisfed.org). There are 30 data series for the unemployment rate (UR), 15 data series for the gross domestic product (GDP), 16 data series for the consumer price index (CPI) and two data series for the consumer sentiment index (CS). The data is then divided into training and testing data sets. The training data set is used to develop the SVR model while the testing data set is used to evaluate the forecasting performance of the SVR model. In this study, 90 percent of the data is used for training (1961 to 2004) and 10 percent is as testing data set (2005 to 2009).

3.2 Model Evaluation

The proposed method, GRA_SVR is compared with multiple linear regression (MLR) to evaluate its performance. MLR is a popular linear model in multivariate analysis. The performances of both models are evaluated based on the testing data set using quantitative error measurements and statistical test. Quantitative error measurements are conducted using four statistical error test, namely root mean square error (RMSE), mean square error (MSE), mean absolute percentage error (MAPE), and mean absolute deviation (MAD). In statistical test, paired sample *t*-test is performed to prove that there is no significant means difference between the forecast values and the actual data.

4.0 RESULT AND DISCUSSION

Table 1 summarized the selected data series and significant economic indicators for each crime rate by GRA_SVR. This study proposes the selection of data series based on the type of crime through the use of GRA. The selection of significant economic indicators for each crime rate was done by using GRA_SVR. The result from the selection of significant economic indicators discovered that robbery and murder and non-negligent manslaughter are influenced by UR. Meanwhile, CPI has been found to influence aggravated assault. However, forcible rape is a violent crime that is not influenced by any of the economic indicators studied.

The UR can produce both crime motivation and crime opportunity that would affect the crime rates [51]. The unemployed are more likely to commit crimes because they do not have income to support themselves. They also tend to reduce spending on goods and spend more time at home. Subsequently, this will reduce opportunities for property crimes. Although robbery is a violent crime, it is more motivated by economic gain. This is the possible reason UR is significant to robbery. Meanwhile, unemployment and loss of income probably causes depression and desperateness for money and may lead to the occurrence of violent crimes such as murder. The annual percentage change in a consumer price index is used as a measure of inflation. Rising inflation will create opportunities for criminals, while a low rate circulation of money and property will reduce the opportunities for property crime. An increase in price of goods can reduce the purchasing power [12]. This may cause stress and lead to violent crime such as aggravated assault.

Table 1 The results of economic indicators selection by GRA_SVR

Violent crime	Significant economic indicator	Selected data series
Robbery	UR	UR - 20 to 24 years
Aggravated assault	CPI	CPI for all urban consumers: Apparel
Forcible rape	No significant economic indicator	
Murder and non-negligent manslaughter	UR	Natural rate of unemployment

Table 2 shows the difference across significant economic indicators selected by the MLR and GRA_SVR models. The economic indicators selected by the MLR and the GRA_SVR are different because both models differs in term of relationships existing among the data. For example, MLR has linear relationships while GRA_SVR has nonlinear relationships. In MLR models, the CS has been found significant to robbery and murder and non-negligent manslaughter. Meanwhile, GDP has been found significant to robbery and forcible rape. In addition, UR was also found significant to murder and non-negligent manslaughter. The results indicate that the relationships between CS and GDP with robbery are linear. In parallel, GDP was found to have linear relationships with forcible rape. While, UR has both linear and nonlinear relationships with murder and non-negligent manslaughter. There is no significant linear relationship between CPI and violent crime rates. Thus, CPI only has a nonlinear relationship with aggravated assault.

Table 2 Selected significant economic indicators by GRA_SVR and MLR

Violent crime	GRA_SVR	MLR
Robbery	UR	GDP CS
Aggravated assault	CPI	No economic indicator
Forcible rape	No economic indicator	GDP
Murder and non-negligent manslaughter	UR	UR CS

