REVIEW OF SOIL MODELS AND THEIR IMPLEMENTATION IN
MULTIBODY SYSTEM ALGORITHMS

Ulysses Contreras\textsuperscript{1}
Guangbu Li\textsuperscript{2}
Craig D. Foster\textsuperscript{3}
Ahmed A. Shabana\textsuperscript{1}
Paramsothy Jayakumar\textsuperscript{4}
Michael D. Letherwood\textsuperscript{4}

\textsuperscript{1} Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, 842 West Taylor Street, Chicago, Illinois 60607
\textsuperscript{2} Department of Mechanical Engineering, Shanghai Normal University, 100 Guilin Road, Shanghai, 200234
\textsuperscript{3} Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 West Taylor Street, Chicago, Illinois 60607
\textsuperscript{4} U.S. Army RDECOM-TARDEC, 6501 East 11 Mile Road, Warren, MI 48397-5000
The mechanical behavior of soils may be approximated using different models that depend on particular soil characteristics and simplifying assumptions. For this reason, researchers have proposed and expounded upon a large number of constitutive models and approaches that describe various aspects of soil behavior. However, there are few material models capable of predicting the behavior of soils for engineering applications and are at the same time appropriate for implementation into finite element (FE) and multibody system (MBS) algorithms. This paper presents a survey of different commonly used terramechanics and continuum-based soil models. The aim is to provide a summary of soil models, compare them, and examine their suitability for integration with large-displacement FE absolute nodal coordinate formulation (ANCF) and MBS algorithms. Special emphasis is placed on the formulations of soils used in conjunction with vehicle dynamic models. A brief review of computer software used for soil modeling is provided and the implementation of these soil models in MBS algorithms used in the analysis of complex vehicle systems is discussed.
ABSTRACT

The mechanical behavior of soils may be approximated using different models that depend on particular soil characteristics and simplifying assumptions. For this reason, researchers have proposed and expounded upon a large number of constitutive models and approaches that describe various aspects of soil behavior. However, there are few material models capable of predicting the behavior of soils for engineering applications and are at the same time appropriate for implementation into finite element (FE) and multibody system (MBS) algorithms. This paper presents a survey of different commonly used terramechanics and continuum-based soil models. The aim is to provide a summary of soil models, compare them, and examine their suitability for integration with large-displacement FE absolute nodal coordinate formulation (ANCF) and MBS algorithms. Special emphasis is placed on the formulations of soils used in conjunction with vehicle dynamic models. A brief review of computer software used for soil modeling is provided and the implementation of these soil models in MBS algorithms used in the analysis of complex vehicle systems is discussed.

Keywords: Soil; Finite element; Multibody systems; Terramechanics.
1. **INTRODUCTION**

The characteristics of soils, as with any other material, depend on the loading and the soil conditions. The response of the soil model to loading conditions depends on the assumptions used in and the details captured by the specific model. Some models are based on simple discrete elastic models that do not capture the soil distributed elasticity and inertia. More detailed soil models employ a continuum mechanics approach that captures the soil elastic and plastic behaviors. Continuum mechanics-based soil models can be implemented in finite element (FE) algorithms. Nonetheless, the integration of these FE soil models with multibody system (MBS) algorithms for modeling vehicle/soil interaction represents a challenging implementation and computational problem that has not been adequately covered in the literature. This integration is necessary in order to be able to develop more detailed and more accurate vehicle/terrain dynamic interaction models.

Depending on the level of detail that needs to be considered in a soil investigation, the parameters that define the soil in a computer model can significantly vary. However, among the many different characteristics of soil behavior, there are a few that must be considered in a soil model. These characteristics are summarized as follows:

1. **Shear strength and deformation characteristics:** The mean stress and change in volume produced by shearing greatly affects the shear strength and deformation characteristics of soil. Soils generally exhibit higher shear strength with increasing mean stress (applied pressure) due to interlocking effects. At very high mean stresses, however, soils may fail or yield due to pore collapse, grain crushing, or other phenomena. The dilatation of soil under shear loading is shown in Fig. 1a. Sand demonstrates interlocking behavior that increases with a corresponding increase in the density of soil.
2. **Plasticity**: An increase of applied stress beyond the elastic limit results in an irrecoverable deformation which often occurs without any signs of cracking or failure. A small elastic region which results in plastic behavior at or near the onset of loading is characteristic of many soils.

3. **Strain-hardening/softening**: This soil characteristic can be defined as change in the size, shape, and location of the yield surface. This can be identified graphically as shown in Fig. 1 (Maugin, 1992). The dilatation of dense granular material, such as sand, and over-consolidated clays is commonly associated with the strain-softening behavior. Likewise, the compaction of loose granular material, such as sand, and normally consolidated clays is commonly associated with the strain-hardening behavior (Fig. 1).

Other characteristics of soil such as tensile strength, temperature-dependency, and drainage effects, etc. are not considered here because they are beyond the scope of this review paper.

