INVESTIGATIONS ON THE PREDICTION OF CONCRETE CARBONATION DEPTH BY ARTIFICIAL NEURAL NETWORKS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES OF NEAR EAST UNIVERSITY

By

IKENNA DESMOND UWANUAKWA

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Civil Engineering

NICOSIA, 2016

Ikenna Desmond UWANUAKWA: INVESTIGATIONS ON THE PREDICTION OF CONCRETE CARBONATION DEPTH BY ARTIFICIAL NEURAL NETWORKS

Approval of Director of Graduate School of Applied Sciences

Assoc. Prof. Dr. Nadire ÇAVUŞ

We certify this thesis is satisfactory for the award of the degree of Masters of Science in Civil Engineering

Examining Committee in Charge:

Prof Dr Adnan Khashman	Committee Chairman, Electrical and Electronic Engineering, FIU.
Asst Prof Dr Ertuğ Aydın	Committee Member, Civil Engineering Department, EUL.
Asst Prof Dr Pınar Akpınar	Supervisor, Civil Engineering Department, NEU.

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name:

Signature:

Date

ACKNOWLEDGEMENTS

Without the help and guidance of some individuals, I may have failed to present this work. My supervisor Asst. Prof. Dr. Pınar Akpınar, brought out in me the confidence to explore more options; *Hocam, çok teşekkür ederim.*

A fully activated network of family members kept me going with their love and prayers among whom are; Nwakudu A. C. (Dad), Uwanuakwa A. C. (Mum), Enyinnaya, Ij, Chidinma, Azo, Edo, Kelechi and not the least Blessing. I have loved and missed you all.

Special thanks to Professor Dr. Adnan Khashman for the insights he made us gain through his valuable comments. Also, many thanks to Asst. Prof. Dr. Boran Şekeroğlu, Mr. Cemal Kavalcıoğlu and Mr Olaniyi Ebenezer for their valuable supports and for the technical information that they have provided.

Assoc. Prof. Dr. Nadire Çavuş, I thank you for valuable input.

Okey U. O. (CEng), Prof. Ukachukwu S.N., Williams E. N. your supports have been not forgotten. My layers of friends, *gracias.*

To Okey Obioma U…….

ABSTRACT

Carbonation problem in concrete occurs as a result of the chemical reaction between the products of cement hydration and CO_2 penetrating into the concrete porosity. This chemical reaction do not only alter the concrete microstructure, but it is also known to cause initiation of reinforcement corrosion. Hence, the service life of the structure is affected. An adequate model capable of considering the effects of influencing factors for the prediction of the progress of carbonation process in concrete would provide benefits in maintaining the designed service life of structures.

This study aims to investigate the feasibility of Artificial Neural Networks (ANN) for the prediction of carbonation depths progressing in concrete as a non-destructive method. A supervised neural network models based on the Feed-Forward backpropagation learning algorithm was used. 18 input parameters including binders' composition, mix design parameters, curing properties and environmental factors, that are known to influence carbonation process, were employed in the model. 225 experimental cases obtained from the related literature were used to train and test the proposed ANN model and carbonation depth was predicted as the output. A combination of 14 different optimization functions with three training/testing ratios and five different numbers of hidden neurons was studied.

The results obtained indicates the feasibility of ANN use for carbonation depth predictions; correlation coefficient (R) values that were greater than 0.9 in all cases, together with the network training mean square error (MSE) converging to a threshold of 0.001 were obtained. The results shows that optimized combination of training/testing ratios, number hidden neurons and the optimization function yielding best performance was found to be Scaled Conjugate Gradient (SCG) under 60:40 traning:testing distribution with 10 hidden neurons.

***Keywords*:** Concrete durability; carbonation problem; factors affecting carbonation depth in concrete; artificial neural networks; feed-forward backpropagation algorithm

ÖZET

Betondaki karbonatlaşma problemi, atmosferdeki CO_2 gazının beton mikrostrüktürüne girerek çimento hidratasyon ürünleri ile reaksiyona girmesiyle oluşur. Bu reaksiyon, hem beton mikrostrüktüründe değişimlerin meydana getirmekte, hem de donatı korozyonunun başlamasına neden olabilmektedir. Donatı korozyonu ile birlikte betondaki karbonatlaşma problemi, betonarme binalarda ciddi hasarlara neden olabilmekte ve böylelikle binaların servis ömrünü etkileyebilmektedir. İlgili tüm faktörleri göz önünde bulundurarak karbonatlaşma probleminin beton içerisinde ilerlemesini öngörebilecek bir modelin oluşturulması ile binaların tasarlanan servis ömürlerini sürdürebilmeleri konusunda yarar sağlaması beklenmektedir

Bu tez çalışması, betondaki karbonatlaşma derinliğinin tahribatsız bir yöntemle belirlenebilmesi ve karbonatlama probleminin ilerlemesinin tahminin yapılabilmesinde "Yapay Sinir Ağları"nın uygulanabilirliğini araştırmaktadır. Literatür taramasından elde edilen 225 deneysel numune bilgileri ileri beslemeli geriye yayınım yapay sinir ağları ile üç katmanlı bir modelde 14 algoritma ile üç değişik eğitim/test dağılımı ve beş farklı gizli noron sayısı kullanılırak çalışıldı.

Bu çalışmada elde edilen sonuçlar yapay sinir ağları ile karbonatlaşma derinliği tahmini çalışmalarının başarılı şekilde yapılabileceğini göstermektedir; çalışmada kullanılan tüm değerler ile bulunan korelasyon katsayıları (R) 0.9'dan daha yüksek, ve karesel ortalama hata (MSE) değerleri 0.001'e yaklaşmış olarak elde edilmiştir. Tüm elde edilen sonuçlara bakıldığında, "Scaled Conjugate Gradient (SCG)" fonksiyonunun 60:40 oranındaki eğitim/test veri dağılımı ve 10 gizli nöron ile kullanılmasıyla en başarılı karbonatlaşma tahmininin elde edildiği gözlemlenmiştir.

***Anahtar Kelimeler*:** Beton dürabilitesi; karbonatlaşma problem; karbonatlaşma derinliğini etkileyen faktörler; yapay sinir ağları; ileri beslemeli geriye yayınım yöntemi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1. Carbonation Problem in Concrete

Concrete is the most widely used construction material in the world. It can be moulded into any desired form and can be used for both offshore and onshore structures. However, some of its advantages may cease in time as certain concrete durability problems may occur throughout the lifetime of a reinforced concrete structure. Carbonation problem in concrete has been identified as one of the potentially severe durability problems since it can lead to the initiation of corrosion in reinforcing steel bars along with the alterations it causes in hydrated cement paste microstructure.

Carbonation in concrete occurs when carbon dioxide from the air penetrates into concrete and reacts with cement hydration compounds, such as calcium hydroxide, to form calcium carbonates. The process is continuous if sufficient CO_2, adequate moisture and favourable temperature are maintained from the external environment. As the carbonation process progresses inwards within the concrete, pH level is reduced as a result of newly formed compound with different alkalinity. The decrease in pH of the concrete leads to destruction of protective layer over the steel bar. Continuous supply of oxygen and moisture that can be easily available from the atmosphere triggers corrosion of steel bars in concrete, and therefore deteriorations on the structure is inevitable.

1.2. Definition of the Problem

Majority of the conventional experimental methods (see section 2.2.3) used to predict carbonation depth in concrete are mainly destructive and they are capable of providing only approximate results.

Other non-destructive methods such as Infrared spectroscopy and x-ray diffraction methods are not cost effective. Moreover, these methods cannot provide information the individual effects of each influencing parameter on the extent of the progress of carbonation.

1.3. The Objectives, Scope and the Significance of the Study

This study aims to carry out preliminary studies for determining the feasibility of Artificial Neural Network as a non-destructive method for the prediction of carbonation depth in concrete. The choice of using Artificial Neural Networks (ANN) was mainly based on the well-known capability of ANN to predict non-linear data by developing experience from the previous examples introduced to the model.

The previous work in the literature (see section 3.10) on the application of ANN application on carbonation depth prediction is found out to be very limited and a need for improved ANN models capable of covering more aspects of carbonation problem is detected.

In this study, the level of accuracy of the ANN model for prediction of carbonation depth was studied with the special focus on the effect of varying training:testing distribution and number of hidden neurons. The study was carried out with an extensive set of experimental data selected from the related literature.

The use of ANN for the prediction of carbonation depth in an efficient way has the potential to provide a reliable and non-destructive alternative to costly and laborious experimental test methods. The application of ANN also has the potential to provide insight on the individual effects of each parameter influencing the progress of carbonation in concrete. Hence, this study may also serve as a basis for a future application of ANN in the mix design stage for designing a concrete with a desired carbonation performance in a defined lifetime.

1.4. The Structure of the Thesis

The problem that was addressed in this thesis, as well as the objectives, scope and the significance of the study are introduced in chapter one.

Chapter two and three are dedicated to review of literature on concrete carbonation and artificial neural networks respectively Chapter four deals with methodology of the study, with details of ANN model used and mode adopted for selection of network parameters.

Results are presented and extensively discussed in chapter five. Finally, conclusions that are drawn from the results, and the recommendations for future studies are presented in chapter six.

The detailed results showing model training performance graph and correlation between measure and predicted normalized carbonation depth are provided at the end of the thesis in the appendix, as an electronic copy.

CHAPTER 2
LITERATURE REVIEW ON CONCRETE CARBONATION

2.1. Concrete Durability in General

The advent of today cement (Portland) in 1824 as patent by Leeds Builder Joseph Aspdin (Neville and Brooks, 2010) introduced the word into the production of reinforced concrete as seen in the works of Joseph Monier in 1867 and many more (Wang and Salmon 1985). Though it proffers solution to the much needed long spanning encountered by classical builders who used pointed Arches (Gothic architecture) or rows of Arches (Aqueduct of C. Sextilius Pollio, Ephesus) to overcome a given span. The durability of reinforced concrete structure has become a major challenge to the scientific community. It is expected of every concrete structure within their service life to be structurally stable. The durability of concrete structures is a function of the amount of free water within concrete pores (Auroy et al., 2015) and therefore require that every concrete structure be manufactured to reduce pore spaces in concrete. Studies on concrete durability have shown that carbonation attack is the principal culprit of chemically induced deterioration in cementitious materials as it open the gate for other attacks such as sulphate, alkali-aggregate reaction and chlorides attacks. Carbonation has been defined by various authors, Houst, (1996) defined carbonation of cement as neutralisation reaction of bases by an acid formed by carbon dioxide in the air. Castellote*et. al.*, (2008) defined carbonation as a slow and complex physicochemical process involving the interaction of atmospheric CO_2 with cementitious materials in presence of water which modifies the structure of the concrete. In summary carbonation is hereby defined as a continuous physicochemical neutralisation reaction of hydrated $Ca(OH)_2$ in the presence of water resulting to precipitate of $CaCO_3$ and modification of micro structure and properties of hardened cement materials.

2.2. Concrete Carbonation Mechanism

Carbonation is a continuous (Castellote et al., 2008), and gradual attack in cement paste and concrete. The process involves physical and chemical processes of diffusion, permeability and absorption of CO_2 and H_2O into cement matrix, which dissolves in the pore solution to form HCO_3^- and CO_3^{2-} ions, and reacts with Ca^{2+} from portlandite ($Ca(OH)_2$), calcium

silicate hydrate (C-S-H) and the hydrated calcium aluminate and ferroaluminate to give a precipitate of calcium carbonate ($CaCO_3$), silica gel and hydrated aluminium and iron oxide. (Parrott, 1987; Borges et al., 2010)

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \quad (2.1)$$

$$H_2CO_3 + Ca(OH)_2 \rightarrow CaCO_3 \quad (2.2)$$

$$H_2CO_3 + CaCO_3 \rightarrow Ca(HCO_3)_2 \quad (2.3)$$

$$Ca(HCO_3) + Ca(OH)_2 \rightarrow 2CaCO_3 + 2H_2O \quad (2.4)$$

Atmospheric air supplies the needed CO_2 that penetrates the pore spaces and react with moisture to give aqueous carbonic acid (H_2CO_3). The chemical reaction creates a modification between the solution and the hydrates with a precipitate of ($CaCO_3$). Calcium carbonate formed fill and increase the densification of the microstructure with a decrease of the hydrated cement pH from 13.5 to 9.5 (Berkely and Pathmanaban, 1990; Ahmad, 2003), resulting to de-passivation hence deterioration of the "*cementitious*" structure set in (Villain et al., 2007).

Physically the process is governed by the absorption and penetration rate of CO_2 into the pore space of the matrix and the amount of exposure, percentage of CO_2 content in air, internal and external relative humidity of the concrete and the influence of temperature (Saetta et al., 1993; Salvodi el al., 2015). Salvodi el

al., (2015) further reviewed that ambient humidity can substitute the internal humidity of the concrete on the assumption that the external relative humidity will reach a steady rate with the internal relative humidity. For measurement of rate of carbonation, it is important to consider these prevailing factors which influences carbonation. CO_2 as an inert gas does not readily react with other compounds except certain conditions are met. CO_2 is a stable gas and goes into reaction in an aqueous solution to form carbonic acid. The reaction of CO_2 with hydrated cement can only take place in solution hence the rate of carbonation is dependent on relative humidity of the atmosphere (Parrott, 1987; Sevelsted and Skibsted 2015).

According to Houst (1996) concluded; the water held in pores of hydrated cement paste (hcp), are in form of absorbed water, condensed capillary water and free water found in large

capillaries pores resulting from a decrease in vapour pressure above a concave liquid meniscus over the pressure above a plane liquid surface. The pressure decrease gives rise to capillary condensation which may occur at a relative humidity lower than 100%. Kelvin's equation estimates the maximum radius of pores (r_k) filled by water in capillary condensation as given below;

$$r_k = -\frac{2rV_m}{RTl_n(^p/_{po})} \quad (2.5)$$

r= surface tension of water

r_k = Maximum radius of pores filled with water

p/po = Relative humidity, T= Absolute temperature

R = Gas constant

Also the condensation is made possible when films of water molecules overlays the matrix pore walls to a thickness t_n, which decrease the real pore radius r_p of the hydrated cement paste.

$$r_p = r_k + tn \quad (2.6)$$

2.2.1. Transportation mechanism of carbonation in concrete

Carbonation transport mechanism is characterized by the physicochemical model. It involves transport of liquid and gas in pore spectrum. The microstructure properties of the cement matrix such as pore size distribution, level of porosity, connectivity of the pores, specific surface area are dependent factors that influence carbonation transport in hydrated cement paste and concrete elements (Morandeau et al., 2014). Carbonation under normal environmental conditions of CO_2 concentration and RH is largely controlled by diffusion through the empty pores in the exposed surface layer (Parrott, 1991) driven by concentration, and in finer pore sizes capillary absorption controls the movement with is based on the surface tension. Apparently in larger pores with high liquid concentration, suction due to pressure gradient is the prime mechanism (Hanžič et al., 2010) the acid attack is normally carried out by a combination of absorption, permeability and diffusion mechanism through the distribution and size of the microstructure in the matrix. Thiery et al. (2012), reported

that the pore matrix is controlled by water/cement (w/c) ratio and that carbonation is capable of producing large capillary pores for high w/c. the transportation process begins with the ingress of CO_2 in into the hydrated cement paste matrix by diffusion mechanism (Houst, 1996) which moves inward in solution with the help of moisture content of the pore spaces producing a poor solute (Castellote et al., 2008).

Initial carbonation is usually faster and its rate begins to decrease as carbonation process modifies the microstructure of the hydrated cement paste. Since the CO_2 diffuses further in the concrete paste to regions having lower concentration, its rate is dependent on the porosity of the hydrated paste and relative humidity. In Houst (1996) investigation, the diffusivity of CO_2 is influenced by cement content, w/c ratio, degree of hydration and independent of the pore size for large pores "$\theta \geq 450$ nm for CO_2, at 20^{o}C and 1 atm" and proportional to the pore diameter for finer pores ($\theta \geq 45$ nm), the ambient relative humidity and pore sizes distribution controls the humidity of the pore space (water content) which governs that gaseous diffusion in free volume. The report further analysed the movement of CO_2 using the Flick's first law. A one dimensional diffusion of a gas passing through a porous system as is given as;

$$J = D_e \frac{\Delta C}{d} \tag{2.7}$$

Where J = flux of the gas; D_e = effective diffusion coefficient; ΔC = concentration of CO_2 in air that makes contact with the material; d = the depth of carbonation.

In Conciatori et al., (2008), estimation of *d* takes into account the molar concertation of carbonation reaction, the atmospheric concentration of carbon dioxide. The CO_2 diffusion coefficient is predominantly influenced by concrete permeability, the moisture content and the chemical reaction rate in the concrete pores.

$$d = \frac{\sqrt{2 \cdot [CO_2] \cdot D_{e\,Co_2}}}{[Ca(OH)_2] + 3 \cdot [CSH]} \sqrt{t} \tag{2.8}$$

Where $[CO_2]$ = molar concentration of CO_2; $[Ca(OH)_2]$ & [CSH] = molar concentration of calcium hydroxide and silicate; $D_{e\,CO2}$ = CO_2 diffusion coefficient; t = time

The rate of permeability governs the ingress rate of CO_2, and the degree at which the pores is said to be permeable is a function quantity of $Ca(OH)_2$ available to react with percolated CO_2 (Lammertijn & De Belie, 2008).

Beside the permeability properties discussed above, relative humidity is another factor that affects the transportation of mechanism of carbonation in hydrate cement paste. All other factors can be said to be directly govern by this two. More so, it has been established that

carbonation processes cannot take place in the absence of water. Thiery et al. (2007), reported that a low relative humidity induces dehydration of the capillary pores and in turn reduces the rate of the CO_2 dissolution-dissociation and the dissolution of hydrates, while high relative humidity causes capillary water condensation in the pores and reduces the rate CO2 diffusion into the concrete pores. However, it has been confirmed that a relative humidity between 50% and 70% enhances an optimum carbonation transportation mechanism which create a partial moisture content in the pores and enhance phase diffusion of CO_2 gas and formation of carbonic acid.

2.2.2. Modification of microstructure of cement paste

Development of microstructure of a concrete matrix starts from the first stage of concrete production, which chemically sets a transformation from fluid to plastic phase within the first few days after mixing with water. The amount of water used in concrete production is important factor to consider which affects first the workability, strength and porosity of a produced concrete structure. Hydroscopic and hydrophilic properties of cement paste and the presence of sub-microscopic pores in cement with respect to ambient humidity also contributes to increased water content of cement paste (Neville, 2005). Water added during mixing is partly consumed in the chemical reaction of cement paste, while the unused water wither bleeds to the surface or trapped within the concrete mixture. Dehydration in hardened concrete paste leads to loss of trapped water either to the atmosphere or used up incurring process thereby creating pores within the concrete matrix. From the forgoing it is evident and has been confirmed that the amount of trapped water is proportional to degree of porosity and transport properties of hydrated cement phase. Also air bubbles trapped within the concrete pore may contribute to pore connectivity. Carbonation reactions leads to restructuring of the microstructure with decrease in porosity caused by formation calcium carbonate crystals. Formation of calcite result to decrease in amount of portlandite, ettringite and C-S-H gel of cement paste phase (Castellote et al., 2008). Carbonation produces a clogging of pores within its zones (Auroy et al., 2015) as a result precipitated carbonates which has a low solubility and causes an expansion in volume of the pores and development micro-cracks in carbonated zones (Johannesson & Utgenannt, 2001), the precipitates causes a loss in pore connectivity.

Castellote et al. (2008) reported that the rate of change which occurred on each phase different in carbonation process (Portlandite, Ettringite and CSH phases) disappeared with respect to an exponential decay of first order, with formation of calcite.

In summary, it is not worthy to state herewith that carbonation destroys the interconnections of porous network with a reduction in gaseous diffusivity coefficient with the carbonation zone (Castellote et al., 2009) and is attributed to increase in the formation of calcite from carbonation of C-S-H at high pressure (Hyvert et al., 2010).

2.2.3. Testing Methods used for carbonation problem in concrete

a. Phenolphthalein: This has been the oldest method known to detecting carbonation in concrete. When core samples are extracted from the concrete structure, a solution of diluted phenolphthalein in alcohol is spread over the sample. Region coloured in pink represents the free $Ca(OH)_2$ while the other uncoloured region is the carbonated portion. According to RILEM (1988), precautions should be taking while measuring depth of carbonation and measurement taken within the a series of tests and to the nearest 0.5 mm, where depths less than 0.5 mm are not differentiated.

b. X-ray Diffraction: one of the early used laboratory techniques. XRD used for identification of atomic and molecular structure of crystals, where a crystalline atom causes a beam of incident X-rays to diffract in different directions. The angles and intensities of diffracted beams imprint a 3-D image of density of electrons within the crystals. Mean positions of atoms in the crystals, their chemical bonds and other information.

c. Infrared Spectroscopy (Spectrophotometer)

Fourier Transformation Infrared Spectroscopy (FT-IR)

The application of infrared spectrum fingerprint of molecules which identifies elements by their unique absorption of infrared radiation. In FT-IR spectroscopy, characterization is done by passing a sample through and infrared radiation each element resonates according to its absorption frequency with the electromagnetic spectrum region and all frequency are measured simultaneously. Fourier transformation is applied to the signals in spectral through plotting absorption against each wavelength.

According to Lo and Lee (2002), FT-IR characterization of concrete element is a complex instrumentation analysis, because large number of characteristic peak may be found and at the same time some overlapping peaks may characterize different functional groups.
IR has also been used in detecting phase transformation of in hardened concrete (Gao et al., 1999), early stage hydration of OPC (Mollah et al., 2000; Yimén et al., 2009), while Lo & Cui, (2004) studied the phase transformation characteristics of the Interfacial Zone using FT-IR.

d. Thermogravimetric Analysis (TGA): TGA is classified under thermal methods of analysis in chemical instrumentation. Thermal analysis is a technique in which physical properties/reaction products of an substance is measured with respect to temperature under a controlled temperature programme (Mackenzie, 1979), while TGA is an analytical method in which the continuous record of the mass of a sample in a controlled atmosphere is made and which is a function of the linearity of temperature/time of the reviewed sample on (Skoog & Leary, 1992) thermal decomposition curve (Earnest, 1984). This method quantifies the calcium carbonates and CH contents in a sample of carbonated concrete (Parrott & Killoh 1989; Platret and Deloye 1994; Villain et al., 2007), it gives accurate quantitative analysis of the chemical phases linearly against depth of carbonation.
Samples are taken from power extracted from sawn slice to avoid mixed with aggregate particles because calcite of limestone sand contaminates the result of calcite resulting from carbonation, hence Chemical analysis is combined with TGA on the sample for evaluation of cement and sand content of the studied concrete mix (Villain et al., 2007; Thiery et al., 2007)

e. Chemical Analysis(CA): This method of analysis is employed to proportion the mineral phase and gradation of carbonated and un-carbonated concrete sample (Villain et al., 2007).

f. Gammadensimetry: It is a non-destructive test procedure is based on absorption of the gamma rays produced by a radioactive source of Cesium Cs^{137} with respect to Lambert law:

$$N = N_o \exp(-\mu\rho l) \quad (2.9)$$

From the above equation, and with a given data of N, N_o, μ, and *l*, the density ρ, can be estimate hereunder;

$$\rho = \frac{-1}{\mu l} \ln\left(\frac{N}{N_o}\right) \quad (2.10)$$

g. Mercury Intrusion Porosimetry (MIP): The mercury intrusion porosimetry method, provides information on the distribution of pore. Because mercury is non-wetting and liquid at normal temperatures, liquid mercury is intruded in the pores under high pressure, its pore sizes are quantify from the relationship between the volume of intruded mercury and applied pressure. However, it is also important to note that this method do not measure the shapes and location of pores (Tanaka & Kurumisawa, 2002).

