SPATIAL PREDICTION OF CAR CRASHES SEVERITY BASED ON DYNAMICALLY UPDATED AUTOREGRESSIVE NEURAL MODELS

Loukas Dimitriou, King Saud University, Prince Mohamed Bin Naif Chair for Traffic Safety Research, P.O. Box 800, Riyadh 11421, Saudi Arabia

Hany M. Hassan, King Saud University, Prince Mohamed Bin Naif Chair for Traffic Safety Research, P.O. Box 800, Riyadh 11421, Saudi Arabia

ABSTRACT

Road safety for many decades lies at the centre of the transportation research interest. One of the important aspects of the relevant research concerns the prediction of future crash propensity and in particular the location and severity of accidents in urban and inter-urban road networks. Such information is of significant importance since could support decisions on planning and scheduling of road safety resources. While the majority of the existing models are aiming either on correlating traffic variables with crash occurrence or modelling crash frequencies both in highways as well in urban network links and intersections, in the current paper, results from an exploratory analysis are presenting, based on spatially distributed time series prediction modelling. The current analysis is based on auto-regressive time-series models and in particular in models of Artificial Neural Networks (ANNs). Such models are belonging to the Artificial Intelligent class of modelling approaches, which have not widely presented in the literature although can be regarded as useful in cases where detailed and reliable information of traffic characteristics is not available. Alternative structures of auto-regressive ANNs have been calibrated and used on a suitable -for such analysis- detailed and realistic database, composed of daily records covering almost 6 years from Riyadh, capital of the Kingdom of Saudi Arabia. The results are providing evidence on the performance of such approaches in predicting spatially distributed time series car crashes.

Keywords: *Car Accidents Prediction, Spatially Distributed Time Series, Dynamic Autoregressive Artificial Neural Networks*

1. INTRODUCTION

In developing countries, one of the common indicators stands for the increment of the car ownership index and the car usage, due to the amelioration of the private economics and the lag of public transportation infrastructure expansions. Additionally to the transport infrastructure development delays, problems with driver training have not been avoided. In brief, these phenomenon are closely related to road safety issues as have been repeatedly reported, leading road safety to be highlighted as among the most significant public safety issues for the years to come (UN 2010).

For addressing the complex issue of road safety, the available means of contemporary technology are expected to contribute significantly and in particular those under the general framework of Intelligent Transport Systems (ITSs) and the Smart Infrastructure. Such 'tools' and methods that have been developed tested and proved in the last decades, providing encouraging results. As so, the current paper aims on providing some evidence on the ability of "computational intelligent" methods previously tested for other road-traffic paradigms to support decisions related to road safety and in particular here some computational experience will be offered on the important matter of crash predictions.

Among the large 'palette' of the advanced prediction methods and frameworks available for decades now, here some applications of crash predictions based on spatially distributed time-series will be presented. Time series can be regarded as a class of models that are evidential-based rather than the explanatory-based modeling approaches typically used for crash risk prediction purposes, relating operational conditions with risk. Despite that fact, time series analysis, if suitably conducted, could be regarded that utilizes all important information of parameters affecting crash risk. For example, urban road networks are typically servicing commuting trips (both in weekdays as well as in weekends) and as so the potential of such models applications can be regarded as a worthwhile prediction practice. Here, an exploratory analysis of dynamic prediction mechanisms based on autoregressive models is conducted, since auto-regressive models possesses the characteristic that can 'transmit' augmented past information (as such information is depicted in time-series) into the future. The results here come from the analysis of Auto-Regressive Dynamic Neural Networks, cardinally due to the fact that these models can handle cases of non-linear and non-stationary dynamic processes. The application of the proposed prediction mechanism is tested over a dataset of weekly crash records from the Riyadh, capital of the Kingdom of Saudi Arabia, an area experiencing severe road safety issues for years now accompanying the rapid development of the urban space.

The paper structure starts with a parsimonious review of previous studies and models used for predicting car crashes as long as predictors of other road traffic-related characteristics. Then, the characteristics of the available dataset are presented, exposing the particularities involved in crash data. The selection of the set of dynamic auto-regressive ANNs models that could be regarded as possible relevant candidates for time-series analysis of such database is presented next, followed by the discussion of these models performance. The final section concludes and gives some points of possible further developments in the time series predictions of car crashes.

2. CRASH DATA AND PRELIMINARY ANALYSIS

The prediction of traffic conditions is one the important elements in the deployment of ITSs and especially in management and control of road networks. Related to the prediction of car crashes the reader may refer to Abdel-Aty and Abdelwahab (2004) and Abdelwahab and Abdel-Aty (2004 *a&b*), where in these studies the value of historical information has been highlighted. ITSs technologies have also been used for predicting car crashes (Abdel-Aty et al. 2004, Abdel-Aty and Pande, 2005, Abdel-Aty et al. 2008 and Christoforou et al., 2011) introducing and augmenting advanced information technologies for predicting crash type/severity and location. Focusing in studies in the Kingdom of Saudi Arabia, Al-Gamdi (2002 *a&b*) and Al-Gamdi and Al-Gadhi (2004) utilized advanced computational methods for purposes of predicting crash risk in alternative circumstances.
Despite the fact that in time-series models have not yet used in crash prediction, a large body of literature exists on the time-series traffic prediction mechanisms (for a brief review see

Stathopoulos et al. 2008, Dimitriou et al. 2008). In this line of research, a major distinction exists between the statistical models and those belonging to the Artificial Intelligent class of models (for a comprehensive review see Karlaftis and Vlahogianni, 2010). As mentioned in the introductory section, for addressing issues of non-stationarity and non-linearity here the analytical framework will be based on hybrid Dynamic Artificial Neural Networks (DANNs). Before getting into details on the structure of the DANNs used here, the first step will be the statistical examination of the crash time series.

The database of the crash reports used for the current analysis correspond to location-specific detailed records of all serious crashes (crashes with at least one fatality or injury but excluding crashes with property damages only) occurred within the period of March 2004-December 2011. A depiction of the location and frequency/density of crashes is provided in Figure 1, providing a typical picture of the distribution of crashes in urban networks. It can be observed in Figure 1 that crashes are mainly concentrated in the city centre and major highways and arterials and more sparsely distributed in the city surroundings.

