Geobrain: Dutch Feasibility Database for Installing Sheet Pile Walls

A. R. (André) Koelewijn
Deltares (formerly GeoDelft), Delft, The Netherlands

A. M. J. (Annemieke) Mens
Delft University of Technology & Deltares (formerly GeoDelft), Delft, The Netherlands

GEOBRAIN: DUTCH FEASIBILITY DATABASE FOR INSTALLING SHEET PILE WALLS

A.R. (André) Koelewijn
Deltares (formerly GeoDelft)
Delft, The Netherlands

A.M.J. (Annemieke) Mens
Delft University of Technology & Deltares (formerly GeoDelft)
Delft, The Netherlands

ABSTRACT

In this paper it is shown how the knowledge embedded in case histories can be used to explicate some of the uncertainties contributing to the gap between theory and practice. With the help of computational intelligence techniques, collections of case histories in databases, as a type of collective memory of the geotechnical profession may be explored to turn this memory into collective brains in geotechnics: a GeoBrain. Regarding the scarcity of soil investigation data and the translation of the available data into a model, the 'schematization factor' has been introduced as a partial safety factor to account for the influence of data availability and the role of human expertise. Using a database of increasing size on the feasibility of installing sheet pile walls, the determination of optimal parameter values for prediction models is illustrated. It is shown that computational intelligence techniques like Bayesian Belief Networks and Genetic Algorithms can be very helpful to improve predictions of what is likely to happen in geotechnical practice.

INTRODUCTION

Of all civil engineering disciplines, geotechnical engineering seems to include the highest material-related uncertainties. Whereas with steel the uncertainty is less than 5% and with concrete it is less than 10%, uncertainties in geotechnics are often more than 50%. Many geotechnical engineers seem to take this for granted, referring to the inherent uncertainties when dealing with heterogeneous material in the subsoil that cannot be removed. Meanwhile, budgets for soil investigations are too small and made available too late. However, our modern society with its increasing level of information on all kinds of details shows less and less acceptance that such issues are not being solved. The intrinsic uncertainty in geotechnics often leads to remarkably high risk and cost (Littlejohn, 1991, van Staveren, 2006), and additional vulnerability in areas prone to hazards. Following these authors, we clearly need to justify the intrinsic uncertainty in geotechnics to the outside world.

This paper deals with two elements in this geotechnical struggle against uncertainty: the introduction of a partial safety factor related to the extent of the soil investigations and the application of databases of case histories to aid the engineer in the design of new projects.

As the uncertainty can be decreased by (additional) soil investigations, this should be reflected by design codes allowing a smaller uncertainty factor in case of more detailed soil investigations (and, vice versa, a higher factor if little or no soil investigations are carried out). In the next section an example of this is given from the new Dutch flood protection guidelines, based on a yet limited number of case histories.

Collecting case histories, as a kind of 'collective memory', can make us learn from the past. The application of tools like Artificial Neural Networks, Bayesian Belief Networks and Fuzzy Logic to a collection of case histories may help to explore the contained knowledge and turn this collective memory into 'collective brains'. These collective brains may eventually know better than individual experts. Such collective brains in geotechnics, or a 'GeoBrain', can be made available to the whole profession – including the distinghuished experts. The main part of this paper is dedicated to the description of the performance of a GeoBrain applied to the feasibility of the vibratory installation of sheet pile walls.

Finally, some concluding remarks and a future outlook are given.

EXPERIENCE INTO MODELS: GEOBRAIN

As pointed out by Barends in the 2005 Terzaghi Oration (Barends, 2005), the initiation, design and construction of large-scale infrastructural projects is becoming more and more complex. Policies concerning multi-functional space should cover sustainable building, integral project approach, procedures and licenses, incorporation of existing infrastructure, pollution control and archeological issues. The situation calls for swift and comprehensive answers with the adoption of all available expertise and experience, presented in a clear and understandable manner at all stages.

GeoBrain provides a toolbox for an integral approach of complex situations where the subsoil is an important risk factor, leading to a comprehensive view on the objective consequences of choices made. Also in a legal context, GeoBrain has already proven its value, as the results are reproducible (Spruit, 2007). This development has a strong parallel with developments in other disciplines, like in medical science where diagnostic systems are used to translate empirical knowledge into generally applicable concepts with the use of present-day information and communication technology.

