LATERAL PRESSURES ON RIGID RETAINING WALLS : A NEURAL NETWORK APPROACH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERSAN YILDIZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF CIVIL ENGINEERING

AUGUST 2003

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Mustafa Tokyay
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. M. Yener Özkan
Supervisor

Examining Committee Members

Prof. Dr. Erdal Çokça

Prof. Dr. Semra S. Uluatam

Prof. Dr. M. Yener Özkan

Prof. Dr. Özdemir Akyılmaz

Dr. Oğuz Çalışan

ABSTRACT

LATERAL PRESSURES ON RIGID RETAINING WALLS :
A NEURAL NETWORK APPROACH

Yıldız, Ersan
M.S., Department of Civil Engineering
Supervisor : Prof. Dr. M. Yener Özkan

August 2003, 101 pages

Lateral pressures on non-yielding walls due to surface strip loads were investigated considering the non-linear stress-strain behaviour of the soil by finite element analyses. Data obtained from the finite element analyses were used to train neural networks in order to obtain a solution to assess the total lateral thrust and its point of application on a non-yielding wall due to a strip load. A 2-layered backpropogation type neural network was used. An artificial neural network solution was obtained, as a function of six parameters including the shear strength parameters of the soil (cohesion and angle of friction). The effects of each input parameter on the lateral thrust and point of application were summarized and the results were compared with the conventional linear elastic solution.

Keywords : Lateral pressure, Strip load, Neural networks, Hardening soil model

ÖZ

RİJİT DUVARLARA UYGULANAN YANAL BASINÇLAR:
YAPAY SİNİR AĞI YAKLAŞIMI

Yıldız, Ersan
Yüksek Lisans, İnşaat Mühendisliği Bölümü
Tez Yöneticisi : Prof. Dr. M. Yener Özkan

Ağustos 2003, 101 sayfa

Yüzeysel şerit yüklerin rijit istinat duvarları üzerinde uyguladıkları yanal basınçlar, zeminin doğrusal olmayan gerilme-deformasyon özellikleri göz önüne alınarak incelenmiştir. Sonlu elemanlar yöntemi ile bulunan veriler, yüzeysel şerit yükten dolayı bir rijit istinat yapısına gelen toplam yanal kuvvetin ve bu kuvvetin tatbik noktasının tahmini amacıyla yapay sinir ağlarının eğitilmesinde kullanılmıştır. Bu amaçla 2 tabakalı geri-yayılımlı tipte (back-propogation type) sinir ağından yararlanılmıştır. Duvara gelen toplam yük ve tatbik noktasının hesaplanması için kullanılabilecek ve 6 girdinin bir fonksiyonu olan, yapay sinir ağı çözümü elde edilmiştir. Her bir girdi parametresinin toplam yanal kuvvet ve tatbik mesafesine olan etkileri belirtilmiş ve sonuçlar dogrusal elastik çözümle karşılaştırılmıştır.

Anahtar kelimeler : Yanal basınç, Şerit yük, Yapay sinir ağı

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In many applications of soil engineering, lateral pressures acting on non-yielding retaining walls due to surface strip loads behind the wall are required.

Retaining structures supporting continuous wall footings, highways, railroads and crane loads are typical examples for the surface strip loading.

In the case of a rigid retaining wall, linear elastic solutions based on Bousinessq's equations(Misra 1981, Jarquio 1981) are frequently used for the determination of the lateral thrust due to strip loads. However, the theory of elasticity does not consider the strength and the variation of the stiffness of the soil with different stress states.

The objective of this study is to investigate the lateral earth pressures acting on rigid retaining walls due to surface strip loading only by modelling the soil as an elastoplastic material having non-linear stress-strain relationship. For this purpose, a finite element program " PLAXIS ver. 7.11" is used for the calculations.

It is proposed to obtain a closed-form solution for the calculation of the total lateral thrust and its point of application considering the non-linear stress-strain relationship of the soil, as an alternative to the linear elastic solution.

The closed form solution is proposed to be obtained by using a new tool, artificial neural networks, based on the results of various finite element analyses. “MATLAB ver. 6.0 Neural Network Toolbox” is used to establish and train the neural networks in this study.

CHAPTER 2

LATERAL EARTH PRESSURE

2.1 LATERAL EARTH PRESSURE

Lateral earth pressure is the lateral force exerted by the soil to an adjoining retaining structure. It is dependent on the soil structure and the interaction of soil with the retaining structure.

The classical solutions of lateral earth pressure are Coulomb's (1773) and Rankine's (1857) earth pressure theories. These fundamental solutions still form the basis of earth pressure calculations today. All earth pressure theories now available have their roots in Coulomb and Rankine's work. (Coduto, 2001)

2.1.1 RANKINE'S THEORY OF EARTH PRESSURE

Rankine's theory considers the state of stress in a soil mass when the condition of plastic equilibirium has been reached. The Mohr circle representing the state of stress at failure in a two dimensional element is shown in Fig. 2.1. Shear failure occurs along a plane at an angle of ($45 + \phi/2$) to the major principal plane. If the whole soil mass is stressed that the principal stresses are in the same directions for every point, there will be a network of failure planes as shown in Fig. 2.1. It should be noted that sufficient deformation is required for the development of the plastic equilibrium.

A semi-infinite mass of soil with a horizontal surface and having a vertical boundary formed by a smooth wall surface extending to semi-infinite depth is considered as shown in Fig. 2.2. The soil is assumed to be homogenous and isotropic. A soil element at a depth z is subjected to a vertical and horizontal stress and since there can be no lateral transfer of weight if the surface is horizontal, no shear stresses exist on horizontal and vertical planes. Therefore the vertical and horizontal stresses are principal stresses.

Active Case :

Suppose that such a soil deposit is stretched in the horizontal direction by a movement of the wall away from the soil. The value of σ_x decreases due to expansion outwards; if the expansion is large enough, the value of σ_x decreases to a minimum value such that a state of plastic equilibrium develops. Therefore σ_x and σ_y are minor and major effective stresses (σ_3 and σ_1) respectively.

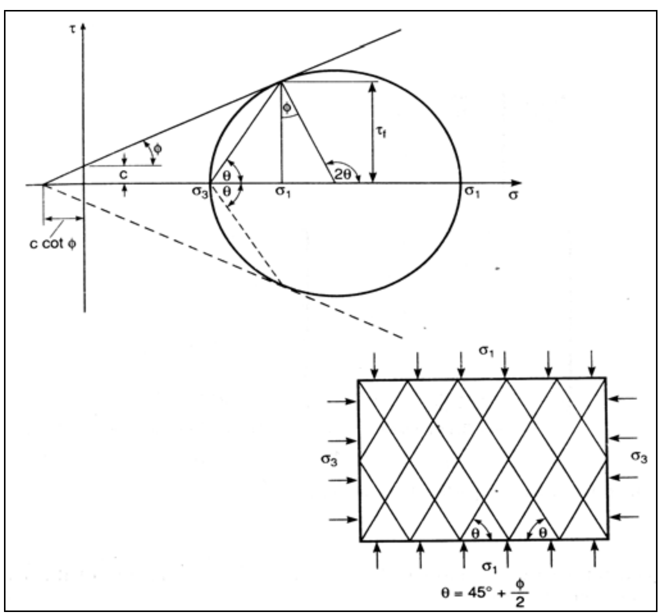


Figure 2.1 State of plastic equilibrium (Craig 1992)

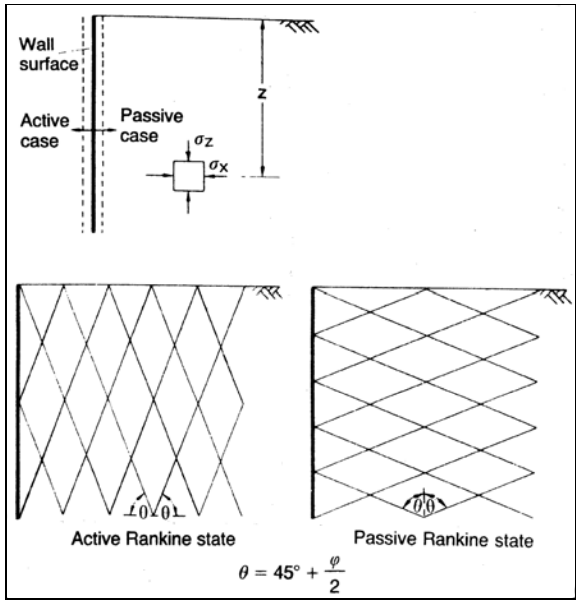


Figure 2.2 Active and passive Rankine states (Craig 1992)

The value of σ_3 (= σ_x) is determined when a Mohr circle through the point representing σ_1 touches the failure envelope. The relationship between σ_1 and σ_3 when the soil reaches a plastic equilibirum state can be derived from this Mohr circle as:

$$\sigma_3 = \sigma_1\left(\frac{1-\sin\phi}{1+\sin\phi}\right) - 2c\sqrt{\frac{1-\sin\phi}{1+\sin\phi}} \qquad (2.1)$$

The horizontal stress for the above condition is defined as the *active pressure* (p_a). If $K_a = \dfrac{1-\sin\phi}{1+\sin\phi}$ is defined as the active pressure coefficient, then Equation 2.1 can be written as:

$$p_a = K_a \gamma z - 2c\sqrt{K_a} \qquad (2.2)$$

When the horizontal stress becomes equal to the active pressure the soil is said to be in the *active Rankine state.*

Passive Case :

If the wall is moved against the soil mass the soil will be compressed in the horizontal direction and the value of σ_x will increase until a state of plastic equilibirum is reached. Therefore σ_x and σ_y are major and minor effective stresses (σ_1 and σ_3) respectively.

The maximum value of σ_1 is reached when the Mohr circle through the point representing σ_3 touches the failure envelope. In this case the horizontal stress is defined as the *passive pressure* (p_p). Rearranging Equation 2.1 :

$$\sigma_1 = \sigma_3\left(\frac{1+\sin\phi}{1+\sin\phi}\right) + 2c\sqrt{\frac{1+\sin\phi}{1-\sin\phi}} \qquad (2.3)$$

If $K_p = \dfrac{1+\sin\phi}{1-\sin\phi}$ is defined as the passive pressure coefficient Equation 2.3 can be written as:

$$p_p = K_p\gamma z + 2c\sqrt{K_p} \qquad (2.4)$$

When the horizontal stress becomes equal to the passive pressure the soil is said to be in the *passive Rankine state.*

If a uniformly distributed surcharge pressure of q per unit area acts over the entire surface of the soil mass, the vertical stress at any depth is increased to ($\gamma z + q$), resulting in an additional earth pressure of $K_a q$ in the active state or $K_p q$ in the passive state.

2.1.2 COULOMB'S THEORY OF EARTH PRESSURE

Coulomb's theory considers the stability of a wedge of soil between a retaining wall and a trial failure plane. The force between the wedge and the wall surface is determined by considering the equilibrium of forces acting on the wedge when it is on the point of sliding up or down the failure plane.

Friction between the wall and the soil is taken into account. The angle of friction between the wall and soil material, denoted by δ and a constant component of shear resistance or wall adhesion, c_w are considered.

The shape of the failure surface is curved near the bottom of the wall due to wall friction in both active and passive cases (Fig. 2.3), but in the Coulomb theory the failure surface is assumed to be plane in each case. The error due to this assumption is relatively small for the active case and for the passive case for values of δ less than $\phi/3$, but for the higher values of δ the error becomes relatively large. (Craig, 1992)

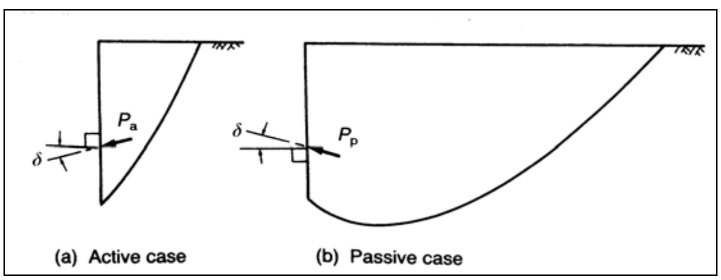


Figure 2.3 Curvature due to wall friction (Craig 1992)

Active Case :

The forces acting on the soil wedge between a wall surface and a trial failure plane are shown in Fig. 2.4 The cohesion parameter c is taken as zero. For the failure condition the soil wedge is in equilibrium under its own weight (W), the reaction force between the soil and the wall (P) and the reaction on the failure plane (R). Because the wedge tends to move down the plane at failure, *P* acts at an angle δ below the normal to the wall. At failure, when the shear strength of the soil has been fully mobilized, the direction of *R* is at an angle ϕ below the normal to the failure plane. The directions of the three forces and the magnitude of *W* are known so that the magnitude of *P* can be determinded from the triangle of forces (Fig. 2.4).

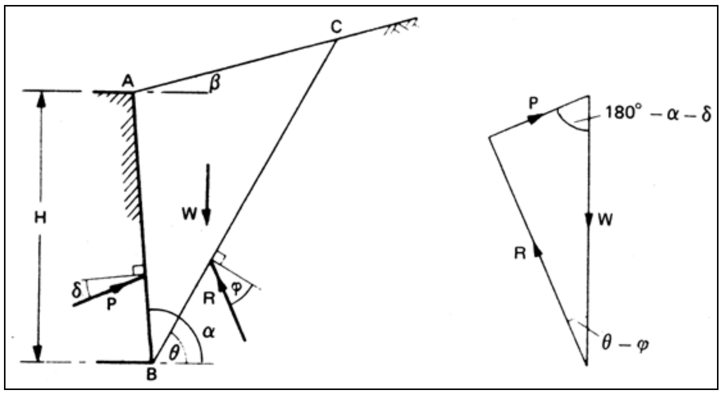


Figure 2.4 Coulomb theory, active case with c = 0 (Craig 1992)

Different failure planes have to be selected to obtain the maximum value of *P*. However the maximum value of *P* can be solved by expressing *P* in terms of *W* and the angles and differentiating with respect to θ, $\partial P/\partial\theta = 0$ as :

$$P_a = \frac{1}{2} K_a \gamma H^2 \qquad (2.5)$$

where

$$K_a = \left[\frac{\sin(\alpha - \phi)/\sin\alpha}{\sqrt{\sin(\alpha+\delta)} + \sqrt{\dfrac{\sin(\phi+\delta)\sin(\phi-\beta)}{\sin(\alpha-\beta)}}} \right]^2 \qquad (2.6)$$

The point of application of the total active thrust is assumed to act at a distance of *H/3* above the base of the wall.

The Coulomb theory can be extended to soils that have a cohesion parameter c greater than zero, a value is then selected for the wall parameter c_w. It is assumed that tension cracks may extend to a depth z_0 , the trial failure plane extending from the heel of the wall to the bottom of the tension zone, as shown in Fig. 2.5. The forces acting on the wedge are the weight of the wedge (W), the reaction between the wall and soil (P), the force due to constant component of shearing resistance on wall ($C_w = c_w$ x EB), the reaction on the failure plane (R), the force on the failure plane due to constant component of shear strength (C = c x BC).

The directions of all five forces are known together with the magnitudes of W, C_w and C so that the value of P can be determined from the force diagram. Again different trial failure planes have to be selected to obtain the maximum value of P. (Craig, 1992)

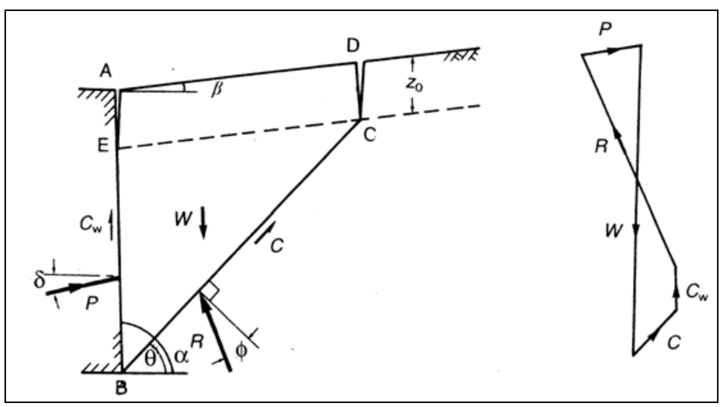


Figure 2.5 Coulomb theory, active case with c > 0 (Craig 1992)

Passive Case :

In the passive case the reaction P acts at an angle δ above the normal to the wall surface and the reaction R at an angle ϕ above the normal to the failure plane. In the triangle of forces the angle between the *W* and *P* is (180 - $\alpha + \delta$) and the angle between *W* and *R* is ($\theta + \phi$). The total passive resistance, equal to the minimum value of *P*, is given by :

$$P_p = \frac{1}{2} K_p \gamma H^2 \tag{2.7}$$

where

$$K_p = \left[\frac{\sin(\alpha + \phi) / \sin\alpha}{\sqrt{\sin(\alpha - \delta)} + \sqrt{\dfrac{\sin(\phi + \delta)\sin(\phi + \beta)}{\sin(\alpha - \beta)}}} \right]^2 \tag{2.8}$$

2.1.3 EARTH PRESSURE AT REST

If lateral strain in the soil is zero the corresponding lateral pressure is called the *earth pressure at-rest* and is expressed as :

$$p_0 = K_0 \gamma' z \tag{2.9}$$

where K_0 is defined as the coefficient of earth pressure at-rest in terms of effective stress.

At-rest condition does not involve the failure of the soil, K_0 can be determined from laboratory or in-situ tests and emprical correlations.

Generally, for any condition intermediate to active and passive states the value of the lateral stress is unknown. The range of possible conditions can only be determined experimentally. Fig. 2.6 shows the relationship between strain and lateral pressure coefficient.The exact relationship depends on the initial value of K_0 and on whether excavation or backfilling is involved in the construction. (Craig, 1992).

For normally consolidated soils the value of K_0 can be found by the formula proposed by Jaky (1944):

$$K_0 = 1 - \sin\phi' \qquad (2.10)$$

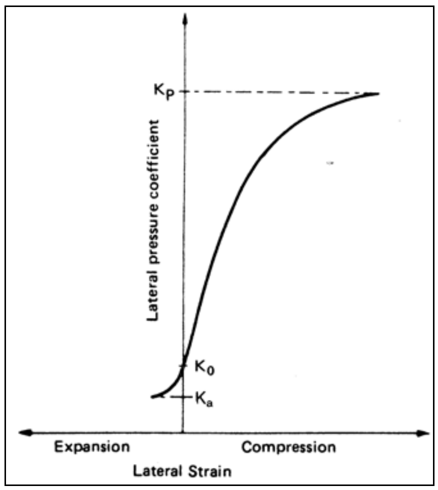


Figure 2.6 Relationship of lateral strain and lateral pressure coefficient (Craig 1992)

2.2 LATERAL PRESSURE DUE TO SURFACE STRIP LOAD

In many earth retaining problems, it is necessary to consider additional earth pressures produced by strip loads acting on the soil surface behind the wall. Retaining structures supporting continuous wall footing, highway and railroad loadings are practical examples in which the strip load type of surcharge is applicable.

Some methods used to calculate the lateral earth pressures on retaining structures due to surface strip load are summarized as :

2.2.1 LINEAR ELASTIC SOLUTION

The stresses within a semi-infinite, homogenous, isotropic mass, with a linear stress-strain relationship, due to a point load on the surface, were determined by Bousinessq in 1885. The stresses due to surface loads distributed over a particular area can be obtained by integration from the point load solutions.

Assuming that the soil behaves as a linear elastic material of constant modulus of elasticity, the lateral earth pressure caused by strip surface load is (Jarquio, 1981) :

$$\sigma_h = 2q/\pi \,.(\beta - \sin\beta\cos2\alpha) \qquad (2.11)$$

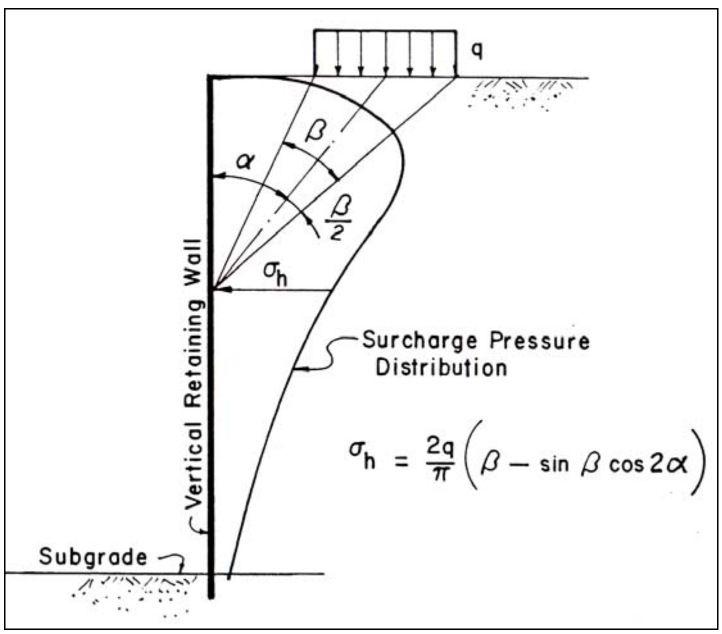


Figure 2.7 Linear elastic solution (Jarquio 1981)

2.2.2 GENERALIZED COULOMB ACTIVE EARTH PRESSURE FOR DISTANCED SURCHARGE

This method is suggested by Motta(1994). The basic assumptions of this method are that the soil is homogenous, dry and cohesionless; the failure surface of the wedge is a plane; and the extesion of the uniformly distributed load q is large enough that it is intersected by the failure plane as shown in Fig. 2.8.

Based on the notations on Figures 2.8 and 2.9, cricital angle (α_c) for the failure plane, the angle that gives the maximum-earth pressure can be found by the following expression :

$$\tan(\alpha_c - i) = \frac{\sin a \ \sin b + \left(\sin^2 a \ \sin^2 b + \sin a \ \cos a \ \sin b \ \cos b + A \cos c \ \cos a \ \sin b\right)^{\frac{1}{2}}}{A \cos c + \sin a \ \cos b} \quad (2.12)$$

where $a = \phi' + \delta - i$; $b = \phi' - i - \theta$; $c = \theta + \delta$ and

$A = [(1+n_q)\sin(i)\cos(i) + \lambda n_q] / [(1 + n_q)\cos^2(i)$

$n_q = 2q / \gamma H$; $\lambda = d / H$

The active earth-pressure coefficient for the self-weight of the soil and the surcharge can be found as follows:

$$K_{a,\gamma q} = \frac{(1+n_q)\cos^2 i[1 - A\tan(\alpha_c - i)]\left[\cos b - \dfrac{\sin b}{\tan(\alpha_c - i)}\right]}{\cos\theta[\cos a + \tan(\alpha_c - i)\sin a]} \quad (2.13)$$

The total active earth pressure due to the self-weight and the surcharge is:

$$S_a = 1/2 * \gamma H^2 K_{a,\gamma q} \quad (2.14)$$

The above equations are valid only if $\alpha_c < \alpha_1$, that is : $\tan(\alpha_c) < \tan(i) + 1 / \lambda$

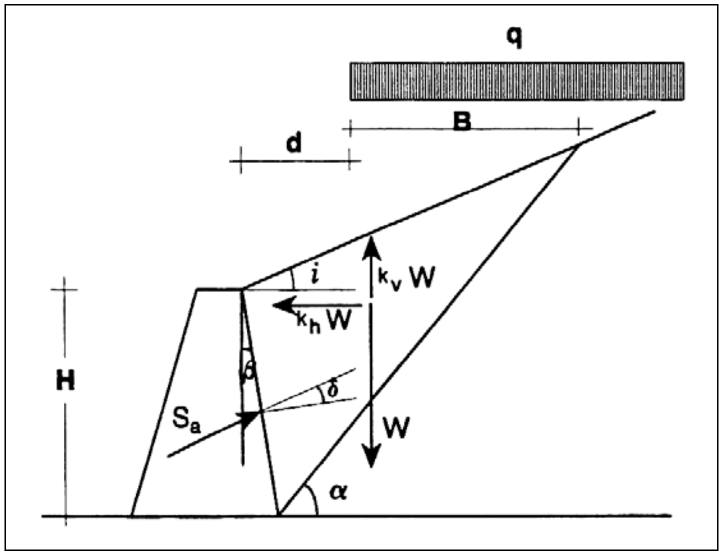


Figure 2.8 Scheme for Earth Pressure Evaluation (Motta 1994)

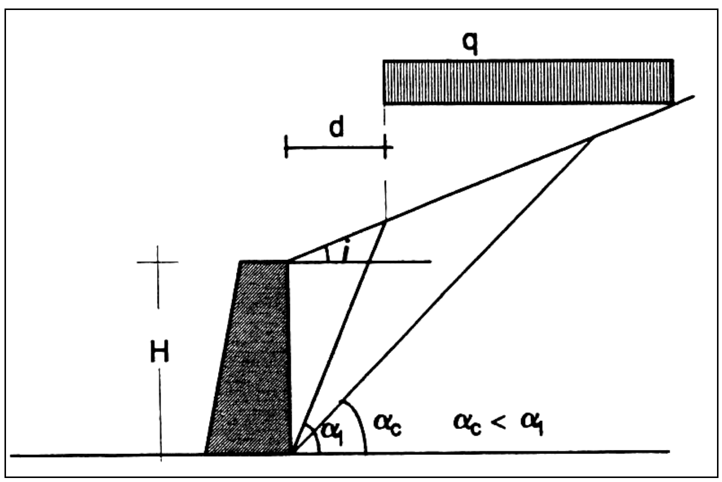


Figure 2.9 Limits for Boundary Conditions (Motta 1994)

2.2.3 BETON KALENDER APPROACH

For yielding retaining walls, Beton Kalender(1983) recommends the use of an approximate method (As cited in Reference 9). As shown in Figure 2.10 a uniform or triangular lateral surcharge pressure distribution can be obtined.