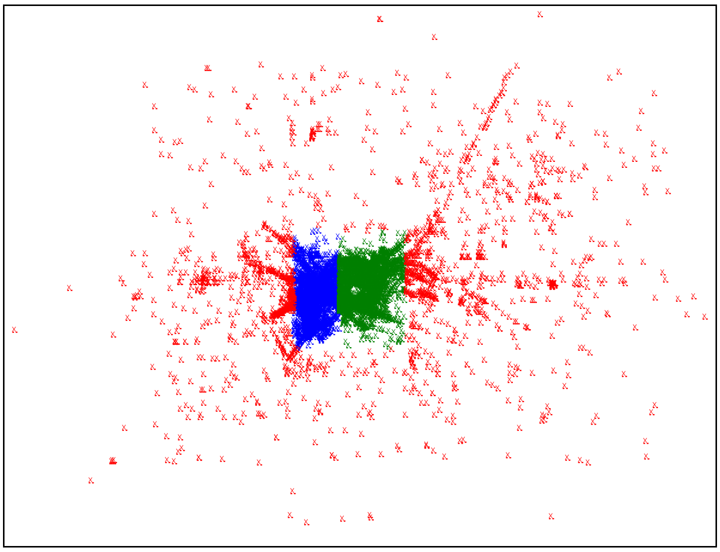


Figure 1: Colour coding of the Riyadh's city area division in three distinctive sectors: two for the city's centre (blue and green dots) and one for the surroundings (red dots).

For the purposes of the current study, the Riyadh city is divided into three distinctive regions: two dividing the city's centre and one covering the surrondings. This division can be broadly correspond to the city's division in patrolling sectors and thus predicted information of future crashes is extremely valuable for allocating traffic enforcement units, abulances etc. This division (and the relevant colour coding) holds for the rest of the paper. It is noted that the simultaneous prediction of the number of crashes per distinctive area leads to a multivariate (spatialy distributed) prediction case which can be regarded as challenging from an analytic percpective.

The dataset used for the curent analysis is composed of daily location-specific records of each individual crash, augmented in weekly subsets. As expected, the evolution of the all variable of car crashes are in this case 'smoother', both in the total city's nombers as well as distributed in regions. Those elements can be identified in Figure 2*a,* where total city's weekly number of accidents (variable name: 'WeekNo#') and in Figure 2*b* where total cars involved in crashes per region (variables name: 'WeekCar#') for the three distincive regions are depicted.

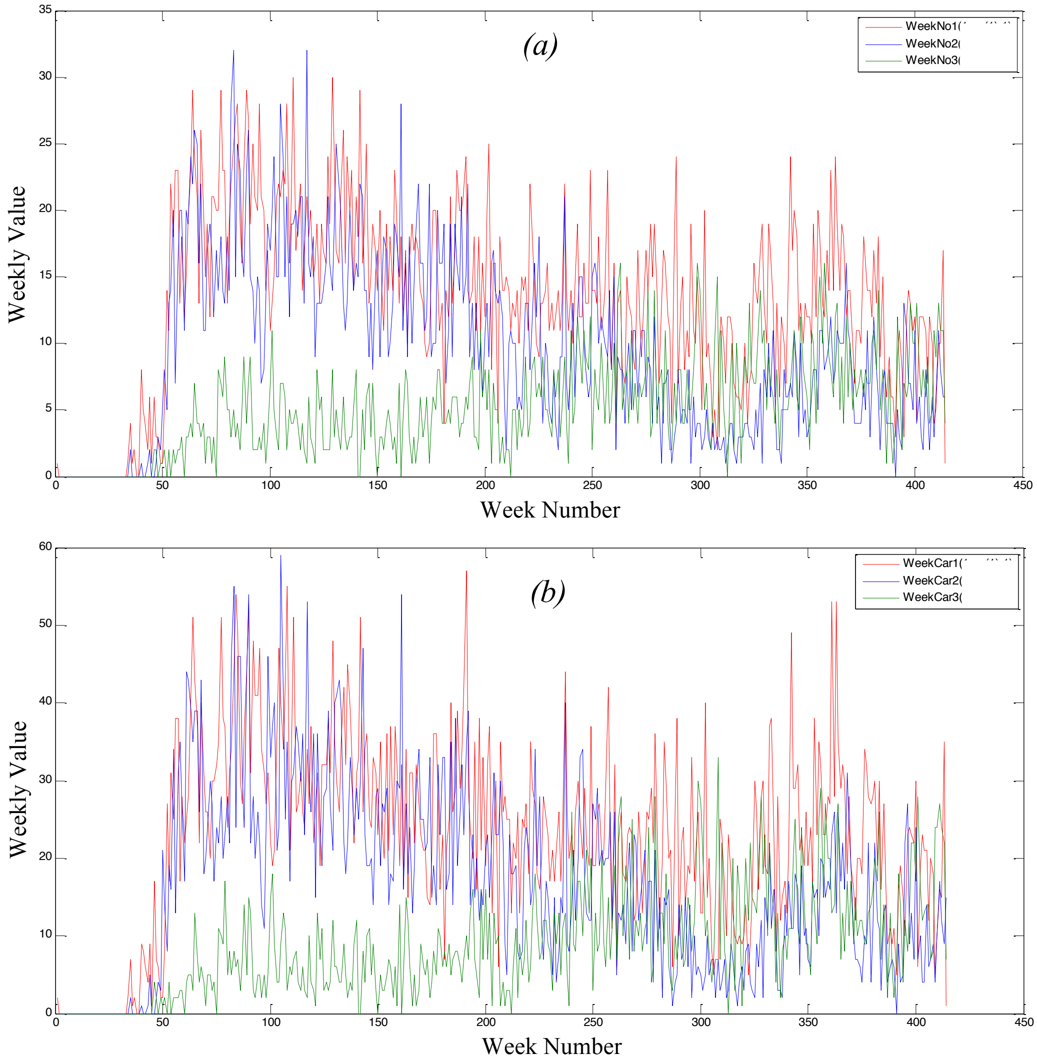


Figure 2: Weekly time series of the number of accidents and cars involved (Fig. 7a) and weekly spatially distributed (Fig. 2b&c).

