Estimating the Maximum Shear Modulus with Neural Networks

Manuel Cruz[1], Jorge M. Santos[1], Nuno Cruz[2]

[1] DMA - School of Engineering, Polytechnic of Porto and LEMA, Portugal
[2] Mota-Engil and Universidade de Aveiro, Portugal
emails: mbc**, jms (@isep.ipp.pt), nbdfcruz@gmail.com

Abstract. Small strain shear modulus is one of the most important geotechnical parameters to characterize soil stiffness. In-situ stiffness of soils and rocks is much higher than was previously thought as finite element analysis have shown. Also, the stress-strain behaviour of those materials is non-linear in most cases with small strain levels. The commun approach for getting the small strain shear modulus is usually based on measure of seismic wave velocities. Nevertheless, for design purposes is very useful to derive that modulus from correlations with in-situ tests output parameters. In this view, the use of Neural Networks seems very appropriate as the complexity of the system keeps the problem very unfriendly to treat following traditional data analysis methodologies. In this work, the use of Neural Networks is proposed to estimate small strain shear modulus for sedimentary soils from the basic or intermediate parameters derived from Marchetti Dilatometer Test.

1 Introduction

Maximum shear modulus, G_0, is nowadays a key geotechnical parameter in soil stiffness evaluation. The standard way to measure it is to evaluate compression and shear wave velocities and thus obtain results supported by theoretical interpretations. Despite the advantages appointed by the scientific community (e.g. [1, 2]), this approach has a drawback that is mainly appointed by the industrial counterpart: the use of seismic measures implies a specific and more expensive test than the ones in old-fashioned way. As a result, many authors have dedicated their efforts to correlate other in-situ test parameters with G_0. Among others, the works from Peck, Lunne, Marchetti or Cruz do it for the Standard Penetration Test (SPT) [3], Piezocone Test (CPTu) [4] or Marchetti Dilatometer Test (DMT) [5–7].

In this context, the DMT seems a very appropriate equipment to accomplish that task with success. That may be explained as follows:

1. DMT measure a load range related with a specific displacement (E_D)
2. E_D may be used to deduce highly accurate stress-strain relationship, supported by the Theory of Elasticity

** The author was partially supported by both DMA and LEMA.

3. The type of soil can be numerically represented by DMT Material Index, I_D
4. The in situ density, overconsolidation ratio (OCR) and cementation influences can be represented by lateral stress index, K_D

which allows for high quality calibration of the stress-strain relationship [7].

In this paper, an estimation of G_0 derived from the DMT basic and intermediate parameters using neural networks is presented.

2 G_0 prediction by DMT

Marchetti dilatometer test, commonly designated by DMT, has been increasingly used and it is one of the most versatile tools for soil characterization. The test was developed by Silvano Marchetti [5] and can be seen as a combination of both Piezocone and Pressuremeter tests with some details that really makes it a very interesting test available for modern geotechnical characterization [7]. The main reasons for its usefulness on deriving geotechnical parameters are related to the simplicity and the speed of execution generating quasi-continuous data profiles with high accuracy and reproducibility.

In its essence, dilatometer is a stainless steel flat blade with a flexible steel membrane in one of its faces. The blade is connected to a control unit on the ground surface by a pneumatic-electrical cable that goes inside the position rods, ensuring electric continuity and the transmission of the gas pressure required to expand the membrane. The equipment is pushed (most preferable) or driven into the ground, by means of a CPTu rig or similar, and the expansion test is performed every 20cm. The (basic) pressures required for lift-off the diaphragm (P_0), to deflect 1.1mm the centre of the membrane (P_1) and at which the diaphragm returns to its initial position (P_2 or closing pressure) are recorded. Due to the balance of zero pressure measurement method (null method), DMT readings are highly accurate even in extremely soft soils, and at the same time the blade is robust enough to penetrate soft rock or gravel. The test is found especially suitable for sands, silts and clays.