1. For uniform pressure distribution :

$$\sigma_{h1} = [q.b.\cos\delta.\sin(45-\phi/2)] / [d.\cos(45 - \phi/2 - \delta] \tag{2.15}$$

2. For triangular pressure distribution :

$$\sigma_{h11} = [q.b.\cos\delta.\sin(45-\phi/2)] / [2d.\cos(45 - \phi/2 - \delta] \tag{2.16}$$

where δ is the angle of wall friction in degrees.

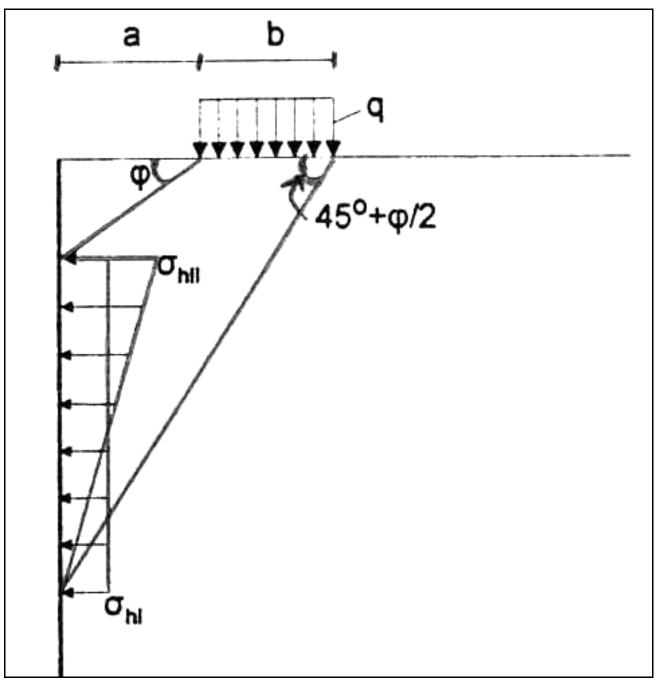


Figure 2.10 Beton Kalender Approach (Georgiadis and Anagnostopoulos 1998)

2.2.4 45? DISTRIBUTION APPROACH

This approach is suggested by Cernica(1995) for distanced surcharge loading and adopted to strip loading case by Georgiadis and Anagnostopoulos (1998). It considers that the strip load *q* is distributed at 45? angles as shown in Figure 2.11. The lateral pressure is obtained by:

$$\sigma_h = K_a \cos\delta \frac{b}{b+2a} q \qquad (2.17)$$

where K_a = active earth pressure coefficient; δ= friction angle between the soil and the wall; a= distance between the load and the wall; b= width of the strip load.

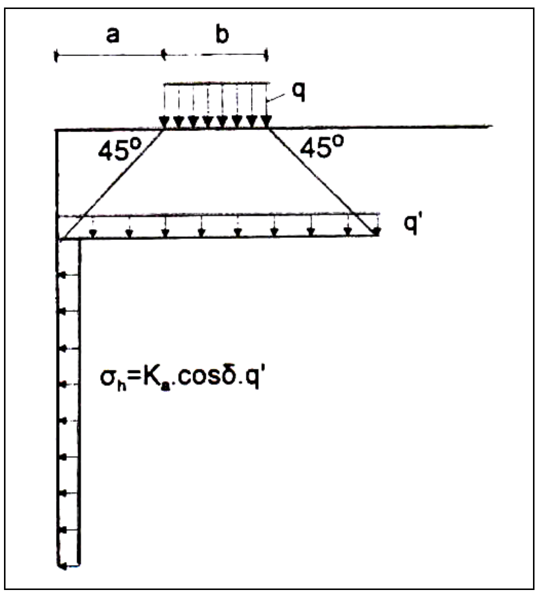


Figure 2.11 45° distribution approach (Georgiadis and Anagnostopoulos 1998)

2.2.5 ADDITIONAL INFORMATION ABOUT LATERAL PRESSSURES DUE TO STRIP LOAD

Georgiadis and Anagnostopoulos (1998) have presented an experimental investigation of the problem, in which model test measurements are compared to bending moments computed using the lateral earth pressure theories mentioned before. A cantilever sheet pile wall installed in a 1200 x 300 x 500mm-deep tank filled with fine to medium sand for the model tests.

Lateral earth pressures obtained with these methods , for a strip load q=12 kPa applied at a distance a=200 mm from a 250 mm deep excavation are presented in Figure 2.12. Total bending moments computed using combined active earth pressures and surcharge pressures are compared to measured bending moments in Figure 2.13. It is clear from the figures that the best predicitions are obtained by the Coulomb and 45^{o} distribution approaches for the model sheet pile wall. The elastic solution gives extremely large bending moments, which were up to eight times larger than the measured values. No information regarding the displacement of the model sheet pile wall is given in the report by Georgiadis and Anagnostopoulos.

Georgiadis and Anagnostopoulos have also performed finite element analyses. A rigid wall, supporting a uniform strip load q = 12kPa at a distance a = 200 mm, was analyzed considering rotation around a hinge placed at a depth of 500 mm. The soil was modelled as an elastic material and friction between the wall and the soil is neglected.

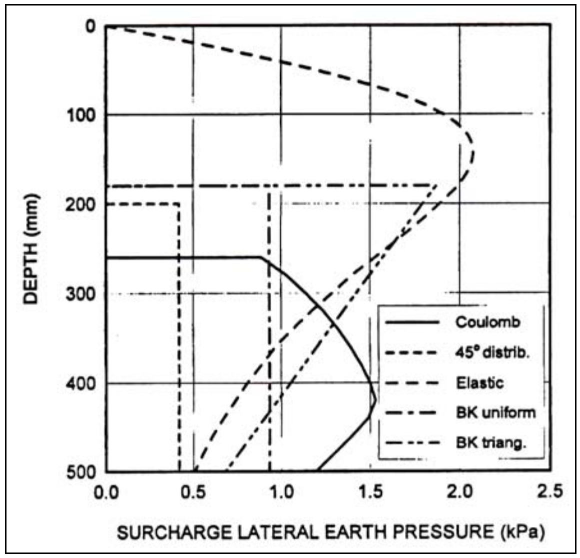


Figure 2.12 Lateral surcharge pressure distributions for various methods (Georgiadis and Anagnostopoulos 1998)

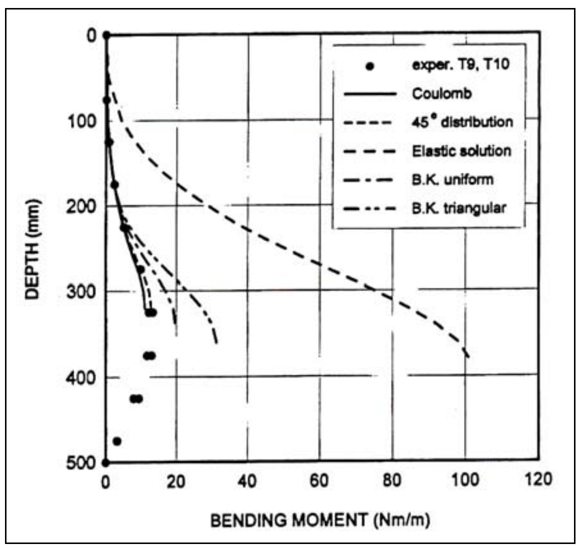


Figure 2.13 Comparison of measured and predicted bending moments (Georgiadis and Anagnostopoulos 1998)

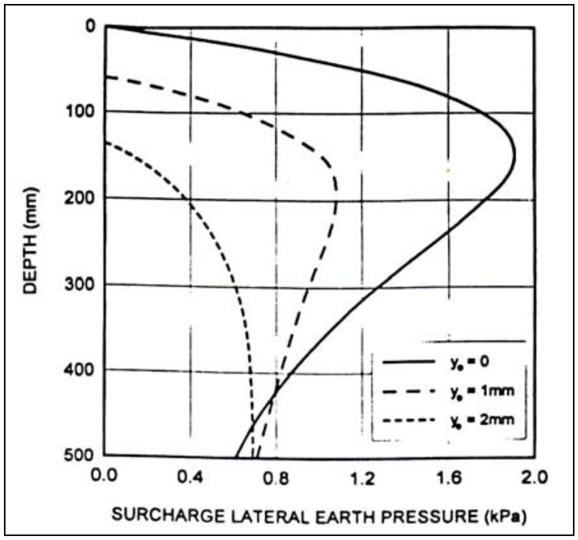


Figure 2.14 Effect of wall movement on lateral surcharge pressures

(Georgiadis and Anagnostopoulos 1998)

Fig. 2.14 shows the lateral earth pressures obtained for various wall rotations and corresponding wall movements (y_0) at ground level. Fig. 2.15 shows the computed bending moments for these lateral pressure distributions. It can be seen from the figures that, even small lateral yielding of the wall significantly reduces lateral surcharge pressures and bending moments determined by elastic theory.

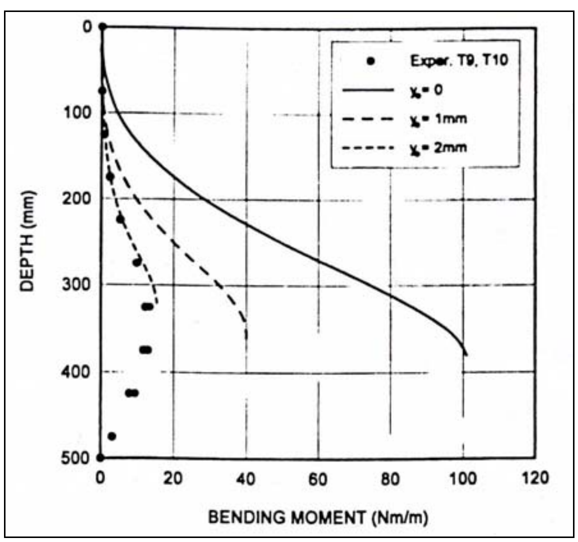


Figure 2.15 Effect of wall movement on bending moments

(Georgiadis and Anagnostopoulos 1998)

Jarquio(1981) has derived a direct solution for the total lateral surcharge pressure and for the location of the centroid of the total lateral surcharge pressure and the point of maximum unit lateral pressure based on Bousinessq's equations.

Based on notations on Figure 2.16, the total lateral surcharge pressure is :

$$P = q[h(\theta_2-\theta_1)]/90 \qquad (2.18)$$

where θ_1 and θ_2 are expressed in degrees.

Dimensions "a" and "b" should be included within the soil wedge defined by θ and x and y axes.

For active pressure condition: $\theta = (45-\phi/2)$

For passive pressure condition : $\theta = (45+\phi/2)$

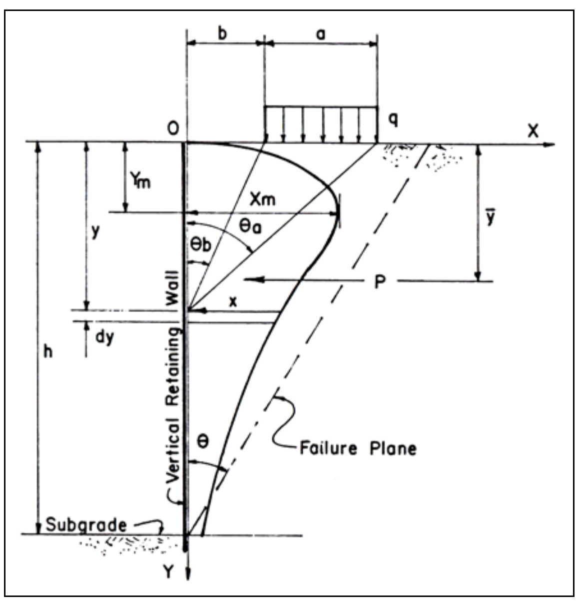


Figure 2.16 Notations for total lateral surcharge pressure (Jarquio 1981)

For further information (about the determination of the point of application) the reader is referred to Jarquio(1981).

Jarquio(1981) suggests the use of direct solutions based on Bousinessq's equations for both yielding or unyielding retaining wall structures with the soil wedge behind it either in the active or passive mode of failure depending on the given condition of the problem.

However, Steenfelt and Hansen(1983) reports that the presented formulas by Jarquio(1981) seem reasonable for only unyielding structures and suggests the use of Coulomb's earth pressure theory for yielding walls.

As shown in Fig. 2.14, wall movement has a considerable effect on the lateral pressure distributions and as shown in Fig. 2.13, and the linear elastic solution is not applicable to yielding walls (as for the cantilever sheet pile wall) that the results can be very conservative.

A non-displacing rigid retaining wall is assumed in this study; the linear elastic solution is the most convenient one for this case among the solution methods mentinoned before. So the results of the obtained neural network solution (that considers the non-linear stress-strain behaviour of the soil) will be compared to linear elastic solution.

CHAPTER 3

MATERIAL MODEL

3.1 INTRODUCTION

A realistic assesment of the behaviour of the soil should be made by the geotechnical engineers in order to carry out a meaningful analysis.Accounting for the highly complex nature of soil behaviour, some models are obtained by simplification of the real behaviour of the soil.

Linear elastic models based on Hooke's law are generally used for the analysis of a soil mass when no failure is involved. This is known as the "elasticity problems". On the other hand, theories of plasticity are used to deal with the conditions of failure of a soil mass. These are called "stability problems". Due to the simplicity in practice and historical development of mechanics of solids, the elasticity problems and the stability problems in soil mechanics are frequently treated seperately in some unrelated ways. The connection between the elasticity problems and the stability problems is known as the progressive failure problems that deal with the elastic-plastic transition from the initial linear elastic state to the ultimate state of the soil by plastic flow. The set of equations for the solutions of progressive failure problems is called the "constitutive equations of soils", which give unique relationship of stress and strain for different geotechnical materials.

Typical stress-strain relationships for soils in the triaxial tests are shown in Fig. 3.1. It can be seen in the figure, the relation of the deviatoric stress $\sigma_3 - \sigma_1$ v.s. axial strain ε_1 for a normally consolidated clay in a drained test and overconsolidated clay in an undrained test is characterized by a non-linear response curve that rises at a slower rate after reaching a certain stress level. This phenomenon is known as "*strain hardening*" which will be discussed later in this chapter. The stress-strain curves for the overconsolidated clay in a drained test and normally consolidated clay in an undrained test have a peak that occurs at a low strain level. And the material becomes weaker for strains beyond the strain corresponding to the peak stress. This phenomenon is known as "*strain softening*".

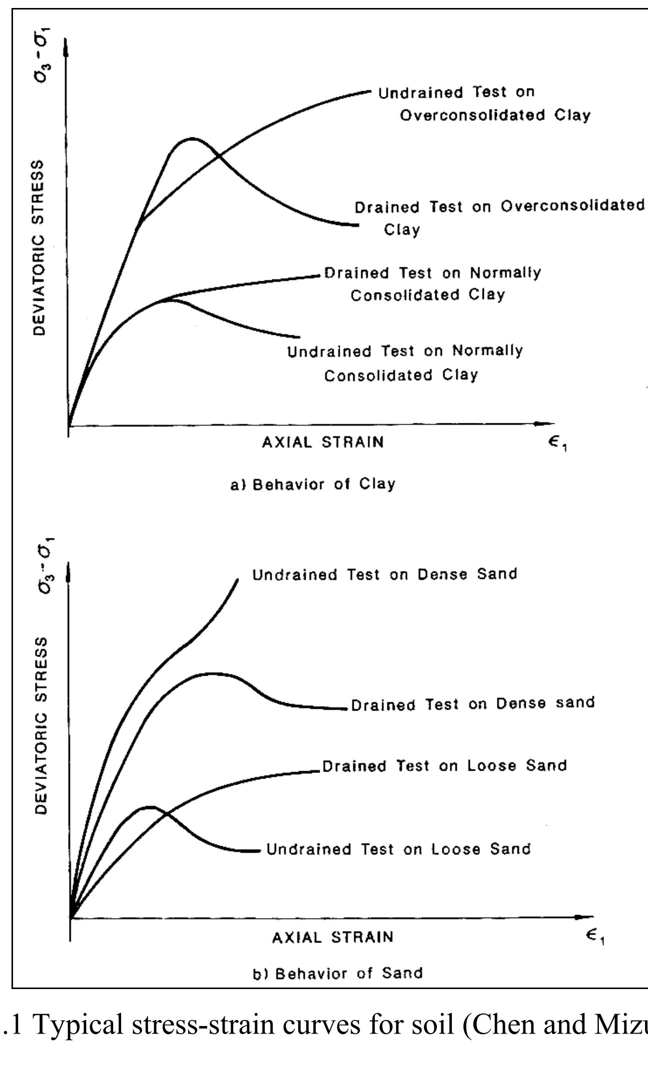


Figure 3.1 Typical stress-strain curves for soil (Chen and Mizuno 1990)

Similar conclusions can be made from Fig. 3.1 for sand. Dense sand in an undrained test and loose sand in a drained test show strain-hardening behaviour. On the other hand, dense sand in a drained test and loose sand in an undrained test show strain softening. (Chen and Mizuno 1990)

Hardening soil model (Schanz et al. 1999) which is used in this study to model the soil will be discussed in this chapter. To understand hardening soil model (which is a non-linear plastic model) better, a review of plasticity and the well known Duncan-Chang hyperbolic model is briefly outlined in this chapter.

3.2 LINEAR ELASTICITY

The linear elastic model (Generalized Hooke's Law) is the simplest model and gives a unique and linear relation between the state of stress and strain. If a linear elastic materail is stressed in the *x* direction only by a normal stress σ_x, then it experiences strains as :

$$\varepsilon_x = \sigma_x / E \qquad (3.1)$$

$$\varepsilon_y = -\upsilon\sigma_x / E \qquad (3.2)$$

$$\varepsilon_z = -\upsilon\sigma_z / E \qquad (3.3)$$

where E is the modulus of elasticity (Young's modulus) and υ is the Poisson's ratio. If a shear stress τ_{xy} is applied, the material experiences shear strain, γ_{xy}, as :

$$\gamma_{xy} = G.\tau_{xy} \qquad (3.4)$$

where G is the shear modulus and can be expressed as :

$$G = \frac{E}{2(1+\upsilon)} \qquad (3.5)$$

Elastic bulk modulus K is expressed as :

$$K = \frac{E}{3(1-2\upsilon)} \tag{3.6}$$

which relates the volumetric strain to mean normal stress as :

$$\frac{(\sigma_x + \sigma_y + \sigma_z)}{3} = K(\varepsilon_x + \varepsilon_y + \varepsilon_z) \tag{3.7}$$

As described above, only two parameters (E and υ) are sufficient to define the stress-strain relationship for soil if it is assumed to be linear elastic. The disadvantage of this model is that, the strength of the soil has no influence on the stress distributions or displacements. Also the linear elastic model lacks of the effect of the variation of the soil properties with depth or different stress states.

3.3 REVIEW ON PLASTICITY

For many materials, the stress-strain relationship in not unique (as in the linear elastic model) and many states of strain can correspond to one state of stress and vice versa. The stress-strain curve obtained from a tension test on a metal bar is shown in Fig. 3.2. The relation between stress and strain is linear for the initial portion OA. The stress-strain relation is reversible in any unloading case from any point on OA. If the bar is loaded beyond A, subsequent unloading is also reversible but non-linear. However, there is a point B beyond which unloading is not reversible. This point is called as the yield point of the material. The points A and B can often be regarded as coincident for practical purposes. If the bar is loaded to point C and unloaded, the path CD is followed, resulting in a permanent strain represented by OD. This permanent strain is called as the plastic strain. When the material is loaded to point C, the total strain is the sum of the plastic strain, OD, and elastic strain, DE. Further loading beyond C continues until the bar fails (at point F). The stress at the point F is often called as the ultimate strength.

Let's consider two identical bars tested. The first has gone through a stress cycle OCD, but the second has not. It can be seen that the first bar has a higher yield point than the second one. Then we can say that the first bar is harder than the second. The raising of the yield point is called as hardening. The term strain-hardening is used to describe this kind of behaviour. (Britto and Gunn 1987)

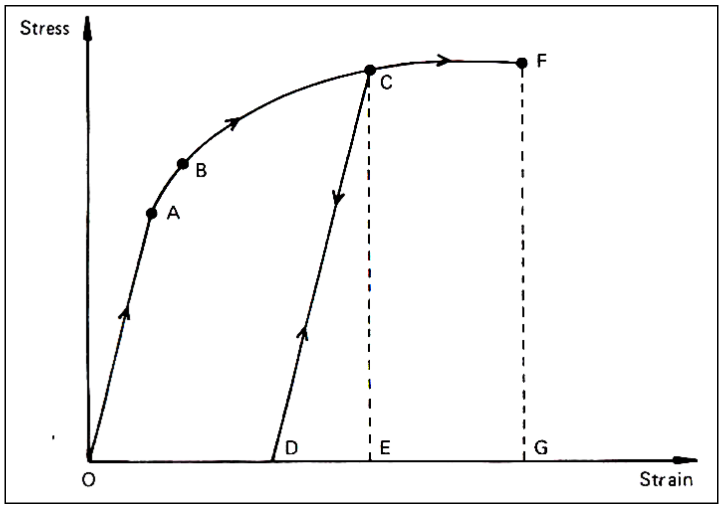


Figure 3.2 Typical stress-strain curve for metals (Britto and Gunn 1987)

To model the materials having plastic behaviour, some idealisations have to be made. In such idealisations, the main features of the behaviour are identified but aspects of secondary importance are ignored.

Fig. 3.3 shows some widely used idealisations of plastic behaviour. In elastic-perfectly plastic model, the material shows linear elastic behaviour until it yields. After yielding, the material continues to deform at constant yield stress. In elastic, strain-hardening plastic model, the stress-strain curve remains linear at a reduced slope after yielding. When only collapse loads are to be considered in a calculation, it is convenient to use rigid-plastic models in which no elastic strain exists.