GeoBrain forms a unique research facility, the brain-side complementary to common physical and numerical research facilities. By applying artificial intelligence to couple numerical prediction models, physical tests and case histories, a complete set of data interpretations, practical experiences, expert views and test results can be translated into objective information. By this approach so-called soft data (e.g. complaints from neighbours) can be combined with hard data (e.g. weight of crane).

There has been hitherto no possibility of systematic learning from case histories of completed projects. Practicing engineers do, from time to time, propose ad-hoc rules and equations based on experience and field observations, but no unified framework for dissemination of knowledge has been available to engineers.

In recent years, the development of computational intelligence tools and the increasing availability of computational power enable engineers to analyze field data during construction and truly apply observational methods as recommended by various codes of practice. Up to now, geotechnical institutes and engineers have concentrated on the development of computational prediction models to simulate the observations of engineering practice, sometimes with limited success.

One of the forms by which a real improvement can be achieved is by strenghtening the evaluation and feedback loops from contractors and subcontractors to the designer, as indicated in Fig. 1.

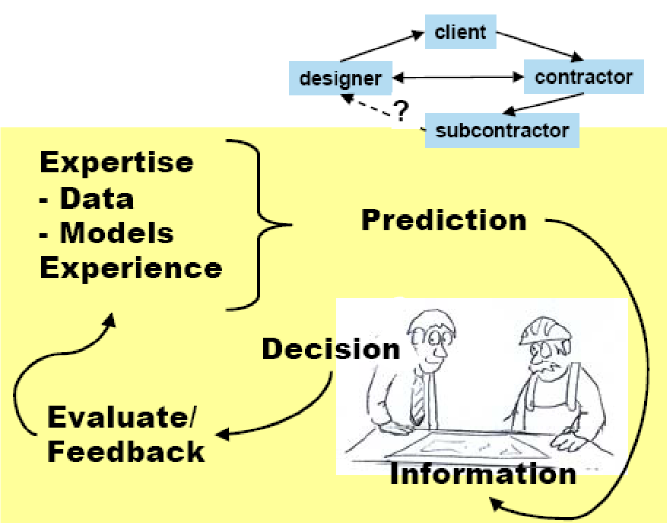

Fig. 1. The decision and control process at the site (Barends, 2007).

In general, objectives are to decrease risk in construction projects, reduce losses, improve the image of contractors and geo-engineers, improve working conditions, ensure completion of these projects without unforeseen delays and last but not least the reduction of insurance fees. Especially in foundation engineering and in drilling technology it is difficult to insure projects. Fees are high and often the policy does not cover major failures.

GeoBrain is addressing these problems directly by developing experience databases from case histories and disseminating these experiences via the Internet. These databases, complemented with expert knowledge, can be used to make predictions with a methodology based on artificial intelligence. There are, therefore, two kinds of output from a GeoBrain system: experiences and predictions, see Fig. 2.

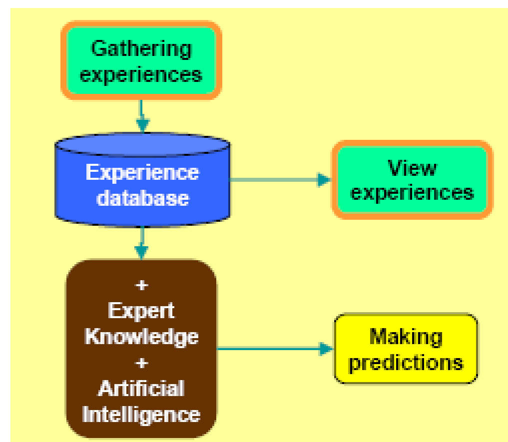

Fig. 2. Flow chart of a GeoBrain system (Barends, 2005).

Experience in its context can be objectivated using obligatory questionnaires, which have been composed together with the users and providers of equipment (e.g. a dropdown list for steel sheet pile profiles). Based on this, predictions are made using artificial neural networks and Bayesian Belief Networks, built from expert knowledge and validated by real case experiences. Its components are indicated in Fig. 3.