This paper aims to review some of the existing basic terramechanics and continuum mechanics based soil models and discuss their suitability for incorporation into FE/MBS simulation algorithms. Section 2 outlines the empirical, analytical, and parametric approaches used in terramechanics. Also reviewed are some of the tools and methodologies which determine the parameters used in the definitions of the terramechanics models. Section 3 describes the continuum mechanics based soil models. These models include elastic-plastic, viscoplastic, and bounding surface plasticity formulations. Section 4 describes three of the most popular particle-based and meshfree methods; the discrete element method, smoothed particle hydrodynamics, and reproducing kernel particle method. Examples of computer programs used to model the behaviors of soils are presented in Section 5. Section 6 offers a comparison of the various soil models presented in the previous sections and a suitable soil model for implementation in a
FE/MBS algorithm is selected. The computer implementation of the selected soil model and solution procedure is outlined in Section 7. Section 8 describes the procedure for the incorporation of the selected soil model with an ANCF/MBS formulation. In Section 8, the structure of the dynamic equations that allows for systematically integrating soil models with FE/MBS system algorithms used in the virtual prototyping of vehicle systems is presented. Section 9 offers a summary and describes the direction of future work.

2. TERRAMECHANICS-BASED SOIL MODELS

Terramechanics is the study of the relationships between a vehicle and its environment. Some of the principal concerns in terramechanics are developing functional relationships between the design parameters of a vehicle and its performance with respect to its environment, establishing appropriate soil parameters, and promoting rational principles which can be used in the design and evaluation of vehicles (Wong, 2010). The standard parameters by which a vehicle performance is compared include drawbar-pull, tractive efficiency, motion resistance, and thrust. If the normal and shear stress distributions at the running gear-soil interface are known, then these parameters are completely defined.

2.1 Empirical Terramechanics Models

One approach used to establish the appropriate parameters, properties, and behaviors of soil involves the determination of empirical relationships based on experimental results which can be used to predict at least qualitatively the response of soils under various conditions (Bekker, 1969). Concerns were raised as to whether the relationships established by this method could be applied in circumstances which were entirely dissimilar to those in which they were established (Bekker, 1969). Bekker proposed using only experiments that realistically simulated the manner
in which the running gear of a vehicle traversed the terrain. This entailed using soil penetration plates comparable in size to the contact patch of a tire (or track), and producing pressures and shear forces of comparable magnitude to those produced by a vehicle. Parametric models, which are based on experimental work and have been widely used, offer practical means by which an engineer can qualitatively evaluate tracked vehicle performance and design. Using these principles, Bekker developed the Bevameter. When a tire or a track traverses a terrain, soil is both compressed and sheared. A Bevameter measures the terrain's response to normal and shear stresses by the application of penetration plates and shear heads. These responses are then used to produce pressure-sinkage and shear stress-shear displacement curves. These curves are then taken as characteristic response curves for each type of terrain.

Another terrain characterization device of importance (due to its widespread use) is the cone penetrometer. A penetrometer applies simultaneously shear and normal stresses. A simplified version of a penetrometer can be visualized as a long rod with a right circular cone on one end. Penetrometers are pushed (at a certain rate) into the soil and the resulting force per unit cone base area, called the cone index (CI), is measured. These cone indices can then be used to establish the trafficability, on a one or fifty pass basis, of vehicles in different types of terrain (Priddy and Willoughby, 2006). Trafficability is the measure of a vehicle's ability to traverse terrain without becoming incapacitated. Hence, a vehicle with a CI on a fifty pass basis can be expected to make fifty passes on a particular route without becoming incapacitated. It is important to note that individual soil parameters cannot be derived from cone penetration tests. It has been established that the cone penetrometer measures different terrain properties in combination and it is impossible to determine to what degree each particular affects the results of cone penetrometer tests (Priddy and Willoughby, 2006).
A collection of data (CI, Vehicle Cone Index, Rating Cone Index, etc.) and algorithms used to predict vehicle mobility on terrain specific to certain parts of the world, as compiled beginning in the late 1970's, is referred to as the NATO Reference Mobility Model (NRMM). Using the NRMM, the cone penetrometer and the cone index derived from it can be used on a “go/no go” basis of vehicle trafficability in a variety of terrains around the world. While the use of the cone index and the NRMM for in situ measurement of soil strength for use in decision making is invaluable, the empirical method is not suited for vehicle development, design, and operation purposes (Schmid, 1995). Design engineers require the use of vehicle parameters which are simply not taken into consideration in the empirical methods.

2.2 Analytical Terramechanics Models

Soils modeled as elastic media can be used to predict the stress distribution in the soil due to normal loads. Figure 2 demonstrates the stress distribution formula for points in the soil due to a point load on the surface. The resulting equation for the normal stress at a point is called the Boussinesq equation and is given below (Bekker, 1969).

\[ \sigma_z = \frac{3W}{2\pi R^3} \cos^3 \theta \]  \hspace{1cm} (1)

where \( W \) is the magnitude of the point load applied at the surface, \( R \) is radial distance at which the stress is being calculated, and \( \theta \) is the angle between the \( z \) axis and the line segment for \( R \). Notice that the Boussinesq equation does not depend on the material; it gives the stress distribution for a homogeneous, isotropic, elastic medium subject to a point load on the surface (Bekker, 1969). Once the stress distribution for a point load is known, then, given the contact area one may integrate the point load stress formula over the contact area to determine the
normal stress distribution in the soil. For a load applied under a circular loading area, as shown
in Fig. 3, one can show that integrating the Boussinesq equation over the contact area leads to

\[
\sigma_z = 3p \int_0^{\gamma(z)} \frac{u \, du}{(1 + u^2)^{3/2}} = p_0 \left[ 1 - \frac{z^3}{(z^2 + r^2)^{3/2}} \right]
\]  

(2)