2.2.4. Factors affecting concrete performance against carbonation problem

Carbonation been a major contributor to deterioration concrete structures especially reinforced concrete in dependant of other physical, chemical and atmospheric factors to undermine the performance of concrete structures. With respect to location, condition of exposure of concrete surfaces and the prevailing micro and macro climatic condition, the effect of carbonation on concrete can be evaluated. These micro factors are grouped into 5 major categories;

1. Climatic
2. Physical
3. Chemical
4. Transportation Mechanism
5. Material

a. Climatic Factors

i. Temperature: Studies has shown that temperature is the one of the principal micro factors affecting the performance of concrete structures. Temperature of mixed, placed and hardening and hardened concrete control its mechanical, physical and chemical properties (Ma et al., 2015; Karagol et al., 2015). Most compound react vigorously at higher temperature, corresponding linearly with the rate of the reaction. Temperature rise, increases the value of collision fraction exponentially needed to produce required kinetic energy to overcome the barrier at *transition state* (Activation Energy E_a) between reactants and products (Burns, 2003; Burrows et al., 2009; McMurry & Fay, 2008). In concrete (cement paste reaction), individual reactions of C_2S, C_3S and C_3A has been evaluated by different researchers (Ciach & Swenson 1971; Kamiński & Zielenkiewicz, 1982; Berhane, 1983; Bensted, 1983). As defined in ASTM C 403-92, higher ambient temperature accelerates the setting time in concrete and increased demand of water in fresh concrete resulting to plastic

shrinkage cracking and crazing with crack ranging from 0.1 to 3 mm developed over a length of 1m (Neville, 2003). More so, concrete placed in lower temperature less than 13 ^{o}C pose a threat to the strength and internal structure development of concrete due to freeing thawing action. In arid regions, ingress of water into these micro cracks under freezing-thawing action will further create more connection between pores. Diffusion of CO_2 in aqueous solution do not follow the trend of temperature is directly proportional to rate of reaction, rather the solubility of CO_2 decreases with increasing temperature at a given pressure (Jödecke et al., 2015), with reduction in concentration of dissolved CO_2, causing the equilibrium to shift to the left (Practical Chemistry, 2015) of the Equation 1.

ii. Relative Humidity: The assessment of concrete structures' durability requires a defined evaluation of moisture transportation throughout the service life of the structure, with the analysis focusing computation of water flow in unsaturated condition, water sorption isotherm and the permeability characteristics of the hardened concrete (Drouet et al., 2015).

The diffusion coefficient of water in porous materials (hardened concrete) is strongly dependent on temperature which involves an activation process, and the characteristic energy depends on the temperature range (Glover & Raask, 1972). But its retention within the pores is also affected by the temperature (Drouet et al., 2015).

Table 2.1: The solubility (S) of CO2 in water at different temperatures and pressures (Liu et al., 2011)

The solubility (S) of CO2 in water at different temperatures and pressures.							
35.0 °C	45.0 °C	50.0 °C	55.0 °C				
P/MPa	S/wt%	P/MPa	S/wt%	P/MPa	S/wt%	P/MPa	S/wt%
2.1	2.25	2.08	1.81	2.1	1.65	2.86	2.01
4.09	3.76	4.1	3.13	4.11	2.81	4.37	2.85
6.08	4.83	6.09	4.08	6.12	3.75	6.11	3.59
8.09	5.3	8.11	4.8	8.1	4.32	8.48	4.28
10.08	5.47	10.08	5.14	10.1	4.75	9.99	4.56
12.05	5.61	12.06	5.28	12.04	4.95	12.2	4.88
14.01	5.79	14.11	5.36	15.99	5.12	13.19	4.99
15.83	5.88	15.86	5.43			15.23	5.05

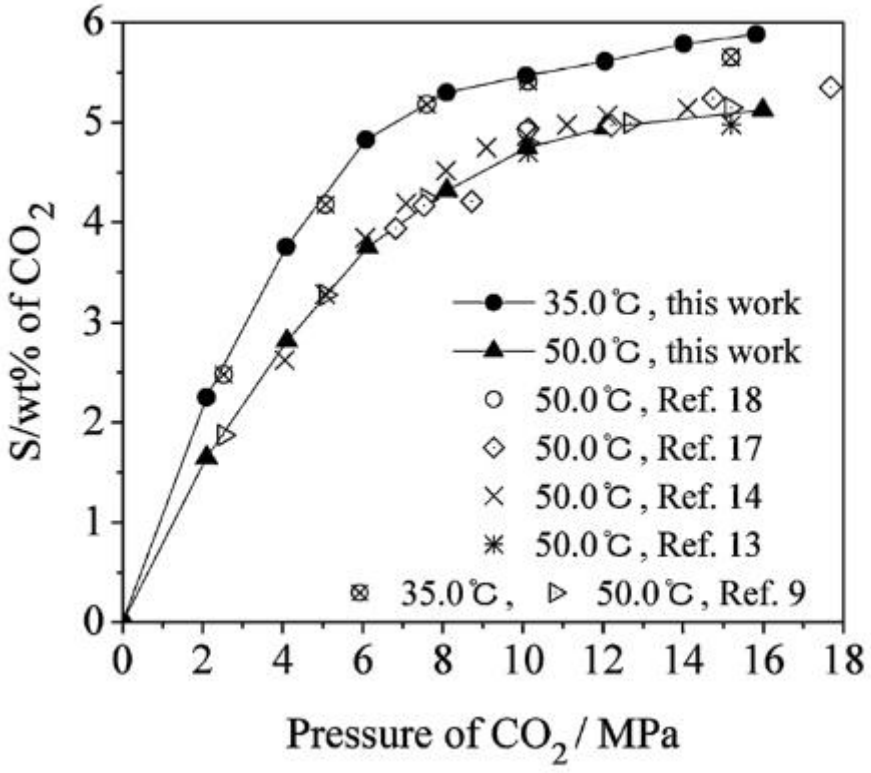

Figure 2.1: Comparison of CO2 solubility in water at 35.0 ◦ C and 50.0 ◦ C and pressures to 18.0MPa determined by different authors (Liu et al., 2011)

Although, water content of both internal and external wall surfaces of concrete pore is been control by ambient relative humidity, and at equilibrium of a given material they at equal (Wu et al., 2014).From the foregoing; that water held within the matrix is grossly a function of the percent of ambient relative humidity and temperature. While the former provided the provide the means (water) the later effect its movement into the pores. As stated earlier, carbonation is largely driven by the amount of water present for the chemical reaction to take place, thus the prevailing relative humidity data must be handy from pre-design to maintenance stage for effective designing of protective measures needed for optimum performance of the structure. In summary, increase in temperature at a constant RH results to equilibrium change between the adsorbed phase (exothermic process) and water vapour (endothermic reaction), water consequently is released (Drouet et al., 2015).

b. Physical Factors: Microstructure: The size, distribution, surface characteristics and the connectivity of hardened concrete pores are of importance in determination its effect on

performance of concrete structure subjected to carbonation. Duan et al. (2013), grouped pore development in hardened cement paste matrix into four;

i. Gel pores: they are micro-pores between 0.5 to 10 nm in dimension.
ii. Capillary pores: meso-pores of radius 5 to 5000 nm.
iii. Entrained air Macro-pores; developed as a result of entrained air in fresh concrete
iv. Compaction macro-pores: resulting from inadequate compaction.

Other microstructural development identified by Duan et al. (2013), are cracks around aggregate, cracks developed due to shrinkage. The microstructure affects the permeability, frost resistance and mechanical characteristics performance of concrete (Duan et al., 2013, Song and Kwon 2007), as it affects the gaseous diffusion and liquid permeability in concrete (Tanaka & Kurumisawa, 2002).

Moreover, among research in cement paste and related field, it has been agreed upon that the microstructure of is generally controlled by water/cement ratio. For a higher w/c produces more pores and connectivity with decrease in strength and a lowered w/c of below 0.4 produces a more durable concrete.

Permeability and Diffusion Coefficient: Permeability and diffusion coefficient has been discussed in 2.2.1. It is also noteworthy to further give details on their effect on concrete durability performance. In hardened cement paste, the coefficient of permeability subjected to carbonation decreases with time, this is due to formation of hydrates from reaction of $Ca(OH)_2$ with carbonic acid formed resulting to densification of the pores (Song & Kwon, 2007).$CO_{2(g)}$requires a gaseous phase of 10^{12} m^2 s^{-2} coefficient for ingress action into the concrete in that in liquid phase, the diffusion coefficient of 10^4 m^2 s^{-2} is smaller when compared to gaseous phase (Baroghel-Bouny, 2007).

Porosity: Porosity in concrete is govern by different parameters such the clinker composition, reactivity of aggregates, modulus of fineness of aggregates, water/cement ratio and admixtures used. In a typical OPC concrete mix, aggregate comprises of 75% by volume while cement takes the 25%. In a traditional mix, pore sizes are highest with high occurrence of porosity at the Paste-aggregate interface (Grattan-Bellew, 1996).

Interfacial Transition Zone (ITZ) in concrete is considered as a weak zone with respect to strength and porosity. Its microstructure is determined by the packing of the anhydrous cement particles against the bulk particles of the aggregates (Scrivener & Kamran 1996), which control the pore sizes in a concrete mix. It has been observed that the due to anhydrous

cement packing at the aggregate surface, the water/cement ratio is found to be higher that the bulk cement paste in concrete. (Laugesen, 1993) and decreased stiffness (Cohen et al., 1994).

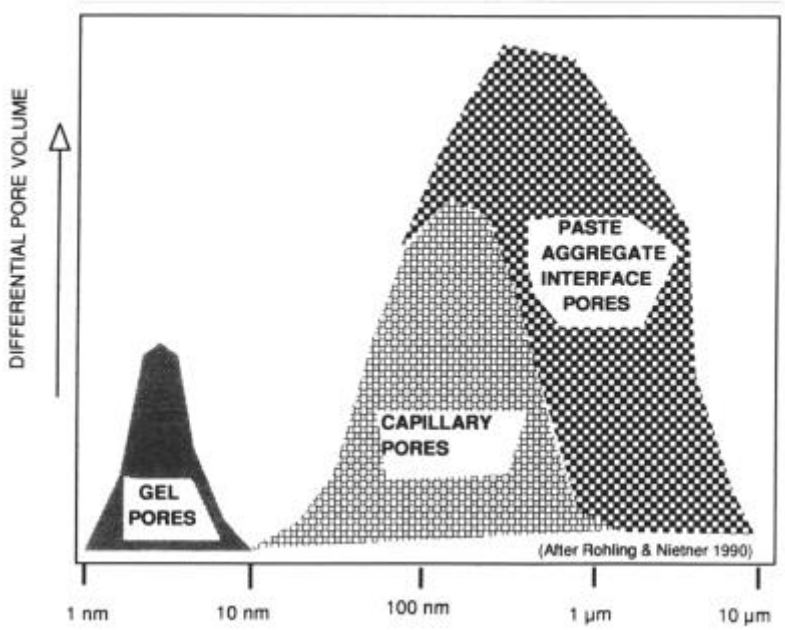

Figure 2.2*:* Diagram showing the large contribution of pores at the paste-aggregate interface to the total porosity in concrete (Grattan-Bellew, 1996)

Ferdi et al. (2008), in Handbook of Porous Solids presented different methods of characterisation of porosity in solids. MIP is widely accepted method of measuring pore profile of harden concrete specimen through analysis of its percolation and pore diameter distribution. Although Diamond, (2004), believes that though MIP may characterize pore profile of less than 0.1 µm, SEM (Backscattered Electron) gives a better validation.

In MIP characterization pore profiles are assumed to be circular and evaluated using Equation 2.11 **(**Washburn, 1921), where radius r (assumed to be cylindrical), is the radius of pores P is the imposed pressure, γ (surface tension) is the interfacial energy of mercury and the contact angle of mercury with the material is θ Lawrence et al. (2007).

$$p = \frac{-2\gamma Cos\theta}{r} \qquad (2.11)$$

SEM examination detects spherical air voids found in cement-paste system with no air-entraining agent, irregular pores in high w/c (Diamond 2004) such found in ITZ (Laugesen, 1993).

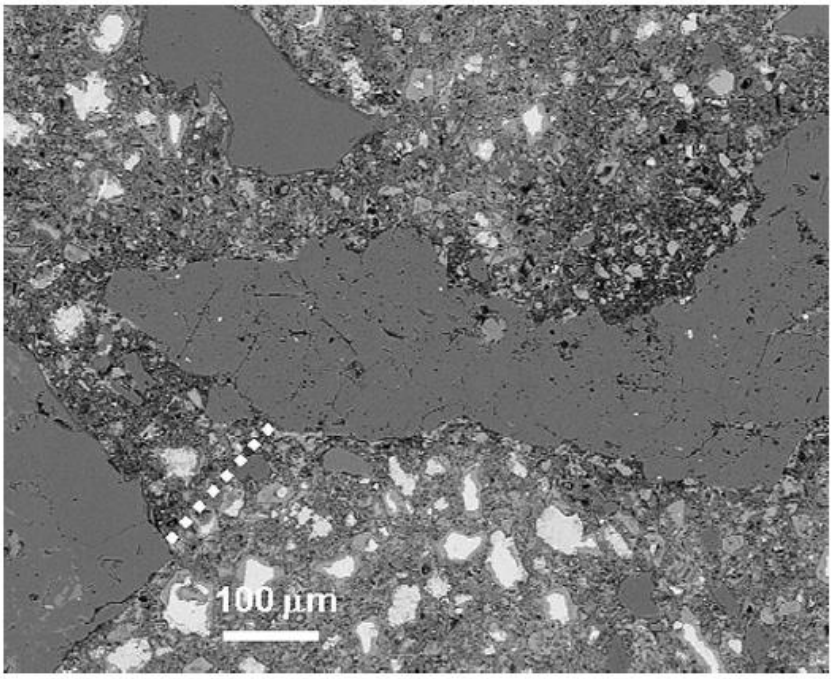

Figure 2.3: Area showing dense and porous patches in a laboratory-mixed w:c 0.50 concrete hydrated for 28 days (Diamond, 2004)

c. Chemical Factors: Chemical Equilibrium Effect on Carbonation: In chemical reactions of both reactants and products that moves towards a dynamic equilibrium has much significant on the concentration of both product and reactants but the concentration remains unchanged in equilibrium mixture (Atkins, 1994). Carbonation chemical reaction is dependent on two major factors; moisture and concentration of the solute.

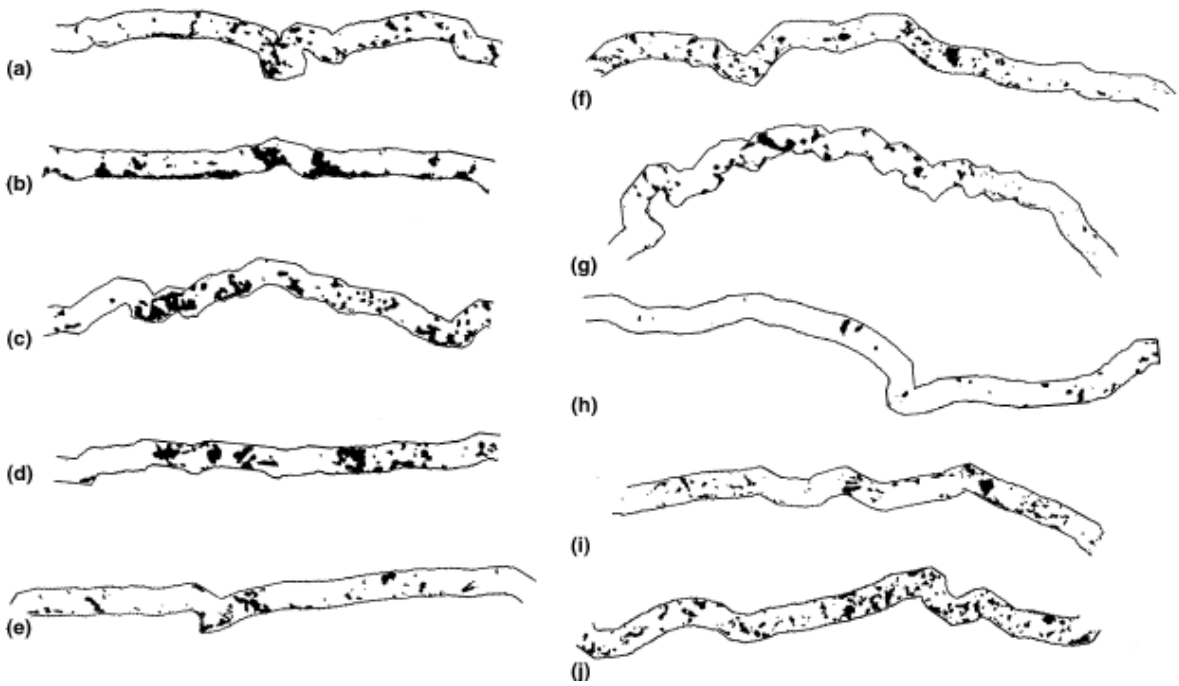

Figure 2.4: Binary segmented images showing the distribution of pores within cement paste sampling units 0 to10 µm from the aggregate surface in three-day-old quartzite aggregate concrete (Diamond, 2001)

Visser (2014) stated that though the carbonation reaction of ambient CO_2 concentration is a slow process, and increase in concentration can only accelerate the rate of diffusion of CO_2 without affecting the chemical reaction and transportation mechanism, provided the concrete maintains a sufficient degree of dryness to allow gaseous diffusion.

According to Burrows et al., (2009), CO_2 is a weak acid which reacts with water to form carbonic acid (H_2CO_3), a further dissociation will produce hydrogen carbonate (HCO_3^-) and carbonate (CO_3^{2-}) anions.

$$CO_{2(g)} \rightleftharpoons CO_{2(aq)} \quad (2.12)$$

$$CO_{2(aq)} + H_2O_{(l)} \rightleftharpoons CO_{2(aq)} \quad (2.13)$$

$$H_2CO_{3(aq)} + H_2O_{(l)} \rightleftharpoons HCO_3^{2-}{}_{(aq)} + H_3O^+{}_{(aq)} \quad (2.14)$$

$$HCO_3^- + H_2O \rightleftharpoons CO_3^{2-}{}_{(aq)} + H_3O^+{}_{(aq)} \quad (2.15)$$

At a constant temperature, the reaction is largely dependent on the magnitude of the equilibrium constant *K*, and when the reaction reaches equilibrium, the concentration of each species remains constant. The chemical equilibrium is delicate in that a change in pH changes the concentration of H_3O^+ ions and alters the equilibria of the 3 other species (H_2CO_3, HCO_3^- and CO_3^{2-}).

From the above equations (10), (11), (12) and (13), it can be deduced that formation of carbonic acid in carbonation reaction is in dynamic equilibrium. K (Thermodynamic equilibrium constant) is given as;

$$K = \frac{\Pi(a(Products))_{eqm}^{V_p}}{\Pi(a(Reactants))_{eqm}^{V_R}} \quad (2.16)$$

$\prod$= Product of

a = Equilibrium activities of the products and reactants

V_p = stoichiometry of products
V_R = stoichiometry of reactants

According to Henry's Law, the solubility of a gas is directly proportional to the pressure over the solution (Burrows et al., 2009), and approximate atmospheric pressure at sea level is 760 mmHg (Ebbing & Gammon, 2005), of which 0.03% represents the partial atmospheric pressure of CO_2 (Castellote et al., 2009; Hyvert et al., 2010.). It has been proven that the increase in pressure (accelerated carbonation) increases the rate of carbonation with respect pressure/CO_2 solubility (Liu et al., 2011).
Therefore, chemical equilibrium of CO_2 polymerisation reaction is a key factor in determining the rate of carbonation as it affects the concentration of aqueous CO_2 with respect to concentration of gaseous CO_2. Also the pressure with the solution and the atmospheric pressure affects the production of aqueous carbon dioxide (Liu et al., 2011).

d. Transportation Mechanism: The distinction between porosity and permeability in concrete is important in order to appreciate structure of transport mechanism and their interdependency. Porosity is associated with the percentage of pore occupied in a given concrete volume. Neville (2005), noted that disconnected pores contributes to low

permeability *vis-à-vis* high permeability and should be between 120 -160 nm in diameter to support fluid transportation, and further defined permeability as the flow due to differential pressure. Furthermore, connectivity of pores in concrete paste is responsibly to transportation in concrete paste. Sequel to presentation of high porosity of ITZ in hardened concrete, Larbi (1993), as presented in Neville (2005), argued that despite the high porosity of the interfacial zone, the permeability is determined by the continuous phase of the bulk hardened cement paste present in concrete. Sidney Diamond in two separate publications in 2001 and 2004, clearly deviated from the typical characterisation of ITZ (*aureole de transition*) in ordinary concrete. Diamond (2004) considered the acclaimed properties ITZ as a function of local deficiency of cement particles closed to the aggregate in fresh concrete after mixing.

Acclaimed properties of ITZ may be as a result of poor quality in concrete production. Diamond (2001) concluded that the effect of heterogeneous nucleation and crystallisation growth of CH along aggregate and porosity is not different from bulk regions away from ITZ and shows no evident of concentration of pores in the cement paste with few µm of the aggregate.

Consequently, transportation mechanism can be view to be a function of connected pores usually due to high w/c leading to bleeding, entrained air and gross packing density. The mechanism controlling the absorption in porous concrete include:

Absorption: This takes place in partially dry concrete at low relative humidity. Since diffusion and capillary action is the key transport mechanism in porous concrete, the can be a very slow process leaving out capillary and the main culprit of transportation of deterioration agents into concrete in partially saturated concrete surface.

Diffusion: Diffusion herewith is as a result of differential concentration gradient in pores. Theoretically diffusivity coefficient of a gas is inversely proportional to square root of its molar mass (Papadakis et al., 1991). $CO_{2(g)}$ ingress into hardened concrete is by diffusion and controlled by relative humidity, porosity and temperature. Also movement of water vapour under differential concentration occurs due to differential humidity on two opposing faces, an increase in relative humidity reduces the available air-filled pores for diffusion. Similarly, chloride and sulfates are transported by diffusion in pore water (Henry & Kurtz 1963; Neville, 2005).

Diffusion coefficient is given by;

$$J = -D\frac{dc}{dL} \qquad (2.17)$$

Where $\frac{dc}{dL}$ = concentration gradient in kg/m^4 or moles/m^4; D = diffusion coefficient in m^2/s; J = mass transport rate in kg/m^2 s (moles/m^2 s); L = thickness of the sample in metre.

e. Material: Materials constitute the essential ingredients of concrete, in fact with materials there will be no concrete. The understanding of properties of individual constituents of concrete is essential prerequisite for designing and construction of concrete structures. According to Neville and Brooks (2010), concrete and steel are the major construction materials in use with the former requiring expertise in its construction since is not premade like the later. It therefore demands an extensive care in all the processes involved from batching to curing of constructed elements.

With respect to carbonation, this research work shall only focus on cement in that the effect of other materials other than cement can be controlled and minimized through appropriate selection of materials.

Modern cement is believed to have been invent in 1756 in the experiment of John Smeaton for production of mortar needed for construction foundation and masonry of the Eddystone Lighthouse, where hydraulic material was produced from mixture of Aberthau blue lias, South Wales limestone and Italian pozzolana. Structure like Brunel's Thames Tunnel and Stephenson's Britannia Bridge foundation were constructed with cement (Illston, 1994). Before the patency of OPC by Joseph Aspdin in 1824 (Neville & Brooks, 2010), other nineteenth century contributors include M. Vcat and James Frost (Raina, 1990).

Cement Composition: cement is made chiefly from Lime and Silica (Illston, 1994), bauxite for high-alumina cement (Neville, 2005). The argillaceous and calcareous including other materials are partially fused at about 1450 ^{o}C (in the kiln) to form clinker, and grinded to range of 2 – 80 µm on cooling with gypsum (Gambhir, 1995; Illston 1994). Chatterjee, (2011) sum up the process as follows

a. Dissociation of limestone
b. Solid-state reactions
c. Liquid-phase sintering
d. Reorganisation of clinker microstructure through cooling.

The review of the following cement type properties will grant us a better understanding in its effect on carbonation process.

i. Ordinary Portland Cement: Portland cement is a product from the partial fusion with the production of "nodules of clinker" between the mixture of limestone and clay or other similar materials at about 1450 ^{o}C. The clinker is further mixed with calcium sulphate which controls the rate off set and also influences the rate of strength development (Talylor, 1997). Since 1843 of its first production by William Aspdin, there has been an evolution in technology of production and cement chemistry. The development of new cementitious binders with view of reducing CO_2 emission for production of sustainable construction materials have in the past and present decay seen production of blended cement; Supplementary Cementitious Materials (SCMs).

SCMs already have partly replaced the conventional Portland cement with Portland cement clinker burnt in a rotary kiln been the main component of Portland-composite cements (Ludwig & Zhang2015).

ii. Clinker: The partial or complete fusion at high temperature (1450 ^{o}C) between mixed chalk, clay and other materials in a rotary kiln.