Table 3 shows the results from paired sample *t*-test. Paired sample *t*-test is used to compare the actual data of crime rates with the forecast values of the proposed model, GRA_SVR and MLR. The test is conducted to ensure that there is no statistically significant difference of means between the actual data and the forecast values from the models. The results show that between the actual data and the forecast values from the GRA_SVR models, the *P-value* is greater than 0.05 and the difference between the upper and lower values for 95% interval is ranging between negatives and positives values. This result indicates that the hypothesis test fails to reject the null hypothesis which implies that there is no statistically significant difference of means between the actual data of violent crime rates and the forecast values of GRA_SVR models. Meanwhile, for MLR models, the hypothesis test fails to reject the null hypothesis, except for robbery rate model. The hypothesis test rejects the null hypothesis for robbery rate model which implies that there is statistically significant difference of means between the actual data of robbery rate and the forecast values of MLR model. GRA_SVR models show smaller absolute value of mean, standard deviation and standard error mean as compared to MLR models. In addition, forecasting errors by GRA_SVR models are found to be smaller than MLR models as shown in Table 4 and Figure 4. Therefore, GRA_SVR has shown better forecasting performance as compared to the MLR.

Table 3 Paired sample test

Pair	Paired differences					Sig.(2-tailed)
	Mean	Std. deviation	Std. error mean	95% confidence interval of the difference		
				Lower	Upper	
Robbery and GRA_SVR	-1.00730	2.88881	1.29191	-4.59422	2.57963	0.479
Robbery and MLR	6.13707	3.97332	1.77692	1.20354	11.07060	0.026
Aggravated assault and GRA_SVR	-0.19296	1.57754	0.70550	-2.15174	1.76582	0.798
Aggravated assault and MLR	-5.65413	5.86676	2.62369	-12.93867	1.63042	0.097
Forcible rape and GRA_SVR	0.01651	0.27268	0.12195	-0.32207	0.35509	0.899
Forcible rape and MLR	0.07321	0.83840	0.37494	-0.96780	1.11422	0.855
Murder and non-negligent manslaughter and GRA_SVR	-0.01323	0.12356	0.05526	-0.16666	0.14019	0.822
Murder and non-negligent manslaughter and MLR	-0.26099	0.32360	0.14472	-0.66279	0.14082	0.146

Table 4 Statistical error test

Violent crime	Model	RMSE	MSE	MAPE	MAD
Robbery	MLR	7.09	50.29	4.31	6.14
	GRA_SVR	2.77	7.69	1.77	2.56
Aggravated assault	MLR	7.71	59.50	2.57	7.10
	GRA_SVR	1.42	2.03	0.39	1.11
Forcible rape	MLR	0.75	0.57	2.38	0.72
	GRA_SVR	0.24	0.06	0.65	0.19
Murder and non-negligent manslaughter	MLR	0.97	0.95	2.99	0.90
	GRA_SVR	0.24	0.06	0.65	0.19

Based on the model evaluation that has been carried out, we can conclude that the proposed method, GRA_SVR is more appropriate to be employed in modeling the violent crime rates with economic indicators as compared to MLR. MLR only identifies features based on a linear relationship, but the GRA_SVR is able to identify important features based on nonlinear relationships in the data. Since most of the relationships that exist between crime rates and economic indicators are nonlinear, then, the GRA_SVR is able to recognize important features better than MLR.

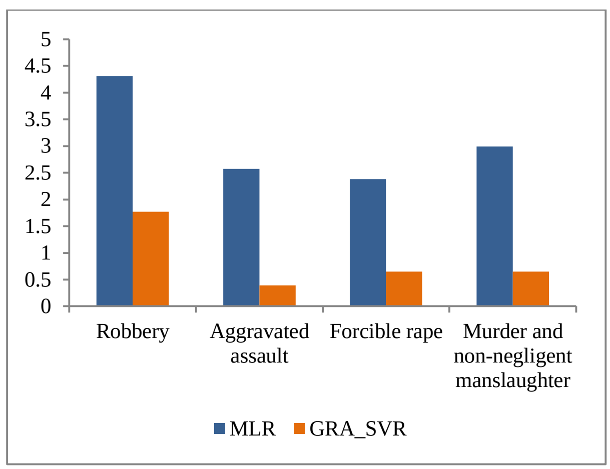


Figure 4 Forecasting performances based on MAPE

5.0 CONCLUSION

The proposed method, GRA_SVR is suitable for time series data without any assumptions about the data, including data with small sample size. Thus, it can be applied to any time series data in criminology and other fields or domains. GRA_SVR is able to recognize nonlinear relationships between the economic indicators and the crime rate, rank the economic indicators based on its importance to the crime rates and select the significant economic indicators. In cases where there is more than one series of data for each economic indicator, GRA_SVR provides the selection of data series for the economic indicator based on the suitability of each data series with the crime rates. The significant economic indicators can simplify the model in order to produce more accurate forecasting results. Thus, this method can assist decision makers in choosing the input factors without the need to refer to the domain experts. In criminology, the application of machine learning in crime forecasting is a new field of study. The results of this study indicate that the proposed model, GRA_SVR is very suitable to be implemented in modeling the violent crime rates. GRA_SVR outperformed MLR in forecasting the violent crime rates. Based on the good performance of GRA_SVR in forecasting the violent crime rates, it can be suggested that GRA_SVR can be applied in the field of criminology as a tool for crime rates forecasting.

