THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

PERFORMANCE INDICATORS

A performance prediction method for moisture safety design

VERONICA YVERÅS

Department of Civil and Environmental Engineering
Division of Building Technology
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2009

PERFORMANCE INDICATORS
A performance prediction method for moisture safety design
VERONICA YVERÅS
ISBN 978-91-7385-308-8

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr. 2989
ISSN 0346-718X

Department of Civil and Environmental Engineering
Division of Building Technology
Chalmers University of Technology
SE-412 96 Göteborg, Sweden
Telephone + 46 (0)31-772 1000
Fax +46 (0)31-18 97 95

Printed by
Chalmers Reproservice
Göteborg, 2009

PERFORMANCE INDICATORS
A performance prediction method for moisture safety design
VERONICA YVERÅS
Department of Civil and Environmental Engineering
Division of Building Technology
Chalmers University of Technology

ABSTRACT

The aims of this research project is to develop and gain knowledge about a different approach to moisture safety design based on AI in order to attain healthy buildings and relate this new approach to current practice and prospective users.

Secondary data (data produced for some other purpose) was used to train Artificial Neural Networks (ANN) to predict the performance of outdoor ventilated crawl-spaces regarding microbiological smell, mould and rot. The best performing ANN managed to predict smell 100%, mould 76%, and rot 92% correctly on 38 validation cases not used in the training process. A reliability test was performed designed as a parameter study. The results highlighted some uncertainties in the trained ANN which are likely to be due to a high level of missing values and skewed data. In addition, the parameter study goes far outside what the ranges of the retrieved training data, forcing the ANN to extrapolate predictions.

The interview study with engineering consultants indicated that experience is considered to be a decision support in moisture safety design even though feedback from past projects rarely is available. In addition the general opinion was that available tools are too demanding. Through a questionnaire a performance prediction comparison was set up to test the competitiveness of the trained ANN. The average prediction result for the respondents (engineering consultants, moisture damage consultants, moisture experts) was 50% correct predictions whereas the ANN had a 93% correct prediction level. There was no notable indication of a correlation of the prediction results with the respondents' background. The same study also revealed that a system to capture experience is highly requested by the respondents.

The results so far are promising but ANN, based on real life experience, must be tested further with better training data, preferably with data designed for this purpose. The method has a potential to capture real life experience in a structured and systematic manner. Moreover, it may be helpful in the decision process during the early stages of design.

Keywords: Performance prediction, Artificial Neural Network, Moisture safety

LIST OF PUBLICATIONS

This thesis is based on the following papers referred to by Roman numerals I-V in the text.

I Burke, S. and Yverås, V. (2003) A Swedish perspective on the prevention of moisture problems during the building's design phase. *Nordic Journal of Surveying and Real Estate Research* **1**(1): 102-113*

II Yverås, V. (2005) Predicting the service life by Artificial Intelligence. Proceedings of the 7th Symposium on Building Physics in the Nordic Countries, Reykjavík, Iceland, 13-15 June, 2005.

III Yverås, V. (2010) Performance prediction method in the early stages of design for outdoor ventilated crawl-spaces based on artificial neural networks *Accepted for publication during 2010 in Journal of Building Physics.*

IV Yverås, V. (2009) Competitiveness of a performance prediction method based on artificial intelligence. *Submitted to Journal of engineering, design and technology*

V Yverås, V. (2009) Reliability test on performance prediction method for outdoor ventilated crawl-spaces based on neural networks. *Submitted to Neural Computing and Applications*

* The input to the paper was equally shared between the authors

OTHER PUBLICATIONS

In addition to the previously listed publications, the author has also written the following papers during the course of this research project which have had an influence on the direction in which the research has proceeded.

Yverås, V (2002) Performance indicators as a tool for decisions in the building process. In. ed. Atkin, B., Borgbrant, J., and Josephson, P-E., *Construction Process Improvement*. Oxford: Blackwell Science

Yverås, V (2002) Performance indicators as a tool for decisions in the building process. *Proceedings of the 6th Symposium on Building Physics in the Nordic Countries, Trondheim, Norway, June 17-19, 2002*

Yverås, V (2002) *The performance indicators as Decision Support Tool.* Göteborg: Chalmers University of Technology

Yverås, V (2002) *Funktionsindikatorer – Ett verktyg för bedömning av tekniska lösningar.* (Performance indicators – A tool to predict technical solutions) FoU-Väst, Sveriges Byggindustrier: Göteborg

Yverås, V (2003) Performance indicator tool vs simulation tool. *Proceedings of the 2nd International Building Physics Conference, Leuven, Belgium, September 14-18, 2003*

Yverås, V (2006) Verktyg kan förutspå risk för framtida fukt. (A tool can predict future moisture problems) *Husbyggaren* nr1:06

Yverås, V (2008) *Aspects of extracting real life data for neural network learning on service life predictions.* Report No. 2008:16, Göteborg: Chalmers University of Technology

Yverås, V (2008) *Predicting service life of outdoor ventilated crawl spaces by neural networks.* Report No. 2008:17, Göteborg: Chalmers University of Technology

PREFACE

Writing this preface makes me realise that this also puts an end to a phase of my life. It has been a mix of learning, writing, doubts, new friendships and endless hours of reasoning. In the end it has resulted in a thesis. Hopefully, my friends and family finally can see and perhaps understand what I have been doing during these years.

I would like to express my sincere gratitude to Prof. Carl-Eric Hagentoft, head supervisor, for all the support during my doctoral studies at Chalmers University of Technology. Your feedback and input has been of great value. Many thanks also go to Docent Bertil Thomas, Chalmers, who helped me through the jungle of Artificial Neural Networks technology. Kent Haglund, former with JM AB, thank you for believing in me and the project.

I wish to acknowledge the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS), the Development Fund of Swedish Construction Industry (SBUF), Centre for Management of the Built environment (CMB), and JM AB who kindly funded the project and also the support provided by Competitive Building. I also thank the reference group in FoU-Väst who has been an enriching discussion partner during the whole research project. Anticimex and Småhusskadenämnden, many thanks for giving me access to your archives.

Finally I would like to thank the two most important persons in my life. My husband Magnus, for all love, support and encouragement during these years. My son Eric, the sunshine in my life who always gives me a reason to smile and is my inspiration. I will always love you.

Veronica Yverås
August, 2009

CONTENTS

1 INTRODUCTION

1.1 BACKGROUND AND RESEARCH CONTEXT

Moisture issues in the building process have been handled differently during the course of time with various results. There is a general frustration reflected by the media that are eager to provide an image of the building industry as a non-serious business, often claiming that poor workmanship in the end results in moisture and mould problems. There is no doubt that the construction industry in many ways is responsible for the moisture problems. It is a well known fact that these problems are burdened with considerable costs. Tolstoy (1994) states that in Sweden roughly SEK 6 billion per year is spent on repairs and maintenance of buildings and of that, approximately half goes to damages caused by moisture. Furthermore, it is estimated that € 9 billion is spent on repair of moisture related damages in the European Union, which approximately equals to 1% of the annual return of the building sector (Adan *et al*, 2004). In British Columbia, Canada, it is predicted that 50-70% of the homes built between 2000 and 2010 will suffer from moisture damage (Barrett, 2000). According to Lisø (2007) the yearly costs of moisture damage in Norway is estimated to 10.5 billion Norwegian kroner. Anticimex, which is a damage consultant company, estimates that 175 000 homes in Sweden have a damaged crawl space (Anticimex, 2004) where the repair costs range from 30 000 to 250 000 SEK. In the US, the direct and indirect costs of treating asthma caused by poor indoor air quality were reported to be about $13 billion US (Weiss & Sullivan, 2001). Thus, moisture problems in the building sector exist in many countries.

Regardless of who is to be blamed, all actors within the construction process have a responsibility to prevent moisture problems from occurring. The engineering consultants involved in the design stage are no exception. It is in this stage that the conditions for a good service life are created. Despite the wide range of tools available, they are not frequently used. One reason is thought to be that they require expert knowledge (Hendriks and Hens, 2000). Tools at the expert level are not only difficult to use in practice, but also provide designers with raw data that often requires some level of expertise to interpret. This data must in turn be translated into information that the clients are able to grasp. By improving the quality of information during the design process, the client is better equipped to understand the different issues implicated in the project (Barrett and Stanley, 1999). There is a need for decision support systems that could provide designers with simplified prediction models, which, in turn, could be used to classify and illustrate results for the client's benefit.

Several of the available moisture calculation tools originate from a research environment and, hence, are based on quantitative determination. These tools are, therefore, primarily not aimed at practical design work. According to Sandin (1998) there are basically three approaches to moisture safety design: quantitative determination, qualitative assessment or tried and documented solutions. The last one is only applicable in a situation where the conditions are the same as those of the new project. Quantitative determination involves calculations of the moisture condition of the design at hand. According to Sandin (1998), the results of such calculations require extensive knowledge to be handled and correctly interpreted. The qualitative assessment offers a somewhat simplified moisture safety design approach. It is about making small changes in a well known design and assess if the performance is improved or not. This approach might for instance also include reading tables indicating expected moisture conditions in different situations.

As the early stages of design contain a low level of information, the input data are, therefore, not that detailed. During this process different designs are evaluated where the components are known but rarely the dimensions. Quantitative tools are therefore not helpful at this point. A decision support tool at this stage must be able to deal with simplified data. Due to simplified data the output data can therefore only provide rough estimates. Currently, there are no such tools available why the engineering consultant has to rely on previous experience and on information found in literature.

This kind of experience is however difficult to capture and externalise when many parameters are involved. Artificial Intelligence (AI) has become rather popular in this kind of context. It provides the possibility to extract knowledge from experience and apply it to a prediction problem. Several areas such as economics, law and medicine have found AI to be helpful for prediction problems which rather often can be based on either real life data or expert knowledge. Thomas (2003) describes it as a software system with in-built knowledge that within a knowledge domain has the ability to resemble a human expert.

In a previous work (Yverås, 2002), performance indicators as a decision support tool have been explored in the context of simplifying the initial moisture safety design process. The essence of performance indicators is to simplify complex relationships by presenting rough estimates. Even though the indicators provide rough estimates, it can be valuable and usable information, especially for those who are not familiar with the complex theoretical background. By using performance indicators as the basis of a decision support tool, the required knowledge to make moisture safe designs can be made more accessible to the engineering consultants. Moreover, it helps the clients to understand the implications of their design decisions.

The hypothesis has therefore been identified and formulated as follows:

A performance prediction method based on artificial intelligence and performance indicators can be applied in the early design stage to initiate the moisture safety design process.

1.2 RESEARCH OBJECTIVES

The aim for the work presented in this thesis is to develop and gain knowledge about a different approach to moisture safety design based on AI in order to construct healthy buildings.

The specific objectives of the research are:

1. Identify a suitable Artificial Intelligence system for the task.
2. Apply the chosen system on a design/building element and validate.
3. Relate and evaluate the approach to current practice and prospective users.

1.3 SCOPE OF THE RESEARCH

The scope of the research is focused on the early stages of design where the first decisions are made. This only concerns moisture related issues as incorrect decisions would lead to a poor indoor environment due to mould or to impaired mechanical properties due to rot or rust.

It is not the objective to analyse and evaluate all available AI-systems as it would require too much time. For the same reason the application of this approach will be limited by available AI-software.

The application of AI was restricted to one design, the outdoor ventilated crawl-space. The results of this study are, therefore, only applicable to designs with similar characteristics. The ventilated crawl-space is, for instance, a rather open design which allows ocular inspections without destruction.

1.4 READING INSTRUCTIONS

This thesis presents a short summary of five papers which the thesis is based. More detailed information can be found in the five appended papers.

After the introduction section the thesis continues to describe the research methodology applied in this research project. The section following, *Data set up and data quality* could have been included in the research methodology chapter. However, as this information can not be found in the appended papers it deserves to be treated in a separate chapter. The thesis is thereafter ended with separate results, discussion and conclusion chapters.

2 RESEARCH DESIGN

2.1 EVALUATION OF AI-SYSTEMS (paper II)

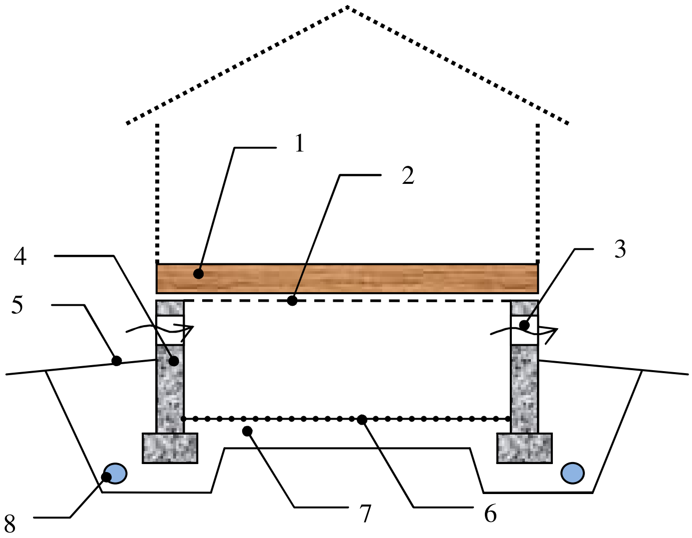
The evaluation of AI-systems involved a literature study of journals, text books, and conference proceedings. The aim of the literature review was to gain general theoretical knowledge of the AI-area and to study different applications. Furthermore, to save time and effort, the study was restricted to two different main AI-systems. Besides comparing these two systems with each other, the characteristics of the moisture safety design application were also taken into consideration.

2.2 APPLICATION OF ARTIFICIAL NEURAL NETWORKS (paper III and V)

The second stage of the study concerned the application of the final chosen AI-system, artificial neural networks (ANN), which was performed in four steps:

1. Retrieval of real life data
2. Data inspection
3. ANN design and training
4. Validation

The outdoor ventilated crawl space design was chosen as test object (Fig.1). In order to build an ANN it is necessary to have access to training data. Several options of sources for data retrieval where considered; human experience, field studies and secondary data in existing archives. The first involves extracting knowledge from experts, who would have to review at least 300 constructed cases with a large amount of variables (20-25). Halford *et al.* (2005) have performed research into the human mind and its capacity of processing information. Their findings indicate that a structure of four variables is at the limit of human processing capacity. From this point of view the expert would be subject to an impossible task in predicting the service life. The risk of receiving contradictory data is high.



1. Floor structure
2. Counter floor
3. Ventilation valve
4. Foundation wall
5. Ground surface
6. Vapour barrier
7. Capillary breaking and draining material layer
8. Drainage pipe

Figure 1. Basic outdoor ventilated crawl-space design (Yverås,2010)

Field studies were another possibility considered. This would involve in situ documentation and assessment of a large amount of outdoor ventilated crawl spaces with a distribution all over the country (Sweden). The quality of the assembled data is likely to be very good as the data would be retrieved in a consistent manner. However, due to much travel, the documentation rate performance would be rather low with perhaps only one case a day. It would require at least a year of documentation activities. More importantly, the travelling expenses do not correspond to available research resources.

The last alternative - existing databases and archives - was therefore finally chosen. Compared to previous source this is rather time efficient due to low travel intensity. However, it is secondary data that is not designed for this purpose. Missing data is therefore an expected issue that will have to be dealt with.

Real life data has been retrieved from two separate sources; *National Organisation for aid to owners of private small houses* (Småhusskadenämnden in Swedish or SNN)) and *Anticimex*. The SSN archive was founded in 1986 and recently closed down. Homeowners that had encountered moisture problems could turn to this fund and apply for financial help. All cases were investigated by the help of consultants who in turn delivered a report of the extent of damage, cause of damage and recommended measures. The reports were thereafter stored in the SNN archives.

The other source of data belongs to the private sector company Anticimex, whose business includes building inspections of different kinds. This project has taken interest in the inspection reports that are made before a house purchase is finalised. In contrast to SSN, the findings are documented by the surveyor on a form which is then stored on a computerized catalogue system. Approximately 1500 inspections are performed all over the country per year of which 31 % concern buildings with crawl-spaces.

The reason for having two sources is that they compensate each others deficiencies. The SNN archive provides cases with a higher degree of completeness but instead there is a lack of healthy cases. It is important to find both good and bad cases, meaning a range of crawl spaces that have different levels of conditions.

Before retrieving training data it is necessary to identify the parameters that are believed to have an impact on the service life of the building element. These parameters will serve as input data. A literature study was therefore performed regarding the outdoor ventilated crawl space which resulted in a data requirement specification. In addition it also had to be decided on how to describe the condition of the crawl-space cases, descriptions which would represent the output data. A preview of the data sources were therefore performed to state available parameters.

Having retrieved the data it was inspected in order to determine whether its quality agrees with the requirements of ANN. Noisy data that are believed to impair the training process of the ANN have to be removed and missing data need to be replaced. Analysing the data quality also helps to interpret the prediction results of the AI-system and maybe also to foresee difficulties. These issues need to be attended before training the ANN. Therefore, there are several steps that need to be taken before the training process of ANN can begin. An estimate reveals that about 60 % of the project effort is spent on data preparation (Qin *et al*, 2006). If the data preparation is well carried out the complexity of the network will be reduced and its' generalisation ability will increase (Lai *et al*., 2006)

Training of the ANN has been performed using the Neural Network Toolbox of Matlab 7.0 (Demuth & Beake, 2000). A back-propagation (Levenberg-Marquardt) algorithm with log-sigmoid transfer function in the nodes is applied to predict the performance of the outdoor ventilated crawl-space. Finding the best performing ANN designs is an iterative process which is why nearly 30 different ANN designs were tried.

The results of ANN-application were validated using a number of cases that were not used during training. The validation process measures the prediction ability by comparing the prediction results with the real outcome in the real life data. This is the traditional way to validate the trained ANN in order to state the accuracy. The results did however not correspond with the stated data quality, which is why a reliability test was performed which was designed as a parameter study comparing with expected results.

2.3 RELATING THE AI-APPROACH TO CURRENT PRACTICE AND PROSPECTIVE USERS

This part was divided into two separate studies where the first was undertaken in the beginning of this research project, and the second when the performance indicator tool had been developed and was ready for testing.

2.3.1 Current practice of moisture safety design (paper I)

The first study was designed as an exploratory survey and was performed in collaboration with another research project. The primary aim for this research project was to gain insight of how moisture safety design is approached in the design process in order to relate it to the applied AI-method. A total amount of eight building consultants with different levels of education and experience were interviewed. The survey was based on five main key questions (see Appendix A).

2.3.2 Performance *prediction* test, AI versus professionals (paper IV)

The second study tested the competitiveness of the AI-approach in comparison with professionals. It was designed as a performance prediction test where the trained ANN and the respondents were given five different outdoor ventilated crawl-space designs with different geographical locations (Fig. 2) and ages, to predict the condition. As the previous study indicated that experience was important as a decision support the results of the predictions were also compared amongst the participants in the study to see if their background had any impact on their prediction ability.

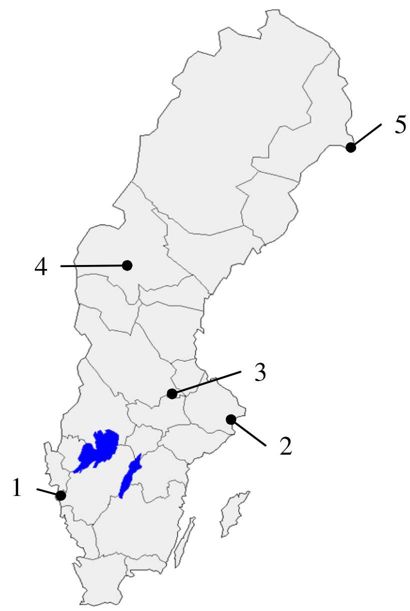


Figure 2 Geographical location of test cases.

The study was conducted through a survey using a web based questionnaire tool, Dialog manager 3.0, which also allows cross-analysis of the retrieved data. It was mailed out to those who are known to handle moisture issues and also to prospective users of a tool based on the method explored in this thesis. The questionnaire can be found in Appendix B.

3 DATA SETUP AND DATA QUALITY

In section 4.1 two different AI-systems are described, of which the Artificial Neural Network (ANN) was finally chosen. This chapter, besides presenting how the data was retrieved and organised, primarily describes how the flaws in the secondary data were handled.

When retrieving the cases the information was arranged as output and input parameters. Three output data, performance indicators, used to describe the condition of the outdoor ventilated crawl-space case: perceived microbiological smell, level of visible mould growth and level of visible rot, Table 1. Table 2 presents the input parameters which influence the performance of the crawl-space.

Table 1 Output representation – performance indicators

Output data	Definition
Y_1 Smell	0 = No smell, 1 = microbiological smell
Y_2 Ocular detected mould	0.125 = Nothing visual, 0.375 = Local spots, 0.625 = Light growth in major part of crawl-space, 0.875 = Extensive / rich growth
Y_3 Ocular detected rot	0.167 = Nothing visual, 0.5 = On surface, 0.833 = In depth

The drawback of using secondary data like in this research project is the level of data quality and that the parameter list and characteristics are limited as the sources are designed for another purpose. In this case this has resulted in missing data (table 5, Paper V), absence of at least one important parameter and a dearth of performance indicators. The performance of the ANN relies very much on the data quality at hand.

Table 2 Input representation

	Parameter	Definition
A	X_1 Capillary breaking layer	1 = yes, 0 = no
	X_2 Drainage system – roof	1 = yes, 0 = no
	X_3 Drainage system – ground	1 = yes, 0 = no
	X_4 Surrounding ground inclination	1 = yes, 0 = no
B	X_5 Insulation - counter floor	1 = yes, 0 = no
	X_6 Level of insulation in floor structure	[mm]
	X_7 Insulation – foundation wall	1 = yes, 0 = no
C	X_8 Ventilation – mechanical	1 = yes, 0 = no
	X_9 Vapour barrier	1 = yes, 0 = no
D	X_{10} Load carrying structure: inorganic	1 = yes, 0 = no
	X_{11} Counter floor: inorganic	1 = yes, 0 = no
	X_{12} Foundation wall: inorganic	1 = yes, 0 = no
	X_{13} Impregnation of wood material	1 = yes, 0 = no
E	X_{14} Floor heating	1 = yes, 0 = no
	X_{15} Organic waste	1 = no, 0 = yes
F	X_{16} Relative humidity	[%]
	X_{17} Precipitation	[mm]
	X_{18} Mean annual temperature	[C°]
	X_{19} Reference wind velocity	[m/s]
	X_{20} Surrounding terrain	1 = Outside urban areas , 0 = Urban
	X_{21} Ground material	Rock, clay = 0 / Moraine, Silt = 0.5 / Gravel, Sand = 1
G	X_{22} Age at inspection	Age at inspection – year of construction

3.1 REPLACEMENT OF MISSING DATA

Missing data needs to be replaced which can be done using various replacement techniques (random values, statistical methods, neural networks). Table 3 displays the final choice of missing values for each parameter where through the following procedure: In the first instance, replace missing data with "basic" values through logical reasoning, or use values from other cases originating from the same house manufacturer during the same year. Otherwise, use mean values where possible (for example for parameters such as insulation). If none of the above alternatives matches the situation: replace the missing data by a neutral value (0.5)

Table 3. Replaced missing values

Parameter		Replacement value
X_1	Capillary breaking layer	0.5
X_2	Drainage system – roof	1
X_3	Drainage system – ground	0.5
X_4	Surrounding ground inclination	0.5
X_5	Insulation - counter floor	0
X_6	Level of insulation in floor structure	Alt 1:.Manufacturer * Alt 2: concrete structure=85 mm Alt 3: else=200 mm
X_7	Insulation – foundation wall	0
X_8	Ventilation – mechanical	0
X_9	Vapour barrier	1
X_{10}	Load carrying structure: inorganic	0
X_{11}	Counter floor: inorganic	0
X_{12}	Foundation wall: inorganic	1
X_{13}	Impregnation of wood material	0.5
X_{14}	Floor heating	0
X_{15}	Organic waste	1
X_{16}	Relative humidity	No missing value
X_{17}	Precipitation	No missing value
X_{18}	Mean annual temperature	No missing value
X_{19}	Reference wind velocity	No missing value
X_{20}	Surrounding terrain	No missing value
X_{21}	Ground material	0.5
X_{22}	Age at inspection	No missing value
Y_1	Smell	0
Y_2	Ocular detected mould	0.125
Y_3	Ocular detected rot/rust	0.167

* Cases with same manufacturer can be considered to have same dimensions

As can be seen in Table 3, the strategy favours logical reasoning; for example if no remark or comments can be found regarding the vapour barrier in a case, it can be assumed that there is a vapour protection on the ground. In this case the logic is based on the conception that there is a prevailing idea of how a basic crawl-space should be composed, e.g. *there should always be a vapour protection on the ground*. Accordingly, parameters that are deviating from the conception of a basic crawl-space are presumed to have been mentioned in the inspection protocol. In order to follow and apply this logical reasoning when replacing missing values, a strong "idea" of what is regarded as necessary in a basic crawl-space is required. Moreover, can the investigators be trusted to have good knowledge about the basic design and what it should contain?

There is, however, one basic parameter where the basic reasoning could not be applied. The impression received when reviewing the retrieved cases is that parameter X_1 (capillary breaking layer) does not seem to be an important element in the crawl-space design as it is never mentioned even where the ground contains clay. This is why the neutral value 0.5 is assigned to parameter X_1. The same goes for *drainage system in ground* (X_3) as it is difficult to inspect.

It was found that the surrounding ground inclination is an important element in the archives. Snow on the ground was however a reason for the inspectors not to give an opinion about the ground inclination. Hence, a neutral value of 0.5 replaces missing values.

Parameter X_5, X_7, X_8, X_{10}, X_{11} and X_{14} are elements that are believed to improve the service life conditions of a crawl-space, but are not considered to be a part of the basic crawl-space design. The assumption is therefore that these solutions would not be left out in the reports if they exist, which is why any missing value are replaced by 0.

Regarding the insulation in the floor structure there are three ways to replace a missing value. Firstly, in many of the reports the name of the house manufacturer is mentioned. From what can bee seen in the retrieved material the insulation thickness in the floor structure changes very little over the years but varies between manufacturers. This knowledge can be used when replacing missing values. Secondly, a general assumption can be made about the insulation of concrete structures: the thickness has a mean value of 85 mm (equivalent mineral wool) which will replace missing data in such cases. Thirdly, most of the designs have a wooden floor structure and an insulation thickness of 200 mm. The remaining missing values are therefore replaced with that value.

Impregnation of wooden materials (X_{13}) and *ground material* (X_{21}) seem not to be included as "basic" elements in the crawl-space when viewing the retrieved data. These two parameters are therefore assigned a neutral value of 0.5 when missing. The situation is the opposite when it comes to vapour barrier (X_9) and presence of organic waste (X_{15}) in the crawl-space. Parameter X_{12} (foundation wall - inorganic) is considered to be present when not mentioned which means a missing replacement value of 1. Parameter X_{15} is given the same missing value but with the opposite meaning – no organic waste present. A value equal to 1 means that the parameter is supposed to have a positive effect on the performance.

The output (performance indicators); smell, mould and rot/rust have received a *healthy* value when they are not mentioned in the reports, which is taken to mean that no smell, mould or rot/rust have been found.

3.2 DATA RELIABILITY

First of all the data is collected from reports that have been created by people with different backgrounds and thereby different knowledge about crawl-spaces. This might influence what information is important for them to include in the reports. Furthermore, the assessment of the crawl-space condition is very subjective: for example, smell can be perceived differently. There can also be different conceptions about the level of visible mould growth. In addition mould growth can also be disguised by the surface it is growing on. Black mould spots are more difficult to detect on a dark surface than on a light coloured surface.

Technical descriptions from the building permit were included in the majority of the reports from the SSN. These descriptions turned out in some cases not to agree with what was really built. For instance the drainage system was found to be missing when the moisture damage consultant had a test pit dug.

Therefore, both input data and performance indicators are associated with some level of uncertainty, which of course also influences the data quality.