Portland cement clinker is a hydraulic material consisting of mainly, of calcium silicates ($(CaO)_3SiO_2$ and $(CaO)_2SiO_2$), aluminium oxide (AL_2O_3), iron oxide (Fe_2O_3) and other oxides (Hewlett, 2004). The mass composition of clinker constituent as reported by Telschow (2012) shows that 40-80% C_3S, 10-50% C_2S, 0-15% C_3A and 0-20% C_4AF.

Production of Portland Cement Clinker (PCC), takes place in a heated rotary kiln inclined to the horizontal (1 – 3 ^{o}C) heated from the lower, calcination process preceding the clinkerisation occurs between 800 to 900 ^{o}C, and as the raw mixed material moves down the kiln and at 1250 ^{o}C solid state reaction occurs with gradual formation of belite, aluminate and ferrite. Towards the lower end of the kiln of between 1300 to 1500 ^{o}C formation of sticky solid particles from granulation/nodulation of molten aluminate, ferrite and ferrite and some quantity of belite. Alite present in clinker is then formed from free CaO and belite. Re-crystallisation of the finely grained aluminate phase is done from fast cooling at 1200-1250 ^{o}C, which enhances slow controllable hydration reaction of cement (Bye, 1983; Taylor, 1997; Telschow, 2012).Furthermore, according to Telschow (2012) clinker it is generally composed crystal phases;

1. Calcium silicate phase contains alite (Ca_3SiO_5) at tricalcium silicate phase and belite (Ca_2SiO_4) at dicalcium silicate phase.

2. Aluminate phase formed from CaO and Al_2O_3 contains aluminate ($Ca_3Al_2O_6$) at tricalcium aluminate phase.
3. Ferrite phase produced from CaO, Al_2O_3 and Fe_2O_3 at tetracalcium aluminoferrite phase ($Ca_4Al_2Fe_2O_{10}$).

Supplementary Cementitious Materials (SCM)

Supplementary cementitious materials (SCMs) are widely used in concrete mixtures as a replacement of a percentage of clinker in cement or as a replacement of a percentage of cement in concrete. It is an age long practice in the construction industry, with evidence in lowered cost of concrete, environmental impact and higher long-term strength and improved durability. (Juenger & Siddique, 2015).

According ENV 197-1: 1992 the term CEM used in describing any cementitious materials containing certain percentage of Portland cement. See table below for percentage of Portland cement clinker in various types of cement as described by EN 197-1:1992. Though ASTM C 1157-94a, describe non Portland Cement as "blended hydraulic cement"

"*which consist of two or more inorganic constituents that contributes to the strength-gaining properties of the cement, with or without other constituents, processing additions and functional additions"* (Neville, 2005).

Table 2.2: Classification of main cements according to EN 197-1:1992 (Neville, 2005)

Types	*Designation*	*Mass as percentage of mass of cementitious materials*			
		Portland Cement Clinker	*Pozzolana or fly ash*	*Silica fume*	*ggbs*
I	Portland	95-100	--	--	--
II/A	Portland slag	80-94	--	--	6-20
II/B		65-79	--	--	21-35
II/A	Portland pozzolana or Portland fly ash	80-94	6-20	--	--
II/B		65-79	21-35	--	--
II/A	Portland silica fume	90-94	--	6-10	--

II/A	Portland	80-94	6-20	6-20	6-20
II/B	composite	65-79	21-35	21-35	21-35
II/A	Blastfurnace	35-64	--	--	36-65
II/B		20-34	--	--	66-80
II/C		5-19	--	--	81-95
II/A	Pozzolanic	65-89	11-35	11-35	--
II/B		45-64	36-55	36-55	--

SCM could be in form of ternary blended and binary blended cements, the later consisting of only the percentage blending of SCM with PC.

Ternary blended cement consisting of Portland cement, granulated blast-furnace slag and fly ash (PC–SL–FA system has been reported to improve the performance of concrete as compared to conventional Portland cement and binary blended cements.(Uchikawa & Okamura, 1993; Khalil & Anwar, 2015). The addition of fly ash can increase workability and reduce bleeding of slag cement concrete. (Berry, 1980; Khalil & Anwar, 2015)

In their inference, Khalil and Anwar (2015) observed that the rate of carbonation is inversely proportional to the strength of concrete irrespective of it cementitious content.

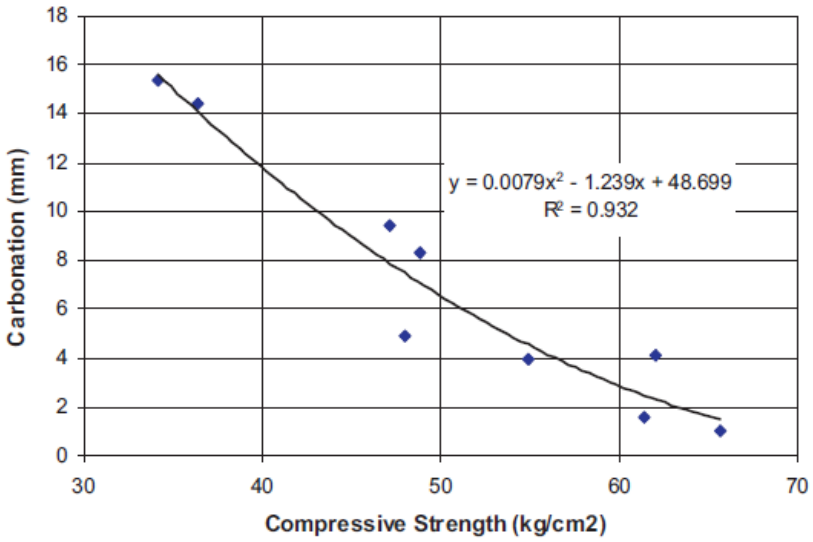

Figure 2.5: Carbonation vs compressive strength.(Khalil & Anwar, 2015)

CHAPTER 3
LITERATURE REVIEW ON ARTIFICIAL NEURAL NETWORKS

3.1. Evolution of Computation

Throughout history the human brain has been the only source of computation, though it might seemed to be slow in performance of complex computation but its values can never be equated to machines.

Origin of computation is dated back around 400 to 300 B.C, when Greek mathematician Euclid invented greatest common divisor (gcd) of two positive integers and in ninth century when Mohammed al-Khowârizmî, provided the step-by-step rules of add/subtract/multiply& divide of ordinary decimal numbers. The name *algorithm* gotten from Latin written *Algorimus*for al-Khowârizmî (Harel & Feldman, 2004).

Other computational methods include abacus, slide rule, the log tables etc.

3.1.1. Artificial intelligence

Overview of Intelligence

Artificial intelligence (AI) borders on computation that requires computer program to perform a similar function to that of a human brain. Depending on the task and amount of input data, AI could be complex and requires advanced reasoning.

For years long, human rely on *general* or *common-sense* knowledge to gain experience and make prediction.

Declaration knowledge and *procedural* knowledge, generally are factual processes employed by humans in identifying and carrying out a specific task. For a particular experience the *general* or *common-sense* knowledge acquired provides the *domain-specific knowledge;* our *general* or *common-sense* knowledge tell us that parents are older than their children (Finlay & Dix, 1996)

Neuroscience

According Mackay (1967), among “all the natural phenomena to which science can turn its attention none exceeds in its fascination the working of the human brain”

The human brain encompasses of complex organic living tissue, of about 100 billion neurons or nerve cells (Beatty, 1995), 10 percent of the cells (neurons) are dedicated for conduction of electrical signal while *glial* or glue cells forms the other 90 percent of support cells to neurons (Bose & Liang, 1996). The neurons are linked together through a cellular contact of over 10 trillion connections (Beatty, 1995).

The neurons is divided into three parts, *the Cell body, (or soma), an axon, and dendrites.*

The cell body is located at the center of the neurons and contains the nucleus of the cell in association with the cell's genetic materials. It also provide the molecular synthesizing mechanism; for transfer of information, repair and maintenance of the cell including excretion of bye-products from the cell (washington.edu; retrieved January 30, 2016).

The axon on the other hand are *excitable membrane* that connects the cell body and the regions of synaptic contact together. They have the capabilities of generating and transmitting a distinctive electrical potential response (*single solitary traveling pulse of action potential*) along the entire length of the axon (Katz, 1966; Hille, 1984; Beatty, 1995).

The dendrites is a treelike form of extensions from the cell body as can be seen in fig 3.1 below. Beatty, (1995) inferred that *dendritic spread* pattern can predict the functional properties of a neuron. Dendrites receives information through chemical in the form of chemicals discharge from axon terminals of adjacent neurons.

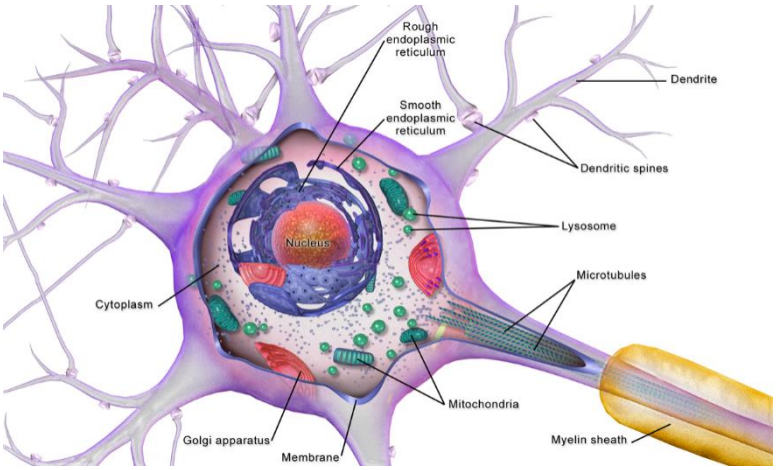

Figure 3.1: Anatomy of a Multipolar Neuron (wikimedia.org-Blausen, 2016)

Synapses is connection point between and two neurons where impulse are exchanged. It can be chemical or electrical synapses depending on the location. According Beatty, (1995) depolarization or hyperpolarization of the cell membrane takes place at the *end foot* of an axon (synapses) as an electrical responses in the receiving cell to chemical inducement of

neurotransmitter (neuromodulator) releases from the axon. Electrically, synapses transmits and provides continuous and direct flow of ionic current between cells.

The complexity transduction (electrical-to-chemical transduction) processes of the transfer of impulse determines the *synaptic strength* while Synaptic plasticity measures the adaptive (learning) ability of nervous system to her environment acquired and strength true time and experience (Benfenati, 2007).

Neurons are highly polarized cells which can transmit only a digital stereotyped signal (the action potential) information over long distances that cannot be modulated in amplitude, rather in frequency, and the synaptic digital-to-analogic process enables the transmission of a highly *modulatable* identical signal across a synaptic cleft (Zucker,1996; Benfenati et al., 1999; Benfenati, 2007).

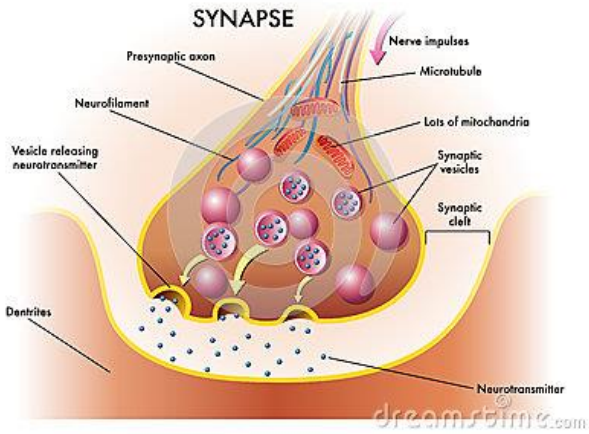

Figure 3.2: Schematic of a Synapse (dreamstime.com, 2016)

Synaptic cleft: is complex fluid secreted gap of about 20-30 nanometer wide between presynaptic and postsynaptic membrane (Beatty, 1995).

3.1.2. Biological neural network (BNN)

BNN consist of the neuron which receives and process information. Their network is complex as compared to ANN.

Biological neural network do not operate as a single independent unit, the co-exist as a large network (MoukamKakmeni & Nguemaha, 2016) of closely packed neuronal network.

The networks of neurons are dynamic but non-linear, which forms electrophysiological spatiotemporal patterns from rapid communication between neurons in the network (Gross et al., 1995).

To generate an input (signals), sensory receptors and effectors collect and transmit input signal from the internal and external environment through constant observation of occurrence changes the brain (neurons) for processing in a "*feed-forward-back system*" via a relay line of axons. Stimulated receptors propagates generates potentials (electrical charge) through ionic movement across the synaptic cleft (from presynaptic to postsynaptic neuron). For a *graded potentials* to occur at the postsynaptic neuron, the summation of transmembrane potential difference across the axon hillock must exceed a threshold in order to trigger excitation (depolarization) along the axon resulting from influx and leakage of Na^+, K^+ and Ca^{2+} within the cell membrane which generates electrical impulse (action potential). The strength of the spike is dependent on frequency. Threshold potential is approximately measured at -55 mV, where resting potential is -70 mV' therefore for depolarization (activation) to take place, an influx of about +35 mV is required from positively charged Na^+ into the cell needed to cause a release of *neurotransmitters* from the synaptic vesicles located at presynaptic membrane. An electrical potential is induced at the postsynaptic membrane due to diffused neurotransmitter across the synaptic cleft. This set of algorithm is a continuous process between neurons both in inflow of input and transmission of processed reaction (output) with the CNS (Kandel et al., 2000; Bose & Liang, 1996; Benfenati, 2007; Beatty, 1995).

3.2. Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) are particularly employed in decision and regression tasks as nonparametric estimator. Coined from mimicking the BNN, it has the capabilities of developing its own experience of the environment and use acquired knowledge in generalizing new targets for a given input set.

It is characterized by pattern connections between neurons (network architecture), calculation of its weight over the connections (training, or learning, algorithm) with its activation function (Fausett, 1994).

ANN consists of a large number of processing elements known as neurons, units or nodes. Individual have direct communication link with other neurons of the network with corresponding connection weights. The connection weights serve as the network memory

for storage and recalling of data or patterns. Within each neuron exist its activation level (activation), a function of input received and it is used to send signal to other neuron one at a time (Fausett, 1994).

Artificial neural networks has been applied to a wide range of problem in pattern and character recognition, performing general mappings, finding solutions to constrained optimization problems (Fausett, 1994).

3.3. Learning Process

Learning which basically means acquiring or receiving knowledge and instructions from the environment.

The Neural Network in mimicking the human pattern, also learning from its environment through a interactive process of adjusting its synaptic weights and bias levels and improves thereupon from successive iteration in accordance with prescribed measures (Haykin, 1999). Therefor learning in this context according to Mendel and McLaren (1970), is define as *"a process by which the free parameters of a network are adapted through a process of stimulation by the environment in which the network is embedded and determined by the manner in which the parameter changes take place".*

Experienced gained remains the control parameter in which the network interacts with its environment.

3.3.1. Learning rules

Error-Correction Learning: implies the minimization of a cost function or index function from sequential corrective adjustment of the synaptic weights (real-valued numbers) of neurons leading to the delta rule (learning rule). The corrective adjustment is applied until output signal converges to a desired response in a step-by-step manner (Belciug & Gorunescu2014; Haykin, 1999).

Memory-Based Learning: Is a learning algorithm who's desired respond variables depends on information (past experience) stored in a large memory from the previous input-output attributes (Haykin, 1999; Bennamoun, lecture note, retrieved on May 31, 2016).

$\{(x_i, d_i)\}_{i=1}^{N}$ where x_i = input vector; d_i = desired responds.

Hebbian Learning: It is feed forward, unsupervised learning algorithm with the learning signal equal to the neuron's output. Prior to learning the weights are initialization at small random values around w_i=0,. Each excitations results to either increase or decrease of the weights. Also in Hebbian rules, if two neurons are simultaneously activated, the connection strength is increase otherwise it decreases in an asynchronously activation. It also important to point out that the Hebbian synapses is time dependent, and learning takes place only when two signals interacts in a concurrent manner (Bennamoun, 2016).

Competitive Learning: Output neurons compete to get active with only a single neuron been activated at a time. Learning by a neuron takes place when synaptic weights shifts from inactive to active input nodes. This is possible if a synaptic weights are distributed randomly to a set of neurons which responds differently to subset of input signals. Only one neuron (winner-takes-all-neuron) is activated within a set after competing with other neurons with a limit imposed on the strength of each neurons. The pattern induces individual neurons of the network to specialize on ensemble of similar patterns in order to become a *feature detectors.* Potential winning neuron *k* induces a local field v_k for a specific input pattern x. The output signal y_k of winning neuron is set to one while others that lost is set to zero. The winning neuron also at its input node releases a some proportion of its synaptic weights which is redistributed equally among the active input nodes (Haykin, 1999).

$$y_k \begin{cases} 1 & if\ v_j > v_j\ for\ all\ j, j \neq k\ other\ wise \\ 0 & \end{cases}$$

$$\sum_j w_{kj} = 1 \quad for\ all\ k$$

$$\Delta w_{kj} = \begin{cases} \eta(x_j - w_{kj}) & if\ neuron\ k\ wins\ the\ competition \\ 0 & if\ neuron\ k\ loses\ the\ competition \end{cases}$$

$$\eta = learning\ rate$$

Boltzmann Learning: It is a stochastic learning algorithm named after Ludwig Boltzmann. A neural network learning based on Boltzmann rule is called *Boltzmann Machine.* It is characterized by an energy function, E, the value is set by a particular state occupied by the individual neurons of the machine. A neuron is choosing at random by BM, and with probability flips its state from x_k to $-x_k$ at a certain *pseudotemperature* temperature T. Learning takes place when the *machine* reaches thermal equilibrium if the process is repeated (Haykin, 1999).

$$E = -\frac{1}{2}\sum_j \sum_k w_{kj} x_k x_j \quad (j \neq k) \tag{3.1}$$

$$P(x_k \rightarrow -x_k) = \frac{1}{1+exp\left(-\Delta E_k / T\right)} \tag{3.2}$$

Where ΔE_k = energy change

3.4. Learning Algorithm

The process or rather the steps of learning herein is referred to as the learning algorithm. Neural networks learning depends on the activities that talks place in a network. It is function of collective actions of all neurons participating in the network.

In Janabi-Sharifi and Wilson (1993), learning is said to be adaptive, self-organized, self-repair and inductive and can be further be defined as a system during it interaction with the environment acquires an information or knowledge and applies acquired new knowledge in optimizing performance in subsequent operation of the system.

Considering Hebbian learning law, a synaptic connection between two neurons *I* and *j* increases its strength w_{ij} when through an input stimulus, the two neurons are repeatedly *simultaneously activated*. The change in weight Δw_{ij} is a function of the product of the two neurons (*i;j*) activation values (*ai; aj*) (Kasabov, 1996).

$\Delta w_{ij} = c.\ ai.\ aj$

The process is repeated with the network reacting positively until there is no more changes in the synaptic weight, hence the network is said to have stopped learning and is to refers to as network *convergence* (Kasabov, 1996).

Three learning algorithm is discussed hereunder;

a. Supervised Learning (Learning with Teacher)

The input and output signals (training data) comprising representing the knowledge of the environment which is unknown to the neural networks. It is an approximate process of a set of *training vectors* towards a desired respond by the network parameters iteratively (Haykin, 1999).

b. Unsupervised Learning (Learning without Teacher)

Unsupervised learning is a learning algorithm used where only the input datasets are used to estimate the desired responses. Is a self-adaptive process where patterns in the input space are identified, without direct feedback from a teacher or supervisor (Watanabe & Tzafestas1990; Kohonen, 1982, 1987). This algorithms use no knowledge of the eventual targets to understand its operation by observing the input data (Becker & Plumbley, 1996).

c. Reinforced Learning.

Experience has been the tool employed by infants to adapt and interact with their environment. Reinforced learning basically has four components; the policy (goal), reward function, value function and the model of the environment. If we consider these components in relation to learning, reinforced learning is learning through mapping (Sutton & Barto, 1998), through continuous interaction with the environment with goal of minimizing the scalar index of performance (Haykin, 1999).

The network systems configured to observe under *delayed reinforcement* a temporal sequence stimuli(*primary reinforcement signal*) from the environment, which is then transformed it a stronger signal know as *heuristic reinforcement signal*(Haykin, 1999).

3.5. The Perceptron

Perceptron as a supervised learning binary classifier, introduced by Rosenblatt, (1958, 1961). Is a single-layer network that can train weights and biases to produce a correct target vector from corresponding input vector (mathworks.com, perceptron-neural-networks, June 15, 2016). The network is a feedforward, three layered connection and graphically represent as a two layer network since the weights between the buffer layer and the second layer (*feature layer*) are constant.

Sensory data are stored in the first layer whose elements are fully or arbitrarily connected to the *feature layer* which function as a linear combiner of signals from different sensory element. The neurons from the output layer (*perceptronlayer*) are connected with neurons from the corresponding features layers. Therefore, learning takes place only when an input vector is misclassified from training example (Kasabov, 1996; Haykin, 1999).

3.6. Major Types of Neural Networks

a. Feedforward Networks

The main features of this network is that the connections between the input and out nodes are only in a forward direction. The networks do not store previous values of activations state of its neurons and output values.

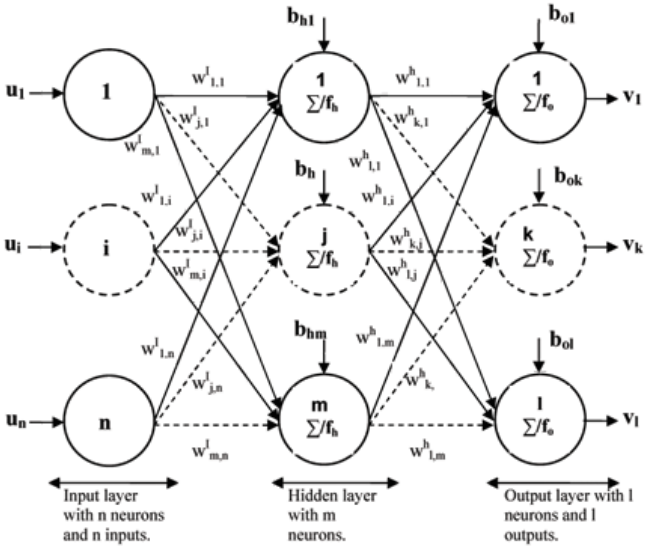

Figure 3. 3 Feedforward Network (Si-Moussa et al., 2008)

It comprises of 3 basic layers, the input, hidden and output layer. The layers are linked during presentation of input signals with a feedforward connection. While an error corrections algorithm is applied hidden units in order to adjust the connection weights (Kapageridis, 1999).

b. Radial Basis Function Networks (RBF)

RBF is a 3 layered neural networks based on interpolation in a multidimensional space with each layer performing a different function. The input layer contains the sources nodes which receives inputs from the environment (Haykin, 1999), the hidden layer neurons are activated by a RBF, where hidden neurons nodes contains a centre c vector (with the same parameter as input vector x) whose distance between the centre and the network input vector (x) is given as;

$$\|x(t) - c_j(t)\| \qquad (3.3)$$

While the output of the hidden layer can be estimated from a nonlinear function (($h_j(t) =$

$$exp\left(-\frac{\|x(t)-c_j(t)\|^2}{2b_j{}^2}\right), j = 1, \ldots\ldots\ldots m \quad (3.4)$$

Where m = number of hide nodes, b_j = positive scalar (Liu, 2013).

c. Recurrent Networks

Recurrent networks is different from feedforward network in that there is at least one feedback loop. The feedback can be a local applied to a single neuron or global to the entire network (Haykin, 1999).

The network keeps records of its previous parameters, hence the subsequent state of the network depends on the combined effect of the previous networks state and the connections weights. It is also important to note that this type of topology can be risky if it is not properly applied.

d. Self-Organisation Networks

Self-organisation networks is an unsupervised learning topology based on vector quantization.

The networks tend to learn from an n-dimensional input space into a two-dimensional space with a topological map (Elhag, 2002).

3.7. Transfer Function

The processing elements (*neurons*) of a network with its adjustable internal parameter (*weights*) Duch and Jankowski (1999), the transformation of the neurons from input space to output space is determined by a non-linear function (*transfer function*).

Their primary function is to define exactly how the scale or magnitude of a neurons response to applied signal.