Four intermediate parameters, Material Index (I_D), Dilatometer Modulus (E_D), Horizontal Stress Index (K_D) and Pore Pressure Index (U_D), are deduced from the basic pressures P_0, P_1 and P_2, having some recognizable physical meaning and some engineering usefulness [5], as it will be discussed below. The deduction of current geotechnical soil parameters is obtained from these intermediate parameters covering a wide range of possibilities. In the context of the present work, besides the basic pressures, only E_D, I_D and K_D have a physical meaning on the determination of G_0, so they will be succinctly described as follows [7]:

1. Material Index, I_D: Marchetti [5] defined Material Index, I_D, as the difference between P_1 and P_0 basic measured pressures normalized in terms of the effective lift-off pressure. In a simple form, it could be said that I_D is a "fine-content-influence meter"[7], providing the interesting possibility of defining dominant behaviours in mixed soils.

2. Horizontal Stress Index, K_D: The horizontal stress index [5] was defined to be comparable to the at rest earth pressure coefficient, K_0, and thus its determination is obtained by the effective lift-off pressure (P_0) normalized by the in-situ effective vertical stress. K_D is a very versatile parameter since it provides the basis to assess several soil parameters such as those related with state of stress, stress history and strength, and shows dependency on several factors namely cementation and ageing, relative density, stress cycles and natural overconsolidation resulting from superficial removal, among others.
3. Dilatometer Modulus, E_D: Stiffness behaviour of soils is generally represented by soil moduli, and thus the base for in-situ data reduction. Generally speaking, soil moduli depend on stress history, stress and strain levels drainage conditions and stress paths. The more commonly used moduli are constrained modulus (M), drained and undrained compressive Young modulus (E_0 and E_u) and small-strain shear modulus (G_0), this latter being assumed as purely elastic and associated to dynamic low energy loading.

Maximum shear modulus, G_0, is indicated by several investigators [2, 7, 10] as the fundamental parameter of the ground. It can be accurately deduced through shear wave velocities,

$$G_0 = \rho v_s^2 \tag{1}$$

where ρ stands for density and v_s for shear wave velocity.

However, the use of a specific seismic test imply an extra cost, since it can only supply this geotechnical parameter, leaving strength and insitu state of stress information dependent on other tests. Therefore, several attempts to model the maximum shear modulus as a function of DMT intermediate parameters for sedimentary soils have been made in the last decade. Hryciw [11] proposed a methodology for all types of sedimentary soils, developed from indirect method of Hardin & Blandford [12]. This methodology ignores dilatometer modulus, E_D, commonly recognized as a highly accurate stress-strain evaluation, and also lateral stress index, K_D, and material index, I_D, which are the main reasons for the accuracy in stiffness evaluation offered by DMT tests [6]. Being so, the most common approaches [13–15] with reasonable results concentrated in correlating directly G_0 with E_D or M_{DMT} (constrained modulus), which have revealed linear correlations with slopes controlled by the type of soil. In 2006, Cruz [6] proposed a generalization of this approach, trying to model the ratio $R_G \equiv \frac{G_0}{E_D}$ as a function of I_D. In 2008, Marchetti [16] using the commonly accepted fact that maximum shear modulus is influenced by initial density and considering that this is well represented by K_D, studied the evolution of both R_G and G_0/M_{DMT} with K_D and found different but parallel trends as function of type of soil (that is I_D), recommending the second ratio to be used in deriving G_0 from DMT, as consequence of a lower scatter. In 2010, using the Theory of Elasticity, Cruz [7] approximate G_0 as a non-linear function of I_D, E_D and K_D, from where a promising median of relative errors close to 0.21 with a mean(standard deviation) around 0.29(0.28) were obtained. It is worth mention that comparing with the previous approach - R_G - this approximation, using the same data, lowered the mean and median of relative errors in more than 0.05 maintaining the standard deviation (Table 2).