On the overall, it can be observed that car crashes tend to concentrate to the city's centre where the most of heavy traffic is observed, while in the latest months the distribution of all crash variables tends to 'spread' evenly along the Riyadh area (more-or-less the variables' values for all regions are of the same magnitude).

The next ster is to identify possible correlations among the datasets composed by the alternative time-series. An initial picture can be drawn by the the matrix of correlation diagrams. In Figure 3 the correlation diagrams matrix is presented, exposing some correlations among variables, an element of value for multivariate analysis and in particular for identifying valuable relationships among datasets. The matrix of correlation diagrams are followed by a standard variance-covariance and correlation analysis. In Table 1*a* the variance-covariance matrix among variables are further investigated from which can be identified the –possibly expeted- covariance between the number of crashes and number of cars involved weekly per region, but also some covariances of these variables between regions. This is further justified for specific pairs of variables in the correlation table (Table 1*b*).

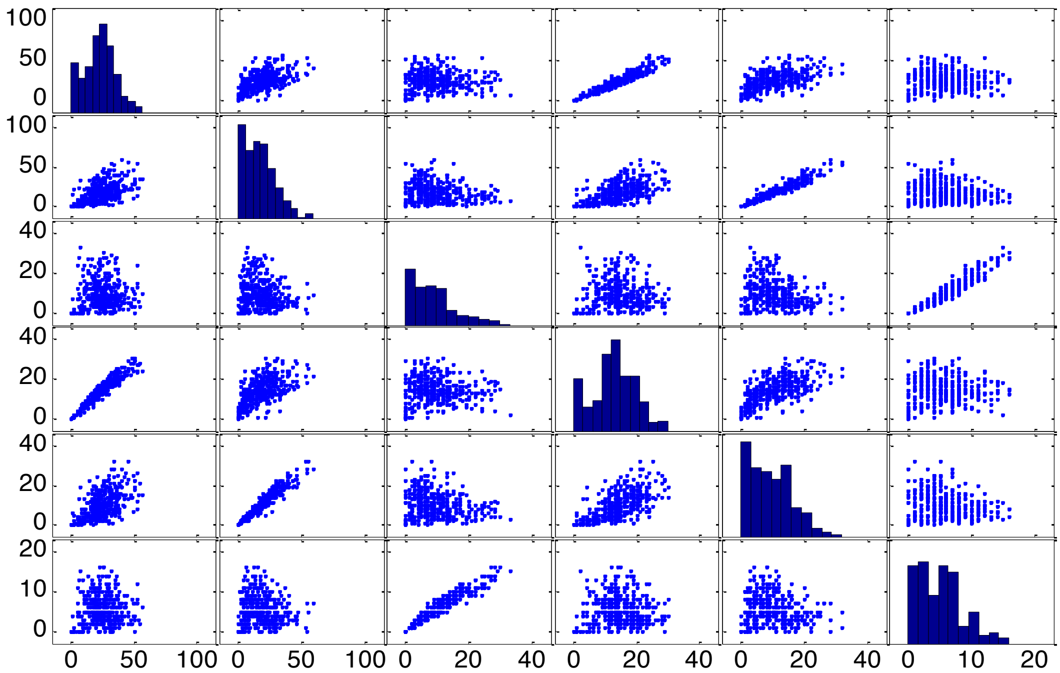
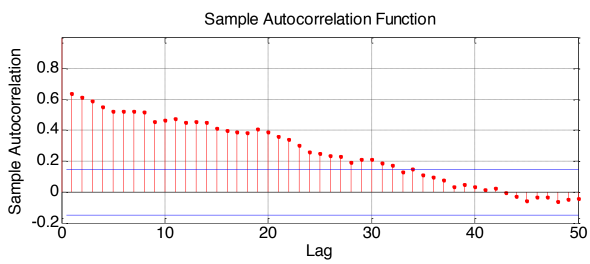


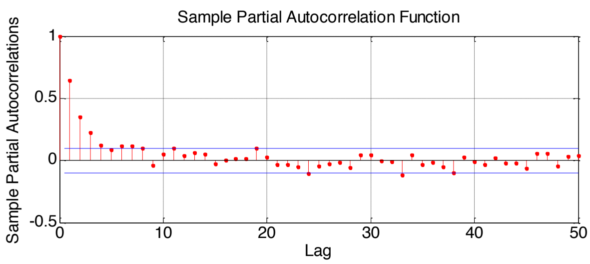
Figure 3: Correlation diagram of the spatially distributed time series.

Table 1: Variance-Covariance and Correlation tables of the weekly crash records.

Variance-Covariance					
156.5323	95.2857	16.9274	84.2496	54.6470	11.2655
95.2857	151.7034	0.3571	55.8090	81.6814	2.6736
16.9274	0.3571	50.8353	8.5287	-0.8847	25.4378
84.2496	55.8090	8.5287	48.7795	32.5921	5.9864
54.6470	81.6814	-0.8847	32.5921	47.0253	1.0085
11.2655	2.6736	25.4378	5.9864	1.0085	13.7437
Correlation					
1.00000	0.61834	0.18976	0.96416	0.63694	0.24288
0.61834	1.00000	0.00407	0.64877	0.96707	0.05855
0.18976	0.00407	1.00000	0.17127	-0.01810	0.96238
0.96416	0.64877	0.17127	1.00000	0.68050	0.23120
0.63694	0.96707	-0.01810	0.68050	1.00000	0.03967
0.24288	0.05855	0.96238	0.23120	0.03967	1.00000

The above evidence among variables is encouraging for justifying spatial correlation of crash occurrence. The next step of the analysis correspond to the identification of temporal evolution and correlation in the dataset and thus to the crash occurrence phenomenon. For investigating this, at first a linear correlation of each variable is estimated, by means of calculating the autocorrelation and partial autocorrelation of all variables (Figure 4&5).