Furthermore, Duch and Jankowski (1999) defined Transfer function as a composition of "*the activation function* and *the output function"* together determines the values of a neuron outgoing signals in the N-dimensional input space (parameter space). The activation function is a linear combiner whose function is to determine the total signal a neuron receives. If a neuron *i* connects to neuron *j* (if *j=1….,N)* with signal sent (x_j) and connection strength of $w_{ij,}$ total activation $I_i(x)$ is given as;

$$I_i(x) = \sum_{j=0}^{N} w_{ij}x_j \quad (3.5)$$

Where $w_{i,0}$= Θ (threshold) and x_0 = 1
Also, the output function which operates on a scalar activations and returns scalar values are responsible neurons signal processing.
Different types of transfer function used in ANNs are discussed hereunder.

Sigmoid Functions: Sigmoid functions isan S-shape curves exhibiting a graceful balance between linear and non-linear behaviours (Menon et al., 1996), they are natural and a good squashing functions for unbounded activation (Duch & Jankowski, 1999).

Hard-limited Threshold function: The neuron is said to be activated and the activation value set to 1 if the net input value to the neuron exceed a particular threshold, otherwise it remains inactive at the activation value of 0 (Kasabov, 1996).

The linear threshold function: Unlike hard-limited threshold function, the activation value linear increase is directly proportional to increase in the net input signal until a certain threshold with saturated output (Kasabov, 1996).

3.8. Other Properties of the Neural Networks

Input Neurons: Neurons are computing elements of the network. Input neurons are external neurons located at the source node of the network which receives input signals from the environment.

Output Neurons: They are neurons at the sink nodes which represents the network target. Neurons at the output nodes are also external neurons because they receive processed signals from the hidden layer in a multilayer learning networks.

Hidden Layers: The hidden layer(s)sometimes called the processing layer are introduced to increase the computation power of the network. In a simple input/output layered network, they are limited and cannot compute an *Exclusive OR* (XOR) function. The hidden neurons or rather the hidden units are processing units contained in the hidden nodes which is responsible for processing input signals.

Weights: They are values associated with each vector and nodes in the network and determines how the input relates to the output. (measures the relevance of input data to output data). Each neuron in the network is able to receive input signals, to process them and to send an output signal. Because individual neurons are connected to at least one neuron, and their connection is evaluated by a real number (weight coefficient), that reveals the degree of importance of a given connection in the network (Svozil et al., 1997). Adjustment of weights in the work represents the learning.

Bias: Weights assigned to individual nodes of the network. Their values are initial set to 1 and updated as the networks is trained.
The weights represent the memory of the network.

3.9. Properties of the Model

Coefficient of Correlation: commonly known as R, it measures the linear correlation between two variables. It is given as;

$$R = \pm\sqrt{\frac{\sum(y_{est}-\hat{y}_i)^2}{\sum(y-\hat{y}_i)^2}} \qquad (3.6)$$

Where $\sum(y_{est} - \hat{y}_i)^2$ = explained variation (Spiegel & Stephens, 1999)

note that a higher R does not simply means that the model is fitted, different parameters can lead to overfitting which include;

1. slope of regression
2. size of dataset and distribution
3. Ratio of independent parameters to number of dataset.

Performance Error: The network performance error will be calculate using Mean Square Error (MSE).

3.10. Review of Related Literature

Works of Kwon and Song (2010), Liu et al. (2008), Lu and Liu (2009) and Taffese et al. (2015) are reviewed hereunder. Due to lack of information provided by BU et al. (2009) the literature will not be reviewed herewith.

Input nodes: in Kwon and Song (2010), 4 units was employed comprising “water/cement ratio, total volume ratio of sand and coarse aggregate and relative humidity”. Liu et al. (2008) modelling was based on “water/gel ratio, cement content and time of exposure”. Stress level of concrete, testing age, water/cement ratio, cement/fine aggregate ratio and cement/coarse aggregate ratio are five input used in Lu and Liu (2009). Taffese et al. (2015) considered 25 variables at input node consisting of carbonation duration, cement type, admixtures, additives, mix properties in including test for workability and environmental condition.

Hidden layer and neurons: all four studies used 1 hidden layer, only Lu and Liu (2009) reported the size of hidden neurons used which was 10 units.

Output node: Within the four studies, one node was established as target unit. All except Kwon and Song (2010) have carbonation depth as output node, while the former established CO_2 diffusion coefficient as target node.

Dataset: 19 datset was used by Lu and Liu (2009), 12 dataset by Kwon and Song (2010), Liu et al. (2008) used 70 samples while Taffese et al. (2015) did not provide number sample used.

Within this studies, high R values (correlation coefficient) was obtained, however input nodes used did not reflect the exact factors influencing the progress of carbonation in concrete with respect to cement composition. Also modelling was based on one optimization function with limited dataset.

CHAPTER 4
METHODOLOGY USED FOR THE PREDICTION OF CARBONATION DEPTH USING ANN

4.1. Introduction

Behaviour of factors influencing carbonation is nonlinear and neural networks is an appropriate approach. Application of ANN for any given problem requires a sufficient dataset for training of the network. Dataset acquired was treated to conform to method used. Since logistic sigmoid function was used as transfer function between the layers, dataset was converted into binary form to avoid earlier saturation in the network. Details of sample input and output sample sizes was also outlined in this section.

4.2. Data Selection

Extracted dataset was based on factors influencing the progress of carbonation. These factors include;

a. Concrete microstructure and Porosity
   1. Binder Properties (input)
      i. Cement and cement additives composition
      ii. Water/cement ratio
   2. Mix properties (input)
      i. Cement quantity
      ii. Cement additives quantity
      iii. Water quantiy
      iv. Water/binder ratio

b. Curing properties (input)
   1. Curing Relative humidity
   2. Duration.

c. Environment influencing parameters (input)
   i.Relative humidity
   ii.Carbon dioxide concentration
   iii.Temperature

d. Age of concrete: duration of carbonation. (input)

e. Carbonation depth in mm (Output)

18 influencing factors was isolated been refer to as input data. Within each test number a corresponding carbonation depth was recorded.

Table 4.1: List of input parameters

No	Description	unit
1	Cement CaO content	%
2	Cement SiO_2 content	%
3	Cement Fe_2O_3 content	%
4	Cement Al_2O_3 content	%
5	Fly Ash CaO content	%
6	Fly Ash SiO_2 content	%
7	Fly Ash Fe_2O_3 content	%
8	Fly Ash Al_2O_3 content	%
9	Total cement content	kg/m^3
10	Total fly ash content	kg/m^3
11	Water content	kg/m^3
12	Water/binder ratio	%
13	Curing time	day
14	Curing relative humidity	%
15	external temperature	^{o} C
16	external CO_2 content	%
17	external relative humidity	%
18	Carbonation duration	day

Table 4.2: List of scientific papers where dataset was extracted

s/n	Name of paper	Author/s	Dataset number
1	"The experimental investigation of concrete carbonation depth Cement and Concrete Research 36 (2006) 1760– 1767"	Cheng-Feng Chang; Jing-Wen Chen	**2**
2	"Analysis of an accelerated carbonation test with severe preconditioning; Cement and Concrete Research, Volume 57, March 2014, Pages 70–78"	Ph. Turcrya, L. Oksri-Nelfiaa, A. Younsib, A. Aït-Mokhtara	**10**
3	"Measurement methods of carbonation profiles in concrete: Thermogravimetry, chemical analysis and gammadensimetry; Cement and Concrete Research 37 (2007) 1182–1192".	Géraldine Villain, MickaëlThiery, Gérard Platret	**3**
4	"Different methods to measure the carbonation profiles in concrete; International RILEM Workshop on Performance Based Evaluation and Indicators for Concrete Durability: Madrid, Spain, 19-21 March 2006"	G. Villain, M. Thiery, V. Baroghel-Bouny, G. Platret	**10**
5	"Accelerated carbonation and testing of concrete made with fly ash; Construction and Building Materials, Volume 17, Issue 3, April 2003, Pages 147–152"	Cengiz Duran Atiş	**40**
6	"A model for predicting carbonation of high-volume fly ash concrete; Cement and Concrete Research, Volume 30, Issue 5, May 2000, Pages 699–702".	LinhuaJianga, BaoyuLinb, YueboCai	**36**
7	"A performance based approach for durability of concrete exposed to carbonation; Construction and Building Materials, Volume 23, Issue 1, January 2009, Pages 190–199"	Emmanuel Rozièrea, Ahmed Loukilia, François Cussighb	**32**
8	"Reactive transport modelling of long-term carbonation; Cement and Concrete Composites, Volume 52, September 2014, Pages 42–53".	O.P. Kari, ,J. Puttonen, E. Skantz	**44**
9	"Effects of curing upon carbonation of concrete; Construction and Building Materials, Vol 9, No. 2, pp. 91-95, 1995".	J.P. Balayssac, Ch.H. Détriché, J. Grandet	**48**

4.3. Data Pre-processing: Normalization

Variables presented in this study are in binary form (0 to 1), since transfers functions used in neural networks gives outputs within a specified range, and faster convergence of training set is needed to reduce number of training iteration. Saturation of neurons leading to slow convergence in a logistic Sigmoid mapped network (Franzini, 1987; Dahl, 1987; Fahlman, 1989; Chen & Mars, 1990; Lee et al., 1991; Balakrishnan & Honavar, 1992; Spartz & Honavar, 1993; Parekh et al., 1993;Vitela & Reifman, 1993;Vitela & Reifman, 1997) are attributed normalization of training set (Sola & Sevilla, 1997). While considering the importance of normalization in input dataset, Khashman (2010), argued that normalizing the whole set with by dividing the it with highest number will cause the some figure to become insignificant since their values after normalization is scaled towards zero. Sola and Sevilla, (1997), observed that network error decreases as normalization minimizes the changes between the variation range of different variables.

Therefore, normalization will be perform within each variables.

4.4. Feedforward Multilayer Perceptron Networks

The works of Rumelhart et al. (1986), Werbos (1990) and many more gave a leeway to research in MLPs. The presence of hidden layers differentiates multilayer networks from single layer networks. Unlike the single-layer networks where computational nodes are located at the output layer, hidden layer with its corresponding nodes (hidden units or neurons) constitutes the computation nodes of the network.

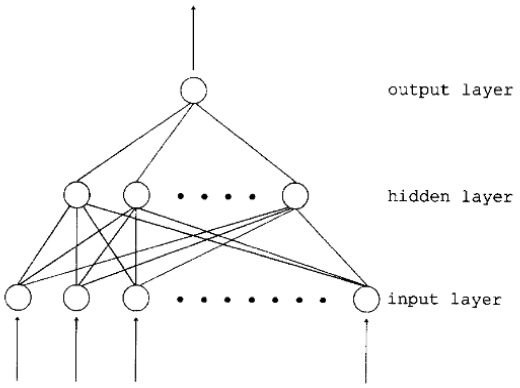

Figure 4.1: Feedforward Multilayer Network (Svozil et al., 1997)

Also, the problem associated with linear separability of the perceptrons with cases of solving OXR problems distinguishes MLPs from single layer netwoks. MLPs neurons contains "continuous value inputs and outputs, summation input function, and nonlinear activation function". In order to minimize global error E, a gradient descent rule is applied to optimize connection weights W_{ij}. Change of weight ΔW_{ij} in the direction of negative gradient of the error at a cycle (t + 1) is given as (Kasabov, 1996).;

$$\Delta W_{ij}(t+1) = -lrate\left(\partial E/\partial W_{ij}(t)\right) \quad (4.1)$$

Irate= learning rate.

If *E* is as a surface weight in the weights' vector space, a global error for all training may be given as (Kasabov, 1996). ;

$$E = \sum_{(p)} \sum_{(f)} {E_{rr}}^{(p)} \quad (4.2)$$

Where $E_{rr} = {\left({y_j}^{(p)} - {0_j}^{(p)}\right)^3}/_2$ *as a MSE*

For every completed cycle representing the network epoch or iteration training propagated through the network, error is calculated.

4.5. Feedforward Backpropagation Algorithm

Backpropagation algorithm also known as "*generalized delta rule"* that apply gradient error function method to minimize the total squared errors of the output computed by the network in three phases;

1. Feedforward of the input training signals
2. Calculation and propagation of associated error
3. Adjustment of connection weights.

Forward pass propagates input signals received at by the input units X_i via the hidden units Z_j unto the output units *Yk.* At the input and hidden layers, respective activation x_i and z_i is computed and sends signals x_i and z_i to connected units in the next layer.

Output unit initiates the *backward pass* by computing its activation y_k and compare it with target value t_k to compute associated error. Thereafter δ_k factor (k= 1, ……, m)is computed and used to distribute error at the output unit Y_k back to all units connected to the output layer from the hidden layer. Also in a similar case but simultaneously, δ_j factor (j= 1, ……,p)

of hidden units Z_j are computed and used to update the weight connections of inputs and hidden layers (Fausett, 1994; Kasabov, 1996).

Fausett, (1994) presented a summarized algorithm for backpropagation neural networks as follows;

Step 0.

"Initialize weights. (set to small random values)".

Step 1.

"While stopping is false, do steps 2-9"

Step 2. "for each training pair, do steps 3-8"

Feedforward:

Step 3. "Each input unit (X_i, i = 1, ……., n) gets input signal x_i and transmits it to all neurons in the hidden layer".

Step 4. 1. "Each hidden neurons (Z_j, j = 1, ……., p) calculations its weighted input signals",

$$z_{-in_j} = v_{oj} + \sum_{i=1}^{n} x_i v_{ij},$$

2. "activation function is applied to calculate output signal"

$z_j = f(z\_in_j),$

3. "activated signal is sent to all neurons in the output layer".

Step 5. "Each output neurons (Y_k, k= 1, ……., m) calculates it weighted applied signals",

$$y_{-in_k} = w_{ok} + \sum_{j=1}^{p} z_j w_{jk},$$

2. "activation function is applied to calculate output signal"

$y_k = f(y\_in_k),$

Backpropagation of error

Step 6. 1. "each output neurons (Y_k, k = 1,….m) obtains a"

"Corresponding input training target pattern, and evaluates the error".

$\delta_k = (t_k - y_k) f'(y\_in_k),$

2. "calculates weight correction term for updating w_{jk}"

$\Delta w_{jk} = \alpha \delta_{kzj},$

3. computes bias correction term for updating w_{ok}

$\Delta w_{ok} = \alpha\delta_k$,

4. applies δ_k to neurons in the layer below.

Step 7. 1. individual hidden neurons (Z_j, j=1,...,p) calculates its delta inputs from the neurons in the output layer.

$$\delta_{-in_j} = \sum_{k=1}^{m} \delta_k w_{jk},$$

2. computes error term by multiplying the derivative of its activation function.

$\delta\_in_j = \delta\_in_j f'(z\_in_j)$,

3. computes weight correction term for updating v_{ij}

$\Delta v_{ij} = \alpha\delta_{jx_i}$,

4. computes bias correction term for updating v_{oj}

$\Delta v_{oj} = \alpha\delta_j$.

Weight and Bias Update

Step 8. 1. individual output neurons (Y_k, k =1,……., m) updates its bias and weight (j = 0, …….p):

$w_{jk(new)} = w_{jk(old)} + \Delta w_{jk}$

2. Individual hidden neurons (Z_j, j=1,…….., p) updates its bias and weights (i= 0, ……., n):

$v_{ij(new)} = v_{ij(old)} + \Delta v_{ij}$

Step 9. Test stopping condition met.

4.5.1. Activation function

Binary sigmoid function (0, 1) is preferred. Since normalization of independent variables are between zero and one, it requires an activation function with continuous, differentiable and monotonically non-decreasing for efficient computation.

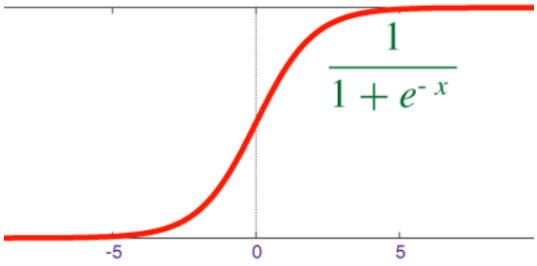

Figure 4.2: Sigmoid function

$f_1(x) = \frac{1}{1+e^{(-x)}}$, with (4.3)

$f_1'(x) = f_1(x)[1 - f_1(x)]$ (4.4)

4.5.2 Training functions (optimization methods)

Conjugate-Gradient method: is suitable for training MLPs owning to numerous adjustable parameters. It differs from method of steepest descent of slow convergence and can be applied in large-scale problems. This I a second-order optimization problem (Haykin, 1999). It was proposed in 1964 by fletcher and Reeves, the method applies successive conjugate direction based on the gradient and the residue. If the objective function is a quadratic function, the search direction is minimized exactly at each epoch. Network weights are updated in the descent direction ***S(n)*** according to unidirectional search (Robitaille et al.,1996).

$\Delta w(n) = \lambda(n)S(n)$ (4.5)

Where $S(n) = G(n) + \frac{\|G(n)\|}{\|G(n+1)\|}S(n+1)$

Scaling factor used include;

Polak – Ribière = $\frac{r^T(n)(r(n)-r(n-1))}{r^T(n-1)r(n-1)}$ (4.6)

Fletcher-Reeves = $\frac{r^T(n)r(n)}{r^T(n-1)r(n-1)}$ (4.7)

Quasi-Newton Method: Quasi-Newton method directly approximate the inverse Hessian matrix $[H(n)]^{-1}$ from the derivative, G(n).

Here the descent direction $S(n)= -[H(n)]^{-1}G(n)$.

4.6. Distribution of Dataset

Percentage of dataset distributed in training-testing-validation minimises the problem of over-fitting cause by poor generalization as a result of network remembering by heart during validation. Here the margin is clear on network accuracy between testing and cross-validation sets.

Yadollahi et al., (2016) adopted 74% -26% in "predicting the optimal mixture of radiation shielding concrete" with 3 inputs and 3 outputs. Taffese and Sistonen, (2016) in their study for prediction for deterioration risk analysis of concrete façade element used 75% training and 25% testing with Outdoor RH and temperature of ambient air as input variables. As input neurons increases the complexity of model also increases making it more likely to have overfitting (Haykin, 1999).

To insure that the model is not generalizing by heart, and adopting Khashman, (2010) suggestion of a close range of 40% - 60%, 50% - 50% and 60% - 40% training-to-testing data distribution. Above distribution ensure that with whatever experience the network gains, it can predict and the accuracy of the result can be justified.

The training set is further divided into estimation (train and select model) and validation (test or validate model) subset (Haykin, 1999). But in this study, data division above testing and validation are grouped together in order to appreciate the size of training subset. Also the data distribution is refers in this study as learning scheme (LS), with LS1, LS2 and LS3 representing 40% - 60%, 50% - 50% and 60% - 40% training/Testing respectively.

4.7. Number of Hidden neurons

There is a perception that increase in hidden neurons is directly proportional to decrease in network error. This is generally the case for most new users of artificial neural networks. The behaviour of ANN is yet to be fully understood, early researchers called the method "the blackbox" referring to unexplained phenomenon associated with it.

The difficulty in selecting the number of hidden neurons and proposed solutions (Zhang et al., 2003; Yuan et al., 2003) are yet to be accepted among researchers (Atici, 2011).

Earlier proposals were not scientific as some suggest number of hidden neurons can be estimated by adding the number of input neurons and output neurons and divide by two. Other suggestions are maximum number of hidden neurons may not exceed number of samples. In Yuan et al. (2003) view these methods, in the first case, factors that affect

network structure cannot be ruled out such number of samples in the training set, noise size of sample and complexity of learning function. While in the second case viewed, it can only solve cases of overlearning in the network.

This is study is not to establish an algorithm or function for selection of number of hidden neurons, it is beyond the scope of this study. Traditional method of training with a series of hidden neurons will be applied and the number with optimum performance will be selected. Hidden neuron sizes used are 5, 10, 15, 20, and 25 representing ANN-1, ANN-2, ANN-3, ANN-4, and ANN-5 respectively.

4.8. Optimisation Methods

a. Steepest Decent Methods

i. Gradient Descent Backpropagation (GD)

Gradient descent is a batch steepest descent algorithm with weights and bias of the network updated along negative gradient. Since the learning rate multiplies the negative of the gradient, in order to estimate weights and bias changes, larger rate for learning force the network to converge quickly while for smaller learning rate, the network is much longer.

ii. Gradient descent with Adaptive Learning Rate Backpropagation (GDA).

GDA is a modified algorithm from GD reported in the previous subsection. Adaptive learning rate here implies that the learning rate of the network is not constant throughout the training. There is modification according to error of the network at each epoch, weight and bias are either retained or discarded with respect to the value of the new and the old error. Increase and decrease of learning rate are performed by multiplying the learning rate with *lr_inc* and *lr_dec* to obtain a new learning rate.

iii. Gradient Descent with Momentum Backpropagation (GDM).

Addition of a momentum filter to learning parameters maintains stability of the algorithm even when larger learning rate is used. Also convergence is accelerates when the trajectory is moving in a consistent direction (Hagan et al., 1996). With addition momentum, the result is expected improve within the learning scheme when compared with other steepest descent.

iv. *Gradient Descent with Momentum and Adaptive Learning Rate Backpropagation (GDX*)

b. Conjugate Gradient Method.

i. Conjugate Gradient Backpropagation with Powell-Beale Restarts (CGB)

Conjugate gradient algorithms with restarts at point when the orthogonality between the current gradient and the previous gradient is small (mathworks.com, retrieved August 31, 2016).

ii. Conjugate gradient backpropagation with Fletcher-Reeves updates (CGF)

This is a steepest descent direction search based algorithm. Search lines are perform to determine the optimal distance required for any given search direction. It is fast converging algorithm suitable for estimating carbonation problem.

iii. Conjugate Gradient Backpropagation with Polak-Ribiére updates (CGP).

This method is similar to Fletcher-Reeves updates method of conjugate gradient. It is difficult to analyses network performance based on number of iteration as there is no pre-determined search direction stipulated for the network. As discussed earlier in GDM and GDA, "gradient descent with momentum and adaptive learning rate backpropagation" combines both methods for optimization of network models

iv. Scaled Conjugate Gradient (SCG)

Unlike conjugate gradient backpropagation with Fletcher-Reeves updates, conjugate gradient backpropagation with Polak-Ribiére updates and gradient backpropagation with Powell-Beale restarts, scaled conjugate gradient method despite been a conjugate direction algorithm do not perform a line search at each iteration. It uses a Levenberg-Marquardt method in order to scale the step size (Moller, 1993). Absence of line search in this method can be seen in the maximum epoch obtained in each set of training. Increase in the number of hidden neurons also decreases the computational cost.

c. Levenberg-Marquardt Backpropagation (LM)

i. Levenberg-Marquardt Backpropagation (LM)

Levenberg-Marquardt algorithm is a non-linear least iterative technique that minimizes a least square function to sum of its square. A search direction is performed and can alternate

between Gauss-Newton direction and the steepest descent direction (www.mathworks.com retrieved on September 15, 2016).

d. Bayesian Regularization Backpropagation (BR)

Proposed by McKay, (1992). Variables are adjusted according to Levenberg-Marquardt algorithm which uses the backpropagation to calculate the Jacobian jX of performance in connection to the weight and bias of the variables X (mathworks.com, retrieved August 29, 2016).

e. BFGS Methods

i. BFGS Quasi-Newton Backpropagation (BFG)

Named after the contributions of Broyden, Fletcher, Goldfarb, and Shanno. It convergences faster than conjugate gradient method and does not require to second derivatives. At each epoch, approximate Hessian matrix is updated (mathworks.com, retrieved August 29, 2016).

ii. One-step Secant Backpropagation OSS

one step secant is a method that is neither fully BFGS algorithm nor fully conjugate gradient algorithm, the algorithm requires less storage computation per epoch than BFGS and more than the conjugate gradient algorithm

f. Random order incremental training with learning functions (R)

Weights and bias are updated incrementally after input are presented randomly (mathworks.com, 2016).

g. Resilient backpropagation (RP)

Proposed in Riedmiller and Braun (1993), to overcome problem of weight updates in gradient-descent. Weight-steps are according to sequence of signs, not on the magnitude of the derivatives causing the learning to be distributed evenly over the entire network.

CHAPTER 5
RESULTS AND DISCUSSION

5.1. Results and Discussion Overview

In the cause of this study, several difficulties were met, overcome and some limited the progress of the study.

- Experimental carbonation dataset was limited owning to limited information provided by some authors on the properties of cement and concrete mix.
- Acquired dataset
- Times was a major factor in this study. Experimental evaluation was replaced and option of extracting data from published literature was adopted.
- Publication on the subject with respect to neural networks is also limited as the time of this study, hence major citations was done from similar problem on ANN prediction.
- Difficultly in estimating the number of hidden neurons was a major problem.

