Neurocomputing 167 (2015) 3–7

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling

Fabio Moretti [a,b,*], Stefano Pizzuti [a,b], Stefano Panzieri [b], Mauro Annunziato [a]

[a] *ENEA (Italian Energy New technologies and sustainable Economic development Agency)*
[b] *Università degli Studi "Roma Tre", Computer Science and Automation Department*

A R T I C L E I N F O

Article history:
Received 7 November 2013
Received in revised form
16 July 2014
Accepted 11 August 2014
Available online 12 May 2015

Keywords:
Traffic flow forecasting
Ensembling
Bagging
Neural networks

A B S T R A C T

In this paper we show a hybrid modeling approach which combines Artificial Neural Networks and a simple statistical approach in order to provide a one hour forecast of urban traffic flow rates. Experimentation has been carried out on three different classes of real streets and results show that the proposed approach outperforms the best of the methods it puts together.

1. Introduction

Transportation is a wide human-oriented field with diverse and challenging problems waiting to be solved. Characteristics and performances of transport systems, services, costs, infrastructures, vehicles and control systems are usually defined on the basis of quantitative evaluation of their main effects. Most of the transport decisions take place under imprecision, uncertainty and partial truth. Some objectives and constraints are often difficult to be measured by crisp values. Traditional analytical techniques were found to be ineffective when dealing with problems in which the dependencies between variables were too complex or ill-defined. Moreover, hard computing models cannot deal effectively with the transport decision-makers' ambiguities and uncertainties. In order to come up with solutions to some of these problems, over the last decade there has been much interest in soft computing applications of traffic and transport systems, leading to some successful implementations [1]. The use of soft computing methodologies for modeling and analyzing traffic and transport systems is of particular interest to researchers and practitioners due to their ability to handle quantitative and qualitative measures, and to efficiently solve complex problems which involve imprecision, uncertainty and partial truth. Soft computing can be used to bridge modeling gaps of normative and descriptive decision models in traffic and transport research. Transport problems can be classified into four main areas: traffic control and management, transport planning and management, logistics, design and construction of transport facilities. The first category includes traffic flow forecasting which is the topic tackled in this work. This issue has been faced by the soft computing community since the nineties [4–10] up today [12–14,11] with Artificial Neural Networks (ANNs) [2,3]. As example, among the most recent work [14] focuses on traffic flow forecasting approach based on Particle Swarm Optimization (PSO) with Wavelet Network Model (WNM). Pamula et al. [11] review neural networks applications in urban traffic management systems and presents a method of traffic flow prediction based on neural networks. Bucur et al. [12] propose the use of a self-adaptive fuzzy neural network for traffic prediction suggesting an architecture which tracks probability distribution drifts due to weather conditions, season, or other factors. All the mentioned applications have one feature in common: they use one single global model in order to perform the prediction. Therefore, the main novelty of the proposed work is to combine different heterogeneous models in order to get a meta-model capable of providing predictions more accurate than the best of the constituent models. In our work we firstly composed of a neural networks ensemble with a simple statistical model and compare the results over the one hour forecast, then we improved ensembling model with BAGGING. Results shown highlight a remarkable decrease of error through the BAGGING learning phase.

* Corresponding author.
E-mail addresses: fabio.moretti@enea.it (F. Moretti), stefano.pizzuti@enea.it (S. Pizzuti), stefano.panzieri@uniroma3.it (S. Panzieri), mauro.annunziato@enea.it (M. Annunziato).

http://dx.doi.org/10.1016/j.neucom.2014.08.100
0925-2312/

2. Methods

2.1. Basic model

In order to perform a meaningful comparison for the forecasting, a basic model should be introduced in order to quantify the improvement given by more intelligent and complex forecasting techniques. For seasonal data a basic model might be defined as

$$x_t = x_{t-s} \tag{1}$$

with S being the appropriate seasonality period. This model gives a prediction at time t presenting the value observed exactly a period of S steps before. For this work we put the value of S=1 which corresponds to the previous hour. It means that to predict the flow rate of the following hour it is used the current flow measure.

2.2. Statistical

One of the simplest and most widely used models when dealing with regular time series (as urban traffic flows) is to build an average weekly distribution of the traffic flow sampled hourly. Thus, from the data we compute for each day the average flow rate hour by hour in such a way that we get an average distribution made of $24.7 = 168$ points.