In this work, to infer about the results quality it will be used some of the same indicators used by Hryciw, Cruz and others that are: the median, the arithmetic mean and standard deviation of the relative errors

$$\delta^{i}_{\widetilde{G_0}} = \frac{|\widetilde{G_0}(i) - G_0(i)|}{|G_0(i)|}; i = 1, 2, ..., N \tag{2}$$

where $\widetilde{G_0}(i)$ stands for the predicted value and $G_0(i)$ for the measured value given by seismic wave velocities (which is assumed to be correct). A final remark to point out that since in this work the no-intercept regression is sometimes used, the R^2 values will not be presented as they can been meaningful in this case [17]. It is also worth to remark that in the context of DMT and from the engineering point of view, median is the parameter of choice for assessing the model quality [7] since the final value for maximum shear modulus relies on all set of results obtained in each geotechnical unit or layer.

3 Data Sets, Experiments and Results

3.1 The WDS and PsS Data sets

In the forthcoming experiments there was used one subset of the WDS data set named PsS data set. The WDS data set was used in the development of the non-linear G_0 approximation done by Cruz in [7], resulting from 860 DMT measurements performed in Portugal by Cruz and world wide by Marchetti et al. [16] (data kindly granted by Marchetti for the work presented in [7]), which included data obtained in all kinds of sedimentary soils, namely clays, silty clays, clayey silts, silts, sandy silts, silty sands and sands. Afterwards was used again as base for the work by Cruz et al [8] where the DMT intermediate parameters are used to estimate G_0. Since the Marchetti data does not include the record of the basic parameters, P_0, P_1 and u_0, only the Portuguese subset (denoted by PsS) will be used when trying to predict G_0 from those parameters.

In order to have some comparisons between the present work and the one made in [8], the WDS main statistical measures with respect to I_D, E_D, K_D and G_0 parameters are given in Table 1 (in parenthesis the same measures for PsS). Figures 1 and 2, where data from WDS and PsS, respectively, is represented using MatLab function *plotmatrix*, aims a clear view of variables dispersion. This is important in a Geotechnical point of view as it may show the (very) different types of soil who serve has base to this work. It should be noted that in Figure 2 the additional parameters P_0, P_1 and u_0 are presented.

3.2 G_0 prediction by DMT Parameters: A NN approach

In addition to the work reviewed in Section 2, in 2011, Cruz *et al* [8] went a little further reported the fitting of G_0 through the DMT intermediate parameters

Values	I_D	E_D	K_D	G_0
min	0.05070 (0.05070)	0.3644 (0.3644)	0.9576 (0.9576)	6.430 (12.71)
max	8.814 (8.814)	94.26 (85.00)	24.61 (24.61)	529.2 (110.6)
median	0.5700 (0.2192)	13.44 (4.372)	3.575 (3.136)	77.91 (34.51)
mean	0.9134 (1.063)	18.83 (9.963)	4.916 (3.808)	92.52 (38.81)
std	1.074 (1.946)	18.83 (13.08)	3.608 (2.791)	69.61 (19.37)

Table 1. Sample WDS (PsS) statistical measures rounded to 4 significant digits.