Acknowledgement

This work is supported by Ministry of Higher Education under Research University Grant (Q.J130000.2628.10J03) and Fundamental Research Grant Scheme (RJ130000.7828.4F179). Authors would like to thank Research Management Centre (RMC) Universiti Teknologi Malaysia, for the research activities, *Soft Computing Research Group* (SCRG) and *UTM Big Data Centre* for the support and motivation in making this study a success.

References

[1] Yufei, Z. and J. Huifeng. 2010. Using GIS to Analyze and Forecast The Chinese Crime Rate. *Information Science and Engineering (*ICISE*)*. Hangzhou, China. 4-6 December 2010. 3352-3354.

[2] Tang, C. F. and H. H. Lean. 2007. Will Inflation Increase Crime Rate? New Evidence from Bounds and Modified Wald Tests. *Global Crime*. 8(4): 311-323.

[3] Tang, C. F. 2011. An Exploration of Dynamic Relationship Between Tourist Arrivals, Inflation, Unemployment and Crime Rates in Malaysia. *International Journal of Social Economics*. 38(1): 50-69.

[4] Wu, D. and Z. Wu. 2012. Crime, Inequality and Unemployment in England and Wales. *Applied Economics.* 44(29): 3765-3775.

[5] Chen, P., H. Yuan and X. Shu. 2008. Forecasting Crime Using the ARIMA Model. *Fuzzy Systems and Knowledge Discovery (FSKD)*. Jinan, Shandong, China. 18-20 October 2008. 627-630.

[6] Shrivastav, A. K. and Ekata. 2012. Applicability of Box Jenkins ARIMA Model in Crime Forecasting: A Case Study of Counterfeiting in Gujarat State. *International Journal of Advanced Research in Computer Engineering & Technology*. 1(4): 494-497.

[7] Noor, N. M. M., A. Retnowardhani, M. L. Abd and M.Y.M. Saman. 2013. Crime Forecasting Using ARIMA Model and Fuzzy Alpha-Cut. *Journal of Applied Sciences*. 13(1): 167-172.

[8] Rosenfeld, R and R. Fornango. 2007. The Impact of Economic Conditions on Robbery and Property Crime: The Role of Consumer Sentiment. *Criminology.* 45(4): 735-769.

[9] Andresen, M. A. 2013. Unemployment, Business Cycles, Crime, and the Canadian Provinces. *Journal of Criminal Justice.* 41(4): 220-227.

[10] dos Santos, M. J. and A. L. Kassouf. 2013. A Cointegration Analysis of Crime, Economic Activity, and Police Performance in São Paulo City. *Journal of Applied Statistics*. 40(10): 2087-2109.

[11] Kapuscinski, C. A., J. Braithwaite and B. Chapman. 1998. Unemployment and Crime: Toward Resolving the Paradox. *Journal of Quantitative Criminology.* 14(3): 215-243.

[12] Rattner, A. 1990. Social Indicators and Crime Rate Forecasting. *Social Indicators Research*. 22(1): 83-95.

[13] Yearwood, D. L., and G. Koinis. 2011. Revisiting Property Crime and Economic Conditions: An Exploratory Study to Identify Predictive Indicators Beyond Unemployment Rates. *Social Science Journal.* 48(1): 145-158.

[14] Fajnzylber, P., D. Lederman and N. Loayza. 2002. What Causes Violent Crime? *European Economic Review*. 46(7): 1323-1357.

[15] Han, M., and R. Wei. 2008. Variable Selection for Multivariate Time Series Prediction with Neural Networks. *Neural Information Processing (*ICONIP*)*. Kitakyushu, Japan. 13-16 November 2007. 415-425.