4 RESULTS

4.1 ARTIFICIAL NEURAL NETWORKS OR CASE BASE REASONING?

Two of the most dominant artificial intelligent systems in the literature are artificial neural networks (ANN) and case base reasoning (CBR), both of which have been investigated here. The aim was to find out which one of them is most appropriate for the problem described in this research project. The crucial difference between these two systems is the level of required knowledge of the area the systems are applied on. The major downside with CBR is that it requires an expert within the knowledge domain to structure the system correctly. This means that the problem at hand can not be flawed, containing unknown attributes. Using the CBR system requires the developer to be very knowledgeable about the area in order to be able to organise the parameters in terms of importance. This is not the case with ANN as this is a self learning system. It is, however, necessary to have enough information to capture the parameters influencing the problem area. The CBR system is based on a library of different cases while ANN creates a memory structure based on what it has learned from the cases presented during a training process. According to Leondes (2002) the application of ANNs has great value when it is difficult or impossible to uncover relationships. The method is also helpful even when the data is noisy or incomplete. One drawback though with ANN is that it acts as a black box where the trained ANN is not available for inspection. More on the comparison between these two systems can be found in Paper II. However a short and simplified description of ANN and CBR will be given below.

An ANN consists of a structure of interconnected nodes arranged in layers, Figure 3. Between the nodes there are weights wherein the result of the training process is captured. In order to train the ANN a number of cases with identified outcomes are needed. The input layer represents the parameters describing the problem which in this case are the parameters influencing the condition of the outdoor ventilated crawl-space. The outcome, in this case the condition of the outdoor ventilated crawl-space, is represented in the output layer. Each case is presented for the chosen ANN and is propagated through the entire network where the output signal is compared with the real outcome. The calculated error is thereafter back-propagated in the network adjusting the weights. This procedure is repeated until the error has reached an acceptable level. This has to be tried on several ANN structures with different amount of nodes and hidden layers in order to find the best ANN design.

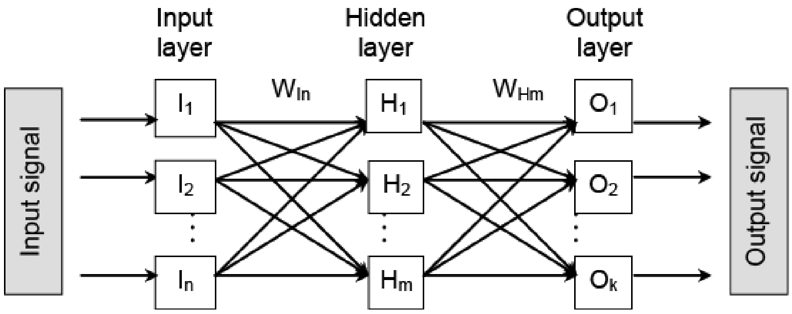


Figure 3 Artificial Neural Network structure.

A CBR system is created through a library of different cases that is arranged in a certain manner. When a prediction case is presented the CBR system looks for the most similar case in the library of cases to match the current case as close as possible. If it cannot match the case exactly an adaption process takes place with the most similar cases found in the library. In order for the CBR system to prioritize and adapt a final suggestion the importance of each parameter must therefore be given when creating the CBR system. This is why the knowledge of the prediction area of interest must be better than when applying ANN.

4.2 ANN DESIGN, TRAINING AND VALIDATION RESULTS

Separate ANNs have been trained on each performance indicator. The ANN designs were tried on each performance indicator using both one and two hidden layers (see Paper III). In all, 27 different ANN designs where tried. The training was stopped through a cross validation. During training the prediction performance of the ANN was tested against 38 cases not used in the training set. When the error increased in the validation cases, the training of the ANN was terminated. The best performing designs and the results are presented in Table 4.

Table 4. ANN prediction performance

Performance indicators	Corrrect classification [%]	ANN Design
Smell	100 (38/38)	17 + 17
Mould	76.3 (29/38)	5 + 10
Rot/Rust	92.1 (35/38)	2 + 17

As can be seen from Table 4, the trained ANNs ability to predict the performance of the outdoor ventilated crawl-spaces is rather good. The validation cases were chosen to represent a wide range of designs. Besides the basic design (Figure 1) and cases with different geographical locations there were cases deviating from the basic design as follows:

- No vapour barrier
- Floor heating
- Increased/decreased insulation in the floor structure
- Insulation of counter floor
- Insulation foundation walls
- Mechanical ventilation
- Concrete structure
- Organic foundation walls
- Impregnation of wood material

The reliability test of the ANN, through a parameter study (Table 7) spanning 50 years, highlighted some uncertainties and flaws in the trained ANN. Foremost of these flaws, is the mould performance indicator which displays a reversed degradation process.

Table 7 Chosen parameters in the study

	Parameters
0	Case base
1	Removed vapour barrier
2	Removed vapour barrier + Low permeability in ground
3	Removed vapour barrier + High permeability in ground
4	Mechanical ventilation
5	Outside urban areas
6	Foundation wall organic
7	Decreased insulation floor structure
8	Increased insulation floor structure
9	Floor heating
10	Insulation foundation wall
11	Insulation counter floor
12	Thermal capacity (low permeability)
13	Impregnation of wood material
14	Load carrying structure of concrete
15	No capillary breaking layer
16	No drainage system ground
17	No drainage system roof
18	Ground inclination to the house

Other displayed uncertainties are associated with the results of 1, 2, 3, 12, 13, 15, 16, and 17 in Table 7. All of the other elements in the parameter study agree in general with the expected results so far. The same goes for the results in the second part of the reliability test concerning different geographical locations. For more detailed results see paper V, Figure 2-5.

4.3 CURRENT PRACTICE IN THE DESIGN STAGE

The interview study explored how consultants evaluate the performance of a building regarding moisture safety and to what extent knowledge about building physics theory is being used during the design process. The eight interviewed consultants had various backgrounds regarding experience and education, se Paper I, Table 1.

Experience was referred to as a decision support. However, it was admitted that they do not get adequate feedback on past projects which decreased the value of experience. Some of the consultants also included safe and well-known designs in the concept of experience.
Few tools were used amongst the consultants and those who did either had very basic tools or built their own design tools. Those who did not use any computer based tools stated that they are:

- too costly to buy
- too difficult to learn
- requires to much time (to run the simulations)
- not enough time allocated to evaluate the performance

In general, the most desired feature of any computer-based tool amongst the consultants is that it needs to be easy to use in terms of low level of input and output.

Other interesting results in the interview study concerned the level of awareness and confidence of the interviewed. A higher level of education appeared to be related to and increase their level of awareness regarding the whole design process, complex performance issues and current levels of the technology base (Paper I, Figure 1). Furthermore, those with less education indicated a great deal of confidence about the complexity of building physics. Confidence is defined as the strength of a person's belief that a specific statement is the best or most accurate response (Peterson and Pitz, 1988).

4.4 PERFORMANCE PREDICTION CHALLENGE

The questionnaire was sent out to 110 people, which resulted in a response level of 50% and had the following distribution by profession:

– Engineering consultant	40% (22)
– Moisture damage consultant	29% (16)
– Moisture expert	31% (17)

In the performance prediction challenge the ANN managed to make 14 (93%) correct predictions of a possible 15, whereas the respondents provided an average of 7.5 (50%) correct predictions. The results of the respondents ranged between 3 and 12 correct predictions.

The prediction results of the questionnaire where cross analysed by profession, educational background, and years of experience regarding moisture damage inspection. The results did not present a distinct indication that would verify any of these correlations, see Paper IV, Table 5-8.. There was however one small exception for those with an additional moisture education who managed to provide 10% better prediction results than those without additional moisture education. Most surprising was that experience of moisture damage inspections did not lead to better prediction results.

Is this kind of tool wanted? According to this survey, the respondents thought that all suggested moisture design decision support tools should be improved or developed (Paper IV, Table 9). Highest on the list though was a system to capture experience (93%).

5 DISCUSSION

In the study it was indicated that ANN was better qualified for this problem than CBR (Paper II). This AI system was finally chosen because of the self-learning feature which does not require full knowledge of the area. The outdoor ventilated crawl-space has been explored in several research projects and the knowledge of the performance of the design is therefore rather extensive. However, it is not good enough to be applied on the CBR-system as this requires the system builder to assign weights of importance on each parameter influencing the performance of the design. The drawback with ANN on the other hand is the black-box behaviour where the user cannot trace the reasoning process (Chua *et al*., 2001).

Secondary data from real life was used to train the ANN and despite poor data quality the trained ANN managed to deliver good prediction results (Paper III). The parameter study on the other hand highlighted some flaws in the prediction results, of which the most conspicuous concerned the reversed mould process (Paper V). There are several explanations for this behaviour of the trained ANN, but the most likely has to do with the composition of training cases where the average age increases with increased mould growth. However, this might not be the sole explanation as mould growth is a rather complex process. According to Hukka and Viitanen (1999) it is possible for wood to partly recover from mould infestation during dry periods when the mould activity is decreased. It is therefore difficult to dismiss the displayed reversed degradation process as completely incorrect.

The deviating results of 2, 3, 12, 13, 15, 16, and 17 in Table 7 can be explained by one common factor, that is the high level of missing values, exceeding the critical level of 20% (Famili et al, 1997). Some of them might, however, have a natural explanation. For instance, the absence of drainage system in ground does not necessarily impair the performance. If the ground is highly permeable there is no need to arrange a drainage system (Nevander and Elmarsson, 1994). The deviating result when removing the vapour barrier (nr 1 in Table 7) is a result of the parameters with a high level of missing values. If the ground condition is uncertain it is difficult for the ANN to predict the performance if the vapour barrier is removed.

Having a vapour barrier on the ground is recommended in the literature but the ANN suggests that this decreases the performance. There are cases found in the training data without the vapour barrier displaying healthy conditions. Removing the vapour barrier does not necessarily always result in a more humid climate in the crawl-space. According to Kurnitski (2000) materials with high moisture capacity can improve the moisture condition in the crawl-space. If there is a dried out layer of clay in the ground this might have a moisture buffering effect.

The results in the second part of the reliability test concerning different geographical locations seem to be correct. If this is the case the crawl-space design must be adapted to local conditions and geographical locations. In all, the parameter study indicates that there are components of skewed data, missing data and possible unrevealed internal relations or underlying mechanisms that makes it difficult to completely dismiss the prediction performance of the trained ANN in this instance. In addition, the parameter study goes far outside the ranges of the retrieved training data. Each design case does not, for instance, have representation over the whole age span, 0-50 years, given in the study. This has forced the ANN to extrapolate the predictions. Having access to complete and flawless data could have shed some more light on why, for instance, some outdoor ventilated crawl-spaces after many years still remain healthy while others don't.

As a final test, the trained ANN prediction ability was compared with the ability of engineering consultants, moisture experts and moisture damage consultants (paper IV). The results indicated that the ANN did provide far better results than the respondents in the survey. This can be caused by other things than just pure skill. The low prediction ability of the respondents might have been caused by a concentration drop as the first cases had a better hit rate than the last cases (see Table 4, Paper IV). Unmotivated respondents can also have influenced the results. As stated in the results experience of moisture damage inspections did not provide for better prediction results. A probable cause can be the high number of parameters which can make it difficult to handle. According to Halford *et al.* (2005) it is impossible for the human brain to process information with more than four variables.

The application of ANN on this kind of problem provides the possibility to create a system that can capture real life experience; experience that in the survey (paper I) was referred to as decision support even though it was rarely or ever followed up. In the prediction challenge survey (paper IV) this kind of tool was requested. The question is, though, if the drive to use such a tool is strong enough if it were developed further and introduced? As noticed in the results of the interviews (paper I) the confidence can be an obstacle. If you think your ability for the task is good enough the motivation to enhance your knowledge area or incorporate new tools for improved decision support can be rather small. That previous tools have been perceived as too costly or to difficult to use may also stop prospective users from trying new tools.

6 CONCLUSIONS

Is it possible to apply Artificial Intelligence, in this case Artificial Neural Networks, as an early decision support system based on real life experience?

The results so far are promising but ANN, based on real life experience, must be tested further with better training data. It would be even better if the data retrieval was designed for this purpose. The parameter study highlighted some uncertainties which largely could be related to data quality, composition and design of parameters. The current archives of SSN and Anticimex are therefore not sufficient to meet the necessary requirements of data quality, in particular, because of the level of missing values. In addition, the performance indicators must be improved in order to avoid noisy data. The trained ANN based on used secondary data is therefore not suited for any further conclusions regarding the performance of the outdoor ventilated crawl-space design to be drawn.

What makes this ANN approach so interesting is the fact that there is a gap to be filled regarding different types of moisture design methods, in this case *well tried solutions*. This approach can be of assistance in the decision process during the early stages of design when the accumulated project information still is small and the detail level is rough. Different design options can be assessed and compared. Previous experience can be captured in the structured and systematic manner as displayed in Figure 4. The method also allows the knowledge to be distributed easily. Another benefit is that both input and output data are straightforward and easy to grasp. The decision maker can easily learn the true outcome of a certain design from experience and how it evolves over time. To compare alternative designs by humidity and temperatures alone, which is offered by traditional tools, is more difficult and requires more knowledge to be handled. Acquiring a large amount of primary data for neural network training in this context requires extensive resources. Considering the costs due to moisture damage every year, the use of this tool should fairly quickly pay off for the building industry and society.

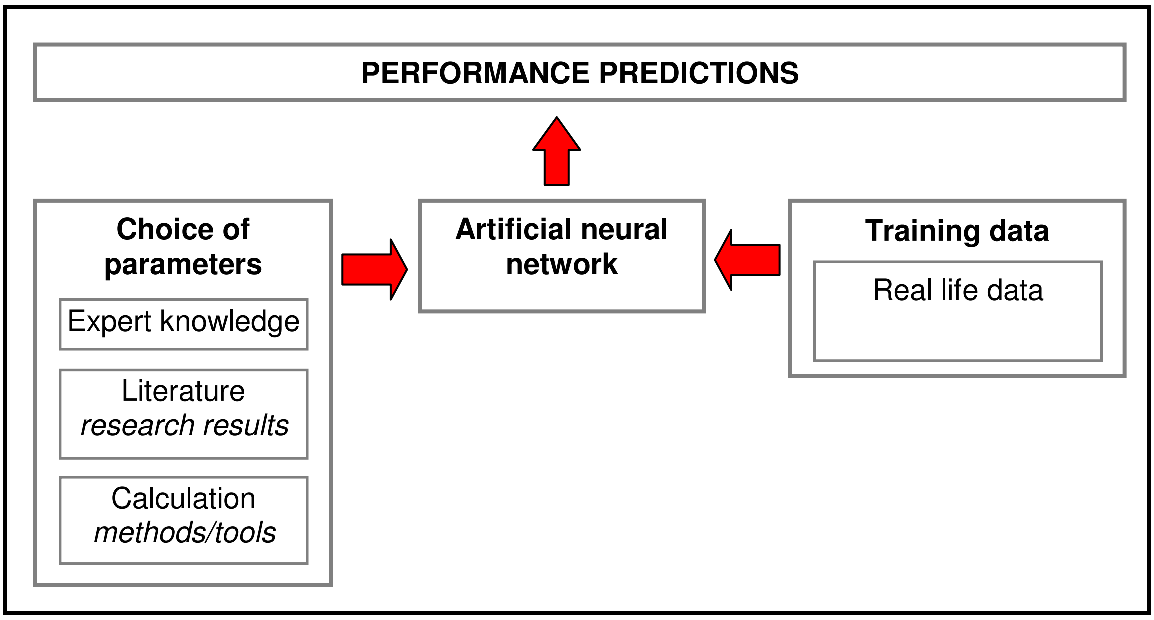


Figure 4. Performance prediction by artificial neural networks.

Amongst professionals, experienced based decisions seem to be common in the design stage even though follow ups are rarely, if ever, done. The questionnaire also indicated that there is a demand for a system to capture experience regardless if they are engineering consultants, moisture damage consultants or moisture experts. In the prediction challenge the trained ANN performed better than any of the respondents which in turn indicates the potential of such a tool for practical use. Design decision based on personal experience, even if it was followed-up, is not reliable – especially when the design includes a large number of parameters to be considered. It is therefore not recommended to use the approach of *well tried solutions* without any computational aid. The question is, however, if a tool based on the ANN-approach can be implemented for practical use still remains to be answered. Can confidence and awareness be obstacles in an implementation process?

7 FURTHER WORK

To explore this subject further would firstly, require better quality of the training data. This is necessary as the results in this thesis could neither completely support nor reject the suggested performance prediction method. Listed below are the important aspects learned in this research project, points which must be considered for further work in order to be able to improve the data retrieval process:

- The design and composition of input data also with respect to data retrieval conditions allowing robust, consistent and reliable data
- A clear and unambiguous sampling strategy should be considered
- Output data (performance indicators) can be sensitive to seasonal changes
- Consider limiting the age span of retrieved cases as it would reduce the required amount of cases
- For the same reason as above - narrow down the variations of the design
- Design separate ANNs for each performance indicator

One aspect not dealt with in this research project is the influence of workmanship. Some designs can be more sensitive to this aspect than others, a fact that should be considered.

The outdoor ventilated crawl-space used as a test case in the application of ANN is a rather open design allowing a non-destructive data retrieval. This is not the case for most other construction designs, which is why the design of performance indicators (output data) should be adapted to each construction design.

One of the more important features of this kind of tool, is that the input and output both are rather simple and that complex relations and knowledge are built in. An alternative to real life data is to use the results of numerical calculations based on reliable physical models. In this way, this complex knowledge is packaged into a tool accessible for those with less theoretical knowledge. The challenge is to make the output (results) easy to understand and grasp.

8 REFERENCES

Adan, O., Brocken, H., Carmeliet, J., Hens, H., Roels, S. & Hagentoft, C.-E. (2004) Determination of liquid water transfer properties of porous building materials and development of numerical assessment methods: Introduction to the EC HAMSTAD project. *Journal of Thermal Envelope and Building Science* 27(4):253-260.

Anticimex (2004) Fukt- och mögelskadad krypgrund. Ett faktum för 175 000 svenska villaägare. Sweden: Anticimex

Barrett, D. (2000) *The renewal of trust in residential construction.* Part II Volume 2. Victoria, BC: Ministry of Social Development and Economic Security

Barrett, P. and Stanley C. (1999) *Better construction briefing*. Oxford: Blackwell Science.

Chua, D.K.H., Li, D.Z. and Chan, W.T. (2001) Case-Base Reasoning approach in bid decision making. *Journal of Construction Engineering and Management* 127(1):35-45

Demuth H. & Beake M. (2000) *Neural network toolbox: for use with MATLAB, user's guide.* Natick, Mass.: MathWorks

Famili A., Shen W.-M., Simoudis E. & Weber R. (1997) Data preprocessing and intelligent data analysis. *Intelligent data analysis* **7**(1-4):3-23.

Halford G. S., Baker R., McCredden J. E. & Bain J. D. (2005) How many variables can humans process? *Psychological Science* 16(1):70-76.

Hendriks, L. and Hens, H. (2000) *Building envelopes in a holistic perspective.* IEA Annex 32, task A. Leuven, Laboratorium Bouwfysica.

Hukka, A. and Viitanen, A. (1999) A mathematical model of mould growth on wooden material. *Wood Science and Technology* 33(6):475-485

Kurnitski, J (2000) *Humidity control in outdoor-air-ventilated crawl spaces in cold climate by means of ventilation, ground covers and dehumidification.* Report A3. Espoo: Helsinki University of Technology.

Lai K.K., Wang S. & Yu L (2006) An integrated data preparation scheme for neural network data analysis. *IEEE Transactions on knowledge and data engineering* 18(2):217-230.

Leondes, C. T. (2002). Expert systems: the technology of knowledge management and decision making for the 21th cenutry. San Diego, Academic Press.

Lisø, K. R. (2007) *Klimatilpasning av byggnader.* SINTEF Byggforsk.

Nevander L. E. & Elmarsson B. (1994) *Fukt handbok.* Stockholm: Svensk Byggtjänst.

Peterson, D.K. and Pitz, G.F. (1988) Confidence, uncertainty, and the use of information. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14:85-92

Qin Y., Zhang C., Zhang J., Zhang S. & Zhu X. (2006) *Semi-parametric optimization for missing data imputation.* Applied Intelligence 27(1):79-88.

Sandin, K. (1998) *Fuktdimensionering ger fuktsäkrare byggnader.* Stockholm: Byggforskningsrådet.

Tolstoy, N. (1994) *The condition of buildings. Investigation methodology and applications.* Stockholm: Royal Institute of Technology (TRITA-BYM

Thomas, B. *(2003). Framtidens Intelligens. En bok om AI, återkoppling och neurala nätverk.* Lund: Studentlitteratur.

Weiss, K. & Sullivan, S. (2001). The health economics of asthma and rhinits: assessing the economic impact. *Journal of Allergy and Clinical Immunology* 107(1), 3-8

Yverås, V. (2002) *The Performance Indicators as Decision Support Tool.* Göteborg: Chalmers University of Technology

APPENDIX A

INTERVIEW QUESTIONS

Key Question 1:
How would you describe the design process of a building?

Design process

1. How much time is spent evaluating buildings? (moisture, lifetime calculations, energy use, thermal comfort, ventilation etc).
   Key: to determine the level of importance of these issues

2. How much time would you like to spend?
   Key: are they aware of the issues or are they restricted in some way?

3. Hindrances?
   Key: Describing the restrictions.

4. How are the various aspects integrated to get the whole picture?
   Key: determine if there is any form of co-operation between the different consultant groups.

5. If you had more money on a project for the evaluation phase, where would you spend it? Why? (Get them to elaborate on the answers.)
   Key: To see if they are aware of building physics aspects. Do they really think it is important?

Key Question 2:
What are the most important performance requirements when designing a building?

Performance requirements

1. What performance requirements do you have and how do you check that they are evaluated?
   Key: Shows if they use the performance concept.

2. Do the customers have specific requirements?
   Key: Shows the level of knowledge of the clients.

3. Does the consultant ask the clients about other requirements (above the minimum required by law)?
   Key: Shows if they understand the performance concepts.

4. Has the clients ever suggested any other solutions that the consultants disagreed with?
   Key: How have they dealt with such clients? How do they show clients their mistakes in the design of the building? Do they point out flaws if the client has approved/designed the design?

5. Do you ever educate your clients on the importance of evaluating a building's long-term performance?
   Key: Empowering the client, is it done?

6. Do the clients assume it's the job of the consultants to evaluate the building?
   Key: What is expected of a consultant?

7. Have you ever made recommendations that would improve a building only to have them dismissed by the client because of the cost/other reasons? (Examples)

Key Question 3:
How do you evaluate the performance of a building?

Tools

1. Hypothetically speaking, what types of decision tools would be useful to you if there were some available?
   Key: Recognise holes in the market.

2. What specific feature would you want these tools?
   Key: Defining the tools needed.

   (If not natural, steer towards performance and building physic-based tools and note reaction)

3. What are the benefits to your company in using these tools? Why not?
   Key: Identify obstacles for the implementation phase of our projects.

   If possible: - Do you currently use any tools?
   Key: Identify the 'good' software on the market.

      Yes – Which ones do you use? Describe their strengths and weaknesses.
      Key: Use this information to improve our own tools.

      No – Is there a reason to not use the tools ex. Are they too difficult to use, do they take too much time to use? Are the results from the current tools worthless?
      Key: Use this information to improve our own tools.

Key Question 4:
What influences do economical aspects, such as market conditions and market trends have on the design on the building?

Economical aspects

1. What are the current market conditions (generally)?
   Key: Historical background.

2. What are the current trends in regards to
   a. customer demands and
   b. industry demands?
   Key: Historical background

3. What do you have to gain by using performance and building physics based tools?
   Key: Identify obstacles for the implantation phase of our projects.

Key Question 5:
Moisture problems are becoming more popular in the media. As you know, it is part of building physics theory. Do you feel comfortable working with building physics issues (heat, ventilation and moisture issues)?

Level of competence
Comfortable

1. Do your co-workers feel comfortable with this?

2. How does the industry in general feel? Nervous?

3. What is needed to solve this problem?

APPENDIX B

QUESTIONNAIRE – PERFORMANCE PREDICTION CHALLENGE (PAPER IV)

Invitation letter

This is a survey that has been developed through a research project on the Chalmers University of Technology which has been financed by Formas, SBUF and CMB.

The main part of the questionnaire contains a test where you are to asses the condition of different crawl spaces and ages. As the study would benefit a high response level I hope that you will find the time to answer this web based questionnaire. All the respondents are guaranteed to be anonymous in the survey.

Looking forward to your participation in the survey before the 5th February, 2008.

With best regards,

Veronica Yverås
Chalmers University of Technology

Question 1

Backward

How many employees does your company have?

- Fewer than 5
- 5-35
- 36-100
- 100-500
- More than 500

Question 2

Backward

Which category of profession agrees best with your current work?

- Engineering consultant
- Moisture damage investigator
- Moisture expert

Question 3

Backward

How many years of experience do you have within the following areas?

	0	1-5	6-10	10-15	15-20	>20 years
Engineering consultant						
Moisture damage investigations						
Research & Development						

Continue

Question 4

Backward

Which of the following educations do you have?

- High school engineer
- Bachelor of science
- Master of science
- Licentiate of engineering / PhD
- Other

Question 5

Backward

Do you have any additional education regarding moisture issues?

When answering this question you can choose more than one option.

- Byggdoktor (House Doctor)
- Diplomerad fuktsakkunnig (Moisture adviser)
- Auktoriserad fuktkontrollant RBK (Authorised moisture inspector)
- Other

Continue

Question 6

Backward

What other additional moisture education do you have?

Continue

Question 7

Backward

There is a need to improve/develop following tools in order to deal with moisture issues in your profession.

	1 Don't agree at all	2	3	4	5	6 Completely agree	Don't know
Handbooks	☐	☐	☐	☐	☐	☐	☐
Systems to capture experience	☐	☐	☐	☐	☐	☐	☐
Product information from material producers	☐	☐	☐	☐	☐	☐	☐
Moisture calculation tools	☐	☐	☐	☐	☐	☐	☐
Moisture education	☐	☐	☐	☐	☐	☐	☐
Guidelines during design	☐	☐	☐	☐	☐	☐	☐

Continue

Question 8

Backward

ASSESSMENT OF CONDITION

You are in a situation where you are to assess the condition of 5 different outdoor ventilated crawl spaces with various ages. In other words - what is their condition within a designated amount of years?

The condition assessment will be done on the basis of 3 categories; smell, mould and rot. Only one answer for each category is allowed. **This means that you will predict the condition for each case - all with different ages.**

Each case will be displayed by a table where marked fields shows the design composition of the crawl space. Empty fields means that these solutions or materials are excluded. This is all the information that is available for you!

You are free to solve the predictions with any tool you like.

Good luck!