2.3. Neural network ensembling

Models ensemble is a technique where many prediction models cooperate on the same task. The aggregation of multiple prediction of the same variable may lead to better results and generalization than using a single model prediction. In order to increase generalization capability, the model learning phase is crucial. The goal obtains better predictive performance than could be obtained from any of the constituent models. In the last years, several ensembling methods have been carried out [17,15,16]. The first one, also known as Basic Ensemble Method (BEM), is the simplest way to combine M neural networks as an arithmetic mean of their outputs. This method can improve the global performance [20,21] although it does not take into account that some models can be more accurate than others. This method has the advantage to be very easy to apply. A direct BEM extension is the Generalised Ensemble Method (GEM) in which the outputs of the single models are combined in a weighted average where the weights have to be properly set, sometimes after an expensive tuning process. Bagging (Bootstrap AGGregatING) [18] technique improves generalization: for each learner replaces part of the training data set with a random combination of training data itself. Thus each dataset may contain duplicated entries of the same sample or not at all. Improvement occurs especially when small changes in dataset may lead to a large changes in prediction. Adaboosting [19] introduces weights on the training points.

2.4. Hybrid model

Hybrid models are an extension of the ensembling approach in the sense that the final goal is to combine different models in such a way that the accuracy of the composition is higher than the best of the single models. The difference is that the combination is performed among highly heterogeneous models, that is models generated by different methods with different properties and thus the composition among them is a complex rule taking into account the peculiarities of the models and/or of the problem itself. Therefore, in this work we propose a novel hybrid model which combines an ANN ensemble with the statistical model. The composition rule is the following : "IF the statistical model has a high error (meaning that for some reason we are out of a normal situation) THEN use the neural model ELSE use the statistical one". This criterion is based on the absolute error of the statistical model, thus the composition rule turns into

$$|x^t - y^t| > \epsilon \Rightarrow y^{t+1} = y_n^{t+1} \tag{2}$$

$$|x^t - y^t| \le \epsilon \Rightarrow y^{t+1} = y_s^{t+1} \tag{3}$$

where y^{t+1} is the outcome (one hour prediction) after the composition rule, y_n^{t+1} is the prediction of the neural ensemble, y_s^t is the current outcome of the statistical model and y_s^{t+1} is its prediction. This basically means that if we are in normal statistical conditions (where the statistical model makes a small error) then use as prediction model the statistical one (which is very accurate in this condition), else (when out of normal statistical situations) take the neural ensembling estimation.

3. Experimentation

In this paragraph we test and compare the methods presented in the previous section. The test case has concerned the short term traffic flow rate of three different streets, shown in Table 1, located in the town of Terni (about 90 km north of Rome). The data set is made of 3 months (13 weeks) of measurement corresponding to 2184 hourly samples.

The data set has been partitioned into training/testing and validation made respectively of 10 and 3 weeks each. We firstly present the result obtained using hybrid model based on Neural Network Basic Ensemble model and statistic model and we show an improvement on the forecasting, then we replace Basic Ensemble Model with a Bagging based one. Results show a further improvement on the forecasting.

3.1. Neural network setup

The ANN are feed-forward MLP with 10 hidden neurons and one output (the one hour flow forecast) with sigmoid as activation function for all the neurons. The number of inputs N has been

Table 1
Street parameters.

Street	Maximum traffic flow rate
Street 1	600
Street 2	800
Street 3	950

Table 2
History length selection.

N (h)	Street 1	Street 2	Street 3
3	5.72	6.88	5.81
5	3.90	5.07	3.99
8	3.29	3.43	3.02
10	3.54	4.12	3.74

Table 3
Hybrid model parameter ϵ tuning. Errors percentage of hybrid model at different values of ϵ parameter.