[16] Song, Q. and M. Shepperd. 2011. Predicting Software Project Effort: A Grey Relational Analysis Based Method. *Expert Systems with Applications.* 38: 7302-7316.

[17] Deng, J. L. 1982. Control problems of Grey Systems. *Systems and Control Letters*. 5: 288-294.

[18] Çaydaş, U., and A. Hasçalik. 2008. Use of the Grey Relational Analysis to Determine Optimum Laser Cutting Parameters With Multi-Performance Characteristics. *Optics and Laser Technology.* 40(7): 987-994.

[19] Dong, F., Q. Tan and X. Li. 2009. The Relationship Between Chinese Energy Consumption and GDP: An Econometric Analysis Based on the Grey Relational Analysis (GRA). *Grey Systems and Intelligent Services (GSIS)*. Nanjing, China. 10-12 November 2009. 153-157.

[20] Sallehuddin, R. 2010. *Hybridization of Nonlinear and Linear Model for Time Series Forecasting*. Doctor Philosophy, Universiti Teknologi Malaysia.

[21] Zhang, G. P. 2003. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing.* 50: 159-175.

[22] Zheng, L., H. Zhou, C. Wang and K. Cen. 2008. Combining Support Vector Regression and Ant Colony Optimization to Reduce Nox Emissions in Coal-Fired Utility Boilers. *Energy and Fuels.* 22(2): 1034-1040.

[23] Ding, Z. 2012. Application of Support Vector Machine Regression in Stock Price Forecasting. *Business, Economics, and Financial Sciences, Management (BEFM)*. Jeju Island, South Korea. 30-31 December 2011. 359-365.

[24] Wu, J. and L. Jin. 2011. Daily Rainfall Prediction with SVR using a Novel Hybrid PSO-SA Algorithms. *High-Performance Networking, Computing and Communications Systems (ICHCC)*. Singapore, Singapore. 5-6 May 2011. 508-515.

[25] Wu, J. and E. Chen. 2010. A Novel Hybrid Particle Swarm Optimization for Feature Selection and Kernel Optimization in Support Vector Regression. *Computational Intelligence and Security* (CIS). Nanning, China. 11-14 December 2010. 189-194.

[26] Zhao, S. and L. Wang. 2010. Support Vector Regression Based on Particle Swarm Optimization for Rainfall Forecasting. *Computational Sciences and Optimization, (CSO), Theoretical Development and Engineering Practice*. Huangshan, Anhui, China. 28-31 May 2010. 484-487.

[27] Liao, R., H. Zheng, S. Grzybowski and L. Yang. 2011. Particle Swarm Optimization-Least Squares Support Vector Regression Based Forecasting Model On Dissolved Gases In Oil-Filled Power Transformers. *Electric Power Systems Research*. 81(12): 2074-2080.

[28] Makridakis, S. 1993. Accuracy Measures: Theoretical and Practical Concerns. *International Journal of Forecasting*. 9(4): 527-529.

[29] Kennedy, J. and R. Eberhart. 1995. Particle Swarm Optimization. *Neural Networks*. Perth, Australia.1942-1948.

[30] Wang, H., H. Sun, C. Li, S. Rahnamayan and J. S. Pan. 2013. Diversity Enhanced Particle Swarm Optimization With Neighborhood Search. *Information Sciences.* 223: 119-135.

[31] He, G. and N. J. Huang. 2012. A Modified Particle Swarm Optimization Algorithm with Applications. *Applied Mathematics and Computation.* 219(3): 1053-1060.

[32] Lu, C. J., C. H. Chang, C. Y. Chen, C. C. Chiu and T. S. Lee. 2009. Stock Index Prediction: A Comparison of MARS, BPN and SVR in an Emerging Market. *Industrial Engineering and Engineering Management (IEEM)*. Hong Kong, China. 8-11 December 2009. 2343-2347.

[33] Lins, I. D., M.D.C. Moura, E. Zio, and E.L Droguett. 2012. A Particle Swarm-optimized Support Vector Machine for Reliability Prediction. *Quality and Reliability Engineering International*. 28(2): 141-158.

[34] Hu, Y., C. Wu and H. Liu. 2011. Prediction of Passenger Flow on the Highway Based on the Least Square Support Vector Machine. *Transport*. 26(2): 197-203.