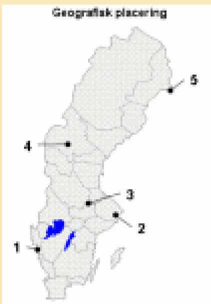
Continue

Question 9

Backward

Case 1

Crawl space design: Age 31 years	
Capillary breaking layer in ground	x
Drainage system roof	x
Drainage system ground	x
Surrounding ground sloping from building	
Counter floor insulated	x
Inside foundation wall insulated	
Mechanical ventilation	
Vapour barrier on ground	x
Inorganic counter floor	x
Inorganic slide protection in foundation wall	x
Floor heating	



Click the image for full size

No smell
Microbiological smell

No visual mould
Local spots of mould
Light mould growth in major part
Extensive/rich mould Growth

No rot
Rot on surface
Rot in depth

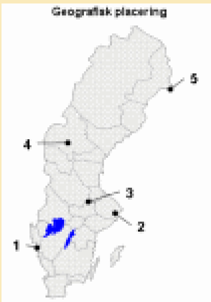
Continue

Question 10

Backward

Case 2

Crawl space design: Age 20 years	
Capillary breaking layer in ground	
Drainage system roof	X
Drainage system ground	X
Surrounding ground sloping from building	
Counter floor insulated	
Inside foundation wall insulated	
Mechanical ventilation	x
Vapour barrier on ground	
Inorganic counter floor	
Inorganic slide protection in foundation wall	X
Floor heating	



Click the image for full size

No smell
Microbiological smell

No visual mould
Local spots of mould
Light mould growth in major part
Extensive/rich mould growth

No rot
Rot on surface
Rot in depth

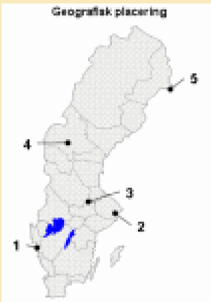
Continue

Question 11

Backward

Case 3

Crawl space design: Age 17 years	
Capillary breaking layer in ground	X
Drainage system roof	X
Drainage system ground	X
Surrounding ground sloping from building	
Counter floor insulated	
Inside foundation wall insulated	x
Mechanical ventilation	
Vapour barrier on ground	
Inorganic counter floor	
Inorganic slide protection in foundation wall	X
Floor heating	



Click the image for full size

- No smell
- Microbiological smell

- No visual mould
- Local spots of mould
- Light mould growth in major part
- Extensive/rich mould growth

- No rot
- Rot on surface
- Rot in depth

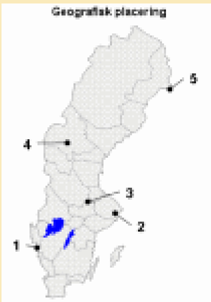
Continue

Question 12

Backward

Case 4

Crawl space design: Age 15 years	
Capillary breaking layer in ground	X
Drainage system roof	
Drainage system ground	X
Surrounding ground sloping from building	
Counter floor insulated	
Inside foundation wall insulated	
Mechanical ventilation	
Vapour barrier on ground	x
Inorganic counter floor	
Inorganic slide protection in foundation wall	X
Floor heating	x



Click the image for full size

- No smell
- Microbiological smell

- No visual mould
- Local spots of mould
- Light mould growth in major part
- Extensive/rich mould growth

- No rot
- Rot on surface
- Rot in depth

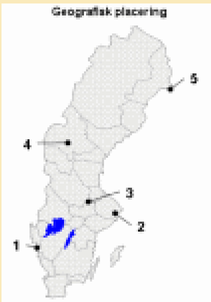
Continue

Question 13

Backward

Case 5

Crawl space design: Age 18 years	
Capillary breaking layer in ground	
Drainage system roof	X
Drainage system ground	
Surrounding ground sloping from building	
Counter floor insulated	
Inside foundation wall insulated	
Mechanical ventilation	
Vapour barrier on ground	x
Inorganic counter floor	
Inorganic slide protection in foundation wall	
Floor heating	



Click the image for full size

- No smell
- Microbiological smell

- No visual mould
- Local spots of mould
- Light mould growth in major part
- Extensive/rich mould growth

- No rot
- Rot on surface
- Rot in depth

Continue

Question 14

Backward

The crawl space cases were easy to assess.

- 1 Don't agree at all
- 2
- 3
- 4
- 5
- 6 Completely agree

Question 15

Backward

What were your assessments based on?

Continue

Question 16

Backward

Kindly complete your contact information below.

Company name:
*
First name:
*
Last name:
*
Address:
*
E-mail:

Continue

Thank you for your cooperation!

Chalmers
Department of Civil and Environmental Engineering

Veronica Yverås
veronica.yveras@chalmers.se

PAPER I

A SWEDISH PERSPECTIVE ON THE PREVENTION OF MOISTURE PROBLEMS DURING THE DESIGN PHASE

Burke, S. and Yverås, V. (2003) *Nordic Journal of Surveying and Real Estate and Research* 1(1):102-113.

A Swedish Perspective on the Prevention of Moisture Problems During the Building's Design Phase

Stephen Burke [1] **and Veronica Yverås** [2]
[1] Department of Building Physics, Lund Institute of Technology,
Box 118, 221 00, Lund, Sweden
stephen.burke@byggtek.lth.se
[2] Department of Building Physics, Chalmers University of Technology, Sweden
veronica.yveras@jm.se

Abstract. *Moisture problems in buildings are increasingly being reported in the mass media in Sweden, often leading to some controversial stories about companies and their building processes. Using building physics and building performance principles during the design stage can often prevent most problems from occurring. One of the big questions is, with all the available knowledge about designing a building, how can these problems still be occurring in new buildings? This paper explores this question by interviewing some engineering consultants on how they evaluate the performance of a building, and to what extent knowledge about building physics theory is being used during the design process to prevent moisture problems from occurring. It was found that building physics is not used extensively in the building industry due to many reasons. The lack of good design tools and the fact that clients do not request it are two main reasons. However, it was revealed that clients do not request it because they either have no interest in spending the extra money for a better design, or they do not know it is optional and just assume everything is taken account of in the final design. Furthermore, the consultants do not advise them on the available options applicable for their particular design. Due to the method used to analyse the interviews, an unexpected relationship between education level and their perceived level of awareness of building performance issues emerged. It appears that the higher the level of education of the consultant, the more they are aware of the impact of performance issues in a building's design. Their experience level does not appear significant in this relationship, however this cannot be proven and will require more studies to verify.*

Keywords: *building physics, building performance, interviews, tools, consultants, education, and economics*

1 Introduction

Moisture design appears to be a growing trend in Sweden. This can be explained by the attention from mass media that various projects around Sweden have been getting. Specifically, projects involving mould in buildings and moisture damage in newly constructed buildings, largely multi-family dwellings (Jelvefors 2002; Luthander 2001). This trend is increasing because the media has brought it to the attention of the public that the consultants do not perform a moisture analysis on a building's design during the design phase. The consultants admit that clients do not request moisture design because the clients assume that it is included in the normal design process (Arfvidsson and Sikander 2002, p. 14).

Building physics in Sweden is defined as the study of the transport of heat, moisture, and air through a building's envelope in relation to both the indoor and outdoor climate (Hagentoft 2001). It is a key area in the development of energy efficient, healthy, moisture safe and durable buildings. It is this field of science that focuses on the prevention of moisture problems during the design phase of a building. Please note that the Swedish definition of building physics does not include lighting and acoustics, unlike most other countries around the world.

In many countries, architects are responsible for the design and detailing of a building. In the Swedish building industry it is common that the architects are only responsible for the form and shape of a building and engineering consultants are responsible for the technical specifications. Recently, Sweden has seen an increase in the amount of mass-media attention that problematic buildings are getting; even to the point of being damaging for the companies involved in all phases of the construction (Luthander 2001; Jelvefors 2002; Samuelson and Wånggren 2002). One of the big questions is, with all the available knowledge about designing a building, how can these problems still be occurring in new buildings?

The aim of this paper is to explore this question by interviewing some engineering consultants on how they evaluate the performance of a building, and to what extent knowledge about building physics theory is being used during the design process to prevent moisture problems from occurring.

The driving forces behind this study are two research projects that are both looking at the use of building physics based design tools for engineering consultants in the building industry. By tools we mean any aid that influences the design. Tools can be either computer or paper based in the form of checklists, graphs, tables, simulations etc.

One project, *Performance indicators as a tool for decisions in the building process*, (Yverås 2003) deals with the problem of developing a design tool that will increase the application of building physics in the early stages of design. Performance indicators can assist in this decision-making and help to avoid failures that would otherwise reduce service life. Even though knowledge about designing a building is widely available, incorrect decisions are all-too common. Consequences from poor decisions can include a reduction in service life arising from conditions such as mould growth, rot and corrosion. These conditions can be avoided, but not without the application of robust knowledge based on the principles of building physics. However, this requires more than knowledge; it demands tools that designers can understand and use. It is important, therefore, to have a clear picture of what is required of any decision support tool, which is why the interview study is important in the further development of the performance indicator tool.

The second project, *Tools for determining the economical effects of building physics aspects during the building process*, (Burke 2003) investigates, studies and quantifies the economical benefits in using the knowledge from building physics as a design and decision tool in the building process. Problems in the building process related to building physics will be identified in co-operation with the building industry. Existing calculation programs, databases, statistical inquiries will be compiled into useful, easy to use tool packages especially designed to give adequate information about the costs and risks associated with different designs. These interviews were necessary to gain insight into what extent building physics is utilised in the building industry, and what types of applications designers want that would enable them to apply building physics theories more easily to designs.

2 Method

As mentioned in the background, the two projects behind this paper are developing design tools to be used during the design phase. Information and insight was needed about the design process in Sweden as well as the types of tools that designers would want to use. Since these tools are intended for designers during the design phase of a project, we focused our information gathering on designers who will potentially have use for our tools.

Of the various methods considered – for example experimental, literature review and surveys – the latter seemed to offer most promise. Due to the nature of our enquiry, we felt that an exploratory survey was more likely to reveal the key features of the underlying problem than either of the other methods.

Questionnaires were considered as the primary method for gathering information. However they have the disadvantage of being too linear. In addition, the information generated could not be anticipated, so it was not considered appropriate to gather the information by questionnaires. Interviews were more appropriate by allowing us to be dynamic, with the ability to probe interesting information to a much deeper level than is possible by questionnaires.

The questions for the interviews were formulated around two themes. One was to get a picture of the consultants' conditions used to evaluate the performance of a building (i.e. their perception of the building process), and the second was to determine their level of comfort and experience in working with building physics issues.

To ensure that all interviews yielded comparable results, they were based on five principal questions with about 26 more specific questions. They consisted of open and closed questions that allowed us to assess various aspects of the interviewees unbeknownst to them. For example, a respondent can be assessed on his or her familiarity with the latest information and technology without directly asking. The closed questions allowed us to categorise the interviewees into predetermined categories. The five principal questions were:

1. How would you describe the design process of a building?
2. What are the most important performance requirements when designing a building?
3. How do you evaluate the performance of a building?
4. What influences do economical aspects, such as market conditions and market trends have on the design of a building?
5. Do you and your co-workers feel comfortable working with building physics issues, i.e. heat, air and moisture issues?

Interviews were conducted with eight building consultants over the span of two weeks and all consultants answered all of the questions. Two consultants declined to be interviewed because they were too busy but were positive to the interviews and recommended alternative people, whom accepted. All but one, the building physics professional, were chosen at random with no information about them prior to the interviews. It was decided to stop conducting interviews at eight because after the 5th or 6th interview very little new information was obtained.

The results were analysed based on the grounded theory approach, which "is a method for discovering theories, concepts, hypotheses, and propositions directly from data rather than from a priori assumptions, other research, or existing theoretical frameworks" (Taylor and Bogdan 1998, p. 137). In other words, there were no assumptions made as to what results we would obtain prior to the interviews.

3 Results and discussion

3.1 Relationships

Table 1 shows the profiles of the interviewees. Category refers to their general level of ability regarding the application of building physics to a design. Category A covers expert engineering consultants, category B covers the average ability expected from a building engineering consultant, and category C covers engineering consultants with very little ability. Some of the consultants indicated that experience is very important when dealing with the performance of a building. However, this was not apparent when analysing the interviews. Arfvidsson and Sikander (2002, p. 16) also found that consultants want more feedback on past projects, which supports our finding that they do not get adequate feedback on past projects, hence decreasing the value of experience. When looking at the experience level compared to the perceived level of awareness, i.e. the whole picture of the design process combined with a comprehension of complex performance issues and an awareness of the current levels of technology base, there did not appear to be any pattern. However, the level of education appeared to be related to their level of awareness. Figure 1 shows how we perceived the level of awareness for each person interviewed.

Table 1. *Profiles of those interviewed*

Category	Education hh	Experience
A	PhD in building physics	20 years
	Civil engineer + extra education building physics	15 years
B	Civil engineer	30 years
	Civil engineer	15 years
	Civil engineer	15 years
	Civil engineer	7 years
C	2-year engineering diploma	6 years
	High school	40 years

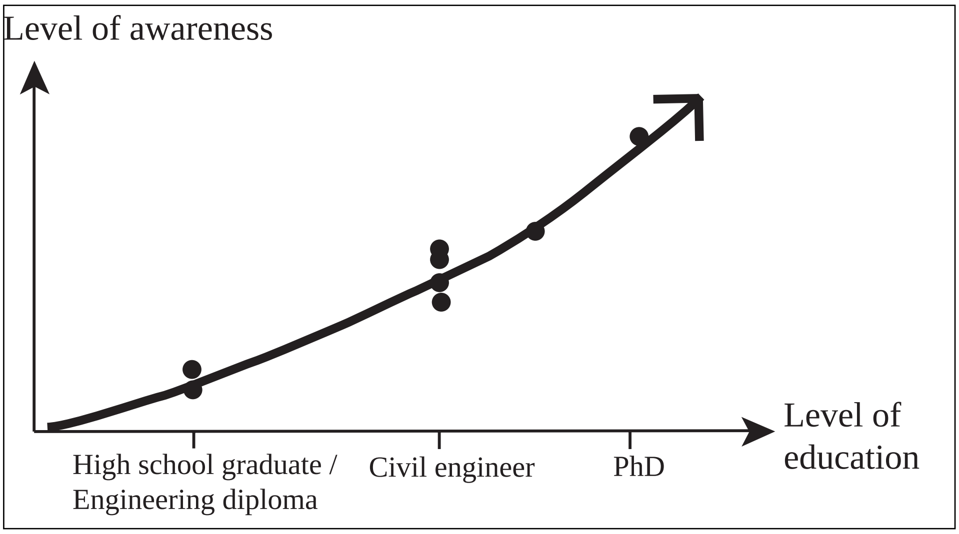


Figure 1. *Perceived correlation between level of education and awareness*

It is important to remember that the engineers in category C, and part of category B, did not have access to an expert. This could affect the results in this study since a lot of education flows internally from the experts in the companies. Other companies with experts and category C employees working together may have a totally different level of awareness due to the expert's influence. More in-depth studies would be needed to investigate this relationship further.

There also seemed to be different attitudes towards the required time directed to handle moisture control issues during design. Those within category A said that they would like to have more time whereas those in category C did not even allocate time especially for these issues. This was stated despite that they stated earlier that these issues are highly prioritised. They did however motivate it by using safe and well-known designs, referring to their own experience. However, their experience on well-known designs can be questioned as the consultants rarely have the time or the opportunity to return to, or follow-up projects that were finished 10 years ago or more. In practice, the long-term design for engineers is 2 years, according to one of the interviewed engineers.

When asked who is responsible for most of the performance problems experienced in buildings today, the consultants in categories A and part of B were also including themselves when asked. This was the opposite of the others (categories C and part of B), who blamed anyone else but themselves. These results partly agree with Arfvidsson and Sikander (2002, p. 13) who concluded that no actor in the building industry is willing to take responsibility for moisture prevention issues when designing a building.

Some of the questions dealt with how comfortable the consultant feels if they must work alone on problems dealing with building physics. In most cases the answer to this question was related to whether or not they have access to an expert in building physics. If the consultant had access to an expert, they were usually not comfortable working with these issues and usually sought advice from their expert before finalising a design. The consultants in this category acknowledged that since the media attention began, they have felt even less comfortable with these issues and rely heavily on their experts. Those without an expert in-house were more prone to saying that they felt very comfortable with building physics issues.

The group within the profession that has lower education level relies mainly on their experience. But if professionals rely mainly on experience, how do they know when there are gaps in their knowledge or whether some of their standard rules are no longer applicable (Barrett and Stanley 1999). Decisions made without knowledge of their consequences can have dire effect (Ellis and Mathews 2001).

One might easily draw the conclusion that people with less knowledge would suffer from insecurity more so than those with expert background. This was not the case during the interviews. Members of group C, showed a great deal of confidence and no worries about the complexity of building physics. Confidence is defined as the strength of a person's belief that a specific statement is the best or most accurate response (Peterson and Pitz 1988). In other words, it is a measure of how strongly they believe what they say. So far, no study has been performed that examines if there is any correlation between mistakes in design and the level of knowledge of the designer. However, there is a great deal of research, which indicates that people are often more confident than they are correct (Blanton *et al.* 2001). Blanton *et al.* (2001) states that educators may meet obstacles from people's overconfidence about their knowledge when trying to educate them. As the individual with the PhD said, "People *think* they can moisture proof a building, but they can't and I have to correct the problems later, which takes a lot of time."

3.2 Consultant/engineer and liability

Noting that moisture analysis requests began increasing after the media reported moisture problems, we began to wonder what the role of a consultant is in the Swedish construction industry and what their liabilities are. One tool used is called ABK 96 (Byggandets kontraktskommité 1996). It is a standard contract template that explains in detail how engineering and architectural consultants should conduct themselves. It also describes the limitations of liability that a consultant has. Most consultant companies use this voluntary contract to guide the consultants and also the client – consultant relationship. Each party is informed of what is expected of them by the other.

Despite this, there also appears to be some confusion around the labels of consultant and engineer for consulting companies, even though it is not spoken of. A consultant is defined as “an expert who gives advice.” (Princeton 1997a) An engineer is defined as “a person who uses scientific knowledge to solve practical problems.” (Princeton 1997b) Paragraph four (Byggandets kontraktskommité, 1996, p. 5) states that the consultant must be competent, professional and have adequate knowledge to consult in the areas of their field. However, overconfidence and lack of awareness in building physics on the part of some consultants, can cloud the issue of a consultant having adequate knowledge for building physics issues.

From the interviews, it was obvious that many consultants expect to be told what to do by the clients without informing the clients of what is available. In this way some of the consultants take on the role of engineer. This change in attitude is reflective of the traditional methods of building design consulting when a lot of information was unknown and the designs were simpler. An example was one consultant who disclosed technical solutions to example problems during the interviews that are proven to lead to mould and moisture problems in houses.

If a client is an experienced buyer or an expert client, they will have predetermined tasks and technical solutions available for the consultant since they are usually aware of all the major problems and their solutions. However, not all clients are fully informed, almost all have some weakness, for instance the science of building physics is not known by a typical client. A statement during one of the interviews, “Clients don’t know enough (about building physics-issues) to have any requirements” supports this idea.

There are occasions where poor decisions have been made that have lead to a failure in performance. This was exemplified during the interviews where one described how she strongly advised the client not to follow the architects’ direction of having the outside wall continue into the ground without a base. Two years later the predicted problems arose and the plaster closest to the ground fell off due to frost erosion. Clearly this was a case where the client was not used to handling these issues, lacked the experience to make a correct decision and the consultant failed to present the information. The reasons are considered to be due largely to the inability of design engineers to encode and present the consequences of a decision. By improving the quality of information during the design process, the client is better equipped to understand the different issues implicated in the project (Barrett and Stanley 1999). The consultant above admitted that by having real life cases to show, including a cost of the consequence, the outcome of this case might have been different.

The consultant in this case was not liable for the damages that incurred later because the consultant, firstly, recognised the problem and secondly, recorded their disagreement with the client in the protocol during the design phase. The consultant would have been liable for the damages if they did not inform the client of the problem, either voluntarily or unknowingly, i.e. was not aware of the consequences of a particular design feature. This case was not typical in that the consultant did a moisture analysis to determine the consequences.

The client usually assumes that the consultants they hired will solve all the known problems. The reality is that most engineering consultants, not all, are actually operating like engineering firms, in that they do not analyse a building from a building physical point of view unless asked specifically. Their reasoning being that changing the design requires more time, hence more money that clients are unwilling to pay. The result of this is that the minimum amount of work is done when analysing a building's design and the clients get very upset when problems occur.

One fact that they are neglecting to consider is that the cost of the building might actually decrease if the design is optimised using building physics. This could be in the way of material substitution, removing unnecessary components, or utilising a quicker construction method. In the U.K., quantity surveyors are able to calculate the cost difference of various designs. This position does not exist in Sweden so it is very difficult for engineering consultants to motivate changing the design based on building physics theory because of the difficulty in calculating the savings or extra costs associated with the changes.

3.3 Design tools

When asked what design tools were used when conducting the evaluation of a building from a building physics perspective, most replied that they did use some very basic ones. Two people, including the expert, built their own design tools from Delphi Pascal or Excel spreadsheets. Only the expert had a 'wish list' for what was desired in future tools. The others said they did not know since either their local expert uses the tools, or they did not use any.

When those who replied that they did not use any design tools were asked why, they replied that they were too costly to buy, too difficult to learn, required too much time to run the simulations, and not enough time was allocated to evaluate a building's design properly. These results follow Hien *et al.* (2000, p. 727) who found that "Most firms view the use of simulation tools as involving extra costs and effort but with little recognition and appreciation from the clients."

The most desired features of any computer-based tool, according to the consultants, were that they had to be easy to use in terms of low level of input and output data. These are statements that contradict with what is typically produced by researchers. Researchers have too often failed to deliver numerical models and tools that are user friendly and that take into account the education and expertise of the likely user (Goodings and Ketcham 2001). Hien *et al.* (2000) reveals that designers regard current tools as user unfriendly with very steep learning curves; moreover, the output generated could be extremely difficult to interpret and utilise for design decision-making. Ellis and Mathews (2001, p. 1011) also confirm this and have identified that tools of today are:

- complicated (not user friendly)
- time consuming (too much input)
- require a high level of theoretical knowledge (to make the input and to interpret the results)
- Information needed is not available during preliminary design.

Regarding the wish list of the tools the answers can be categorised after what level of education the respondents have. Those within category C had no wish list. Category B directed their interest to simplify computer programs in order to make use of such programs, whereas category A people had a bigger picture and directed the use of wish tools that could be used to persuade the clients for better performance. Examples of these are tools that can show the consequences of a chosen design in terms of reduced service life due to mould, rot or corrosion and cost analysis programs. Energy calculation, heat flow and airflow programs were not mentioned by any of the interviewees despite the fact that these topics all fall under the area of building physics.

Building industry related journals were also mentioned as being a tool that provides them with useful information. However, the interviewee did not state what specific types of journals they referred to.

In another civil engineering area, geotechnics, a trend is the growing number of experts (post doctoral) joining conventional firms instead of making a career within the university (Goodings and Ketcham 2001). This trend helps bring existing research into practice where it is most needed. Augenbroe (2002, p. 891) agrees with the idea of making more use of experts in the industry stating, “The latter trend recognizes that the irreplaceable knowledge of domain experts and their advanced tool sets is very hard to match by ‘in-house’ use of ‘dumbed down’ designer friendly variants”. This difference between having a design tool, versus having an expert in the company is significant, and this was reflected in the results of the interviews. All consultants who had access to an expert made use of them constantly, and all stated that they would be uncomfortable working with moisture control problems if they did not have access to their expert. They much prefer having the expert than using a simplified tool.

3.4 The bigger picture

Despite advances and knowledge in the construction industry in the past decades, it appears that this knowledge is not generally implemented until it becomes a requirement. This was explained by Becker (1999, p. 526) who states, "incorporation of new concepts into an existing professional activity field can be accomplished only if the right infra-structure, composed of some basic conditions, is present:

- the acting parties recognize the significance of these concepts and their contribution to improving the results of their work,
- clear routines and friendly working tools for smooth incorporation of the new concepts are available, and
- young new professionals are educated to regard the new concepts as an integral part of the profession."

These statements can be seen in the Swedish construction industry today. From the interviews, we saw that some recognise the significance of the concepts of building physics and building performance. Most indicated that there were no good design tools available for designing a performance building. Some did not even know that there were tools available on the market today.

With the third point, compliance and company tradition will quickly change the young professionals into operating like the other members of a company. Even if they want to make changes according to what was learned in school, a higher power can quickly overrule any decisions that they feel are unnecessary. The younger workers learn quickly not to make these decisions again in the future.

4 Conclusions

The interviews conducted with the engineering consultants in the Swedish construction industry suggest that experience might not necessarily be important when it comes to consultants and the topic of building physics performance. In addition, the higher educated consultants felt less comfortable and showed less confidence when working with these issues than their less educated counterparts. Their comfort and confidence levels were also inversely related to their amount of access to an expert in building physics, i.e. the more access they had, the less confident they were in working with these issues. The consultants with no expert support felt very confident and comfortable in working with these issues, however the quality of their work could be questionable due to a lack of feedback loops in the system. Awareness, education, and a view of the bigger picture are all needed to effectively deal with performance problems in the current construction industry. However, even if they possess all of these traits, there are many obstacles out of their control that can prevent an effective analysis of a building's design. Some of these obstacles include having to make do with the amount of time allocated to the analysis phase of a building, meeting the client's demands, the architect's demands, the level of competence of the consultant, whether or not they have access to an expert in building physics, and the types of tools they have at their disposal.

The interviews indicate that problems are still occurring in new buildings today because either clients do not request the correct design options, the designers do not include these options in their designs due to the extra time it takes, or the constructors disregard some basic issues which lead to problems during the construction phase. Sometimes clients do not request extra design work because they believe it increases the total cost and they will not be personally affected by the improvements, for example clients who build public housing, or apartments.

Further research is needed to determine if there is a relationship between the level of education and the level of awareness in building engineering consultants and the effect that their confidence levels have on clients.

Acknowledgements

The authors would like to thank the funding bodies for these projects, FORMAS (The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning), SSF (Swedish Foundation for Strategic Research), SBUF (Development Fund of the Swedish Construction Industry), Competitive Building, Lund Institute of Technology and JM AB. We would also like to thank all the consultants who gave their time and experience to us and all the people who have given us feedback for the improvement of this work.

References

Arfvidsson, J. and Sikander, E. (2002). Fuktsäkert byggande – Enkätstudie om kunskapsläget (Moisture-safe building – Survey of the level of knowledge) Gothenburg: FoU-Väst Report # 0205.

Augenbroe, G. (2002). Trends in building simulation. Building and Environment 37, p. 891.

Barrett, P. and Stanley, C. (1999). Better construction briefing. Oxford: Blackwell Science.

Becker, R. (1999). Research and development needs for better implementation of the performance concept in building. Automation in Construction 8, p. 525-532.

Blanton H., Carvallo M., DeHart T. and Pelham B. W. (2001). Overconfidence as Dissonance Reduction. Journal of Experimental Social Psychology 37, p. 373-385.

Burke, S. (2003). Reducing the Risk of Failure in Performance within Buildings In: ed. Atkin, B., Borgbrant, J. & Josephson, P.-E., Construction Process Improvement, p. 82-92, Oxford: Blackwell Science.

Byggandets kontraktskommité. (1996). Allmänna bestämmelser för konsultuppdrag inom arkitekt- och ingenjörsverksamhet av år 1996: ABK 96 (General rules for the consulting profession within architectural and engineering fields from 1996: ABK 96.), Stockholm: FörlagETT, p. 4-10.

Ellis, M.W. and Mathews, E.H. (2001). A new simplified thermal design tool for architects. Building and environment 36(9), p. 1009-1021.

Goodings, D.J. and Ketcham, S. A. (2001). Research versus practice in transportation geotechnics: Can we bridge the chasm? Journal of professional issues in engineering education and practice 128(1), p. 26-31.

Hagentoft, C.-E. (2001). Introduction to Building Physics. Lund: Studentlitteratur.

Hien, W.N., Pho, L.K. and Feriadi, H. (2000). The use of performance-based simulation tools for building design and evaluation - a Singapore perspective. Building and Environment 35, p. 709-736.

Jelvefors, A. (2002). Lägenheter vattenskadade på Bo01 (Waterdamaged apartment at Bo01). In Metro, Malmö.

Luthander, P. (2001). Skanskas byggslarv orsakade mögel (Skanska'sconstruction carelessness caused mould), Dagens Nyheter, Stockholm.

Peterson, D. K., and Pitz, G. F. (1988). Confidence, uncertainty, and the use of information. Journal of Experimental Psychology: Learning, Memory, and Cognition 14, p. 85–92.

Princeton University, (1997a). Wordnet 1.6 as found at http://www.dictionary.com/search?q=consultant accessed April 24, 2002.

Princeton University, (1997b). Wordnet 1.6 as found at http://www.dictionary.com/search?q=engineer accessed April 24, 2002.