Street	$\epsilon = 10$	$\epsilon = 20$	$\epsilon = 30$	$\epsilon = 40$	$\epsilon = 50$	$\epsilon = 60$
Street 1	2.98	2.83	2.81	2.80	2.88	2.99
Street 2	2.85	2.69	2.65	2.66	2.68	2.75
Street 3	3.25	3.13	3.08	3.04	3.03	3.04

Table 4
Bagging factor parameter tuning. Results are averaged over 10 runs with different bagging factors.

bf	0.75	0.8	0.85	0.9	0.95	1
Street 1	3.56 ± 0.11	3.41 ± 0.14	3.39 ± 0.17	3.21 ± 0.17	3.12 ± 0.14	3.11 ± 0.20
Street 2	4.06 ± 0.15	4.03 ± 0.21	3.87 ± 0.17	3.73 ± 0.19	3.65 ± 0.18	3.52 ± 0.21
Street 3	3.14 ± 0.14	2.98 ± 0.15	2.94 ± 0.19	2.86 ± 0.11	2.76 ± 0.13	2.69 ± 0.11

Table 5
Comparison of models percentage error. Basic is the 1-step forward model, Stat is the weekly average profile, ANN is Artificial Neural Networks model, BEM is Basic Ensemble Method obtained averaging each neural network output, HBEM is the hybrid statistical and BEM model, BEG is the bagging ensemble model and finally HBAG is the hybrid statistical and Bagging ensemble model

Street	Basic	Stat	ANN	BEM	HBEM	BAG	HBAG
Street 1	8.92	5.90	4.93 ± 0.25	4.58	3.29	3.01 ± 0.11	2.42
Street 2	9.99	7.14	5.22 ± 0.31	4.59	4.15	3.52 ± 0.21	3.01
Street 3	7.66	5.56	3.70 ± 0.28	3.71	2.93	2.69 ± 0.11	2.30

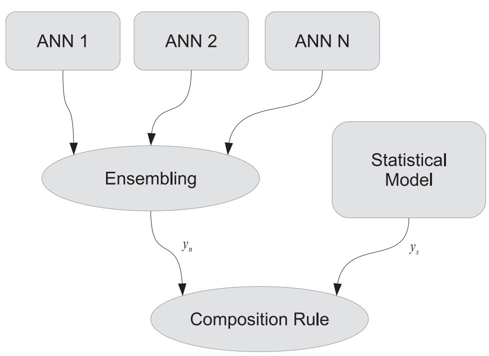


Fig. 1. Proposed hybrid model approach.

chosen with a preliminary analysis by calculating the validation prediction error after ensembling for different values of N as shown in Table 2.

By this analysis it turned out the optimal number of input neurons (namely the length of the history window) to be eight. Training has been performed through the Back-Propagation algorithm with adaptive learning rate and momentum stopping after 100,000,000 iterations and a save best strategy to avoid overfitting. The reported results are averaged over 10 different runs (with standard deviation in brackets) and the ensemble is therefore made by the same 10 models.

The reported errors are measured as

$$e = \frac{|x-y|}{M-m} \tag{4}$$

where x is the real value to be predicted, y is the output model, M is the real maximum value and m is the minimum.

Afterwards, it has been tuned parameter ϵ of the hybrid model. Table 3 shows error obtained with the hybrid model composed of statistical and BEM models, ϵ value is the threshold expressed in number of vehicles.

3.2. Bagging ensemble setup

Bagging factor parameter determines the percentage of the original data set that is going to be replaced by a recombination. Therefore we carried out a tuning of the parameter, evaluating mean ensemble error for different bagging factors. Table 4 shows that increasing bagging factor implies an almost negligible

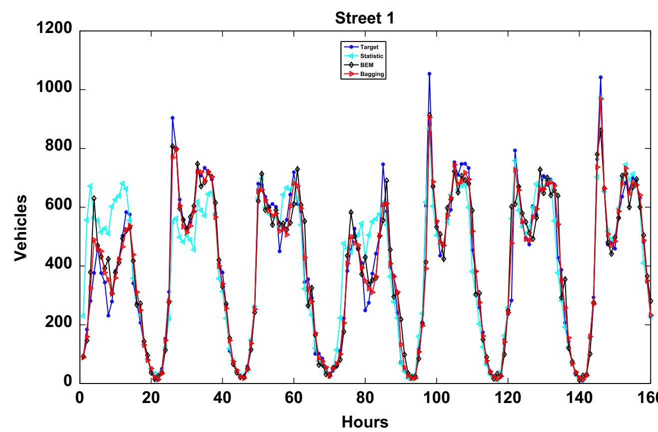


Fig. 2. Street 1 models comparison.

