13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13
Seoul, South Korea, May 26-30, 2019

Comparison Between Two Surrogate Models for Embankment Earthquake-Liquefaction-Induced Settlements Prediction.

Claudia Aristizábal
Postdoctoral Researcher, CentraleSupélec - Laboratoire MSSMat, Univ. Paris Saclay, Gif-Sur-Yvette, France

Fernando Lopez-Caballero
Associate Professor, CentraleSupélec - Laboratoire MSSMat, Univ. Paris Saclay, Gif-Sur-Yvette, France

ABSTRACT: The development of quick, easy-to-handle, surrogate models of complex simulations is a key issue to derive a rapid approximation of more expensive models. For instance, Lopez-Caballero and Khalil (2018), performed multiple finite element (FE) numerical simulations to assess the effect of the liquefaction-induced settlements of the soil foundation of an embankment due to real earthquakes. However, due to computational limitations of the FE analysis, they proposed a Gaussian Process (GP) emulator to represent the output of the more expensive FE model. Furthermore, they derived the analytical fragility curves constructed on the basis of the output of the nonlinear dynamic surrogate model. The aim of this work is to confront two different kinds of meta-models techniques, based on the same earthquake-liquefaction-induced settlements data. Hence, we proposed a comparison between two different surrogate models: (1) the GP model proposed by the authors and (2) an Artificial Neural Network (ANN) model proposed under the scope of this work. A comparison between the resultant fragility curves of the levee using both surrogate models is discussed, together with the impact of both meta-models in terms of fragility curves and its corresponding uncertainty. Finally, the main advantages and drawbacks of the surrogate models are highlighted.

Keywords: surrogate models, artificial neural networks, Gaussian process, fragility function, damage levels.

1. INTRODUCTION

The quick development of highly sophisticated computer codes to recreate the behavior of complex physical phenomenon is currently possible thanks to the availability of accessible, affordable and faster computational resources. Nowadays, state-of-the-art simulators are parameterize in a more elaborated way, with large number of input parameters describing the complexity of the model such as: material properties, initial conditions, boundary conditions, constitutive laws, among other. Nevertheless, and despite the fact that computational capacity of computers has increased exponentially in the past decades [Hilbert and López (2011)], there still exist computational limitations for plenty of applications. Where performing uncertainty quantification (UQ) tasks with Monte Carlo (MC) methods is almost always infeasible because of the need to perform hundreds of thousands or even millions of forward model evaluations in order to obtain convergent statistics [Tripathy and Bilionis (2018)]. Hence, the development of quick, easy-to-handle, surrogate models of complex simulations is a key issue to derive a rapid approximation of more expensive models. For instance, Lopez-Caballero and Khalil (2018), performed multiple finite ele-

ment (FE) numerical simulations to assess the effect of the liquefaction-induced settlements of the soil foundation of an embankment due to real earthquakes. However, due to computational limitations of the FE analysis to properly perform UQ analysis, a Gaussian Process (GP) emulator to represent the output of the more expensive FE model was proposed by the authors. Furthermore, the authors derived the fragility model of the embankment constructed on the basis of the output of the nonlinear dynamic surrogate model proving its use for uncertainty quantification. The aim of this work is to confront two different kinds of meta-models techniques, based on the same earthquake-liquefaction-induced settlements data. Hence, a comparison between two different surrogate models is discussed: (1) the GP model proposed by the authors and (2) an Artificial Neural Network (ANN) model proposed under the scope of this work. A comparison between the resultant fragility curves of the levee using both surrogate models is provided, together with the discussion on the impact of both meta-models in terms of fragility curves and its corresponding uncertainty. Finally, the main advantages and drawbacks of the surrogate models are highlighted .