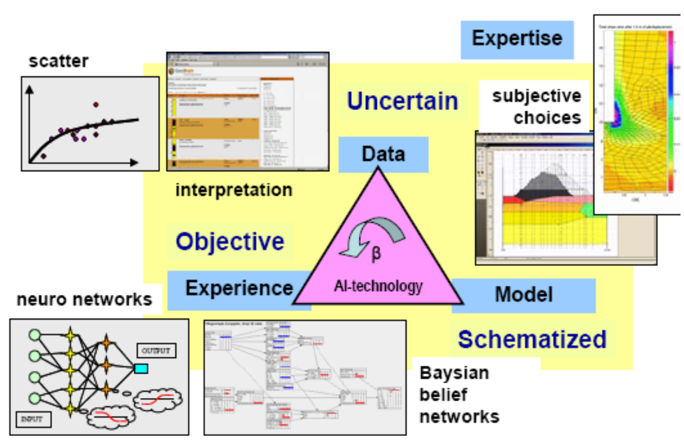

Fig. 3. GeoBrain matching expertise and experience (Barends, 2007).

The Internet is an ideal medium to display experiences and results of a prediction. Users can search through experiences by archetype or via a map on location, allowing refinement of queries. Predictions can be made on the same website.

As yet, the toolbox has to be filled for a large part. Examples of already implemented sections, mainly on foundation engineering and horizontal directed drilling technology can be found at www.geobrain.nl.

Although focussing on geotechnical engineering, the scope of GeoBrain is rather diverse. The combination of data and models, traditionally requiring expertise, by artificial intelligence-technology with a potentially large number of experiences from the past in combination with the proper types of visualization enables to make the right decision in an efficient way, and providing feedback as a new experience, as depicted in Fig. 4.

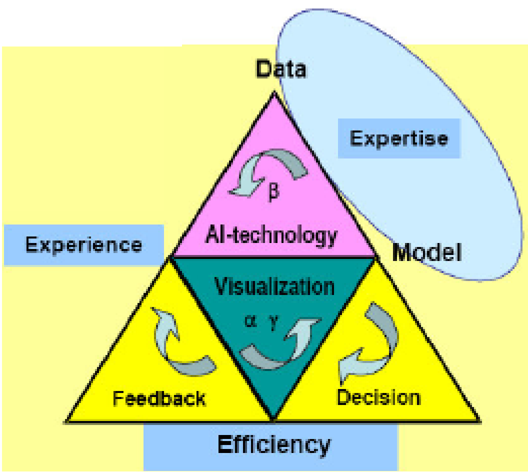

Fig. 4. The diversity of GeoBrain (Korff, 2007).

One of the factors contributing to the apparent uncertainty is the translation of the available data into a model. This does not only involve the data and the model, but also the translator, i.e. the engineer. More or less at the start of GeoBrain, the first author made five experienced geotechnical consultants independently determine the stability factor of a river embankment according to Bishop's method (Bishop, 1955) and a similar method developed by Van (2001). As detailed by Koelewijn (2002), this lead to mutual differences of more than 20% - in spite of the fact that they were given the same information regarding the subsoil, geometry and boundary conditions. In this case, it appeared that all consultants advised on the safe side, as the considered embankment only failed in a large-scale field test at a significantly higher loading level than predicted according to the evaluation guidelines these consultants had to use.

With this experience in mind, the Dutch commission responsible for the design and evaluation guidelines for the flood protection embankments has decided to introduce a 'schematization factor' as a partial safety factor in design (ENW, 2007). This factor has initially been set at 1.3. In the explanation it is stated that this factor may be reduced to as low as 1.0 if a sufficient reduction of the uncertainties concerning the composition of the subsoil and the pore pressures can be demonstrated, but the procedure for this still has to be developed. Simultaneously with the introduction of this schematization factor, the partial safety factor for the material properties has been reduced to arrive on average at the same results. Although the details still have to be worked out, in many cases this procedure will make additional efforts in soil investigations pay off in the design, even if adverse soil conditions are found.