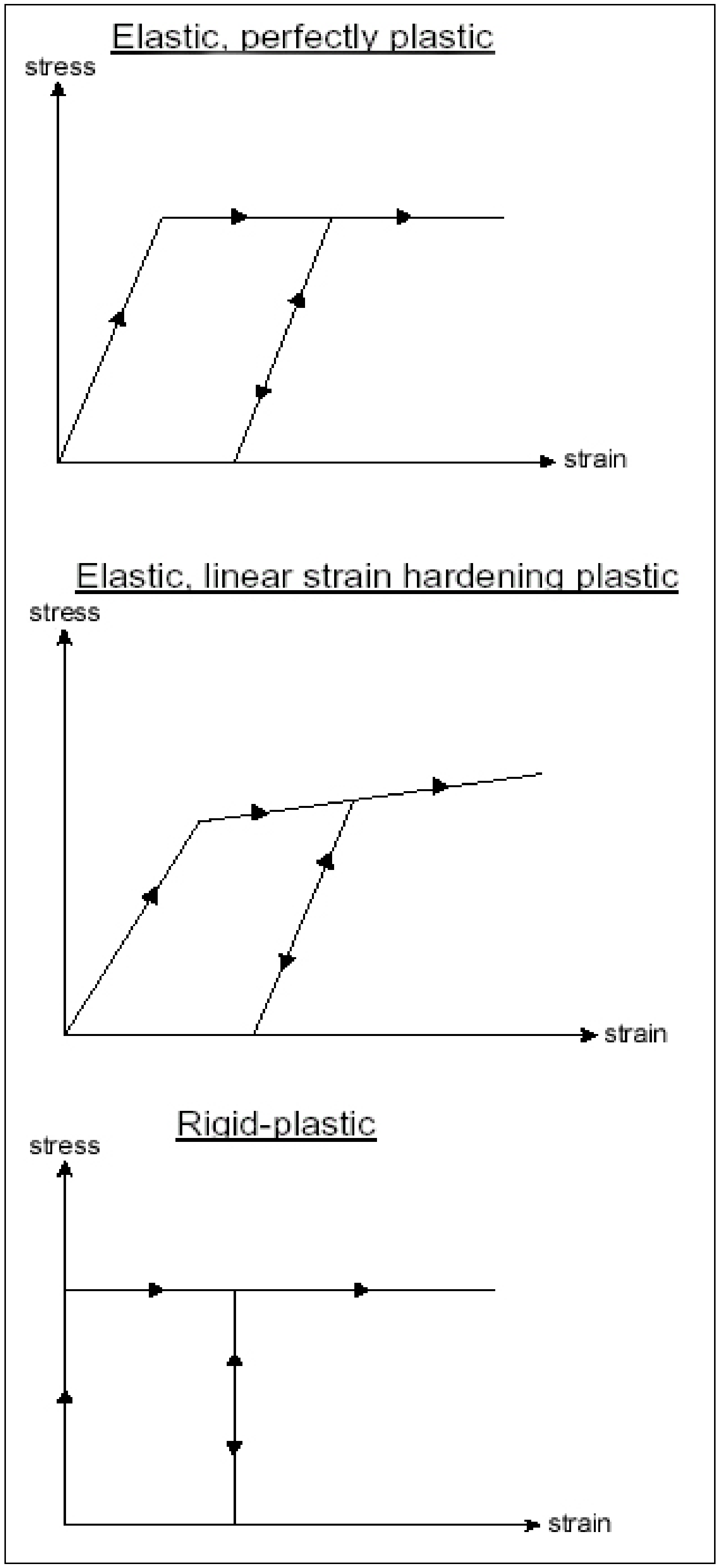


Figure 3.3 Idealisations of plastic behaviour (Britto and Gunn 1987)

To completely describe the stress-strain relations for an elasto-plastic material, four different types of statement are required : (Britto and Gunn 1987)

1- A yield function : This generalises the concept of yield stress for one-dimensional loading to two or three dimensional stress states.

2- A relationship between the directions of the principal plastic strain increments and the principal stresses.

3- A hardening rule : This is the relationship between the amount of hardening and plastic strain when the material is yielding. Thus the hardening rule determines the changes in the yield surface.

4- A flow rule : This specifies the relative magnitudes of the incremental plastic strains when the material is yielding.

3.3.1 YIELD FUNCTION

If a material is subjected to two or three dimensional states of stress, the state of the material (elastic or plastic) depend on all the stress components (six in the fully three dimensional case). If the material is isotropic, then it's sufficient to consider only the principal stresses (σ_a , σ_b and σ_c), and generally the yield functions are expressed in terms of them.

In general a yield function is written as:

$$f(\sigma_{a,} \sigma_b, \sigma_c) = 0,$$

this equation representing a surface in three-dimensional stress space. Generally yield function is written in such a way that, the negative value of the function for the current stress state indicates that the behaviour is elastic (inside the yield surface). A zero value of the function indicates that yielding takes place and positive values are not allowed.

3.3.2 A PARTICULAR YIELD FUNCTION : MOHR-COULOMB FAILURE CRITERION

The Mohr-Coulomb failure criterion by Coulomb (1773) is the most commonly used yield function for soils. According to this criterion :

$$\tau_f = c + \sigma \tan(\phi) \tag{3.8}$$

where τ_f is the shear strength (maximum shear stress) at a point on any plane within a soil mass and σ is the normal stress at the same point on the same plane. c and ϕ are the shear strength parameters described as the cohesion and the internal angle of friction respectively. Generally it is preferred to write this equation in terms of effective stresses :

$$\tau_f = c' + \sigma' \tan(\phi') \tag{3.9}$$

This equation is generally treated by a Mohr's circle, however it can be represented in three-dimensional stress space as :

$$\sigma_1' - \sigma_3' = \sin(\phi').(\sigma_1' + \sigma_3' + 2c' \cot(\phi') \tag{3.10}$$

where σ_1' and σ_3' are the major and minor effective stresses respectively. There can be six possible permutations of the magnitudes of the three principal stresses such as: $\sigma_a > \sigma_b > \sigma_c$, $\sigma_b > \sigma_c > \sigma_a$, ... etc. Therefore six planes are generated in principal effective stress space and the Mohr-Coulomb criterion is represented by the surface of an irregular hexagonal pyramid as shown in Fig. 3.4.

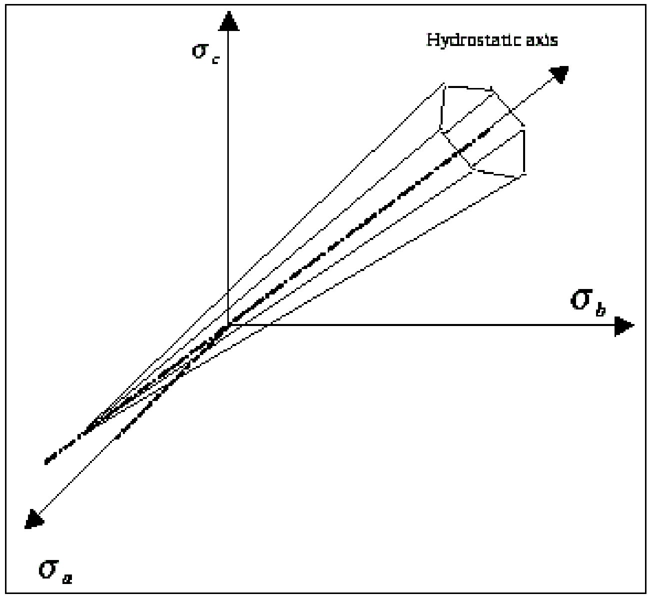


Figure 3.4 Mohr-Coulomb yield surface (Britto and Gunn 1987)

When the stress state of the material is inside the yield (failure) surface, the behaviour is elastic. The yielding occurs if the stress state is described by a point on the yield surface. Stress states outside the failure surface are impossible to attain.

The disadvantage of the use of Mohr-Coulomb criterion for soils is that, soils show evidence of volumetric yielding under isotropic stress changes where Mohr-Coulomb suggests elastic behaviour. (Britto and Gunn 1987)

3.3.3 HARDENING RULE

The hardening rule is used to define the motion (changes in size,shape and location) of the yield surface during plastic loading. Hardening rules are classified as isotropic hardening, kinematic hardening and mixed hardening. The yield surface expands uniformly in isotropic hardening, while it moves as a rigid body in stress space in kinematic hardening. (See Fig. 3.5) Mixed hardening combines both of these types of hardening and permits the yield surface to expand or contract unifomly and to translate in stress space.

If the loading is monotonic, then the isotropc hardening rule is adequate to describe the material behaviour. The kinematic hardening rule is suitable for materials under cyclic and reversed type of loadings. (Chen and Mizuno 1990)

3.3.4 FLOW RULE

The flow rule defines the ratios of plastic strain increments for a yielding material at a particular stress state. It defines only the relative sizes of individual strain increments, not their absolute sizes. The following expression is used to define the flow rule :

$$\delta\varepsilon^{p} = \delta\lambda \frac{\partial g}{\partial \sigma} \tag{3.11}$$

where $\delta\varepsilon^{p}$ is the plastic strain increment, $\delta\lambda$ is the proportionality factor and g is the plastic potential function.

The plastic potential function, $g(\sigma_a, \sigma_b, \sigma_c) = 0$ defines such a surface in principal stress space that, the plastic strain increment vectors are normal to this surface. (See Fig. 3.6) The yield function can be used as a potential function for many materials. This is called as the normality condition or associated flow rule. If a potential function different than the yield function is used, then it is called as non-associated flow.

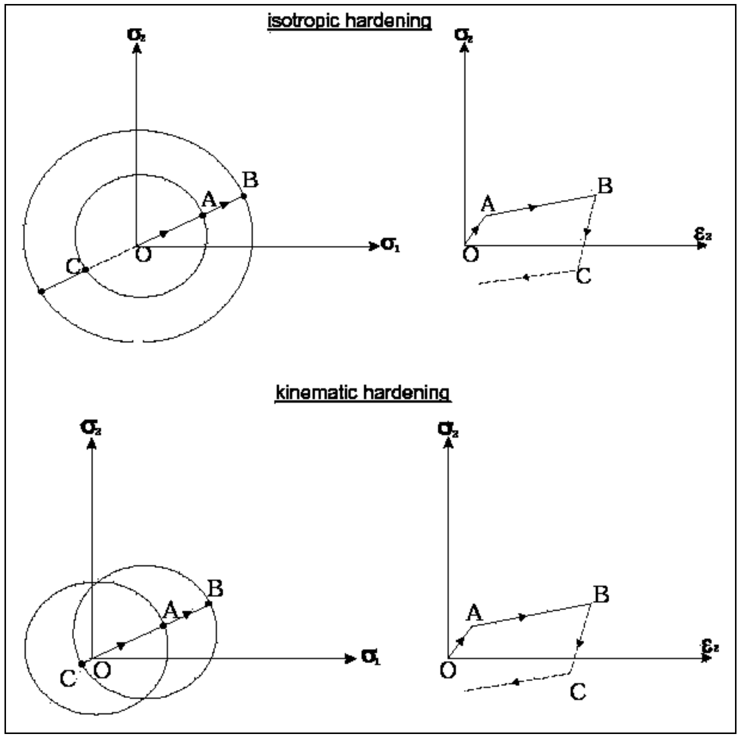


Figure 3.5 Isotropic and kinematic hardening (Britto and Gunn 1987)

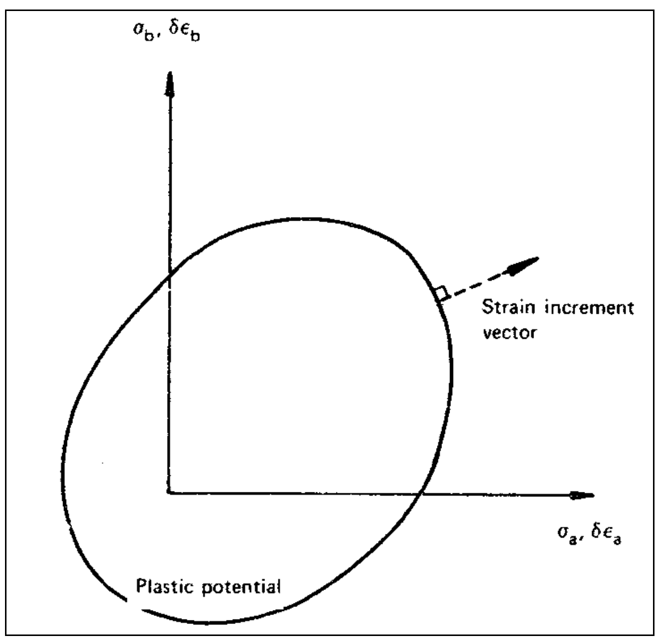


Figure 3.6 The plastic potential (Britto and Gunn 1987)

3.4 DUNCAN & CHANG HYPERBOLIC MODEL

Konder and his co-workers approximated the stress-strain curves for both clays and sands by the hyperbolic relation :

$$(\sigma_1 - \sigma_3) = \frac{\varepsilon}{\dfrac{1}{E_i} + \dfrac{\varepsilon}{(\sigma_1 - \sigma_3)_{ult}}} \tag{3.12}$$

where σ_1 and σ_3 are the major and minor principal stresses respectively, ε is the major principal strain, E_i is the initial tangent modulus and $(\sigma_1 - \sigma_3)_{ult}$ is the asymptotic value at infinite strain. This kind of hyperbola is shown in Fig. 3.7.

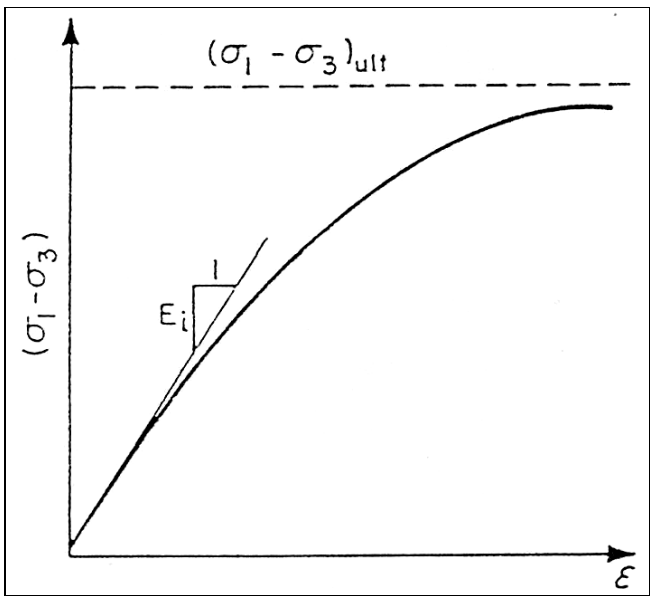


Figure 3.7 Hyperbolic representation of a stress-strain curve (Duncan et al. 1980)

Duncan and Chang (1970) suggested that E_i should be dependent on the confining pressure, σ_3, and vary in the following manner :

$$E_i = KP_a\left(\frac{\sigma_3}{P_a}\right)^n \tag{3.13}$$

where K is the modulus number and n is the modulus exponent, both of which are dimensionless numbers. P_a is the atmospheric pressure that should have the same units as σ_3. They further suggested that :

$$(\sigma_1 - \sigma_3)_f = R_f(\sigma_1 - \sigma_3)_{ult} \qquad (3.14)$$

where $(\sigma_1 - \sigma_3)_f$ is the compressive strength or principal stress difference at failure and R_f is the failure ratio. $(\sigma_1 - \sigma_3)_f$ is always smaller than $(\sigma_1 - \sigma_3)_{ult}$ and R_f changes from 0.5 to 0.9 for most soils.

The compressive strength can be calculated by Mohr-Coulomb failure criterion as :

$$(\sigma_1 - \sigma_3)_f = \frac{2(c.\cos\phi + \sigma_3.\sin\phi)}{1 - \sin\phi} \qquad (3.15)$$

where c is the cohesion intercept and ϕ is the internal angle of friction for the soil. The tangent modulus, E_t can be calculated by differentiating Eq. 3.12 with respect to ε and substituting Eqs. 3.13, 3.14 and 3.15 into the resulting expression as :

$$E_t = \left[1 - \frac{R_f(1-\sin(\phi))(\sigma_1 - \sigma_3)}{2(c.\cos(\phi) + \sigma_3.\sin(\phi))}\right]^2 .K.P_a\left(\frac{\sigma_3}{P_a}\right)^n \qquad (3.16)$$

In the case of unloading and reloading, Duncan and Chang (1970) proposed the use of unloading-reloading modulus, E_{ur}, for both cases. E_{ur} is expressed similarly to E_i as :

$$E_{ur} = K_{ur}.P_a.\left(\frac{\sigma_3}{P_a}\right)^n \qquad (3.17)$$

where K_{ur} is the unloading-reloading modulus number. The modulus exponent, n, is assumed to be the same for both unloading-relading and primary loading.

Duncan and Chang (1970) assumed the second elastic constant, Poissons's ratio, to be constant. This assumption is modified by Duncan et al. (1980) by introducing a bulk modulus (B) for the soil. The bulk modulus is assumed to be independent of stress level (σ_1 - σ_3) and vary with the confining pressure. The variation of B with σ_3 approximated similarly to variation of E_i with σ_3 as :

$$B = K_b . P_a . \left(\frac{\sigma_3}{P_a} \right)^m \tag{3.18}$$

where K_b is the bulk modulus number and m is the bulk modulus exponent both of which are dimensionless.

As a summary, the non-linear and stress dependent stress-strain characteristics of soils are defined by : (Duncan et al. 1980)

1- Tangent values of Young's modulus (E_t) which vary with confining pressure and the percentage of the strength mobilized.

2- Values of bulk modulus (B) which vary with confining pressure and are independent of the percentage of the strength mobilized.

The reader is referred to Reference 8 for more details (including determination of the hyperbolic parameters, advantages, limitations and typical values for various soil types) on the Duncan & Chang hyperbolic model.

3.5 HARDENING SOIL MODEL

Reviews of plasticity and Duncan & Chang hyperbolic model are briefly outlined in Sections 3.3 and 3.4. Reading these sections before will help the reader to understand Hardening Soil model better.

The hardening soil model by Schanz et al. (1999) is an advanced model based on the theory of plasticity for simulating the stress-strain behaviour of different types of soil.

This model is used to represent the stress-strain characteristics of the soil in the finite element analyses using PLAXIS ver. 7.11 in this study.

Isotropic hardening rule is used for the model, and distinction is made between two types of hardening : Shear hardening is used to model irreversible strains due to primary deviatoric loading. Compression hardening is used to model irreversible strains due to primary compression in oedometer loading and isotropic loading. Soil dilatancy and a yield cap is included in the model.

3.5.1 CONSTITUTIVE EQUATIONS

In the case of a primary deviatoric loading, soil shows a decreasing stiffness and irreversible plastic strains develop. In the special case of a drained triaxial test, the relationship between the axial strain and the deviatoric stress can be approximated by a hyperbola that can be described by :

$$-\varepsilon_1 = \frac{q_a}{2E_{50}} \frac{(\sigma_1 - \sigma_3)}{q_a - (\sigma_1 - \sigma_3)} \quad \text{for } q_a < q_f \qquad (3.19)$$

The ultimate deviatoric stress, q_f, is derived from the Mohr-Coulomb failure criterion as :

$$q_f = (c.\cot(\phi) - \sigma'_3)\frac{2\sin(\phi)}{1 - \sin(\phi)} \qquad (3.20)$$

where c and ϕ are the strength parameters of the soil. Note that σ'_3 is assumed to be negative for compression. The asymptotic value, q_a, is defined as :

$$q_a = q_f / R_f \tag{3.21}$$

where R_f is the failure ratio, which is always smaller than 1. $R_f = 0.9$ is often a suitable value to use for most soils. This hyperbolic relationship is shown in Fig. 3.8.

The parameter E_{50} is the confining stress dependent stiffness modulus for primary loading. It is expressed as :

$$E_{50} = E_{50}^{ref}\left(\frac{c.\cot(\phi) - \sigma_3'}{c.\cot(\phi) + p^{ref}}\right)^m \tag{3.22}$$

E_{50}^{ref} is the reference stiffness modulus corresponding to the reference stress (effective confining pressure) p^{ref}. It is determined from a triaxial stress-strain curve for the mobilization of 50 % of the maximum shear strength q_f (See Fig. 3.8). The amount of stress dependency is given by the power m.

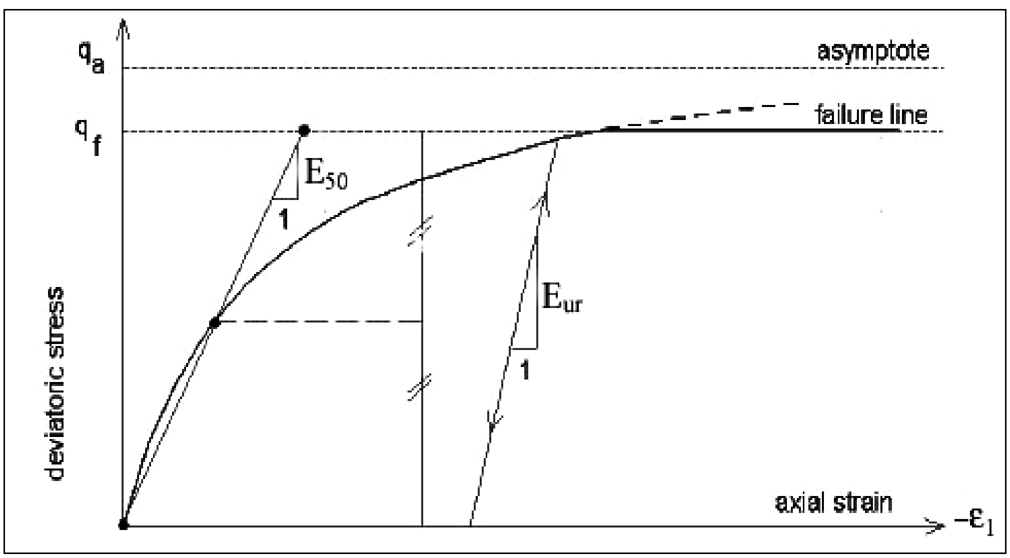


Figure 3.8 Hyperbolic stress strain relationship in primary loading for a standart drained triaxial test (Schanz et al. 1999)

For unloading and reloading, another stress-dependent stiffness modulus, E_{ur} is used and expressed as :

$$E_{ur} = E_{ur}^{ref}\left(\frac{c.\cot(\phi) - \sigma_3'}{c.\cot(\phi) + p^{ref}}\right)^m \tag{3.23}$$

where E_{ur}^{ref} is the reference modulus for unloading and reloading corresponding to the reference pressure p^{ref}.

To simulate the oedometer loading or isotropic loading, the tangent stiffness modulus for oedometer loading, E_{oed} is used as :

$$E_{oed} = E_{oed}^{ref}\left(\frac{c.\cot(\phi) - \sigma_1'}{c.\cot(\phi) + p^{ref}}\right)^m \tag{3.24}$$

where E_{oed}^{ref} is the tangent stiffness at a vertical stress of p^{ref} as shown in Fig. 3.9.

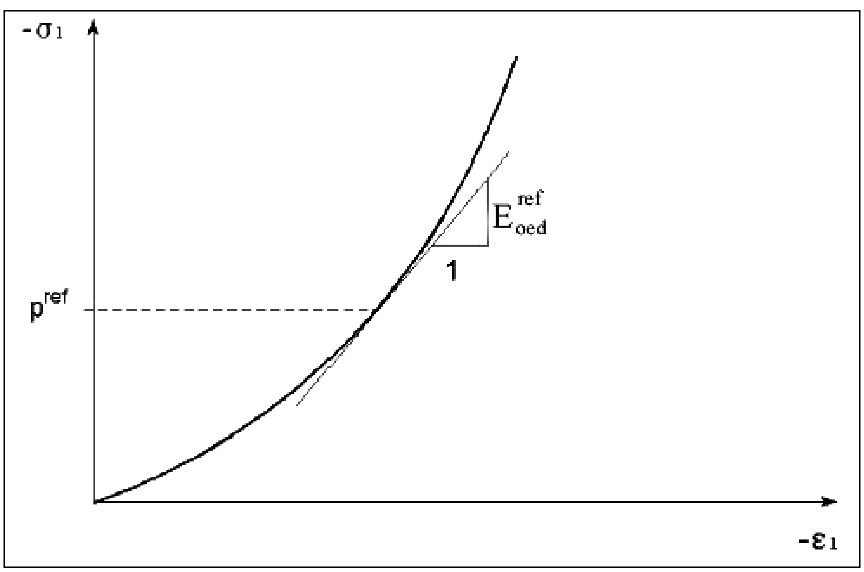


Figure 3.9 Definition of E_{oed}^{ref} in oedometer test results (Schanz et al. 1999)

3.5.2 YIELD SURFACE, FAILURE CONDITION, HARDENING LAW

Considering the triaxial case, two yield functions f_{12} and f_{13} are defined as :

$$f_{12} = \frac{q_a}{E_{50}} \frac{(\sigma_1 - \sigma_2)}{q_a - (\sigma_1 - \sigma_2)} - \frac{2(\sigma_1 - \sigma_2)}{E_{ur}} - \gamma^p \tag{3.25a}$$

$$f_{13} = \frac{q_a}{E_{50}} \frac{(\sigma_1 - \sigma_3)}{q_a - (\sigma_1 - \sigma_3)} - \frac{2(\sigma_1 - \sigma_3)}{E_{ur}} - \gamma^p \tag{3.25b}$$

with the definition of γ^p as :

$$\gamma^p = -(\varepsilon_1^p - \varepsilon_2^p - \varepsilon_3^p) = -(2\varepsilon_1^p - \varepsilon_v^p) \approx -2\varepsilon_1^p \tag{3.26}$$

Here, the measure of plastic strain, γ^p, is used as the relevant parameter for frictional hardening. In reality, plastic volumetric strain ε_v^p will never be zero, but it is small compared to axial strain for hard soils, so it is neglected in the determination of γ^p.