2. METHODOLOGY

The methodology followed in this article is shown in Figure 1, and described in the following steps: (1) Select a robust earthquake database; (2) Propagate the different time histories (TH) through the soil media using the FEM model; (3) Train both surrogate models (GP and ANN) using the learning database (LDB) and the validation database (VDB). (4) Predict the crest settlements (δ_u/H[

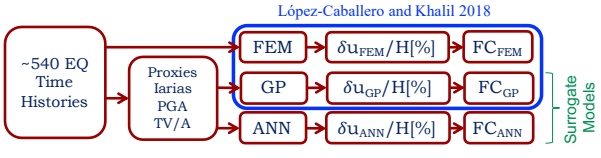

Figure 1: Methodology followed in this article.

2.1. Surrogate Models

A metamodel, an emulator or a surrogate model is an analytical function used to provide rapid approximations of a more expensive model (e.g. analitical model, finite-element model) [Lopez-Caballero and Khalil (2018)] while being computationally cheaper to evaluate . Even though there are multiple surrogate models that could be used for this purpose, for this case study a Gaussian Process Metamodel (GP) and an Artificial Neural Network (ANN) were used to predict the crest settlement of the levee based on predefined database containing the input an output parameters of the costly FEM model of the embankment.

2.1.1. Gaussian Process (GP)

A Gaussian Processes is an stochastic process (a collection of random variables indexed by time or space) that offer a mathematically funded and versatile framework to building statistical models. The main assumptions are: the phenomenon output is Gaussian and is a functional choice of covariance function (kernel). The statistical model needs physical knowledge: through data and expertise guiding the choice of kernel (which may come from the physical model). The statistical model is in essence complementary to the physical model and typically useful for decision making(e.g. optimization, uncertainty propagation, distributions shape, among others). Detailed Gaussian process references include Cressie (1993), Williams and Rasmussen (1996), Stein (1999), Rasmussen and Williams (2006).

2.2. Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) approach was inspired by investigations into the structure of the human brain, which consists of interconnected neurons McCulloch and Pitts (1943). It consist of a mathematical process that uses different layers to deal with the information its fed with. ANN's learn from examples and capture subtle functional relationships among the data even if the underlying relationships are unknown or hard to describe with more physical models. Thus ANNs are well suited for problems whose solutions require knowledge that is difficult to specify but for which there are enough data or observations [Zhang et al. (1998)]. The simplest ANN is composed of an input, a hidden and an output layers, but they

can become as complex of any order by adding additional hidden layers, which will turn the problem into what its better know as deep learning [Schmidhuber (2015)]. Currently, there exist countless of excellent ANN references and Zhang et al. (1998) provides an excellent overview of the ANN technique. On the other hand, Kong et al. (2018) provides an overview of current applications of machine learning (ML) in seismology, while Paolucci et al. (2018) presents an example of application for earthquake engineering purposes.

3. CASE STUDY

The selected case study consists of an embankment of 9 m height, composed of dry dense sand, founded over a 10 m layers of loose and dense sand overlaying bedrock, as shown in Figure 2. The first top 4 m of foundation consists of liquefiable loose-to-medium sand on top of 6 m of saturated dense sand. The underlying bedrock at the bottom of the dense sand has a mass density (ρ_{bd}) of 2,000 kg/m^3 and a shear wave velocity (Vs) of 1,000 m/s. The water table is located at 1 m below the foot of the embankment and it was kept dry. A 1:3 (vert:horiz) slope characterize the embankment inclination, and the geometry used on the FEM model is based on the one proposed by Rapti et al. (2018).

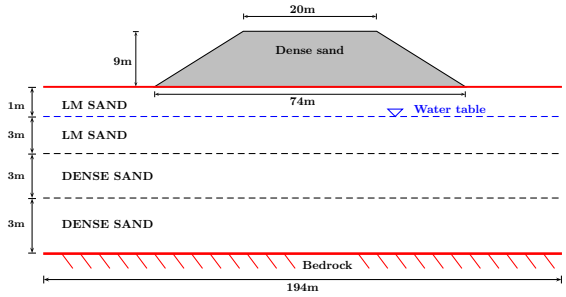

Figure 2: Embankment geometry and soil behavior used in the numerical FEM model (after Lopez-Caballero and Khalil (2018)).