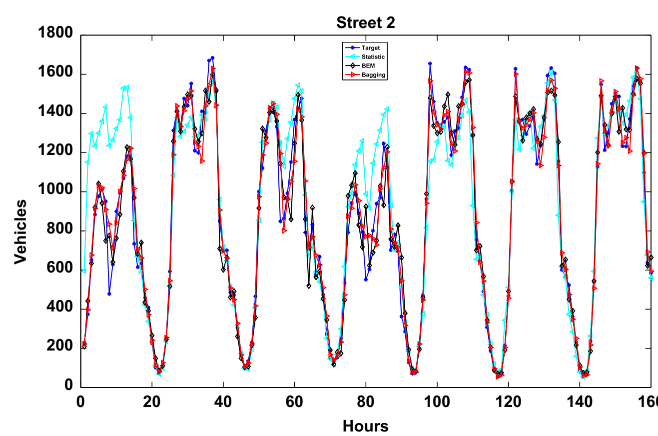


Fig. 3. Street 2 models comparison.

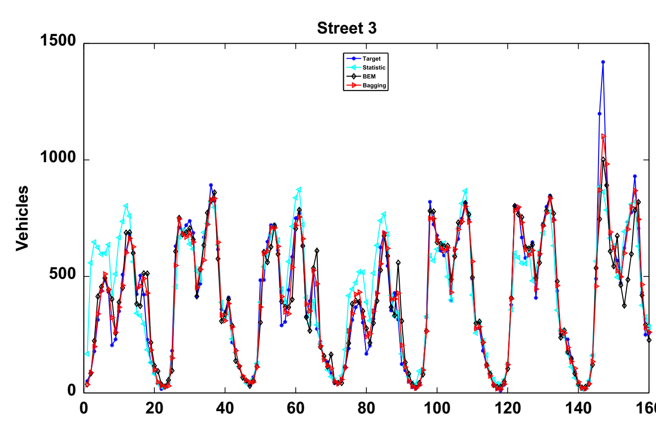


Fig. 4. Street 3 models comparison.

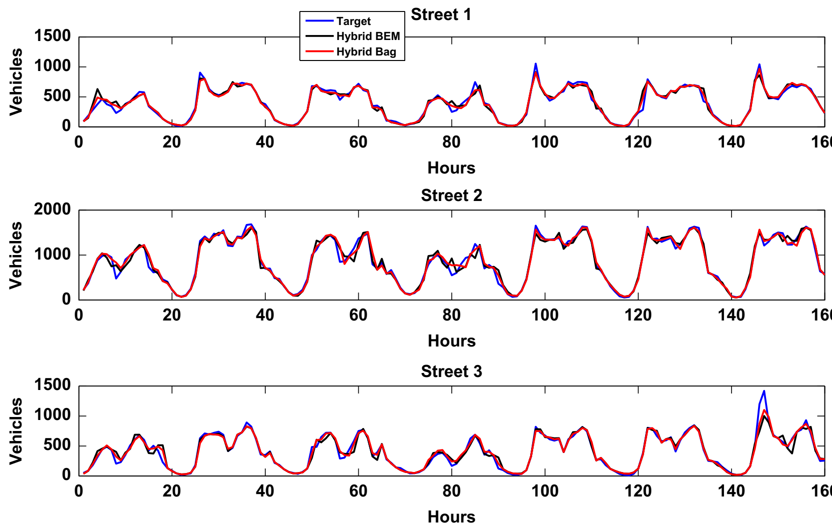


Fig. 5. Comparison of the statistical and BEM hybrid model with statistical and Bagging hybrid model on the 3 streets case study.

betterment of the performances. In our tests we used bagging factor 1, that is a recombination of the whole learning dataset.