As explained in the previous chapters, distribution of number of hidden neurons forms the ANN models. Five different models were proposed based on trial analysis; ANN-1, ANN-2, ANN-3, ANN-4 and ANN-5 which refer to 5, 10, 15, 20 and 25 hidden neurons, respectively contained in the hidden layer. While LS1, LS2 and LS3 representing 40:60, 50:50 and 60:40 % of data distribution, respectively.

Table 5.1 below presents have results according to learning scheme, Table 5.1 to 5.3, are results of LS1, LS2 and LS3, respectively. Correlation coefficient (R), mean square error (MSE) and iteration are presented as the network results. R values presented in the cross-validation table 5.4 - 5.16 are training R, testing R and total R of the network.

Network performances are discussed according to method of optimization with respect to changes in learning scheme and number of hidden neurons.

Table 5.1: Results of learning scheme-1 with changing ANN model

Learning Scheme	Optimization Method	ANN-1 (5H)			ANN-2 (10H)			ANN-3 (15)			ANN-3 (20)			ANN-3 (25)		
		R	**MSE**	**iter**	**R**	**MSE**	**iter**	**R**	**MSE**	**iter**	**R**	**MSE**	**iter**	**R**	**MSE**	**iter**
40:60	GD	0.944	0.0038	20000	0.936	0.0042	20000	0.931	0.0047	20000	0.933	0.0047	20000	0.908	0.0069	20000
	GDA	0.943	0.0039	20000	0.943	0.0040	20000	0.945	0.0038	20000	0.925	0.0059	3848	0.953	0.0032	20000
	GDM	0.933	0.0045	20000	0.919	0.0054	20000	0.936	0.0045	20000	0.928	0.0048	20000	0.927	0.0052	20000
	GDX	0.936	0.0045	8434	0.940	0.0041	20000	0.932	0.0047	20000	0.936	0.0045	20000	0.939	0.0046	20000
	CGB	0.966	0.0025	1120	0.953	0.0032	419	0.965	0.0024	390	0.950	0.0035	538	0.942	0.0043	417
	CGF	0.939	0.0041	1077	0.950	0.0035	392	0.942	0.0040	607	0.963	0.0044	103	0.951	0.0034	447
	CGP	0.963	0.0025	1828	**0.966**	**0.0023**	**464**	0.946	0.0037	487	0.956	0.0031	597	0.951	0.0037	456
	SCG	0.947	0.0036	688	0.960	0.0027	489	0.941	0.0043	523	0.935	0.0045	311	0.930	0.0050	277
	LM	0.959	0.0029	28	0.944	0.0046	15	0.942	0.0041	50	0.949	0.0036	16	0.940	0.0041	8
	BR	0.943	0.0039	429	0.937	0.0043	1091	0.941	0.0040	26	0.946	0.0037	1206	0.938	0.0043	1248
	BFG	0.954	0.0032	367	0.954	0.0032	167	0.948	0.0036	52	0.931	0.0048	198	0.933	0.0049	266
	OSS	0.956	0.0033	3189	0.943	0.0039	1766	0.942	0.0040	2222	0.944	0.0038	1506	0.929	0.0048	376
	R	0.917	0.0063	20000	0.944	0.0044	20000	0.946	0.0041	20000	0.942	0.0040	20000	0.904	0.0070	349
	RP	0.951	0.0034	20000	0.949	0.0035	11638	0.939	0.0045	655	0.947	0.0036	9489	0.937	0.0042	8660

***** different colours above delineates grouping of optimization function.

Table 5.2: Results of learning scheme-2 with changing ANN model

Learning Scheme	Optimization Method	ANN-1 (5H)			ANN-2 (10H)			ANN-3 (15)			ANN-3 (20)			ANN-3 (25)		
		R	MSE	iter	R	MSE	iter	R	MSE	iter	R	MSE	iter	R	MSE	iter
50:50	GD	0.915	0.0057	20000	0.937	0.0043	20000	0.952	0.0033	20000	0.936	0.0043	20000	0.953	0.0033	20000
	GDA	0.950	0.0036	20000	0.948	0.0037	20000	0.945	0.0039	20000	0.948	0.0037	20000	0.941	0.0041	20000
	GDM	0.951	0.0034	20000	0.935	0.0046	20000	0.945	0.0038	20000	0.938	0.0042	20000	0.939	0.0041	20000
	GDX	0.945	0.0037	20000	0.951	0.0034	20000	0.946	0.0037	20000	0.952	0.0034	20000	0.956	0.0030	20000
	CGB	0.967	0.0023	683	0.970	0.0021	643	0.964	0.0025	841	0.968	0.0023	559	0.963	0.0025	524
	CGF	0.939	0.0043	1141	0.966	0.0024	661	0.960	0.0028	697	0.962	0.0026	807	0.958	0.0029	509
	CGP	0.970	0.0020	994	**0.975**	**0.0018**	**547**	0.956	0.0031	395	0.946	0.0037	562	0.954	0.0032	596
	SCG	0.967	0.0023	550	0.969	0.0022	675	0.963	0.0029	600	0.965	0.0024	719	0.972	0.0020	559
	LM	0.967	0.0023	26	0.963	0.0026	42	0.958	0.0030	53	0.965	0.0024	53	0.958	0.0030	29
	BR	0.952	0.0033	173	0.948	0.0036	909	0.953	0.0033	22	0.946	0.0037	2486	0.946	0.0037	575
	BFG	0.966	0.0026	339	0.951	0.0034	354	0.966	0.0024	258	0.948	0.0037	272	0.951	0.0034	98
	OSS	0.961	0.0027	2007	0.951	0.0035	2432	0.965	0.0025	2115	0.957	0.0031	2317	0.947	0.0039	2638
	R	0.945	0.0038	20000	0.862	0.0091	20000	0.910	0.0063	20000	0.934	0.0048	11266	0.916	0.0057	1322
	RP	0.964	0.0025	18630	0.956	0.0030	20000	0.944	0.0038	20000	0.958	0.0029	6783	0.939	0.0042	10184

***** different colours above delineates grouping of optimization function.

Table 5.3: Results of learning scheme-3 with changing ANN model

Learning Scheme	Optimization Method	ANN-1 (5H)			ANN-2 (10H)			ANN-3 (15)			ANN-3 (20)			ANN-3 (25)		
		R	**MSE**	**iter**	**R**	**MSE**	**iter**	**R**	**MSE**	**iter**	**R**	**MSE**	**iter**	**R**	**MSE**	**iter**
60:40	GD	0.947	0.0037	20000	0.953	0.0032	20000	0.955	0.0031	20000	0.948	0.0036	20000	0.954	0.0031	20000
	GDA	0.953	0.0032	20000	0.954	0.0031	20000	0.952	0.0034	20000	0.952	0.0033	20000	0.960	0.0028	20000
	GDM	0.945	0.0038	20000	0.941	0.0040	20000	0.948	0.0036	20000	0.943	0.0039	20000	0.942	0.0040	20000
	GDX	0.945	0.0038	20000	0.964	0.0025	20000	0.961	0.0027	20000	0.966	0.0023	20000	0.962	0.0026	20000
	CGB	0.971	0.0020	618	0.974	0.0018	572	0.973	0.0019	423	0.967	0.0022	537	0.968	0.0022	597
	CGF	0.976	0.0017	1385	**0.980**	**0.0014**	**1561**	0.976	0.0017	738	0.962	0.0027	554	0.962	0.0027	540
	CGP	0.976	0.0017	1009	0.977	0.0016	537	0.970	0.0021	555	0.970	0.0022	888	0.972	0.0019	631
	SCG	0.970	0.0021	890	**0.980**	**0.0014**	**676**	0.961	0.0027	514	0.969	0.0022	422	0.970	0.0022	611
	LM	0.960	0.0028	49	0.959	0.0028	29	0.973	0.0019	41	0.972	0.0020	50	0.972	0.0020	23
	BR	0.952	0.0033	240	0.954	0.0032	1106	0.950	0.0034	1746	0.954	0.0032	960	0.947	0.0036	1691
	BFG	0.971	0.0020	714	0.960	0.0027	318	0.957	0.0029	294	0.963	0.0025	306	0.968	0.0022	284
	OSS	0.970	0.0021	3789	0.973	0.0020	2778	0.969	0.0021	2932	0.967	0.0023	2125	0.956	0.0031	1599
	R	0.951	0.0034	20000	0.968	0.0023	20000	0.953	0.0032	18972	0.950	0.0034	20000	0.959	0.0039	20000
	RP	0.967	0.0023	20000	0.967	0.0023	12749	0.957	0.0029	16806	0.958	0.0029	12524	0.967	0.0023	14950

***** different colours above delineates grouping of optimization function.

Steepest gradient descent functions.

Conjugate gradient functions.

Resilient backpropagation function.

Bayesian Regularization Backpropagation function.

Levenberg-Marquardt backpropagation function.

BFGS functions

Random order incremental training with learning function.

Table 5.4: Cross validation analysis for steepest gradient descent method for varying hidden neurons at constant 40:60 train:test distribution (LS1)

LS1	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	GD	0.950	0.942	0.944	38	0.972	0.961	0.936	42	0.966	0.893	0.931	47	0.976	0.889	0.933	47	0.977	0.898	0.908	69
	GDA	0.935	0.965	0.943	39	0.979	0.923	0.943	40	0.956	0.927	0.945	38	0.985	0.885	0.925	59	0.947	0.968	**0.953**	32
	GDM	0.956	0.929	0.933	45	0.969	0.889	0.919	54	0.976	0.918	0.936	45	0.946	0.921	0.928	48	0.979	0.907	0.927	52
	GDX	0.974	0.891	0.936	45	0.975	0.963	0.94	41	0.977	0.922	0.932	47	0.966	0.921	0.936	45	0.980	0.916	0.939	46

Table 5.5: Cross validation analysis for steepest gradient descent method for varying hidden neurons at constant 50:50 train:test distribution (LS2)

LS2	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
50:50	GD	0.931	0.913	0.915	57	0.950	0.813	0.937	43	0.979	0.898	0.952	33	0.953	0.897	0.936	43	0.968	0.941	0.953	33
	GDA	0.962	0.940	0.95	36	0.968	0.915	0.948	37	0.952	0.820	0.945	39	0.982	0.883	0.948	37	0.960	0.915	0.941	41
	GDM	0.953	0.879	0.951	34	0.957	0.886	0.935	46	0.976	0.935	0.945	38	0.961	0.873	0.938	42	0.970	0.878	0.939	41
	GDX	0.963	0.943	0.945	37	0.976	0.878	0.951	34	0.982	0.874	0.946	37	0.967	0.920	0.952	34	0.979	0.925	0.956	30

Table 5.6: Cross validation analysis for steepest gradient descent method for varying hidden neurons at constant 60:40 train:test distribution (LS3)

LS4	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
60:40	GD	0.953	0.919	0.947	37	0.966	0.910	0.953	32	0.978	0.917	0.955	31	0.968	0.952	0.948	36	0.950	0.951	0.954	31
	GDA	0.970	0.883	0.953	32	0.950	0.960	0.954	31	0.970	0.830	0.952	34	0.972	0.912	0.952	33	0.956	0.943	0.96	28
	GDM	0.958	0.940	0.945	38	0.955	0.862	0.941	40	0.973	0.800	0.948	36	0.977	0.913	0.943	39	0.984	0.869	0.942	40
	GDX	0.969	0.855	0.945	38	0.973	0.942	0.964	25	0.973	0.959	0.961	27	0.984	0.908	0.966	23	0.971	0.931	0.962	26

5.2. Discussion of Results for Steepest Gradient Method

Generally results obtained in this method shows that all steepest gradient descent method require additional iteration for optimization of network performance. Although total R values were high within the four sub-methods, training subset did not converge to 0.001 MSE.

In LS1, total R values within the methods was decreasing as number of hidden neurons increases. When the size of training subset set was increased from 40% to 50% for LS2, total R values within the sub-methods increased slightly, also training subset R (train R), increased also.

Performance of network in LS3 having 60% of sample size in the training subset, showed a significant improvement in total R values and MSE. Although within 5 hidden neurons, changes in R values when compared with LS1 and LS2 did not show a sharp increase. However, total R values and MSE in 10, 15 and 20 hidden neurons in LS3 increased significantly and slightly in 25 hidden neurons.

Network epoch (iteration) remains 20,000 except in GDX (8434 iteration) and GDA (3848 iteration) in LS1 at 5 and 20 hidden neurons respectively. The training subset was found to have converge to 0.001 MSE in both GDX and GDA mentioned above. When the R value of training subset and testing set was compared, it was discovered that less iteration in both cases was as a result of overfitting (learning by heart).

Consequently, it should be noted that steepest descent method for predicting carbonation depth in concrete require additional computation cost other than the limit set in this study. Notwithstanding, since R values are high and validation MSE are within 0.001, this method can be effectively used in predicting carbonation depth in concrete. Within this method combination of 60:40, 20 hidden neurons and with GDX optimization method was the best obtained method based total R value and MSE. Fig. 5.1 shows the behavour of training in GDX method. Also it can also be seen on blue line, that the training subset did not converge to set threshold. Corresponding fig. 5.2 shows the cross-validation R values of the 3 subsets and the total network R.

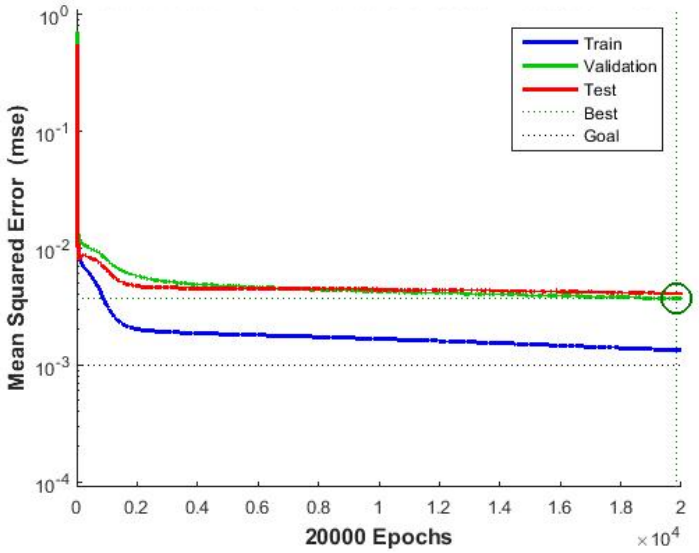

Figure 5.1: MSE graph for descent method at 60:40, GDX-20H

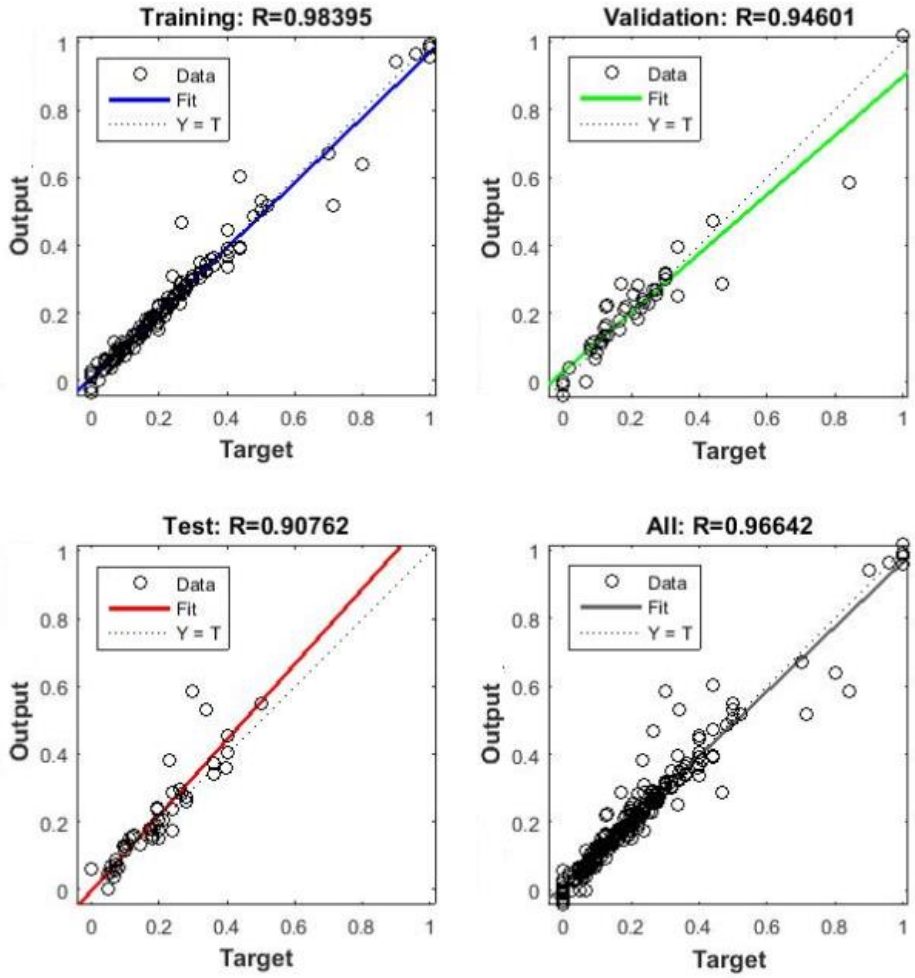

Figure 5. 2: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in Steeped Descent method at 60:40, GDX-20H

5.3. Discussion of Results for Conjugate Gradient Descent Method

Results in conjugate gradient decent method are present in Table 5.7 -5.9 Additional graphs are attached in appendix.

Reliability of this method for predicting carbonation in concrete increased as training subset is increased from 40% to 60%. In learning scheme-1, there was an increase in total R value when hidden neurons was increased from 5 to 10 units. Further increase of hidden neurons from 10 to 25 units caused a decrease in total R values, vice versa increase in MSE.

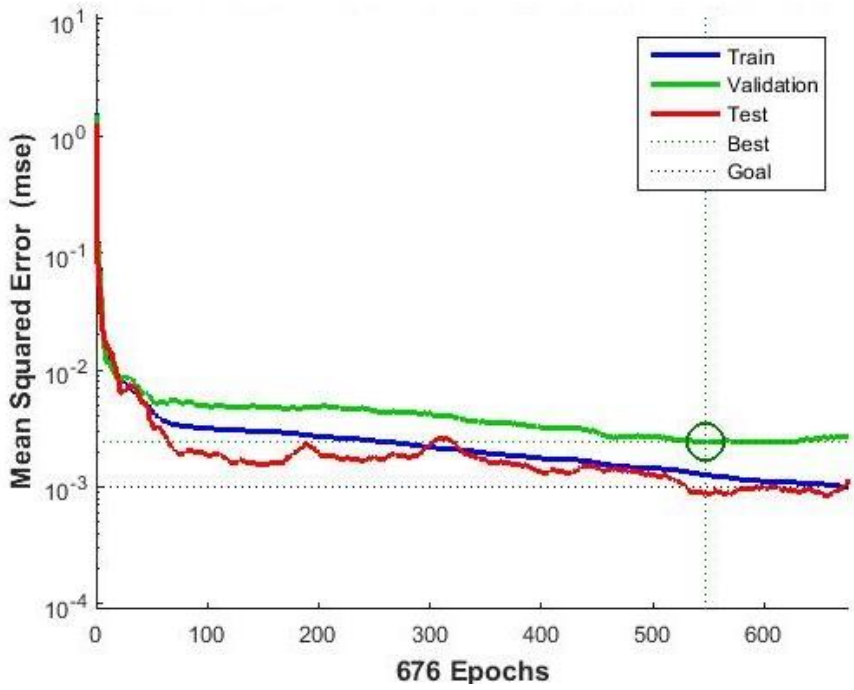

Figure 5.3: MSE graph for conjugate gradient descent method at 60:40, SCG-10H

Results obtained in LS2 shows an increase in total R value and MSE across the sub-methods within conjugate gradient descent group. Network performance improved when hidden neurons was increased from 5 units to 10 units, and later decreases with increase of hidden units from 10 to 25 units except in SCG where reliability increases with increased in hidden neurons. Similar, increase in hidden neurons in LS2 did not affect the network iteration linearly.

However, in this method linear relationship could not be established between iteration and number of hidden neuron, since the network is not likely to choose the search direction for every training.

Table 5.7: Cross validation analysis for conjugate gradient descent method (LS1)

LS1	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	CGB	0.989	0.922	0.966	25	0.983	0.980	0.953	32	0.980	0.945	0.965	24	0.952	0.939	0.950	35	0.987	0.937	0.942	43
	CGF	0.967	0.942	0.939	41	0.986	0.906	0.95	35	0.976	0.915	0.942	40	0.981	0.915	**0.963**	44	0.986	0.920	0.951	34
	CGP	0.985	0.961	0.963	25	0.984	0.973	**0.966**	23	0.979	0.841	0.946	37	0.985	0.926	0.956	31	0.982	0.941	0.951	37
	SCG	0.959	0.920	0.947	36	0.986	0.943	0.96	27	0.972	0.925	0.941	43	0.984	0.882	0.935	45	0.985	0.922	0.93	50

Table 5.8: Cross validation analysis for conjugate gradient descent method (LS2)

LS2	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
50:50	CGB	0.985	0.954	0.967	23	0.980	0.934	0.970	21	0.988	0.886	0.964	25	0.989	0.941	0.968	23	0.988	0.926	0.963	25
	CGF	0.978	0.896	0.939	43	0.984	0.949	0.966	24	0.981	0.961	0.96	28	0.987	0.906	0.962	26	0.983	0.942	0.958	29
	CGP	0.984	0.963	**0.97**	20	0.983	0.954	**0.975**	18	0.978	0.919	0.956	31	0.959	0.959	0.946	37	0.980	0.893	0.954	32
	SCG	0.984	0.941	0.967	23	0.989	0.960	0.969	22	0.986	0.888	0.963	29	0.963	0.954	0.965	24	0.985	0.971	**0.972**	20

Table 5.9: Cross validation analysis for conjugate gradient descent method (LS3)

LS3	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
60:40	CGB	0.978	0.961	0.971	20	0.986	0.947	0.974	18	0.984	0.946	0.973	19	0.977	0.922	0.967	22	0.983	0.951	0.968	22
	CGF	0.988	0.941	**0.976**	17	0.986	0.981	**0.98**	14	0.983	0.944	**0.976**	17	0.980	0.961	0.962	27	0.985	0.954	0.962	27
	CGP	0.986	0.930	**0.976**	17	0.982	0.981	0.977	16	0.988	0.933	0.97	21	0.987	0.957	0.97	22	0.987	0.947	**0.972**	19
	SCG	0.975	0.974	0.97	21	0.984	0.978	**0.98**	14	0.985	0.923	0.961	27	0.986	0.968	0.969	22	0.987	0.960	0.97	22

LS3 result shows that increase in training subset increases network performance. Also in this learning scheme 3, 10 hidden neuron model, was found to be optimized hidden unit size and it decreases as hidden units increases.

60:40, 10 hidden neurons and SCG, was seen to have performed better than the other method within this group. SCG cross-validation was also better when compared to other methods of conjugate gradient method. Fig. 5.4 shows the obtained R values with 60:40, SCG-10H, while fig 5.3 illustrates the network progressive performance. Both training (blue line) and testing (testing) subsets converged to MSE. This shows that the network experience was sufficient to predict carbonation depth in concrete effectively .