All specific details related to the FEM analysis and the elastoplastic multimechanism of the soil used to represent the soil behavior in Figure 2 are extensively described in Lopez-Caballero and Khalil (2018) and will no further be discussed in this work.

4. EARTHQUAKE INPUT MOTIONS

Under the scope of this work the same case study used by Lopez-Caballero and Khalil (2018) is considered, hence, the same input ground motions used to build the surrogate model are also considered. A total of 540 unscaled records were selected by the authors from the PEER database [Ancheta et al. (2013)], the Center for Engineering Strong Motion Data, and the Kiban Kyoshin strong-motion network (KIK-NET) [Aoi and Okada (2000)]. The events range between 5.2 and 7.6 in moment magnitude (Mw). The recordings have site-to-source distances from 15 to 50 km and concern dense-to-firm soil conditions (i.e., 360 m/s $< V_{S30} <$ 800 m/s). All input signals have a baseline correction and a sampling time (Δt) equal to 0.005 s.

Table 1: Correlation matrix of the input and output variables for the 95 accelerograms used as the learning database (LDB).

CORR	IArias	PGA	TV/A	$\delta_u/H[\%]$
IArias	1.00	0.83	0.11	0.85
PGA	0.83	1.00	-0.01	0.76
TV/A	0.11	-0.01	1.00	0.27
$\delta_u/H[\%]$	0.85	0.76	0.27	1.00

The signals were filtered using a noncausal fourth-order Butterworth bandpass filter (i.e., zero-phase digital filtering), between 0.1 and 25.0 Hz. The correlation plots between the embankment crest settlement, $ln(\delta_u/H[\%])$, and the three input variables, X_i, are shown in Figure 3 and its corresponding correlations coefficients are shown in Table 1.

5. LEARNING, VALIDATION AND TEST DATABASES.

The earthquake database was split as follows: 95 signals concern the learning database (LDB), 50 ground motions are used for the validation set (VDB), and the test database (TDB) is composed of 395 unscaled records. The ranges of the input earthquake features obtained for each database (LDB, VDB and TDB) are summarized in Table 2. The earthquake features used to derive the surrogate models are: maximal outcropping acceleration (*PGA*), Arias Intensity (*Iarias*) and period of

Table 2: Statistics of the different input parameters for each database (LDB, VDB and TDB).

Index	Variable	Definition	Units	LDB (95)	VDB (50)	TDB (395)
X_1	IArias	Arias intensity.	m/s	0.001 - 20.64	0.04 - 4.13	0.0036 - 20.64
X_2	PGA	Peak ground acceleration.	g	0.01 -1.93	0.03 - 1.16	0.03 - 1.93
X_3	TV/A	Period of equivalent harmonic wave.	s	0.09 - 1.91	0.13 - 1.42	0.13 - 1.43

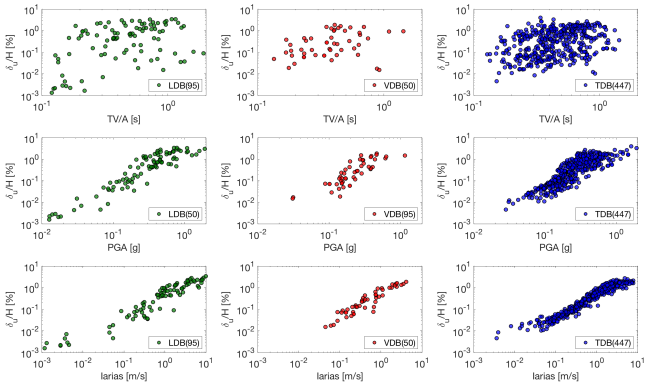

Figure 3: Correlation plots for the three different databases, LDB (green), VDB (red), TDB (blue), between the input variables (IArias, PGA and TV/A) and the output variable ($\delta_u/H[\%]$) used to derived the GP and ANN surrogate models.

equivalent harmonic wave ($T_{V/A} = \alpha \cdot$ PGV/PGA with α=4.89). Figure 4 displays the FEM crest settlement ($\delta u_{FEM}/H[\%]$) versus the input features (*PGA*, *Iarias* and $T_{V/A}$) plots for the learning (LDB) and validation (VDB) databases, showing similar correlation and variability as the one discussed in Figure 3. The VDB was selected in such a way that it covers a large range of the LDB and avoids any extrapolation from it. Finally, the TDB was also selected in a way that no extrapolation was allowed and with a considerable amount of events (395 events) to perform uncertainty quantification.