FEASIBILITY OF INSTALLING SHEET PILE WALLS

Introduction

Sheet piles are long structural sections with a vertical interlocking system that creates a continuous wall. The walls are most often used to retain either soil or water and usually made of steel. The ability of a sheet pile section to perform is dependent upon its geometry and the soils it is driven into. The pile transfers pressure from the high side of the wall to the soil in front of the wall.

There are permanent and temporary applications. Permanent sheet piles remain in the ground and serve as permanent retaining structures. Temporary sheet piles are designed to provide safe access for construction, and are then removed.

Database on foundation engineering

The GeoBrain experience database on foundation engineering and the one on drilling technology (Hemmen, 2005) give a framework for storing case histories in order to provide engineers and designers with extra information to come to fast, reproducible and objective decisions. This database is complemented with an prediction tool based on a Bayesian Belief Network. Therefore the GeoBrain database essentially provides for two kinds of output: experiences and predictions, which tend to reduce the gap between theory and practice.

Experiences available

Together with the Dutch Association of Contractors in Foundation Engineering (NVAF) which has substantial experience and a good reputation in the application of geotechnical know-how, this on-line database is continuously filled with up to date information about ongoing projects. The total number of entries approached 1300 projects by the end of 2007, of which 381 concern the vibratory installation of steel sheet piles and 514 the installation of prefabricated concrete piles. An experience is uniquely defined by the type of element (for example a sheet pile or prefabricated concrete pile), the type of equipment used and the soil conditions. In addition to this numerical data, also details concerning the building pit, the crew and the surroundings are included.

For the validation of the prediction model for the vibratory installation of steel sheet piles described hereafter, only 191 out of the 381 available case histories have been used, as for the other cases additional measures were used to reach the planned depth, or essential data, like the results from a cone penetration test, were missing.

Bayesian Belief Network

Bles et al. (2003) developed a Bayesian Belief Network (BBN), based on professionals' experience, to model the risks during construction of pile foundations. The model can be used to forecast the drivability of a steel sheet pile walls. BBNs use probabilistic theory for reasoning under uncertainty and risk in expert systems. Bayes' theorem is the cornerstone in this way of reasoning, because it provides a way to calculate the posterior probability $P(h|D)$, from the prior probability $P(h)$, together with $P(D)$ and $P(D|h)$:

$$P(h\,|\,D) = \frac{P(D\,|\,h)P(h)}{P(D)}. \qquad (1)$$

$P(h|D)$ is also called the conditional probability of h, given D (Mitchell, 1997).

The method transforms joint probability functions to a set of stochastic variables, ordered in a network. The network itself consists of two parts. The qualitative part shows the relations between the variables in a graphical representation (the network). The quantitative part assigns conditional probabilities to all variables, using likelihood-tables, which describe the effect of preceding variables on the underlying ones. Finally the BBN provides the user with a number between 0 and 100, describing the amount of risk not reaching depth. In the prediction model available on the Internet the number is translated into a color code ranging from green (no risk) to red (unacceptable), to avoid discussions about the threshold value, see Fig. 5. The threshold for the BBN should be a score ranging from 30 to 45, according to experts. The experiences in the database make it possible to optimize this threshold value.

Fig. 5. Example of color code as a result from the Bayesian Belief Network.

Example of the practical use of the BBN in design

A practical example of the use of the BBN is the following. Consider a polluted site with an aquifer of dense sand at a depth of 15 to 20 metres. To avoid spreading of the pollution through the groundwater, one may think of a steel sheet pile wall extending through the sand layer. Once installed, this sheet pile wall may be very thin, as hardly any loading will act upon it. A prediction by the BBN will yield an unacceptable risk that it will not be possible to install the elements. Application of a thicker profile will reduce the number yielded by the BBN. For a heavy steel sheet pile profile, a small risk and thus a 'green score' will be obtained. Of course, in this case one may also think of installing a slurry wall instead.