Samuelson, I. and Wånggren, B. (2002). Fukt och mögelskador Hammarby Sjöstad (Moisture and mould damage Hammarby Sjöstad). Borås: Swedish National Testing and Research Institute.

Taylor, S.J. and Bogdan, R. (1998). Introduction to Qualitative Research Methods - A Guidebook and Resource. 3rd Edition, New York: John Wiley & Sons, Inc. p. 87-116.

Yverås, V. (2003). Performance Indicators as a Tool for Decisions in the Construction Process In: ed. Atkin, B., Borgbrant, J. & Josephson, P.-E., Construction Process Improvement, p. 68-81, Oxford: Blackwell Science.

PAPER II

PREDICTING SERVICE LIFE BY ARTIFICIAL INTELLIGENCE

Yverås, V. (2005) Proceedings of the 7th Symposium on Building Physics in the Nordic Countries, Reykjavík, Iceland, 13-15 June, 2005.

PREDICTING THE SERVICE LIFE BY ARTIFICIAL INTELLIGENCE

Veronica Yverås, M.Sc,
Dept. of Civil and Environmental engineering, Chalmers University of Technology
veronica.yveras@bt.chalmers.se

KEYWORDS*: Service life prediction, Neural Network, Case-Based reasoning, Crawl space*

SUMMARY*:*
The use of Artificial Intelligence (AI) has a vast area of application. In this research project its application is explored to package existing experience from physical objects, building parts, in order to predict the service life. The focus is to facilitate decision during the early stages of design in order to prevent moisture and mould problems. A prototype is going to be developed by using the crawl space design as an example why a literature review of two different AI-systems, Case-Based Reasoning (CBR) and Neural Networks (NN), has been performed. Due to the fact that the area of service life prediction is a fragmented area of knowledge, the NN was found to be the best choice. The requirement of a user-friendly tool that does not need expert knowledge to be handled also made the NN-method more favourable.

1 Introduction

1.1 Background and research context

It has been found that practitioners tend to rely on their own experience (Burke and Yverås, 2004), (Leondes, 2002), where new design problems are often solved by the reuse of similar previous cases (Andrade et al., 2003). However, their experience base is not completely documented. They have drawings, computer files, personal notes or memories but the most important ingredient is missing – the outcome. The designer has no knowledge if the design has been affected by moisture and mould problems and as long as no complaints reaches the designer the designer takes it as a confirmation that nothing is wrong. This is a very dangerous assumption to make, as the effects of moisture problems tend to occur after a longer period of time. By that time, the guarantee has expired and the customer does not have a reason to turn to the consultant company for any compensation. This is an attitude that is forced by lack of time, which later might have turned into an accepted way of working.

There are basically three sources of knowledge to assess a design or a part of a building, fig 1. In this project the focus is to capture the knowledge from one of them - the experience from physical objects, by using Artificial Intelligence and also to find out which the implementation parameters are.

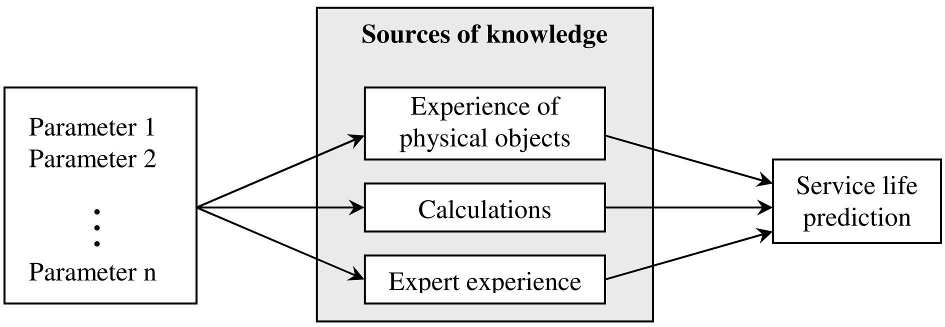


FIG. 1: Three sources of knowledge

The research questions are formulated as follows:

- How to make service-life predictions of building parts by using the technology of Artificial Intelligence? (1)
- What are the conditions to attain an implementation of an AI-system in decision processes concerning moisture related service life issues? (2)

1.2 Research objectives
The aim is to examine the application and implementation possibilities of an AI-technique to predict the service life, by capturing experience from physical objects. The specific objectives of the research are:

- (A) Identify an AI-systems that could be of interest for the research task and choose best suited.
- (B) Exemplify the chosen system by developing a prototype.
- (C) Examine the potential of an AI-tool in a design process with respect to users, needs and reliability of the tool.

However, only the result of the first objective (A) is here presented in this paper.

1.3 Scope of the research – objective A
The scope of the research is directed at exploring an AI-tool to predict the service life of a design relating to moisture and mould problems. A crawl-space design will be used in order to build a prototype that can be tested on users further on in the project. This means that it can be difficult to draw some conclusions on how it will respond to a less extensive building part. However, this might provide some general conclusions about data extraction of building parts in order to generate service life knowledge. The system will not include the aspects of maintenance or workmanship, which to some extent affects the service life.

1.4 Methodology – objective A
In order to choose an AI-system, a literature study to explore different AI-systems was carried out. The goal was to find two systems and then compare them to each other. Also a survey of previous applications within the construction area was made. The intention was to find out how different AI-applications behaved, which might be helpful in selecting an AI-system. Prior the decision of which is the more appropriate system, some general requirements had to be stated. With support from literature, these requirements were made with respect from a user perspective and the nature of the area of interest. Some consideration has to be taken to which building part that is chosen to be represented in the prototype. In this project the choice fell on the crawl space. It is favourable design as the service life is not heavily dependent of maintenance. By this it is possible to minimise the effect of maintenance on the service life

2 Results
AI has been defined as a computer or a software system, with built-in-knowledge that has the ability to imitate a human expert within an area (Thomas, 2003). There are mainly two branches of AI, Expert systems and Machine learning, fig 1.

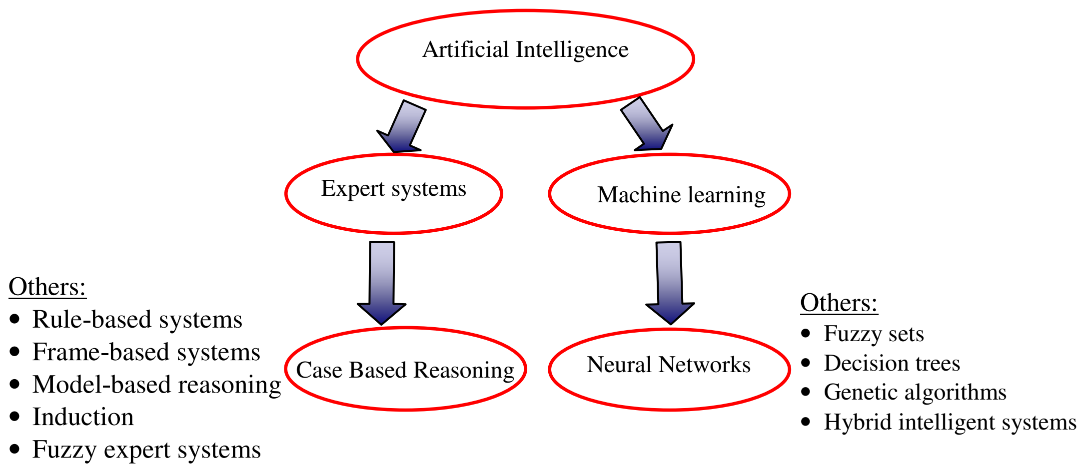


FIG. 2: Different systems of Artificial Intelligence

Expert systems are defined as extracting knowledge from human experts. This knowledge is a combination of a theoretical understanding of the problem and their own experience. This reliance on human expert knowledge is a major feature of expert systems (Luger and Stubblefield, 1998). Expert systems are also known as knowledge-based systems. However there is a method that does not have to solely rely on human expertise, which could be applied on this research project, called Case-Based Reasoning (CBR). The difference is that the knowledge is not compiled in rules, but stored as a set of structured cases (Chen and Burrell, 2001).

Another branch of artificial intelligence is called machine learning. Instead of extracting knowledge from human experts this system extract knowledge from data. This means that it is possible to discover relationships within the collected data area and we can determine the factors that influence the outcome (Negnevitsky, 2002). It is also called knowledge discovery. This technique is becoming increasingly more popular, especially in areas where there is large amount data available but the knowledge is poor (Cios *et al.* 1998).

In this paper a closer look is taken on Case-based Reasoning and Neural Networks.

2.1 Neural Network (NN)

The application of NN has grown to be a very popular method when solving different kinds of problems in various areas, like finance, engineering and medicine. The most appreciated feature is its learning and generalization ability (Cios *et al.* 1998). NN have proven to be successful in prediction, classification and clustering problems. It can address prediction problems when the output is continuous or act as a classifier when the output is binary (Negnevitsky, 2002). According to Leondes (2002) the application of NNs have great value when it is difficult or impossible to uncover relationships. The method is also helpful even when the data is noisy or incomplete.

The NN technique is inspired by the way a human brain works when solving problems and how it learns from experience. A network contains of several nodes, called neurons that are organised in one or several layers. Usually there is an input layer and an output layer, layers in between are called hidden layers. The neurons are all linked together between the layers where the connections have a weight (w) attached, which is set through a learning process of the network. They express the strength/importance, of each neuron input and have been referred to as the 'intelligence of the network' (Kauko, 2003). The structure of neurons, process a numerical signal coming from outside and sends it through the system and produces an output signal.

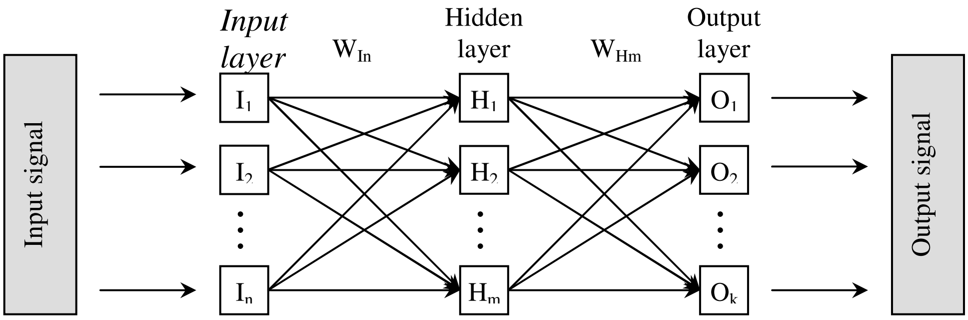


FIG. 3: A simple model of a neural network.

The process of creating a NN includes a three-step procedure. First of all it has to be decided upon what kind of architecture to be used. According to Russel (2003) there are three different classes; single-layer feedforward networks, multilayer feedforward networks, and recurrent networks. Kauko (2003) adds a fourth called competitive networks. The multilayer feedforward network, as in fig 3, is usually preferred as it can handle non-linear data. A structure with one hidden layer can handle continuous functions and discontinuous functions by two hidden layers (Negnevitsky, 2002). There are no guidelines of how many neurons and layers should be used, as it is a problem not yet well understood (Russel, 2003). If a structure has too few hidden neurons the network will overgeneralize (Arditi and Tokdemir, 1999). A general rule of thumb seems to be that the network tends to grow with the size of input parameters, which can result in several neurons in several layers (Negnevitsky, 2002). However, if the number of hidden neurons is too big there might be a problem of overfitting. The network might just memories all the training examples and thereby prevent it from generalising (Negnevitsky, 2002).

To create the memory in a NN, the system has to undergo a learning process. It is the most important feature with NN - the ability to learn and being able to improve with training and experience (Borrow, 1996). The second step in developing a NN involves the choice of an appropriate learning algorithm, which by nature can be classified as either supervised or unsupervised learning (Kauko, 2003). Among hundred different learning algorithms available, the most popular is back-propagation (Negnevitsky, 2002). It is a supervised training method of multilayer neurons (Russel, 2003). The general principle of learning is presenting an amount of cases containing input data with associated output data.

When the learning algorithm and architecture of the network is decided upon the final step takes place, which is the training process. In a back-propagation network the input pattern is propagated through the network, which has been assigned initial weights, where after an output pattern is generated. If the resulting output pattern differs from the desired output an error is calculated. This error is then sent back through the network where the weights are adjusted as the error is propagated (Negnevitsky, 2002). It is through repeated adjustment of these weights that is the core of the neural network learning process. Wieland (1987) describes this as a ‘curve-fitting’ problem. This viewpoint allows us to look on generalization as a non-linear interpolation of the input data (Russel, 2003). This process continues with a set of cases, often in several loops, until the error is sufficiently small. At this point the network is trained managing solving new cases that has not been presented to the network before. However, a network can also be over trained with too many cycles of the same cases and thereby have an over-fitting problem as described earlier. Although, this is not a big problem as there are procedures to overcome this problem (Russel, 2003). Before any system is used in a “sharp situation” it has to be validated. This is usually done by presenting an amount of new cases and then observe how the system respond.

Despite all the advantages of using NN, there are some weaknesses to be aware of. Conflicting training sets, when there are two or more identical input patterns with different outputs, prevents the system from ever learning the application (Skapura, 2002). Too many parameters can cause an over-fitting problem (Russel, 2003). Another disadvantage often mentioned, is that NN acts as a black-box. It is stated that neural networks are not transparent; a user will not be able to understand and explain the prediction result (Kim, G.-H. *et al.*, 2004), (Kauko, 2003). NN are also time consuming to develop as it is afflicted with a trial and error process (Arditi and Tokdemir, 1999), (Thomas, 2003).

2.2 Case-Based Reasoning (CBR)

This is a method that uses past experience of different cases, just like the NN. However this method does not create a memory structure, it is founded on a library of different cases. A CBR system looks for the most similar past case in a database system to match the current problem. Leondes (2002) describes this as a four-step process that starts with an input of information about the current problem, the new case. The system then tries to find a similar case within the case base, which is followed by an adaptation process of the found case/cases to fit the new case. Finally, the system provides the user a suggested solution, see figure 4. This tool is frequently used in help-desk applications when companies, in contact with users, needs to solve product problems. Also the juridical area takes interest in this system especially in situations when formulating legal reasoning through precedent (Luger and Stubblefield, 1998).

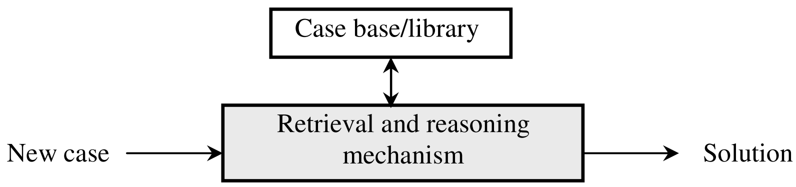


FIG. 4: The structure of a case-base reasoning system / case-based reasoning approach

A CBR study can be depicted through three major activities. The first step is related to developing the case library. The cases to be stored are experience represented by different cases with different outcomes. This first step has a major impact on how the system finally will behave. It is a structural issue of the case library, which is known as the indexing problem. According to Chua *et al.* (2001) it has two aspects; appropriate *labels* to each case and the *organizing* of the cases. The labeling involves defining features names, deciding feature value definitions and deciding feature matching approach. The organization of cases in the case library can be done in basically two ways, either trough a flat case base or a hierarchical structure. A hierarchical structure generates a more efficient retrieval process than a flat case base, that has to search through the whole case library. However a hierarchical structure demands more knowledge about the area as the cases are grouped into categories (Sankar, 2004), (Chua *et al.*, 2001). The indexing of a case library is of great importance in order to create an efficient and accurate search through the case library and retrieval of cases. This is done by assigning weights to the features, which are the parameters of each case. (Sankar, 2004), (Arditi and Tokdemir, 1999) According to Sankar (2004) and Arditi and Tokdemir (1999) the assigning of indexes is still largely a manual process and relies on human experts, however, various attempts at using automated methods (algorithms) have been proposed. By algorithms it is possible to generate weights that reflects the characteristics of the case library (Yau, 1998).

The second step involves the retrieval process. Of several retrieval methods, the most common is the k-nearest neighbor. The case retrieved, is the one that have the largest weighted sum of its features matching the current case. (Sankar, 2004) To have more than one case retrieved it is possible to use a similarity score which have a value range from 0 to 1. The similarity score describes how well it agrees with the target case, where value 1 corresponds to an exact match (Chua *et al.*, 2001). When a new case is presented to the system, the CBR system retrieves one or more stored cases similar to the new case. This according to the percentage similarity, similarity score, calculated by a user-defined similarity function (Kim, G.-H. *et al.*, 2004). A similarity score target has to be set before a search is initiated. A higher target will narrow down the number of retrieved cases.

In the final step, a case adaptation of retrieved cases has to be performed. Sankar (2004) defines the case adaptation as the process of transforming a solution retrieved into a solution appropriate for the current problem. If the system is unable to find an exact match, the tool has to make a case adaptation. This can be done either manually or by a rule-based system. The latter usually requires experts or system designers to be handled. As a result, many CBR systems work primarily as a case retrieval and proposal system and leave the adaptation process to be undertaken by the user (Chen and Burrell, 2001).

Also CBR has some weaknesses to consider. It is easy to add new cases, as they appear, however this feature can also be a disadvantage for the system. If a case is added to the system already existing in the case base with a different solution, it might cause a conflict. Maintenance of the case base must therefore be performed regularly, as it safeguards the stability and accuracy of CBR systems (Sankar, 2004). Like NN, CBR also lacks guidelines - indexing and retrieval combination has to be tested, in order to obtain best possible results (Arditi and Tokdemir, 1999).

2.3 Requirements of the AI-tool

Before choosing any AI-system to work further on, it is necessary to have a closer look on what might be required. The main goal is to translate existing experience into a useful format of information for the end-user. Focus is directed at the early stage of design where choice of systems is made. Designers usually apply their own experience to new cases (Burke and Yverås, 2004), however, their experience is not documented. A system that could predict the condition of a design, concerning moisture issues, would therefore benefit safer choices. Service life predictions described in years is not possible, due to the nature of available sources from where data can be extracted. The conditions are depicted in terms of presence and extent of rot, mould and smell.

It is of great importance that the prediction results are easy to understand and does not require further processing. Also the input procedure should require minimum of time. Several studies have proved that implementation of new tools are inhibited by steep learning curves, difficult interpreted results, complicated and the need of expert knowledge to be handled (Hien *et al.*, 2002), (Ellis and Mathews, 2001). Itard (2003) stresses the importance of winning confidence in the tool by users. If the user does not immediately understand the results of the tool, the user will be inclined to reject the software. An aim would therefore be to have a tool, which allows elaborating with pre-chosen parameters to view how it affects the outcome of chosen parameters after a number of years.

The area of service life prediction is in it self not well understood, as all the relationships between durability of materials, agents and mechanism are far from completely mapped out. It is therefore difficult to have any prior reliable knowledge in how a design will behave and in what kind of condition it will be, after a number of years. Consequently the AI-tool must have the ability to distinguish important parameters from less important parameters. Another issue is the ability to have a tool that can be developed over time, in order to add new experience. As time goes by, more cases are available and also cases further down the time line. It is important to have a range of cases that can cover the whole service life.

The system must also be reliable in delivering predictions and prove to be helpful in the early stages of design. It is an important aspect when considering the implementation process, as designers believes that they can handle moisture issues without any aid (Burke and Yverås, 2004). An AI-tool must, at least, be as good as human experts or even better in order to gain trust from the expected users of the tool.

2.4 Previous application of AI within the area of construction

Literature provides a great range of examples where AI has been applied. Behaviour of materials can be predicted through AI, such as concrete strength, cracking risk of concrete (Lai, 1997), (Dalmagioni, 2001), (Kim, J.-I. *et al.*, 2004). NN seem to be the most popular approach here, which might be due to the extensive available amount of data. The service life area has also been subject to testing AI-tools, so far only within the area of infrastructure. In this area Melhem and Cheng (2003) and Morcous (2002) managed to reach a level of 50% respectively 70% correct predictions by using CBR. NN seem to gain more interest when it comes to energy and predicting i.e. the heating demand of a building (Olofsson, 1998), (Mihalakakou *et al.*, 2002). These systems proved to perform very well with an error of 5-8%. The CBR-technique has been frequently used within the design area (Andrade *et al.*, 2003), (Yau, 1998). Usually this can be done by regular calculations and therefore it is merely a way of choosing a design more efficiently in order to avoid iterative calculations. The AI-technique is not used here to find knew knowledge. This is not the case when it comes to bidding and predicting construction costs. The area is highly complex and is usually handled by experts who learned through experience. Chua *et al.* (2001) uses CBR to facilitate the decision-making in a bidding process in order to be the winning contractor. The system provided 55% winnings. Kim, G.-H. *et al*. (2004) tested both CBR and NN in developing a construction cost estimating model, which resulted in an error rate of 4,8% and 3,0%. Another research project aimed at predicting the outcome of construction litigations (Arditi and Tokdemir, 1999) where both CBR and NN where applied on the problem. The best prediction result was obtained from the CBR system by 83%. The NN-application only delivered 67% correct predictions. In table 1, seven of the applications from above are presented, the others did not provide enough information.

Application	Method	Number of input/output	Number of cases	Correct predictions
1. Concrete strength (Lai and Serra, 1997)	NN	8 / 1	17	95%
2. Concrete strength (Kim, J.-I. *et al.*, 2004)	NN	9 / 1	24	96%
3. Litigation prediction (Arditi and Tokdemir, 1999)	NN	45 / 1	102	67%
	CBR			83%
4. Bidding (Chua *et al.*, 2001)	CBR	30 / 2	-	55%
5. Construction cost (Kim, G.-H. *et al.*, 2004)	NN	12 /1	530	97%
	CBR			95%
6. Design (Yau, 1998)	CBR	21 / 12	254	62%
7. Infrastructure deterioration (Morcous, 2002)	CBR	17 / 1	289	70%

TABLE 1: Comparison of different applications

It is difficult to draw some general conclusion out of table 1, as the chosen architecture of NN and indexing of CBR plays a major part in how well the systems performs. However, one might observe that the number inputs and number of cases has an impact on the prediction results. A larger number of inputs require a larger number of cases for both AI-systems. Application 1 and 2 seems to cope with a rather few cases and still generate good prediction results. This might depend on the fact that the cases are retrieved from laboratory measurements, which can be considered as high quality data. Both input and output data are generated in a controlled environment. Another observation is that application 5 shows that with a great number of cases the difference in prediction results, between CBR and NN, is rather small.

3 Discussion

There are a number of differences between CBR and NN-systems. The major difference is how they both operate in order to make predictions. The NN uses a self-learned memory structure whereas CBR stores cases in library. If one studies the nature of the areas of application, it is noticed that CBR is often preferred where the knowledge is well covered. In order to structure and indexing a CBR-tool, it is necessary to have some knowledge of which parameters and to what extent they will have an impact on the outcome. However, if the knowledge is fragmented, NN would be preferred, as it has the ability to extract knowledge from an amount of data. The same reasoning goes with setting the weights. When using the CBR-technique the assigning of weights are often done manually. To decide weights manually requires expert knowledge, which is not recommendable. As Leondes (2002) remarks, extracting knowledge from human experts is associated with many problems and shortcomings. One can obtain substantially different answers from different experts and even the same expert can provide different answers over a period of time. Altogether, a CBR system requires a well-known knowledge domain, whereas a NN does not require any prior knowledge of the domain. “Unlike expert systems, neural networks learn without human intervention.” Negnevitsky (2002)

One drawback of NN is the lack of transparency, as the user cannot trace the reasoning process (Chua *et al.*, 2001). The knowledge is embedded in the network and cannot be broken down in pieces, to be studied in detail. It is argued that CBR better can explain how it arrives at a particular solution by retrieving a description of a similar case, Kim, G.-H *et al.* (2004). This is however not completely true, as the user has the possibility of elaborating with input parameters and thereby study the different outcomes.

When looking from a user perspective, a CBR-system requires more prior knowledge than the NN-system to handle the prediction results. Cios *et al*. (1998) states that the application of NN is useful when the end user lack of experience. The CBR leaves the adaptation of retrieved cases to be handled by the user. In the NN-system the adaptation is considered to be done through its' generalisation ability. This feature can though be a drawback if the input pattern is out of range from the data sets the NN is trained for, which is a result of the black-box behaviour (Olofsson, 1998).

There is no given map on how a NN nor a CBR-system should be designed. Applying these methods on a problem may therefore result in several attempts before a suitable architecture or indexing is found. Kim, G.-H. *et al.* (2004) evaluated 75 different NN models to establish the best combination. Another issue is the updating of the systems. Kim, G.-H. *et al.* (2004) claims that it is easier to do on a CBR system as the NN system requires retraining and retesting. The retesting process is something that a CBR-system also should undergo, to validate the stability of the tool. Both systems must go through a maintenance procedure every time new cases are added to eliminate the risk of conflicting cases.

4 Conclusions and future work

"Forecasts of service life should be viewed as indicative and decisions should be guided, but not dictated, by the results." (ISO 15686-1, 2000) The reason for this statement is that it is impossible to calculate the service life through existing theoretical knowledge. It is still difficult to generate safe and precise predictions, due to lack of knowledge still existing within this area. It is like predicting the end of a book with missing chapters. This is the main reason why the neural network is considered to be the best choice for further work in the project. The second reason is that CBR requires more prior knowledge, than NN, of the user when it comes to evaluating the predictions.

The NN will be built in a software shell, called Neural network toolbox in MATLAB. This will be preceded by an extensive data collection, where 300 cases is needed. It is limited to documented experience, which is usually rare to be found within the construction sector. However, two different sources have been found. One of them is a governmental fund (Småhusskadenämnden) where homeowners can turn to for financial aid when their homes have been moisture damaged. The fund has an extensive archive of documented cases, where every case has been evaluated and described by damage controllers. Anticimex, which is a consulting firm within moisture damage evaluations, is the other source of experience. Both sources have documented cases covering entire Sweden.

5 References

Andrade, J.C.-R, Bento, J. and Virtuos, F. (2003). Design of Highway Bridges: Natural Place for CBR. *Journal of computing in civil engineering* 17(3).

Arditi, D. and Tokdemir, O.B. (1999). Comparison of case-based reasoning and artificial neural networks. *Journal of computing in civil engineering* 13(3): 162-169.

Borrow, H. (1996). Connectionism and neural networks. *Artificial Intelligence*: 135-155.

Burke, S. and Yveras, V. (2004). A Swedish Perspective on the Prevention of Moisture Problems During the Building's Design Phase. *Nordic Journal of Surveying and Real Estate Research* 1(1): 102-113.

Chen, D. and Burrell, P. (2001). Case-based reasoning system and artificial neural networks: a review. *Neural Computing & Applications* 10: 264-276.

Chua, D.K.H., Li, D.Z. and Chan, W.T. (2001) Case-Based Reasoning approach in bid decision making. *Journal of construction engineering and management* 127(1): 35-45

Cios, K., W. Pedrycz, and Swiniarski, R. (1998). *Data mining, Methods for Knowledge Discovery*. Boston, Kluwer Academic Publishers.

Dalmagioni, P., Pellegrini, R. and Lazzari, M. (2001). *Final report of the Task 6. Improved Production of Advanced Concrete Structure - Ipacs.* Luleå, Department of Civil & Mining Engineering.

Ellis, M. W. and Mathews, E.H. (2001). A new simplified thermal design tool for architects. *Building and Environment* 36(9): 1009-1021.

Hien, W. N., Pho, L.K. and Feriadi, H. (2002). The use of performance-based simulation tools for building design and evaluation - a Singapore prespective. *Building and Environment* 35: 709-736.

ISO15686-1 (2000). *Buildings and constructed assests - Service life planning - Part 1: General principles.* Geneva, ISO.

Itard, L. (2003). H.e.n.k., A software tool for the integrated design of buildings and installations in the early design stage. *The 8th International IBPSA Conference, Eindhoven, Netherlands.*

Kauko, T. (2003). On current neural network applications involving spatial modelling of property prices. *Journal of housing and the built environment* 18: 159-181.