6. SURROGATE MODELS TRAINING

Once the LDB, VDB and TDB had been properly selected. The next step consist on training the two different surrogate models by using the LDB to calibrate the models, and the VDB to validate the results. The validation phase consist on comparing the predicted crest settlements of the surrogate models with those of the FEM model, once an acceptable accuracy is obtained after multiple sensibility analysis and optimization procedures, it is possible to proceed with the predictive phase. On the predictive phase, the TDB is used to predict quick approximation of the FEM ($\delta u_{FEM}/H[\%]$) by using the surrogate models instead ($\delta u_{GP,ANN}/H[\%]$).

6.1. *Gaussian Process (GP)*

As mentioned previously, the aim of this work is to confront two different kinds of meta-models techniques, based on the same input and output data. Nevertheless, the metamodel using a GP for this case study has been widely discussed in Lopez-Caballero and Khalil (2018),. Hence, the methodology concerning the GP will not longer be discussed, only the results provided by the authors will be used for comparison purposes.

6.2. *Artificial Neural Network (ANN)*

The most significant contribution of this work is the use of an artificial neural network (ANN) to predict the crest settlements of an embankment in order to to replace the most costly FEM model.

The ANN used in the present study is the feed-forward fully conected neural network as proposed by Hwang and Hu (2001).The input layer contains the external information or input variables, which for this case study are: maximal outcropping acceleration (*PGA*), Arias Intensity (*Iarias*) and period of equivalent harmonic wave ($T_{V/A}$). The number of hidden layers and neurons was finally set after an optimization process. Figure 5 shows the simplest architecture found under the scope of this work: which consists of an input layer, a single hidden layer of 7 neurons and an output layer.

The best architecture was selected trough a sensitivity analysis where the different parameters of the ANN (e.g. learning rate, activation function, number of hidden layers, number of hidden neurons) were modified in order to maximizes the squared correlation coefficient (Q^2)[Eq. 1] and minimizes

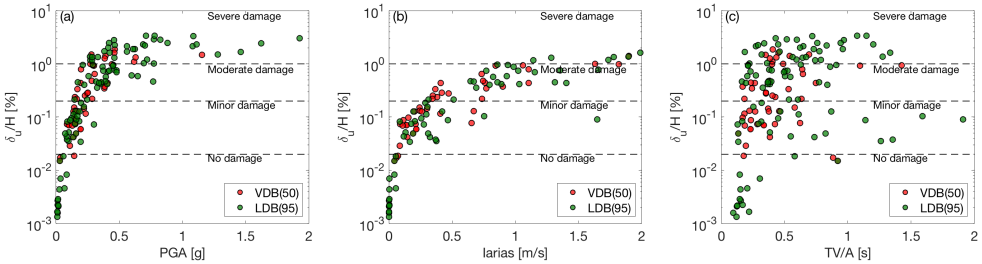

Figure 4: FEM crest settlement ($\delta_u/H[\%]$) versus input features (X_i) plots of the learning (LDB in blue) and validation (VDB in red) database: (a) $X_i = PGA$; (b) $X_i = Iarias$; (c) $X_i = TV/A$;

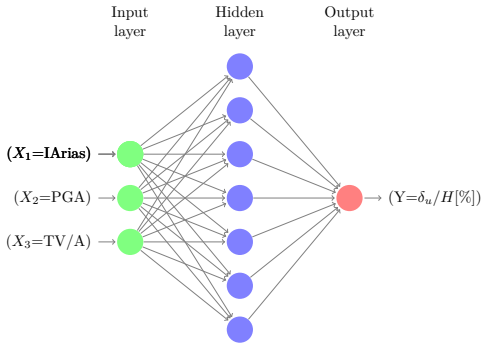

Figure 5: Final Artificial Neural Network architecture after the optimization process used as surrogate model to predict the crest settlements (δ_{uFEM}/H) of the embankment.