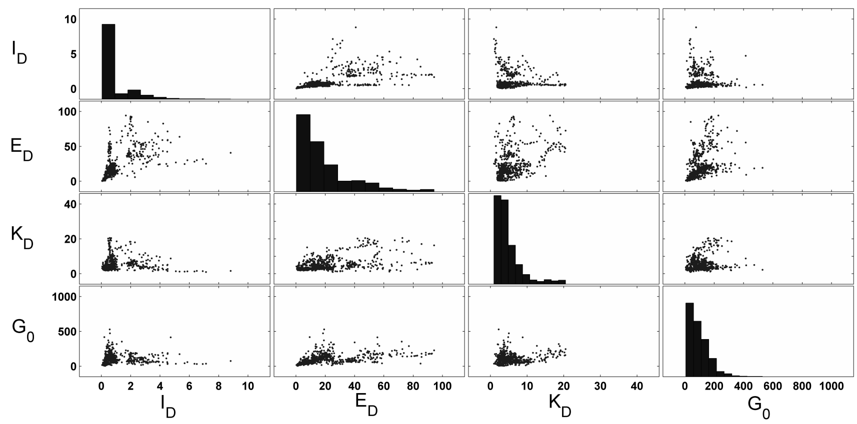

Fig. 1. Sample WDS: values for I_D,E_D,K_D and G_0

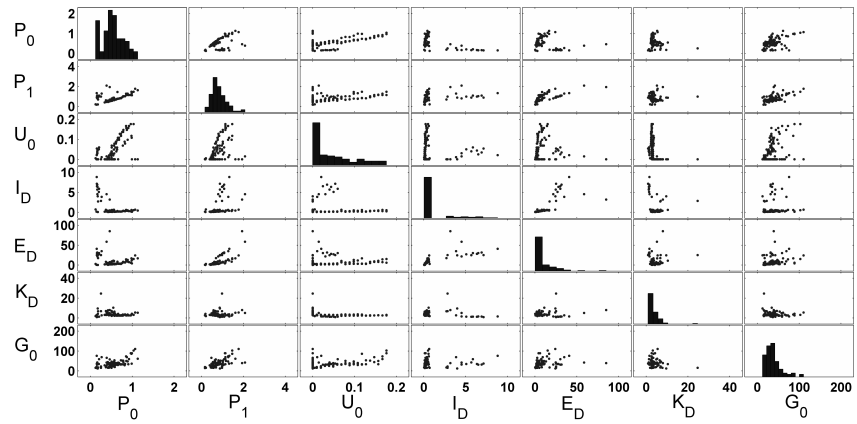

Fig. 2. Sample PsS: values for P_0,P_1,u_0,I_D,E_D,K_D and G_0

E_D, I_D and K_D based on the use of different types of Least Square Non-Linear Regression and Neural Networks (NN). Using the WDS dataset, an attempt to improve the quality of these results was carried out by using Support Vector Regression (SVR). Support Vector Machines [20] are based on the statistical learning theory from Vapnik and are specially suited for classification. However, there are also algorithms based in the same approach for regression problems known as Support Vector Regression. The performed experiments with SVRs were carried out using LIBSVM [21] for Matlab. Two different kinds of SVR algorithms: ϵ-SVR, from Vapnik [22] and ν-SVR from Schölkopf [23] were applied, which differ in the fact that ν-SVR uses an extra parameter $\nu \in (0,1]$ to control the number of support vectors. For these experiments a search for the best results was made in the C, ϵ (ν) space and so different values for the parameter C (cost) and for parameters ϵ and ν were used.

The best results obtained with both ϵ-SVR and ν-SVR with the radial basis function kernel reveal slightly better results when compared with those obtained with the fitting neural network and better than those obtained with the other MLP's and the traditional regression algorithms.

In order to have an easier reading of the present paper, a summary of the results achieved in [8] is presented in Tables 2 and 3.

	Type	Hidden neurons	Median/Mean(std)
Non-Linear Regression	$G_0 = \alpha\, E_D\, (I_D)^{\beta}$	-	0.28/0.34(0.29)
	$G_0 = E_D + E_D\, e^{(\alpha + \beta I_D + \gamma \log(K_D))}$	-	0.21/0.29(0.28)
Neural Networks	Quasinewton	50	0.20/0.38(0.72)
	Conj.Grad.	100	0.19/0.30(0.38)
	SCG	40	0.20/0.28(0.33)
	MLP-Bayesian	20	0.20/0.29(0.30)
	RBF	200	0.20/0.31(0.39)
	Fitting	60	0.17/0.27(0.29)

Table 2. Sample WDS: Relative Error Results (Median/Mean(std)) obtained with $\widetilde{G_0} = f(I_D, E_D, K_D)$.[8]

Type	Cost/$\epsilon(\nu)$	Median/Mean(std)
ϵ-SVR	200/0.1	0.16/0.27(0.43)
ν-SVR	200/0.8	0.16/0.27(0.41)

Table 3. Sample WDS: Relative Error Results (Median/Mean(std)) obtained with Support Vector Regression $\widetilde{G_0} = f(I_D, E_D, K_D)$.