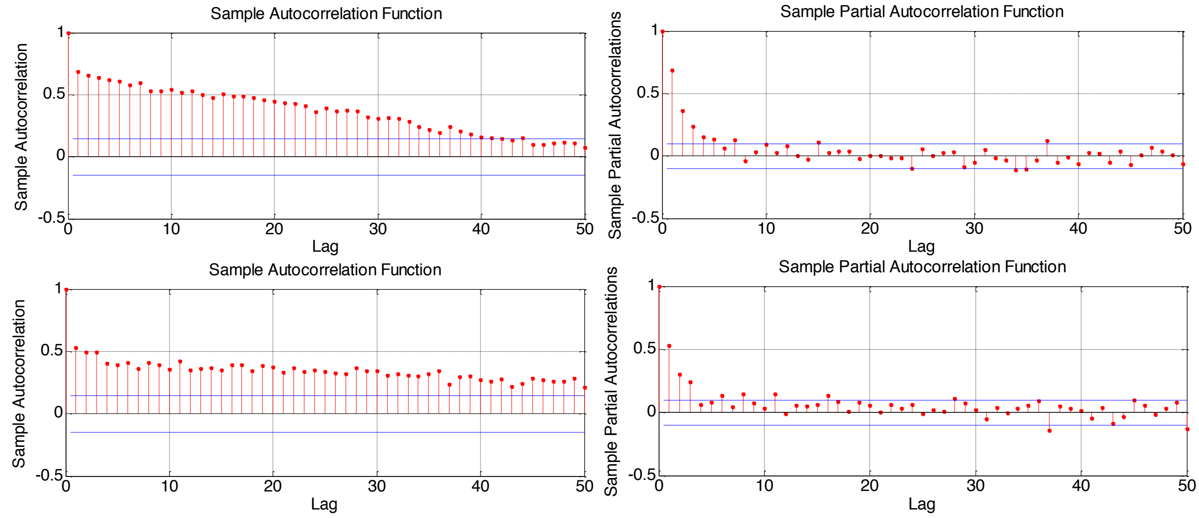


Figure 4: Autocorrelation and Partial Autocorrelation of weekly time series for the number of cars involved in car crashes per region.

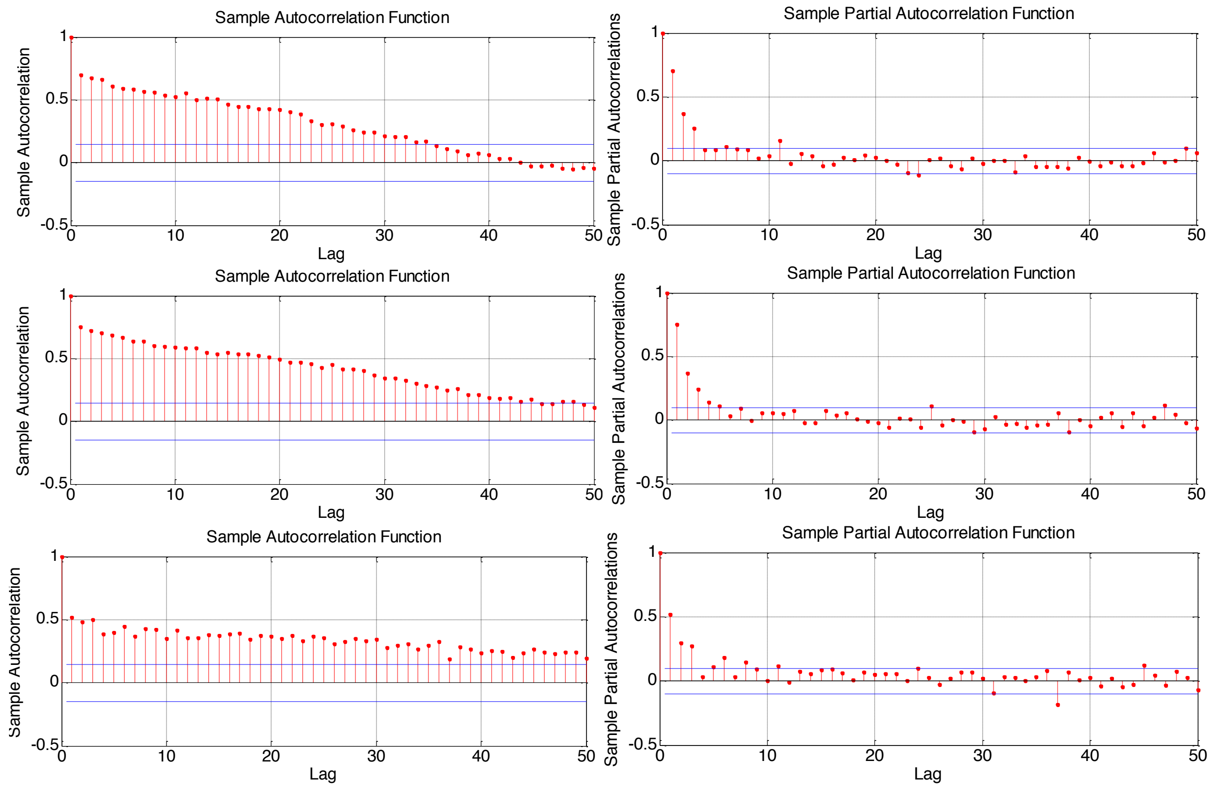


Figure 5: Autocorrelation and Partial Autocorrelation of weekly time series for the number of crashes occurred per region.

In order to investigate cross-correlations among the variables and especially not only linear but also possible higher order correlations, here the analysis will be based on examining the resulting Andrews plot (Andrews 1972) of the dataset. Andrews plots provide a comprehensive mean to depict relevance (if not correlation) in multivariate datasets.

It is explicitly noted here that cross-correlations are also estimated, exhibiting significant temporal cross-correlations but introducing 15 more diagrams are omitted here for brevity.

Moreover, the Andrews plots are providing evidence on possible correlations in multivariate datasets and since the current datasets corresponds to time series, if Andrews plots are revealing possible correlations this spans both to the time as well as to the spatial domain.

