A decision making system using Deep Learning for earthquake prediction by means of electromagnetic precursors

S.M. Potirakis[1], P. Kasnesis[1], C.Z. Patrikakis[1], Y. Contoyiannis[1], N.A. Tatlas[1], S.A. Mitilineos[1], T. Asano[2], M. Hayakawa[2]

[1]*Department of Electrical and Electronics Engineering, University of West Attica, Campus 2, 250 Thivon and P. Ralli, Aigaleo, Athens, GR-12244, Greece; spoti@puas.gr , pkasnesis@yahoo.gr, bpatr@puas.gr, yiaconto@puas.gr, ntatlas@puas.gr, smitil@puas.gr;* [2]*Hayakawa Institute of Seismo Electromagnetics Co. Ltd., UEC (University of Electro-Communications) Alliance Center #521, 1-1-1 Kojimacho, Chofu, Tokyo, 182-0026, Japan; asano@hi-seismo-em.jp, hayakawa@hi-seismo-em.jp*

When a catastrophic event such as an earthquake (EQ) occurs, general guidelines related to the specific event particular location, must be dynamically adapted in near real time by ad-hoc team of experts in order to identify the most urgent recovery actions for the specific emergency. The procedure for hazard risk management, includes several pre and post disaster interventions to be considered in the preparedness and response phases of risk management. (STORM, 2018). In the case of EQs, the pre-disaster interventions, and level of preparedness are undoubtedly the most important phases in risk management. Therefore, possible EQ prediction, could help avoid disastrous effects caused by an EQ on vulnerable structures, such as these often located in cultural heritage sites. I this work, we propose a strategy for the calculation of the probability for a significant ($M \geq 5.5$) EQ occurrence in order to be co-evaluated within a decision making mechanism which could assist in reaching a high level of

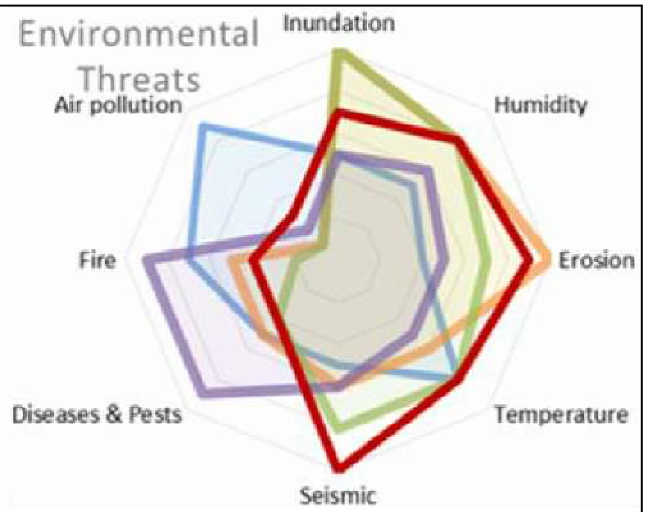

Fig. 1: Pilot sites profiles according to environmental threats.

preparedness. Starting from the theoretical presentation of the methodology proposed, a practical implementation through the integration in a decision making system supported by a computer cloud infrastructure for sensory data is presented. As a reference framework under which the proposed methodology can be applied, the STORM project Cloud infrastructure has been used. STORM project (**S**afeguarding Cultural Heritage through **T**echnical and **O**rganisational **R**esources **M**anagement) project is an ongoing H2020 European research project aiming at providing critical decision making tools to all European Cultural Heritage (CH) stakeholders charged to face climate change and natural hazards (STORM, 2018). The project improves existing processes related to three identified areas: Prevention, Intervention and Policies, planning and processes, and has selected several pilots in CH sites, among which seismic risk is the most common threat, as seen in Fig. 1.

The proposed strategy presupposes the existence of a dense-enough network of VLF/LF receivers for the recording of subionospheric propagation data covering the areas of cultural interest. As an exemplary model of such a network, the network of 8 VLF/LF receivers operating during the last few years throughout Japan which receive subionospheric signals from different transmitters located both in the same and other countries is considered. Based on data collected during a three-year period of operation of the specific network for specific subionospheric propagation quantities we intend to investigate deep-learning (DL) methods for the estimation of the probability for a significant EQ to occur.