Despite all the work reviewed in Section 2 it hasn't been already tried to model G_0 as a straightforward function of the DMT basic parameters P_0, P_1

and P_2. In addition, the promising results showed in Table 3 led the authors to go further and to try that approach. However, there are some difficulties in the interpretation of P_2 values, since it can represent very distinctive situations in different type of soils, as explained below:

- In sands the parameter can be roughly compared to the pore pressure resulting from the hydrostatic level, in equilibrium. In fact the pressure on the membrane is that of the water in the pores.
- In clays P_2 parameter represents a mixed of both water and soil pressures, and thus it should only be used qualitatively, as sustained by Marchetti [16].
- Furthermore, in soils with intermediate behaviours (silts, sandy clays or clayey sands) the problem is even worse than with clays creating some important problem for a reasonable interpretation [7].

As a consequence of these, it was considered more appropriate to work with equilibrium pore-pressures (u_0), calculated from the position of water level externally obtained, instead of P_2. Thus, in the next experiments the objective is to model G_0 as function of P_0, P_1 and u_0 parameters, avoiding the need for special interpretations, which turns to be much more efficient to include in mathematical operations. With the characterization of the PsS data set presented on Table 1 and Figure 2 it can be seen that this subset is comparable to the WDS in terms of variables distribution and limits in exception of the G_0 parameter where the available data is restricted to the range 12-110, where in WDS it goes 6-530. This is relevant, as the conclusions about this experiments must take this into account.

The straight application of the expressions calculated in [7] for the regression applied to this subset returned the relative error parameters shown in Table 4, and the recalculation of the regression constants and subsequent relative error evaluation lead to the results shown in Table 5. Comparing the variability of these results with the ones showed in Table 2 highlights the advantage of using cross validation in experiments.

Concerning the G_0 prediction using the (P_0,P_1,u_0) parameters, the schema was similar to the one described in the previous subsection for the intermediate parameters. A traditional regression approach was first used and then several Neural Network experiments were made. Two sets of input parameters were used: one using P_0, P_1 and u_0 and other neglecting u_0.

Regarding traditional regression, the least squares method returned some interesting results that can be seen on Table 6. Those results are the best when considering all the possible combinations of the transformations exponential, square root, logarithmic and square to the dependent and independent variables.

It should be noted that in Table 6, $\delta \approx 0.448$, which combined with the range of values for u_0 (roughly say, [0,0.2]) results on a multiplicative effect in the prediction $\widetilde{G_0}$ - that is $e^{\delta u_0} \times f(P_0, P_1)$ - of approximately [1,1.1]. Thus, it was expectable that the introduction of the u_0 parameter didn't bring too much improvement to our previous result as it happened.

For all the experiments using NN's or SVR's the 10 fold cross validation method with 20 repetitions was used, since this is the most common and widely

	Type	Median/Mean(std)
Non-Linear	$G_0 = \alpha\, E_D\, (I_D)^{\beta}$	0.32/0.55(0.63)
Regression	$G_0 = E_D + E_D\, e^{(\alpha\,+\,\beta I_D\,+\,\gamma\,\log(K_D))}$	0.34/0.50(0.49)

Table 4. Subset PsS: Relative Error Results (Median/Mean(std)) obtained with non-linear regression $\widetilde{G_0} = f(I_D, E_D, K_D)$ using the (α, β, γ) calculated in Table 2.

	Type	Median/Mean(std)
Non-Linear	$G_0 = \alpha\, E_D\, (I_D)^{\beta}$	0.26/0.34(0.33)
Regression	$G_0 = E_D + E_D\, e^{(\alpha\,+\,\beta I_D\,+\,\gamma\,\log(K_D))}$	0.14/0.18(0.16)

Table 5. Subset PsS: Relative Error Results (Median/Mean(std)) obtained with non-linear regression $\widetilde{G_0} = f(I_D, E_D, K_D)$ revaluating the (α, β, γ) parameters.