For a given value of γ^p, the yield condition $f_{12} = f_{13} = 0$ can be visualized in p'-q plane by means of a yield locus. The shape of the yield loci depend on the exponent m. Fig. 3.10 shows the shape of successive yield loci for m=0.5 being typical for hard soils. For increasing loading, the failure surfaces approach the linear failure condition according to Eq. 3.20.

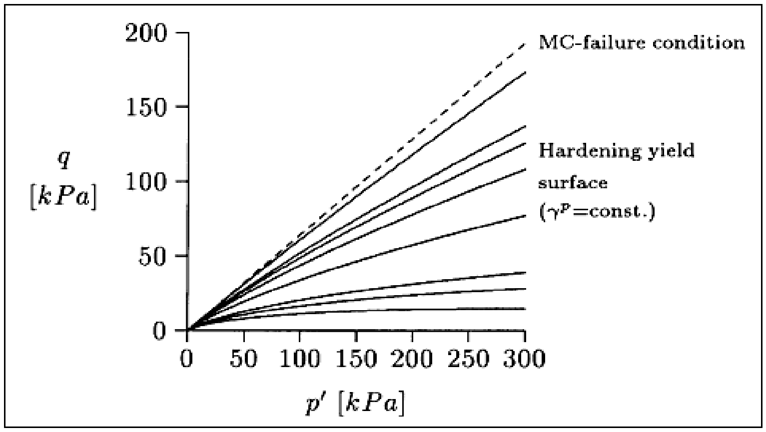


Figure 3.10 Successive yield loci for various values of γ^p and failure surface (Schanz et al. 1999)

3.5.3 FLOW RULE, PLASTIC POTENTIAL FUNCTIONS

As for all plasticity models, the hardening soil model involves a relationship between rates of plastic strain. This flow rule has the linear form :

$$\dot{\varepsilon}_v^p = \sin(\psi_m)\dot{\gamma}^p \qquad (3.27)$$

where ψ_m is the mobilized dilatancy angle. The reader is referred to References 20 and 23 for details on ψ_m and the plastic potential functions.

3.5.4 ON THE CAP YIELD SURFACE

Plastic volumetric strains due to isotropic compression can not be explained in a 2-D plot like the one in Fig. 3.11 that shows the shear yield surfaces. A second type of yield surface must be introduced to close the elastic region in the direction of the p-axis. A cap type of yield surface is required to formulate a model with independent inputs of E_{50} and E_{oed}. The triaxial modulus, E_{50}, largely controls the shear yield surface; and the oedometer modulus, E_{oed}, controls the cap yield surface. In fact, E_{50}^{ref} largely controls the magnitude of the plastic strains associated wiith the shear yield surface. Similarly, E_{oed}^{ref} controls the magnitude of plastic strains that originate from the yield cap.

The cap yield surface is considered as:

$$f^c = \frac{\tilde{q}^2}{\alpha^2} + p^2 - p_p^2 \tag{3.28}$$

where α is an auxiliary model parameter related to K_0^{nc} as will be discussed later. Furthermore we have $p = -(\sigma_1 + \sigma_2 + \sigma_3)/3$ and $\tilde{q} = \sigma_1 + (\delta - 1)\sigma_2 - \delta\sigma_3$ with $\delta = (3 + \sin(\phi))/(3 - \sin(\phi))$. $\tilde{q}$ is a special stress measure for deviatoric stresses. In the special case of triaxial compression $(-\sigma_1 > -\sigma_2 = -\sigma_3)$, it yields $\tilde{q} = -(\sigma_1 - \sigma_3)$ and for triaxial extension $(-\sigma_1 = -\sigma_2 > -\sigma_3)$, $\tilde{q}$ reduces to $\tilde{q} = -\delta(\sigma_1 - \sigma_3)$. The magnitude of the yield cap is determined by the isotropic pre-consolidation stress p_p. We have a hardening law relating p_p to volumetric cap strain ε_v^{pc} as :

$$\varepsilon_v^{pc} = \frac{\beta}{m+1}\left(\frac{p_p}{p^{ref}}\right)^{m+1} \tag{3.29}$$

The volumetric cap strain is the plastic volumetric strain in isotropic compression. α and β are cap parameters which are not direct input parameters. Instead, we have relationships of the form :

$\alpha \leftrightarrow K_0^{nc}$

$\beta \leftrightarrow E_{oed}^{ref}$

such that K_0^{nc} and E_{oed}^{ref} can be used to determine α and β respectively. For understanding the shape of the yield cap, first it should be realised that it is an ellipse in p-$\tilde{q}$ plane as shown in Figure 3.11.

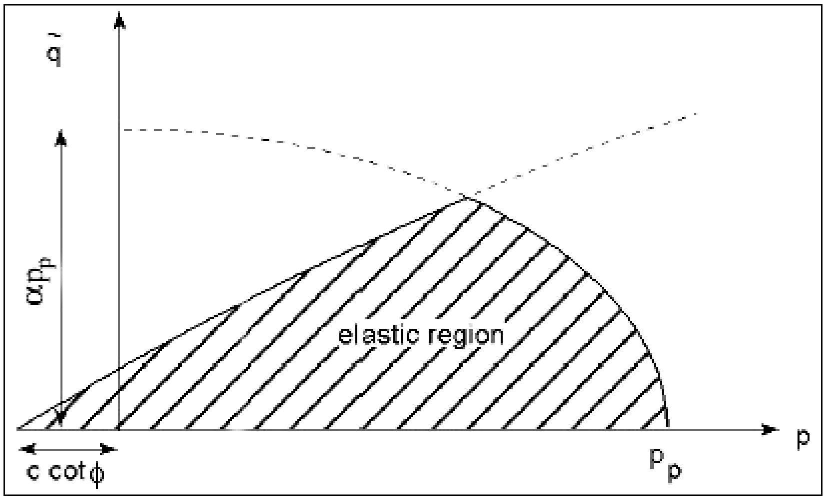


Figure 3.11 Yield surfaces in p-$\tilde{q}$ plane (Schanz et al. 1999)

The ellipse has length p_p on the p-axis and αp_p on the $\tilde{q}$-axis. Hence, p_p determines its magnitude and α its aspect ratio. The ellipse is used both as a yield surface and a plastic potential. Hence :

$$\dot{\varepsilon}^{pc} = \lambda \frac{\partial f^c}{\partial \sigma} \quad \text{with} \quad \lambda = \frac{\beta}{2p}\left(\frac{p_p}{p^{ref}}\right)^m \frac{\dot{p}_p}{p^{ref}} \qquad (3.30)$$

This expression for λ derives from the yield condition $f^c = 0$ and Eq. 3.29 for p_p.

To understand the yield surfaces in full detail, one should consider both Figs 3.11 and 3.12. Fig. 3.11 shows simple yield lines whereas Fig. 3.12 depicts yield surfaces in principal stress space. Both the shear locus and the yield cap have the hexagonal shape of the classical Mohr-Coulomb failure criterion. In fact, the shear yield locus can expand up to the ultimate Mohr-Coulomb failure surface, the yield surface expands as a function of the pre-consolidation stress p_p. (Schanz et al. 1999)

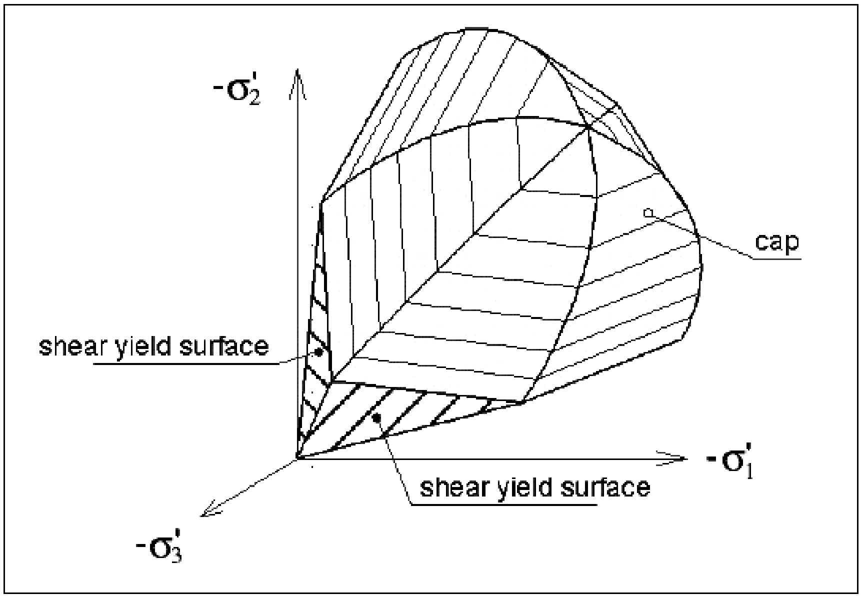


Figure 3.12 Representation of total yield contour in principal stress space for cohesionless soil (Schanz et al. 1999)

For further information about the Hardening Soil model, the reader is referred to References 20 and 23.

CHAPTER 4

ARTIFICIAL NEURAL NETWORKS

4.1 INTRODUCTION

Artificial neural networks (ANNs) are computational devices which are inspired by the networks of nerve cells in the brain. Although ANNs do not approach the complexity of the brain, there are two key similarities. First, the building blocks of both networks are simple computatinal devices that are highly interconnected. Second, the connections between neurons determine the function of the network.(Hagan et al. 1996)

Neural networks utilize a parallel processing structure that has large numbers of processors and many interconnections between them. Each processor is linked to many of its neighbours so that there are many more interconnects than processsors. The power of the neural network lies in the tremendous number of interconnections.(Dayhoff 1990)

ANNs can be trained to perform a particular function by adjusting the interconnections (weights) between neurons. Neural networks are trained to perform complex functions in different fields of application such as pattern recognition, identification, classification, speech, vision and control systems.

Engineers have used various tools to perform casual modeling (mapping from cause to effect for estimation and predicition) and inverse mapping (mapping from effects to causes) which include statistics, regression, probability, optimization, knowledge-based systems, and others. The nature of a neural network is to map from the input patterns to output patterns. Therefore an artificial neural network is another tool for engineers to perform both casual modeling or inverse mapping. (Kartam et al. 1997)

Neural networks are used in various fields of geotechnical engineering including parameter assesment, underground openings, foundations, site investigation, liquefaction, retaining structures, slopes and ground movement. (Toll 1996)

4.2 NEURON MODEL AND NETWORK ARCHITECTURES

4.2.1 NEURON MODEL

SIMPLE NEURON :

A simple neuron with a single input is shown in Fig. 4.1. The scalar input (p) is multiplied by the weight (w) and wp is obtained. A bias (b) is added to wp and the net input (n) is formed. A transfer function (activation function), f, is used to obtain the scalar neuron output (a) from the net input. Then the neuron output is calculated as $a = f(wp + b)$.

Bias may be considered as a weight that has an input value of 1. The bias can be omitted in some neural networks.

The transfer function may be a linear or non-linear function of n. Transfer functions are used to satisfy some specification of the problem that the neuron is attempting to solve. One of the most commonly used transfer functions is the log-sigmoid transfer function shown in Figure 4.2. This transfer function transfers the input into the range 0 to 1 as an output. The log-sigmoid transfer function is generally used in backpropogation because that it is differentiable. Various transfer functions are used in neural networks to achieve the desired goal.

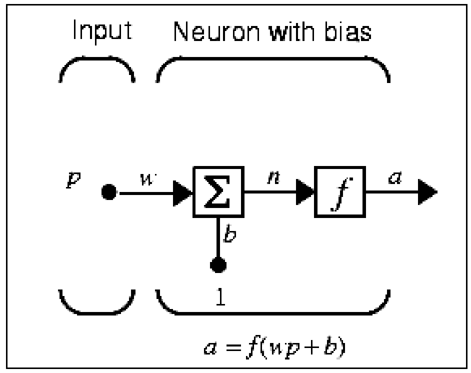


Figure 4.1. Single neuron architecture (Hagan et al. 1996)

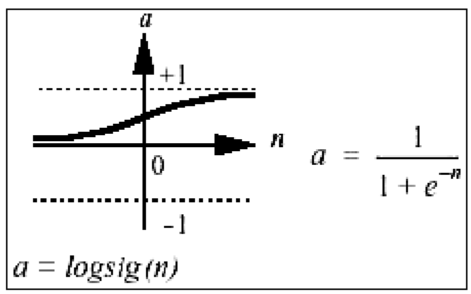


Figure 4.2. Log-sigmoid transfer function (Hagan et al. 1996)

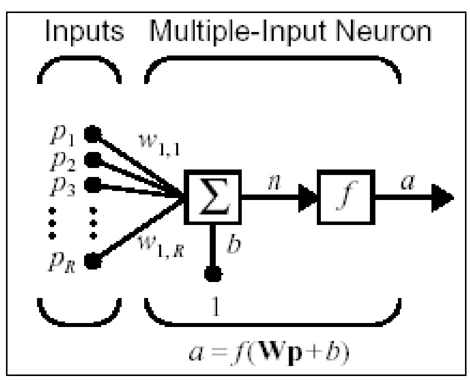


Figure 4.3. Multiple input neuron (Hagan et al. 1996)

NEURON WITH MULTIPLE INPUT :

Generally, a neuron has more than one input. A neuron with a R-element input vector is shown in Fig. 4.3. The individual inputs p_1, p_2, ..., p_R are multiplied by the corresponding weights $w_{1,1}$, $w_{1,2}$, ..., $w_{1,R}$ of the weight matrix W. The net input, n, is calculated as: $n = w_{1,1}p_1 + w_{1,2}p_2 + ... + w_{1,R}\,p_R + b$. This expression can also be written in the matrix form as : $n = Wp + b$. Then the output can be expressed as : $a = f\,(Wp + b\,)$.

4.2.2 NETWORK ARCHITECTURES

A LAYER OF NEURONS :

Multiple neurons are combined in parallel to form a layer. A single layer of S neurons with R input elements is shown in Fig. 4.4. Each input is connected to each neuron with the weight matrix that has S rows and R columns. Each neuron has a bias b_i, a summer, a transfer function f, and an output a_i.

The weight matrix is expressed as : $$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ & & & \\ w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix}$$

where the row indices indicate the destination neuron, and the column indices indicate the input source for that weight. Thus $w_{1,2}$ indicates that this weight represents the connection to the second neuron from the first input.

MULTIPLE LAYERS OF NEURONS :

A single layered network is rarely capable of solving the problems. So generally several layers take place in neural networks. A three layer netwok is shown in Fig. 4.5.

As shown in Fig. 4.5, the layer number is indicated as a superscript to the names of the variables. There are R inputs and S^1 neurons in the first layer, S^1 inputs and S^2 neurons in the second layer. It can be seen that the outputs of the first and second layers are inputs for the second and third layers respectively. Different layers can have different number of neurons.

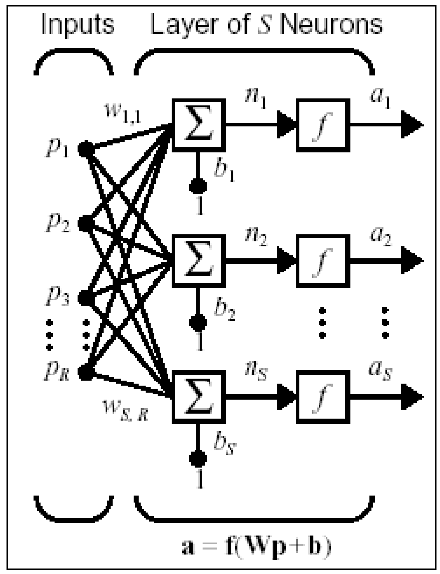


Figure 4.4. A layer of neurons (Hagan et al. 1996)

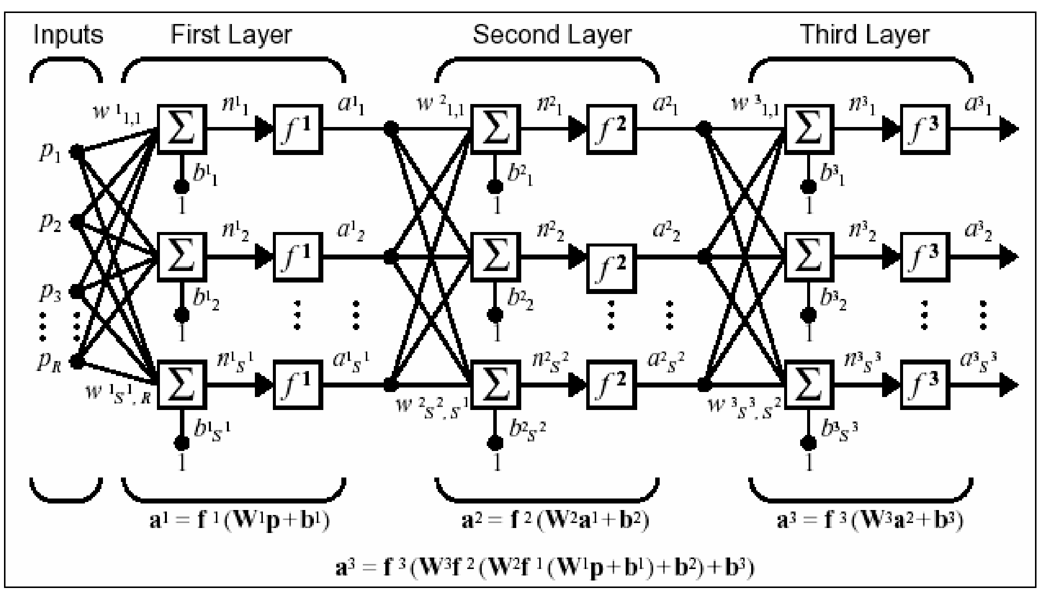


Figure 4.5. Three-layer network (Hagan et al. 1996)

A layer whose output is the network output is called as the output layer. The other layers are called as hidden layers. The network in Fig. 4.5 has two hidden layers which are layers 1 and 2, and one output layer that is layer 3.

Multilayer networks are powerful. In general, a network of two layers where the first layer is sigmoid and the second layer is linear can be used to approximate any function. This kind of network is widely used in *backpropogation* which is discussed later in this chapter.

4.3 TRAINING OF THE NETWORK

Training can be defined as the modification of the connection strengths (weights) of the network by a specified learning rule to reach the desired solution. The learning rule defines how the network is modified in response to experience.

A learning rule is defined as a procedure for the modification of the weights and biases of the network. (This procedure is often referred as training algorithm.) The learning rule is applied to train the network to perform a particular task. In *supervised learning*, the learning rule is provided with the inputs and the outputs. In *unsupervised learning*, no target outputs are available and the weights are modified in response to inputs only.

4.3.1 THE LMS ALGORITHM

The least mean square (LMS) learning rule (also called as standart delta rule) by Widroff-Hoff (1960) is used in the training of *linear filters* (See Fig. 4.6) which are one layered neural networks with linear transfer functions.

The LMS algorithm adjusts the weights of the linear network to minimize the squares of differences between the actual and the desired (target) output values summed over the output layers and all pairs of input/output vectors.

Let's define

$$E_p = \frac{1}{2}\sum_j \left(t_{pj} - o_{pj}\right)^2 \tag{4.1}$$

as the measure of error on the input/output pattern *p* and let $E = \Sigma E_p$ be the overall measure of error (the error function or the performance function). The index *p* ranges over the set of input patterns, *j* ranges over the set of output units, and E_p represents the error on pattern *p*. The variable t_{pj} is the desired output and o_{pj} is the actual output of the *j*'th output unit for pattern *p*. It is desired to find the weights that minimize the error function which is described above.

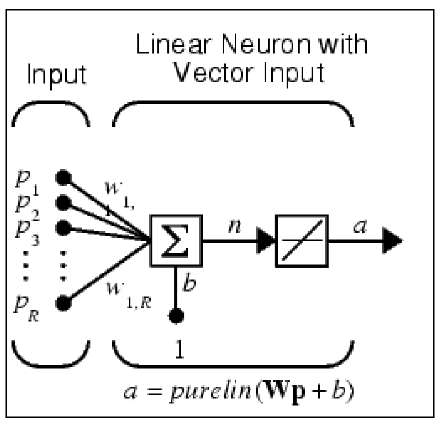


Figure 4.6 An example of linear filter with R input elements (Hagan et al. 1996)

For this purpose, it is useful to consider how the error varies as a function of any weight. The LMS procedure finds the weights that minimize the error function using a method called as *gradient descent.* That is, after each pattern is presented, the error is computed and each weight is moved down the error gradient toward its minimum value for that pattern. Since the entire error function on each pattern presentation can not be mapped, a simple procedure to determine how much to increase or decrease each weight must be found. The idea of gradient descent is to make a change in the weight proportional to the negative of the derivative of the error, as measured on the current pattern with respect to each weight. Thus the learning rule becomes :

$$\Delta w_{ji} = -k \frac{\partial E_p}{\partial w_{ji}} \tag{4.2}$$

where *k* is the proportionality constant. To take the derivative of the performance function, the chain rule can be used to write the derivative as the product of two parts as :

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ji}} \tag{4.3}$$

The first part can be found from Equation 4.1 as :

$$\frac{\partial E_p}{\partial o_{pj}} = -(t_{pj} - o_{pj}) = -\delta_{pj} \tag{4.4}$$

Since we have linear layers,

$$o_{pj} = \sum_i w_{ji} i_{pi} \tag{4.5}$$

from which it can be concluded that :

$$\frac{\partial o_{pj}}{\partial w_{ji}} = i_{pi} \tag{4.6}$$

where i_{pi} is the i'th element of the input pattern *p*. Substituting back into Eq. 4.3,

$$-\frac{\partial E_p}{\partial w_{ji}} = \delta_{pj} i_{pi} \tag{4.7}$$

Therefore Δw_{ji} can be found as :

$$\Delta w_{ji} = \mu.\delta_{pj}.i_{pi} \tag{4.8}$$

where $\mu = 2k$

According to this learning procedure, each weght is changed until it reaches its minimum error value. When all the weights reach their minimum points, the system reaches equilibrium. Then the problem is entirely solved or the set of weights that produce as small an error as possible is obtained. (Mccelland and Rumelhart 1998)

4.3.2 BACKPROPOGATION

Backpropogations are a kind of neural networks which are widely used in solving problems that require pattern mapping (an input pattern is given and the network produces associated output pattern). They are cretaed by generalizing the standart delta rule to multiple-layer networks with nonlinear differentiable transfer functions.

Backpropogation learning rules are based on the simple concept as in the delta rule : the error between the actual output and the desired output is lessened by modifying the weights and as a result future responses are more likely to be correct. When the network is given an input, the output units are obtained by simulation of the network. The output layers then provide the network's response. When the network corrects its internal parameters, the correction mechanism starts with the output layers and back-propogates backward through each internal(hidden) layer. Hence the term backpropogation is used for this kind of networks. (Dayhoff 1990)

The power of backpropogation lies in its ability to train hidden layers and therefore escape the restricted capabilities of single layer networks (like linear filters in which the LMS learning procedure is used).

4.3.3 BACKPROPOGATION ALGORITHM (THE GENERALIZED DELTA RULE)

It is shown how the standart delta rule implements gradient descent in sum-squared error for linear activation(transfer) functions. There is no hidden unit in this case and the error surface is shaped like a bowl with only one minimum. However if hidden units exist, there is a possibility of getting stuck in local minima. Also, linear systems using LMS algorithms can not compute more in multiple layers than they can in a single layer.

The basic idea of the backpropogation learning method is to combine a non-linear system capable of making decisions with the objective error function of LMS and gradient descent. To do this, the derivative of the error function with respect to any weight in the network is calculated and then the weight is changed according to the rule :

$$\Delta_p w_{ji} = -k \frac{\partial E_p}{\partial w_{ji}} \qquad (4.9)$$

With an appropriate choice of non-linear transfer function, the backpropogation learning rule can be derived. (See Reference 22 for derivation) The results of this derivation are summarized in three equations. (Rumelhart et al. 1986) :

First, the generalized delta rule has exactly the same form as the standart delta rule: The weight on each line should be changed by an amount proportional to the product of an error signal, δ, availbale to the layer receiving input along that line and the output of the layer sending activation along that line. In symbols,

$$\Delta_p w_{ji} = \mu.\delta_{pj}.o_{pi} \qquad (4.10)$$

The other equations specify the error signal. The determination of the eror signal, δ_{pj}, is a recursive process and starts with the output layers. The error signal of an output layer is similar to the standart delta rule and can be expressed as :

$$\delta_{pj} = (t_{pj} - o_{pj}) f_j'(net_{pj}) \qquad (4.11)$$

where net_{pj} is the net output and $f_j'(net_{pj})$ is the derivative of the non-linear activation function that maps the total input to the layer to an output value. The error signal for hidden layers for which there is no specified target is determined recursively in terms of the error signals of the layers to which it directly connects and the weights of these connections. That is :

$$\delta_{pj} = f_j'(net_{pj}) \sum_k \delta_{pk} w_{kj} \qquad (4.12)$$

where the layer is not an output layer.