For a contact strip (shown in Fig. 4), which may be taken as the idealization of the contact area
under a track, one can show that the equations for the stresses at a point are

\[
\begin{align*}
\sigma_x &= \frac{p_0}{\pi} \left( \theta - \theta_1 + \sin \theta \cos \theta_1 - \sin \theta_2 \cos \theta_2 \right), \\
\sigma_z &= \frac{p_0}{\pi} \left( \theta - \theta_1 - \sin \theta \cos \theta_1 + \sin \theta_2 \cos \theta_2 \right), \\
\tau_{xz} &= \frac{p_0}{\pi} \left( \sin^2 \theta_2 - \sin^2 \theta_1 \right)
\end{align*}
\]  

(3)

These equations are derived with the assumptions that the contact patch is an infinitely long strip
with constant width, the track links are rigid, and a uniform pressure is applied. Models based on
the theory of elasticity which do not take into account the effect of plastic deformations cannot,
in general, be used to predict the shear stress distribution at the soil-tire interface. Another
shortfall of these elasticity models is that they may not be applied when loads become too large.

The most widely known methods for analytical analysis of tracked vehicle performance are
based on the developments initiated by Bekker. A modified Bekker’s pressure sinkage
relationship is given by (Wong, 2010)

\[
z_0 = \left( \frac{p}{(k_c/b) + k_\phi} \right)^{1/n} = \left( \frac{W/b}{(k_c/b) + k_\phi} \right)^{1/n}
\]  

(4)

In this equation, \( z_0 \) is the sinkage, \( p \) is the pressure, \( W \) is magnitude of the applied load, \( b \) is
contact depth, \( l \) is the contact patch length, \( k_c \) and \( K_\phi \) are the pressure-sinkage parameters for
the Reece equation (Reece, 1965). This pressure-sinkage relationship together with a criterion for
shear failure (most often the Mohr failure criteria) can be used to predict the performance of the vehicle.

A variety of pressure sinkage relationships exist; these pressure sinkage models attempt to capture and correct for behavior that was not considered in the original formulation. Response to cyclic loading, the use of elliptical contact areas, and the extension to small diameter wheels are examples of some of the modifications made to the pressure sinkage formulation.

Other analytical models have been proposed by Wong (2010) including the NTVPM, RTVPM, and NWVPM models. Wong’s models are based on the design parameters of vehicles and an idealization of the track terrain interface. These idealizations, for the case of the flexible track NTVPM model, can be seen in Fig. 5. With this configuration and variable definitions, the following pressure-sinkage relationship was given (Wong, 2010):

$$z_{l+1} = z_{ai} - \frac{T}{Rk_u} + \sqrt{\left(z_{ai} - z_{ri}\right)^2 + \frac{2T}{Rk_u} \left[ \frac{T}{2Rk_u} - z_{ai} + R\cos\varphi_{ri} + z_{ci+1} \right]} \tag{5}$$

where $z_{l+1}$ is the sinkage at point $F$ shown in Fig. 5, $z_{ai}$ is the sinkage of road wheel $i$, $T$ is the tension in the track per unit width, $R$ is the radius of the road wheel, and $k_u$ and $\varphi_{ri}$ are modal parameters. The associated shear-displacement relationship is given by

$$s(x) = \left(c + p(x)\tan\varphi\right)\left[1 - \exp\left(-\frac{(l-(1-i)x)}{K}\right)\right] \tag{6}$$

where $p(x)$ is the normal pressure on the track, $l$ is the distance between the point at which shearing begins and the corresponding point on the track, $K$ is the shear deformation parameter, and $c$ and $\varphi$ are the Mohr-Coulomb failure criteria parameters. Experimental and analytical terramechanics models tend to be simple and do not capture many modes of the soil
deformations that can be captured using the more general continuum mechanics-based soil models.

3. CONTINUUM MECHANICS-BASED SOIL MODELS

In the literature, there are many continuum-based soil models that employ different assumptions. Most of these models are suited for implementation in a finite element framework, as will be discussed in Section 7. These models are briefly reviewed in this section.