3.3. Results

Experimantations carried on shows a comparison of the performance of the models described. Basic model method simply forecast next hour traffic flow according to the previous measured one, so $x_{t+1} = x_t$. Statistic model builds a weekly profile of the traffic flow, so in our case the profile is an average over the 10 weeks of the hourly traffic flow. The result is a model of 168 points representing the average profile of 24 h for 7 days. Since such models do not involve stochastic processes standard deviation is not expressed. Artificial Neural Networks result is the averaged error over 10 runs. BEM is the Basic Ensemble Method, built averaging the output of the Neural Networks previously evaluated. Since BEM error is evaluated on the averaged output also in this case standard deviation is not expressed. HBEM is the hybrid statistic and neural network ensemble based model, in our tests we used $\epsilon = 40$ that means that if the error of the statistical models exceeds ϵ parameter value, used ensemble model, otherwise statistical one. Bagging ensemble results are evaluated over 10 runs with bagging factor parameter set to 1. HBAG is the result of hybrid statistic and bagging ensemble based model. In order to be comparable with HBEM, bagging ensemble output is averaged over 10 runs, then we used the same logic: if statistical error exceeds ϵ parameter, bagging ensemble model is used. Table 5 shows that hybrid model with bagging outperforms previous ones Fig. 1.

Figs. 2, 3, 4 show graphical comparison of statistical, BEM and bagging models and highlight that the use of bagging ensemble outperforms previous techniques.

Fig. 5 shows comparison between the two hybrid models on the 3 streets: target is followed slightly better by the hybrid statistical and bagging model.

4. Conclusions

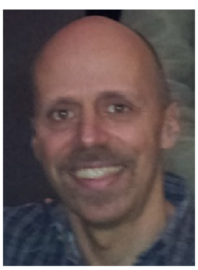
In this paper we showed a novel hybrid modeling approach which combines Artificial Neural Networks and a simple statistical approach in order to provide a one hour forecast of urban traffic flow rates. Experimentation has been carried out on three different classes of real streets and results showed that the proposed approach clearly outperforms the best of the methods it puts together achieving a prediction error lower than 3%. The reason for that is that the neural ensembling model is capable to provide more reliable estimations when out of standard conditions because it considers the real traffic dynamics. The accuracy of the proposed hybrid modeling approach is such that it can be applied for intelligent monitoring, diagnostic systems and optimal control. Future work will focus on further modeling improvements using different composition methods for the hybrid model based on fuzzy sets rather than fixed thresholds. As application, we are going to use these models in public lighting control in order to reduce energy consumption.

References

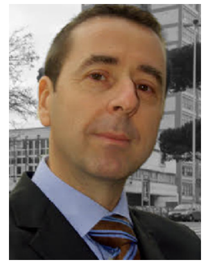
[1] Erel Avineri, Soft computing applications in traffic and transport systems: a review, in: Soft Computing: Methodologies and Applications, Springer, Berlin, Heidelberg, 2005, pp. 17–25.
[2] M.A. Arbib, The Nandbook of Brain Theory and Neural Nnetworks, The MIT Press, Cambridge, MA, 1995.
[3] S. Haykin, Neural Networks, A Comprehensive Foundation, second ed., Prentice Hall, New Jersey, 1995.
[4] D. Canca, J. Larrañeta, S. Lozano, L. Onieva, Traffic intensity forecast in urban networks using a multilayer perceptron, in: Joint International Meeting, Barcelona, 1997.
[5] M.S. Dougherty, M. Cobbett, Short term inter-urban traffic forecasts using neural networks, in: Proceedings of the Second DRIVE-II Workshop on Short-Term Traffic Forecasting, Delft, The Netherlands, 1994, pp. 65–79.
[6] S. Ishak, P. Kotha, C. Alecsandru, Optimization of dynamic neural network performance for short-term traffic prediction, Transp. Res. Rec. 1836 (2003) 45–56.
[7] D. Park, L.R. Rilett, G. Han, Spectral basis neural networks for real-time travel time forecasting, J. Transp. Eng. 125 (6) (1999) 515–523.
[8] C. Taylor, D. Meldrum, Freeway traffic data prediction using neural networks, in: Proceedings of the 6th VNIS Conference, Seattle, WA, 1995, pp. 225–230.
[9] J.W.C. Van Lint, S.P. Hoogendoorn, H.J. Van Zuylen, Freeway travel time prediction with state–space neural networks, Transp. Res. Rec. 1811 (2003) 30–39.
[10] W. Zheng, L. Der-Horng, Q. Shi, Short-term freeway traffic flow prediction: Bayesian combined neural network approach, J. Transp. Eng. 132 (2) (2006) 114–121.
[11] T. Pamula, Road traffic parameters prediction in urban traffic management systems using neural networks, Transp. Probl. 6 (3) (2011) 123–128.
[12] L. Bucur, A. Florea, B.S. Petrescu, An adaptive fuzzy neural network for traffic prediction, control & automation (MED), in: 18th Mediterranean Conference, Marrakech, 2011, pp. 1092–1096.
[13] B.G. Cetiner, M. Sari, Q. Borat, A neural network based traffic flow prediction model, Math. Comput. Appl. 15 (2) (2010) 269–278.
[14] S. Jawanjal, P. Bajaj, A design approach to traffic flow forecasting with soft computing tools, emerging trends in engineering and technology (ICETET), in: 3rd International Conference, Goa, 2010, pp. 81–84.
[15] A. Krogh, J. Vedelsby, Neural network ensembles, cross validation and active learning, Adv. Neural Inf. Process. Syst. 7 (1995) 231–238.
[16] Y. Liu, X. Yao, Ensemble learning via negative correlation, Neural Netw. 12 (10) (1999) 1399–1404.