Therefore, the generalized delta rule involves two phases : During the first phase, the input is presented and propogated forward through the network and the output value o_{pj} is calculated for each layer. This output is then compared with the target values and an error signal δ_{pj} is computed for each output layer. In the second phase, a backward pass through the network is done during which the error signal is passed to each layer and the appropriate weight changes are made. This second phase involves the recursive computation of δ as indicated above.

MOMENTUM :

The generailized delta rule requires only that the change in weights be proportinal to (∂E_p / ∂w). The constant of proportionality is the learning rate. The larger this constant, the larger the changes in the weights. Generally, a learning rate as large as possible is chosen without leading to oscillation that offers the most rapid learning. One way to increase the learning rate without leading to oscillation is to modify the generalized delta rule by adding a momentum term. This can be accomplished by :

$$\Delta w_{ji}(n+1) = \mu(\delta pj.o_{pi}) + \alpha\Delta w_{ji}(n) \qquad (4.13)$$

where n is the presentation number, μ is the learning rate and α is a constant that determines the effect of post weight changes on the current direction of movement in weight space. This provides a kind of momentum in weight space that effectively filters out high-frequency variations of the error surface in the weight space. (Rumelhart et al. 1986)

4.3.4 A BRIEF INFORMATION ABOUT QUASI-NEWTON METHODS

As described, standart backpropogation is a gradient descent algorithm in which the network weights are moved along the negative of the gradient of the performance function. One iteration of the standart backpropogation algorithm can be written as:

$$X_{k+1} = X_k - \alpha_k.g_k \qquad (4.14)$$

where X_k is the vector of current weights and biases, g_k is the current gradient and α_k is the learning rate.

There are some methods which have variations on the basic algorithm and based on other standart optimization techniques. Newton's method is one of them and is an alternative to the standart gradient descent algorithms. The basic step of the method is :

$$X_{k+1} = X_k - A_k^{-1} g_k \qquad (4.15)$$

where A_k is the Hessian matrix (second derivatives) of the performance function at the current values of the weights and biases. The disadvantage of this method is that the calculation of the Hessian matrix is very complex. However, there are some algorithms based on Newton's method which don't require the calculation of the second derivatives. These are called quasi-Newton (secant) methods. These algorithms make an approximation of the Hessian matrix at each iteration of the algorithm.

4.3.5 A BRIEF INFORMATION ABOUT THE LEVENBERG-MARQUARDT ALGORITHM

The Levenberg-Marquardt algorithm is similar to quasi-Newton methods and designed to approcah the second order training speed without the computation of the Hessian matrix. If the performance function is in the form of a sum of squares (as in the delta rule), then the Hessian matrix can be approximated as :

$$H = J^T J \qquad (4.16)$$

and the gradient can be computed as :

$$g = J^T e \qquad (4.17)$$

where J is the Jacobian matrix that contains the firts derivatives of the network errors with respect to the weights, and e is a vector of network errors.

The basic step of Levenberg-Marquardt algorithm is:

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e \qquad (4.18)$$

where μ is a positive scalar and I is a unit matrix. If μ is large, then the algorithm becomes gradient descent with a small step size.

The Levenberg-Marquardt algorithm is very fast and effective compared to the standart gradient descent algorithms for some input/output patterns. Also the neural network training algorithm used in this study is the Levenberg-Marquardt algorithm with the performance function chosen as mean square of the errors.

The reader is referred to References 10 and 11 for more information about the Levenberg-Marquardt algorithm.

CHAPTER 5

ANALYSES AND RESULTS

5.1 ASSUMPTIONS AND THE FINITE ELEMENT MODEL

The following assumptions are made in the calculation of the lateral pressures due to surface strip load :

- Plane strain condition is assumed
- The ground surface is assumed to be horizontal
- The wall is assumed to be vertical
- The wall is assumed to be non-displacing and rigid
- The backfill is assumed to be underlain by a rigid base
- The friction between the wall and backfill is neglected and the wall is assumed to be perfectly smooth
- Surface strip loading is assumed to be flexible

These assumptions are valid for all of the analyses in this study. Fig. 5.1 shows a sketch of the problem geometry. In the figure; h is the height of the wall, a is the distance of the strip load to the wall, w is the width of the strip load, q is the magnitude of the strip load, σ_h is the lateral pressure due to strip load, P is the total lateral thrust on the wall due to only strip load and d is the distance between the point of application of P and the ground surface.

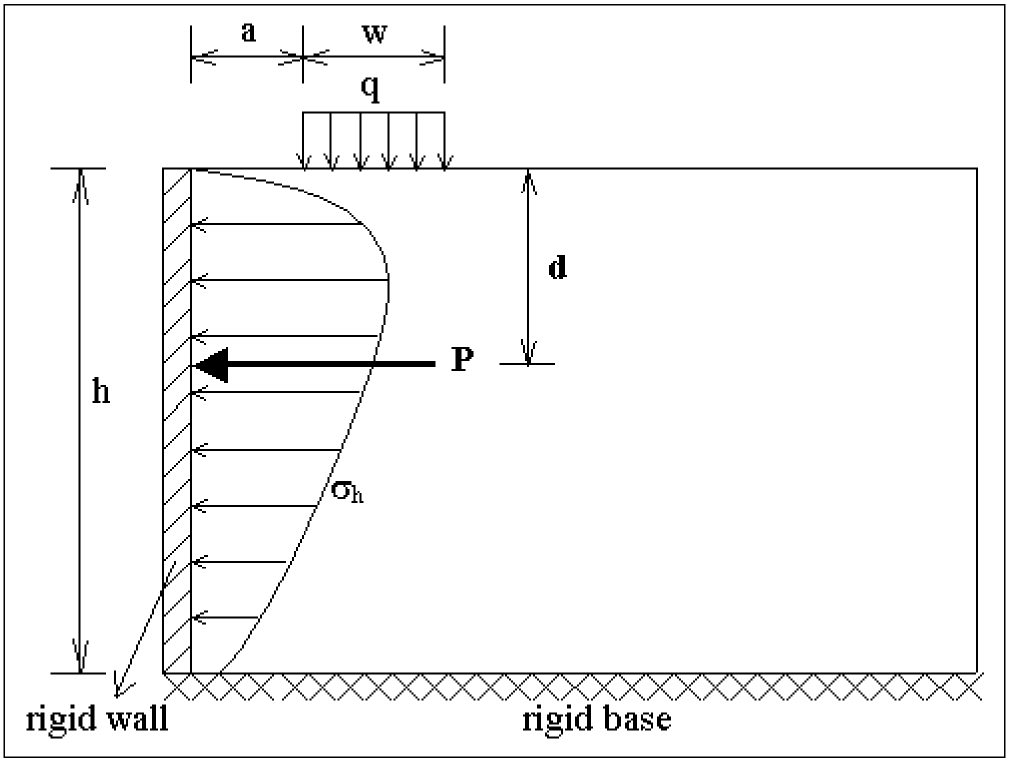


Figure 5.1 Problem Geometry

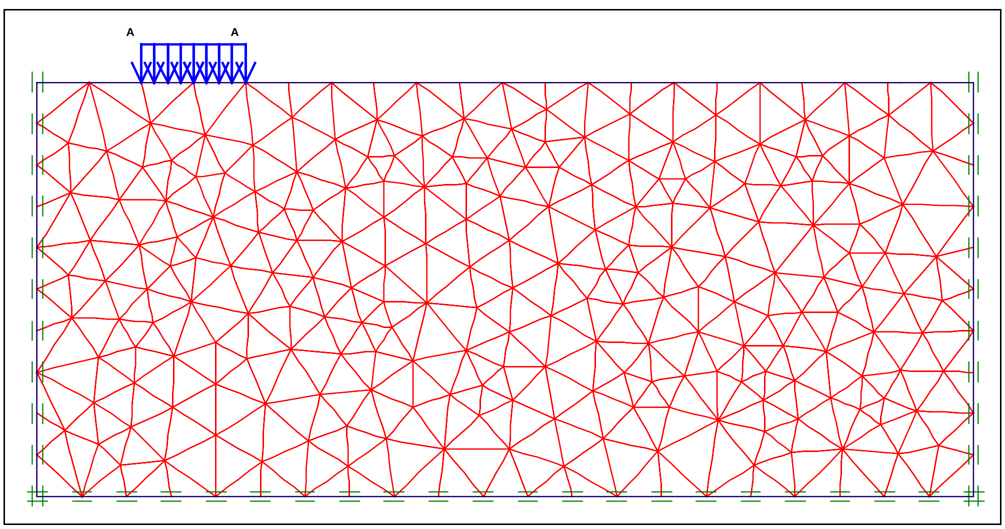


Figure 5.2 Typical Finite Element Model

Fig. 5.2 shows a typical finite element model of the problem used in this study. Because the wall is assumed to be rigid and perfectly smooth, it is sufficient to set only horizontal fixities (which means that the nodes can move freely in the vertical direction and can not move in the horizontal direction) in the wall boundary for the simulation of the wall. The rigid base assumption is simulated by setting vertical fixities (nodes can move freely in the horizontal direction and can not move in the vertical direction) in the lower boundary of the finite element model. A medium-coarse finite element mesh is used for all of the analyses (See Section 5.3 for the effect of the mesh coarseness).

5.2 CALCULATION PROCEDURE

The computer program "PLAXIS ver. 7.11" in which finite element method is used is employed in the calculations. In the finite element method, the domain(continuum) is divided into a number of subdomains called finite elements which consist of a number of nodes. Each node has a number of degrees of freedom (which are the displacement components in this case) that correspond to discrete values of the unknowns in the boundary value problem to be solved.

Two kind of triangular elements are present in PLAXIS for the plane strain condition which are 6-node and 15-node elements. 15-node triangular elements are used in this study for a more accurate calculation of the stresses. The 15-node triangular element consists of 15 nodes as shown in Fig. 5.3. the displacements are calculated at these nodes. In contrast to displacements, stresses are calculated at individual stress points rather than at the nodes. The 15-node element has 12 stress points as shown in Fig. 5.3.

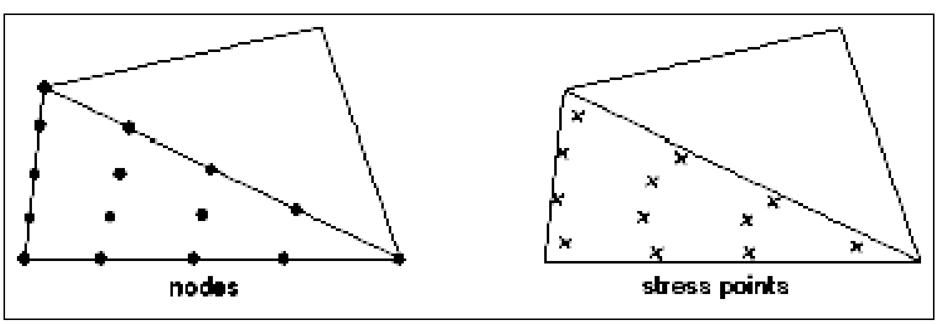


Figure 5.3 Nodes and stress points of the 15-node element

The lateral pressures due to strip load are calculated according to the following procedure :

1- The initial stresses are calculated according to Jaky's formula ($K_0 = 1\text{-}\sin\phi$).
2- The total stresses are calculated after the surface strip load is applied.
3- The lateral pressures due to surface strip load are obtained by substracting the initial stresses (obtained at step 1) from the total stresses (obtained at step 2).

5.3 VERIFICATION OF THE PROGRAM

At the beginning, the results of PLAXIS for a linear elastic soil model are compared to the results of the elastic solution for surface loading of a finite layer underlain by a rigid base recommended by Poulos&Davis(1974). The problem is summarized as :

- 5 m long ,50 kN/m^2 uniform pressure on a 10 m thick layer that is underlain by a rigid base
- Poisson's ratio (ν) is taken as 0.4.
- Horizontal and vertical stresses beneath the edge of the uniform load are calculated by Poulos&Davis(1974) method.
- The same problem is solved by PLAXIS by using linear elastic model. (The geometry and the finite element mesh is given in Fig.5.4)
- The results for the both solution methods are shown in Fig. 5.5.

It can be seen from Fig. 5.5 that the finite element solutions by PLAXIS are very similar to the analytical solution by Poulos&Davis (1974).

The effect of the mesh coarseness is also investigated. For this purpose, the same problem is solved for three different mesh sizes included in the program as : very coarse mesh, medium mesh and very fine mesh.

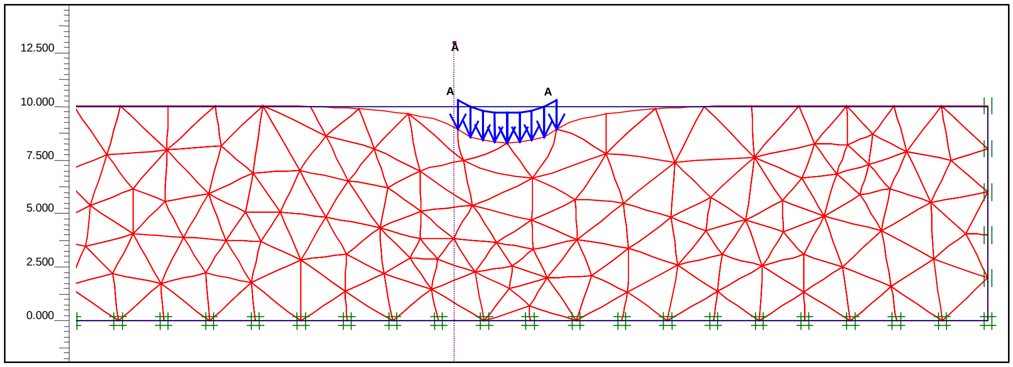


Figure 5.4 Deformed mesh after loading

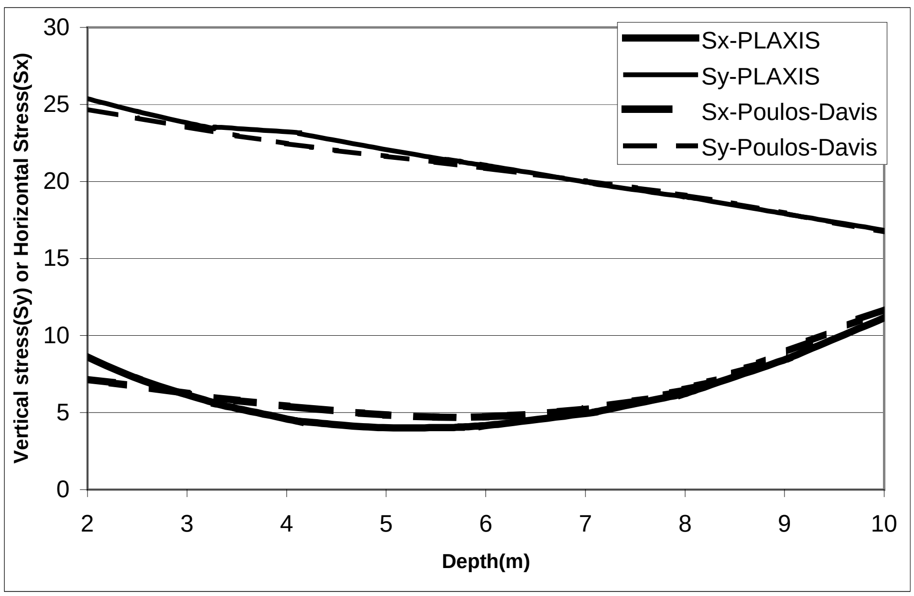


Figure 5.5 Comparison of the PLAXIS and Poulos&Davis solutions

The analyses are made for a 4 m high wall and a strip load at a distance of 1 m to the wall. The width of the strip load is 1 m and its magnitude is 25 kN/m. The finite element models are shown in Fig. 5.6.

The lateral pressure distributions due to strip load for three different mesh coarseness are shown in Fig. 5.7. It can be concluded from the figure that the the results are very similar and the mesh coarseness is not effective on the pressure distributions for this type of loading.

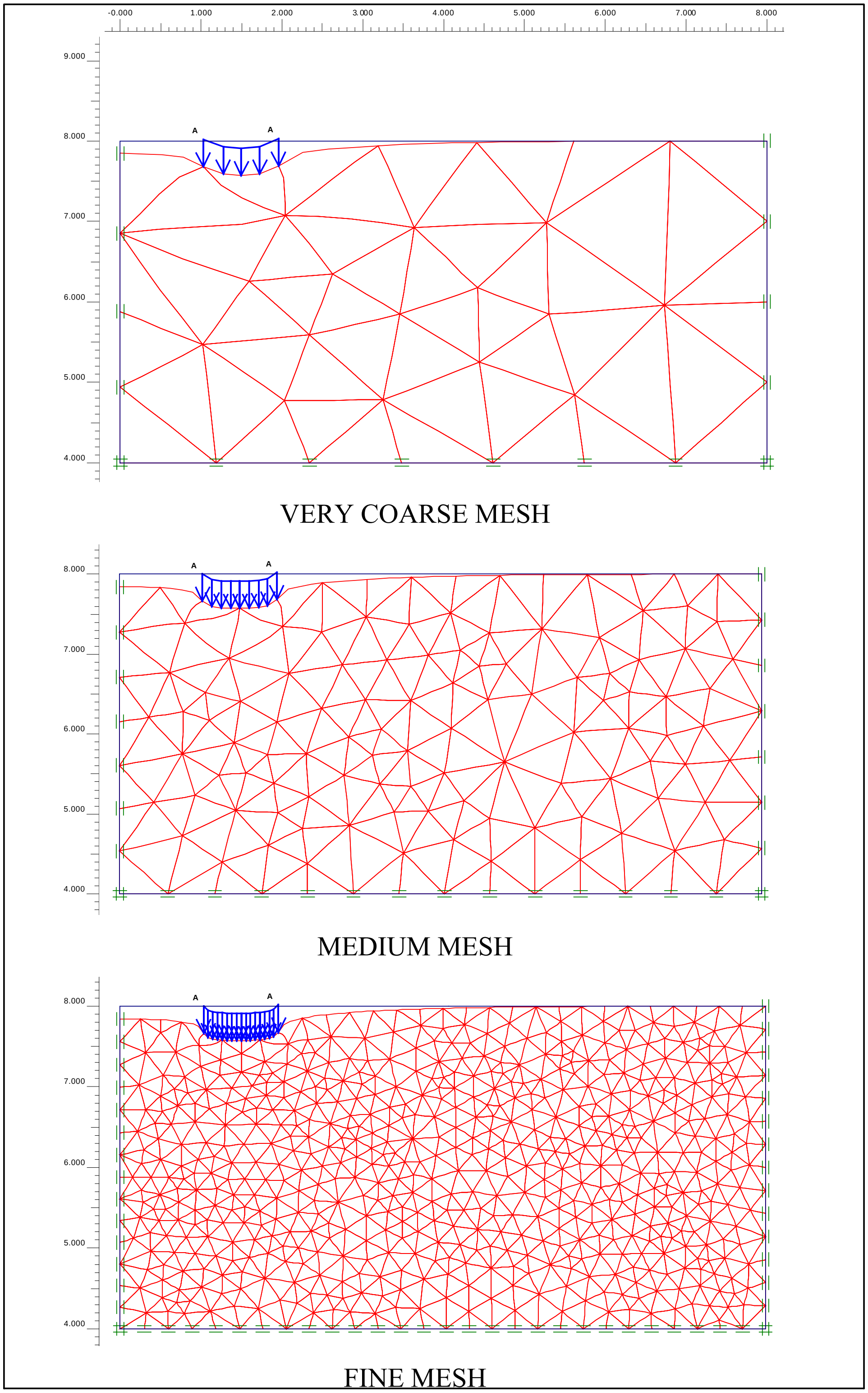


Figure 5.6 Deformed mesh plots for different mesh sizes

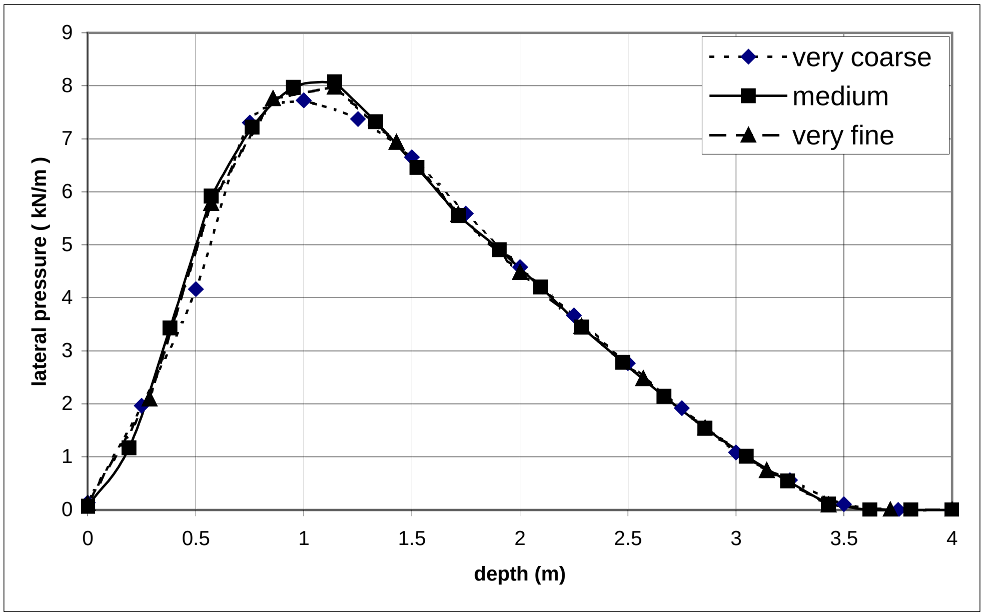


Figure 5.7 Lateral pressure distributions for different mesh sizes

It is decided to use medium-coarse mesh in this study.

5.4 THE EFFECT OF SOIL PARAMETERS

Hardening Soil model (which is discussed in detail in Chapter 3 Section 3.5) is used to represent the stress-strain relationship of the soil in the finite element calculations made by using PLAXIS. The effect of the hardening soil parameters(that are used to represent stress-strain behaviour of the soil) on lateral pressure distributions is investigated.

To investigate the effect of a specific parameter, the other parameters are taken as constant typical values. Then the analyses are made for various values of the specific parameter.

The aim of this study is to find a general solution for the calculation of the total lateral thrust and its point of application. So the stiffness parameters, E_{50}^{ref} and

E_{oed}^{ref} are taken as the same values as proposed for most soils by the program manual. So $E_{50}^{ref} = E_{oed}^{ref} = 30000$ kN/m^2 is taken for all the cases that is suggested for medim sand by the program manual. Similarly R_f is taken as 0.9 for all analyses which is the default value in the program.

The investigated parameters are :

- **c** : (Effective) cohesion of the soil (kN/m^2)

- ϕ : (Effective) internal angle of friction of the soil (in degrees)

- ψ : Angle of dilatancy of the soil (in degrees)

- **m** : Power for stress-level dependency of stiffness (unitless)

- **q** : Strip load magnitude (kN/m^2)

5.4.1 THE EFFECT OF "m"

The stiffness parameters E_{50} and E_{oed} are dependent on the confining stress by the power *m* as in the Equations 3.22 and 3.24. The typical values of *m* are about 0.5. So three values of *m* as 0.4,0.5 and 0.6 are considered to see the effect of *m* on the lateral pressure distributions. The other parameters are taken as :

-height of the wall (h) : 8 m.
-distance of the strip load to the wall (a) : 0 m
- width of the strip load (w) : 1 m
- cohesion of the soil (c) : 10 kN/m^2
- internal angle of friction of the soil (ϕ) : 30
- dilatancy angle of the soil (ψ) : 0

- magnitude of the strip load (q) : 50 kN/m^2

The lateral pressure distributions for different *m* values are shown in Fig. 5.8. It can be seen from the figure that the *m* parameter has no significant effect on the lateral pressure distributions in the case of a strip loading. So it is concluded that the *m* parameter can be omitted for the rest of the calculations and a taking a typical value of *m* as 0.5 is acceptable for a general solution.