3.1 Theory of Elastoplasticity

Given that soils typically experience both recoverable and non-recoverable deformation under loading, elastoplastic theory and several augmentations of the theory have been widely applied to soils. Elastoplasticity theory is based on the decomposition of the strain into elastic and plastic parts. In the case of small strains, the additive strain decomposition \( \varepsilon = \varepsilon^e + \varepsilon^p \) is used, where \( \varepsilon \) is the total strain, \( \varepsilon^e \) is the elastic strain, and \( \varepsilon^p \) is the plastic strain. In the case of large strains, the following multiplicative decomposition of the deformation gradient \( J \) is used. This decomposition is defined as \( J = J^e J^p \), where subscripts \( e \) and \( p \) in this equation refer, respectively, to the elastic and plastic parts. The stress is related to the elastic strain. Since the elastic region, is often relatively small in soils, the linear stress-strain relationship \( \sigma = C^e \varepsilon^e \) is often sufficient, where \( C^e \) is the fourth order tensor of elastic coefficients, \( \sigma \) is the stress tensor, and \( \varepsilon^e \) is the elastic strain tensor. While the linear stress-strain relationship has been widely used in many soil models, it is important to point out that some models have incorporated nonlinear elastic relationships in both the small strain (e.g. Fossum and Brannon, 2004) and large deformation (e.g. de Souza Neto et al, 2008) cases. Such models help correct the amount of elastic strain during plastic deformation.
The elastic region is defined by a yield function \( f(\sigma) \). When \( f < 0 \), the stress state is within the elastic region. Plasticity can only occur when \( f = 0 \), which defines the yield surface. Stress states where \( f > 0 \) are inadmissible. However, the yield surface may evolve or translate, as discussed below, allowing initially inadmissible stress states after some plastic deformation. The evolution of plastic strain is governed by the flow rule (Araya and Gao, 1995; Mouazen and Nemenyi, 1999) \( d\varepsilon^p = d\lambda (\partial g/\partial \sigma) \), where \( d\lambda \) is the plastic multiplier and \( g \) is a plastic potential function that determines the direction of plastic flow. If \( f = g \), then the flow rule is said to be associative. Associative flow follows from the principle of maximum plastic dissipation, allowing the body to reach the lowest possible energy state; hence it is commonly employed in the plasticity theory of metals. However, the principle of maximum plastic dissipation tends to overestimate dilatation in soils and other cohesive-frictional materials, and hence many soil models use nonassociative laws. While any plasticity model may experience a loss of ellipticity condition that leads to spurious mesh dependency in numerical solutions during softening, nonassociative models may experience loss of ellipticity even during the hardening phase (Rudnicki and Rice, 1975), adding a necessity to check for this condition.

The yield function and plastic flow rule together operate under the Kuhn-Tucker optimality conditions \( f \leq 0, \; d\lambda \geq 0, \; f \cdot d\lambda = 0 \), which are important to the solution of the problem. Section 7 will outline the importance of the return mapping algorithms typically employed in the solution to the plasticity equations.

The last element needed to define a plasticity model is the evolution of internal state variables. The yield surface and plastic potential may not be constant but may evolve with plastic work or strain. For example the size of the yield surface may increase, allowing plastic hardening. The elastic constitutive equation, yield function, flow rule, and hardening laws, together, define
the mechanical behavior for a particular model. They are often reformulated in rate form in order to define a solution procedure which can be subjected to the Kuhn-Tucker optimality conditions. Some of the more common soil models are detailed in the following subsections.

3.2 Single Phase Plasticity

In this section, single phase plasticity models are discussed. Here, the soil is treated as a homogenized medium of solid and fluid mass. These models include the Mohr-Coulomb model, the Drucker-Prager and uncapped three-invariant models, modified Cam-Clay and Cap models, and viscoplastic soil models.

**Mohr-Coulomb Model** The Mohr-Coulomb model is one of the oldest and best-known models for an isotropic soil (Goldscheider, 1984). Initially the yield surface was used as a failure envelope, and still is in geotechnical practice. It was later adopted as a yield surface for plasticity models. In two dimensions, the yield surface of the Mohr-Coulomb model is defined by a linear relationship between shear stress and normal stress which is written as (An, 2010)

\[ f = \tau - (c - \sigma \tan \phi) = 0 \]  

(7)

where \( \tau \) and \( \sigma \) are, respectively, the shear and normal stresses, and the constants \( c \) and \( \phi \) are the cohesion and internal friction angle, respectively. In three dimensions, the yield surface is more complicated and is defined by the following equation (An, 2010):

\[ f = \frac{1}{3} I_1 \sin \phi + \sqrt{J_2} \sin \left( \theta + \frac{\pi}{3} \right) + \sqrt{\frac{J_2}{3}} \cos \left( \theta + \frac{\pi}{3} \right) \sin \phi - c \cos \phi = 0 \]  

(8)

where \( I_1 = \text{tr}(\sigma) \) is the first invariant of the stress tensor \( \sigma \), \( J_2 = (S:S)/2 \) is the second invariant of the deviatoric stress tensor \( S = \sigma - (1/3) I_1 I \), and \( \theta \) is equal to the Lode angle defined by (An, 2010):
\[ \cos(3\theta) = \frac{3\sqrt{3}}{2} \frac{J_3}{J_{3/2}^2} \]  

(9)

where \( J_3 = \det(S) \) is the third invariant of deviatoric stress tensor. A Mohr-Coulomb yield surface forms a hexagonal pyramid in principal stress space, as shown in Fig. 6a. As can be seen from Fig. 6a, the failure envelope defined by the Mohr-Coulomb model includes discontinuous slopes between failure surfaces. These discontinuities add complexity to the return-mapping algorithm. While multi-surface plasticity algorithms have been used to handle this situation, such algorithms are complex and more time consuming.

**Drucker-Prager and Uncapped Three-Invariant Models**  

A simpler method to handle the discontinuities is to use a smooth approximation to the yield surface. Drucker and Prager (1952) initially proposed a cone in principal stress space (Fig. 6b), by adding a pressure-dependent term to the classical von Mises yield surface, resulting in the yield function:

\[ f = \sqrt{J_2} + \eta p - \xi c \]  

(10)

where \( J_2 \) and \( p = I_3/3 \) are invariants of the stress tensor, \( c \) is the cohesion, \( \eta \) and \( \xi \) are parameters used to approximate the Mohr-Coulomb criterion. Like von Mises plasticity, one-step return-mapping can be achieved for linear hardening, making the model quite efficient to implement. While the associative model over predicts dilatation, nonassociative versions correct this (Drucker et al., 1952). Initially developed as an elastic-perfectly plastic model, i.e. with no change in the yield surface on loading, researchers later added hardening of the yield surface parameters to the model in various forms. See, for example, Vermeer and de Borst (1984) for a relatively sophisticated phenomenological hardening model.