Kim, G.-H., An, S.-H. and Kang, K.-I. (2004). Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Building and Environment* 39: 1235-1242.

Kim, J.-I., Kim, D.K., Feng, M.Q. and Yazdani, F. (2004). Application of neural networks for estimation of concrete strength. *Journal of materials in civil engineering* 16(3): 257-264.

Lai, S. and Serra, M. (1997). Concrete strength prediction by means of neural network. *Construction and Building Materials* 11(2).

Leondes, C. T. (2002). *Expert systems: the technology of knowledge management and decision making for the 21th cenutry.* San Diego, Academic Press.

Luger, G. F. and Stubblefield, W. A. (1998). *Artificial intelligence, Structures and strategies for complex problem solving.* Reading, USA, Addison Wesley Longman, Inc.

Melhem, H. G. and Y. Cheng (2003). Prediction of remaining service life of bridge decks using machine learning. *Journal of computing in civil engineering* 17(1): 1-9.

Mihalakakou, G., M. Santamouris and Tsangrassoulis, A. (2002). On the energy consumption in residential buildings. *Energy and buildings* 34: 727-736.

Morcous, G., Rivard, H. and Hanna, A.M. (2002). Case-based reasoning system for modeling infrastructure deterioration. *Journal of computing in civil engineering* 16(2).

Negnevitsky, M. (2002). *Artificial Intelligence. A guide to intelligent sytems*. Essex, Pearson Education Limited

Olofsson T., Andersson, S. and Östin R. (1998). A method for predicting the annual building heating demand based on limited performance data. *Energy and buildings* 28: 101-108.

Russel, S. and Norvig, P. (2003). *Artificial intelligence a modern approach*. Upper saddle river, New Yersey, Pearson Education, Inc.

Sankar K.P. and Shiu, S.C.K. (2004). *Foundations of Soft Case-Based Reasoning*. New Yersey, John Wiley & sons, Inc.

Skapura, D. M. (2002). *Building neural networks*. New York, ACM Press.

Thomas, B. *(2003).* Framtidens Intelligens. En bok om AI, återkoppling och neurala nätverk. *Lund, Studentlitteratur.*

Wieland, A. and Leighton, R. (1987). Geometric analysis of neural network capabilities. *First IEEE International Conference on Neural Networks, San Diego, CA.*

Yau, N. J. and Yang, J.-B. *(1998). Applying case-based reasoning technique to retaining wall selection.* Automation in construction *7: 271-283.*

PAPER III

PERFORMANCE PREDICTION METHOD IN THE EARLY STAGES OF DESIGN FOR OUTDOOR VENTILATED CRAWL-SPACES BASED ON ARTIFICIAL NEURAL NETWORKS

Yverås, V. (2010) Accepted for publication in the Journal of Building Physics.

PERFORMANCE PREDICTION METHOD IN THE EARLY DESIGN STAGE FOR OUTDOOR VENTILATED CRAWL-SPACES BASED ON ARTIFICIAL NEURAL NETWORKS

Veronica Yverås*,
Department of Civil and Environmental Engineering, Division of Building Technology,
Chalmers University of Technology, 412 96 Gothenburg, Sweden

ABSTRACT

The purpose of this paper is to explore the possibility of using a tool based on Artificial Intelligence (AI) and real life data. The aim is to develop and analyse *one* AI-method for *one* design part of a single-family house. Real life data from documented experiences have been used as training data to develop a neural network to predict the performance of a specified design part, in this case, the outdoor ventilated crawl-space.

The results of this study indicate that this is an approach that could usefully be developed and investigated further. The tool managed to predict smell 100%, mould 76%, and rot 92% correctly.

Keywords

Performance prediction, Artificial Neural Networks, Crawl-space

INTRODUCTION

Many design mistakes could have been prevented through an understanding of how different designs perform, by applying basic knowledge of building physics and/or by learning from experience. It is, however, difficult to *capture* and *learn* from real life data, as this would require the use of many cases involving several parameters and a large capacity to process. The retrieval of real life data regarding moisture performance requires time as moisture problems can take several years before they are revealed. Even if all of this was managed, the distribution of the knowledge remains to be solved.

This research project is focussed on developing a moisture design tool - the *Performance Indicator* (PI) tool - using a different approach than that of regular moisture design tools. It incorporates theoretical knowledge and real life data into an expert tool. The target audiences are engineers working in the early design stage. The aim is to improve the decision support in order to secure moisture safe solutions.

The objectives of this paper are twofold. The first is to develop and analyse *one* AI-method to predict the performance of *one* building element of a single-family house. The second is to evaluate the approach of using real life data.

* fax +46 31 772 1993, tel +46 31 772 5466, E-mail veronica.yveras@chalmers.se

METHOD

There are a range of AI-tools available. In a previous paper (Yverås, 2005) two main categories were investigated: Case Base Reasoning and Artificial Neural Networks, where the Artificial Neural Network (ANN) approach was found to be best suitable for this problem. By using ANN it is not necessary to know or understand the causalities between different parameters affect the final performance of the design. The method has been tested in several instances: to predict construction cost (Kim *et al,* 2004), to estimate energy performance (Issa *et al*, 1998), for indoor temperature prediction (Thomas and Soleimani-Mohseni, 2007), to predict the prevalence of building-related symptoms (Sofuoglu, 2007), and a number of other prediction challenges. A more closely related subject where ANN has been applied is service life assessment on timber as a building material based on real life data (Yatim *et al*, 2005). However, this particular study did not manage to present an ANN with good prediction ability. This failure is difficult to explain due to lack of information in that paper about, for example, the assessment of some of the chosen parameters and the ANN-design.

The outdoor ventilated crawl-space was chosen as a test case in this research project, because it is a rather usual foundation method in Sweden and also much discussed as it is recognized to be associated with moisture problems. Applying ANN requires a large number of documented cases and this is a design which is well covered in the literature and also well represented in the archives of moisture damage consultants due to its extensive history of moisture problems. Another important consideration is that the design in most cases is easy to inspect in order to determine its condition.

THE TEST DESIGN AND DATA SOURCES

There are mainly four different crawl-space designs that can be found in Sweden: plinth foundation, outdoor ventilated crawl-space, indoor ventilated crawl-space and unventilated crawl-space (Burke, 2007). To simplify, this study focusses on the outdoor ventilated crawl-space (Figure 1) in order to limit the number of parameters. It is a foundation method where the ground floor (1) is separated from the ground by an unheated space which is outdoor ventilated (3). The ground is covered by a vapour barrier to keep the ground moisture from evaporating into the crawl-space (6).

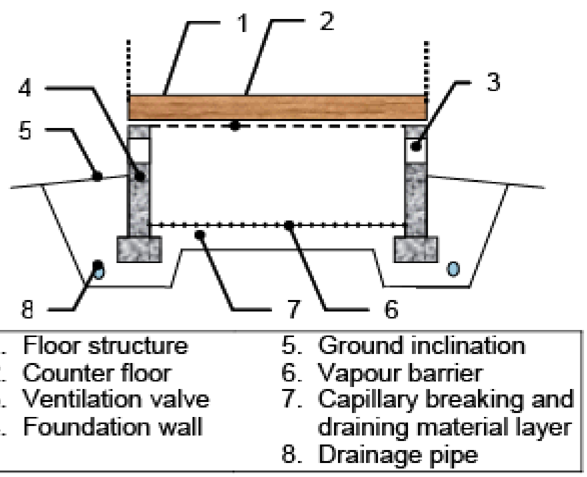


Figure 1. *An outdoor ventilated crawl-space*

The key to understanding the performance of this particular design lies in the seasonal changes in northern countries. During winter the crawl-space is cooled down and, during the summer, the surface temperature in the crawl-space is lower than the outdoor temperature due to the large heat capacity of the crawl-space. As a result the relative humidity in the crawl-space easily exceeds the risk level of relative humidity (80%) for microbiological growth. In the northern parts of Sweden the risk is higher albeit lasting for a shorter period of time than in the south of the country where the risk season is longer (Svensson, 1999).

All parameters that are likely to have an influence on the performance of the design must therefore be captured in the real life data sources. Two data sources have been used, that of the SSN (Småhusskadenämnden: National aid organisation for moisture damaged single family houses) and of Anticimex. The reason for having two data sources is because they can be used to correct each other's deficiencies. The SSN archive provides more fully documented cases, although most of them are "bad" cases where the crawl-space is in a poor condition. The good cases are those where the crawl-space is healthy and unaffected by any mould problems. It is important to find both good and bad cases, or rather crawl-spaces which show a range of conditions to minimize the risk of minority cases being disregarded in the ANN-training. In the Anticimex archive there is a larger amount of good cases to be found which is why this data is included.

NEURAL NETWORKS

The notion of AI first appeared in 1943 in the work done by Warren McCulloch and Walter Pitts, and in 1951 the first neural network computer was built by two graduate students (Russel & Norvig, 2003). Since then the area of AI and ANN has been developed and is today a popular technique with practical application in many fields: medicine, law, economics etc. The basic idea of ANN is to imitate the learning structure and process of the human brain. ANN provides the ability of learning from examples in order to make predictions without having to know the underlying relationships.

Briefly simplified, a neural network consists of a number of interconnected neurons between an input and an output layer, where weights are attached to each connection, Figure 2. Between the input and the output layer there can be one or more hidden. The process of neural network learning starts by introducing pairs of input and output data from real life data. The input data of each case is processed through the net and its hidden layers producing an output data. This is then compared with the target data (true data) of the introduced input data. After each session the error is calculated whereupon the weights in the net are adjusted. This procedure is repeated several times for the whole training data set until the error has reached an acceptable and predetermined level. By now, a memory structure that can recognise and predict cases has been accomplished. Not only that, it is also able to predict cases outside the training data due to the generalisation ability of the net. More about ANN can be found in Fausette (1994), Skapura (1995), Callan (1999), Gurney (1997) and Haykin (1999).

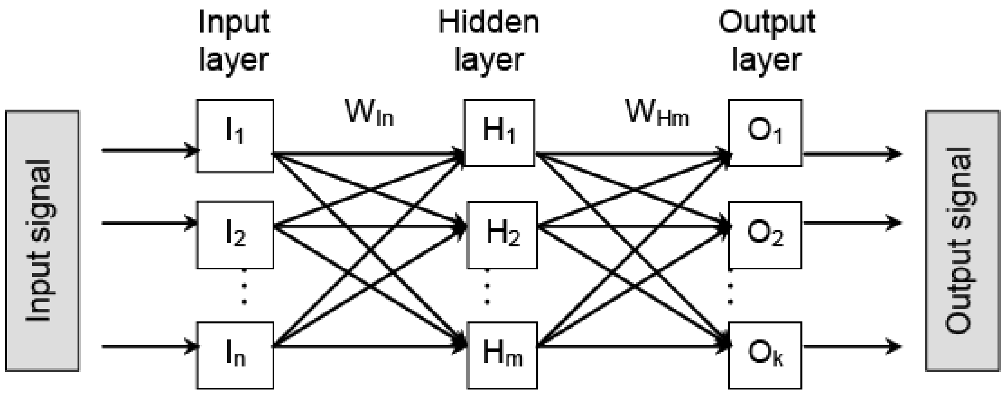


Figure 2. A Neural Network model

There are, however, some issues that need to be addressed when applying ANN. First of all, secondary data sources seldom provide complete data. If it is not possible to reject cases with missing data, a replacement strategy needs to be adopted. Furthermore, each parameter needs to be normalised, for instance in the range of 0-1, in order to prevent some parameters outperforming others in the training. Finally, it is not recommended to train the ANN too many times as this would result in over training, with the ANN adapting too much on the training data, which would have a negative effect on its generalisation ability.

APPLIED ARTIFICIAL NEURAL NETWORKS

In this research project a feed-forward neural network model is applied. The Neural Network Toolbox of Matlab 7.0 (Demuth & Beake, 2000) has been used where a back-propagation (Levenberg-Marquardt algorithm) with log-sigmoid transfer function in the nodes is applied to predict the performance of outdoor ventilated crawl-spaces. In order to prevent overtraining, the training of the net has been stopped through cross validation. As the network weights are initialised randomly the results from any two training trials will differ. Therefore, the best result out of 20 trials is presented in this paper. The prediction performance measure is calculated through the mean absolute error, *MAE*
.

$$MAE = \frac{\sum_{i=1}^{n} |actualvalue - predictedvalue|}{n} \qquad (1)$$

In order to create a memory structure of ANN to predict the performance examples are needed. The training of the ANN has been based on secondary real life data found in the two different archives described earlier: SNN and Anticimex. The retrieved training data contain variables that are believed to affect the condition and in turn the performance of the outdoor ventilated crawl-space. These variables represent the input data. The output of the training data describes the condition of the outdoor ventilated crawl-space design.

Approximately 1500 reports were searched. The data retrieval process was carried out as follows:

1. Decide on which features are representative service life parameters – both input and output
2. Create and structure a preliminary recording unit (template)

3. Preview of archives – which features can be found, in what format and to what extent
4. Evaluate the results and if necessary readjust the recording unit
5. Collect the required amount of data sets from the archives
6. Data inspection

The data retrieval comprises outdoor ventilated crawl-spaces built in Sweden and involves only single-family houses. The study is focussed on parameters that are moisture related.

In order to avoid unnecessary data noise, some cases need to be rejected from the data collection. Seasonal houses such as winter cabins or summerhouses are not included as they might have a different service life than permanently lived-in houses. Rejected are also those that have been subject to water damage such as leaking water pipes or flooding. Furthermore, if the design of the crawl-space has been altered during its life time the case should not be used as any change might have an impact of the service life.

TEST CASES

Separate ANN have been trained on each performance indicator for two reasons. The first is to minimise the network size and the second to make it possible to discover if there are any separate training difficulties for the indicators. The network architecture has been tried using both one and two hidden layers with the following number of hidden units in each layer:

- First hidden layer: 2, 5, 10, 17, 22, 45 units
- Second hidden layer: 2, 5, 10, 17, 22, 45 units

Output training data

The output data describe the condition of the outdoor ventilated crawl-space and are the performance indicators of the design. In the archives the condition is reported using the presence of microbiological smell, visible mould, and rot or rust if there is a floor structure made of concrete. From time to time microbiological activities were measured and lab results presented in the reports. This information, however, has not been used as a condition indicator, firstly because a sampling of two or three spots is unlikely to be representative, and, secondly, because such reports are rare in any case. The output data has therefore been selected and categorised according to Table 1.

As noted each output is differently classified, using 2 to 4 classes. This is an adaption to what has been found in the archives and to how the condition of the crawl-spaces has been described. The output data of smell has binary representation whereas the categories of mould and rot/rust and their internal classification have been assigned a value in the span of 0 – 1 according to Equation 2.

$$\frac{n-0.5}{m} = Y \qquad (2)$$

m: number of classes in the category
n: classification 1,2,3…m

The performance indicators are not flawless. Microbiological smell may not be present during the cold season even if there is mould present. Most microorganisms (bacteria and fungi) found in buildings can not grow in temperatures below 0°C (Flannagan *et al.*, 2001). Such cases, with heavy mould growth but no smell detected, were found during the data inspection. A majority of these cases were reported during the winter season. Additionally, these cases were very few which made them even less dependable and hence they were rejected from the training data. Furthermore, in some instances the mould can be difficult to detect due to the colour of the building material it is growing on. Black spots on a light-coloured counter floor is easy to observe by the human eye but difficult if the counter floor is black. This can result in faulty data that might obstruct an effective ANN-training.

Input training data

The input data are summarised and displayed in Table 2 showing the parameters that in the literature are believed to have an effect on the condition of the outdoor ventilated crawl-space (Nevander & Elmarsson, 1994; Kurnitski, 2000a and 2000b; Svensson 1999a, b; Yverås, 2002).

Table 1. Output representation

Output data	Definition
Y_1 Smell	0 = No smell, 1 = microbiological smell
Y_2 Ocular detected mould	0.125 = Nothing visual, 0.375 = Local spots, 0.625 = Light growth in major part of crawl-space, 0.875 = Extensive / rich growth
Y_3 Ocular detected rot	0.167 = Nothing visual, 0.5 = On surface, 0.833 = In depth

Table 2. Input data (X_1-X_{22}) list

	Parameter	Definition
A	X_1 Capillary breaking layer	1 = yes, 0 = no
	X_2 Drainage system – roof	1 = yes, 0 = no
	X_3 Drainage system – ground	1 = yes, 0 = no
	X_4 Surrounding ground inclination	1 = yes, 0 = no
B	X_5 Insulation - counter floor	1 = yes, 0 = no
	X_6 Level of insulation in floor structure	[mm]
	X_7 Insulation – foundation wall	1 = yes, 0 = no
C	X_8 Ventilation – mechanical	1 = yes, 0 = no
	X_9 Vapour barrier	1 = yes, 0 = no
D	X_{10} Load carrying structure: inorganic	1 = yes, 0 = no
	X_{11} Counter floor: inorganic	1 = yes, 0 = no
	X_{12} Foundation wall: inorganic	1 = yes, 0 = no
	X_{13} Impregnation of wood material	1 = yes, 0 = no
E	X_{14} Floor heating	1 = yes, 0 = no
	X_{15} Organic waste	1 = no, 0 = yes
F	X_{16} Relative humidity	[%]
	X_{17} Precipitation	[mm]
	X_{18} Mean annual temperature	[C°]
	X_{19} Reference wind velocity	[m/s]
	X_{20} Surrounding terrain	1 = Outside urban areas , 0 = Urban
	X_{21} Ground material*	Rock, clay = 0 / Moraine, Silt = 0.5 / Gravel, Sand = 1
G	X_{22} Age at inspection	Age at inspection – year of construction

* Describes the permeability of the ground

The parameters have been divided into groups where A, B and C represents technical solutions that are considered to meet performance requirements. The solutions of group A prevent precipitated water from reaching the crawl-space, while group B represents solutions preventing warm, humid air from condensing in the crawl-space. Group C deals with solutions that prevent ground vapour from reaching the crawl-space and evacuation of humid air in the crawl-space. Ground evaporation and initial construction damp cause an increase of vapour in the crawl-space air. The parameters in group D describe the composition of materials which is of importance for their durability. Wood, for instance, is more sensitive to moisture exposure than concrete in this context. Group E consists of parameters which were not intended to influence the performance of the crawl-space, but whose side-effect may nevertheless do so. Floor heating is such an example in that may provide heat to the crawl-space, decreasing the risk of condensation (Svensson, 1999a). However, it is uncertain to what degree a floor heating system can affect the design as it is highly dependent on the insulation degree in the floor structure. The organic waste (x_{15}) is not a design parameter but it is present in many of the retrieved cases and is very likely to influence the performance indicator of smell. It is therefore included to decrease the risk of complications in the ANN training.

The local conditions (F) arising from climate and topography is another group of parameters influencing the condition of the crawl-space. From the geographical locations given in each case, it was possible to identify the nearest weather station of SMHI (national weather data authority in Sweden) and thereby obtain data of relative humidity, precipitation, wind and temperature. The annual mean value of relative humidity is based on daily mean values calculated from observations at 00, 06, 12 and 18 hrs. The annual mean temperature can be based on hourly observations or less, depending on the type and age of the nearest weather station. The annual wind velocity mean value is calculated from 10 minute observations made every third hour. All in all this data are far from consistent as the observation frequency varies between stations and when the observations were made. Some weather stations have a long record of observations whereas others only can present data from a few years. However, as this project has a rough estimation perspective these data are regarded to have some value in indicating the climate of each case.

There is one important parameter missing which is the one describing the ventilation capacity in the crawl-space. When retrieving the training data the intention was to collect information about the area of the ventilation gaps in relation to the volume of the crawl-space. This intention had to be abandoned as this information was largely missing in the cases. However, this parameter is not the only one to influence the air exchange in the crawl-space – wind speed and local topography are others which in this case are included as input data.

Those parameters that have a yes/no-answer have binary definition. The remaining training data set is linearly scaled in order to avoid parameters with large values overriding parameters with smaller values. The linear scaling has a range of 0-1. Following parameters were scaled by Equation 3:

- Insulation with range 20-250mm
- Relative humidity: 70-87 %

- Precipitation: 284-1001 mm
- Mean annual temperature: -1.0 – 9.3 ºC
- Reference wind velocity: 1.1-8.1 m/s
- Age at inspection: 5-96 years

$$\frac{(I - I_{\min})}{(I_{\max} - I_{\min})} = X \qquad (3)$$

X is the normalised value of the input I, I_{min} and I_{max} are the minimum and maximum values of the parameter range.

Missing data

As this research project does not have access to data primarily designed for this purpose, and the use of secondary data is usually difficult because of missing data, great effort has been directed at the handling of missing data. For seven of the parameters the level of missing data varies between 21 and 63%. It is important to describe this process as it can have great influence on the results. A poor strategy to handle the missing data will be reflected in the neural network training. More on how this was handled can be found in Yverås (2008).

Results from the validation

The trained network was cross validated against 38 cases which had not been used to train the network. These cases have various designs, geographical locations and local conditions within an age span of 3-46 years. Besides the base case design described in Figure 1 there are also cases that are deviating from the basic design as follows:

- No vapour barrier
- Floor heating
- Increased/decreased insulation in the floor structure
- Insulation of counter floor
- Insulation foundation walls
- Mechanical ventilation
- Concrete structure
- Organic foundation walls
- Impregnation of wood material

Another important characteristic of the selected validation cases is that they show a validation data set with representation in all the found condition combinations of the output data, see Table 3, which also shows that the distribution of chosen validation data is approximately 10% of the total amount of retrieved data.

In all, 27 different neural network designs where tried. Table 4 shows the result of three different designs of the best performing networks for each performance indicator. The prediction performance varies with smell being easiest to predict. It should be remembered,

however, that smell has only two classes (smell/no smell) to consider whereas mould has four and rot/rust has three. It is very likely that the prediction results would improve for mould and rot/rust if the number of classes were reduced.
However, this might not be the only reason why the mould indicator does not perform as well as the others. All three indicators are based on how the condition of each crawl-space was perceived by the consultant who assessed the case. It is of course likely that the consultants have different frame of reference of how to assess the degree of mould. What is *light mould growth* for one consultant might be *rich mould growth* for another. Another issue that might make it difficult for the network to predict mould is the cases where there is mould but which has not been observed by the consultant. It can, for instance, be difficult to detect dark mould spots on a dark counter floor, which in turn can impair the neural network training.

Table 3. Output representation of training and validation data

	Smell		**Mould**				**Rot**					
Condition combination	No smell	Microbiological smell	No visible mould observed	Spots of mould growth	General mould cover	Heavy mould growth	No rot observed	Rot on surface	Rot on surface	Number cases	Train	Validation
1	●		●				●			73	67	6
2	●			●			●			34	30	4
3		●	●				●			54	51	3
4		●		●			●			32	26	6
5	●				●		●			31	28	3
6		●			●		●			58	55	3
7		●				●	●			28	27	1
8	●		●					●		2	1	1
9*		●	●					●		2	2	0
10		●		●				●		3	2	1
11		●			●			●		6	5	1
12		●				●		●		5	4	1
13		●		●					●	6	5	1
14		●			●				●	19	16	3
15		●				●			●	38	34	4

*These two cases were initially categorised into category 8 by mistake. It was discovered when the results were compiled which is why no validation has been done.

Table 4. Prediction results from the best performing network with cross validation

Performance Indicators	MAE (training)	MAE (test)	Corrrect classification [%]	ANN Design
Smell	0.0409	0.0324	100 (38/38)	17 + 17
Mould	0.1588	0.1187	76.3 (29/38)	5 + 10
Rot/Rust	0.0608	0.0839	92.1 (35/38)	2 + 17

CONCLUSION AND DISCUSSION

The performance prediction method presented in this paper indicates a potential for further development. It is an interesting approach where the results demonstrate how to capture both theoretical knowledge and real life experience into a common system. Complex and unknown causalities can be evaded with ANN and yet provide performance predictions, which is the main problem when developing traditional moisture design tools.

The method can be useful in the early stages of design when the accumulated project information still is small and the detail level is rough. Different design options can be assessed and compared. Previous experience can be captured in the structured and systematic manner displayed in Figure 3. The method also allows the knowledge to be distributed easily. Another benefit is that both the input and output data is straightforward and easy to grasp. The decision maker can easily learn the true outcome of a certain design from experience and how it evolves over time. To compare alternative designs by humidity and temperatures alone, which is offered by traditional tools, is more difficult and requires more knowledge to be handled.

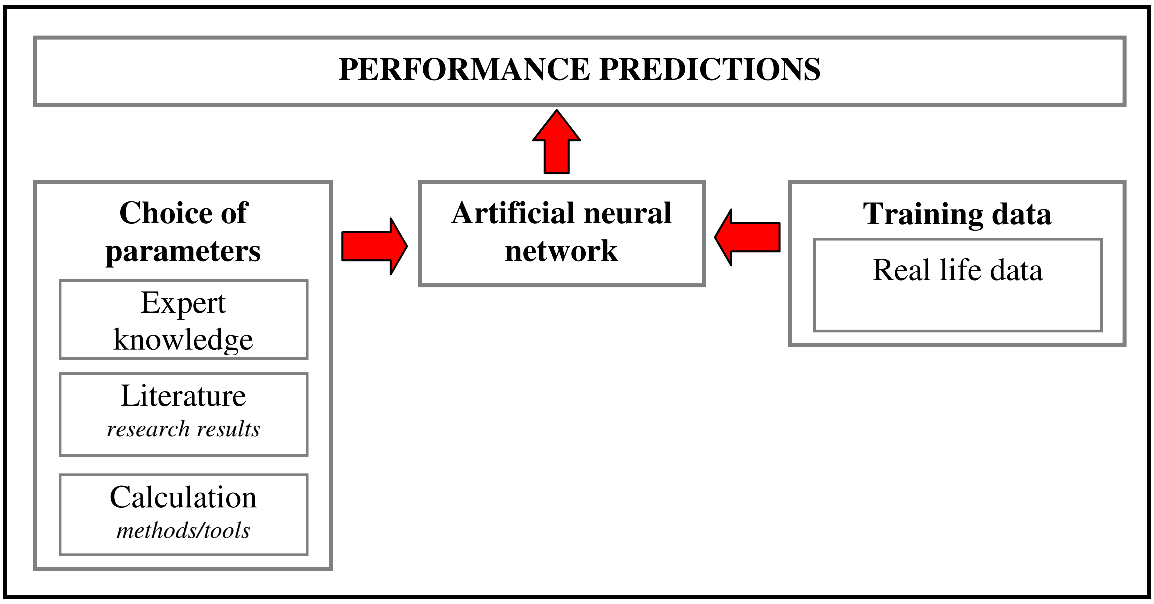


Figure 3. Performance prediction by artificial neural networks.

The positive validation results must still be carefully handled. It should be noted that the validation data set is originating from the same sources as the training data. This means that they lie within the limits of the source of which the neural network has created a memory structure. The ability to predict the condition of crawl-spaces with parameter combinations that does not exist in the training data needs to be further evaluated.

However, when a tool like this should ever be developed for practical use, it should not be based on secondary data from Anticimex and SSN due to the large amount of missing data and dearth of performance indicators. Even though good prediction results were attained, this type of training data should be avoided. If a parameter with a high level of missing values has a significantly stronger influence than other parameters on the performance, it will be impossible to obtain a good prediction level. Furthermore, the performance indicators (output data) need special attention as they are based on individual observations that might have different frames of reference. In a primary data source this can be better controlled. The real challenge is to find real life data in sufficient amount and quality with representative performance indicators. Training an ANN is the easy part.

Acquiring a large amount of primary data for neural network training in this context requires extensive resources. The more input parameters a design needs to be described, the more training data is required. The PI-tool has, however, a potential to be helpful in the decision process during the early stages of design. Another paper, not yet published, compares the prediction ability of the PI-tool with humans where the PI-tool outperforms human experts. Considering the costs due to moisture damage every year, the use of this tool should fairly quickly pay off for the building industry and society. In all, this paper demonstrates an approach worth further investigation in order to develop the reliability of this tool in itself, and also its application on other building elements (roofs and facades).