the root mean square error ($RMSE$)[Eq. 2] simultaneously.

$$Q^2(y_i, y_i^{pred}) = 1 - \frac{\sum_{n=1}^{N}(y_i^{pred} - y_i)^2}{\sum_{n=1}^{N}(y_i - \mu_y)^2} \quad (1)$$

$$RMSE(y_i, y_i^{pred}) = \sqrt{\sum_{n=1}^{N}\frac{(y_i^{pred} - y_i)^2}{N}} \quad (2)$$

Once the ANN architecture is defined, then the intrinsically random variability of the regression process of the ANN is studied by training 300 ANN with the exact same architecture and input and output learning and validation databases (LDB and VDB). This uncertainty is related to the convergence process of the ANN due to the feed-forward propagation process that it uses to correct the network until the ANN outputs approach the FEM outputs, nevertheless multiple solutions will converge, but with different levels of accuracy. This sensibility analysis was performed based on the resemblance of the fragility curves of the surrogate models with respect to those of the FEM model. The fragility curves represent the probability of failure of a system, associated with a specified criterion, for a given intensity measure (IM) of the earthquake motion. Failure herein represents the exceedence of a certain limit of the crest settlement (δ_{uFEM}/H), and the IM in this case is the peak ground acceleration (PGA). Thus, the fragility function can be expressed as follows [Eq. 3][Mai et al. (2017)]:

$$Frag(IM; \delta_o) = P[\Delta \geq \delta_o | IM] \quad (3)$$

$$\widehat{Frag}(IM; \delta_o) = \Phi\left(\frac{lnIM - ln\alpha}{\beta}\right) \quad (4)$$

The classical approach for establishing fragility curves consists in assuming a log-normal shape for the curves described in Eq. 3. This expression can be written in more general form by using the maximum likelihood approach Eq. 4. Where $\Phi(\cdot)$ denotes the standard Gaussian cumulative distribution function (CDF), α is the "median" and β is the "log-standard deviation" of the log-normal curve.

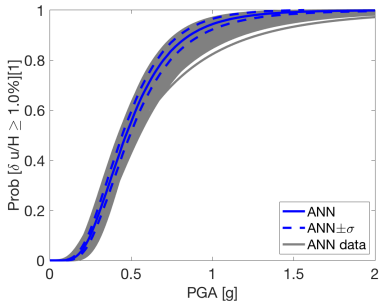

Figure 6: Fragility curves of the embankment for severe damage level ($Prob[\delta_u/H \geq 1.0\%]$), for the 300 ANN models and its $\mu \pm \sigma$ (blue).

Now, following this definition, Figure 6 displays the 300 fragility curves built using the results of

each one of the 300 trained ANN sets together with its median (μ) and its corresponding standard deviation (σ). Considering the uncertainty of the regression procedure is important, since it is possible that if only a single train of the architecture is made, even tough the solution converges, the solution could fall on the lower or upper bounds of the 300 fragility curves, affecting considerably the accuracy of the ANN surrogate model. Finally, the closest ANN to the median (μ) of the 300 models is selected as the final surrogate model: in this particular case model 234.

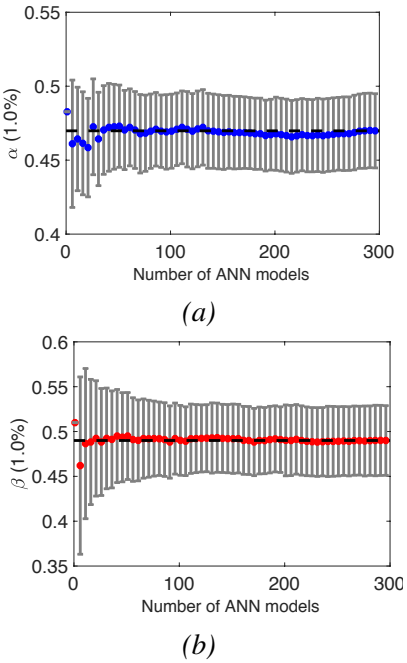

Figure 7: Fragility curve regression parameters derived using the maximum likelihood approach for the severe damage limit state (Prob$[\delta_u/H \geq 1.0\%]$): (a) α (median) and (b) β (standard deviation).