Limitations of the model include: hydrostatic loading and unloading produces considerable hysteresis which cannot be predicted using the same elastic bulk modulus of
loading and unloading and a yield surface which does not cross the hydrostatic loading axis (DiMaggio and Sandler, 1971), and that the cone does not approximate the Mohr-Coulomb hexagonal pyramid well for low friction angles. To account for this last issue, researchers have developed smooth yield surfaces that better approximated the Mohr-Coulomb yield surface. These yield surfaces have different yield points in triaxial extension versus compression, like the Mohr-Coulomb yield surface, but are smooth. The Matsuoka and Nakai (1974) model actually captures both the extension and compression corners of the Mohr-Coulomb yield surface, unlike the Lade-Duncan model (Fig. 6d). While this fact does not necessarily make the Matsuoka-Nakai yield surface more correct, it does make it easier to fit to standard geotechnical strength tests.

The differences in triaxial extension and compression strength can also be captured by modifying a Drucker-Prager type yield surface using a smooth third-invariant modifying function. Two of these functions are developed by Gudehus (1973) and William and Warnke (1975). While the former is simpler in form, it is only convex when the ratio of triaxial extension to compression strength, $\psi$, is greater than 0.69. The William-Warnke function is convex until $\psi = 0.5$. Convexity is essential in yield surfaces to ensure proper return mapping.

A shortcoming of the above models is that they assume a constant ratio between pressure and deviatoric stress, or normal and shear stress, during yielding, that is, a constant friction coefficient. Research in soils shows that this is not the case and the friction angle decreases with increasing pressure. Furthermore, at high confining pressures, soils may exhibit compactive plasticity due to pore collapse, grain crushing, and other phenomena.

**Modified Cam-Clay and Cap Models** The original Cam-Clay model has not been as widely used for numerical predictions as the modified Cam-Clay (MCC). The qualifier “modified” is often dropped when referring to the modified Cam-Clay model (Wood, 1990). The
modified Cam-Clay model by Roscoe et al. (1968) is based on the critical state theory and was meant to capture the properties of near-normally consolidated clays under triaxial compression test conditions. The yield surface is assumed to have an elliptical shape that may be expanded with the increase of volumetric strain, as shown in Fig. 7. The function for the yield surface of the MCC model is defined as

\[ q^2 - M^2[p(p_c - p)] = 0 \]  

(11)

Here, \( p \) is the effective mean stress, the pre-consolidation stress \( p_c \) acts as a hardening parameter, and the stress ratio \( M = q/p \) at critical state is related to the angle of friction through the relationship \( M = 6\sin(\phi)/(3 - \sin(\phi)) \). The modified Cam-Clay model has been extended to the finite deformation case in Borja and Tamagnini (1996).

Cam-Clay models can predict failure and the nonlinear stress-path dependent behaviors prior to failure accurately, especially for clay type soils (DiMaggio and Sandler, 1971). This model, however, still has some disadvantages (DiMaggio and Sandler, 1971): the slope discontinuity at the intersection with the \( p \) axis predicts behavior not supported by experiments (Fig. 8); points on the yield surface above the critical state line do not satisfy Drucker’s postulate of stability; and the shear strain predicted by Cam-Clay models is too high at low stress ratios (Karim and Gnanendran, 2008).

There are several advanced derivatives of the Cam-Clay type soil models that include the three-surface kinematic hardening model and the \( K \)-hypoplastic model. The three-surface kinematic hardening (3-SKH) model employs the following kinematic surfaces: the first surface is defined as the yield surface, the second surface is named the history surface and is the main feature of the 3-SKH model, and the third surface is the bounding surface. The bounding surface
is taken to be the MCC yield surface, the history surface defines the influence of recent stress history, and the yield surface defines the onset of plastic deformations. Kinematic hardening allows it to better predict load reversals. It has been found that the 3-SKH model can acceptably predict over-consolidated compression behavior for clay but can have difficulty modeling pore pressure variations (Bryson and Salehian, 2011). The K-Hypoplastic model employs critical state soil mechanics concepts that can be applied to the modeling of fine-grained soils. It can be formulated in two manners; by enhancing the model with the intergranular strain concept, it can be extended to the case of cyclic loading and further improve the model performance in the range of small-strains. Even without the above enhancement, the K-Hypoplastic model is suitable for fine-grained soils under monotonic loading at medium to large strain levels (Masin et. al., 2006).

Cap-plasticity models were developed to address the shortcomings of the Cam-Clay type models. Drucker et al. (1957) first proposed that “successive yield surfaces might resemble an extended Drucker-Prager cone with convex end spherical caps” as shown in Fig. 6c (Chen and Baladi, 1985). As the soil undergoes hardening, both the cone and the end cap expand. This has been the foundation for numerous soil models.