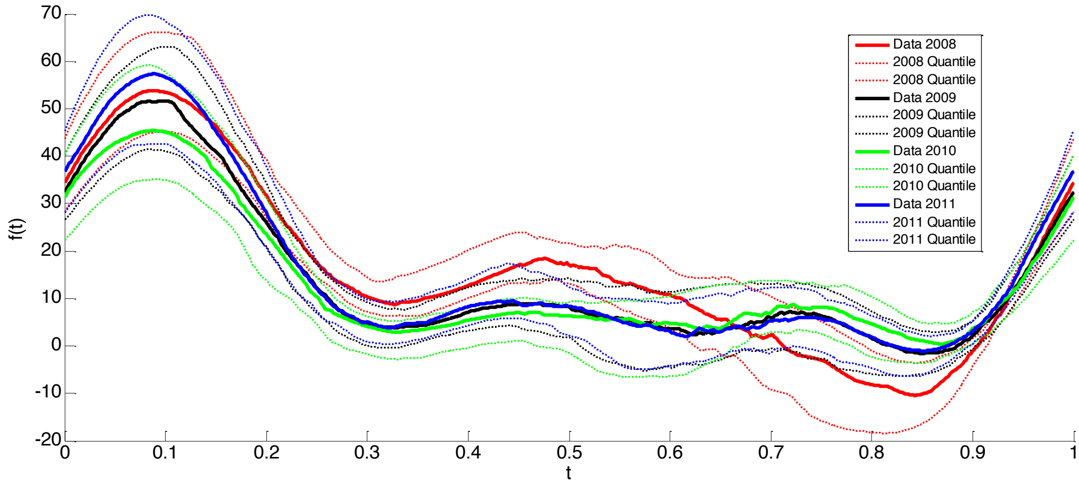


Figure 6: Andrews plot for the mean values of the spatially distributed time series for 4 consecutive years (2008-2011).

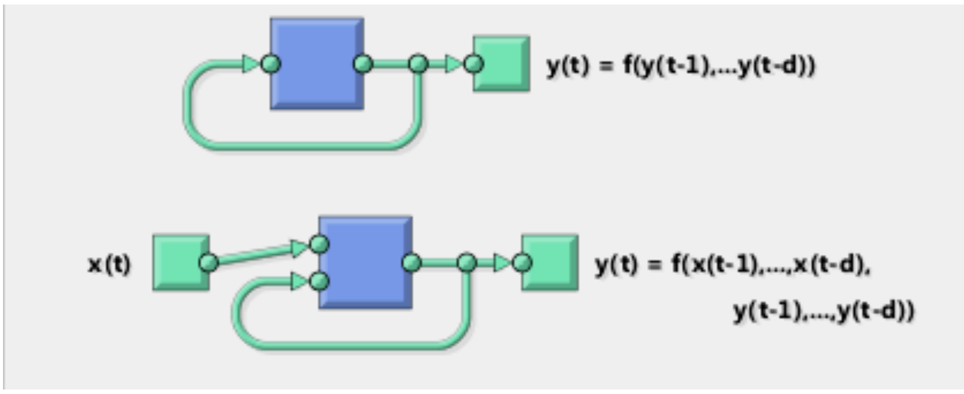
In Figure 6, the Andrews plot of the mean value of the each set of the weekly observations are plotted for 4 consecutive years (2008-2011), accompanied with the plots of the quartile of the mean. As can be observed all lines are following the same pattern and especially those of the three last years are closely placed in the diagram. This observation provide evidence on the correlation of the multivariate dataset and since the dataset is composed of spatial and temporal observations this is an evidence on the correlation of the crash occurrence phenomenon, encouraging for performing time series analysis for predicting future crash risk in the urban context.

In the current paper a special paradigm of time series prediction mechanism will be developed and tested, namely that of the DANNs following an auto-regressive form with and without exogenous stimulus. In the following section, the description and performance of the proposed DANN predictor is presented, along with results from test-runs.

3. METHODS

Artificial Neural Networks have been used extensively in mapping complex input-output interrelations, based in their capability to identify patterns within datasets. Since the monitoring of traffic networks involves the analysis of extensive datasets (typically time series of traffic variables, like volume, occupancy, speed, etc) which are subject to fluctuations (especially those coming from urban arterials), ANN have been tested for modeling and prediction purposes (for a comprehensive review see Karlaftis and Vlahogianni, 2010).

Especially when dealing with dynamic and fluctuating datasets suggesting alterations in the operational characteristics of the underline system, I suitable model setup refers to regularly updated dynamic ANNs. An approach of autoregressive DANNs with error feedback has recently proposed by Dimitriou and Hassan (2013) for predicting crashes times series, providing interesting insights on the time series-based prediction of such highly stochastic datasets. Extending such approaches, in the current paper alternative dynamic network architectures are deployed for predicting the earlier described spatially distributed crash dataset, featuring autoregression and moving average characteristics with feedback loops. In brief, at Figure 5 the two general classes of the proposed DANNs are presented, where the input layer is composed of records from historical time, with a feedback loop without exogenous inputs (Figure 7*a*) and with exogenous inputs (Figure 7*b*).



(a)

(b)

Figure 7. The Structure of the Proposed Dynamic Autoregressive Artificial Neural Networks without exogenous inputs *(a)* and with exogenous inputs *(b)*.

For the case of the DAAN without exogenous inputs, the input-output mapping that is endeavored is based on the calibration of an ANN, in which the inputs corresponds to *d* (d is termed as time-lag) prior observations of the variables that are subject to projection in order to predict the next observation. Here available dataset correspond to three variables, namely, the weekly number of crashes in each of the earlier described three regions of Riyadh, while the expected output is of order 3 (3 outputs: the number of cars involved in serious accident in each region). For the available dataset and configuration, when selecting a lag operator of order *d*, the number of inputs of the DANN results to (d+1)×3. Then at each time *t* the inputs are composed from the triplet observations at *t*-1 plus the *d* number of triples (of the selected time lag) that will provide prediction information. This class of models is solely autoregressive, in the sense that the predictions are based only on the historical information that can be identified and propagated 'inside' the variables set under projection (i.e. the time evolution of each variable and the correlation among variables). For brevity, this class of DANN for now on will be termed as NAR-*d*, where d stands for the lag operator, N for the order of variables' vector under projection and AR for the autoregressive feature of the DAAN calibrated.