Determining the optimal threshold

In order to determine the optimal threshold value x, the field experiences from the database have been compared to the forecasts made by the BBN. Since we are interested in whether a sheet pile will reach its planned depth or not, for each case the percentage of sheet piles not reaching the planned depth has been calculated. For practical reasons, this should be very low, but not necessarily zero. Here, the threshold is put at 1%. For more detailed information on the consequences of using another percentage, see Mens et al. (2008). The results of the BBN for different threshold values will be presented in so-called 'confusion matrices' (Gardner and Urban, 2003) to simplify the comparison. An example is shown in Fig. 6. In this figure, the correspondance between a prediction (horizontal axis) and a field experience (vertical axis) is given. For instance, X_{11} states the percentage of cases with a negative forecast and a positive field experience.

Field experience	**i = 1 (+)**	X_{11}	X_{12}
	i = 2 (-)	X_{21}	X_{22}
Model Specification		**j =1 (-)**	**j = 2 (+)**
		forecast	
		fitness	

Fig. 6. Example of a confusion matrix.

In the ideal case, X_{11} and X_{22} are both zero. As long as the ideal prediction model is not found, the following fitness function is used to compare all results:

$$f = \left(\frac{x_{22}}{x_{21}+x_{22}}\right)^2 + \left(\frac{x_{11}}{x_{11}+x_{12}}\right)^2. \qquad (2)$$

The smaller the value of f, the better the model prediction.

Results

The results for various values of the threshold value of the BBN x are shown in Fig. 7. The best results are found for a value of x = 32, which is near the lower bound of the range indicated by experts, as mentioned before.

BBN x = 30	forecast -	forecast +
Field experience +	35%	53%
Field experience -	5%	7%
	f = 0.48	

BBN x = 32	forecast -	forecast +
Field experience +	41%	47%
Field experience -	7%	5%
	f = 0.37	

BBN x = 38	forecast -	forecast +
Field experience +	62%	26%
Field experience -	8%	4%
	f = 0.61	

BBN x = 45	forecast -	forecast +
Field experience +	66%	22%
Field experience -	9%	3%
	f = 0.61	

Fig. 7. Results of BBN for different values of threshold value x for 191 cases.

These results are a bit surprising when compared to earlier results using only 50 cases, but including three other models, as detailed in an earlier paper by Mens et al. (2008). For clarity, the main results from that study will be repeated in the following.

Comparison with other models

For 50 case histories only, the results of the BBN have been compared with two rules of thumb for the installation of sheet piles, viz. the CUR-rule (CUR, 2005) and the rule of Azzouzi, which is based on a numerical model (Azzouzi, 2003), and the Vibdrive model, which is a numerical model to calculate the penetration speed of a sheet pile (Holeyman et al., 1996). The latter has been improved using a genetic algorithm (Mens et al., 2008). The results are given in the confusion matrices of Fig. 8.

In spite of the remarkable improvement of the Vibdrive-model when applying the optimization by the genetic algorithm, the best results are achieved with the BBN, with threshold values more or less in the middle of the range indicated by the experts.

Discussion

Apparently, fifty cases are not sufficient to determine an optimal value for the threshold of the BBN, as with almost four times more cases a significantly different result is obtained.

Although it has not yet been investigated, it may also be questioned whether for the larger number of cases the best results are achieved with the BBN or with one of the other models.

CUR-rule	forecast -	forecast +
Field experience +	34%	48%
Field experience -	8%	10%
	f = 0.48	

Azzouzi	forecast -	forecast +
Field experience +	20%	62%
Field experience -	6%	12%
	f = 0.50	

Vibdrive	forecast -	forecast +
Field experience +	8%	74%
Field experience -	2%	16%
	f = 0.80	

Vibdrive with GA	forecast -	forecast +
Field experience +	8%	74%
Field experience -	6%	12%
	f = 0.45	

BBN x = 37 to 40	forecast -	forecast +
Field experience +	34%	48%
Field experience -	10%	8%
	f = 0.37	

Fig. 8. Results for different models for 50 cases.

This indicates that when databases of case histories are used to make decisions regarding the optimal parameters of a model, or even regarding which model is the best, these decisions should be checked regularly while expanding the number of case histories in the database.

CONCLUSIONS AND FUTURE OUTLOOK

So far, a few pieces more of the geotechnical puzzle have been solved by the application of the GeoBrain concept.

Regarding the scarcity of soil investigation data and the translation of the available data into a model, the influence of the human factor has been recognized and a start has been made with the introduction of a partial safety factor called 'schematization factor'.