The plastic yield function \( f \) in the inviscid cap model of DiMaggio and Sandler (1971) is formulated in terms of the first stress invariant \( I_1 \) and the second deviatoric stress invariant \( J_2 \) (Sandler and Rubin, 1979; Simo et al., 1988). As shown in Fig. 8, the static yield surface is divided into three regions. The cap is a hardening elliptical surface defined as

\[
f(I_1, \sqrt{J_2}, k) = \sqrt{J_2} - F_c(I_1, k) = \sqrt{J_2} - \frac{1}{R} \sqrt{[X(k) - L(k)]^2 - [I_1 - L(k)]^2} = 0 \tag{12}
\]

where \( J_2 \) is the second invariant of the deviatoric stress \( S \), \( R \) is a material parameter, and \( k \) is a hardening parameter related to the actual plastic volumetric change \( \varepsilon_v^p = \text{tr} (\varepsilon^p) = \varepsilon_{11}^p + \varepsilon_{22}^p + \varepsilon_{33}^p \).
In Equation 1, $L(k)$ is the value of $I_1$ at the location of the start of the cap; $L(k) = k$ if $k > 0$, and $L(k) = 0$ if $k \leq 0$. The yield surface is of a Drucker–Prager type modified for nonlinear pressure dependence and is defined by the function

$$f \left( I_1, \sqrt{J_2} \right) = \sqrt{J_2} - F_c(I_1) = \sqrt{J_2} - \left[ \alpha - \gamma \exp(-\beta I_1) + \theta I_1 \right] = 0$$

(13)

where $\alpha$, $\beta$, $\gamma$, and $\theta$ are material parameters. The tension cutoff surface is defined by $f(I_1) = I_1 - (-T)$, where $T$ is the tension cutoff value. Eleven material parameters are necessary for the elastoplastic cap model: $\eta$, $N$, $f_0$ in the viscous flow rule to be discussed later; $W, D, R, X_0$ in the cap surface; $\alpha$, $\beta$, $\gamma$, $\theta$ in the failure surface; and $T$ in the tension cutoff surface. In addition, the bulk modulus $K$ and the shear modulus $G$ are needed for the elastic soil response.

The Sandia GeoModel builds on the Cap model with some modifications. It is capable of capturing a wide variety of linear and nonlinear model features including Mohr-Coulomb and Drucker-Prager plasticity depending on the model parameters incorporated. Unlike the Cap model, the cap surface and shear yield surface are connected in a smooth manner, and the model also accounts for differences in triaxial extension and compression strength using either Gudehuis or William-Warnke modifying function described above. The yield function can be written as

$$f = \left( \Gamma^\xi \right)^2 J_2^\xi - F_c \left( F_f - N \right)^2 = 0$$

(14)

where $\Gamma^\xi$ accounts for the differences in material strength in triaxial extension and triaxial compression, $J_2^\xi$ is the second invariant of the relative stress tensor $\sigma - \alpha$ (here $\alpha$ is a back stress state variable), $F_c$ is a smooth cap modifying function, $F_f$ represents the ultimate limit on the amount of shear the material can support, and $N$ characterizes the maximum allowed
translation of the yield surface when kinematic hardening is enabled (Fossum and Brannon, 2004). The plastic potential function is given by (Foster, et al. 2005)

\[
g = \left( F^z \right)^2 J^z - F^c \left( F^f - N \right)^2
\]

(15)

where \( F^c \) and \( F^f \) play analogous roles in the plastic potential function as their counterparts in the yield function. The Sandia GeoModel suffers from the following limitations: the triaxial extension/compression strength ratio does not vary with pressure and it is computationally intensive when compared to similar idealized models (Fossum and Brannon, 2004). This model has been further adapted to the Kayenta model (Brannon et. al., 2009).

Soil is not always an isotropic material. Layering and fracture networks, as well as compaction and other history effects may give the soil higher strength or stiffness in certain directions. Often the effects impart different strength and stiffness in one plane, and there is a transversely anisotropic version of the Kayenta model. Anisotropy may also be addressed using fabric tensors (Wan and Guo, 2001). Other anisotropic models include the work of Whittle (1994), and the S-CLAY 1 model (Wheeler et. al., 2003), which builds on the MCC model. Aside from the kinematic hardening mentioned in some of the models, detailed review of anisotropic soil models is beyond the scope of this article, however, and the reader is referred to the above references.

**Viscoplastic Soil Models**

Plasticity models such as those described above do not include strain-rate dependent behavior often observed in soils under rapid loading. These viscous effects are more pronounced in the plastic region of most clay soils and rate independent elastic response is generally adequate for practical engineering applications (Perzyna, 1966; Lorefice, 2008). The models described above can be modified to account for rate-dependent plastic effects.
Such viscoplastic models are more accurate under fast loading conditions. However, it is difficult to determine the correct value of the material time parameter if the stress history is not known.

Two major types of viscoplastic overlays are the Perzyna and Duvaut-Lions formulations. Perzyna's formulation is among the most widely used viscoplasticity models (Darabi et al., 2011). In this model, the associative time-rate flow rule is used to describe viscous behavior leading to a viscoplastic potential which is identical if not at least proportional to the yield surface (Katona, 1984; Chen and Baladi, 1985; Simo et al., 1988). In Perzyna’s viscoplasticity formulation (Perzyna, 1966), the viscoplastic flow rule can be expressed as

\[ \dot{\varepsilon}^v = \eta \langle \phi(f) \rangle \frac{\partial g}{\partial \sigma} \]  

(16)

where \( \eta \) is a material constant called the fluidity parameter, the Macauley bracket \( \langle \ \rangle \) is defined as \( \langle x \rangle = x + |x|/2 \), \( g \) is the plastic potential function, and \( \phi(f) \) is a dimensionless viscous flow function commonly expressed in the form \( \phi(f) = (f/f_0)^N \), where \( N \) is an exponent constant and \( f_0 \) is normalizing constant with the same unit as \( f \). The Cap model has been extended to the viscoplastic case using Perzyna’s formulations. The viscoplastic cap model is adequate for modeling variety of time dependent behaviors such as high strain rate loading, creep, and stress relaxation (Tong and Tuan, 2007).