REFERENCES

Burke, S. (2007) Crawl spaces in wood framed single family dwellings in Sweden: unwanted yet popular. *Structural Survey* **25**(1): 51-60)

Callan R. (1999) *The essence of neural networks.* London: Prentice Hall Europe.

Demuth H. & Beake M. (2000) *Neural network toolbox: for use with MATLAB, user's guide.* Natick, Mass.: MathWorks

Fausett L. (1994) *Fundamentals of neural networks*. London: Prentice-Hall.

Flannagan B., Samson R. A. & Miller J. D.(2001) *Microorganisms in home and indoor work environments*. Boca Raton: CRC Press.

Gurney K. (1997) *An introduction to neural networks.* London: UCL Press.

Haykin S. (1999)*Neural Networks a comprehensive foundation.* Upper Saddle River, N.J: Prentice Hall.

Hendriks, L. and Hens, H. (2000) *Building envelopes in a holistic perspective.* IEA Annex 32, task A. Leuven, Laboratorium Bouwfysica.

Issa, R.R., Flood, I. and de Martini, A. (1998) Estimating the energy performance index of buildings. In: Flood, I and Kartam, N. (Ed.) *Artificial neural networks for* *civilengineers: Advanced features and applications*. Reston: American Society of Civil Engineers. 260-269.

Kim, G.-H., An, S.-H. and Kang, K.-I. (2004). Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Building and Environment* **39**(10): 1235-1242.

Kurnitski, J. (2000a) Moisture conditions of outdoor air-ventilated crawl spaces in apartment buildings in a cold climate. *Energy and Buildings* **33**(1):15-29

Kurnitski, J. (2000b) Crawl space air change, heat and moisture behaviour. *Energy and Buildings* **32**(1):19-39

Nevander L E, Elmarsson B. (1994) *Fukt handbok*. Stockholm, Svensk Byggtjänst.

Russel, S. and Norvig, P (2003) *Artificial Intelligence A modern approach.* Upper Sadler River, New Jersey: Pearson Education Inc.

Thomas, B. and Soleimani-Mohseni, M (2007) Artificial neural network models for indoor temperature prediction: investigations in two buildings. *Neural computing & applications.* **16**(1):81-89

Skapura D. M. (1995) *Building neural networks.* New York: ACM Press.

Sofuoglu, S.C. (2008) Application of artificial neural networks to predict prevalence of building-related symptoms in office buildings. *Building and Environment* **43**(6):1121-1126

Svensson C. *Effekter av åtgärder i uteluftsventilerade krypgrunder med fukt- och mögelskador.* Report TVBH-303;1999, Lund: Lunds Tekniska Högskola. 1999a

Svensson C. *Utvärdering av Småhusskadenämndens arkiv avseende krypgrunder*. Report TVBH 99/3035. Lund, Lunds Tekniska Högskola. 1999b

Yatim, J.M., Tapir, S.H. and Usman, F. (2005) Evaluation of building performance using artificial neural network: Study on service life planning in achieving sustainability. *Proceedings of the 8th International conference on the application of artificial intelligence to civil, structural and environmental engineering*, Rome, Italy 30 August-2 September.

Yverås, V. (2002) *Performance indicators as a decision support tool.* 2002:1 Göteborg: Chalmers University of Technology.

Yverås, V. (2008) Predicting service life of outdoor ventilated crawl spaces by neural networks. *Report No. 2008:17 Göteborg: Chalmers University*

PAPER IV

COMPETITIVENESS OF A PERFORMANCE PREDICTION METHOD BASED ON ARTIFICIAL INTELLIGENCE

Yverås, V. (2009). Submitted to *Journal of engineering, design and technology*

COMPETITIVENESS OF A PERFORMANCE PREDICTION METHOD BASED ON ARTIFICIAL INTELLIGENCE

Veronica Yverås*,
Department of Civil and Environmental Engineering, Division of Building Technology,
Chalmers University of Technology, 412 96 Gothenburg, Sweden

ABSTRACT

Purpose – Based on artificial intelligence, a performance indicator tool (the PI-tool), has been developed to predict moisture-related conditions of building elements. The objective of this paper is to test the competitiveness of the PI-tool against professionals – in other words can artificial intelligence provide better performance predictions than professionals? The paper also explores how much experience and education influenced the prediction results. Last but not least – would professionals be interest in using this tool?

Design/methodology/approach – A performance prediction comparison has been performed on line where respondents were asked to predict the condition of five different outdoor ventilated crawl spaces. The same cases were then submitted to PI-tool prediction.

Findings – The PI-tool predictions were 93% correct whereas the test persons achieved an average of 50% correct predictions. No correlation between the test persons' results and their professional or educational background were found. Finally, the PI-tool was the most requested of suggested decision support tools.

Originality/value – A tool like this can be very helpful in the early design stage in order to ensure moisture safety and thereby prevent future moisture problems. As the PI-tool is better at predicting the performance than professionals, it is worth developing this PI-tool.

INTRODUCTION

It is indisputable that it is necessary to allocate recourses in the design stage to create moisture safe solutions. If disregarded, the consequences can become expensive both from an economical and a health point of view. There are different options available to handle moisture safety during design one of which is known as *well-tried solutions* (Sandin, 1988). To ensure moisture safe solutions well-tried solutions require documented experience of real life cases. When applying a design used in a previous project the surrounding conditions must agree with the previous project. Burke and Yverås (2004) suggested that engineering consultants relied on this experience to solve moisture design problems. However, their design decisions were rarely, if ever, followed up. Generally, decisions are reused in similar projects even though their outcome is often unknown.

* fax +46 31 772 1993, tel +46 31 772 5466, E-mail veronica.yveras@chalmers.se

The performance indicator tool (PI-tool) can in this instance become helpful as it is able to learn through real life experience and thereby captures the real outcome of design decisions. Basically it is a system for experience which would allow experienced knowledge to be easily extracted to be used in the design process. In this research project, a new moisture design tool has been developed to make performance predictions regarding the condition of a building element. It is based on Artificial Neural Networks (ANN) within the area of Artificial Intelligence (AI) and thereby explores a different approach in comparison with regular moisture design tools. Real life experience of more than 350 cases of the outdoor ventilated crawl-space design has been captured in this tool. A prediction by the PI-tool describes the future condition through prevalence of microbiological smell, mould and rot/rust. The tool is aimed to be used as a decision support during the early design stage where the information level is low and rough.

Developing a tool like this further would require a great deal of effort and recourses. A justified question is therefore if prospective users such as engineering consultants, moisture experts and moisture damage consultants would take interest in such a tool.

The objectives of this paper is to compare the performance prediction ability of the tool that of professionals and to measure if experience and education have significance on their results for this particular prediction task. Furthermore, is this kind of system capturing experience desired by the professionals?

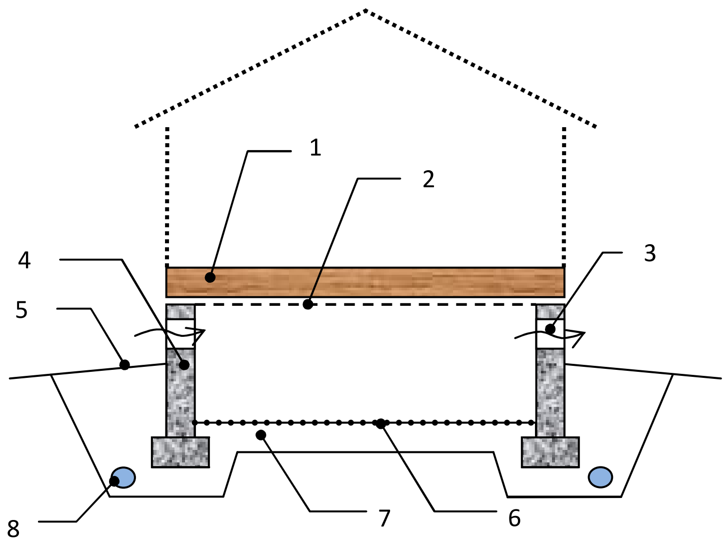
THE PERFORMANCE INDICATOR TOOL

The tool is based on Artificial Neural Network (ANN) which is a method to extract knowledge by learning from real life data. The basic idea of ANN is to imitate the learning process of the human brain. By presenting cases with known outcomes to the computer program, a memory structure is created where the internal weights in the net are adjusted during the training process. When the training is completed the memory structure has the ability to make predictions in cases not trained on.

The main benefit of applying this technique is the possibility of making predictions without having knowledge of the underlying causalities of a certain problem. Unknown causalities regarding performance predictions of building elements are not yet fully solved regarding the biological, chemical and physical relationships. When developing traditional moisture design tools, these relationships need to be known in order to describe the degradation. By the application of ANN the causality problem can be evaded. More on ANN can be found in Fausette (1994), Skapura (1995), Callan (1999), Gurney (1997) and Haykin (1999).

This performance prediction method based on ANN has in this research project only been applied on the performance prediction of outdoor ventilated crawl-spaces, Figure 1. In all, 22 parameters have been used to describe the design, which are regarded as influential on the moisture safety performance have been used to describe the design. The vapour barrier on the ground is one of the chosen parameters. If it were missing, the relative humidity, and subsequently the risk of mould growth, would increase.

The training data have been retrieved from inspection reports of outdoor ventilated crawl-spaces describing the design and current condition. The performance indicators reflect the condition through the presence of smell, mould and rot. Table 1 shows how the indicators are described at different levels. In the training of the ANN each case is presented for the ANN by the parameters describing the crawl-space case with the documented outcome in the inspection reports. The validation results of the PI-tool indicated fairly good prediction ability, as it correctly predicted smell, mould and rot by 100%, 76%, and 92% (Yverås, 2010).



1. Floor structure
2. Counter floor
3. Ventilation valve
4. Foundation wall
5. Ground surface
6. Vapour barrier
7. Capillary breaking and draining material layer
8. Drainage pipe

Figure 1. *A basic outdoor ventilated crawl-space design.*

Table 1. *Performance indicators of the outdoor ventilated crawl-space*

Performance Indicators	Level of condition
Microbiological smell	1. No smell 2. Microbiological smell
Ocular detectable mould	1. Nothing visual 2. Local spots 3. Light growth in major part of crawl-space 4. Extensive / rich growth
Ocular detectable rot	1. Nothing visual 2. On surface 3. In depth

METHOD

It was desired in the study to reach a large amount of possible respondents which is why the prediction test is set up as a web based questionnaire. The study was conducted through the questionnaire tool Dialog manager 3.0 which provides the possibility to cross-analyse the results.

To cover different types of experience the sampling frame encompasses respondents that belong to one of the following professions: *engineering consultant, moisture damage consultant and moisture experts*. These are considered to be dealing with moisture safety issues. 110 e-mails where sent out eliciting participation in this survey where additionally two reminders where sent out. The survey was limited to Sweden.

The main focus of the questionnaire is the performance prediction test where the respondents were asked to predict the performance of five real life cases of outdoor ventilated crawl-space designs. All of the cases have a different geographical location in Sweden, and their age (15-31 years old) and condition vary. The five cases are summarised in Table 2. In all, the respondents have to consider a large range of parameters, including the geographical location, in each case. The information for each case is given without dimensions except for the age. Instead each parameter is given as *present* or *absent*. The PI-tool and the respondents are given the same information

Tabel 2 *Composition of test cases*

Case:	1	2	3	4	5
Capillary breaking layer	X		X	X	
Drainage system – roof	X	X	X		X
Drainage system – ground	X	X	X	X	
Surrounding ground inclination					
Insulation - counter floor	X				
Insulation - foundation wall			X		
Mechanical ventilation		X			
Vapour barrier	X			X	X
Counter floor – inorganic	X				
Foundation wall – inorganic	X	X	X	X	
Floor heating				X	
Outside urban areas			X	X	
Ground permeability – high					X
Age [years]	31	20	17	15	18

The respondents had to make three predictions for each of the five cases, 15 predictions in all. This task requires a rather large effort of the respondents. The risk of having too many cases is that it could either result in fewer responses or a concentration drop during the survey. The respondents were not informed about the PI-tool and thereby unaware of the performance prediction challenge.

Respondents were also asked to give background information concerning current profession and years of experience as engineering consultant, moisture damage consultant or moisture expert. Furthermore, they also had to state their educational background which also includes specific moisture safety related education.

As it was not certain how the respondents would go about solving the prediction test they were requested to describe, in their own words, how they solved the prediction task.

The test does not aim to describe the knowledge level of building physics of the respondents. Most professionals would know by experience that outdoor ventilated crawl-spaces are likely to develop mould problems. Probably, most of them would strongly advise against using the design at all. Therefore the results of this paper regarding the prediction ability of the respondents (professionals) only apply to this kind of prediction problem.

Finally, the respondents were asked if they would request a system to capture experience such as the PI-tool? The question was included with several other options that are able to offer some kind of decision support during moisture safety design. In the questionnaire the respondents were asked to state if these needed to be improved/developed.

RESULTS

Background of respondents

The questionnaire was sent out to 110 people where 55 answered which resulted in a response level of 50% and with the following distribution by profession :

- Engineering consultant (A) 40% (22)
- Moisture damage consultant (B) 29% (16)
- Moister experts (C) 31% (17)

The educational background varied between the different professional groups, Table 3. The respondents predominantly had a Master of Science degree (Civil engineer). The second largest group is those with a high school engineering degree. Only 7% (4) have a research background, with either a Licentiate of engineering or a PhD. The remaining respondents are not considered to have a traditional theoretical education regarding building technology.

Table 3 *Educational background of the respondents*

Education:	A[%]	B[%]	C[%]	All[%]
High school engineer	13.5	43.7	23.6	25.4
Bachelor of science	13.5	25.0	11.8	16.4
Civil engineer	73.0	12.5	29.4	41.8
Licentiate of engineering / PhD	0	6.3	17.6	7.3
Other	0	12.5	17.6	9.1

The professional background of the respondents was more varied than expected, which of course makes it more difficult to make straightforward conclusions regarding the correlation between profession and prediction ability. For instance, almost 60% of the engineering consultants have had working experience with moisture damage investigations, Table 4.

Table 4 *Experience with moisture damage investigations*

Years of experience:	A[%]	B[%]	C[%]	All[%]
0	40.9	6.2	0	18.2
1-5	27.4	18.8	23.6	23.6
6-10	13.6	37.5	29.4	25.4
11-15	4.5	6.2	0	3.6
16-20	4.5	12.5	29.4	14.6
>20	9.1	18.8	17.6	14.6

The PI-tool challenge

Of 15 predictions the PI-tool managed to predict 14 correctly (93%). The incorrect answer concerned a rot prediction in case 2 where the PI-tool predicted "nothing visual" when the correct answer (real life outcome) should have been "on surface". The respondents did not succeed as well as the PI-tool as their average was only 7.5 correct predictions (50%). In Figure 2 the distribution of correct predictions is displayed which range from 3 to 12 correct predictions. The best three results are provided by two moisture damage consultants and one moisture expert (11 and 12 correct predictions). The lowest results of only 3 and 4 correct predictions are represented by one moisture expert (4 correct predictions) and three engineering consultants.

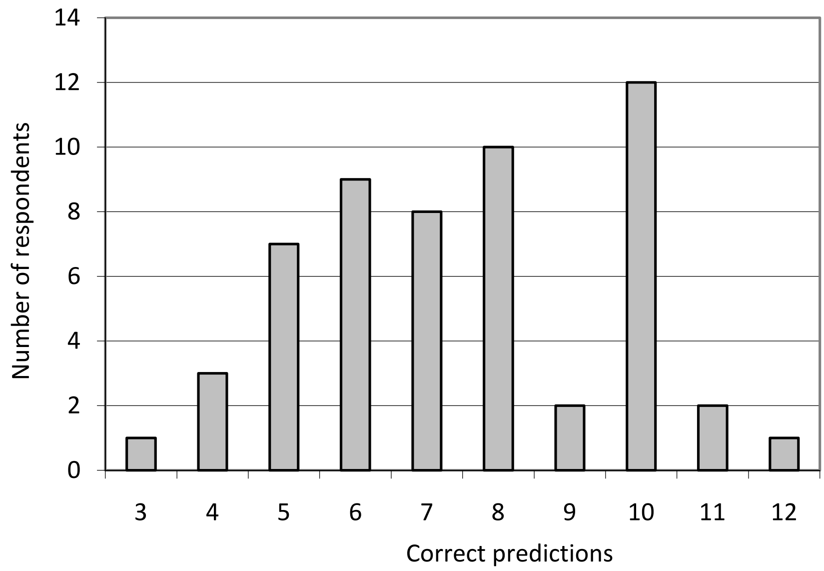


Figure 2. *Distribution of correct predictions.*

In all, the respondents tended to overestimate the degradation rate of the crawl-space cases. Of the incorrect predictions, 30% overestimated whereas 20% underestimated the degradation rate. The case 1 is healthy at age 31 years but half of the respondent group predicted it to have mould problems. In case 4, which is also a healthy case, even fewer believed the case to be free from any mould problems, Table 5.

	Correctly assessed [%]			
Case:	Smell	Mould	Rot	All
1	58	49	92	66
2	85	55	45	62
3	15	49	62	42
4	35	22	76	44
5	84	11	13	36

Tabel 5. *Respondents' prediction results for each case*

The prediction results of the questionnaire were cross analysed by profession, educational background, and years of experience regarding moisture damage inspection, see Table 6, 7, 8, and 9. It is difficult to draw distinct conclusions as neither years of experience concerning moisture damage inspections nor formal education could be related to the prediction ability of the respondents. Table 6 indicates slightly better prediction ability for those who consider themselves to be moisture experts but the difference between the professionals is not that large. The same goes for those who have additional moisture education, Table 9. In Sweden there are two well established educations which are known as: Moisture Adviser or House Doctor. The test results of those who had this additional education were compared with the results of those who did not have it, regardless of former education or experience. Having this additional education improves the prediction ability slightly.

Table 6 *Results of condition assessment by respondents categorised by profession*

	Correctly assessed [%]			
Respondents:	Smell	Mould	Rot	All
All (55)	55	37	58	50
Engineering consultants (22)	49	37	51	46
Moisture damage consultant (16)	54	35	63	51
Moisture expert (17)	65	38	62	55

Table 7 *Correlation of correct prediction and years of experience with moisture damage inspections*

Years of experience	**Smell [%]**	**Mould [%]**	**Rot [%]**	**All [%]**
0 (10)	50	36	54	47
1-10 (27)	58	41	61	53
11-20 (10)	60	36	56	51
>20 (8)	48	28	55	44

Table 8 *Correlation between correct prediction and education*

Education	Smell [%]	Mould [%]	Rot [%]	All [%]
High school engineer (14)	57	41	59	52
Bachelor of Science (9)	64	44	67	58
Master of Science (23)	50	35	54	46
Licentiate of engineering / PhD (4)	40	30	50	40
Other (5)	68	28	64	53

Table 9 *Correlation of correct prediction and additional moisture education*

Additional moisture education	Smell [%]	Mould [%]	Rot [%]	All [%]
No (26)	50	32	51	44
Yes (29)	60	41	63	55

The respondents were also requested to rate how difficult they found the prediction test to be. The majority thought it was rather difficult but 16% (9) found the prediction task easy to solve. However, they did not provide better prediction results than the average result of all respondents. Among them were two moisture experts, six moisture damage consultants, and one engineering consultant who had 10-15 years of experience of moisture damage inspections.

Means to solve the performance prediction

The respondents could use any available means to solve the prediction test e.g. moisture calculation tools, literature or colleagues. They were therefore asked to include information concerning how they solved the prediction task. Only one respondent stated that he/she had used a moisture calculation tool (Crawl 2.0) to calculate moisture and temperature levels in the crawl-space. By using this tool the respondent, a moisture expert, achieved a prediction result of 67% (10/15) correct predictions which was above average. Below are some of the respondents' comments on how they solved the prediction test:

- Experience and theoretical knowledge
- Guess
- Lack of time to solve the predictions
- Intuition – no calculation due to lack of time
- Given information incomplete
- 100 crawl space inspections
- Knowledge and experience
- Rules of thumb
- Theoretical knowledge
- 20 years of experience

Of the 55 respondents, 27 explicitly stated that they used their experience to solve the performance prediction task. Theoretical knowledge was also mentioned also in combination with experience and in only a few cases was literature referred to as a prediction base. Lack of time to solve the prediction task was mentioned by six of the respondents as an explanation as to why they did not exert themselves to perform the predictions. Some of them also thought the provided information in each case description too scant.

Is the PI-tool requested?

The question was posed in a more general manner in order to relate to other possible options of requested tools. The aim of this question is to find out if there is a need for a tool like the PI-tool without being too explicit as the respondents are not informed about the PI-tool. In the questionnaire this is named as *system to capture experience* which reflects the core of the PI-tool. As mentioned earlier the PI-tool aims to capture real life experience and transform this knowledge into accessible information during the design stage as performance predictions. According to the results in Table 10, *system to capture experience* is the most requested alternative by the respondents (93%) followed by moisture calculation programs (87%). The less popular alternative in all were the improvement/development of handbooks with 76% which still is a significant number.

Table 10 *Areas of improvement/development*

Moisture design decision support	A [%]	B [%]	C [%]	All [%]
Handbooks	74	80	75	76
System to capture experience	95	81	100	93
Product information	91	75	88	86
Moisture calculation tools	91	80	87	87
Moisture educations	91	53	97	81
Guidelines during design	91	60	81	80

DISCUSSION

What this prediction test indicates is that real life experience is difficult to handle without computational aid, even for the most experienced. In the results there were no major indications that could verify a correlation between prediction ability and education, experience or present profession. Education and experience did not according to the results give any correlation at all. When cross-analysing the results of prediction ability and experience of moisture damage inspections, it did not provide any clear indication that experience was reliable in the prediction task. Instead it might have to do more with the human capacity to handle a range of variables. According to Halford *et al.* (2005) it is impossible for the human brain to process information with more than four variables. As the number of parameters exceeds four this may partly explain the results. The background of the respondents could therefore have been ruled out as having any impact on their prediction ability. However, the study of Halford *et al.* (2005) is based on test persons learning a new problem, in contrast to the respondents in this survey who have knowledge and experience over a longer period of time. This might increase the number of variables that they are able to handle but not enough to handle the number presented in the prediction task.

The small correlation indicated regarding prediction ability and profession could have been more significant but may have been distorted by the fact that the experience base of the respondents was rather wide. For instance, 61% (14) of the consultant engineers have had experience of moisture damage consulting. It may, however, not be the sole explanation. Interestingly, a small indication on the correlation between prediction results and additional moisture education was noticed in this survey. Those with additional education performed somewhat better than those without one. However, the difference is too small to draw any conclusions on without further investigation.

One of the concerns when designing the study was the number of cases in the prediction test. Having too many cases could repel invited participants of the study as each case requires some effort. The other implication is that the respondents are not fully committed to the performance prediction task. The commitment issue was partly confirmed by the respondents when referring to lack of time. But even though there was enough time allocated for this prediction task it is also possible that the concentration level drops with the last cases. In Table 5 it can be noted that the prediction ability of the respondents was almost reduced to half between the first and the last case. When looking at the results, Figure 3, for only the first two cases – 17 of the respondents managed to predict as well as the PI-tool (100% correct predictions). The average prediction ability of all respondents is improved to 64% but still no difference regarding the correlation with education, profession or experience of moisture damage inspections. However, the first two cases differ from the others which concern their locations. Both of them are situated in large city areas in the south of Sweden where most of the respondents are active. Their experience is therefore limited to these geographical areas. If the performance of outdoor ventilated crawl-spaces is sensitive to the geographical location it can have an impact on the prediction ability of the respondents.

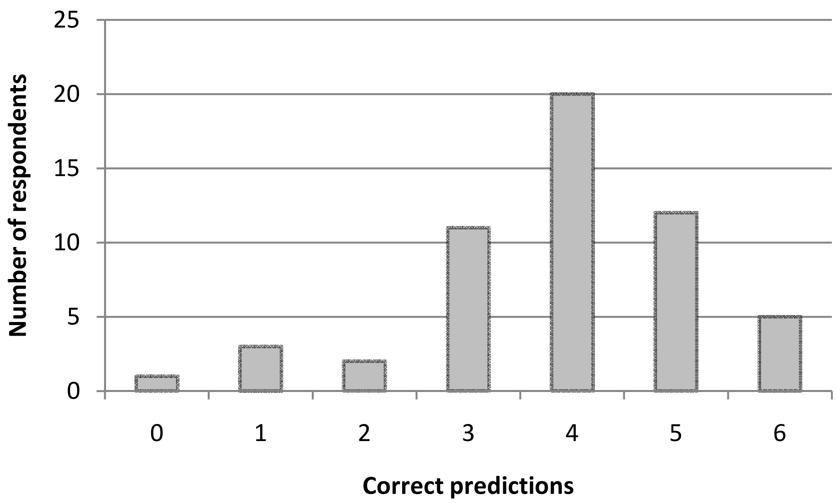


Figure 3. *Distribution of correct predictions for the first two cases.*

It should be noted that the degree of difficulty varies between cases and does not escalate after each case, in fact the last one should be the easiest to predict as it is a typical failing design according to real life experience found in inspection reports. The design is a foundation wall that partly consists of wood boards which keep the surrounding ground from caving in to the crawl-space. For the respondents with experience of moisture damage inspections it should be common knowledge that these boards rather soon evolve mould and rot degradation. This is yet another signal that the respondents did not manage to keep the same motivation or concentration to solve the prediction task to the last case.

Very few of the respondents thought the prediction task was easy. Only one respondent used a moisture calculation tool on crawl space designs for guidance to estimate the moisture condition in the cases. A moisture design calculation tool such as Crawl 2.0 can be helpful in describing the moisture and thermal condition over the year in a crawl-space. However the given information in the task is rather rough and comparable to the characteristics of the information in the early stages of design. This might impair the use of moisture design calculation tools as they usually require rather detailed information and the respondents would be forced to guess or estimate some of the input data. But even if they were able to provide moisture and temperature conditions and thereby predict if it become critical it would still be difficult to predict the rate of the degradation process.

It is not possible to draw any conclusions whether or not a tool (Crawl 2.0) like this made any difference on the prediction ability. In general, the respondents used their experience to solve the prediction task. From the information given of how they solved the predictions, the experience was based on real life experience from inspections, literature/theoretical knowledge and intuition.

The survey also aimed to find out if the professionals would take interest in the PI-tool. What can be noticed in general from the results in Table 10, is a need for improvement and/or development of different tools that are related to moisture design issues. It is indicated that systems to capture experience is acknowledged as a very important tool. Even though a *systems for experience* is on the top of the wish list it does not follow that the PI-tool would be widely applied in the design process if it was developed further and introduced to potential users. Being interested in a tool and using it in reality is two different things. External motivation, such as the new Swedish building code (Boverket, 2008), requiring improved moisture design can be an important factor to encourage the use of moisture safety design tools in general. However, a new tool means extra expense and the client of the consultants may not see the benefit of incorporating it as decision support which implies cost sensitivity in relation to perceived value for the client. Furthermore, a new tool might require new procedures in the design process which demands some level of motivation to succeed. If there already are moisture design tools used in the design process less motivation is probably required than if no such tools are involved at all.

CONCLUSION

The performance prediction challenge showed that the PI-tool was better at predicting the condition of outdoor ventilated crawl spaces than were the respondents: 97% correct predictions and 50% respectively. It is important to note that this does not measure the respondents' abilities to handle moisture issues in general. Though the PI-tool was better, it is difficult to tell how much better due to the way in which the survey was constructed. The number of cases and number of variables probably decreased the respondents' motivation or concentration and are therefore likely to have skewed the results. The study could not identify a distinct correlation between the respondents' prediction ability and education, experience of moisture damage inspections or additional moisture education. Trying to make conclusions in a design process regarding future performance, without any aid, may therefore be difficult to handle even if previous similar projects has been followed up.