Once defined the median (μ) fragility curve and its corresponding ANN model (model 234), it is important to verify if the the number of ANN training sets selected to evaluate the uncertainty: here arbitrarily selected as 300; is sufficient to obtain robust statistics. Figure 7 a,b show the median ($\mu_{\alpha,\beta}$) and standard deviation ($\sigma_{\alpha,\beta}$) of the α and β parameters, respectively, when considering 1 to 300 models. As it is possible to observe in both in both figures, above 100 ANN model, neither the median ($\mu_{\alpha,\beta}$) nor standard deviation ($\sigma_{\alpha,\beta}$) vary as a function of the number of considered fragility curves. Hence in this particular case, the 300 considered models were adequately to predict robust statistics, nevertheless this conclusion is structure dependent, hence, this validation needs to be performed for each structure and for the different considered damage level.

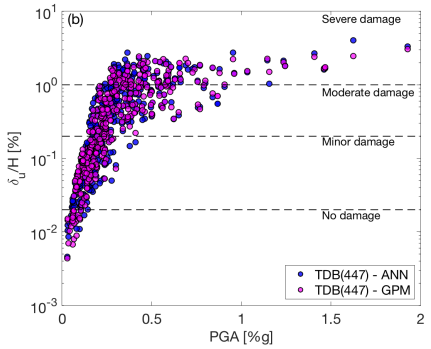

Figure 8: Crest settlement predicted using the VDB and the TDB databases for the GP model, $\delta_{uGP}/H[\%]$, (magenta) and the ANN model, $\delta_{uANN}/H[\%]$, (blue).

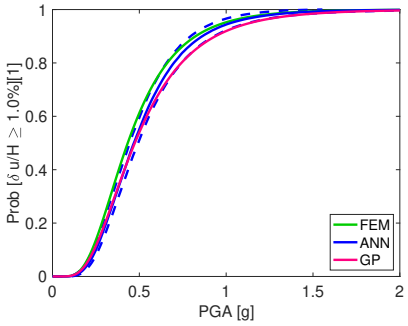

Figure 9: Fragility curves: FEM (green); ANN ($\mu \pm \sigma$) (blue); GP (magenta).

7. SURROGATE MODELS PREDICTIONS

The next step consist on predicting the crest settlements ($\delta_u/H[\%]$) by using the test database (TDB) composed of 395 unscaled records and the validation database (VDB) composed of 50 records. Figure 8 shows the predicted ($\delta_u/H[\%]$) by the ANN and the GP with respect PGA for the combined databases.

The three damage level thresholds are superposed in the same plot and they correspond to $\delta_u/H[\%]$=0.02, 0.2, and 1.0%. The damage state limits or the performance levels of the levee are those proposed by Swaisgood (2003). From visual inspection, both models covered approximately the