	Type	Median/Mean(std)
Non-Linear	$G_0 = \alpha\, e^{\beta P_0\,+\,\gamma P_1}$	0.22/0.28(0.23)
Regression	$G_0 = \alpha\, e^{\beta P_0\,+\,\gamma\sqrt{P_1}\,+\,\delta u_0}$	0.22/0.28(0.23)

Table 6. Subset PsS: Relative Error Results (Median/Mean(std)) obtained with non-linear regression $\widetilde{G_0} = f(P_0, P_1)$ and $\widetilde{G_0} = f(P_0, P_1, u_0)$.

accepted methodology to guarantee a good neural network generalization [19]. For each NN a huge set of experiments was performed, varying the involved parameters such as the number of neurons in the MLP hidden layer, the number of epochs or the minimum error for stopping criteria. The results here presented are therefore the best ones for each regression algorithm and represent the mean of the 10×20 performed tests for each best configuration. It is also important to stress the fact that, when compared to traditional approaches where all the data is used to build the model, this methodology tends to produce higher standard deviations since in each experiment only a fraction of the available data is used to evaluate the model. Several exploratory experiments were performed with different kinds of MLPs and SVRs. Results from these preliminary experiments show that the best ones were also obtained with SVRs with the radial basis function kernel and for that reason we focus on more detailed experiments using this combination. Results from the SVRs with radial basis function kernel are presented in Table 7, where the u_0 parameter also seem to be negligible in terms of G_0 prediction.

4 Conclusions

Figure 3 summarizes the results presented in the previous subsections and represents the quality parameters of some of the best results on the estimation of G_0 via DMT's basic and intermediate parameters.

Input	Type	Cost/$\epsilon(\nu)$	Median/Mean(std)
(P_0, P_1)	ϵ-SVR	40/0.0001	0.24/0.29(0.22)
(P_0, P_1)	ν-SVR	20/0.9	0.25/0.31(0.25)
(P_0, P_1, u_0)	ϵ-SVR	40/0.0001	0.24/0.29(0.22)
(P_0, P_1, u_0)	ν-SVR	20/0.9	0.25/0.31(0.25)

Table 7. Sample PsS: Relative Error Results (Median/Mean(std)) obtained with SVR's.

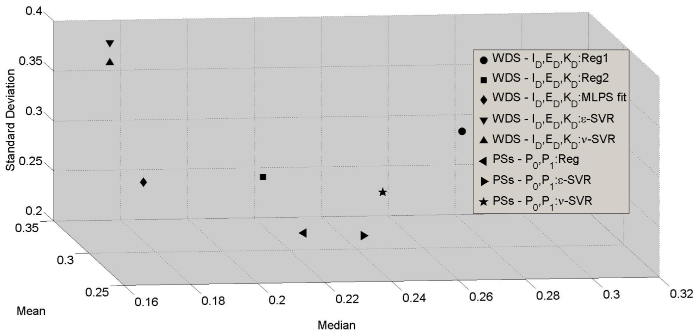

Fig. 3. Best Results: values for median, mean and std of $\frac{|G_0-\widetilde{G_0}|}{G_0}$

This emphasizes the good results of applying Neural Networks to predict maximum shear modulus by DMT. Based on performed experiments it is possible to outline the following considerations:

- Neural Networks and/or SVR's improve the state-of-the-art in terms of G_0 prediction. The results show that, in general, NNs and/or SVR's lead us to much smaller medians, equivalent means and higher standard deviations in respect to relative errors, when compared to traditional approaches.
- Regarding the problem characteristics the SVR approach gives, on the prediction with DMT intermediate parameters, the best results considering the median as the main quality measure as discussed earlier.
- When compared with the intermediate parameters, the results show that the basic input parameters (P_0, P_1) does not improve the fitness of G_0.
- In addition to the previous sentence, the inclusion of u_0 as third input parameter does not seem to improve the fitness. Future work should consider other auxiliar data, mainly measured depth, depth of water level, and/or P_2.
- The available unbalanced data, regarding G_0 distribution, suggests that more tests should be made using G_0 values of higher magnitude (>110).