Using a database of increasing size on the feasibility of installing sheet pile walls, the determination of optimal parameter values for prediction models has been illustrated. It has been shown that computational intelligence techniques like Bayesian Belief Networks and Genetic Algorithms can be very helpful to improve predictions of what is likely to happen in practice.

As yet, only small parts of geotechnical practice are covered. In the coming years, this way of collecting case histories to turn the embedded knowledge into collective geotechnical brains will be expanded to more applications where at present a significant gap exists between theory and practice.

REFERENCES

Azzouzi, S. [2003]. *"Intrillen van stalen damwanden in niet-cohesieve gronden – welke predictie is (on)juist?"* ["Vibratory installation of steel sheet pile walls in non-cohesive soils – which prediction is(n't) wrong?"], M.Sc. thesis, Delft University of Technology, Delft [in Dutch].

Barends, F.B.J. [2005]. "Associating with advancing insight – Terzaghi Oration 2005." *XVI Int. Conf. on Soil Mech. and Geot. Eng.* Osaka, 217-248.

Barends, F.B.J. [2007]. "Geotechnics in Modern Societies." *Krebs Ovesen Memorial, XIII European Conf. on Soil Mech. and Geot. Eng.* Madrid, 8 pp.

Bishop, A.W. [1955]. "The use of the slip circle in the stability analysis of slopes." *Géotechnique,* 5:7-17.

Bles, T., Al-Jibouri, S., and Adel, J. van den [2003]. "A risk model for pile foundations." *ISARC 2003, 20th Int. Symp. on Automation and Robotics in Construction.*

CUR [2005]. *"Damwandconstructies, 4e druk"* ["Sheet pile wall constructions, 4th edition"], CUR handbook 166, CUR-NET, Gouda.

Expertise Netwerk Waterkeren (ENW) [2007]. *"Addendum bij het technisch rapport waterkerende grondconstructies"* ["Addendum to the technical report water retaining soil constructions"], Ministerie van Verkeer en Waterstaat, Den Haag, ISBN 978-90-369-1411-6 [in Dutch].

Gardner, R.H., and Urban, D.L. [2003]. "Model validation and testing: Past lessons, present concerns, future prospects." *Models in Ecosystem Science* (Cole, J.J., and Lauenroth, W.K., eds), Princeton University Press, Princeton, 184-203.

Hemmen, B.R. [2005]. "The synergy between theory and practice in geo-engineering." *XVI Int. Conf. on Soil Mech. and Geot. Eng.* Osaka, 2809-2811.

Holeyman, A., Legrand, C., and Rompaey, D.V. [1996]. "A method to predict the driveability of vibratory driven piles." *3rd Int. Conf. on the Application of Stress-Wave Theory to Piles*, Orlando, 1101-1112.

Koelewijn, A.R. [2002]. "The practical value of slope stability models", *Learned and Applied – Soil Mechanics out of Delft* (Barends, F.B.J., and Steijger, P.M.P.C., eds), Balkema, Lisse, 107-114.

Korff, M. [2007]. "How to make use of expertise, experience and predictions for decision making and control?", *DMI Workshop*, Delft University of Technology, Delft.

Littlejohn, G.S. [1991]. *Inadequate site investigation*, Thomas Telford, London.

Mens, A.M.J., Tol, A.F. van, and Koelewijn, A.R. [2008]. "Optimizing Foundation Engineering, Validating models against experience using Artificial Intelligence", *12th IACMAG*, Goa. *To appear.*

Mitchell, T.M. [1997]. *Machine Learning*, McGraw-Hill, Singapore.

Spruit, J.W. [2007]. *"Model voor risicoanalyse open stalen buispalen"* ["Model for risk analysis open tubular steel piles"], B.Sc. thesis, Hogeschool Utrecht, Utrecht [in Dutch].

Staveren, M. van [2006]. *Uncertainty and ground conditions: a risk management approach,* Elsevier, Burlington.

Van, M.A. [2001]. "New approach for uplift induced slope failure." *XV Int. Conf. on Soil Mech. and Geot. Eng.* Istanbul, 2285-2288.