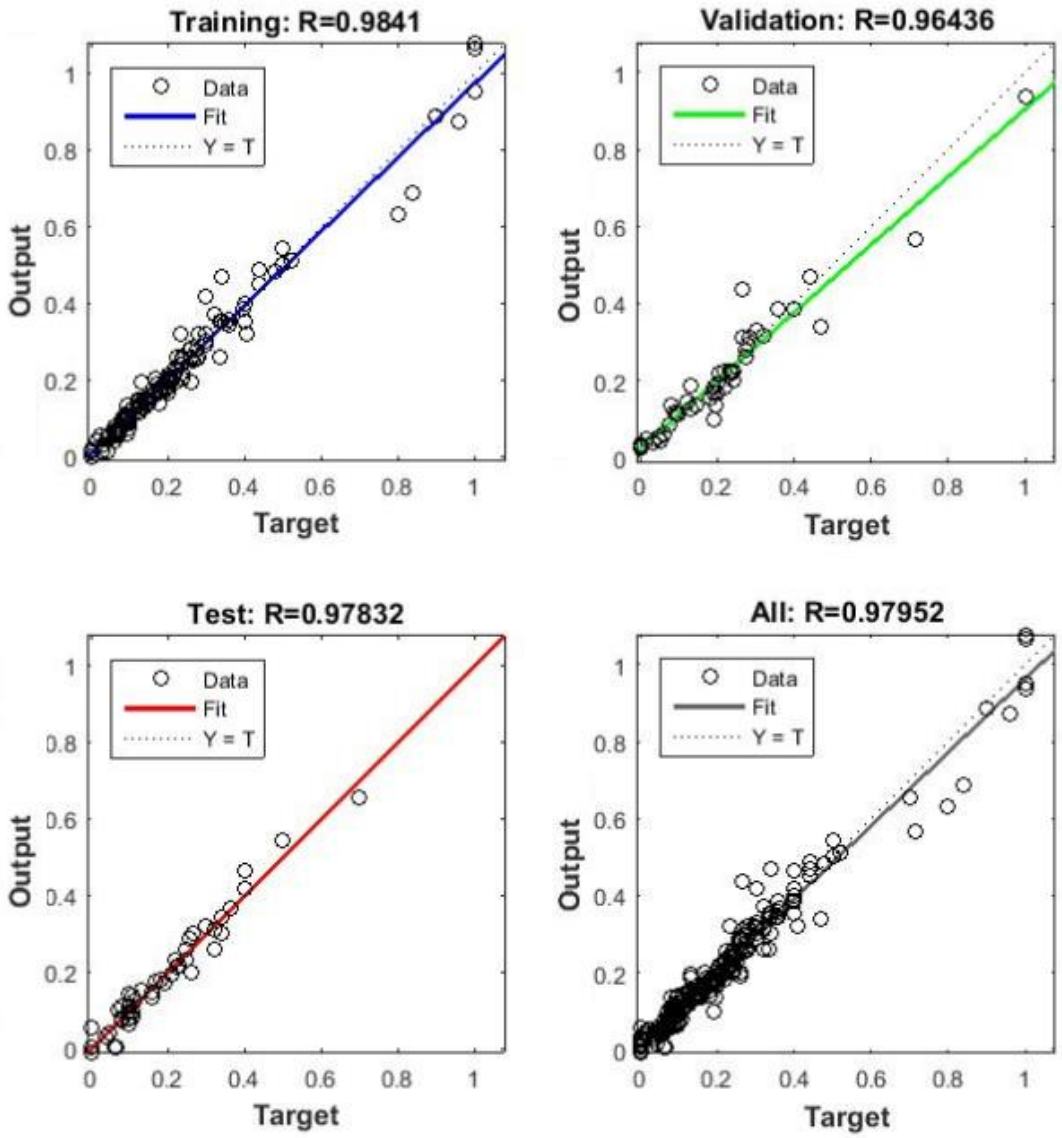

Figure 5. 4: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in conjugate gradient descent method at 60:40, SCG-10H

5.4. Discussion of Results for Levenberg-Marquardt Method

This method presents an optimized computation cost. Comparison of the result within this method as presented in table 5.1 -5.3 and 5.10, shows that network performance increases as training subset is increased. The behaviour of LM is somewhat different form the other methods, network optimization within LS3 increasing as size of hidden units are increased. Moreover, increase in training subset size increases gross total R value of the model. Fig. 5.5 and 5.6 for display optimised network performance and R values respectively.

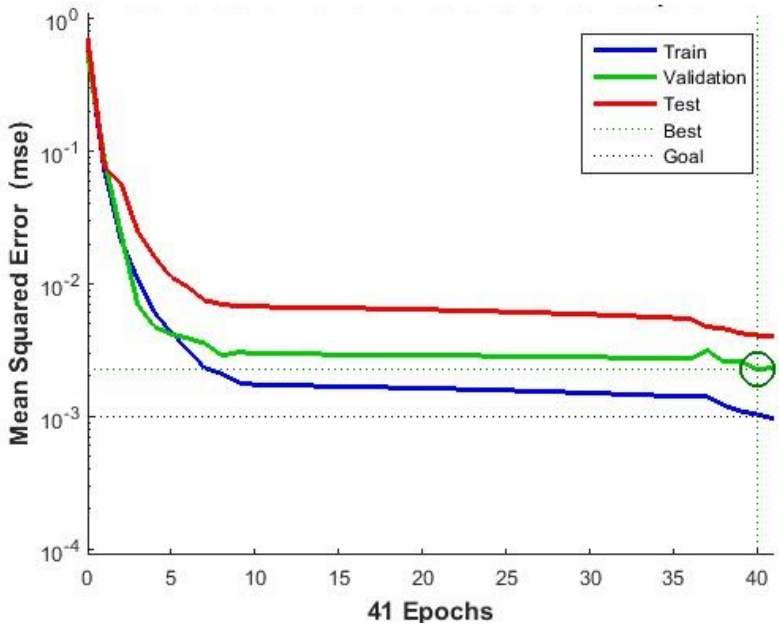

Figure 5.5: MSE graph for Levenberg-Marquardt method at 60:40, LM-15H

In all three learning schemes, rate of convergence was fast, causing the network to learn by heart (overfitting). Table 5.10 and fig. 5.6 shows a significant difference in train and testing R values.

Also, reliability of this method increases with increase in training size.

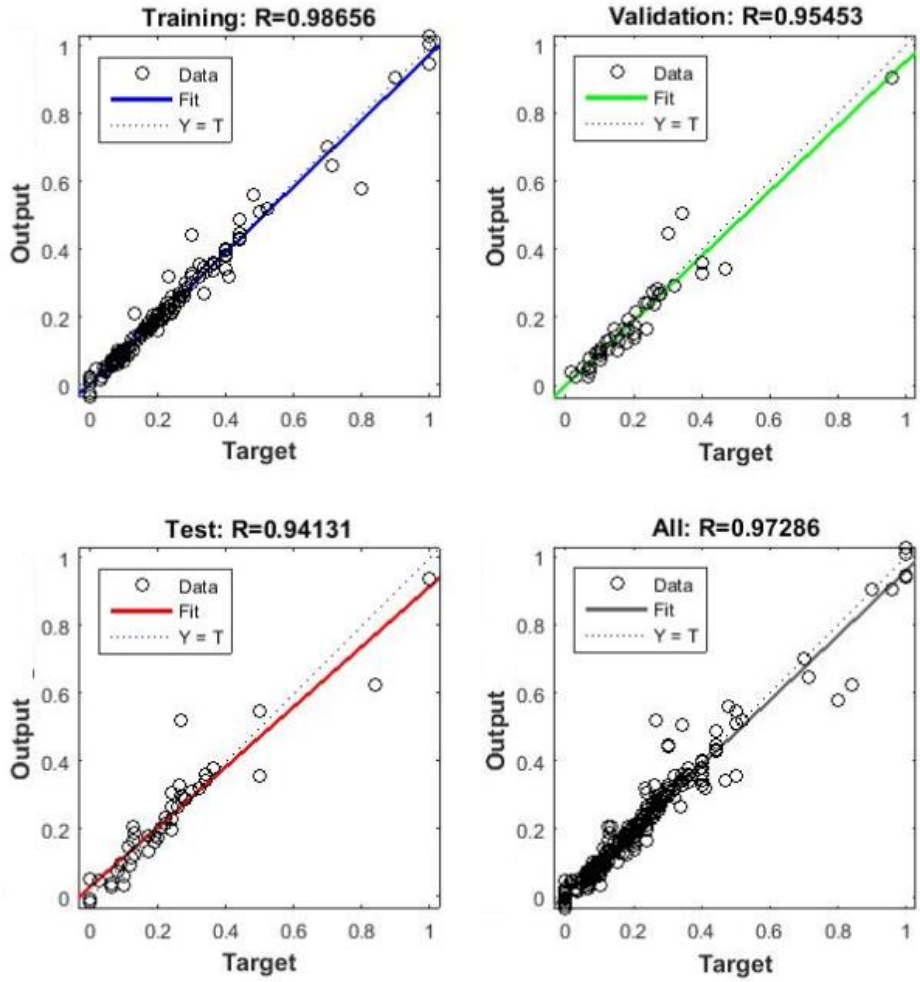

Figure 5. 6: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in Levenberg-Marquardt Method at 60:40, LM-15H

5.5. Discussion of Results for Bayesian Regularization Backpropagation (BR) Method

All Levenberg-Marquardt are known to have faster convergence when compare with gradient decent algorithms. BR results obtained as presented in Table 5.11, shows that LS1 generally did not show a remarkable change in its MSE and total R value. When size of hidden neurons are changed from 40% to 50%, network performance was found to have increase sharply, and decreased as number of hidden neurons increases further from 5 units except with a sharp increase at 15 units. Furthermore within LS3, total R values in hidden neuron sizes of 5, 10, 15 and 20 units are higher than values in 25 hidden neurons. Model for predicting carbonation depth in concrete using Bayesian regularization

backpropagation method require smaller amount hidden units for optimal network performance. Since total R values recorded in the learning schemes are high with LS3 as the optimal learning scheme, this can be used for prediction of carbonation depth in concrete. In fig. 5.7 and 5.8 optimised network performance and R values combination is presented. In the performance graph, it can be seen that the training subset did not converge to 0.001 MSE unlike other methods, rather network was stop when minimum gradient was reached.

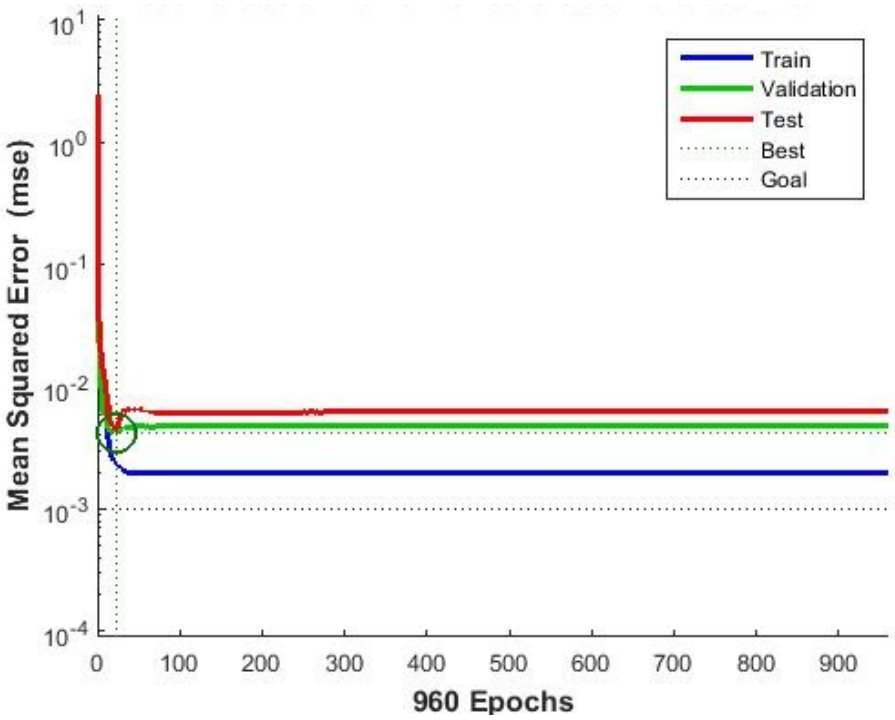

Figure 5.7: MSE graph for Bayesian regularization backpropagation method at 60:40, BR-20H

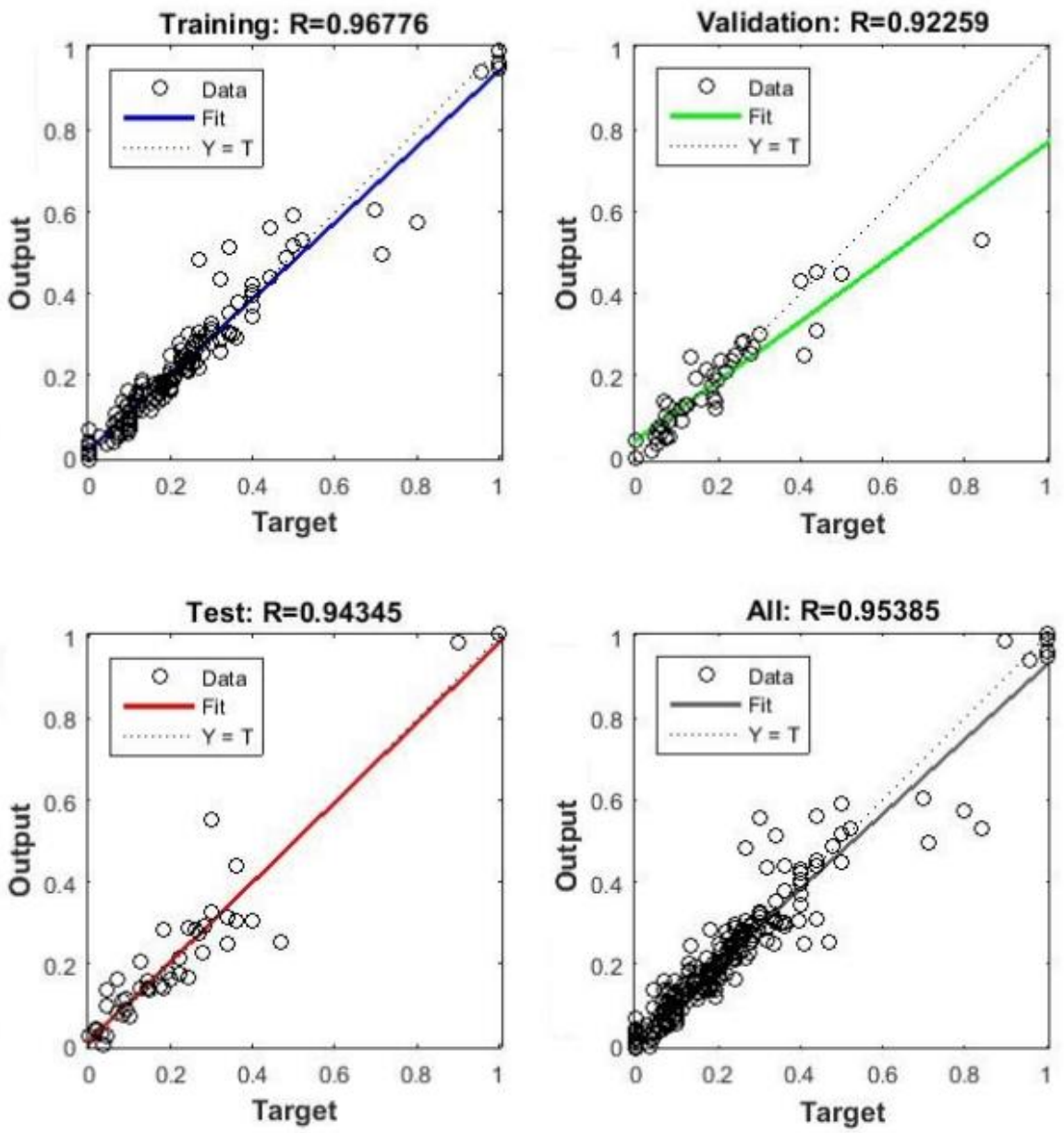

Figure 5.8: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in Bayesian regularization backpropagation method at 60:40, BR-20H

Table 5.10: Cross validation analysis for Levenberg-Marquardt Method

LS	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	0.987	0.916	0.959	29	0.963	0.919	0.944	46	0.977	0.963	0.942	41	0.982	0.891	0.949	36	0.986	0.891	0.94	41
50:50	0.980	0.953	0.967	23	0.983	0.930	0.963	26	0.967	0.964	0.958	30	0.989	0.827	0.965	24	0.990	0.901	0.958	30
60:40	0.990	0.898	0.96	28	0.976	0.936	0.959	28	0.987	0.941	0.973	19	0.988	0.929	**0.972**	20	0.977	0.949	**0.972**	20

Table 5.11: Cross validation analysis for Bayesian "Regularization backpropagation" (BR) Method

LS	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	0.967	0.912	0.943	39	0.971	0.920	0.937	43	0.976	0.917	0.941	40	0.978	0.918	0.946	37	0.972	0.933	0.938	43
50:50	0.966	0.952	0.952	33	0.972	0.973	0.948	36	0.983	0.946	0.953	33	0.962	0.907	0.946	37	0.977	0.940	0.946	37
60:40	0.966	0.925	0.952	33	0.967	0.959	0.954	32	0.949	0.935	0.950	34	0.968	0.943	0.954	32	0.948	0.948	0.947	36

Table 5.12: Cross validation analysis for BFGS method (LS1)

LS1	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	BFG	0.986	0.902	0.954	32	0.983	0.9376	0.954	32	0.979	0.922	0.948	36	0.97	0.934	0.931	48	0.962	0.848	0.933	49
	OSS	0.986	0.950	0.956	33	0.971	0.924	0.943	39	0.983	0.916	0.942	40	0.986	0.892	0.944	38	0.972	0.901	0.929	48

Table 5.13: Cross validation analysis for BFGS Method (LS2)

LS2	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
50:50	BFG	0.988	0.948	0.966	26	0.9803	0.8865	0.951	34	0.987	0.922	0.966	24	0.958	0.956	0.948	37	0.981	0.904	0.951	34
	OSS	0.980	0.952	0.961	27	0.961	0.929	0.951	35	0.979	0.952	0.965	25	0.987	0.926	0.957	31	0.970	0.933	0.947	39

Table 5.14: Cross validation analysis for BFGS Method (LS1)

LS3	Opt fcn	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
		Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
60:40	BFG	0.978	0.943	0.971	20	0.986	0.9204	0.960	27	0.972	0.894	0.957	29	0.973	0.951	0.963	25	0.981	0.945	0.968	22
	OSS	0.985	0.987	0.97	21	0.985	0.965	0.973	20	0.983	0.951	0.969	21	0.986	0.929	0.967	23	0.973	0.823	0.956	31

Table 5.15: Cross validation analysis for "random order incremental training with learning functions" (R) Method

LS	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	0.981	0.915	0.917	63	0.972	0.875	0.944	44	0.974	0.942	0.946	41	0.982	0.916	0.942	40	0.988	0.831	0.904	70
50:50	0.981	0.892	0.945	38	0.942	0.850	0.862	91	0.953	0.894	0.91	63	0.989	0.850	0.934	48	0.984	0.879	0.916	57
60:40	0.972	0.893	0.951	34	0.972	0.967	0.968	23	0.987	0.880	0.953	32	0.976	0.949	0.95	34	0.980	0.867	0.959	39

Table 5.16: Cross validation analysis for "Resilient backpropagation" (RP) Method

LS	ANN-1 (5H)				ANN-2 (10H)				ANN-3 (15H)				ANN-4 (20H)				ANN-5 (25H)			
	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}	Train R	Test R	Total R	MSE 10^{-3}
40:60	0.973	0.967	0.951	34	0.977	0.881	0.949	35	0.988	0.860	0.939	45	0.967	0.948	0.947	36	0.946	0.927	0.937	42
50:50	0.982	0.961	0.964	25	0.955	0.898	0.956	30	0.963	0.952	0.944	38	0.981	0.964	0.958	29	0.975	0.928	0.939	42
60:40	0.970	0.987	0.967	23	0.980	0.841	0.967	23	0.982	0.936	0.957	29	0.987	0.939	0.958	29	0.984	0.920	0.967	23

5.6. Discussion of Results for BFGS Methods

OSS is partially BFGS, result in table 5.1-3 OSS and BFG have one common trend; increase in hidden neurons decreases the number of network epoch.

In LS1 both OSS and BFG, 5 hidden neurons was the optimal network hidden unit size, its performance (total R value and MSE) decrease as the size of hidden neuron were increased. Moreover, when the training subset was increased from 40% to 50 and 60 % respectively, network was seen to have improved total R values and MSE at 15 hidden neurons in LS2 for both OSS and BFG, and at 5 hidden neurons at LS3.

Generally, reliability within the learning scheme improved with increase in training subset. However this method is also suitable for predicting carbonation depth in concrete owning to its performance behavior as shown in fig 5.9. Training subset was progressive while validation and testing subset exhibited similar trend in their response to training experience.

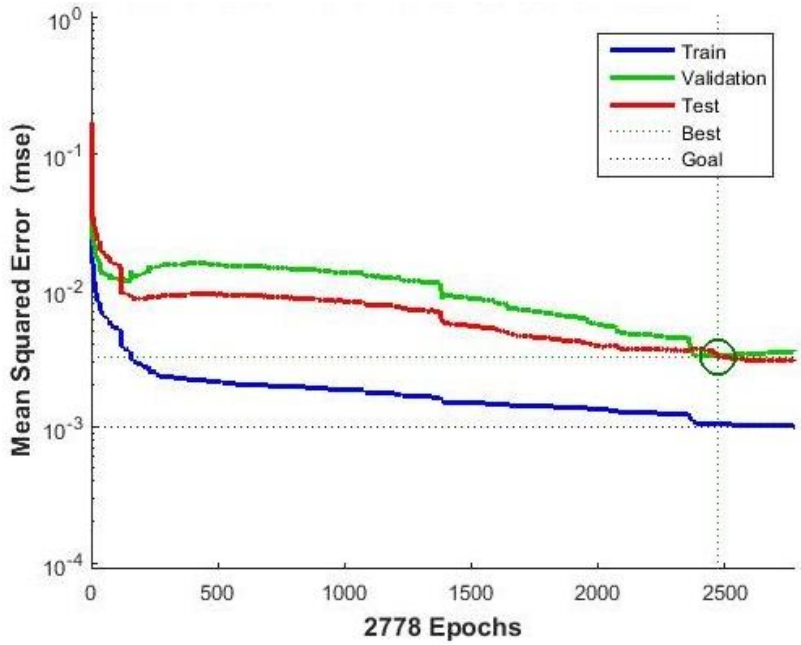

Figure 5.9: MSE graph for BFGS Methods at 60:40, OSS-10H.

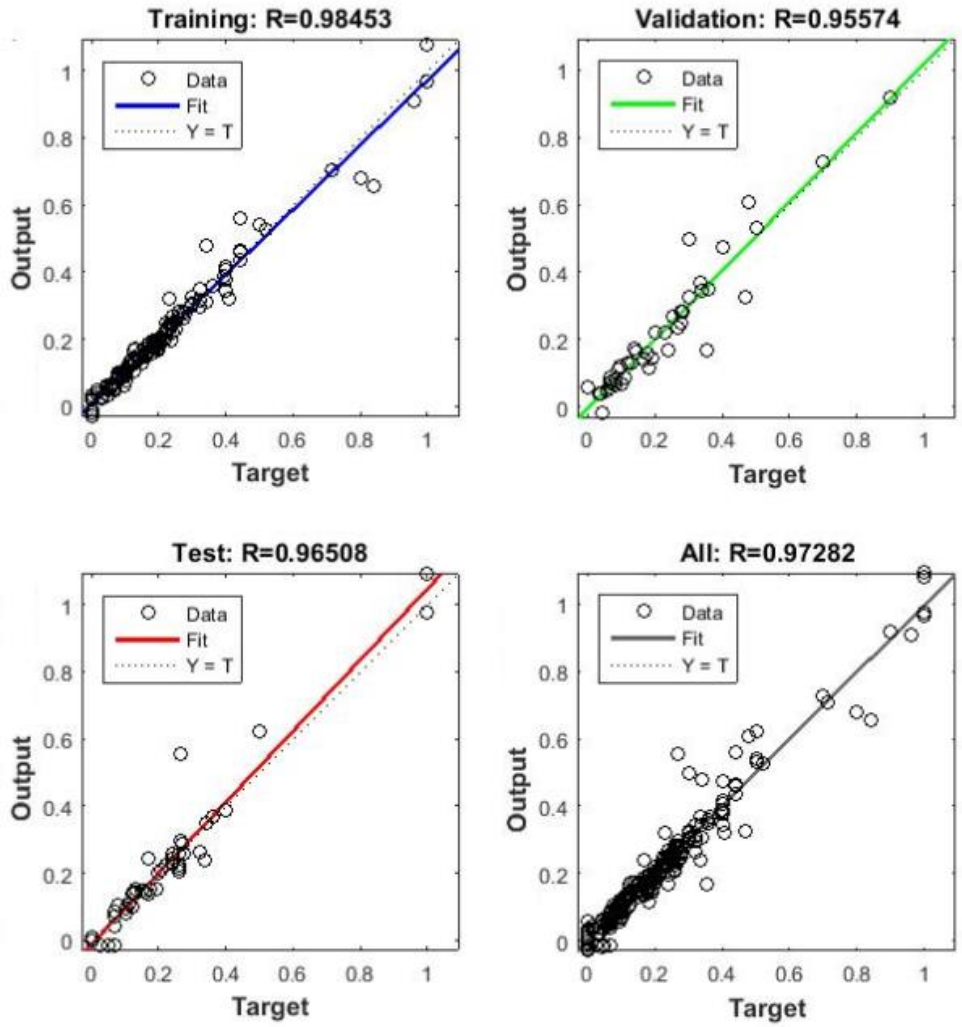

Figure 5. 10: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in BFGS Methods at 60:40, OSS-10H

5.7. Discussion of Results for "Random order incremental training with learning functions" (R) Method

"Random order incremental training with learning functions" is very expensive to apply in carbonation problem. A minimum of 6 hours is for each training set as against an average of 5mins when compared to other optimisation. Despite a slower training rate per each epoch approximately 1.08 sec/epoch, R values recorded in the three learning schemes are low when compared to other optimization methods. Additional cost is required to optimize the model.