According to the conventional nighttime fluctuation method (Hayakawa, 2011), the daily (1 per day) normalized values $DP*$, $TR*$, and $NF*$ of the quantities "trend": $TR = \sum_{N_s}^{N_e} dA(t) \big/ (N_e - N_s)$, "dispersion": $DP = \sqrt{(1/N_e - N_s)\sum_{N_s}^{N_e}(dA(t) - TR)^2}$, and "nighttime fluctuation": $NF = \sum_{N_s}^{N_e}(dA(t))^2$, of the VLF/LF subionospheric propagation data are usually studied. A value exceeding the $\pm 2\sigma$ threshold ($TR* \le 2\sigma$, $DP* \ge 2\sigma$, $NF* \ge 2\sigma$) is considered a candidate precursor. The normalized values are calculated as $X* = \left(X - \langle X \rangle_{\pm 15 days}\right) \big/ \sigma_{\pm 15 days}$, where $\langle X \rangle_{\pm 15 days}$ and $\sigma_{\pm 15 days}$ denote the mean value and standard deviation ± 15 days around the day of interest, respectively. In the abovementioned equations $dA(t)$ is the residue between the received signal amplitude $A(t)$ and an average signal amplitude $\langle A(t) \rangle$ calculated by means of a running average over ±15 days as $dA(t) = A(t) - \langle A(t) \rangle$, and N_s and N_e are the time points of the start and end of the nighttime depending of the period of the year.

Based on the variation of the abovementioned subionospheric propagation quantities normally a prediction of an upcoming EQ determining time period, position and magnitude is made. Considering that a prediction is successful in the cases that (a) the EQ occurred ±2 days from the predicted period, (b) the EQ epicenter was within a radius of 50km around the predicted position, and (c) the magnitude was up to 0.5 different from the predicted one, the success rate of the conventional nighttime fluctuation method is ~ 65% based on the results during the last five years.

DL is a particular branch of machine learning that is based on Artificial Neural Networks (ANNs). In contrast to other machine learning techniques, DL algorithms are capable of extracting features that are a non-linear combination of the input features and are represented in the hidden layer. Previous works based on shallow ANNs (Popova et al., 2013) have proven to be efficient in predicting seismic events based on low-frequency signal monitoring. However, in case of time-series data conventional ANNs cannot capture the local dependencies (Zeng et al., 2014).

Recurrent Neural Networks (RNNs) are a family of neural networks for processing a sequence of values (Goodfellow et al., 2016), and are applied broadly to natural language processing and time-series analysis. In particular, a value x_i depends on a set of previous n values $\{x_{i-1}, x_{i-2}, \dots, x_n\}$. This sequential dependency is represented by adding weighted connections between the hidden states h (Fig. 2A). Moreover, in order to enhance the memory of the network, a mechanism named LSTM (Long Short-Term Memory) (Hochreiter and Schmidhuber, 1997) is applied to it (Fig. 2B).

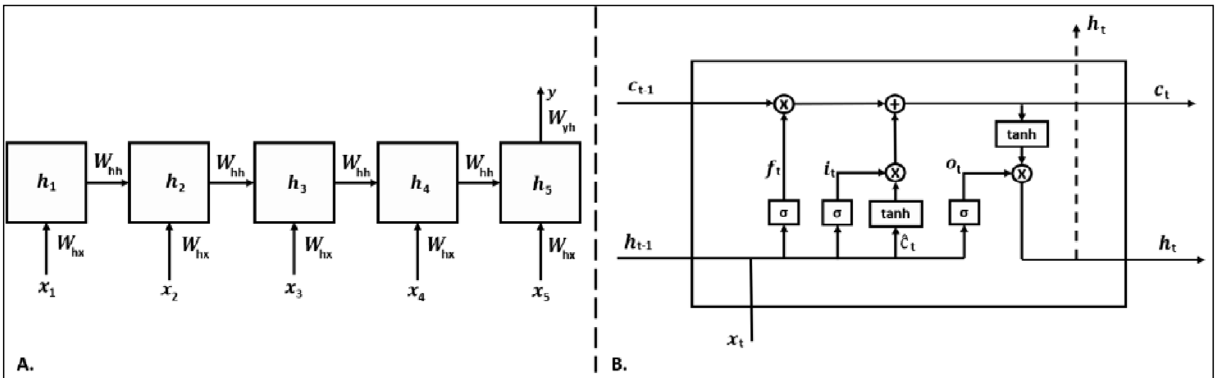

Fig. 2: Graphical representations of an RNN (A) and the LSTM module (B).