For comparative purposes, an enhanced model architecture has been developed, where a set of additional -but correlated in some sense- variables is augmented to the inputs set aiming on contributing additional information possibly valuable for the prediction and termed as exogenous variables. The general architecture is depicted on Figure 7*b*. Here, the 'exogenous dataset' are the number of car crashes per region, also introduced by the lag operator *d,* while the exogenous variables corresponds to the d×3 inputs. At this DANN the total number of inputs equals 2×(d+1)×3, a significantly larger ANN configuration. This class of DANN accordingly is termed as NARX-*d*, adding symbol X which stands for the introduction of exogenous variables. All the above-proposed architectures of the DAAN can be calibrated/trained by the optimization routines typically used in dynamic ANNs.

4. COMPUTATIONAL EXPERIENCE

In order to test the performance of the earlier presented configurations of the DANNs, a highly nonlinear neural network configuration has been selected. In detail, for all types of DANNs, a network of one hidden layer has been used with 50 neurons, each of which use a log-sigmoid transfer function, able to be calibrated by a non-linear optimization routine (details on the calibration of ANNs are omitted here, since can be found in relevant textbooks widely available). Here, the Levenberg–Marquardt algorithm has been used for training the specific type of DANNs. The calibration process utilizes 3 parts of the available dataset. In particular this is divided in three sections, namely, the training set (in which the optimization algorithm is using for updating the direction of the ANN calibration), the validation set (which is used for stopping the training process by avoiding biased fitting or over-fitting) and the test set (which corresponds to a randomly selected subset used for performance evaluation of all calibrated ANNs architectures). The scheme used here for distributing the available dataset (of 414 weekly records) into the three above-mentioned parts is 70%-training, 15%-validating and 15%-testing.

The first attempt to investigate the performance of the earlier introduced DANNs to the available dataset stands for a parsimonious autoregressive model with a lag operator d=2 (NAR-2). This model of 'short memory', at time t utilizes information of the current and the past two observations of the weekly number of cars involved in serious accidents (in total at time t, t-1 and t-2) in each on the 3 regions (see Section 2) in order to predict the number of cars involved in crashes per region at time t+1.

The training convergence diagram and characteristics are presented in Figure 8*a*&*b*. As it can be observed in Figure 8*a* the training process selects the network's configuration resulted from the 9th epoch of training, since after that the network begins to 'drift' toward the values of the training set (the error metric of the training set is reduced) with negative effects on its generalization (the error on the validation set increases). A 'picture' on quality of the resulted NAR-2 can be deduced by the errors distribution depicted in Figure 8*b*, where the these are concentred around zero, an indication of unbiased prediction properties of the trained network.

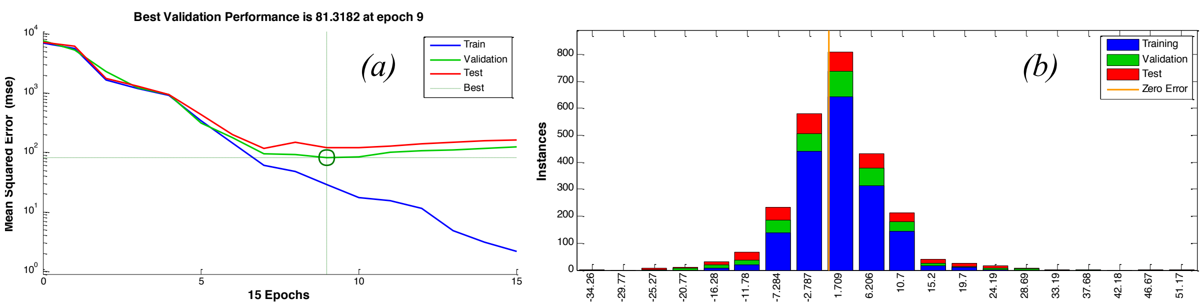


Figure 8. Typical training results for a NAR-2 model: convergence diagram (a) and histogram of errors (b).

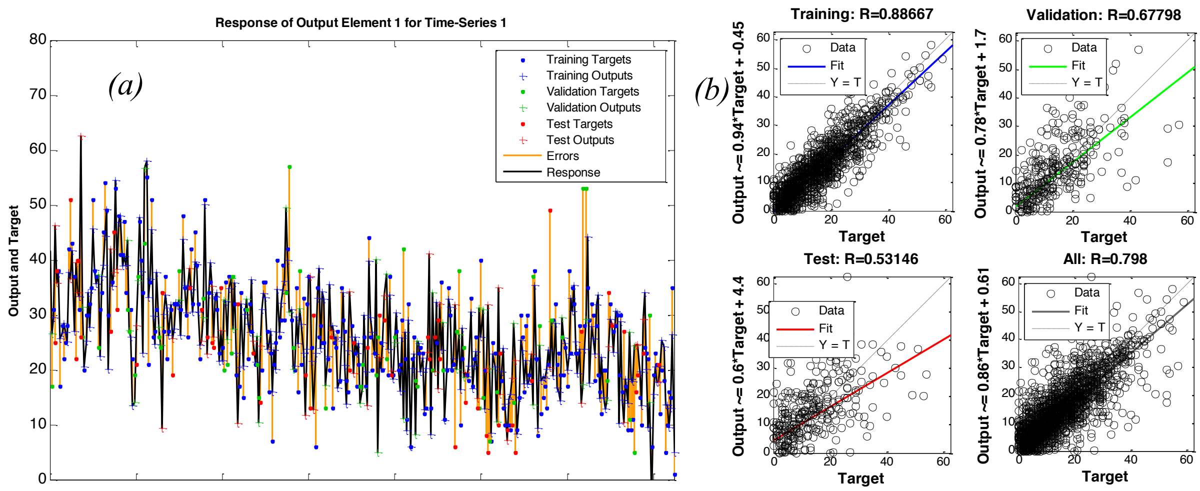


Figure 9. Predictions vs. Observations time series for region #3 (a) and correlation diagram (b) for a typical NAR-2 model.