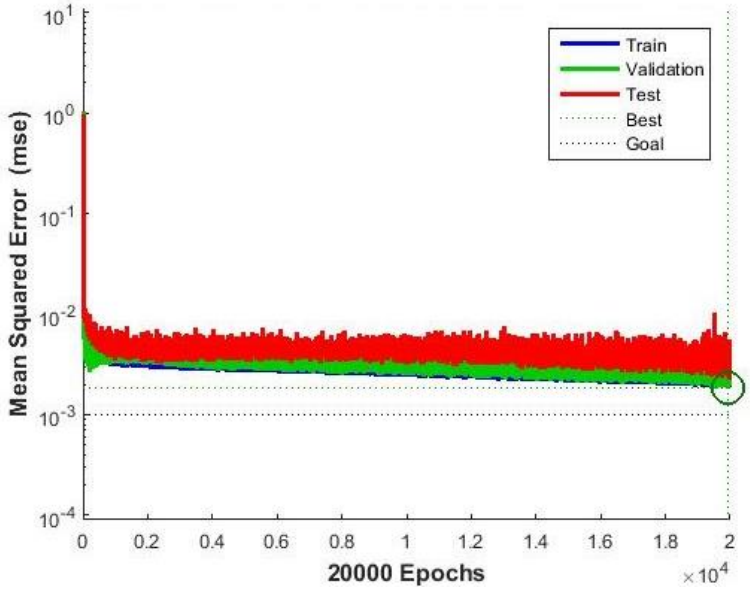

Figure 5.11: MSE graph for "Random order incremental training with learning functions" Method at 60:40, R-10H

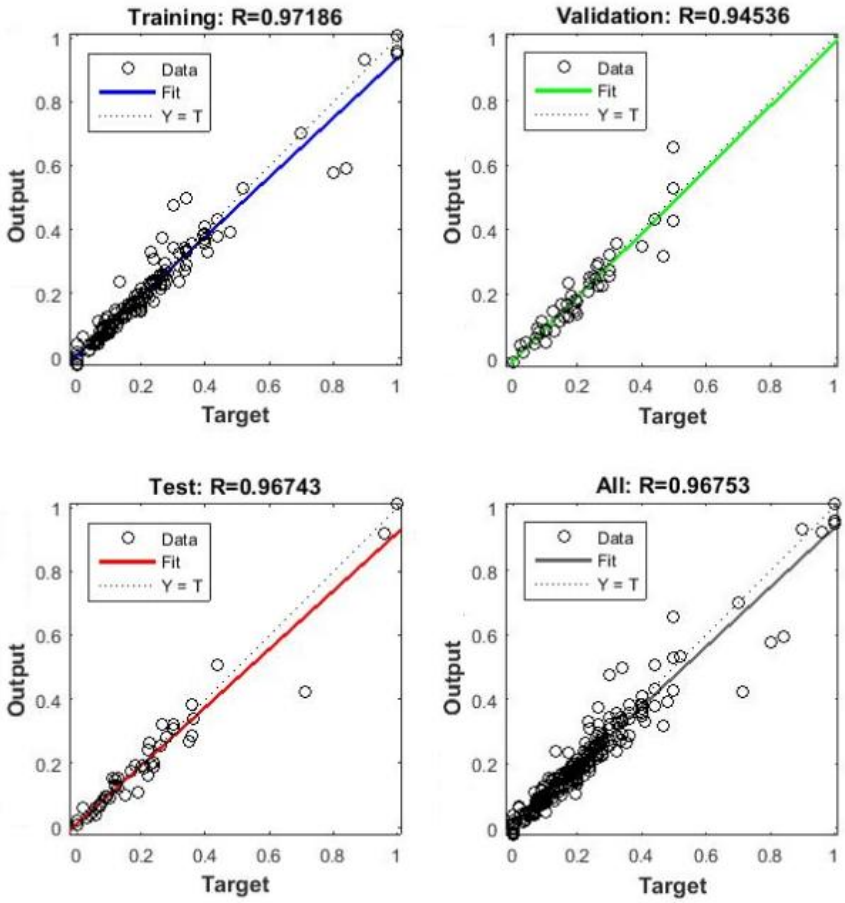

Figure 5.12: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in for "Random order incremental training with learning functions" at 60:40, R-10H

Increase in training subset also increased network performance between LS1 and LS3. With exception of learning scheme 2, network total R values was optimized between 10 and 15 hidden neurons. Overfitting within this method can be observed in all the learning schemes, since the method did not have adjustable learning parameters as seen in the other methods, to control the performance of the network. Where other methods are available, this method is not suitable for carbonation depth prediction. But fig. 5.11 shows clearly that the 3 subsets exhibit similar trait, although none converged to 0.001 MSE, but the case of overfitting was minimised.

5.8. Discussion of Results for "Resilient backpropagation" (RP) Method

Results obtained when compared with steepest descent methods, shows improvement in total R values and updated MSE. Within the learning scheme, increase in training subset from 40-50-60 %, increases network performance.

Increase in size of hidden neurons did not affect network iteration. Rather in LS1 &2, increase in hidden neurons decreases network performance. This is seen in sharp decreases in total R values and corresponding decrease in MSE. Resilient backpropagation application in predicting carbonation depth in concrete is reliable since total R values are high and training subset converged to 0.001 MSE. Care should be taken why training this model, network should be stopped when there is a significant decrease is R values of testing and validating subset as can be seen in Fig. 5.13.

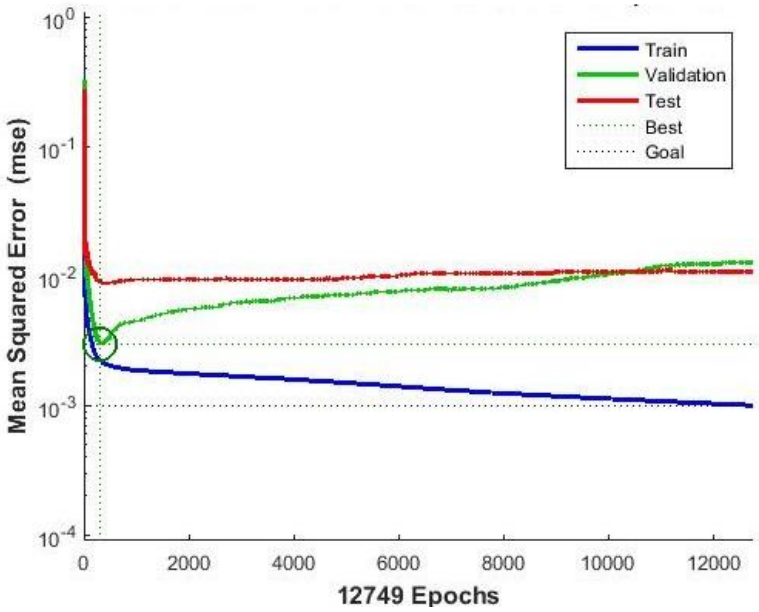

Figure 5.13: MSE graph for "Resilient backpropagation" (RP) Method at 60:40, RP-10H

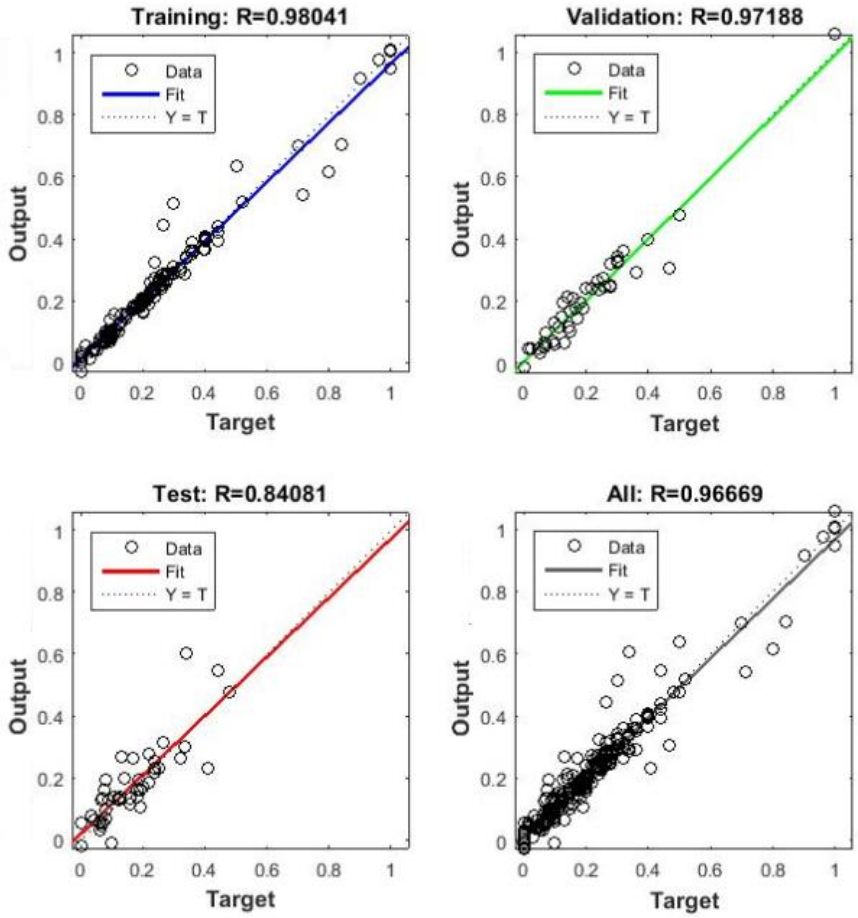

Figure 5.14: Regression plot for predicted against measured carbonation depth for training, testing and validation dataset for highest R value obtained in "Resilient backpropagation" Method at 60:40, RP-10H

5.9. Discussion of Result for Comparison with Existing Literature

Kwon and Song (2010), in their work used backpropagation neural networks with four inputs of water/cement ratio, unit weight of cement, total aggregate volume and relative humidity. Diffusion coefficient was set as output neuron and a total of 12 dataset was trained using tansig as transfer function. Network threshold for epoch and MSE was set at 1000 and 0.0001 respectively. However, details of optimization function, number of hidden layer and hidden units used was not given. In comparison, 225 dataset and 18 input neurons was provided within this study giving a detailed and extensive input units. There was also increase within input nodes accounting for better learning experience for the network. Besides cement content, this study also provided cement compound composition responsible for progress of carbonation in concrete.

With respect to aforementioned information on ANN architecture and training methods, output neuron of Kwon and Song (2010) output neuron was diffusion coefficient which differs from output of this study. Although it is an important parameter showing the progress of carbonation in concrete, maximum error obtained was 6.3% which is within comparable region with error values obtained in this study.

In CaPrM, at study presented Taffese et. al. (2015), 25 input neurons was extracted from concrete mix properties, curing properties, environmental conditions age of concrete and properties of fresh and hardened concrete. Carbonation depth in concrete was network output. The study trained a feedforward backpropagation neural network using Levenberg-Marquardt and tansig as optimization and transfer functions respectively. When compared to results in this study, Taffese et. al. (2015) had R value of 0.9701 is within comparable region for LM method since 0.97286 was recorded against LM method in this study.

CHAPTER 6
CONCLUSION AND RECOMMENDATION

6.1. Conclusion

225 datasets were extracted from nine different research papers presenting experimental studies on carbonation problem in concrete. Dataset used, includes both accelerated test and natural carbonation. Three learning scheme of sample distribution was adopted along with five different sets of hidden neurons. The combination of three learning scheme and five hidden neuron sizes was training using 14 different optimization algorithm in a feedforward backpropagation algorithm of neural networks.

210 sets of training were performed on 2015a MATLAB. Results obtained were analyzed based on MSE, total R value of the network and cross-validation behavior of the set. It was observed that learning scheme-1 containing 40:60 % of training:testing sample performed poorly in all the optimization methods and sizes of hidden neurons. MSE within LS1 was the highest. Thereafter, increase in training subset from 40 through 50 and 60 % increased network performance.

In predicting carbonation depth in concrete, optimized hidden neuron range was found to be within 10 to 15 hidden neurons' yielding effective results with all the learning scheme (see table 5.17). Moreover, models with higher hidden neuron sizes tend to have lower computation cost but not necessarily improving the network. R values decreases when size of hidden neurons increases from 15 to 25 units. Likewise the MSE was also on the increase within the above mention hidden neuron range.

Steepest gradient-descent (GD, GDA, GDM and GDX) and Random order incremental training with learning functions algorithms performance were poor, network computation cost was also expensive. Generally all 5 algorithms training were stopped when 20,000 iteration was reached without training subset converging to 0.001 MSE.

Conjugate-gradient descent results shows a consistent improvement alone the learning scheme. Besides LM, BFG and BR, the conjugate gradient descent have less iteration. This optimized the network performance.

Table 5.17: Summary of results obtained within the learning scheme

Learning scheme	Best optimization fun	R value	MSE	Iteration	Hidden neuron size
LS1	CGP	0.966	0.0023	464	10
LS2	CGP	0.975	0.0018	547	10
LS3	SCG	0.980	0.0014	676	10

For carbonation depth prediction using artificial neural networks, with respect to obtained results, dataset feed to the network should be sufficient enough to minimize overfitting and extrapolation in the model. Where it is necessary for steepest gradient descent method is to be adopted, training epoch should be increased beyond 20,000. Oscillation should be avoided by using smaller learning rate and where adaptation learning is applicable. Adequate attention should be paid to size of *lr_inc* and *lr_dec* (multiplication constant).

However, Batch steepest descent algorithms was found to have higher computational cost, while search line based algorithms and Levenberg-Marquardt step size methods have lowered cost and provided better performance within the set threshold.

A combination of conjugate gradient descent and Levenberg-Marquardt methods (SCG) was found to yield the best optimized result with respect to R value of 0.980, MSE of 0.0014 and computational cost 676 iterations under 10 hidden neurons with 60:40 testing:training distribution.

6.2. Recommendations for Further Studies

In this study three layer feedforward backpropagation was used with one hidden layer. Although various ANN references, which are not directly related to carbonation problem, used in this study suggested the used one hidden layer is sufficient, it is important to carry out future studies on influence of using two or more hidden layer sizes for carbonation studies.

Furthermore, the behaviour of the model when training sample exceed 60% and alternating testing and validating ratio should also be investigated with larger hidden neuron size.

Although backpropagation algorithm is widely used, other algorithms in neural networks should be studied and result compared to establish a more reliable ANN algorithm for carbonation studies.

Comparison with multi linear regression analysis is believe to provide an external independent assessment of the network performance and a better perspective capabilities of ANN in prediction of carbonation in concrete. Such study is also recommended.

REFERENCES

Ahmad, S (2003). Reinforcement corrosion in concrete structures, its monitoring and service life prediction—a review. *Cement and Concrete Composites, 25*(4), 459-471.

Atici, U. (2011). Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network. *Expert Systems with Applications, 38*(8), 9609-9618.

Atkins, P. W. (1994). *Physical Chemistry, Fifth Edition*. Oxford: Oxford University Press.

Auroy, M., Poyet, S., LeBescop, P., Torrenti, J.-M., Charpentier, T., Moskura, M., & Bourbon, X. (2015). Impact of carbonation on unsaturated water transport properties of cement-based materials. *Cement and Concrete Research, 74*, 44–58.

Baroghel-Bouny, V. (2007). Water vapour sorption experiments on hardened cementitious materials. part I: essential tool for analysis of hygral behaviour and its relation to pore structure. *Cement and Concrete Research, 37*(3), 414–437.

Beatty, J. (1995). *Principles of behavioral neuroscience*. McGraw-Hill

Becker, S., & Plumbley, M. (1996). Unsupervised neural network learning procedures for feature extraction and classification. *Applied Intelligence, 6*(3), 185-203.

Belciug, S., & Gorunescu, F. (2014). Error-correction learning for artificial neural networks using the bayesian paradigm. application to automated medical diagnosis. *Journal of Biomedical Informatics, 52*, 329-337.

Benfenati, F., Onofri, F., & Giovedi, S. (1999). Protein-protein interactions and protein modules in the control of neurotransmitter release. *Philosophical transactions of the Royal Society of London, B*(354), 243-57.

Benfenati, F. (2007). Synaptic plasticity and the neurobiology of learning and memory. *ACTA BIOMED, 78*(1), 58-66.

Bennamoun, M. (2015). CS407 Neural Computation, Lecture 3: Neural Network Learning Rules. Retrieved May 31, 2016 from http://teaching.csse.uwa.edu.au/units/CITS4210/lectureNotes/Lect3-UWA.pdf

Bensted, J. (1983). Hydration of portland cement. *Advances in Cement Technology*, 307-347.

Berhane, Z. (1983). Heat of hydration of cement pastes at different temperatures. *Cement and Concrete Research, 13*(1), 114-118.

Berkely, K.G.C., & Pathmanaban, S. (1990). *Cathodic protection of reinforcement steel in concrete*. London: Butterworths & Co. Ltd.

Berry, E.E. (1980). Strength development of some blended cement mortars. *Cement Concrete Research*, *10*(1), 1–11.

Blausen 0657 MultipolarNeuron by BruceBlaus - Own work. Licensed under CC BY 3.0 via Commons. Retrieved on January 30, 2016 from https://commons.wikimedia.org/wiki/File:Blausen_0657_MultipolarNeuron.png#/media/File:Blausen_0657_MultipolarNeuron.png

Bose, N. K & Liang, P. (1996). *Neural network fundamentals with graphs, algorithms, and applications*. New Jersey, NJ: McGraw-Hill, Inc. Hightstown.

Burns, R. A. (2003). *Fundamentals of Chemistry, Fourth Edition*. New Jersey, NJ: Prentice Hall.

Burrows, A., Holman, J., Parsons, A., Pilling, G., & Pricce, G. (2009). *Chemistry*^{3}, *Introducing inorganic, organic and physical chemistry*. Oxford: Oxford University Press.

Bye, G.C. (1983). *Portland Cement Composition, Production and Preparation*. Oxford: Pergamon Press Ltd.

Castellote, M., Andrade, C., Turrillas, X., Campo, J., & Cuello, G. J. (2008). Accelerated carbonation of cement pastes in situ monitored by neutron diffraction. *Cement and Concrete Research, 38*, 1365–1373.

Castellote, M., Fernandez, F., Andrade, C., & Alonso, C. (2009). Chemical changes and phase analysis of OPC pastes carbonated at different CO_2 concentrations. *Materials and Structures, 42*, 515–525.

Chatterjee, A. K. (2011). Chemistry and engineering of the clinkerisation process — incremental advances and lack of breakthroughs. *Cement and Concrete Research, 41* (7), 624-641.

Chen, J.R., & Mars, P. (1990). Stepsize variation methods for accelerating the back-propagation algorithm. In M. Caudill (Ed.), *In Proceedings of the International Joint Conference on Neural Networks*, (pp. 601–604). Piscataway, New Jersey: IEEE Neural Networks Council.

Ciach, T. D., & Swenson, E.G. (1971). Morphology and microstructure of hydrating portland Cement and its constituents IV. changes in hydration of a C3S, C3S, C3A, C4AF and gypsum paste with and without the admixtures triethanolamine and calcium lignosulphonate. *Cement and Concrete Research, 1*(4), 367-383.

Cohen, M.D, Goldman, A., & Chen, W.F. (1994). The role of silica fume in mortar: transition zone vs. bulk paste modification. *Cement and Concrete Research*, *24*, 95-98.

Conciatori, D., Sadouki, H., & Brühwiler, E., (2008). Capillary suction and diffusion model for chloride ingress into concrete. *Cement and Concrete Research, 38*, 1401–1408.

Dahl, E. D. (1987). Accelerated learning using the generalized delta rule. In M. Caudill & C.B. Butler (Eds.), *In Proceedings of the International Joint Conference on Neural Networks*, (Vol. II, pp. 523–530). Piscataway, New Jersey: SOS Printing.

Diamond, S. (2004). The microstructure of cement paste and concrete—a visual primer. *Cement and Concrete Composites, 26*(8), 919-933.

Skoog, D. A., & Leary, J. J. (1992). *Principles of instrumental analysis*. Saunders College Publishing.

Drouet, E., Poyet, S., & Torrenti, J. (2015). Temperature influence on water transport in hardened cement pastes. *Cement and Concrete Research*, *76*, 37–50.

Duan, P., Shui, Z., Chen, W., & Shen, C. (2013). Efficiency of mineral admixtures in concrete: Microstructure, compressive strength and stability of hydrate phases. *Applied Clay Science*, 83–84.

Duch, W., & Jankowski, N. (1999). Survey of neural transfer functions. *Neural Computing Surveys*, 2, 163-212.

Earnest C. M. (1984). Modern Thermogravimetry. *Analytical Chemistry*. *56*(13), 1471-1486.

Ebbing, D. D., & Garmmon, S. D. (2005). *General Chemistry, Eighth Edition*. Boston: Houghton Mifflin Company.

Elhag, T. M. S (2002). *Tender price modelling: Artificial neural networks and regression techniques* (Doctoral dissertation, University of Liverpool). Retrieved on June 13, 2016 from http://ethos.bl.uk

Fahlman, S.E. (1989). Faster-Learning variations on back-propagation: an empirical study. In D. Touretzky, G. Hinton, & T. Sejnowski (Eds.), *Proceedings of the Connectionist Models Summer School* (pp. 38–51). San Mateo, CA: Morgan Kaufmann.

Fausett, L. V. (1994). *Fundamentals of neural networks: Architectures, algorithms, and applications*. Prentice-Hall international editions. Prentice-Hall.

Finlay, J. & Dix, A. (1996). *An Introduction to artificial intelligence*. CRC Press, London.

Franzini, M. A. (1987). Speech recognition with back propagation. *In Proceedings of the Ninth Annual Conference of the IEEE Engineering in Medicine and Biology Society*. (Vol. 33 , pp. 1702–1703). New York, NY: IEEE.

Gambhir, M. L. (1995). *Concrete Technology, Second Edition*. New Delhi: Tata McGraw-Hill.

Glover, G. M., & Raask, E. (1972). Water Diffusion and Microstructure of Hardened Cement Pastes, *Matériaux et Construction*, *5*(5), 315-322.

Grattan-Bellew, P.E. (1996). Microstructural investigation of deteriorated Portland cement concretes. *Construction and Building Materials*, *10*(1) p. 3-16.

Gross, G. W., Rhoades, B K., Azzazy, H. M.E., & Ming-Chi Wu (1995). The use of neuronal networks on multielectrode arrays as biosensors. *Biosensors and Bioelectronics*, *10*(6), 553-567.

Gross, G. W, Harsch, A., Rhoades, B. K, & Göpel, W. (1997). Odor, drug and toxin analysis with neuronal networks in vitro: extracellular array recording of network responses. *Biosensors and Bioelectronics*, *12*(5), 373-393.

Hagan, M. T., Demuth, H. B. & Beale, M. H. (1996). *Neural network design*. Boston: PWS Publishing, MA.

HanŽiČ, L., Kosec, L., & Anžel, I. (2010). Capillary absorption in concrete and the Lucas–Washburn equation. *Cement & Concrete Composites*, *32*, 84–91.

Harel, D. & Feldman, Y. (2004). *Algorithmics: The Spirit of Computing (3rd Edition) 3rd Edition*. Addison-Wesley.

Haykin, S. (1999). *Neural networks: A comprehensive foundation*. Prentice Hall.

Henry, R. L. & Kurtz, G. K. (1961). *Water vapor transmission of concrete and of aggregates*. Technical Report R-244. U. S. Naval Civil Engineering Laboratory. Retrieve on September 20 15, 2016 from http://www.dtic.mil/dtic/tr/fulltext/u2/257789.pdf

Hewlett, P. (2004). *Lea's Chemistry of Cement and Concrete, Fourth Edition*. Elsevier Science & Technology Books.