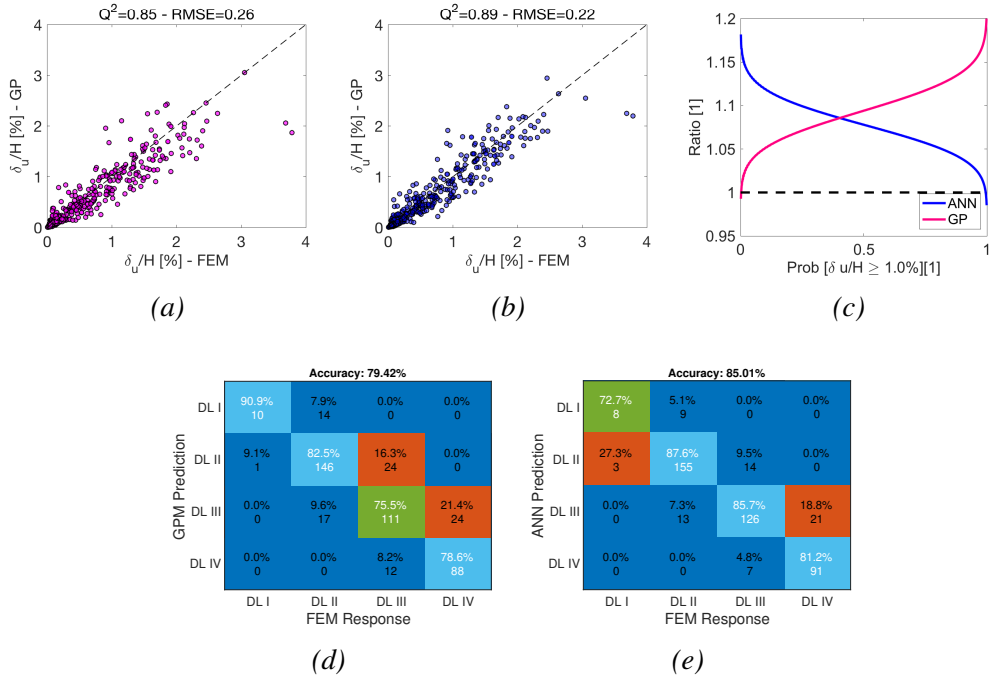

Figure 10: (a) FEM Vs. GP crest settlements ($\delta_u/H[\%]$); (b) FEM Vs. ANN crest settlements ($\delta_u/H[\%]$)(blue); (c) Surrogate models and the FEM model fragility curves ratio (GP or ANN/FEM); Confusion matrices between the FEM and : (d) GP emulator; (e) ANN emulator.

same surface of the plot and with similar density distributions of sample points. Nevertheless in the next section a quantitative comparison between models is presented.

8. SURROGATE MODELS COMPARISON

The fragility curves of the FEM, GP and ANN crest settlements ($\delta_{uGP}/H[\%]$) for the severe damage level limits state ($Prob[\delta_u/H \geq 1.0\%]$) are shown in Figure 9. The goodness of the final surrogate models with respect to the FEM model has been quantitatively estimated in three different ways: (1) Comparison of ($\delta_u/H[\%]$) values obtained with FEM and with the surrogate models, Figure 10a,b; (2) Ratio between the surrogate models fragility curves and the FEM model, Figure 10c. (3) Comparison of the damage limit states confusion matrix obtained considering FEM with respect to the surrogate models, GP and ANN, respectively Figure 10d,e. Firstly, the regression plot of the FEM Vs. GP/ANN Figure 10a,b allows to calculate the squared correlation coefficient (Q^2) and the root mean square error ($RMSE$)of the final surrogate models, where a $Q^2 = 1.0$ and an $RMS = 0$ will represent the perfect model. Secondly, the ration between fragility curves, Figure 10c, allows to estimate if the surrogate models are over estimating or underestimating the real model: for this particular case, an overestimation of the fragility curve is preferable since is on the safety side. For this case example, for a given probability of exceedance ($Prob[\delta_u/H \geq X]$), both surrogate models overestimates the IM (PGA) up to a value of 20% with respect to the FEM model. Nevertheless, the ANN tends to predict better results at larger probabilities, while the GP does it at small probabilities, being particularly preferable the ANN, since for risk purposes it is more important to properly predict the larger return periods. Despite this particular finding, that could motivate future users to prefer the ANN over the GP models, no general conclusion can be drawn from this particular exercise. Finally, the confusion matrices are a very powerful tool to evaluate the accuracy of surrogate models, Figure 10d,e. For instance, the ANN performs slightly better in terms of accuracy than the GP: 85% accuracy for the former, against 79% for the later. Also, even though both models generally tend to overestimate damage levels III and IV, the ANN overestimates less than the GP model, hence, with a better resemblance to the FEM model than the GP model. But again, no general conclusion about which model is better than the other can be drawn based on this particular case example.