Quantitative evidence about the performance of the trained network can be drawn from the predictions accuracy. In particular, in Figure 9*a* the time series of observations vs. predictions are depicted (by distinguishing training, validation and testing points for checking the randomness of each point selection) from where it can be observed that this network configurations can follow and predict –with sufficient accuracy- the evolution of regional car crashes. This is further supported by the correlation diagrams (distinguished per set used in the calibration process) and R values (Figure 9*b*). It is noted, that although the R value of the training set quite large (R=0.88), in the test set the R is significantly reduced (R=0.53). The results of this model highlight two issues: first the dataset represents a highly stochastic phenomenon (even for the case of weekly data), and, additionally the specific model has reduced prediction abilities (a model with R=0.53 is considered low for practical use). Although several performance metrics and model configurations could be tested, here the R-value will be used for the rest of the paper for comparative purposes.

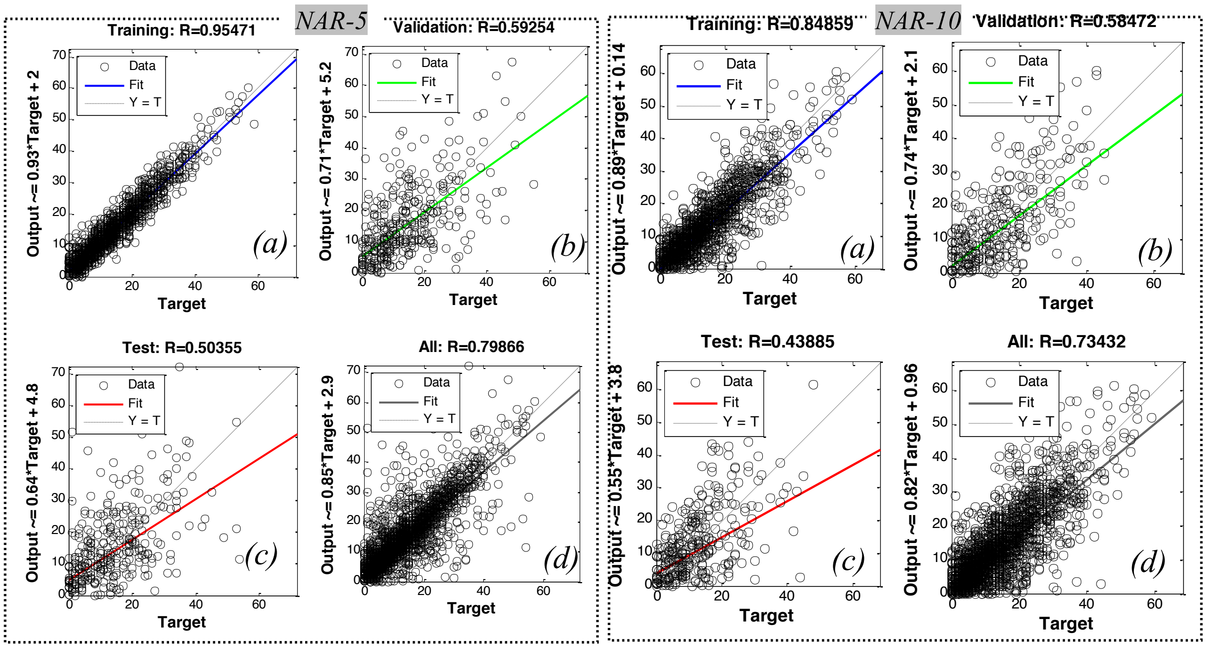


Figure 10. Correlation diagrams of Predictions vs. Observations for the Training set (a), Validation set (b), Test set (c) and for all dataset (d) for NAR-5 and NAR-10 models.

In Figure 10*a&b*, two additional autoregression models with 'longer memory' are tested, namely a NAR-5 and a NAR-10). Although these models are utilizing larger information source (of 5 and 10 lags respectively), as can be observed by the R value of their test sets have reduced prediction accuracy (R=0.50 and R=0.43 respectively). It is noted that for both networks their performance on the training set are comparable to the NAR-2 (suggesting that the optimization algorithm fully exploit the information available in the training set). On the overall, it can be observed that the network size is not the determinant factor in the model's generalization and that parsimonious models can perform better than larger and more complicated ones. Also, the results suggest that additional sources of information should be investigated in order to come up with results of practical use.

The next step of the experiments correspond the investigation of the role of additional information on the autoregressive models performance. Thus NARX-2 and NARX-10 models are tested while the results are presented in Figure 11*a&b*.

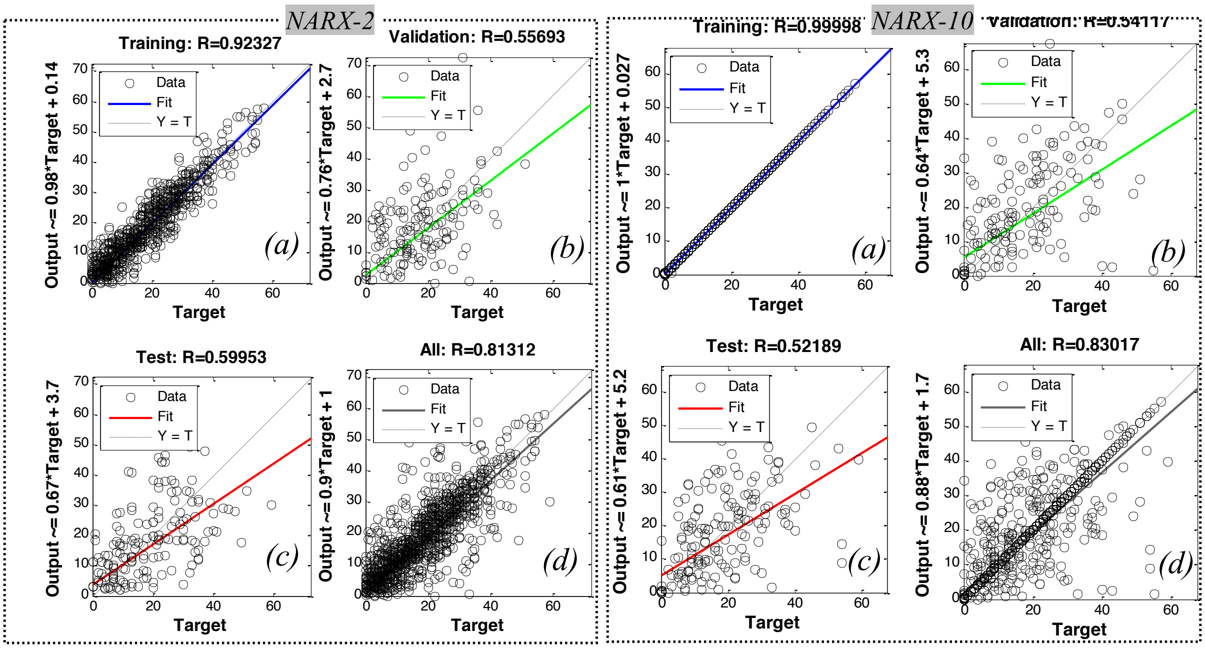


Figure 11. Correlation diagrams of Predictions vs. Observations for the Training set (a), Validation set (b), Test set (c) and for all dataset (d) for NARX-2 and NARX-10 models.