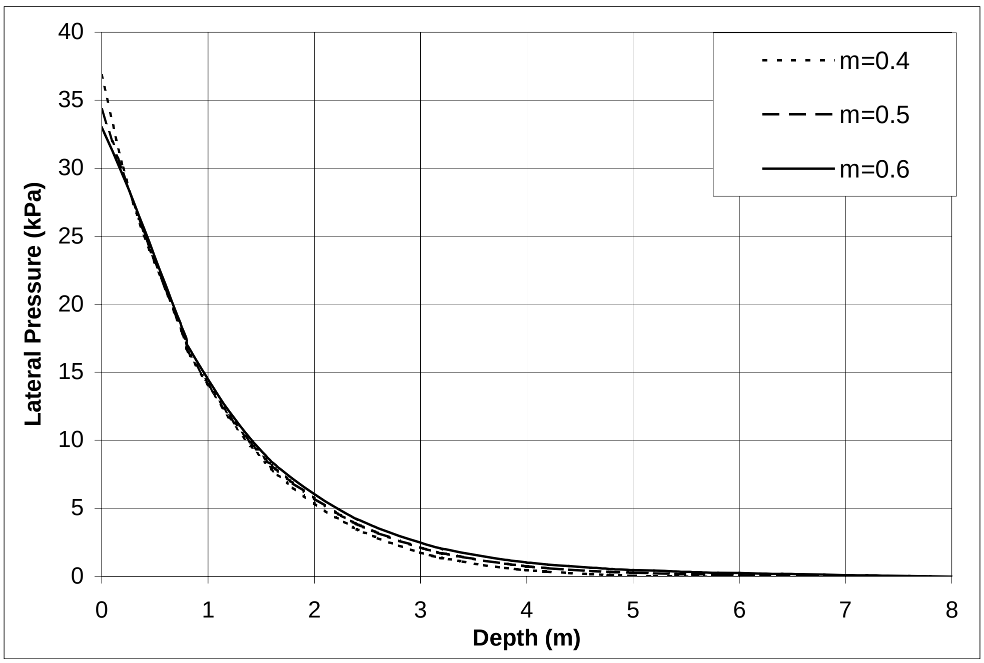


Figure 5.8 Lateral pressure distributions for different *m* values

5.4.2 THE EFFECT OF DILATANCY ANGLE

The effect of dilatancy angle, ψ, is investigated by considering three typical values as 0, 5 and 10. For the three analyses, the other parameters are taken as:

-height of the wall (h) : 8 m.
-distance of the strip load to the wall (a) : 0 m

- width of the strip load (w) : 1 m
- cohesion of the soil (c) : 0 kN/m^2
- internal angle of friction of the soil (ϕ) : 40
- parameter of stress dependency (m) : 0.5
- magnitude of the strip load (q) : 50 kN/m^2

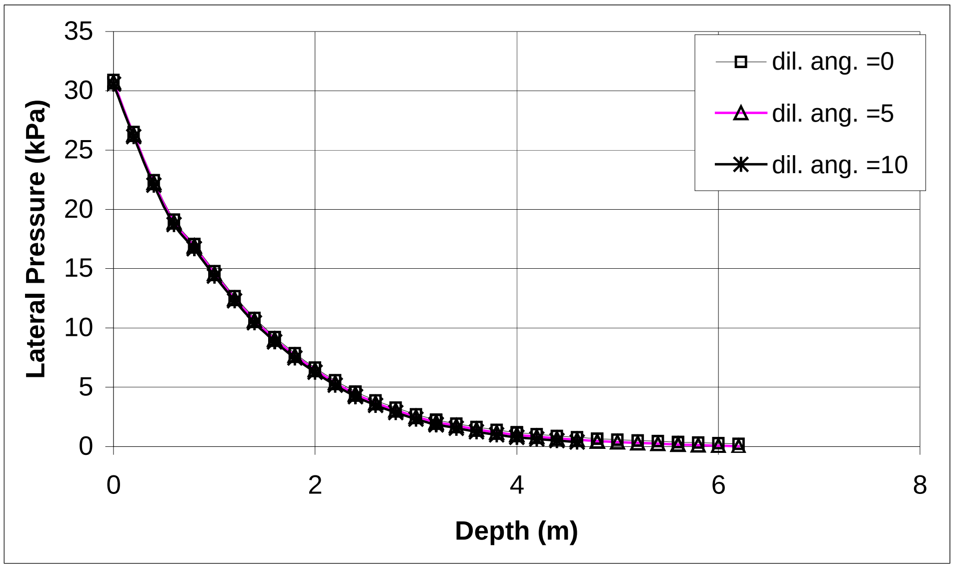


Figure 5.9 The effect of dilatancy angle on lateral pressures

Figure 5.9 shows the lateral pressure distributions for different values of dilatancy angle. It is clear from the figure that the results for the three cases nearly coincide and the angle of dilatancy has no effect on the lateral pressure distributions for this type of loading. Therefore it is concluded to omit the dilatancy angle and to take the dilatancy angle as zero for the analyses to find a general solution.

5.4.3 THE EFFECT OF LOAD MAGNITUDE

For a linear elastic analysis, the lateral pressure value is affected by only two parameters : the strip load magnitude and the poisson's ratio. The relationship between the lateral pressure at a point and the magnitude of the strip load is linear. Thus if the lateral pressure at a specific point is *p* for a load magnitude *q*, the lateral pressure at the same point will be 2*p* for a load magnitude 2*q*. However, for the hardening soil model, the non-linear behaviour of the soil effects the relation between the lateral pressure and the load magnitude.

To investigate the effect of non-linearity, the calculated lateral pressure values are normalized with the strip load magnitude. For this purpose, four different cases which are : q = 5 kN/m^2, q = 10 kN/m^2 , q = 25 kN/m^2 and q = 50 kN/m^2 are considered. The other parameters for these analyses are taken constant as :

-height of the wall (h) : 8 m.
-distance of the strip load to the wall (a) : 0 m
- width of the strip load (w) : 1 m
- cohesion of the soil (c) : 10 kN/m^2
- internal angle of friction of the soil (ϕ) : 30
- dilatancy angle of the soil (ψ) : 0
- parameter of stress dependency (m) : 0.5

The results are shown in Fig. 5.10. It can be seen in the figure that the normalized values do not coincide (which would coincide for a linear elastic analysis). So it is concluded that the load magnitude must be take part as an effective parameter in the general solution.

5.4.4 THE EFFECT OF COHESION

The cohesion of the soil is one of major parameters of the model. The ultimate deviatoric stress (q_f) is calculated by using the cohesion value. which is used in the calculation of the hyperbolic relationship between the deviatoric stress and the vertical strain by Equations 3.19 and 3.21. The cohesion is also used in the calculation of the confining stress dependent modulus as in Equations 3.22 and

3.24. So the cohesion of the soil is expected to have a significant effect on the lateral pressure distributions.

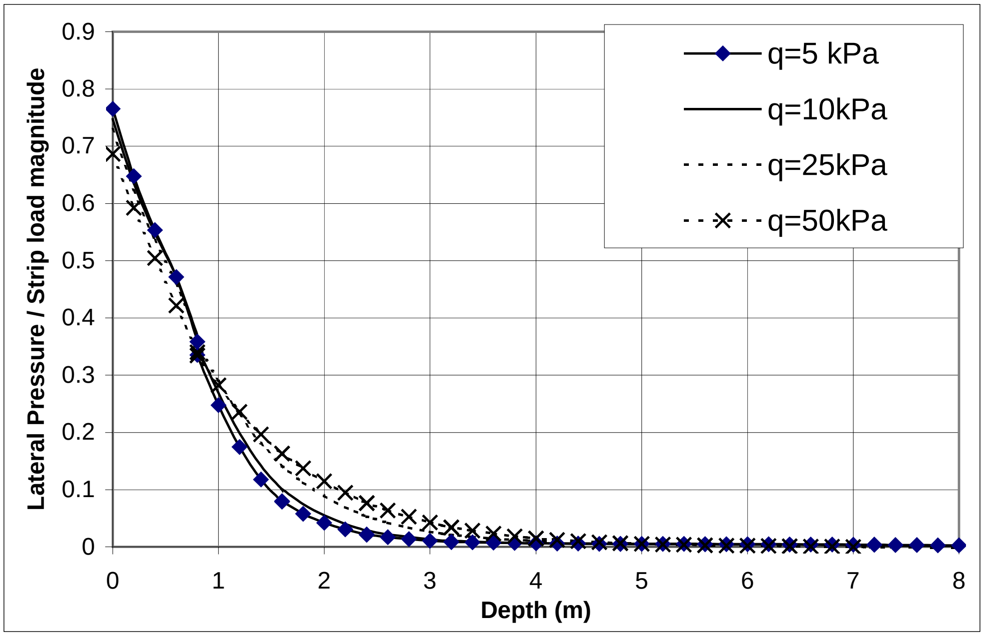


Figure 5.10 Normalized lateral pressure distributions for different load magnitudes

To investigate the effect of cohesion, four different cohesion values : c=0 kN/m^2, c=5 kN/m^2, c=10 kN/m^2 and c=20 kN/m^2 are used. The other parameters are held constant and taken as :

-height of the wall (h) : 8 m.

-distance of the strip load to the wall (a) : 0 m

- width of the strip load (w) : 1 m

- internal angle of friction of the soil (ϕ) : 30

- parameter of stress dependency (m) : 0.5

- dilatancy angle of the soil (ψ) : 0

- magnitude of the strip load (q) : 50 kN/m^2

Fig. 5.11 shows the effect of cohesion on the lateral pressures due to surface strip load. It is clear from the figure that the cohesion has significant effect on the lateral pressure distributions as expected. Therefore it is decided to take the cohesion another effective parameter on the lateral pressures in the general solution.

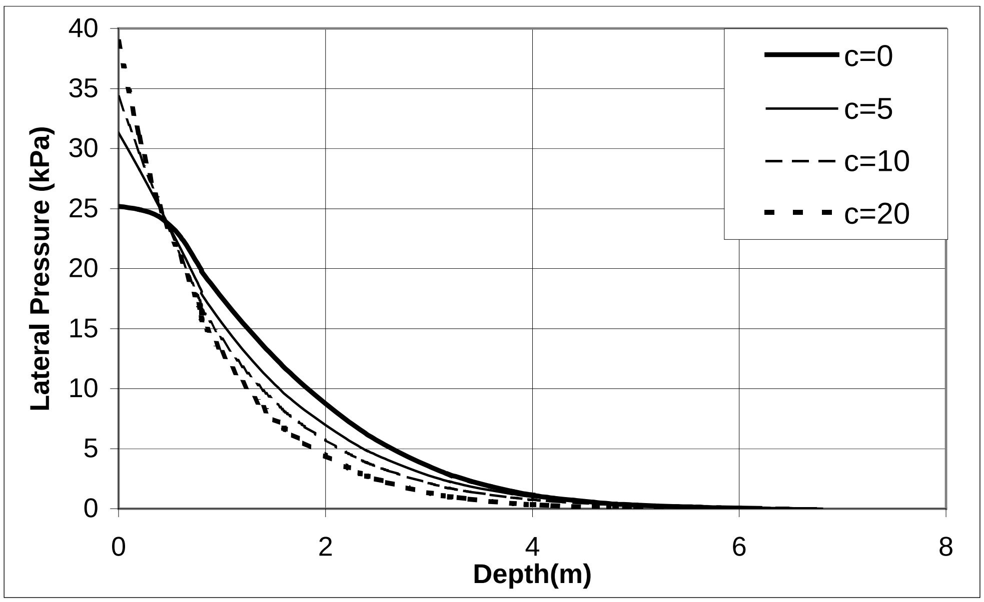


Figure 5.11 Lateral pressure distributions for various cohesion values

5.4.5 THE EFFECT OF THE ANGLE OF FRICTION

Similar to cohesion, the internal angle of friction (ϕ) is used in the calculation of the confining stress dependent modulus and the hyperbolic relationship between the strain and the deviatoric stress. Angle of friciton is also used in the determination of the initial stresses by using Jaky's formula. So the internal angle of friction is also expected to have a major effect on the lateral pressure distributions.

The lateral pressure distributions for four different angles of internal friction as : ϕ=25, ϕ=30, ϕ=35 and ϕ=40 are calculated and shown on Fig. 5.11. The other parameters for these analyses are taken as:

-height of the wall (h) : 8 m.

-distance of the strip load to the wall (a) : 0 m

- width of the strip load (w) : 1 m

- cohesion of the soil (c) : 0 kN/m^2

- parameter of stress dependency (m) : 0.5

- dilatancy angle of the soil (ψ) : 0

- magnitude of the strip load (q) : 50 kN/m^2

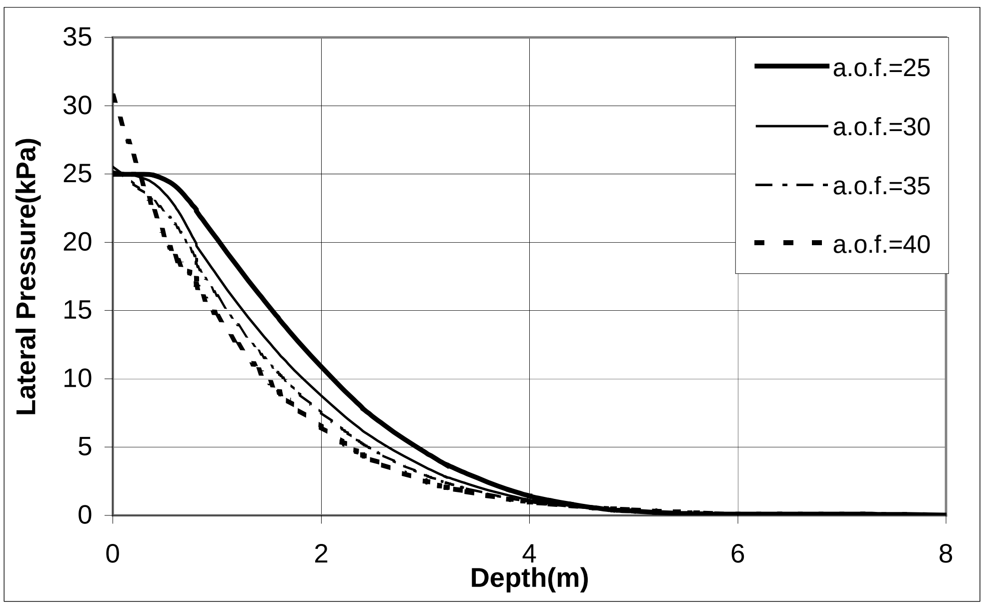


Figure 5.12 The effect of angle of friction on lateral pressure

From Fig. 5.12 it can be concluded that the internal angle of friction has a significant effect on the lateral pressure distributions. Therefore ϕ is also taken as a major effective parameter in the general solution.

5.5 SELECTION OF THE PARAMETERS

It is shown that the major effective soil parameters on lateral pressure due to surface strip load are the cohesion and the internal angle of friction. The load magnitude is found to be another parameter that should be considered because of the non-linearity in stress-strain behaviour of the soil.

The height of the wall, the distance of the strip load to the wall and the strip load width are the dimensional parameters used in the determination of the lateral pressures.

Therefore we have six parameters to consider when calculating the lateral pressures. It is decided to obtain a solution dependent on these six parameters which are :

* Height of the wall (h)
* Distance of the strip load to the wall (a)
* Width of the strip load (w)
* Magnitude of the strip load (q)
* Internal angle of friction for the soil (ϕ)
* Cohesin of the soil (c)

As discussed previously, the other required parameters for the analyses using Hardening Soil model are taken as typical values that are acceptable for a general solution. These parameters are summarized as :

- E_{50}^{ref} : 30000 kN/m^2 (for $p_{ref} = 100$)
- E_{oed}^{ref} : 30000 kN/m^2 (for $p_{ref} = 100$)
- m : 0.5
- R_f : 0.9
- ψ : 0
- γ : 20 kN/m^3

5.6 ANALYSES FOR THE GENERAL SOLUTION

To obtain a general solution for the calculation of the total lateral thrust on the wall and its point of application, artificial neural networks are proposed to be used. To obtain a solution by neural networks, an input-output pattern is given to the network and the network is trained to find the relation between the input and output data.

The input pattern consists of six parameters as discussed previously. The input data is prepared to cover sufficient cases of different material and geometrical parameters. For this purpose, the range of the parameters are set as :

- Height of the wall (h) : 2m – 10m
- Distance of the strip load (a) : 0m – 5m
- Width of the strip load (w) : 0.5m – 3m
- Magnitude of the strip load (q) : 2.5 kN/m^2 – 50 kN/m^2
- Cohesion of the soil (c) : 0 kN/m^2 – 20 kN/m^2
- Angle of friction (ϕ) : 25 – 40

When strip loads exceeding 50 kPa are applied on soils with small strength parameters (such as c = 0 and ϕ = 30), it is seen that the computer program used is unable to make the calculations due to failure in load-advancement procedure for the elements beneath the strip load. So the range of the strip load magnitude is set between 2.5 kPa and 50 kPa.

It is noted that; in the manual of the program,for bearing capacity and collapse load calculations, it is recommended to use prescribed displacements instead of prescribed loads, so prescribed displacements are increased until failure instead of load magnitudes. Using prescribed displacements is not possible in this study, because the load magnitude can not be controlled since the stress-strain behaviour of soil is non-linear and also dependent on the strength parameters (c and ϕ) which vary in each case.

Table 5.1. Considered cases and results

h	a	q	c	Φ	w		p	d
(m)	(m)	(kPa)	(kPa)	(^{o})	(m)		(kN)	(m)
2.00	0.00	5.00	0.00	30.00	1.00		2.30	0.48
2.00	0.00	20.00	10.00	40.00	1.00		5.60	0.37
2.00	0.00	50.00	15.00	35.00	1.00		16.00	0.40
2.00	1.00	10.00	5.00	25.00	1.00		2.68	1.26
2.00	1.00	25.00	20.00	30.00	1.00		5.26	1.26
2.00	1.00	50.00	10.00	40.00	1.00		11.37	1.33
2.00	2.00	20.00	15.00	35.00	1.00		3.60	1.50
2.00	2.00	35.00	12.50	37.50	1.00		6.73	1.49
2.00	2.00	12.50	2.50	30.00	1.00		3.18	1.50
2.00	3.00	25.00	5.00	35.00	2.00		8.87	1.29
2.00	3.00	2.50	0.00	30.00	3.00		1.29	1.30
2.00	5.00	30.00	20.00	25.00	2.00		9.79	1.11
2.50	1.50	42.50	7.50	27.00	0.50		6.98	1.41
3.00	0.50	17.50	2.50	27.00	0.50		4.58	0.73
3.00	1.50	32.50	12.50	25.00	1.50		16.43	1.80
3.00	2.50	47.50	7.50	32.00	1.50		20.06	2.15
3.50	0.50	32.50	17.50	25.00	1.50		19.17	1.11
4.00	0.00	2.50	0.00	25.00	1.00		1.73	0.64
4.00	0.00	40.00	20.00	35.00	1.00		20.00	0.53
4.00	0.00	20.00	10.00	30.00	1.00		11.55	0.60
4.00	1.00	30.00	5.00	32.50	1.00		13.05	1.40
4.00	1.00	10.00	12.50	27.50	1.00		3.86	1.32
4.00	1.00	5.00	0.00	25.00	1.00		2.70	1.40
4.00	2.00	25.00	20.00	37.50	1.00		6.83	2.38
4.00	2.00	50.00	17.50	40.00	1.00		14.16	2.35
4.00	2.00	35.00	5.00	30.00	1.00		13.92	2.17
4.00	3.00	10.00	5.00	30.00	3.00		7.91	2.93
4.00	3.00	30.00	15.00	25.00	2.00		17.46	2.78
4.00	5.00	20.00	10.00	30.00	2.00		8.79	2.93
5.00	0.50	42.50	2.50	27.00	1.50		41.37	1.49
5.00	1.50	7.50	17.50	35.00	2.50		5.24	2.66
5.00	2.50	17.50	7.50	30.00	0.50		3.23	2.57
6.00	0.00	10.00	0.00	25.00	2.00		13.81	1.29
6.00	0.00	5.00	15.00	25.00	3.00		6.82	1.25
6.00	1.00	15.00	20.00	35.00	1.00		5.98	1.41
6.00	1.00	40.00	15.00	35.00	2.00		31.55	1.93
6.00	2.00	25.00	10.00	30.00	3.00		28.24	3.21
6.00	2.00	30.00	0.00	30.00	1.00		16.07	2.21
6.00	3.00	35.00	0.00	30.00	3.00		42.74	3.77
6.00	5.00	25.00	5.00	25.00	2.00		16.56	4.27
6.00	5.00	20.00	5.00	35.00	3.00		16.67	4.45
6.50	2.50	47.50	12.50	30.00	2.50		45.25	3.40
7.00	3.50	42.50	12.50	25.00	0.50		9.27	3.28
7.00	2.50	42.50	12.50	25.00	1.50		29.51	2.87

Table 5.1. (continued)

7.00	1.50	17.50	7.50	27.00	2.50		19.98	2.63
7.50	1.50	27.50	7.50	32.00	0.50		7.21	1.73
8.00	0.00	7.50	15.00	25.00	1.00		5.19	0.62
8.00	0.00	50.00	2.50	27.50	1.00		43.92	1.14
8.00	0.00	22.50	7.50	35.00	1.00		16.06	0.82
8.00	1.00	12.50	2.50	32.50	1.00		8.17	1.84
8.00	1.00	27.50	0.00	40.00	1.00		17.38	1.96
8.00	1.00	37.50	17.50	37.50	1.00		19.80	1.71
8.00	2.00	5.00	20.00	27.50	1.00		2.61	2.55
8.00	2.00	25.00	0.00	32.50	1.00		15.96	2.67
8.00	2.00	35.00	10.00	37.50	1.00		18.10	2.69
8.00	3.00	10.00	5.00	35.00	2.00		8.37	4.04
8.00	3.00	2.50	5.00	25.00	3.00		3.48	4.20
8.00	5.00	50.00	10.00	30.00	3.00		53.47	5.24
9.00	0.50	2.50	17.50	32.00	0.50		0.63	0.86
9.00	1.50	32.50	2.50	30.00	2.50		45.44	2.84
9.00	2.50	7.50	17.50	35.00	1.50		4.54	3.30
10.00	0.00	10.00	20.00	30.00	2.00		11.10	1.05
10.00	1.00	15.00	0.00	30.00	3.00		25.43	2.73
10.00	1.00	45.00	5.00	25.00	2.00		59.03	2.25
10.00	2.00	25.00	0.00	25.00	2.00		30.71	2.92
10.00	2.00	40.00	15.00	30.00	1.00		19.93	2.39
10.00	3.00	40.00	15.00	30.00	3.00		50.77	4.29
10.00	3.00	10.00	10.00	35.00	2.00		8.15	4.07
10.00	5.00	45.00	5.00	35.00	3.00		51.65	5.93
10.00	5.00	15.00	10.00	30.00	2.00		12.16	5.65

Analyses are made for various values of these parameters for the preparation of data for neural network solution. Seventy different cases are considered in this study. The parameters and the calculated results for each case are given in Table 5.1. These results are proposed to be used for the determination of the neural network solution of the problem.

CHAPTER 6

NEURAL NETWORK STUDY

6.1 INTRODUCTION

The computer program " MATLAB ver. 6.0 Neural Network Tolbox " is employed for the neural network models is this study. The advantage of using this program is many types of networks are included in the program and many training algorithms with different properties can be used for a specific network model.

It is proposed to find the relationship between the input data consisting of six parameters and output data consisting of two parameters as discussed in Chapter 5. A feed-forward backpropogation type neural network is a convenient one for this case. A two-layered network having a sigmoid transfer function in the first layer (hidden layer) and a linear transfer function in the second layer (output layer) is recommended by the program manual for the approximation of any function given sufficient neurons in the hidden layer.

The modification of weights and biases to reach the best solution for the input-output pattern is performed by training of the network. So the training algorithm is very important for the solution. In this study, the Levenberg-Marquardt training algorithm (LM), which is faster and more accurate than the other backpropogation algorithms based on standart gradient descent methods, is used in the training of the networks. (See Chapter 4 Section 4.3 for backpropogation algorithms)

The basic characteristics of the neural network model used in this study can be summarized as :

* Network type : Feed-forward backpropogation

* Training algorithm : Levenberg-Marquardt algorithm (TRAINLM)

* Adaption learning function : Gradient descent with momentum (LEARNGDM)

* Performance function : Mean square error (MSE)

* Number of Layers : 2

* Transfer function (1st layer) : Sigmoid

* Transfer function (2nd layer) : Linear or sigmoid

6.2 NEURAL NETWORK MODEL

A two-layered network which has a sigmoid transfer function in the first (hidden) layer and a linear or sigmoid transfer function in the second (output) layer is proposed to be used for the data given in Table 5.1.