Hille, B. (1984). Ionic channels of excitable membranes: *Sinauer Associates.*

Houst, F. Y., (1996). The role of moisture in the carbonation of cementitious materials. *Internationale Zeitschrift für Bauinstandsetzen*, *2*(1), 49-66.

Matheworks: Neural networks perceptron. Retrieved on June 15, 2016 from mathworks.com/help/nnet/ug/perceptron-neural-networks.html

Hyvert, N., Sellier, A., Duprat, F., Rougeau, P., & Francisco, P. (2010). Dependency of C–S-H carbonation rate on CO2 pressure to explain transition from accelerated tests to natural carbonation. *Cement and Concrete Research*, *40*(11), 1582-1589.

Illston, J. M. (1994). *Construction Materials: Their nature and behavior*. New York, NY: E & FN Spon.

Janabi-Sharifi, F. & Wilson, W. J. (1993). A multi-layered learning model. *Journal of Intelligent and Robotic Systems, 8*(3), 399–423.

Jödecke, M., Xia, J., Kamps, Á. & Maurer, G. (2015). Influence of (phenol and sodium sulfate) on the solubility of carbon dioxide in water. *The Journal of Chemical Thermodynamics*, *86*, 123-129.

Johannesson, B., & Utgenannt, P. (2001). Microstructural changes caused by carbonation of cement mortar. *Cement and Concrete Research, 31*, 925-931.

Juenger, M. C. J., & Siddique, R. (2015). Recent advances in understanding the role of supplementary cementitious materials in concrete. *Cement and Concrete Research, 78*, 71–80.

Kamiński, M. & Zielenkiewicz, W. (1982). The heats of hydration of cement constituents. *Cement and Concrete Research*, *12*(5), 549-558.

Kandel, E., Schwartz, J., & Jessell, T (2000). *Principles of Neural Science, Fourth Edition (International Edition)*. McGraw-Hill Companies, Incorporated.

Kapageridis, I. K. (1999). *Application of Artificial Neural Network Systems to Grade Estimation from Exploration Data*. (Doctoral dissertation, University of Nottingham). Retrieved on June 15, 2016 from http://ethos.bl.uk

Karagol, F., Demirboga, F., & Khushefati, W. H. (2015). Behavior of fresh and hardened concretes with antifreeze admixtures in deep-freeze low temperatures and exterior winter conditions. *Construction and Building Materials*, *76*, 388–395.

Kasabov, N. K. (1996). *Foundations of neural networks, fuzzy systems, and knowledge engineering*. London: Marcel Alencar.

Katz, B. (1966). *Nerve, Muscle, and Synapse*. New York, NY: McGraw-Hill.

Khalil, E. A. B., & Anwar, M. (2015). Carbonation of ternary cementitious concrete systems containing fly ash and silica fume. *Water Science, 29*(1), 36-44.

Khashman, A. (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Systems with Applications*, *37*(9), 6233–6239.

Kohonen, T. (1987). Adaptive, associative, and self-organizing functions in neural computing. *Applied Optics, 26*(23), 4910-4916.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*, 59-69.

Kwon, S. & Song, H., (2010). Analysis of carbonation behavior in concrete using neural network algorithm and carbonation modeling. *Cement and Concrete Research, 40*(1), 119-127.

Lammertijn, S. & DeBelie, N. (2008). Porosity, gas permeability, carbonation and their interaction in high-volume fly ash concrete. *Magazine of Concrete Research, 60*(7), 535–545.

Larbi, L. A. (1993). Microstructure of the interfacial zone around aggregate particles in concrete. *Heron, 38*(1), 69.

Laugesen, P. (1993). *Presentation at 4th Euro Seminar on Microscopy of Building Materials*, Visby, Sweden.

Lee, Y., Oh, S. H., & Kim, M. W. (1991). The effect of initial weights on premature saturation in back-propagation training. *In Proceedings of the International Joint Conference on Neural Networks*: (Vol. II, 765–770). Piscataway, NJ: IEEE Neural Networks Council.

Liu, J. (2013). *Radial basis function (RBF) neural network control for mechanical systems: Design, analysis and MATLAB simulation*. Springer Science & Business Media.

Liu, Y., Hou, M., Yang, G., & Han, B. (2011). Solubility of CO2 in aqueous solutions of NACL, KCL, CACL2 and their mixed salts at different temperatures and pressures. *The Journal of Supercritical Fluids, 56*(2), 125-129.

Ludwig, H-M., & Zhang, W. (2015). Research review of cement clinker chemistry. *Cement and Concrete Research, 78*, 24-37.

Ma, Q., Guo, R., Zhao, Z., Lin, Z., & He, K. (2015). Mechanical properties of concrete at high temperature—A review. *Construction and Building Materials, 93*, 371–383.

MacKay, D. J. C. (1992). Bayesian interpolation. *Neural Computation, 4*(3), 415–447.

Mackay, D. M. (1967). *Freedom of Action in a Mechanistic Universe***.** Cambridge: University Press.

MacKenzie, R. C. (1979). Nomenclature in thermal analysis, Part IV. *Thernzochim. Acta, 28*(1), 1-6.

Martys, N. S., & Ferraris C. F. (1997). Capillary transport in mortars and concrete. *Concrete and Research*, *27*(5), 747 – 760.

McMurry, J.E., & Fay, R.C. (2008). *Chemistry, Pearson International Edition, Fifth Edition*. New Jersey, NJ: Prentice Hall.

Mendel, J. M., & McLaren, R.W. (1970). Reinforcement learning control and pattern recognition systems. *Adaptive, Learning and Pattern Recognition Systems: Theory and Applications* (pp. 287-317). New York, NY: Academic Press.

Menon, A., Mehrotra, K., Mohan, C., & Ranka, S. (1996). Characterization of a class of sigmoid functions with applications to neural networks. *Neural Networks, 9*(5), 819-835.

Morandeau, A., Thiéry, M., & Dangla, P. (2014). Investigation of the carbonation mechanism of CH and C-S-H in terms of kinetics, microstructure changes and moisture properties. *Cement and Concrete Research*, *56*, 153–170.

MoukamKakmeni, F. M., & Nguemaha, V. M. (2016). Enhancement of synchronization in inter–intra-connected neuronal networks. *Physics Letters*, *380*(1), 200-206.

Neville A.M. (2003). *Properties of Concrete, Fourth Edition*. Harlow: Prentice Hall.

Papadakis, V. G., Vayenas, C. G., & Fardis, M. N. (1991). Physical and chemical characteristics affecting the durability of concrete. *American Concrete Institute Materials Journal*, *88*(2), 186-196.

Parekh, R., Balakrishnan, K., & Honavar, V. (1993). An empirical comparison of flat-spot elimination techniques in back-propagation networks. In M.L. Padgett (Ed.), *Proceedings of the Third Workshop on Neural Networks*: Academic/Industrial/NASA/Defense (pp. 55–60). San Diego, CA: Society for Computer Simulation.

Parrott, L. J. (1987). Carbonation, moisture and empty pores. *Cement and Concrete Association*, *4*, 111–118.

Parrott, L. J. (1991). A Review of carbonation in reinforced concrete. *Advances in Cement Research*, 4(15), 111-118.

Parrott, L.J., & Killoh, D.C. (1989). Carbonation in a 36 year old in-situ concrete. *Cement and Concrete Research*, *19*(4), 649–656.

Platret, G., & Deloye, F. X. (1994). Thermogravimetry and carbonation of cements and concretes, *Actes des Journées des Sciences de l'Ingénieur du réseau des Laboratoires des Ponts et Chaussées* (237–243). Paris, Publication LCPC.

Practical Chemistry. The reaction between carbon dioxide and water. Retrieved on August 18, 2015 from http://www.nuffieldfoundation.org/practical-chemistry/reaction-between-carbon-dioxide-and-water

Raina, V. K. (1990). *Concrete for construction: Facts and practice.* New Delhi: Tata McGraw-Hill.

Riedmiller, M., & Braun, H (1993). a direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks*, (586–591), San Francisco, CA.

RILEM (1988), CPC-18 Measurement of hardened concrete carbonation depth. *RILEM Publications SARL, 21*(126), 453 – 455.

Robitaille, B., Marcos, B, Veillette, M., & Payre, G. (1996). Modified quasi-newton methods for training neural Networks. *Computers & Chemical Engineering*, *20*(9), 1133-1140.

Rosenblatt, F (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *6*, 386–408.

Rosenblatt, F. (1961). *Principles of Neurodynamics*, Washington, D.C.: Spartan Press.

Schuth, F., Sing, K.S.W., & Weltkamp, J. (2008). *Handbook of Porous Solids*, Wiley-VCH, Weinheim.

Scrivener, K. L., & Nemati, K. M. (1996). The percolation of pore space in the cement paste/aggregate interfacial zone of concrete. *Cement and Concrete Research*, *26*(1), 35-40.

Sevelsted, T. F., & Skibsted J. (2015). Carbonation of C–S–H and C–A–S–H samples studied by ^{13}C, ^{27}Al and ^{29}Si MAS NMR spectroscopy. *Cement and Concrete Research*, *71*, 56–65.

Si-Moussa, C., Hanini, S., DerricheI, R., Bouhedda, M., & Bouzidi, A. (2008). Prediction of high-pressure vapor liquid equilibrium of six binary systems, carbon dioxide with six esters, using an artificial neural network model. *Brazilian Journal of Chemical Engineering*, *25*(1), 183-199.

Sola, J., & Sevilla (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Transaction on Nuclear Science, 44*(3), 1464- 1468.

Song, H., & Kwon, S. (2007). Permeability Characteristics of Carbonated Concrete Considering Capillary Pore Structure. *Cement and Concrete Research, 37*(6), 909-915.

Spartz, R., & Honavar, V. (1993). An empirical analysis of the expected source values rule. In M.L. Padgett (Ed.), *In Proceedings of the Third Workshop on Neural Networks*: Academic/Industrial/NASA/Defense (pp. 95–100). San Diego, CA: Society for Computer Simulation.

Stark, J. (2011). Recent advances in the field of cement hydration and microstructure analysis. *Cement and Concrete Research, 41*, 666–678.

Structure And Cell Biology Of The Neuron: Chapter 4 Psych333 –Handouts-Course Pack. University of Washington. Retrieved on January 30, 2016 from http://courses.washington.edu/psych333/handouts/coursepack/ch04-Neuron_structure_and_cell_biology.pdf

Sutton, R. S., & Barto, A. G (1998). *Reinforcement Learning: An Introduction*. MIT Press. Michigan.

Svozil, D., Kvasnicka, V. & Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, *39*(1), 43-62.

Taffese, W & Sistonen, E. (2016). Neural network based hygrothermal prediction for deterioration risk analysis of surface-protected concrete façade element. *Construction and Building Materials*, *113*, 34-48.

Taffese, W.; Sistonen, E., & Puttonen, J., (2015). CaPrM: Carbonation prediction model for reinforced concrete using machine learning methods. *Construction and Building Materials*, 100, 70-82.

Tanaka, K., & Kurumisawa, K. (2002). Development of technique for observing pores in hardened cement paste. *Cement and Concrete Research*, *32*(9), 1435-1441.

Taylor, H.F.W. (1997). *Cement Chemistry, Second Edition*. London: Academic Press Thomas Telford.

Telschow, S. (2012). Clinker Burning Kinetics and Mechanism (Doctoral dissertation, Technical University of Denmark, DTU). Retrieved on June 30, 2016 from http://orbit.dtu.dk/ws/files/51216918/PhD_thesis_Samira_Telschow.PDF

Thiery, M. , Villain, G., Dangla, P., & Platret, G. (2007). Investigation of the carbonation front shape on cementitious materials: Effects of the chemical kinetics. *Cement and Concrete Research*, *37*, 1047–1058.

Uchikawa, H., & Okamura, T. (1993). *Binary and ternary components blended cement. In: Sarkar, S.L. (Ed.), Mineral Additives in Cement and Concrete*. New Delhi: ABI Books Private.

Visser, J.H.M. (2014). Influence of the carbon dioxide concentration on the resistance to carbonation of concrete. *Construction and Building Materials, 67,* 8-13.

Vitela, J.E., & Reifman, J. (1997). Premature saturation in backpropagation networks: mechanism and necessary conditions. *Neural Networks*, *10*(4), 721–735.

Washburn, E.W. (1921). The dynamics of capillary flow. *Phys*. *Rev*, *17*(3), 273–283.

Watanabe, K., & Tzafestas, S. G. (1990). Learning Algorithms for Neural Networks with the Kalman Filters .*Journal of Intelligent and Robotic Systems, 3,* 305-319.

Wu, M., Johannesson, B., & Geiker, M. (2014). A study of the water vapor sorption isotherms of hardened cement pastes: Possible pore structure changes at low relative humidity and the impact of temperature on isotherms. *Cement and Concrete Research, 56,* 97–105.

Yadollahi, A., Nazemi, E., Zolfaghari, A. & Ajorloo, A. M. (2016). Application of artificial neural network for predicting the optimal mixture of radiation shielding concrete. *Progress in Nuclear Energy*, *89,* 69-77.

Yuan, H., Xiong. F., & Huai, X. (2003). A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy. *Computers and Electronics in Agriculture*, *40*(1), 57-64.

Zaragoza, H., & Alché-Buc, F. (1998). Confidence measures for neural network classifiers. *Information Processing and Management of Uncertainty Conference*. Paris, France.

Zhang, Z.; Ma, X. & Yang, Y. (2003). Bounds on the number of hidden neurons in three-layer binary neural networks. *Neural Networks, 16*(7), 995-1002.

Zucker, R. S. (1996). Exocytosis: A molecular and physiological perspective. *Neuron,17,* 1049-55.

APPENDIX A
MATLAB TRAINING FUNCTION CODE

trainbfg

```
%Maximum number of epochs to train
net.trainParam.epochs=20000;
%Show training window
net.trainParam.showWindow=TRUE;
%Epochs between displays (NaN for no
displays)
net.trainParam.show=50;
%Performance goal
net.trainParam.goal=0.005;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-06;
%Maximum validation failures
net.trainParam.max_fail=20000;
%Name of line search routine to use
net.trainParam.searchFcn='srchbac';
net.trainParam.searchFcn='srchbac';
net.trainParam.scal_tol          20=;
net.trainParam.alpha=0.0001;
net.trainParam.beta=0.1;
net.trainParam.bmax=26;
net.trainParam.delta=0.01;
net.trainParam.gama=0.1;
net.trainParam.low_lim=0.1;
net.trainParam.up_lim=0.5;
net.trainParam.maxstep=100;
net.trainParam.minstep=1.00E-006;
```

trainbr

```
%Network parameters
%Maximum number of epochs to train
net.trainParam.epochs=20000;
%Performance goal
net.trainParam.goal=0.005;
%Marquardt adjustment parameter
net.trainParam.mu=0.0005;
%Decrease factor for mu
net.trainParam.mu_dec=0.1;
%Increase factor for mu
net.trainParam.mu_inc=10;
%Maximum value for mu
net.trainParam.mu_max=1.00E+10;
%Maximum validation failures
net.trainParam.max_fail=20000;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-07;
%Epochs between displays (NaN for no displays)
net.trainParam.show=50;
%Maximum time to train in seconds
net.trainParam.time=inf;
```

traincgb

%Maximum number of epochs to train
net.trainParam.epochs=20000;
%Epochs between displays (NaN for no displays)
net.trainParam.show=50;
%Show training GUI
net.trainParam.showWindow=TRUE;
%Performance goal
net.trainParam.goal=0.005;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-10;
%Maximum validation failures
net.trainParam.max_fail=20000;
%Name of line search routine to use
net.trainParam.searchFcn='srchcha';
%Divide into delta to determine tolerance for linear search.
net.trainParam.scal_tol=20;
%Scale factor that determines sufficient reduction in perf
net.trainParam.alpha=0.001;
%Scale factor that determines sufficiently large step size
net.trainParam.beta=0.1;
%Initial step size in interval location step

net.trainParam.delta=0.01;
%Parameter to avoid small reductions in performance, usually set to 0.1 (see srch_cha)
net.trainParam.gama=0.1;
%Lower limit on change in step size
net.trainParam.low_lim=0.1;
%Upper limit on change in step size
net.trainParam.up_lim=0.5;
%Maximum step length
net.trainParam.maxstep=100;
%Minimum step length
net.trainParam.minstep=1E-0006;
%Maximum step size
net.trainParam.bmax=26;

traincgf

%Maximum number of epochs to train
net.trainParam.epochs=20000;
%Epochs between displays (NaN for no displays)
net.trainParam.show=1;
%Performance goal
net.trainParam.goal=0.001;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-10;
%Maximum validation failures
net.trainParam.max_fail=20000;
%Name of line search routine to use
net.trainParam.searchFcn='srchcha';
%Divide into delta to determine tolerance for linear search.
net.trainParam.scal_tol 20=;
%Scale factor that determines sufficient reduction in perf
net.trainParam.alpha=0.0001;
%Scale factor that determines sufficiently large step size
net.trainParam.beta=0.1;
%Initial step size in interval location step
net.trainParam.delta=0.001;
%Parameter to avoid small reductions in performance, usually set to 0.1 (see srch_cha)
net.trainParam.gama=0.1;
%Lower limit on change in step size
net.trainParam.low_lim=0.1;
%Upper limit on change in step size
net.trainParam.up_lim=0.5;
%Maximum step length
net.trainParam.maxstep=100;
%Minimum step length
net.trainParam.minstep=1E-006;
%Maximum step size
net.trainParam.bmax=26;

traincgp

%Maximum number of epochs to train
net.trainParam.epochs=20000;
%Epochs between displays (NaN for no displays)
net.trainParam.show=1;
%Performance goal
net.trainParam.goal=0.001;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Minimum performance gradient
net.trainParam.min_grad=1E-10;
%Maximum validation failures
net.trainParam.max_fail=20000;
%Name of line search routine to use
net.trainParam.searchFcn='srchcha';
%Divide into delta to determine tolerance for linear search.
net.trainParam.scal_tol=20;
%Scale factor that determines sufficient reduction in perf
net.trainParam.alpha=0.001;
%Scale factor that determines sufficiently large step size
net.trainParam.beta=0.1;
%Initial step size in interval location step

net.trainParam.delta=0.01;
%Parameter to avoid small reductions in performance, usually set to 0.1 (see srch_cha)
net.trainParam.gama=0.1;
%Lower limit on change in step size
net.trainParam.low_lim=0.1;
%Upper limit on change in step size
net.trainParam.up_lim=0.5;
%Maximum step length
net.trainParam.maxstep=100;
%Minimum step length
net.trainParam.minstep=1E-0006;
%Maximum step size
net.trainParam.bmax=26;

traingd

```
%Maximum number of epochs to train
net.trainParam.epochs=20000;
%Performance goal
net.trainParam.goal=0.001;
%Learning rate
net.trainParam.lr=1.4;
%Maximum validation failures
net.trainParam.max_fail=20000;
%Minimum performance gradient
net.trainParam.min_grad=1E-07;
%Epochs between displays (NaN for no displays)
net.trainParam.show=1;
%Maximum time to train in seconds
net.trainParam.time=inf;
```

traingdm

```
%Maximum number of epochs to train
net.trainParam.epochs=1000;
%Performance goal
net.trainParam.goal=0;
%Learning rate
net.trainParam.lr=0.01;
%Maximum validation failures
net.trainParam.max_fail=1000;
%Momentum constant
net.trainParam.mc=0.9;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-05;
%Epochs between showing progress
net.trainParam.show=25;
%Generate command-line output
net.trainParam.showCommandLine=FAL
SE;
%Show training GUI
```

net.trainParam.showWindow=TRUE;

%Maximum time to train in seconds

net.trainParam.time=inf;

traingdx

%Maximum number of epochs to train

net.trainParam.epochs=1000;

%Performance goal

net.trainParam.goal=0;

%Learning rate

net.trainParam.lr=0.01;

%Ratio to increase learning rate

net.trainParam.lr_inc=1.05;

%Ratio to decrease learning rate

net.trainParam.lr_dec=0.7;

%Maximum validation failures

net.trainParam.max_fail=1000;

%Maximum performance increase

net.trainParam.max_perf_inc=1.04;

%Momentum constant

net.trainParam.mc=0.9;

%Minimum performance gradient

net.trainParam.min_grad=1.00E-05;

%Epochs between displays (NaN for no displays)

net.trainParam.show=25;

%Generate command-line output

net.trainParam.showCommandLine=FALSE;

%Show training GUI

net.trainParam.showWindow=TRUE;

%Maximum time to train in seconds

net.trainParam.time=inf;

trainlm

%Maximum number of epochs to train

net.trainParam.epochs 1000

%Performance goal

net.trainParam.goal 0

%Maximum validation failures

net.trainParam.max_fail 6

%Minimum performance gradient

net.trainParam.min_grad 1.00E-07

%Initial mu

net.trainParam.mu 0.001

%mu decrease factor

net.trainParam.mu_dec 0.1

%mu increase factor

net.trainParam.mu_inc 10

%Maximum mu

net.trainParam.mu_max 1.00E+10

%Epochs between displays (NaN for no displays)

net.trainParam.show 25

%Generate command-line output

net.trainParam.showCommandLine FALSE

%Show training GUI

net.trainParam.showWindow TRUE

%Maximum time to train in seconds

net.trainParam.time inf

trainoss

%Maximum number of epochs to train
net.trainParam.epochs=1000;
%Performance goal
net.trainParam.goal=0;
%Maximum validation failures
net.trainParam.max_fail=6;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-10;
%Name of line search routine to use
net.trainParam.searchFcn='srchbac';
%Epochs between displays (NaN for no displays)
net.trainParam.show=25;
%Generate command-line output
net.trainParam.showCommandLine=FALSE;
%Show training GUI
net.trainParam.showWindow=TRUE;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Divide into delta to determine tolerance for linear search.
net.trainParam.scal_tol=20;
%Scale factor that determines sufficient reduction in perf
net.trainParam.alpha=0.001;
%Scale factor that determines sufficiently large step size
net.trainParam.beta=0.1;
%Initial step size in interval location step
net.trainParam.delta=0.01;
%Parameter to avoid small reductions in performance, usually set to 0.1 (see srch_cha)
net.trainParam.gama=0.1;
%Lower limit on change in step size
net.trainParam.low_lim=0.1;
%Upper limit on change in step size
net.trainParam.up_lim=0.5;
%Maximum step length
net.trainParam.maxstep=100;
%Minimum step length
net.trainParam.minstep=1.00E-06;
%Maximum step size
net.trainParam.bmax=26;

trainr

```
%Maximum number of epochs to train
net.trainParam.epochs=1000;
%Performance goal
net.trainParam.goal=0;
%Maximum validation failures
net.trainParam.max_fail=1000;
%Epochs between displays (NaN for no displays)
net.trainParam.show=25;
%Generate command-line output
net.trainParam.showCommandLine=FALSE;
%Show training GUI
net.trainParam.showWindow=TRUE;
%Maximum time to train in seconds
net.trainParam.time=inf;
```

trainrp

```
%Maximum number of epochs to train
net.trainParam.epochs=1000;
%Epochs between displays (NaN for no displays)
net.trainParam.show=25;
%Generate command-line output
net.trainParam.showCommandLine=FALSE;
%Show training GUI
net.trainParam.showWindow=TRUE;
%Performance goal
net.trainParam.goal=0;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-05;
%Maximum validation failures
net.trainParam.max_fail=1000;
%Learning ratenet.trainParam.lr=0.01;
%Increment to weight change
net.trainParam.delt_inc=1.2;
%Decrement to weight change
net.trainParam.delt_dec=0.5;
%Initial weight change
net.trainParam.delta0=0.07;
%Maximum weight change
net.trainParam.deltamax=50;
```

trainscg

```
%Maximum number of epochs to train
net.trainParam.epochs=1000;
%Epochs between displays (NaN for no displays)
net.trainParam.show=25;
%Generate command-line output
net.trainParam.showCommandLine=FALSE;
%Show training GUI
net.trainParam.showWindow=TRUE;
%Performance goal
net.trainParam.goal=0;
%Maximum time to train in seconds
net.trainParam.time=inf;
%Minimum performance gradient
net.trainParam.min_grad=1.00E-06;
%Maximum validation failures
net.trainParam.max_fail=1000;
%Determine change in weight for second derivative approximation
net.trainParam.sigma=5.00E-05;
%Parameter for regulating the indefiniteness of the Hessian
net.trainParam.lambda=5.00E-07;
```