In the beginning of this paper one moisture design approach was mentioned – *well tried solutions*. To adopt this approach in the design stage requires documented experience with the exact design with the exact same climate. A design in a cold and dry climate is likely to perform differently in a different surrounding. Having access to the PI-tool can therefore be of assistance in the decision process during early stages of design or when suggesting measures in a moisture damage investigation. This kind of tool to support the approach of *well-tried solutions* is not yet available but it seems there is a need for a system to capture experience (93%). The PI-tool represents this kind of system that can incorporate real life experience into the design stage. However, other influencing factors have to be investigated further in order to attain a comprehensive picture of what influences the implementation of a new tool.

REFERENCES

Boverket (2008) Regelsamling för byggande, BBR 2008. [Swedish building code] Karskrona: Boverket

Burke, S. & Yverås, V. (2004) A Swedish perspective on the prevention of moisture problems during the building's design phase. *Nordic Journal of Surveying and Real Estate Research* **1** (1), 102-113

Callan R. (1999) *The essence of neural networks.* London: Prentice Hall Europe.

Fausett L. (1994) *Fundamentals of neural networks*. London: Prentice-Hall.

Gurney K. (1997) *An introduction to neural networks.* London: UCL Press.

Haykin S. (1999)*Neural Networks a comprehensive foundation.* Upper Saddle River, N.J: Prentice Hall.

Halford G. S., Baker R., McCredden J. E. & Bain J. D. (2005) How many variales can humans process? *Psychological Science* **16**(1):70-76.

Sandin K. (1998) *Fuktdimensionering ger fuktsäkrare byggnader.* T19:1988, Byggforskningsrådet: Stockholm

Skapura D. M. (1995) *Building neural networks.* New York: ACM Press.

Yverås, V. (2010) Performance prediction method in the early design stage for outdoor ventilated crawl spaces base on artificial neural networks. *Journal of Building Physics*, Accepted for publication.

PAPER V

RELIABILITY TEST ON PERFORMANCE PREDICTION METHOD FOR OUTDOOR VENTILATED CRAWL-SPACES BASED ON ARTIFICIAL NEURAL NETWORKS

Yverås, V. (2009). Submitted to *Neural computing and applications.*

RELIABILITY TEST OF PERFORMANCE PREDICTION METHOD OF OUTDOOR VENTILATED CRAWL-SPACES BASED ON ARTIFICIAL NEURAL NETWORKS

Veronica Yverås*,
Department of Civil and Environmental Engineering, Division of Building Technology,
Chalmers University of Technology, 412 96 Gothenburg, Sweden

Abstract
When a performance prediction method for outdoor ventilated crawl-spaces was explored, the cross-validation results were rather good. These results were, however, unexpected due to the low quality of used training data (secondary). The objective of this paper was, therefore, to investigate the trained network further in order to state the level of performance reliability. This was done through a parameter study, where the results were compared with an expected outcome as the parameters of a basic outdoor ventilated crawl-space design were changed. The results showed that the deficiencies in the training data did have a large negative effect on the performance ability of the trained ANN, a negative effect that was not picked up in the cross-validation. Foremost, it was the skewed distribution of cases regarding their level of degradation (mould level), which probably caused the reversed degradation process. Another important defect was the high level of missing values for some parameters, which lead to some unreasonable results. However, there were some results from the parameter study that did give reasonable predictions. In all, the application should be investigated further, but with a data source designed for this purpose. Besides having access to complete data, it would give the possibility to improve the indicators describing the performance of the building element such as the crawl-space.

Keywords: *performance reliability, artificial neural network, outdoor ventilated crawl-space, performance prediction method*

Introduction
In the design stage of a building, it is important to create solutions that provide for a long service life without mould problems. The presence of such problems results in an unhealthy indoor environment for people who live or work in moisture damaged buildings and the economic consequences are significant. One approach to create moisture safe solutions during the design stage is to use *well-tried-solutions* (Sandin, 1998). This is a method that has been referred to be applied by engineering consultants, according to Burke & Yverås (2003). The approach requires that the performance of the previous design cases that are to be used again, are followed up and documented. However, the engineering consultants do rarely, if ever, have the time to make the follow-ups, nor do they have any system of experience to

* fax +46 31 772 1993, tel +46 31 703 5735, E-mail veronica.yveras@chalmers.se

use. When using this method, the reused and new design must exactly agree and have the same surrounding conditions that can influence the performance (climate, ground material, topography). A design is likely to behave differently in cold and warm climates, and the difference does not need to be that significant to result in a totally separate moisture behavior. Even if this is a fairly simple approach to create moisture safe solutions there are no tools available in the design process to capture and handle this kind of experience of real life design cases.

In this research project the application of an artificial neural network has been tested on this problem with the aim to create a performance prediction tool that can provide rough estimates regarding the future condition of a design in the early design stage. The study is based on secondary real life data of more than 350 outdoor ventilated crawl-space cases found in Sweden. The input data are represented by 22 parameters which are believed to have an impact on the performance. As output data three different performance indictors are used to describe the conditions of the outdoor ventilated crawl-space: smell, level of visible mould growth, and level of visible rot/rust. The prediction results of the validation turned out well, with a correct prediction level of: 100% smell, 76.3% mould and 92.1% rot/rust (Yverås, 2010). The results do not, however, stand in relation to the quality of the retrieved training data. Some parameters have low representation, and are at risk of being ignored in the ANN training. Another deficiency concerns the level of missing values that for 7 parameters is large levels (21-63%). As the training data is secondary, the reliability of the retrieved data can be questioned. Furthermore, the validation cases (38) originate from the same data sources as the training data and the cases therefore lie within the limits of the training data.

In another study not yet published, the competitiveness of the trained ANN was tested against professionals who handle moisture safety issues within their line of work. The comparison was designed as a performance prediction test of five different real life crawl space designs. There was a notable difference in the prediction results where the ANN outperformed the professionals.

The objective is, therefore, to analyse the prediction performance of the ANN further as the first validation results do not relate to the quality of the retrieved training data. This will be done trough studying the behaviour of the trained ANN when exposed to a parameter study in order to see if the prediction results are reasonable.

The performance of outdoor ventilated crawl-spaces

The outdoor ventilated crawl-space is fairly frequently used in Sweden even though it has a history of moisture problems. A study in the archives of SSN (national organization for aid to owners of single-family homes) states that this is a design that should be avoided (Björk *et al*, 2001). Still, this is widely used by several manufacturers of single-family homes as it is more cost effective to produce this type of foundation (Burke, 2007). Of the single-family homes built between 1991 and 2007, 40% rest on the crawl-space design (Anticimex, 2008). The costs to rectify mould damages in a crawl-space design are estimated to range between 15.000-250.000 SEK (Anticimex, 2008).

A basic outdoor ventilated crawl-space design is shown in Figure 1. The floor structure is separated from the ground by foundation walls. Usually, the floor structure consists of wooden beams, insulation and a counter floor holding up the insulation. Another option for the load-bearing structure is to use concrete, but this is nowadays rarely, if ever, used. It was found, when retrieving real life training data, that the cases with concrete were built between 1958 and 1986 with a concentration in the mid-seventies (Yverås, 2008). The foundation walls can also consist of concrete pillars with concrete beams in between. In these instances, the surrounding ground needs to be prevented from caving which is why boards of wood are placed between the pillars. This is of course not a recommended solution as the boards will decay fairly quickly.

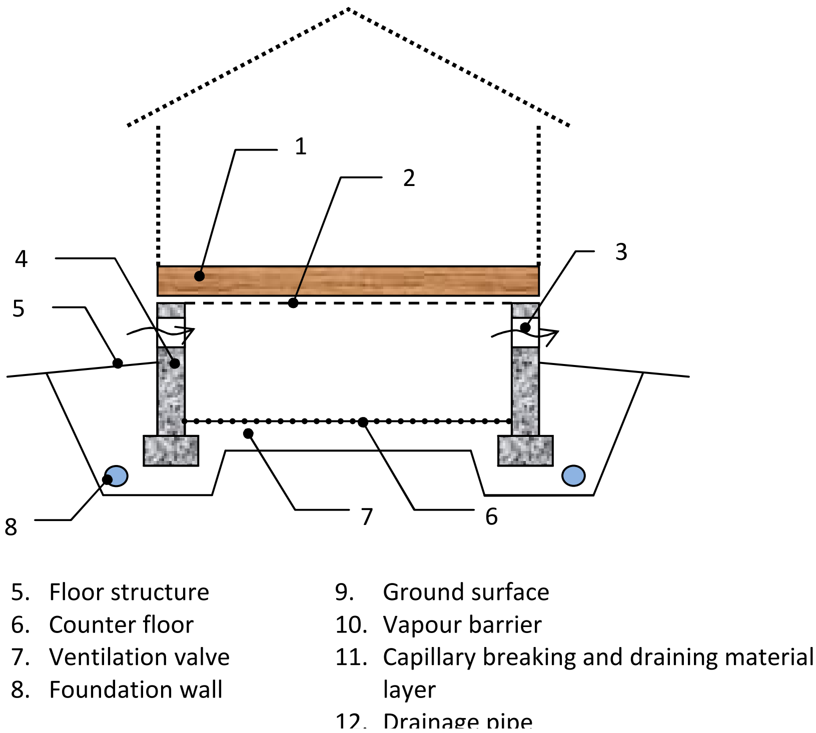


Figure 1. *Basic outdoor ventilated crawl-space design (Yverås,2010)*

The ground beneath the foundation needs a drainage system to lead away precipitated water. Having open water in the crawl-space would accelerate the degradation process due to a more humid climate. It is also recommended that the surrounding ground slopes away from the house to lead rainwater away. A capillary breaking layer prevents capillary transportation of water from the groundwater level. Ground vapour can be prevented by a vapour barrier on the ground. If excluded, the type of ground material can influence moisture conditions in the crawl-space due to evaporation. Clay is almost as bad as an open water surface, with a moisture evaporation rate 25 times higher than crushed stone (Kurnitski, 2000a). However, if the top layer of the clay is dry, it can have a moisture buffering effect, which would be positive for the climate condition in the crawl-space. Kurnitski (2000b) revealed that ground-covers with a moisture capacity like that of lightweight expanded clay aggregate decrease the relative humidity in summer in the crawl space. Perhaps it is a quality that is lost if a vapour barrier is placed on the ground surface in the crawl-space.

The ventilation of the foundation can be done by either natural or mechanical ventilation. In the moisture handbook by Nevander & Elmarsson (1994) it is stated that the required design of the ventilation gap depends on where the building is situated. A single building in an open terrain is more exposed to wind than in urban surroundings, which is why urban crawl spaces need a larger ventilation area. The material composition of the crawl space also affects the need for ventilation. A floor structure of lightweight concrete needs a larger area than wood due to construction damp.

Even if this design were perfectly constructed, this is still a risk design in respect of moisture damage. During the winter season, the crawl-space is cooled down and when the warmer season arrives, warm humid air is ventilated into the crawl-space. This air is off by the surface in the crawl-space which gives rise to an increased relative humidity, exceeding the levels for when mould starts to grow. Due to the heat capacity of the crawl-space, the cooling effect can stretch from spring to early autumn. In the south of Sweden the risk period is rather long, and in the north it is shorter but the risk instead is higher (Svensson, 2001). The solution to increase the air-exchange rate with mechanical ventilation in order to evacuate superfluous moisture is, therefore, not always a good option. According to measurements by Kurnitski (2000), a mechanically ventilated crawl-space can have a 3-5°C lower winter temperature than one with natural ventilation. Instead, a seasonally adjusted air-exchange rate, where the rate is reduced during the cold season and increased in the warm (Svensson, 2001), would be more beneficial for the crawl-space climate.

There are some solutions to improve the climate condition in the crawl-space. One is to add insulation on the ground in the crawl-space and/or to insulate the inside of the foundation walls. This solution would decrease the cooling effect causing a higher moisture condition (Matilainen and Kurnitski, 2003). Having a foundation standing on rock means a somewhat higher mould risk than if the ground material consists of sand or gravel due to differences in heat capacity. Insulation on the counter floor decreases the risk of condensation on the surface. Another option is to add heat into the ground (Matilainen *et al.,* 2003), (Svensson, 2001) during the risk season or to have a dehumidifier installed into the crawl-space. Both decrease the RH level in the crawl-space. Having floor heating in the house may, therefore, have a positive impact on crawl-space climate due to heat loss through the floor structure.

Reducing the insulation level in the floor structure is another way to improve the climate conditions in the crawl-space (Matilainen and Kurnitski, 2003).

Impregnation of the wood may increase the durability in terms of preventing rot. However, it will not stop mould growth, but instead this mycelium produces a more intensive smell than mycelia on wood without impregnation (Nevander & Elmarsson, 1994).

Due to the sensitive climate condition in the crawl-space it is important to keep it free from any organic waste. In several cases of the retrieved data, organic waste such as timber leftovers from the construction period have caused microbiological smell due to mould growth on it.

Applied ANN

A feed-forward neural network model was applied on this performance prediction problem. The Neural Network Toolbox of Matlab 7.0 (Demuth & Beake, 2000) has been used where a back-propagation (Levenberg-Marquardt) algorithm with log-sigmoid transfer function in the nodes is applied to predict the performance of outdoor ventilated crawl-spaces.

The training data is secondary data originating from two different sources, SSN and Anticimex, which consist of inspection reports. SSN was until recently the national organisation for aid to owners of private small houses. Homeowners who had encountered moisture damage could apply for financial help for remedial measures. In each application there is an inspection report describing the damaged part of the house together with its cause and a suggested solution of the problem. Anticimex is a private company and the business encompasses building inspections of different kinds. This project has taken interest in the inspection reports that are made before a house purchase is finalized.

The data retrieval comprises outdoor ventilated crawl-spaces built in Sweden and involves only single family houses. Only enclosed crawl spaces are included, which means that open plinth foundations do not take part in this research project. In order to avoid unnecessary data noise some cases was rejected from the data collection. Seasonal houses such as winter cabins or summer-houses cannot participate as they may have a different degradation process than houses with permanent living. Also excluded are those that have been subject to water damage, such as leaking water pipes or submersions due to extreme weather situations. Furthermore, if the design of the crawl space has been altered during its lifetime, the case has not been used as any change might have had an impact on the performance.

Three different networks for each performance indicator have been trained with two hidden layers where the input layer consists of 22 parameters (see Table 1). The performance indicators are a measure of the condition at a certain point of time for a certain design of an outdoor ventilated crawl-space. Here it is represented by microbiological smell, visible mould and rot in different levels as noted in Table 2. For smell, mould and rot the best performing ANN-design were 17+17, 5+10 respectively 2+17 nodes respectively in the two hidden layers, Table 3. The prediction results are classified according to the spans given in Table 4. More information can be found in Yverås (2010).

Table 1. *Input representation and replaced missing values*

Parameter	Definition	Replacement value
X_1 Capillary breaking layer	1 = yes, 0 = no	0.5
X_2 Drainage system – roof	1 = yes, 0 = no	1
X_3 Drainage system – ground	1 = yes, 0 = no	0.5
X_4 Surrounding ground inclination	1 = yes, 0 = no	0.5
X_5 Insulation - counter floor	1 = yes, 0 = no	0
X_6 Level of insulation in floor structure	[mm]	Alt 1:.Manufacturer * Alt 2: concrete structure=85mm Alt 3: else=200 mm
X_7 Insulation – foundation wall	1 = yes, 0 = no	0
X_8 Ventilation – mechanical	1 = yes, 0 = no	0
X_9 Vapour barrier	1 = yes, 0 = no	1
X_{10} Load bearing structure:inorganic	1 = yes, 0 = no	0
X_{11} Counter floor: inorganic	1 = yes, 0 = no	0
X_{12} Foundation wall: inorganic	1 = yes, 0 = no	1
X_{13} Impregnation of wood material	1 = yes, 0 = no	0.5
X_{14} Floor heating	1 = yes, 0 = no	0
X_{15} Organic waste	1 = no, 0 = yes	1
X_{16} Relative humidity	[%]	No missing value
X_{17} Precipitation	[mm]	No missing value
X_{18} Mean annual temperature	[C°]	No missing value
X_{19} Reference wind velocity	[m/s]	No missing value
X_{20} Surrounding terrain	1 = Outside urban areas 0 = Urban	No missing value
X_{21} Ground material	0 = Rock, clay 0.5 = Moraine, Silt 1 =Gravel, Sand	0.5
X_{22} Age at inspection	[years]	No missing value

* Cases with same manufacturer can be considered to have same dimensions

Table 2. *Output representation (Yverås, 2010)*

Output data	Definition
Y_1 Smell	0 = No smell, 1 = microbiological smell
Y_2 Visible mould	0.125 = Nothing visual, 0.375 = Local spots, 0.625 = Light growth in major part of crawl-space, 0.875 = Extensive / rich growth
Y_3 Visible rot	0.167 = Nothing visual, 0.5 = On surface, 0.833 = In depth

Table 3. *Prediction results from the best performing networks (Yverås 2010)*

Performance Indicators	MAE *(training)	MAE* (test)	Corrrect classification [%]	ANN Design
Smell	0.0409	0.0324	100 (38/38)	17 + 17
Mould	0.1588	0.1187	76.3 (29/38)	5 + 10
Rot/Rust	0.0608	0.0839	92.1 (35/38)	2 + 17

*Mean Absolute Error

Table 4. *Classification of prediction results*

Output data	Definition
Y_1 Smell	<0.5= No smell, ≥ 0.5 = microbiological smell
Y_2 Visible mould	< 0.25 = Nothing visual, 0.25 - <0.5 = Local spots, 0.5 - <0.75 = Light growth in major part of crawl-space, ≥ 0.75 = Extensive / rich growth
Y_3 Visible rot	< 0.33 = Nothing visual, 0.33 - <0.66 = On surface, ≥ 0.66 = In depth

Deficiencies in training data

One major drawback of using secondary real life data is that there are to some extent missing data that needs to be replaced. In this case some of the parameters could be replaced through implicit information, as assumptions can be made due to the general perception of a basic outdoor ventilated crawl-space design. The basic design represented in Figure 1 reflects the general notion of how an outdoor ventilated crawl-space should look like. So if, for instance, the vapour barrier is not mentioned in the report, it is very likely that it is present as it otherwise would have been noted as missing. The opposite goes for parameters that are not compulsory, but are believed to have a positive impact on the performance. In these cases, they are assumed not present in the design if not mentioned. Adding insulation on the counter floor or mechanical ventilation are examples of such parameters. But for 6 of the parameters (X1, X3, X4, X6, X13, X21 – Table 1) none of the above reasoning, for different reasons, could be applied and they were therefore were replaced with neutral values. According to Famili *et al.* there is an upper limit for missing values and if more than 20% of the attributed values are missing the entire record has to be eliminated (Famili et al., 1997). As shown in Table 5 several exceed this level.

Table 5. *Level of missing values in the training data*

Parameter (Input)	**Level of missing values [%]**
X_1 Capillary breaking layer	35
X_2 Drainage system – roof	21
X_3 Drainage system – ground	63
X_4 Surrounding ground inclination	25
X_6 Level of insulation in floor structure	45
X_{13} Impregnation of wood material	54
X_{21} Ground material	50

It is often stressed in the literature of ANN that the training data quality is crucial for a successful ANN training. The training data must be adequately extensive and representative as it will affect the performance of the neural network (Gardner & Dorling, 1998). The size of the training set has an influence on the generalisation capability of the network (Haykin, 1999). The higher the number of variables used, the higher the level of complexity created, which in turn requires more training data and training time (Swingler, 1996; Verleysen et al., 2003). The retrieved data sets must also be representative of the whole problem, and the training patterns should be evenly distributed within the domain (Flood & Kartam, 1994). In this case, 353 training cases were used as training data where the prediction task was represented by 22 input parameters with different levels of presence. Three of them are only represented in 5-8 cases (X8, X5 and X14). These training patterns with a low population are therefore at risk of being overshadowed by training patterns with high representation. In addition, the output representation is uneven. In Table 6 this is an issue for all three performance indicators, in particular, the rot indicator.

Table 6: *Average age and distribution of cases*

Indicator	Performance indication	Percentage of total number of cases	Average age of cases [years]
Smell	No microbiological smell	35	22.5
	Microbiological smell	65	23.0
Mould	No visible mould	34	25.4
	Local mould spots	18	24.6
	Light growth on major part	29	20.4
	Extensive / rich growth	19	21.7
Rot	No visible rot	80	22.5
	On surface	4	25.7
	In depth	16	23.7

Over time technical solutions or elements have been improved or refined in order to improve the performance. During data retrieval this was found to be the case for the vapour barrier and drainage system in ground. In older cases the vapour barriers are not age resistant, and will during the course of time degrade with a decreased vapour resistance as a result. Nowadays the ground drainage system consists of plastic pipes evacuating infiltrated water due to precipitation. Plastic pipes replaced the ceramic pipes during the seventies. Plastic drainpipes that are used nowadays are 4 times more efficient than ceramic drainpipes.

On site, the drainage system in the ground is difficult to inspect and, indeed, it is difficult to verify if there is one at all. Rather often the inspection reports had access to a technical description originating from the documents of the building permit. When the training data were retrieved this information was in some cases found to be incorrect. When a test pit was dug in order to confirm the technical description of a drainage system it was missing. The reliability of the retrieved cases can therefore be questioned. The same goes for the ground type information which is also provided by the technical description.

The input data list of parameters is, unfortunately, incomplete. Foremost, this is the case for natural ventilation and its capacity, as there were rarely any information in the reports describing the ground volume in relation to the ventilation area. In addition, it was not possible to capture if there was any vegetation adjacent to the foundation that would decrease the ventilation of the crawl-space. The absence of these parameters can therefore have impaired the ANN training. In each case it is therefore assumed that the ventilation capacity is designed according to recommendations. Svensson (2001) has found that an increased ventilation rate from low levels has a substantial positive effect on the climate conditions whereas higher ventilation rates has less effect on the mould risk compared to an increased rate.

Another issue that has not been brought up so far is the maintenance aspect of the involved elements of the outdoor ventilated crawl-space. Drainage and sometimes vapour barrier have a limited service life. There are recommendations that the drainage should be redone after 20-30 years. This has not been addressed in the input data and it can only be assumed that the drainage is redone when required which of course is not an entirely correct assumption.

The performance indicators, which are represented by the output data, and described through the presence of microbiological smell, visible mould and visible rot/rust, can also be questioned. Depending on during which season (winter/summer) the inspections are made the microbiological smell can be perceived differently. As the temperature falls during winter fungal activity slows down and finally the fungus hibernates. Most microorganisms (bacteria and fungi) found in buildings can not grow in temperatures lower than 0°C (Flannagan *et al.*, 2001); their activity is therefore lower in lower temperatures, which in turn, reduces the gases that are produced during mould growth. Detecting visible mould during inspection can also be difficult. Black spots on a light coloured counter floor is easy to observe by the human eye, but difficult if the counter floor is dark. For all these reasons the output data cannot be expected to be flawless.

In all, the prediction ability of the neural network depends very much on the data quality in terms of completeness, true values, even distribution of cases, correctly chosen parameters and sufficient amount of training sets. As noted, the retrieved data do have issues on all points.

Method

The evaluation has been based on a parameter study applied on the best performing ANN achieved in earlier work (Yverås, 2010). The study was performed in two steps were the first explored the parameters according to Table 5 and how the condition evolves over time from 0-50 years. The case base represents a basic outdoor ventilated crawl-space situated in Gothenburg, SW Sweden, in an urban area in a coastal environment. It has a wooden floor structure and conforms with the case illustrated in Figure 1. The aim is to see if the ANN can deliver reasonable results regarding the modifications in the case base for each parameter. A total amount of 18 parameters, Table 7, were predicted by the trained ANNs, one network for each one of the performance indicators; smell, mould and rot.

Table 7 also indicates the expected results of the parameter study. Removing the vapour barrier (1) will probably decrease the performance, but would there be any difference if the ground has a *low* (2) or *high permeability* (3)? The *mechanical ventilation* (4) is difficult to foresee as too much ventilation may have a damaging cooling effect on the crawl-space. Having the crawl-space *outside urban areas* (5) can improve the effect of natural ventilation. *Decreasing the insulation* (7) would have a positive impact on the ground as a result of an increased heat flow into the ground. Hence an *increased insulation* (8) level would have the opposite effect on the performance. Having *floor heating* (9) could benefit the climate in the crawl-space using the same reasoning as for decreasing insulation level. To decrease the cooling effect the *foundation walls* can be *insulated* (10) which, in turn, would improve the performance. *Insulation on the counter floor* (11) prevents damaging condensation of humid air on the counter floor. Crawl-space elements of *wood with impregnation* (13) prevent rot but not mould growth, which would instead enhance unwanted smell. Replacing the wood with a load bearing-structure of concrete (14) removes organic elements on the ground, which in turn decrease the risk of microbiological growth. However, a too humid climate can cause the reinforcement to rust. Rock and clay are examples of ground material with *low permeability* (12). They also have a larger heat capacity than other ground materials, which increases the cooling effect on the crawl-space. Leaving out the *capillary layer* (15) on the ground, *drainage system of roof* (17) and *ground* (16), and having a surrounding *negative ground inclination* (18) will increase the moisture supply in the crawl space ground.

Table 7: *Expected results of parameter study*

	Parameters	Expected results
0	Case base	0
1	Removed vapour barrier	-
2	Removed vapour barrier + Low permeability	-
3	Removed vapour barrier + High permeability	-
4	Mechanical ventilation	+/-
5	Outside urban areas	+
6	Foundation wall organic	-
7	Decreased insulation floor structure	+
8	Increased insulation floor structure	-
9	Floor heating	+
10	Insulation foundation wall	+
11	Insulation counter floor	+
12	Thermal capacity (low permeability)	-
13	Impregnation of wood material	+/-
14	Load carrying structure of concrete	+
15	No capillary breaking layer	-
16	No drainage system ground	-
17	No drainage system roof	-
18	Ground inclination to the house	-

*Parameters in the study with expected results, - impairs the performance, +improves the performance

The next part of the study looked at the impact of different climate conditions, where a case base was given 3 different locations with climate data as displayed in Table 8. The locations are spread over Sweden from south to north with inland and coastal climate. As before, the prediction is studied over time.

Table 8: *Climate data of the four different locations*

Location	RH [%]	Precipitation [mm]	T [ºC]	v [m/s]
Gothenburg (A)	81	804	7.3	4.2
Stockholm (B)	78	518	6.5	4.4
Boden (C)	77	485	1.6	2.6

Results

When studying the parameters over time it is revealed that the training data is too narrow and skewed, Figures 2,3, and 4. This is especially noticeable in the mould prediction results, Figure 3. According to the trained ANN the designs become less affected by mould when ageing which would seem to be incorrect. Even though the smell and rot indicators do not display this kind of problem, the rate of the degradation process differs in between. According to the ANN smell evolves rather quickly, and in some cases, immediately. The rot prediction in Figure 4 displays a rather flat progress for organic foundation walls.

The results make it difficult to draw any conclusions of how different design options will evolve over time. However, the relations between the curves can be studied in order to see if the results are reasonable in terms of overall improved or worsened performance for each parameter. The performance of each parameter is compared with the base case design.