There are two sigmoid functions in “ Matlab Neural Network Toolbox ”. These are the log-sigmoid and the hyperbolic tangent-sigmoid transfer functions.The log-sigmoid function (logsig) takes the input and squashes it between 0 and 1. The graph and the symbol of logsig is shown in Fig. 6.1. The logsig function can be expressed as :

logsig(n) = 1 / (1+exp(-n)) (6.1)

where *n* is any value between minus and plus infinity. The other sigmoid function, hyperbolic tangent-sigmoid function (tansig) takes the input and squashes it between –1 and 1. Fig. 6.2 shows the graph and symbol of tansig function. It is expressed as :

tansig(n) = [2 / (1 + exp(-2n))] – 1 (6.2)

The linear transfer function (purelin) does not change the input, takes the input and returns the same value. The graph and symbol of purelin is shown in Fig. 6.3.

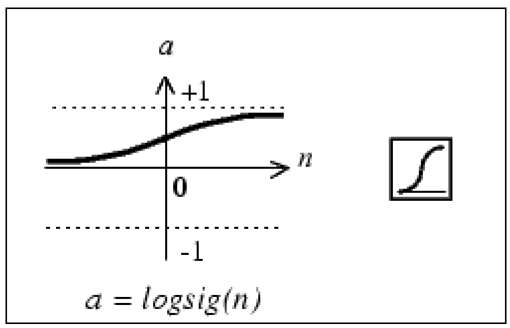


Figure 6.1 Log-sigmoid transfer function (Hagan et al. 1996)

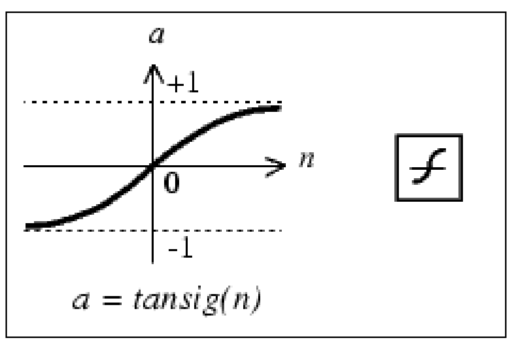


Figure 6.2 Tan-sigmoid transfer function (Hagan et al. 1996)

Figure 6.4 shows a typical neural network model used in this study for an K element input pattern and M element output pattern. In the figure, the dimensions of the vectors and matrices are shown below the symbols indicating the corresponding vector or matrix. Superscripts are used to indicate the layers which the corresponding vector or matrix is associated with.

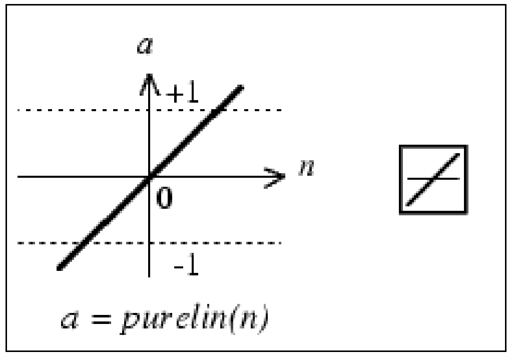


Figure 6.3 Linear transfer function (Hagan et al. 1996)

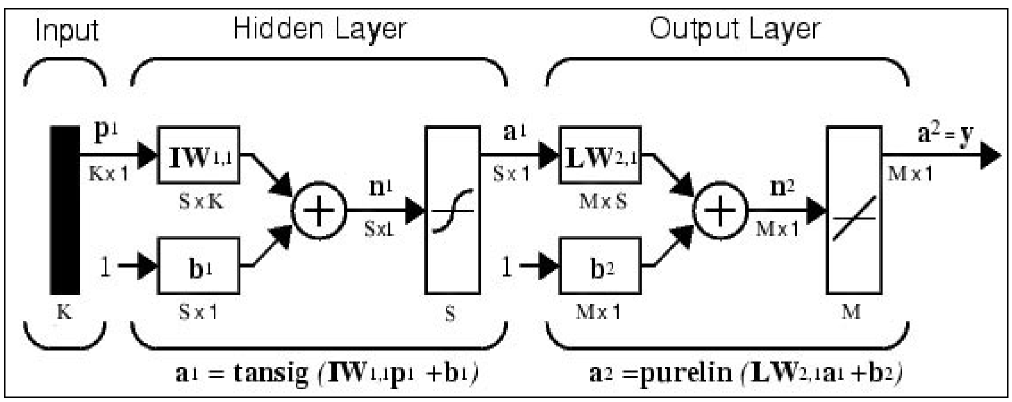


Figure 6.4 Typical feed-forward backpropogation type neural network

The input vector p^1 is represented by the solid dark vertical bar at the left. The input vector is multiplied by the input weight matrix ($IW^{1,1}$). A constant 1 is multipled by the scalar bias vector b^1. The net input to the transfer function (which is tansig for the first layer) , n^1, is the sum of the bias b^1 and the product $IW^{1,1}p^1$. This sum is passed to the transfer function to get the first layer's output vector a^1. Note that, to obtain a Sx1 ouput vector, the dimensions of $IW^{1,1}$ and b^1 should be SxK and Sx1 where the input consists of K elements. The output of the first layer, a^1, can be accepted as an input vector for the second layer. Similar to the first layer, a^1 is multiplied by the layer weight ($LW^{2,1}$) and the bias vector b^2 is added and the net input to the transfer function (which is purelin for the second layer), n^2, is obtained. The output of the second layer and the network is obtained by passing the n^2 to the transfer function. Again it is noted that, to obtain an output vector of M elements, the dimensions of $LW^{2,1}$ and b^2 are taken as MxS and Mx1 respectively.

6.3 TRAINING

Several networks with different properties are trained to find the relationship between the input-output pattern given in Table 5.1. For this purpose, networks with different transfer functions (1st tansig – 2nd purelin, 1st logsig – 2nd purelin , 1st tansig – 2nd logsig) are established and trained. Another important property of a network that effects the accuracy of the network is the number of neurons in the hidden layer (the dimensions of the weight and bias matrices in the hidden layer). So networks with various number of neurons in the hidden layer, having the same transfer functions are also trained.

The graph of a trained network is shown Fig. 6.5. 15 neurons are used in the hidden layer and the tansig and purelin transfer functions are selected for the first and second layer respectively.

The results of this network are given in Figures 6.6 and 6.7. In the figures, the error for each case (network output – actual or desired output) is given as percentage of the actual output values.

Figure 6.6 shows the error in the total lateral thrust on the wall (p). It can be seen that the maximum error is about 20% of the actual value.The error varies between -5% and 5% for except 8 cases and for only two cases an error greater than 10% is obtained. The neural network results are acceptable for 62 cases of total 70 cases but it is not acceptable for a general solution.

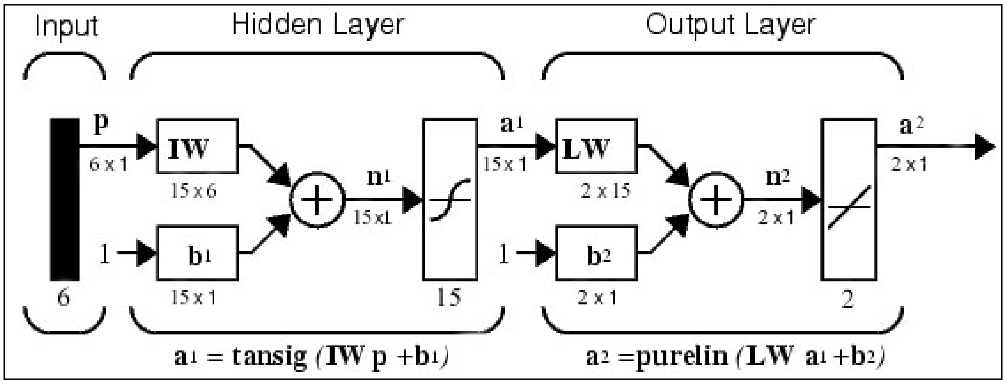


Figure 6.5 Neural network model with 15 neurons in the hidden layer

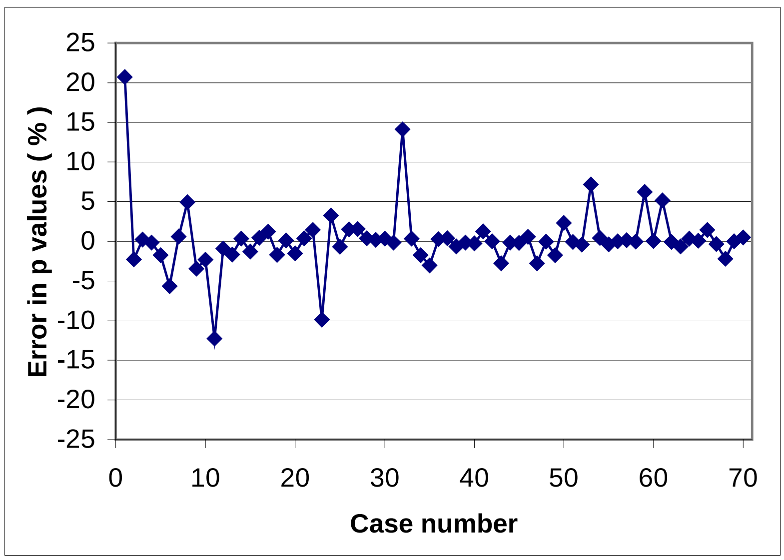


Figure 6.6 Error in *p* values as percentage of the real value

The error of the neural network results for the distance between the point of application of p and the ground surface (d) are shown in Figure 6.7. From the figure, it can be concluded that the neural network results are very different from the real values. The error is over 20% for 18 cases. Even this network gives a result having an error of –111% that means that the distance d is negative. So the neural network solution for determination of d is unrealible.

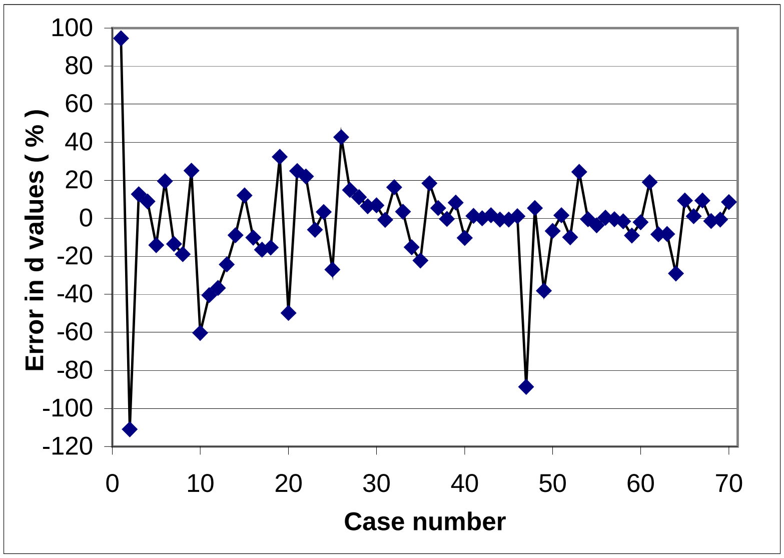


Figure 6.7 Error in *d* values as percentage of the real value

From the results of different neural networks (including the one in Fig. 6.5 and similar networks as discussed before), it is seen that the neural networks are unable to find an acceptable solution for the input-output pattern as the one in Table 5.1.

It is shown that the neural network results do not reach the desired outputs when the six input parameters (h, a, q, c, ϕ, w) and the two output parameters (p, d) are given directly as input-output data to the network. However, it may be a good idea first to modify these parameters, and then use in the network for training.
For this purpose, the dimensional parameters : a (distance of the strip load to the wall), w (width of the strip load) and d (distance betwen the point of application of the total thrust and surface) are normalized by the height of the wall (h). The total lateral thrust is also normalized by the strip load magnitude q. Therefore the six input and the two output parameters take the form as :

Input : h , a/h , q , c , ϕ , w/h

Output : p/q , d/h

The modified values of the 70 cases are given in Table 6.1

Again networks with different properties are established and trained for the modified input-ouput data given in Table 6.1. The results are much better than the ones obtained from the direct input-ouput data given in Table 5.1.

A very accurate solution is obtained for the determination of *p* and *d* using a network with 15 neurons in the hidden layer, using the tansig transfer function in the first layer and the purelin transfer function in the second layer. The graph of this network (will be called as the solution network from now on) is shown in Fig. 6.5.

The results of this network (solution network) are shown in Figures 6.8 and 6.9. Again the error values are given as percentage of the real (desired output) values.

Figure 6.8 shows the percentage of the errors in *p/q* values. It can be seen that the maximum error does not exceed 0.25% and the neural network is very succesful for the calculation of the total lateral thrust on the wall. The errors are below 0.1% for 67 cases and below 0.1% for 60 cases.

We can see in Figure 6.9 that the solution network is also succesful in the determination of the point of application. The errors in *d/h* do not exceed 1%. The errors are above 0.2% for only 16 cases of 70. Although the solution network is not as succesful in the calculation of the point of application as in the calculation of the total lateral thrust, both results have acceptable accuracy for a general solution.

Table 6.1 Modified parameters of the cases and results

h	a/h	q	c	Φ	w/h		p/q	d/h
(m)	-	(kPa)	(kPa)	(^{o})	-		-	-
2.00	0.00	5.00	0.00	30.00	0.50		0.46	0.24
2.00	0.00	20.00	10.00	40.00	0.50		0.28	0.19
2.00	0.00	50.00	15.00	35.00	0.50		0.32	0.20
2.00	0.50	10.00	5.00	25.00	0.50		0.27	0.63
2.00	0.50	25.00	20.00	30.00	0.50		0.21	0.63
2.00	0.50	50.00	10.00	40.00	0.50		0.23	0.67
2.00	1.00	20.00	15.00	35.00	0.50		0.18	0.75
2.00	1.00	35.00	12.50	37.50	0.50		0.19	0.75
2.00	1.00	12.50	2.50	30.00	0.50		0.25	0.75
2.00	1.50	25.00	5.00	35.00	1.00		0.35	0.65
2.00	1.50	2.50	0.00	30.00	1.50		0.52	0.65
2.00	2.50	30.00	20.00	25.00	1.00		0.33	0.56
2.50	0.60	42.50	7.50	27.00	0.20		0.16	0.56
3.00	0.17	17.50	2.50	27.00	0.17		0.26	0.24
3.00	0.50	32.50	12.50	25.00	0.50		0.51	0.60
3.00	0.83	47.50	7.50	32.00	0.50		0.42	0.72
3.50	0.14	32.50	17.50	25.00	0.43		0.59	0.32
4.00	0.00	2.50	0.00	25.00	0.25		0.69	0.16
4.00	0.00	40.00	20.00	35.00	0.25		0.50	0.13
4.00	0.00	20.00	10.00	30.00	0.25		0.58	0.15
4.00	0.25	30.00	5.00	32.50	0.25		0.44	0.35
4.00	0.25	10.00	12.50	27.50	0.25		0.39	0.33
4.00	0.25	5.00	0.00	25.00	0.25		0.54	0.35
4.00	0.50	25.00	20.00	37.50	0.25		0.27	0.60
4.00	0.50	50.00	17.50	40.00	0.25		0.28	0.59
4.00	0.50	35.00	5.00	30.00	0.25		0.40	0.54
4.00	0.75	10.00	5.00	30.00	0.75		0.79	0.73
4.00	0.75	30.00	15.00	25.00	0.50		0.58	0.70
4.00	1.25	20.00	10.00	30.00	0.50		0.44	0.73
5.00	0.10	42.50	2.50	27.00	0.30		0.97	0.30
5.00	0.30	7.50	17.50	35.00	0.50		0.70	0.53
5.00	0.50	17.50	7.50	30.00	0.10		0.18	0.51
6.00	0.00	10.00	0.00	25.00	0.33		1.38	0.22
6.00	0.00	5.00	15.00	25.00	0.50		1.36	0.21
6.00	0.17	15.00	20.00	35.00	0.17		0.40	0.24
6.00	0.17	40.00	15.00	35.00	0.33		0.79	0.32
6.00	0.33	25.00	10.00	30.00	0.50		1.13	0.54
6.00	0.33	30.00	0.00	30.00	0.17		0.54	0.37
6.00	0.50	35.00	0.00	30.00	0.50		1.22	0.63
6.00	0.83	25.00	5.00	25.00	0.33		0.66	0.71
6.00	0.83	20.00	5.00	35.00	0.50		0.83	0.74
6.50	0.38	47.50	12.50	30.00	0.38		0.95	0.52
7.00	0.50	42.50	12.50	25.00	0.07		0.22	0.47
7.00	0.36	42.50	12.50	25.00	0.21		0.69	0.41

Table 6.1 (continued)

7.00	0.21	17.50	7.50	27.00	0.36		1.14	0.38
7.50	0.20	27.50	7.50	32.00	0.07		0.26	0.23
8.00	0.00	7.50	15.00	25.00	0.13		0.69	0.08
8.00	0.00	50.00	2.50	27.50	0.13		0.88	0.14
8.00	0.00	22.50	7.50	35.00	0.13		0.71	0.10
8.00	0.13	12.50	2.50	32.50	0.13		0.65	0.23
8.00	0.13	27.50	0.00	40.00	0.13		0.63	0.25
8.00	0.13	37.50	17.50	37.50	0.13		0.53	0.21
8.00	0.25	5.00	20.00	27.50	0.13		0.52	0.32
8.00	0.25	25.00	0.00	32.50	0.13		0.64	0.33
8.00	0.25	35.00	10.00	37.50	0.13		0.52	0.34
8.00	0.38	10.00	5.00	35.00	0.25		0.84	0.51
8.00	0.38	2.50	5.00	25.00	0.38		1.39	0.53
8.00	0.63	50.00	10.00	30.00	0.38		1.07	0.66
9.00	0.06	2.50	17.50	32.00	0.06		0.25	0.10
9.00	0.17	32.50	2.50	30.00	0.28		1.40	0.32
9.00	0.28	7.50	17.50	35.00	0.17		0.61	0.37
10.00	0.00	10.00	20.00	30.00	0.20		1.11	0.11
10.00	0.10	15.00	0.00	30.00	0.30		1.70	0.27
10.00	0.10	45.00	5.00	25.00	0.20		1.31	0.23
10.00	0.20	25.00	0.00	25.00	0.20		1.23	0.29
10.00	0.20	40.00	15.00	30.00	0.10		0.50	0.24
10.00	0.30	40.00	15.00	30.00	0.30		1.27	0.43
10.00	0.30	10.00	10.00	35.00	0.20		0.82	0.41
10.00	0.50	45.00	5.00	35.00	0.30		1.15	0.59
10.00	0.50	15.00	10.00	30.00	0.20		0.81	0.57

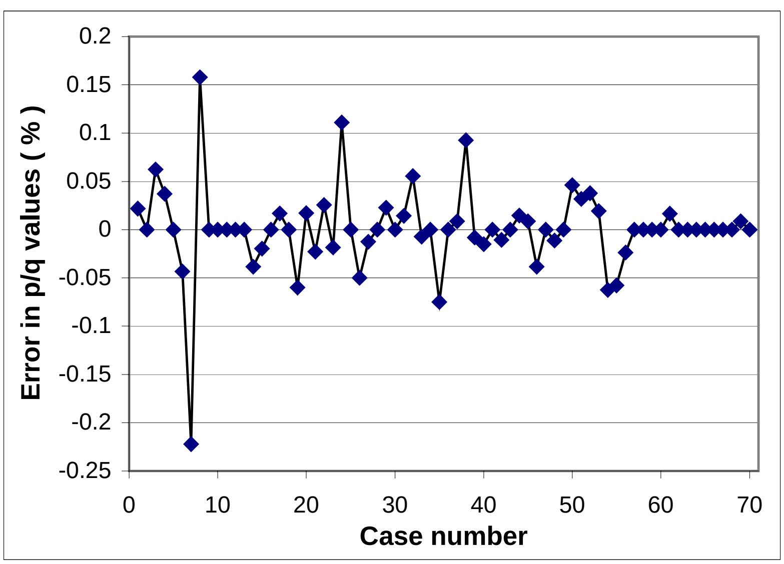


Figure 6.8 Error in *p/q* values for the solution network

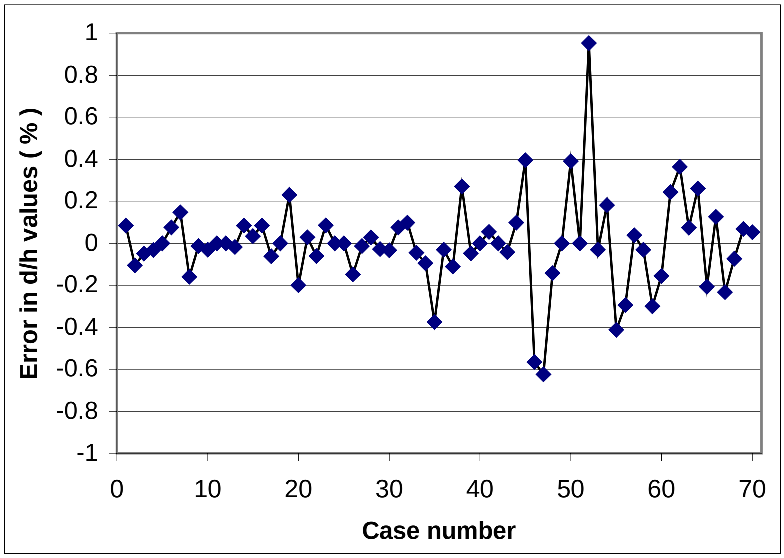


Figure 6.9 Error in *d/h* values for the solution network

Table 6.2 gives a summary of the errors of the two identical networks one of which is trained using the direct input-output parameters and the other trained using modified input-output data.

Table 6.2 Summary of the network results

	Determination of p-(p/q)		Determination of d-(d/h)	
	Direct data	Modified data	Direct data	Modified data
Maximum error (%)	20.69	0.22	111	0.95
Average error (%)	2	0.02	16.9	0.14

It is clear from Table 6.2 that the errors for the case of direct input of the parameters as in Table 5.1 are about 100 times the errros for the case of the modification of the data as in Table 6.1.

Therefore we can say the response of the neural networks can be very different for two numerically different input-output pattern even they represent the same data and are identical from an engineering point of view.

6.4 CLOSED FORM SOLUTION

The solution obtained by neural network shown in Fig. 6.5 can be expressed as :

$$T = LW * (\text{tansig}(IW*K + B1)) + B2 \qquad (6.3)$$

where * denotes the scalar multiplication of two matrices. The input vector, K, consists of the 6 input parameters (knowns) and by using the above equation, the output vector,T, consisting of the two output parameters (unknowns) can be calculated. The input vector K and the output vector T are :

$$K = \begin{bmatrix} h \\ a/h \\ q \\ c \\ \phi \\ w/h \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} p/q \\ d/h \end{bmatrix}$$

where :

h : Height of the wall (m)

a : Distance of the strip load to the wall (m)

w : Strip load width (m)

q : Strip load magnitude (kPa)

c : (Effective) cohesion of the soil (kPa)

ϕ : (Effective) angle of friction of the soil (in degrees)

p : Total lateral force on the wall due to only strip load (kN)

d : Distance between the point of application of P and the ground surface (m)

The weight and bias matrices required for the solution (LW, IW, B1 and B2) are given in Appendix A.

The expression of tansig function is given by Equation 6.2.

Therefore a closed-form solution is obtained for the calculation of the total lateral thrust and its point of application as given by Equation 6.3. This solution is valid under the assumptions given in Section 5.1. Also it is noted that; to obtain accurate and reasonable results from this neural network solution, the input parameters must lie between the range of the parameters used in the training of the network, as given in Section 5.6. For example, it is not recommended to use this solution for a strip load width of 5m, because that the input-ouput pattern used in the solution consists of width parameters between 0.5m and 3m.

6.5 NEURAL NETWORK RESULTS

6.5.1 CHECK OF NETWORK RESULTS FOR DIFFERENT CASES

It is shown that the obtained neural network solution gives very accurate results for the considered 70 cases. However, the response of the network in any case excluding the considered 70 ones is also investigated to see the validity of the solution.

For this purpose, 4 cases different than the considered 70 ones used in the solution are considered. The parameters for the cases are given in Table 6.3.

These 4 cases are solved by both the solution network and PLAXIS and *p* and *d* values are obtained. The results of *p*, *d* and the moment at the base due to the total lateral thrust *p* (that can be calculated as : moment = p x (h-d)) are shown in Table 6.4 for the both solution methods.