Interestingly, NARX-2 performs better than all models on the test set (R=0.60) while the prediction accuracy can be regarded as comparable to other relevant crash prediction models presented in the literature review section. Also, it is worth mentioning that although the relatively large NARX-10 model (of 66 inputs-3 outputs) in the training process was able to replicate the training set, its generalization was limited compared with that of the NARX-2 model (of the 18 inputs-3 outputs), further supporting the evidence that larger and complicated models cannot guaranty benefits in the prediction accuracy. Finally, it is evident from the experiments, although the results were encouraging for the deployment of exploratory models (such spatially time series models) for predicting car crashes phenomena (and car crash risk), the necessary information that these should be provided is a determinant factor in their performance and usefulness.

5. CONCLUSIONS AND RECOMMENDATIONS

In the current paper a prediction model of car crashes has been presented, belonging to the Artificial Intelligent class of models, namely in Dynamic Artificial Neural Networks. In particular, alternative hybrid forms of Dynamic Artificial Neural Networks (DANN) have been developed and tested for the case of time series predictions of crashes, in a spatially distributed format useful for surveillance and operational management of large-scale urban networks. The application of the proposed prediction mechanisms are tested over a dataset of weekly crash records from the Riyadh, capital of the Kingdom of Saudi Arabia, an area experiencing severe road safety issues for years now accompanying the rapid development of the urban space.

A number of future extensions can be regarded in this line of research, ranging from testing alternative statistical and AI models to the application and use of such analytical methods for operational purposes (e.g. daily datasets, exact location of crashes etc). Also, it is regarded that the augmentation of other traffic characteristics (traffic volume, occupancy), environmental information (visibility, whether, etc) and time-depended variables (time of day, seasonal events, etc) will enhance the models predicting power. Such experiments have been already scheduled for the near future.

ACKNOWLEDGEMENTS: The authors would like to thank the Municipality of Riyadh for providing the Saudi crash dataset used in this study. All opinions and results are those of the authors.

6. REFERENCES

Abdel-Aty M. and Abdelwahab H. (2004) Analysis and Prediction of Traffic Fatalities Resulting from Angle Collisions Including the Effect of Vehicles' Configuration and Compatibility, Accident Analysis and Prevention, Elsevier, Volume 36, Issue 3, pp 457-469.

Abdel-Aty, M., and Pande, A. (2005). Identifying crash propensity using specific traffic speed conditions. Journal of Safety Research, 36, 97–108.

Abdel-Aty, M., Pande, A., Das, A., and Knibbe, W. (2008). Assessing safety on Dutch freeways with data from infrastructure-based intelligent transportation systems. Transportation Research Record, 2083, pp. 153-161.

Abdel-Aty, M., Uddin, N., Abdalla, F., Pande, A., and Hsia, L. (2004). Predicting freeway crashes based on loop detector data using matched case–control logistic regression. Transportation Research Record, 1897, 88–95.

Abdelwahab H. and Abdel-Aty M. (2004) Investigating the effect of LTV percentages on Rear-End fatal crashes (Part I), ASCE Journal of Transportation Engineering, Vol.130, No. 4.

Abdelwahab H. and Abdel-Aty M. (2004) Investigating the effect of LTV percentages on head-on fatal traffic crashes (Part II), ASCE Journal of Transportation Engineering, Vol. 130, No. 4.

Al-Ghamdi, A.S. (2002a). Pedestrian–vehicle crashes and analytical techniques for stratified contingency tables. Accident Analysis and Prevention, 34, pp. 205–214.

Al-Ghamdi, A.S. (2002b). Emergency medical service rescue times in Riyadh. Accident Analysis and Prevention, 34, pp. 499–505.

Al-Ghamdi, A.S., AlGadhi, S. A. (2004). Warning signs as countermeasures to camel–vehicle collisions in Saudi Arabia. Accident Analysis and Prevention, 36, pp. 749–760.

Christoforou, Z., Cohen, S., and Karlaftis, M. (2011). Identifying crash type propensity using real-time traffic data on freeways. Journal of Safety Research, 42, 43–50.

Dimitriou, L., and H.M. Hassan (2013). “Dynamic Autoregressive Neural Networks for Spatially Distributed Time Series Prediction of Car Crashes in Urban Networks”. Transportation Research Board Annual Meeting, Washington DC.

Dimitriou, L., T. Tsekeris, and A. Stathopoulos (2008). “Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow”, Transportation Research Part-C, Elsevier, Vol. 16 (5), pp. 554-573.

Karlaftis, M.G., and E.I. Vlahogianni (2010). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. Transportation Research Part C, 19 (2011) 387–399.

Kwon, E. and Stephanedes, Y. J. (1994), Comparative evaluation of adaptive and neural-network exit demand prediction for freeway control, Transportation Research Record, 1446, 66–76.

Stathopoulos, A., Dimitriou, L., and Tsekeris, T. (2008). Fuzzy Modeling Approach for Combined Forecasting of Urban Traffic Flow. Computer-Aided Civil and Infrastructure Engineering, 23 (7), 521–535.

United Nations 2010. Road Safety Collaboration’s "Decade of Action for Road Safety 2011-2020: Global Launch". http://www.who.int/roadsafety/decade_of_action/en/.

Vlahogianni, E.I., Karlaftis, M.G. and Golias, J.C. (2005), Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach, Transportation Research, 13C(3), 211–234.

Williams, B. M. (2001), Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling, Transportation Research Record, 1776, 194-200.