A joint bounding surface plasticity and Perzyna viscoplasticity constitutive model has been developed for the prediction of cyclic and time-dependent behavior of different types of geosynthetics (Liu and Ling, 2007). This model can simulate accelerating creep when deviator stresses are close to the shear strength envelope in a \( q \) creep test and it can also model the behavior in unloading–reloading and relaxation (Yin and Graham, 1999). It has been noted that
for multi-surface plasticity formulations the Perzyna type models have uniqueness issues (Simo and Hughes, 1998).

Another widely used formulation for viscoplasticity is based on Duvant-Lions theory (Duvaut and Lions, 1972). In this formulation, the viscoplastic solution is constructed through the relevant plastic solution. An advantage of the Duvant-Lions model is that it requires the simple addition of a stress update loop to incorporate it into existing plasticity algorithms. Another advantage is that the viscoplastic solution is guaranteed to deteriorate to the plastic solution under low strain rate (Simo et al., 1988).

Viscoplasticity is thought to simulate physical material inelasticity behavior more accurately than the plasticity approach. It eliminates potential loss of ellipticity condition associated with elasto-plastic modeling (Abdullah, 2011). A viscoplastic version of the GeoModel has been developed with separate viscous parameters for volumetric and shear plasticity.

### 3.3 Multiphase Models

Soils may either be treated as homogenized continua or as a mixture in which each phase (solid, liquid, and gas) is treated separately. The latter approach is considered to be more accurate, but more complicated to implement. Mixture theory can be used at the continuum level to account for each phase, by tracking the total stress $\sigma$, fluid pore pressure $p_w$, and a pore air pressure $p_a$.

For saturated soils (no gas phase), an effective stress $\sigma'$ is defined, typically as $\sigma' = p_w I$ (though variations exist). The deformation of the soil skeleton is taken to be a function of the effective stress. Any of the plasticity models above can then be implemented using the effective stress in place of the total stress to determine the solid deformation.
In the unsaturated case, two independent variables are usually used to determine the mechanical response, due to apparent cohesion created by menisci in fluid phase. The total stress may be broken down into a net stress $\sigma^\prime$ for the solid skeleton, and suction stress $p_c$ defined as

$$\sigma^\prime = \sigma - p_a \mathbf{I}, \quad p_c = p_a - p_w$$

(17)

An effective stress and suction may also be used, typically $\sigma^\prime = \sigma - p_a \mathbf{I} + \chi(p_a - p_w)\mathbf{I}$, where $\chi$ is a parameter that varies from 0 for dry soil to 1 for fully saturated. The advantage of this formulation is that it reduces to the standard effective stress at saturation. The solid phase may then be modeled using the effective stress in place of the total stress. Many of the above plasticity and viscoplasticity models have been used to model solid deformation in this framework. In rapid loading, the fluid may be thought of as moving with the solid in the saturated case (undrained), but otherwise fluid flow through the solid matrix needs to be accounted for. In the limit where the fluid has enough time to return to steady state conditions, the material is said to be fully drained. Standard coupled fluid flow-solid deformation finite elements often fail due to volumetric mesh locking phenomena. This shortcoming can be solved by either using a lower order interpolation scheme for fluid flow equations (Brezzi, 1990), or by stabilizing the element (White and Borja, 2008, and references therein).

The Barcelona Basic Model (BBM) proposed by Alonso et al. (1990) remains among the fundamental elasto-plastic models for unsaturated soils. The BBM model is an extension of the modified Clam-Clay model that captures many of the mechanical characteristics of mildly or moderately expansive unsaturated soils. As originally proposed by Alonso, utilizing a critical state framework, the BBM is formulated in terms of the hydrostatic pressure $p^\prime$ associated with the net stress tensor $\sigma^\prime$, suction $s$, and the deviatoric stress $S$. One may write a yield function for the model as follows:
where \( q \) is the difference between the maximum and minimum principal stresses, \( M \) is the slope of the critical state lines, \( k \) is parameter that describes increase in apparent cohesion with suction, \( P_c \) is the pre-consolidation pressure, and the function \( \tilde{g}(\theta) \) is given by

\[
\tilde{g}(\theta) = \sin\phi' / \left( \cos\theta + \left( \sin\theta \sin\phi' / \sqrt{3} \right) \right),
\]

where \( \phi' \) is the friction angle, and \( \theta \) is the Lode angle.

The hardening law follows the following relationship:

\[
dP_0 = \frac{P_0}{\lambda_0^* - \kappa^*} d\varepsilon^p
\]

where \( P_0 \) is the hardening parameter defined by the location of the yield surface at zero suction, \( \lambda_0^* \) is the slope modified at the normal compression line, and \( \kappa^* \) is the modified swelling index that is assumed to be independent of suction. The plastic potential is a slight modification of the yield function given by

\[
g = 3\alpha J_2 - \left( \frac{g(\theta)}{g(-30^\circ)} \right)^2 M^2 \left( p'' + (k) s \right) (P_c - p'')
\]

where \( \alpha \) is defined as \( \alpha = M (M - 9)(M - 3) \left( 1 / (1 - k^*/\lambda_o^*) \right) / 9(6 - M) \) (Alonso et al., 1990).