Table 6.3 Input parameters for the considered 4 cases

case	**h**(m)	**a**(m)	**q**(kPa)	**c**(kPa)	**Φ**(°)	**w**(m)
1	3	1.5	15	5	30	1.5
2	5	2.5	25	15	35	2.5
3	7	3	40	10	25	1
4	9	1	30	0	30	2

Table 6.4 Results for two solution methods

	neural network results			**plaxis results**		
case	**p**(kN)	**d**(m)	**Moment**(kN.m)	**p**	**d**	**Moment**(kN.m)
1	7.0	2.0	14.0	7.1	1.9	13.5
2	15.3	3.5	53.6	18.1	3.3	59.7
3	20.6	3.2	65.9	18.7	3.1	58.0
4	41.2	2.0	82.4	38.1	2.3	87.6

It can be seen from Table 6.4 that the neural network results are very similar to the results calculated from PLAXIS analyses. So it is concluded that the neural network solution given by Equation 6.3 is acceptable for any case excluding the considered ones used in the training of the neural network.

6.5.2 THE EFFECT OF INPUT PARAMETERS BY NETWORK SOLUTION

The individual effect of input parameters a, w, q, c and ϕ are investigated by using the obtained neural network solution that is also given by Equation 6.3. To investigate the effect of each input paramater, the other paramaters are held constant and several cases are solved for various values of the specific parameter.

The results are also calculated according to the linear elastic solution and plotted on the same graphs with the neural network solutions.

To investigate the effect of the distance of the strip load (a), the other parameters are held constant as : h = 6, w = 1, q = 25, c = 10, ϕ = 30 and the results are obtained for different a values. The results are shown in Figures 6.10 and 6.11.

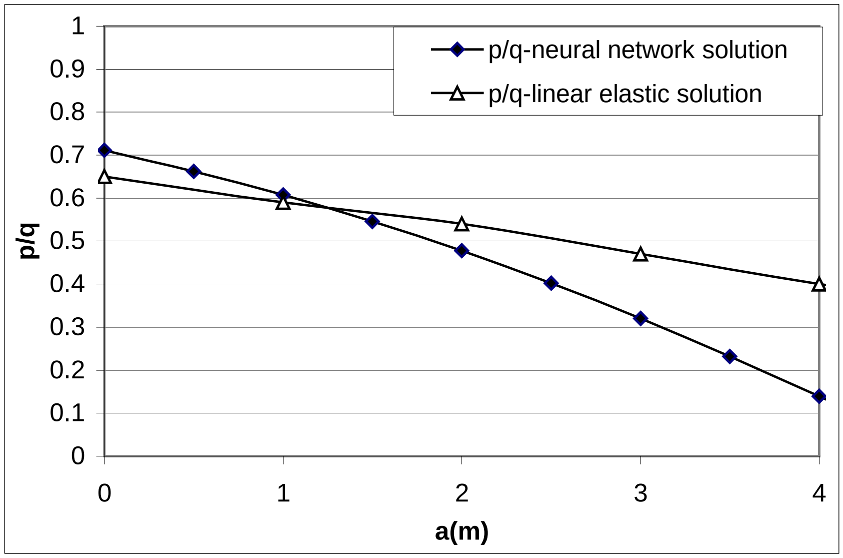


Figure 6.10 The effect of strip load distance (a) on *p/q*

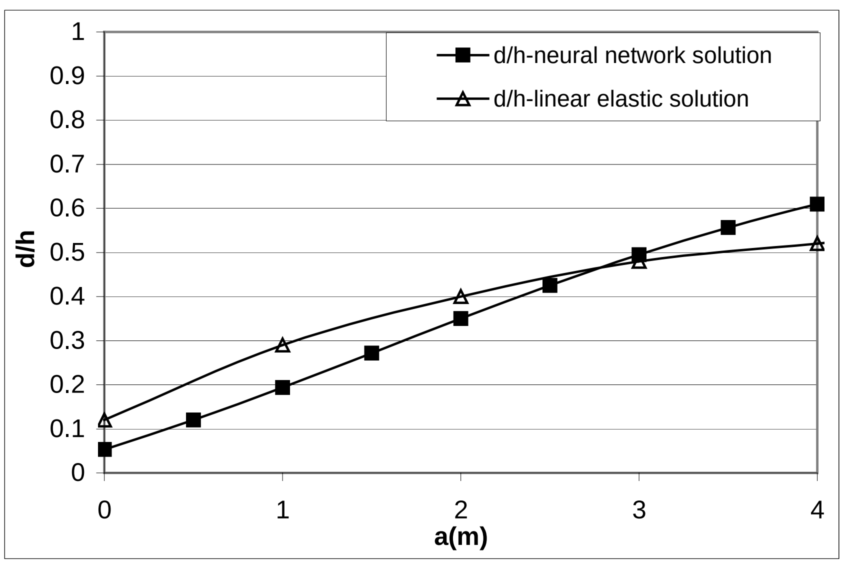


Figure 6.11 The effect of strip load distance (a) on *d/h*

It can be seen from the figures that *p/q* decreases and *d/h* increases with incresing *a* as expected. The linear elastic solution and the neural network solution give very different results. Especially in *p/q* values, for $a > 2$ m the linear elastic solution gives much higher results than the neural network solution.

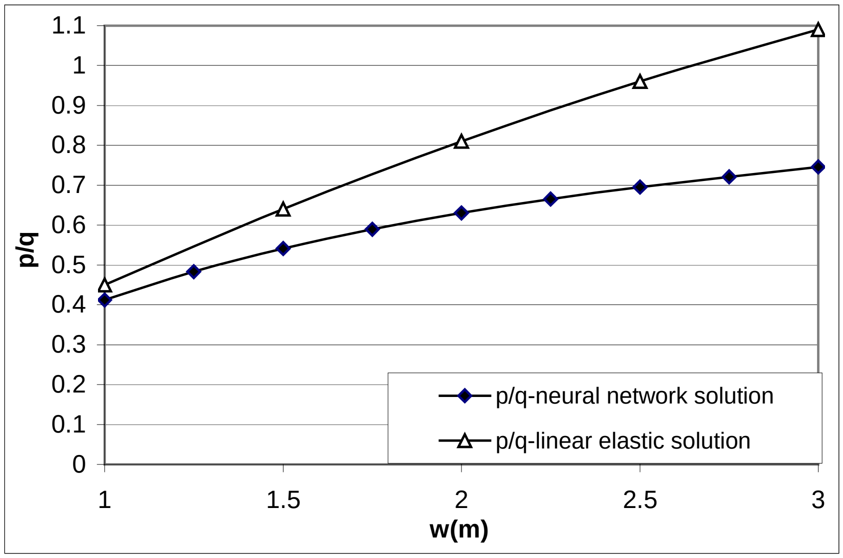


Figure 6.12 The effect of strip load width (*w*) on *p/q*

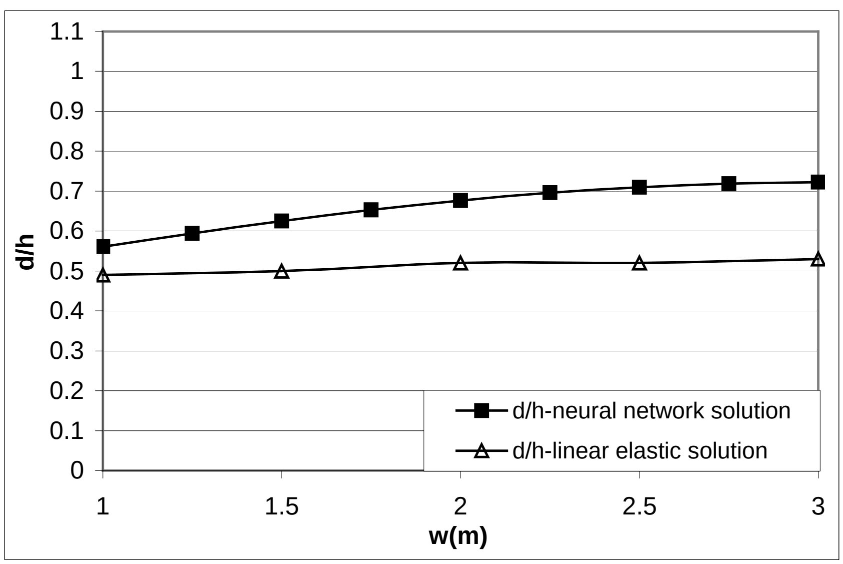


Figure 6.13 The effect of strip load width (*w*) on *d/h*

Figures 6.12 and 6.13 show the effect of strip load width (w) on *p/q* and *d/h* for the case : *h*=4, *a*=2, *q*=25, *c*=10, ϕ =30. It can be concluded that the results are very different . The linear elastic solution gives greater *p/q* values especially for larger values of *w*. On the other hand, the neural network results are greater for *d/h* values.

The effect of *q* on *p/q* and *d/h* is investigated for the case: *h*=6, *a*=2, *w*=2, *c*=10 and ϕ=30. As discussed before, the load magnitude has no effect on the normalized value *p/q* and *d* for the linear elastic solution.

The results are shown in Fig. 6.14. It is clear from the figure that the neural network results are very similar to linear elastic solution adn nearly constant for *d/h*. However, there is a significant decrease of *p/q* for increasing *q* in the neural network solution while it is constant for linear elastic solution.

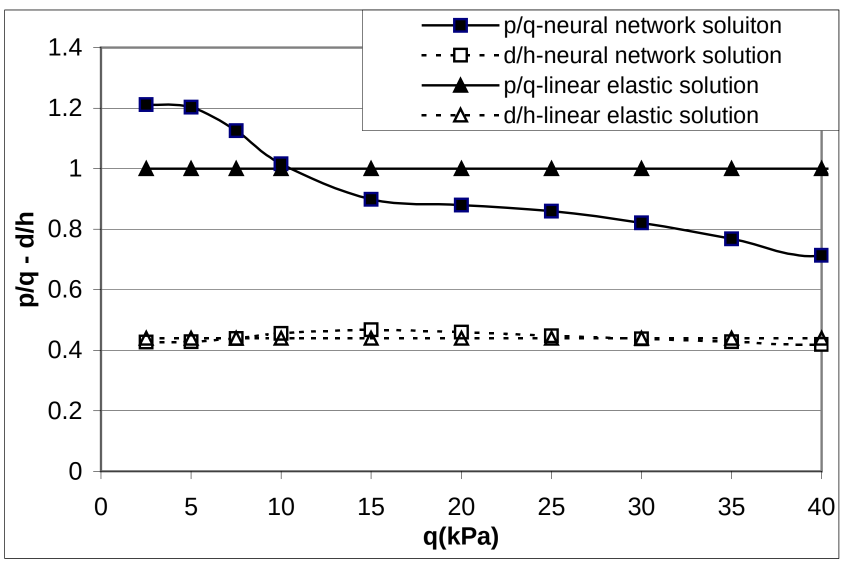


Figure 6.14 The effect of strip load magnitude (*q*) on *p/q* and *d/h*

The effect of cohesion (*c*) while the other input parameters are held constant as : *h*=4, *a*=2, *w*=2, *q*=25, ϕ=30 is shown in Fig. 6.15. It can be seen that the *d/h* is nearly constant for varying cohesion, but the value of *d/h* obtained by neural network is significantly greater than the one obtained according to linear elasticity. It is also clear that the *p/q* significantly decreases for increasing cohesion and the linear elastic solution gives much higher values for greater values of cohesion.

The effect of angle of friction is very similar to cohesion as shown in Fig. 6.16. To investigate the effect of angle of friction, the case : *h*=5, *a*=1.5, *w*=2, *q*=25, *c*=5 is considered. Similar to cohesion results, the *d/h* values are nearly constant and greater than the linear ones obtained by linear elastic solution, and there is a significant decrease in *p/q* for increasing angle of friction resulting in a big difference between the neural network and elastic solutions.

Therefore it can be concluded that the cohesion and angle of friction have neglicible effect on the shape of the lateral pressure distribution, but the pressure values decrease with increasing soil strength (cohesion or angle of friction).

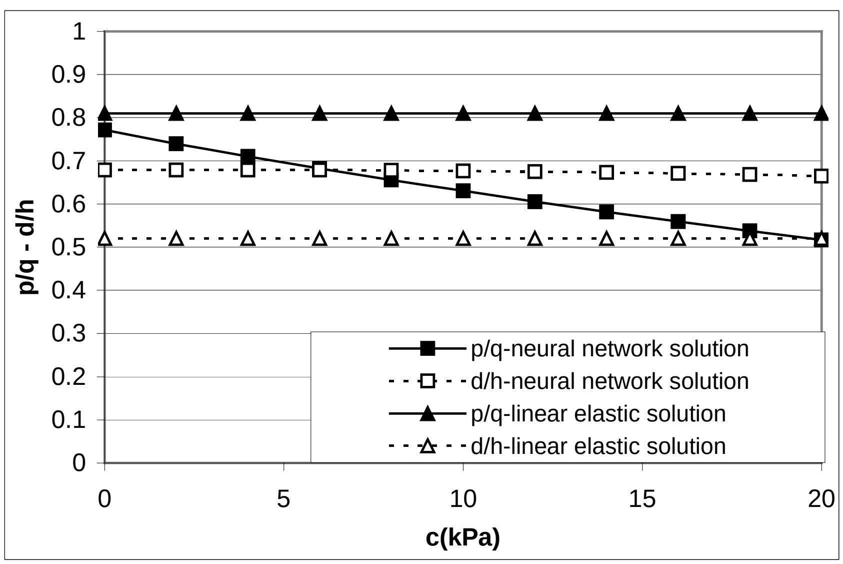
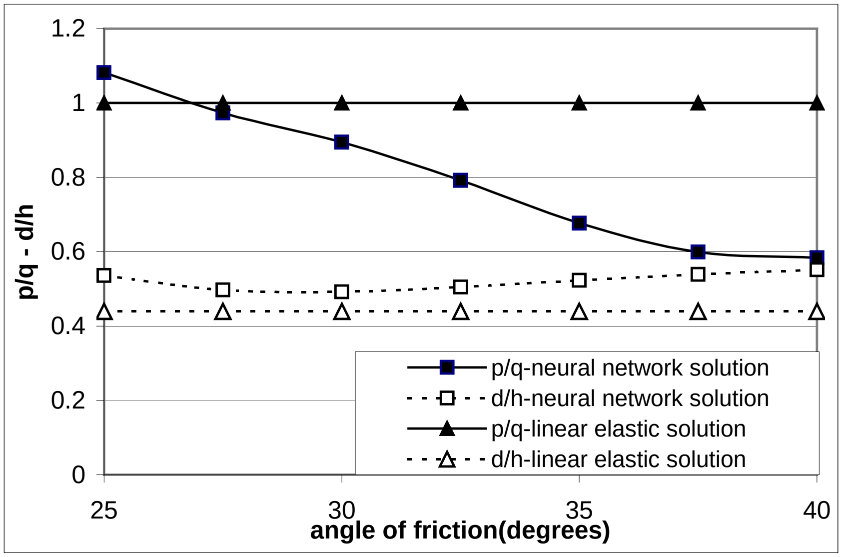


Figure 6.15 The effect of cohesion (*c*) on *p/q* and *d/h*



The effect of angle of friction on *p/q* and *d/h*

CHAPTER 7

CONCLUSIONS

In this study, an investigation of the lateral pressures acting on rigid retaining walls due to surface strip loading has been made.

Analyses are made by a finite element program, “PLAXIS ver.7.11” and the soil is modelled as a non-linear elasto-plastic material. The effects of the material parameters used to represent the stress-strain relationship of the soil, on lateral pressures due to surface strip loading are investigated.

It is concluded that shear strength parameters (cohesion and angle of friction) of the soil are major effective parameters affecting lateral pressures which can be determined by conventional laboratory or in-situ tests.

A closed-form solution is obtained for the calculation of the total lateral thrust and its point of application, as a function of six parameters given below :

- Height of the wall (h)
- Distance of the strip load (a)
- Width of the strip load (w)
- Magnitude of the strip load (q)

- Cohesion of the soil (c)
- Angle of friction (ϕ)

For this purpose, 70 cases including various values of these parameters are considered and analysed by the finite element program. The total lateral thrust and distance of its point of application to the surface are calculated for each case.

Artificial neural networks are used to find a closed from solution. The artificial neural network is trained by inputting the results obtained from the solutions of above mentioned 70 cases. Various combinations of input parameters are tried in order to reach a solution within 1% accuracy.

A check of the obtained solution is made for the cases excluding the 70 cases used in the training of the network. It is seen that the closed from solution has sufficient accuracy for these additional cases.

The effects of input parameters on the total lateral thrust and the point of application are investigated individually by the obtained solution and the results are compared with the linear elastic method. This investigation has led to the following conclusions:

- The shear strength of the soil has a considerable effect on the total lateral thrust. An increase in the shear strength parameter (cohesion or angle of friction) results in a significant decrease in the total lateral thrust. The linear elastic solution is independent of the change in shear strength and gives relatively higher values than the neural network solution.

- Although the shear strength parameters affect the total lateral thrust, they have negligible effect on the distance of the point of application to the surface. This shows that the shear strength of the soil does not change the shape of the lateral pressure distribution but affects the lateral pressure magnitudes due to surface strip loading.

- The results of the linear elastic solution for the total lateral thrust are generally higher than the results obtained from the neural network solution. The difference increases as strip load width, cohesion and friction angle increase.

- The distance of the point of application to the surface determined by the neural network solution is generally greater than the distance obtained by the linear elastic solution.

The neural network solution is valid for the assumptions given in Section 5.1. In the future, new geometries or assumptions can be investigated. For example, a flexible wall instead of a rigid wall can be considered and the effect of the stiffness of this wall can be studied. Another approach may be the investigation of the effect of K_0, thus initial stresses.

REFERENCES

1. Britto, A. M., and Gunn, M. J. (1987), Critical State Soil Mechancis Via Finite Elements, Ellis Horwood Ltd.

2. Cernica, J. N. (1995), Geotechnical engineering: Foundation and Design, John Wiley & Sons, Inc.

3. Chen, W., and Mizuno, E. (1990), Nonlinear Analysis in Soil Mechanics, Elsevier Science Publishers.

4. Coduto, D.P. (2001), Foundation Design Principles and Practices, 2'nd Edition, Prentice Hall.

5. Craig, R. F. (1992), Soil Mechanics, Chapman & Hall

6. Dayhoff, J.E. (1990), Neural Network Architectures, Van Nostrand Reinhold Press.

7. Duncan, J. M., and Chang, C. Y. (1970), " Nonlinear analysis of stress and strain in soil." J. Of the Soil Mech. And Found. Div., ASCE, 96, 1629-1653.

8. Duncan, J. M., Byrne, P., Wong, K.S., and Babry, P. (1980), "Strength, stress-strain and bulk modulus parameters for finite element analyses of stresses and movements in soil masses." Report No: UCB/GT/80-01, University of California, Berkeley.

9. Georgiadis, M., Anagnostopoulos, C. (1998), "Lateral pressure on sheet pile walls due to strip load." J. Geotech. and Geoenv. Engrg., ASCE, 124, 95-98.

10. Hagan, M. T., and Menhaj, M. (1994), " Training feedforward networks with the Marquardt algorithm." IEEE Transactions on Neural Networks, Vol:5, no:6, 989-993

11. Hagan, M. T., Demuth, H. B., and Beale, M. H. (1996), Neural Network Design, Boston, MA: PWS Publishing.

12. Jarquio, R. (1981), " Total lateral surcharge pressure due to strip load" J. Geotech. Engrg., ASCE, 107, 1424-1428.

13. Kartam, N., Garret, J. H., and Flood, I. (1997), Artificial Neural Networks for Civil Engineers: Fundamentals and Applications, ASCE Press.

14. Konder, R. L. (1963), "Hyperbolic Stress-Strain Response: Cohesive Soils," Journal of the Soil Mechanics and Foundations Division, ASCE, 89(SM1), 115-143.

15. Konder, R. L., and Zelasko, J. S. (1963), "A Hyperbolic Stress-Strain Formulation for sands," Proceedings of the 2nd Pan American Conference on Soil Mechanics and Foundation Engineering, 1, 289-324.

16. MATLAB ver. 6.0, Neural Network Toolbox Manual

17. McClelland, J. L., Rumelhart, D. E. (1988), Explorations in Parallel Distributed Processing, The MIT Press.

18. Misra, B. (1981). “Lateral pressures on retaining walls due to loads on surfaces of granular backfills.” Soils and Foundations., Tokyo, Japan, 20(2), 33-44.

19. Motta, E. (1994). “Generalized Coulomb active earth pressure for distanced surcharge.” J. Geotech. Engrg., ASCE, 120, 1072-1079.

20. PLAXIS ver. 7 - Material Models Manual.

21. Poulos, H.G, and Davis, E. H. (1974), Elastic Solutions for Soil and Rock Mechanics, John Wiley and Sons, New York.

22. Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986), Parallel Distributed Processing, Vol:1, The MIT Press.

23. Schanz, T., Vermeer, P. A. And Bonnier, P. G. (1999), “The hardening soil model : formulation and verification.” Beyond 2000 in Computational Geotechnics-10 years of Plaxis, Balkema, Rotterdam.

24. Steenfelt, S., and Hansen, B. (1983), “Disc. of Total lateral surcharge pressure due to strip load” J. Geotech. Engrg., ASCE, 109, 271-273.

25. Toll, D. (1996), “Artificial intelligence systems for geotechnical engineering with specific reference to ground improvement.” Tenth European Young Geotechnical Engineers Conference, İzmir, Turkey.

26. Wood, D. M. (1990), Soil Behaviour and Critical State Soil Mechanics, Cambridge University Press.

APPENDIX A

WEIGHT AND BIAS MATRICES OF THE SOLUTION NETWORK

$$
IW = \begin{bmatrix}
0.707584 & 1.669341 & -0.012751 & 0.096040 & -0.344573 & -0.006084 \\
0.290101 & -0.285203 & -0.020045 & -0.016270 & 0.025231 & 1.034413 \\
0.025184 & 1.675913 & 0.003187 & -0.011426 & -0.013588 & 0.989599 \\
-0.218160 & -0.334959 & -0.017237 & 0.027217 & 0.123004 & 0.904165 \\
0.236921 & 1.428380 & -0.062918 & 0.054190 & -0.064349 & -3.940876 \\
0.401155 & -0.765628 & -0.009593 & -0.025510 & 0.010007 & -0.210672 \\
0.043779 & 0.449645 & 0.015901 & -0.013441 & -0.074459 & -0.846834 \\
-0.507689 & -0.227599 & -0.008705 & 0.016337 & -0.015705 & 1.468541 \\
-1.375848 & -5.436067 & 0.646734 & 0.134772 & 0.119914 & 18.519199 \\
0.363906 & 2.833471 & -0.186603 & 0.100430 & -0.101867 & -3.190632 \\
-0.418049 & 0.099324 & -0.020282 & 0.065495 & 0.154286 & -0.840888 \\
0.512810 & 0.166012 & 0.007541 & -0.016510 & 0.021023 & -1.449474 \\
0.355588 & -0.228980 & -0.009691 & -0.008272 & 0.115341 & 0.026844 \\
-3.261537 & -1.695312 & 0.612241 & -0.075086 & 0.599992 & -15.115130 \\
1.586981 & -2.223672 & 0.079581 & -0.178878 & 0.927314 & 8.396568
\end{bmatrix}
$$

$$
B1 = \begin{bmatrix}
-0.387163 \\ -2.698834 \\ -0.234439 \\ -2.926085 \\ 0.829085 \\ -2.426979 \\ 1.867292 \\ 1.387967 \\ -29.263770 \\ 1.504808 \\ -0.598197 \\ -1.504389 \\ -3.650848 \\ 9.602564 \\ -31.920627
\end{bmatrix}
\qquad
B2 = \begin{bmatrix} 11.043619 \\ 1.207034 \end{bmatrix}
\qquad
LW = \begin{bmatrix} LW1 & LW2 & LW3 \end{bmatrix}
$$

$$
LW1 = \begin{bmatrix}
10.590286 & -0.014567 & 0.017771 & -1.880257 & -0.605940 \\
0.861373 & -0.431648 & 0.722998 & -0.114303 & 0.133397
\end{bmatrix}
$$

$$
LW2 = \begin{bmatrix}
-0.678839 & -3.097798 & -7.987459 & 0.117518 & 0.376596 \\
0.289137 & -0.579776 & 4.443063 & 0.028310 & -0.066399
\end{bmatrix}
$$

$$
LW3 = \begin{bmatrix}
-1.154965 & -7.790873 & -0.459319 & 0.223971 & 0.252342 \\
-0.088841 & 4.795814 & -0.597022 & -0.001774 & 0.081540
\end{bmatrix}
$$