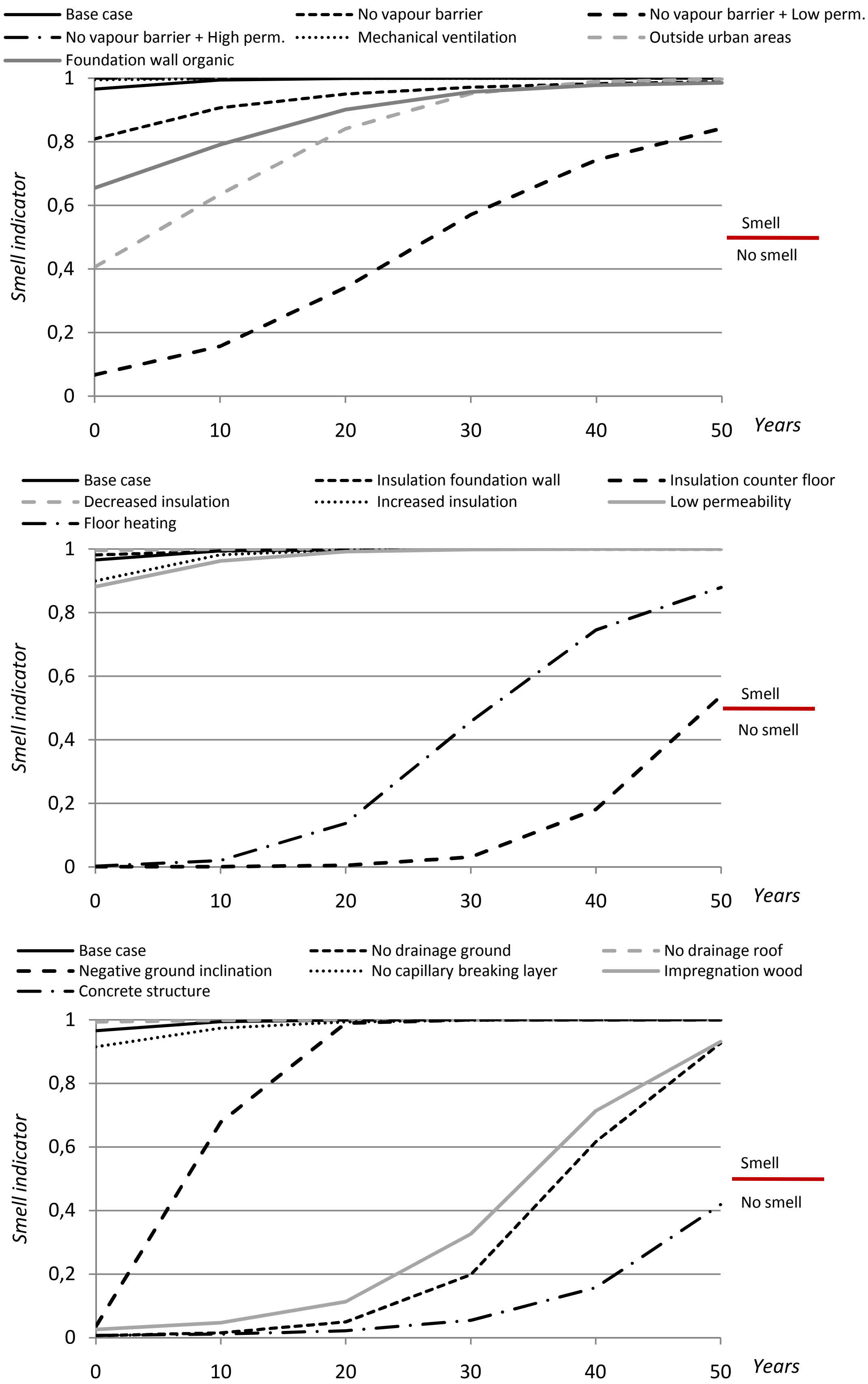


Figure 2. *Results of parameter study – Smell prediction*

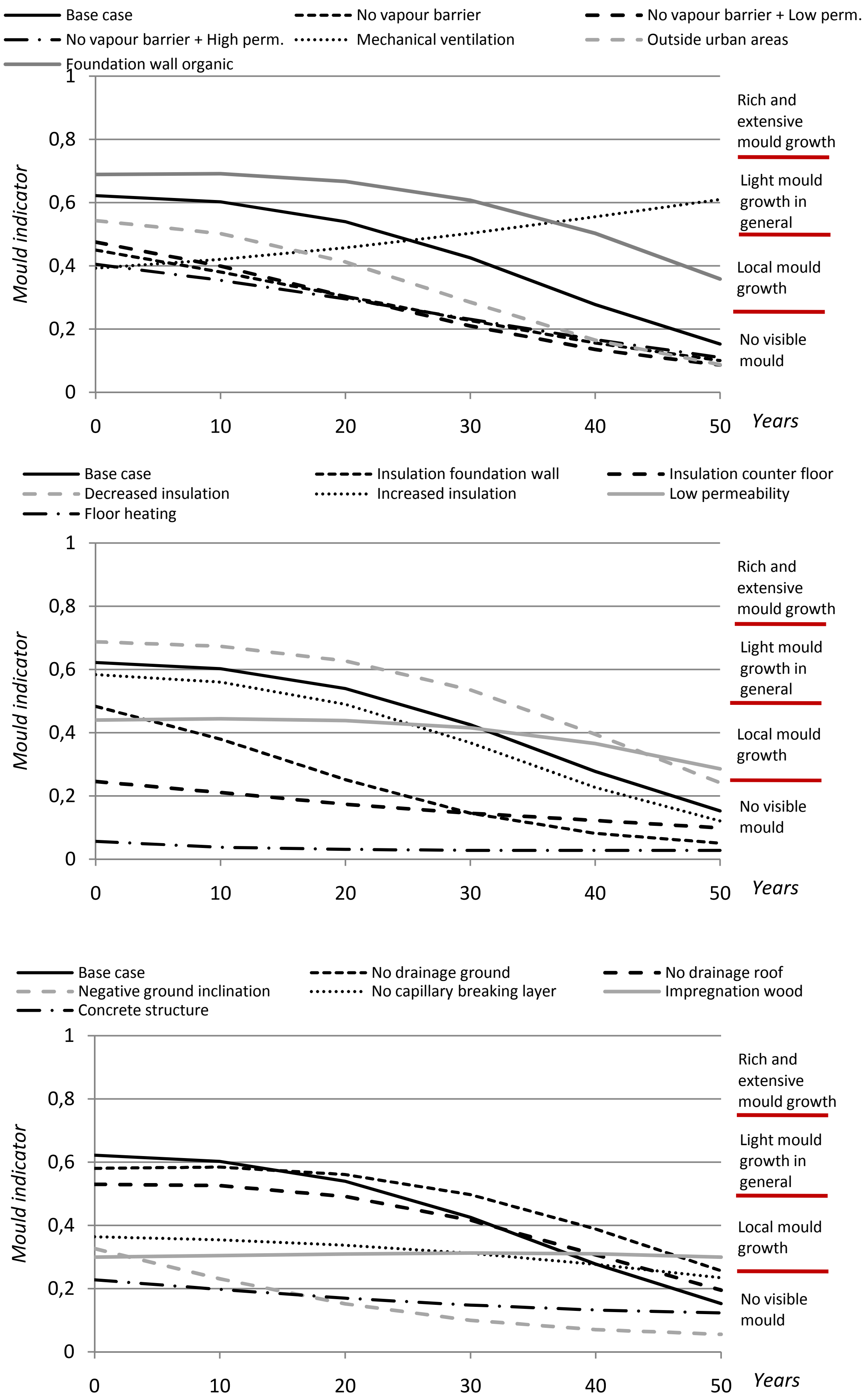


Figure 3. *Results of parameter study – Mould prediction*

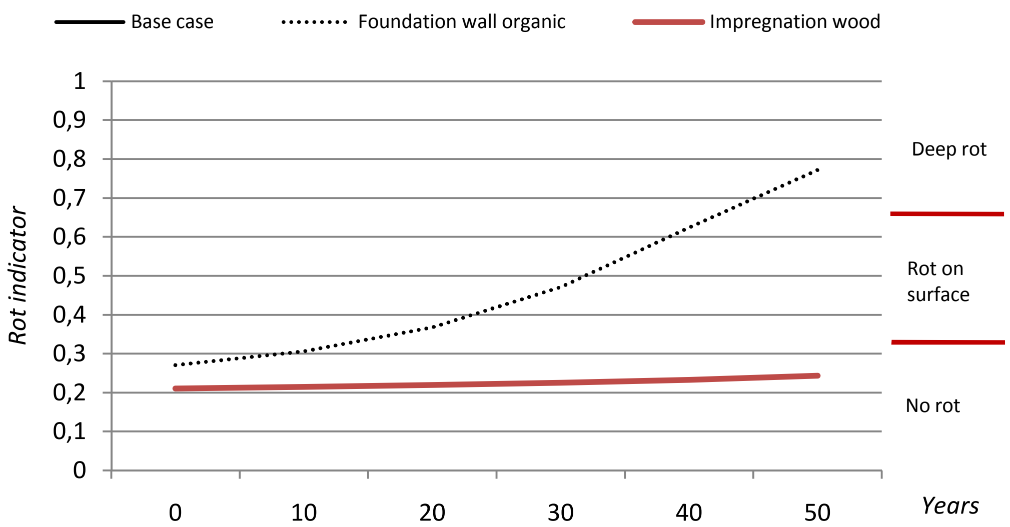


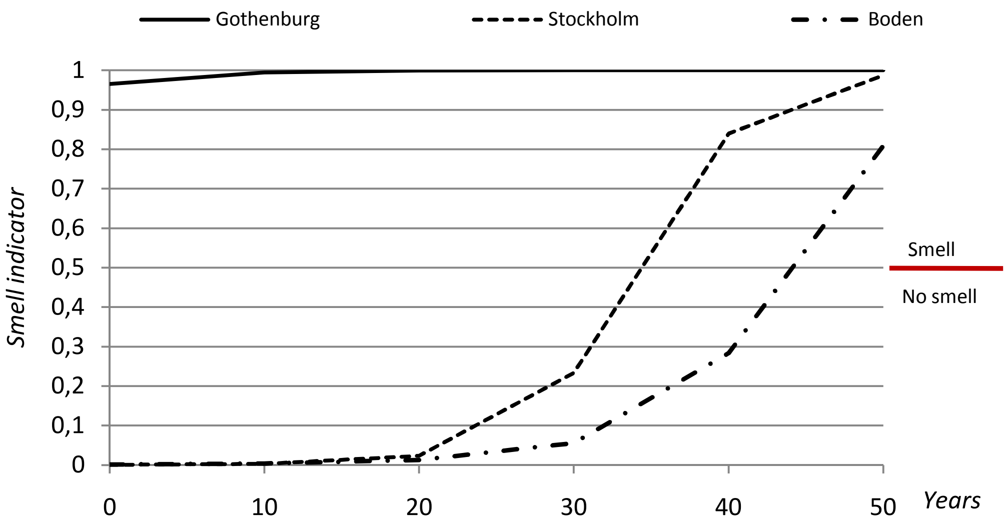
Figure 4. *Results of parameter study – Rot prediction*

Having the vapour barrier removed (1), Table 7, is according to the ANN improving the performance. Other deviating prediction behaviours are also noticed when the drainage system roof (17) and ground (16), and capillary breaking layer (15) are excluded from the design. Having a negative ground inclination should increase the moisture supply, but the ANN still predicts an improved performance. Impregnation of wood (13) should decrease the rot and give rise to smell problem. Instead the rot prediction curve coincides with the base case curve and the smell indicator is improved. The permeability of the ground (2,3) and thermal capacity (12) do not behave as predicted, and the same goes for changing the insulation level (7,8).

Adding mechanical (extract) ventilation (4) does not provide any clear prediction results by the ANN. As opposed to the other parameters regarding the mould indicator, this predicts an increased degradation when ageing. The smell indicator stays unchanged, which all together makes it difficult to say anything about the prediction behaviour of the ANN.

The rest of the tested parameters agree with expected performance. Floor heating (9), insulation foundation wall (10) and counter floor (11), and load carrying structure of concrete (14) are parameters that ANN predicts will improve the performance of the outdoor ventilated crawl-space. Having the crawl-space located outside urban areas is according to the ANN beneficial for the performance. But, due to the mould indicator behaviour with reverse performance caused by skewed training data, it is inappropriate to draw any further conclusions.

When exploring the impact of climate conditions on the performance of the outdoor ventilated crawl-space, a rather large difference can be noted in Figure 5. According to the ANN the indicator level for smell is higher in location A than in the other locations, which agrees with the fact that location A has the most humid climate of the three locations. It is difficult to draw any conclusioins from the mould indicator since the curves, as before, displays a reversed degradation process. In addition, the curves for location A and C cross each other at age 23 years. However, it can be noticed that location B has the lowest indicator level of visible mould. These results do to some extent indicate that a dryer climate, regarding humidity and precipitation, provide for a better performance. However, according to the ANN, the outdoor temperature can also be decisive for the performance of the outdoor ventilated crawl-space. Svensson (2001) states that the period for mould growth is longer in the south of Sweden in comparison with the northern parts, which instead display a higher risk. Probably, it is the cooling effect of the ground that explains this relation, that the crawl-space temperature becomes lower in a colder climate. The difference between outdoor and crawl-space temperature during summer is larger in the north than in south, which gives rise to a higher relative humidity in the crawl-space.



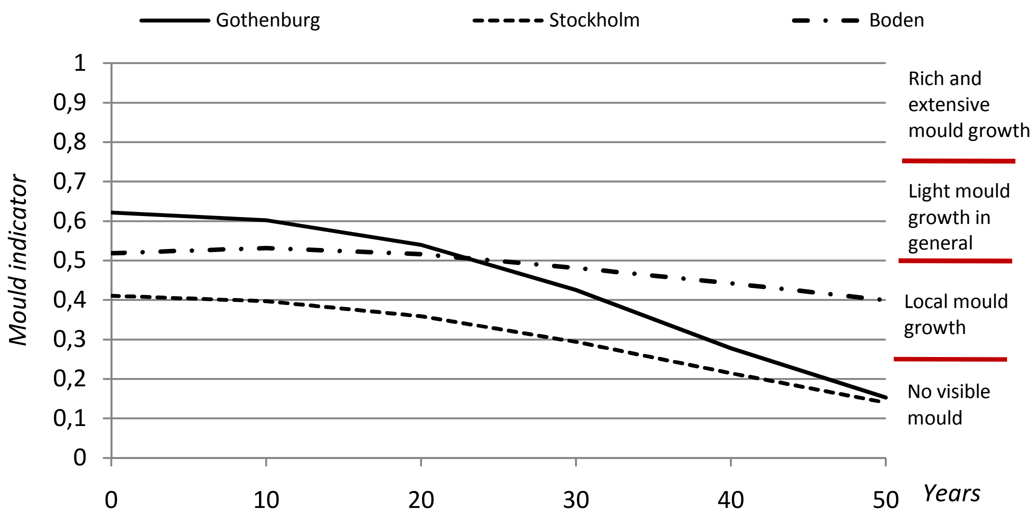


Figure 5. *Performance with different geographical locations*

Discussion

Despite the good results in the cross-validation the ANN did not perform as well in the parameter study. The largest defects concern the reversed degradation in the mould indicator. A reasonable explanation for this prediction behaviour may be found in Table 6. It displays how the average age of the training cases with respect to mould decreases with increasing mould damage condition. This can have been misleading during the ANN training. However, this might not be the sole explanation. According to Hukka and Viitanen (1999) it is possible for wood to partly recover from mould infestation during dry periods when the mould activity is decreased. It is therefore difficult to dismiss the displayed reversed degradation process as completely incorrect.

The unexpected results of 12, 13, 15, 16, and 17 in Table 7 can have been caused by the high level of missing values, Table 5. *These parameters are above the critical level of 20%* (Famili et al., 1997*)* for missing values and lie in the span of 21-63%. As a result, this might have affected the prediction when removing the vapour barrier (1) as it is dependent on the ground conditions and the solutions to evacuate precipitated water and groundwater from the crawl-space. In addition there is no parameter describing the moisture buffering effect of the ground which, as stated earlier, can be positive for the climate conditions in the crawl-space. Furthermore, the unexpectedly improved behaviour when drainage system of roof (17) and ground (16) and capillary breaking layer (15) is removed can depend on the ground characteristics. It does not necessarily have to impair the performance. Natural ground conditions can have just as good or even better drainage effect as a constructed drainage system. In highly permeable grounds it is therefore not necessary to apply a drainage system (Nevander and Elmarsson, 1994). In this instance it is difficult to tell if this is the case as 50% of the values are missing regarding ground type in the training data. In other words, the prediction results might not be entirely incorrect or solely dependent on the level of missing values.

The training process of the ANN seemed to have been sensitive to the skewed data in Table 6, which besides having an effect on mould prediction, also had an effect on the smell and rot predictions. The smell evolves quickly and in some cases immediately, which to some extent could have been prevented by having a representation of cases at age 0 which most crawl-spaces are healthy when newly constructed. However, this assumption only covers those cases that are still recognized as healthy in the inspection reports of real life cases. Wooden materials used in the crawl-space can during construction have been unprotected from precipitation which later can cause and evolve mould growth. Rot does, on the other hand, progress slowly (Fig. 4) which may be caused by the skewed data in Table 6 where the major part of the training cases is unaffected by rot. Having an uneven distribution of different conditions such as the rot performance indicator can make the ANN to choose the most common outcome of the training cases. The larger representation of healthy cases could have made the ANN to underestimate the degradation process.

Improved characteristics of a parameter can have created difficulties when training the ANN. An example of that is the insulation level in the floor structure. Over the years the insulation level in all foundations has been increased due to energy efficient measures. This is also known to have impaired the climate in the outdoor ventilated crawl-space. Older cases therefore have less insulation than more recent cases but due to the age difference the older cases have a higher degradation level. This composition of cases in the training data could have been misleading during training of the ANN. This can also be the case for the drainage system in the ground (plastic/ceramic pipes).

The results in the second parameter study regarding geographical location seem to correspond with previous research if the reversed mould degradation is disregarded. If the geographical location is important for the performance, as is indicated in the study, it may be argued that the outdoor ventilated crawl-space design needs to be adapted for each location. This is not the case today as this foundation has the same design regardless of geographical location.

The parameter study goes outside the limits of the training data which makes it somewhat unfair to apply a parameter study on to the trained ANN to assess the reliability. It has a timeline that does not correspond with the training data. However, it does provide some information on how the training data can be improved if extended in number, involved parameters and design of parameters and performance indicators. The usefulness of the trained ANN to predict the performance is therefore very restricted and not appropriate for parameter studies as it would force the ANN to extrapolation. The probable cause for the good prediction ability in the cross-validation was the same composition of the test data and the training data. In Table 8 the same similar decreasing of age with increasing mould degradation can be noted.

Table 8: *Average age of mould indicators in test cases*

Indicator	Performance indication	Average age of test cases [years]
Mould	No visible mould	26.3
	Local mould spots	23.0
	Light growth on major part	17.6
	Extensive / rich growth	20.5

Conclusions

A performance indicator (PI) tool with successful ANN training would be helpful in the early design stage. Capturing real life knowledge is valuable since the designers rarely if ever are given the opportunity to follow up of previous projects. This kind of tool would allow engineering consultants to evaluate several options without being required to have deep knowledge of building physics. Furthermore, predicting the condition is rather complex as the causalities affecting the performance such as the degradation process are not yet fully understood. ANN makes it possible to evade this which is why it makes so interesting to explore on this prediction problem.

The parameter study provided some unexpected results that are not necessarily completely incorrect. Besides skewed data and the high level of missing values, there may be parameters not captured in the training data that influence the results. Unexpected prediction results do not have to be unreasonable, it can also be new knowledge not yet recognized. But due to the data quality, it is not possible to draw any conclusions. Hence, the results did so far neither reject nor confirm the possibility of applying ANN on this prediction problem. It is important to understand that the results primarily reflect the composition of the training data rather than the ANN as the ANN itself is a product of the training data. The difficulty lies in the available data sources which in this case were inadequate due to a high level of missing data, absence of a certain parameter, and the reliability and consistency of provided information in each training case. Furthermore, relying on cross validation results alone, in this context, has in this study been showed to be insufficient.

In all, better training data is required in order further to evaluate the potential of applying ANN on performance predictions of building elements further. Even if one has access to data sources primarily designed for this purpose, it is still important to consider the choice of performance indicators and to investigate these thoroughly. Mould growth is a very complex process which needs to be handled carefully. Therefore the performance indicators have to be uncomplicated to obtain yet being representative, robust, and reliable.

Further work

But even if one has access to primary complete data and necessary parameters describing the prediction problem at hand, there are some aspects to consider when creating an ANN for predicting the condition of an outdoor-ventilated crawl space design.

The output data is another issue to discuss. In this research project, the performance representation was decided by the available information found in the data sources which in their present form are afflicted with some uncertainties. Starting with the smell indicator, the data collected from the inspection reports stems from inspections made both during the cold and the warm seasons. Furthermore, the microbiological smell is noted by a human sense and is therefore rather likely to be perceived differently. Furthermore, as the mould has a reduced activity during the cold season performance indicators should not be retrieved during that time. Even though microbiological smell is a difficult indicator to capture it can not be rejected as a performance indicator. It has a negative impact if the air from the crawl-space is allowed to leak into the house and contaminate the indoor air. It is identified as causing the sick-building-syndrome (SBS) which gives rise to different symptoms such as headache, fatigue, and irritation in throat, nose and eyes. Mould, as well, is in general not dangerous for the load-bearing structure and can be accepted as long as it does not affect the indoor air quality. The mould indicator was based on ocular findings in the crawl-space during inspections which also can be queried. First of all, mould does not have to be visible in order to develop a microbiological smell. Furthermore, it can also be difficult to detect depending on the material it grows on. As with smell, the extent of mould growth can also be perceived differently. Rot and rust do not, on the other hand, not give rise to large uncertainties. If it is limited to the surface it does not pose any threat to the load-bearing capacity. Deep rot/rust, on the other hand, can risk the element to collapse. It might be difficult to define the breaking point, though this was not an issue in the retrieved training cases in this instance.

If this was to be developed further into a performance indicator tool based on primary data, it is important to design the performance indicators to make the assessment of them secure. There should be clear and unambiguous instructions of a sampling strategy and the results should be assessed with the aim to attain consistent training data. In addition the solution must be feasible when assessing the condition of the crawl-space at the site. If measuring equipment were to be used, it must not be high precision instruments presenting decimal values. It is more important to attain indicative and robust values that are reliable enough to provide consistent data that reflects the condition of the crawl-space in terms of degradation level.

It is rather difficult to attain even distributed training data and this becomes even more complicated when age is involved. One way to evade this problem and the risk of reversed degradation process, is to include only cases of a certain age with a margin of perhaps +/-2 years. However, the chosen age needs to be considered thoroughly. If the chosen age is too young, too few unhealthy cases will be found. The effects of moisture problems can take several years to develop before they are discovered. In the case of outdoor ventilated crawl-spaces the design is very sensitive for the yearly weather changes. It may work very well for several years but one year with increased humidity can boost the degradation process, which is why it is important to have climate data covering such extreme weather years. Choosing a too old age will on the other hand exclude cases with improving design features that have recently introduced.

The input data list is rather long and can be reduced if one only accepts training data from cases with the lowest acceptable standard corresponding with the basic design of an outdoor ventilated crawl-space. The remaining parameters describe performance improving parameters (X_2-X_5), site and climate (X_{11}-X_{17}) conditions, Table 9. As noted, there are two other parameters added that are considered to improve the performance (X_4-X_5). Parameters reflecting recent and improving design will, however, limit the age span that the ANN can be trained on. The mechanical ventilation (X_7) needs some reconsideration as it can be described in capacity and seasonal adapted ventilation if present.

Table 9. *Suggested input representation for future work*

Parameter	
X_1	Level of insulation in floor structure
X_2	Insulation - counter floor
X_3	Insulation – foundation wall
X_4	Insulation cover on ground – *NEW!*
X_5	Dehumidifier – *NEW!*
X_6	Degree of natural ventilation – *NEW!*
X_7	Ventilation – mechanical
X_8	Load carrying structure: concrete/wood
X_9	Counter floor: inorganic/organic
X_{10}	Floor heating
X_{11}	Relative humidity
X_{12}	Precipitation
X_{13}	Mean annual temperature
X_{14}	Reference wind velocity
X_{15}	Surrounding terrain
X_{16}	Surrounding ground inclination
X_{17}	Ground material
X_{18}	Age at inspection

The three performance indicators were separated by having separate ANN for each indicator. They should remain separated even with a good quality of retrieved training data. The number of input data is still rather large and the relation between the indicators does not need to be dependent of each other. For instance, mould does not have to be visible by the human eye to develop microbiological smell.

Finally, the crawl-space design is available for inspections without damaging any materials or elements in order to make an assessment of the condition or to state the composition of the design. This is not always the case for other parts of the building, such as exterior walls. Applying ANN on closed elements can therefore require another approach than that for outdoor ventilated crawl-spaces.

References

Anticimex (2008) *300 000 villaägare med krypgrund riskerar att drabbas av fukt- och mögelskador.* [300,000 home owners with crawl-space foundation may be hit by moisture and mould problem] Stockholm: Anticimex

Björk, F., Mattson, B. and Jóhannesson, G. (2001) *Skador i småhus – Gamla beprövade misstag?* (In Swedish) Report 2001:6 Stockholm: Royal Institute of Technology.

Burke, S. (2007) Crawl spaces in wood framed single family dwellings in Sweden: unwanted yet popular. *Strucutral Survey* **25**(1):51-60

Burke, S. and Yverås, V. (2003) A Swedish perspective on the prevention of moisture problems during the building's design phase. *Nordic Journal of Surveying and Real Estate Research* **1**(1): 102-113

Demuth H, Beake M. Neural network toolbox: for use with MATLAB, user's guide. Natick, Mass., MathWorks; 2000

Famili A., Shen W.-M., Simoudis E. & Weber R. (1997) Data preprocessing and intelligent data analysis. *Intelligent data analysis* **7**(1-4):3-23.

Flannagan B., Samson R. A. & Miller J. D. *Microorganisms in home and indoor work environments*. Boca Raton: CRC Press.

Flood I, Kartam N. Neural networks in civilengineering. I: Principles and understanding. Journal of Computing in Civil Engineering 1994:8(2):131-148.

Gardner M. W. & Dorling S. R. (1998) Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences. *Atmospheric Environment* **32**(14):2627-2636.

Haykin S. (1999) *Neural Networks a comprehensive foundation.* New Jersey: Prentice-Hall.

Hukka, A. and Viitanen, A. (1999) A mathematical model of mould growth on wooden material. *Wood Science and Technology* **33**(6):475-485

Kurnitski, J. (2000a) Crawl space air exchange, heat and moisture behavior. *Energy and buildings* **32** (1): 19-39.

Kurnitski, J (2000b) *Humidity control in outdoor-air-ventilated crawl spaces in cold climate by means of ventilation, ground covers and dehumidification.* Report A3. Espoo: Helsinki University of Technology

Matilainen, M. and Kurnitski, J (2003) Moisture conditions in highly insulated outdoor ventilated crawl spaces in cold climates. *Energy and Buildings* **35** (): 175-187

Matilainen, M., Kurnitski, J. and Seppänen, O. (2003) Moisture conditions and energy consumption in heated crawl spaces in cold climates. *Energy and Buildings* **35**(): 203-216

Nevander L. E. & Elmarsson B. (1994) *Fukt handbok* [Handbook of moisture]. Stockholm: Svensk Byggtjänst.

Sandin K. (1998) *Fuktdimensionering ger fuktsäkrare byggnader.* [Moisture design will result more moisture proof buildings] T19:1988, Byggforskningsrådet: Stockholm

Svensson, C. (2001) *Effects of remedial measures in outdoor air ventilated crawl-spaces with mould and moisture problems*. Report TVBH-3038. Lund: Lund University, Faculty of Engineering

Swingler K.(1996) Applying Neural Networks, A Practical Guide. San Diego: Academic Press Limited.

Yverås, V. (2008) *Aspects of extracting real life data for neural network learning on service life predictions*. Report No. 2008:16 Göteborg: Chalmers University of Technology

Yverås, V. (2010) Performance prediction method in the early design stage for outdoor ventilated crawl-spaces based on artificial neural networks. Accepted for publication in *Journal of Building Physics* in 2010

Verleysen M., Francois D., Simon G. & Wertz V. (2003) On the effects of dimensionality on data analysis with neural networks. (2003) The 7th International work-conference on Artificial and Natural Neural Networks, IWANN 2003, Maó, Menorca, Spain 3-6 June.