9. CONCLUSIONS

Two surrogate models (GP and ANN) were used as alternative to predict the liquefaction-induced

settlements of a levee initially modeled with a more expensive FEM. The comparative results shown here have been evaluated in terms of the fragility functions, damage levels, and crest settlement predictions. The main conclusions drawn from this study are: (1) The ANN slightly predicts better results than the GP. Nevertheless, both the GP and the ANN surrogate models have shown to be good tools to predict accurately enough the nonlinear FEM response. Both surrogate models represent an economy in CPU time with respect to the FEM models proving their value for UQ purposes. Finally, it is remarkable, that with only 3 input proxies of complex waveforms it was possible to retrieve a very close equivalent model with almost zero computation cost once the ANN or GP models are trained.

10. ACKNOWLEDGMENTS

This work has been funded by the French SINAPS@ project (Earthquake and Nuclear Facilities : Ensuring Safety and Sustaining).

11. REFERENCES

Ancheta, T. D., Darragh, R. B., Stewart, J. P., Seyhan, E., Silva, W. J., Chiou, B. S., Wooddell, K. E., Graves, R. W., Kottke, A. R., Boore, D. M., et al. (2013). "Peer nga-west2 database.

Aoi, S., O. K. H. S. K. K. and Okada, Y. (2000). "New japanese uphole-downhole strong-motion observation network: Kik-net.." *Seismol. Res. Lett.*, 72, 239.

Cressie, N. A. (1993). "Statistics for spatial data: Wiley series in probability and mathematical statistics." *Find this article online*.

Hilbert, M. and López, P. (2011). "The world's technological capacity to store, communicate, and compute information." *science*, 1200970.

Hwang, J.-N. and Hu, Y. H. (2001). *Handbook of neural network signal processing*. CRC press.

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., and Gerstoft, P. (2018). "Machine learning in seismology: Turning data into insights." *Seismological Research Letters*.

Lopez-Caballero, F. and Khalil, C. (2018). "Vulnerability assessment for earthquake liquefaction–induced settlements of an embankment using gaussian processes." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 4(2), 04018010.

Mai, C., Konakli, K., and Sudret, B. (2017). "Seismic fragility curves for structures using non-parametric representations." *Frontiers of Structural and Civil Engineering*, 11(2), 169–186.

McCulloch, W. S. and Pitts, W. (1943). "A logical calculus of the ideas immanent in nervous activity." *The bulletin of mathematical biophysics*, 5(4), 115–133.

Paolucci, R., Gatti, F., Infantino, M., Smerzini, C., Güney Özcebe, A., and Stupazzini, M. (2018). "Broadband ground motions from 3d physics-based numerical simulations using artificial neural networks." *Bulletin of the Seismological Society of America*, 108(3A), 1272–1286.

Rapti, I., Lopez-Caballero, F., Modaressi-Farahmand-Razavi, A., Foucault, A., and Voldoire, F. (2018). "Liquefaction analysis and damage evaluation of embankment-type structures." *Acta Geotechnica*, 1–19.

Rasmussen, C. E. and Williams, C. K. (2006). "Gaussian processes for machine learning. 2006." *The MIT Press, Cambridge, MA, USA*, 38, 715–719.

Schmidhuber, J. (2015). "Deep learning in neural networks: An overview." *Neural networks*, 61, 85–117.

Stein, M. (1999). "Interpolation of spatial data: Some theory for kriging springer-verlag new york google scholar.

Swaisgood, J. (2003). "Embankment dam deformations caused by earthquakes." *7th Pacific conference on earthquake engineering*, New Zealand Society for Earthquake Engineering, Wellington, New Zealand.

Tripathy, R. and Bilionis, I. (2018). "Deep uq: Learning deep neural network surrogate models for high dimensional uncertainty quantification." *arXiv preprint arXiv:1802.00850*.

Williams, C. K. and Rasmussen, C. E. (1996). "Gaussian processes for regression." *Advances in neural information processing systems*, 514–520.

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