[35] Xin, N., X. Gu, H. Wu, Y. Hu and Z. Yang. 2012. Application of Genetic Algorithm-support Vector Regression (GA-SVR) for Quantitative Analysis of Herbal Medicines. *Journal of Chemometrics.* 26(7): 353-360.

[36] Wang, J., L. Li, D. Niu and Z. Tan. 2012. An Annual Load Forecasting Model Based on Support Vector Regression With Differential Evolution Algorithm. *Applied Energy.* 94: 65-70.

[37] Cao, H. and M. Ahmed. 2010. Application of Support Vector Regression Trained by Particle Swarm Optimization in Warrant Price Prediction. *Industrial Mechatronics and Automation (ICIMA)*. Wuhan, China. 30-31 May 2010. 358-361.

[38] Min, Z. and T. Huanqi. 2011. Short Term Load Forecasting with Least Square Support Vector Regression and PSO. *Applied Informatics and Communication (ICAIC)*. Xi'an, China. 20-21 August 2011. 124-132.

[39] Lin, Z. and C. Tian. 2013. Soft Measurement Technology and Implementation Based on PSO-SVR. *Applied Mechanics and Materials*. 392: 774-778.

[40] Lin, S. W., Z. J. Lee, S. C. Chen and T. Y. Tseng. 2008. Parameter Determination of Support Vector Machine and Feature Selection Using Simulated Annealing Approach. *Applied Soft Computing Journal*. 8(4): 1505-1512.

[41] Zhou, H., J.P. Zhao, L.G. Zheng, C.L. Wang and K.F. Cen. 2012. Modeling NO x Emissions from Coal-Fired Utility Boilers Using Support Vector Regression With Ant Colony Optimization. *Engineering Applications of Artificial Intelligence*. 25(1): 147-158.

[42] Jiang, H., Z. Yan and X. Liu. 2013. Melt Index Prediction Using Optimized Least Squares Support Vector Machines Based On Hybrid Particle Swarm Optimization Algorithm. *Neurocomputing*. 119: 469-477.

[43] Wu, J., M. Liu and L. Jin. 2010. A Hybrid Support Vector Regression Approach for Rainfall Forecasting Using Particle Swarm Optimization and Projection Pursuit Technology. *International Journal of Computational Intelligence and Applications.* 9(2): 87-104.

[44] Safarzadegan Gilan, S., H. Bahrami Jovein and A. A. Ramezanianpour. 2012. Hybrid Support Vector Regression - Particle Swarm Optimization for Prediction of Compressive Strength and RCPT of Concretes Containing Metakaolin. *Construction and Building Materials.* 34: 321-329.

[45] Deng, J. L. 1989. Introduction to Grey System Theory. *The Journal of Grey System*. 1: 1-12.

[46] Fu, C., J. Zheng, J. Zhao and W. Xu. 2001. Application of Grey Relational Analysis for Corrosion Failure of Oil Tubes. *Corrosion Science*. 43(5): 881-889.

[47] Sallehuddin, R., S. M. Shamsuddin and S. Z. M Hashim. 2008. Hybridization Model of Linear and Nonlinear Time Series Data for Forecasting. *Modelling and Simulation (AMS)*. Kuala Lumpur, Malaysia. 13-15 May 2008. 597-602.

[48] Xuerui, T. and L. Yuguang, L. 2004. Using Grey Relational Analysis to Analyze the Medical Data. *Kybernetes.* 33(2): 355-362.

[49] Huang, Y., H. Wang, G. Xing and D. Sun. 2010. A Hybrid Grey Relational Analysis and Support Vector Machines Approach for Forecasting Consumption of Spare Parts. *Artificial Intelligence and Education* (ICAIE). Hangzhou, China. 29-30 October 2010. 602-605.

[50] Kung, C. Y., T.M. Yan, S.C. Chuang and J.R. Wang. 2006. Applying Grey Relational Analysis to Assess the Relationship Among Service Quality Customer Satisfaction and Customer Loyalty. *Cybernetics and Intelligent Systems*. Bangkok, Thailand. 7-9 June 2006.

[51] Cantor, D. and K. C. Land. 2001. Unemployment and Crime Rate Fluctuations: A Comment on Greenberg. *Journal of Quantitative Criminology.* 17(4): 329-342.