LSTM mechanism is capable of learning long-term dependencies, since it uses cell states c^t to transfer information among the hidden units h_t. Furthermore, LSTM, also, makes use of activation gates in order to forget information from the cell state (forget gate f_t), enter new information to the cell state (input gate i_t), and pass information (output gate o_t) to the next hidden state. It should be noted, that similarly to the RNNs the values of the hidden state h_t are updated at every time step t, which in our case represents a day. The equations for representing the update of an LSTM layer are as follows:

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{1}$$

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \tag{2}$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{3}$$

$$\hat{c}_t = tanh(W_{\hat{c}x}x_t + W_{\hat{c}h}h_{t-1} + b_{\hat{c}}) \tag{4}$$

$$c_t = f_t * c_{t-1} + i_t\hat{c}_t \tag{5}$$

$$h_t = o_t * \tanh{(c_t)} \tag{6}$$

where $\hat{c}_t$ denotes the new candidate values for the cell state, and the terms W are the weight matrices, with subscripts representing the connections between the gates and the input vector x_t. Finally, b denotes the bias term that is related to a particular gate.

Having in mind the effectiveness of the Deep RNNs on signal processing tasks (Ordóñez and Roggen, 2016), using an RNN enhanced with the LSTM cell seems to be a well-suited candidate for predicting seismic events. As it is illustrated by Fig. 2A, the per day extracted features x_t ($DP*$, $TR*$, and $NF*$) are going to feed the LSTM layer that after being trained it will output a value y_t between 0 and 1, showing how much the subionospheric perturbations, which occurred within the previous days, are correlated with a seismic event.

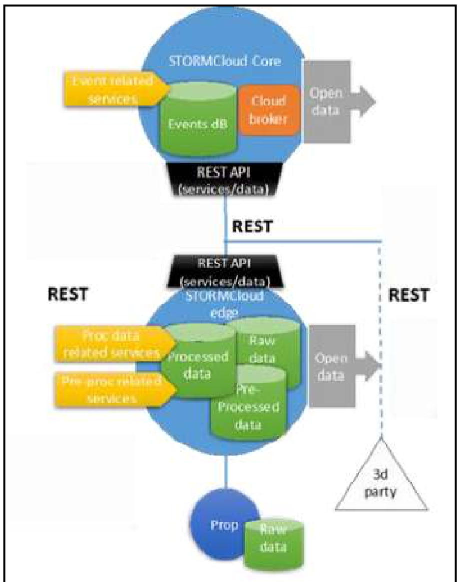

Fig. 3: STORM two level cloud architecture

Preliminary results of the proposed strategy are discussed.

Coming to the practical application of the work presented in this paper, the proposed methodology can be applied in the context of a two level architecture for the detection and mitigation of risks, as the one adopted in STORM project (Fig. 3). The architecture using a two level cloud based infrastructure can deploy the pre-processing of data and the corresponding algorithms for seismic detection, at an edge cloud level, providing results in the form of identified risk events at a second (core-cloud) level, where they can be assessed and also evaluated across other sources for decision making.

Acknowledgment

Work presented in this paper has received funding from the European Union's Horizon 2020 research and innovation Programme under STORM project, grant agreement n°700191.

References

- STORM Project web site, available at http://storm-project.eu, last accessed May 10, 2018.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning (Adaptive Computation and Machine Learning series). MIT Press. Ch. 10. pp. 373-422.
- Hayakawa, M., 2011. Probing the lower ionospheric perturbations associated with earthquakes by means of subionospheric VLF/LF propagation. Earthq. Sci. 24, 609–637, doi:10.1007/s11589-011-0823-1.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Computation 9(8), 735-1780, doi: 10.1162/neco.1997.9.8.1735.
- Ordóñez, F. J., Roggen, D., 2016. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. Sensors 16(1), 1-25, doi:10.3390/s16010115.
- Popova, I., Rozhnoi, A., Solieva, M., Levin, B., Hayakawa, M., Hobara, Y., Biagi., P. F., Schwingenschuh, K., 2013. Neural network approach to the prediction of seismic events based on low-frequency signal monitoring of the Kuril-Kamchatka and Japanese regions. Annals of Geophysics 56(3), R0328, doi:10.4401/ag-6224.
- Zeng, M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P. Juyong Zhang, J., 2014. Convolutional neural networks for human activity recognition using mobile sensors. In Proc. 6th IEEE Int. Conf. on Mobile Computing, Applications and Services (MobiCASE), 6-7 Nov. 2014, pp. 197-205, doi:10.4108/icst.mobicase.2014.257786.