References

1. C. Clayton and G. Heymann. Stiffness of geomaterials at very small strains. *Géotechnique*, 51(3):245–255, 2001.

2. M. Fahey. Soil stiffness values for foundation settlement analysis. In *Proc. 2nd Int. Conf. on Pre-failure Deformation Characteristics of Geomaterials*, volume 2, pages 1325–1332. Balkema, Lisse, 2001.
3. Ralph B. Peck, Walter E. Hanson, and Thomas H. Thornburn. Foundation Engineering (2nd Edition). John Wiley & Sons, 1974.
4. T. Lunne, P. Robertson and J. Powell. Cone penetration testing in geotechnical practice. Spon E & F N, 1997.
5. S. Marchetti. In-situ tests by flat dilatometer. *Journal of the Geotechn. Engineering Division*, 106(GT3):299–321, 1980.
6. N. Cruz, M. Devincenzi, and A. Viana da Fonseca. Dmt experience in iberian transported soils. In *Proc. 2nd International Flat Dilatometer Conference*, pages 198–204, 2006.
7. N. Cruz. Modelling geomechanics of residual soils by DMT tests. In *PhD thesis, Universidade do Porto, 2010.*
8. M. Cruz, J.M. Santos, and N. Cruz. Maximum Shear Modulus Prediction by Marchetti Dilatometer Test Using Neural Networks. In *Engineering Applications of Neural Networks, volume 363, Springer, 335-344, Eds: Iliadis, Lazaros, and Jayne, Chrisina, 2011.*
9. S. Marchetti. The flat dilatometer: Design applications. In *Third Geotechnical Engineering. Conf. Cairo University*, 1997.
10. P. W. Mayne. Interrelationships of dmt and cpt in soft clays. In *Proc. 2nd International Flat Dilatometer Conference*, pages 231–236, 2006.
11. R.D. Hryciw. Small-strain-shear modulus of soil by dilatometer. *Journal of Geotechnical Eng. ASCE*, 116(11):1700–1716, 1990.
12. B.O. Hardin and G.E. Blandford. Elasticity of particulate materials. *J. Geot. Eng. Div.*, 115(GT6):788–805, 1989.
13. B.M. Jamiolkowski, C.C. Ladd, J.T. Jermaine, and R. Lancelota. New developments in field and laboratory testing of soilsladd, c.c. In *XI ISCMFE*, volume 1, pages 57–153, 1985.
14. J.P. Sully and R.G. Campanella. Correlation of maximum shear modulus with dmt test results in sand. In *Proc. XII ICSMFE*, pages 339–343, 1989.
15. H. Tanaka and M. Tanaka. Characterization of sandy soils using CPT and DMT. *Soils and Foundations*, 38(3):55–65, 1998.
16. S. Marchetti, P. Monaco, G. Totani, and D. Marchetti. *From Research to Practice in Geotechnical EngineeringD.K. Crapps*, volume 180, chapter In -situ tests by seismic dilatometer (SDMT), pages 292–311. ASCE Geotech. Spec. Publ., 2008.
17. Yufen Huang and Norman R. Draper *Transformations, regression geometry and R2*, Computational Statistics & Data Analysis, 42(4), pages 647–664, 2003
18. Netlab. http://www1.aston.ac.uk/eas/research/groups/ncrg/resources/netlab/
19. C. Bishop. Neural Networks for Pattern Recognition. Oxford university Press, 1996.
20. C. Cortes and V. Vapnik. Support-vector Networks. Journal of Machine Learning, 20(3), pages 273–297, 1995.
21. C.-C. Chang and C.-J. Lin. LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011
22. V. Vapnik. Statistical Learning Theory. Wiley, New York, NY, 1998.
23. B. Schölkopf, A. Smola, R. C. Williamson, and P. L. Bartlett. New support vector algorithms. Neural Computation, 12, pages 1207–1245, 2000.