[17] L. Breiman, Combining Predictors, Combining Artificial Neural Nets—Ensemble and Modular Multi-net Systems, Springer, Berlin, 1999, pp. 31–50.
[18] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.
[19] Y. Freund, R.E. Schapire, Experiments with a new boosting algorithm. in: ICML, Vol. 96, July 1996, pp. 148–156.
[20] Michael P. Perrone, Leon N. Cooper, When Networks Disagree: Ensemble Methods for Hybrid Neural Networks, No. TR-61, Institute for Brain and Neural Systems at Brown University, Providence, RI, 1992.
[21] C.M. Bishop, Neural networks for pattern recognition, Oxford University Press, 1995, pp. 364–369. R. Kohavi, E. Bauer, An empirical comparison of voting classification algorithms: bagging, boosting and variants, Mach. Learn. 36 (1999) 105–142.

Fabio Moretti received in 2009 the B.Sc. degree in Computer Science and Automation from the University of Roma Tre in Rome, Italy where he is currently Ph.D. student. He is a research fellow in ENEA (Energy New technologies and sustainable Economic development Agency). His research interests includes data fusion, data mining, evolutionary computation, optimization and computer vision applied to energy saving issues. He is currently involved in several projects concerning Smart Cities activities, in particular buildings diagnostics, control and optimization and public lighting control.



Stefano Pizzuti, Dr. Master degree in Information Sciences (final voting : 107/110) obtained in 1996 at University of Rome 'La Sapienza'. Since 1997 researcher at ENEA in the fields of advanced monitoring and control systems applied to energy production plants. In the last five years the main research activities have focused on Smart Cities application, namely smart building network management, smart lighting, smart communities and integrated infrastructures. Involved as work-package and task leader in several national smart cities projects and in the Joint Program EERA Smart Cities. Author of more than 60 national and international publications, program committee member of several international conferences and reviewer of many Impact Factor international journals.



Stefano Panzieri received the Laurea degree in Electronic Engineering in 1989 and the Ph.D. in Systems Engineering in 1994, both from the University of Roma "La Sapienza". Since February 1996 he is within the Engineering Department of University of "Roma Tre", as Associate Professor. Research interests are in the field of industrial control systems, robotics and sensor fusion. He is author of more the one hundred papers involving mobile and industrial robots. In particular, in the area of mobile robots, some attention has been given to the problem of navigation in structured and unstructured environments with a special attention to the problem of sensor based navigation and sensor fusion. His research interests includes Interdependency Modeling; Modeling and simulation of complex systems; SCADA vulnerabilities; Data fusion; Distributed algorithms in Sensor Networks; Smart Energy Management; Building Automation Systems.

Mauro Annunziato is currently Director of the Smart Energy Division of the Energy Dept. of ENEA (95 researchers). The activities of the Division include the research and development of new technologies for smart cities, sustainable mobility, critical infrastructures, energy efficiency, smart buildings, smart homes, intelligent systems, robotics, public lighting, smart appliances, ICT city platform. More than 120 scientific publications on Journals, book chapters and international conferences. More than 60 seminars, academic teachings (masters), many participations in Scientific Committees of Int. Journals, Conferences and Associations, organization of workshops and editor of journal issues, many national and international TV/newspaper interviews. More that 5000 web links on international web pages referring to works of M. Annunziato