Some of the shortcomings of the BBM model are as follows: the BBM cannot completely describe hydraulic hysteresis associated with wetting and drying paths, it does not give the possible ranges of suction over which shrinkage may occur, and it does not include a nonlinear increase in shear strength with increasing suction.

**Elasto-Plastic Cap Model of Partially Saturated Soil**

This section deals with the extension of a cap model which can describe the material behavior of partially saturated soils, in
particular, of partially saturated sands and silts. The soil model is formulated in terms of two stress state variables; net stress $\sigma''$, and matric suction $p_c$ (Fig. 9). These stress state variables are defined in Eq. 17.

The yield surface, consisting of a shear failure surface and a hardening cap surface, the plastic potentials for the non-associated flow rule and the hardening law for the cap are extended by taking into account the effects of matric suction on the material behavior. Furthermore, the third invariant of the deviatoric stress tensor is incorporated in the formulation of the yield surfaces (Kohler, 2007). Using net stress and matric suction as stress state variables allows modeling independently the effects of a change in the skeleton stress and of a change in suction effects on the mechanical behavior of the soil skeleton (Kohler, 2007). The functional form of the shear failure surface is

$$f_1(\sigma'', p_c) = L(\theta)\sqrt{2J_2} - F_e(I''_1) - F_s(p_c)$$

(21)

where $I''_1$ denotes the first invariant of the net stress tensor $\sigma''$. In the preceding equation, $L(\theta) = ((1 - \omega \cos 3\theta)/(1 - \omega))^{-\eta}$, where $\omega$ and $\eta$ are parameters defining the shape of the yield surface with respect to the Lode angle $\theta$. In Eq. 22, $F_e(I''_1)$ defines the shear failure envelope at vanishing matric suction, and $F_s(p_c)$ accounts for the dependence of the shear strength on the matric suction (Kohler, 2007). These two functions are defined as $F_e(I''_1) = \alpha + \eta I''_1$ and $F_s(p_c) = k p_c$, where $k$ is a parameter controlling the increase of the shear failure envelope with increasing matric suction, and $\alpha$ is a material parameter (Kohler, 2007). The functional form of the strain hardening cap is

$$f_2(\sigma'', \kappa(p_c), p_c) = F_e(\sqrt{2J_2}, I''_1, \theta, \kappa(p_c)) - F_e(\kappa(p_c)) - F_s(p_c)$$

(22)
with \( \kappa(p_c) \leq I_1^* \leq X(\kappa(p_c)) \), and

\[
F_c\left(\sqrt{2J_2}, I_1^*, \theta, \kappa(p_c)\right) = \sqrt{L'\left(\theta\right)\left(\sqrt{2J_2}\right)^2 + \left(\frac{I_1^* - \kappa(p_c)}{R}\right)^2}
\tag{23}
\]

The plastic strain rate is determined by the non-associative flow rule \( \dot{\varepsilon}^p = \sum_{i=1}^{2} \dot{\gamma}_i \left( \frac{\partial g_i}{\partial \sigma^*} \right) \), where \( \dot{\gamma}_i \) are the plasticity consistency parameters. The direction of the plastic flow is determined by means of a plastic potential

\[
g_1(\sigma^*, p_c) = \sqrt{2J_2} - \alpha - \psi I_1^* - F_c(p_c)
\tag{24}
\]

In this equation, \( \psi \) is a parameter that governs the amount of plastic dilatation. The plastic potential for the strain hardening cap is assumed as

\[
g_2(\sigma^*, \kappa(p_c), p_c) = \sqrt{\left(\sqrt{2J_2}\right)^2 + \left(\frac{I_1^* - \kappa(p_c)}{R}\right)^2} - F_c(\kappa(p_c)) - F_c(p_c)
\tag{25}
\]

The plastic volumetric strain rate is

\[
\dot{\varepsilon}_v^p = \lambda(p_c) \frac{\dot{X}(\kappa(p_c))}{X(\kappa(p_c))}
\tag{26}
\]

where \( X(\kappa(p_c)) \) corresponds to the apex of the elliptical cap.

**Bounding Surface Plasticity Unsaturated Soil Model**  
Dafalias and Popov (1976) developed bounding surface plasticity for metals. This approach was later applied to clays by Dafalias and Herrmann (1982), extended to pavement based materials by McVay and Taesiri (1985), and to sands by Hashigushi and Ueno (1977), Aboim and Roth (1982), and Bardet (1985). Bounding surface plasticity provides a framework with which to capture the cyclic behavior of engineering materials. The advantages of this framework over conventional plasticity theory have been investigated for monotonic and cyclic loads. Wong, Morvan, and Branque (2009)
developed a new bounding surface plasticity model, which includes an evolving bounding surface, for unsaturated soils with a small number of parameters based on Bardet’s model (Bardet, 1985).

The bounding surface plasticity model developed by Wong et al. (2009) is elliptical in the plane of effective mean-stress $p'$ and deviatoric stress $q$ with $p' = (\sigma_1' + \sigma_2' + \sigma_3')/3$ and using cylindrical symmetry $q = \sigma_1' - \sigma_3'$. The bounding surface can be defined as

$$f\left(\bar{p}', \bar{q}', \varepsilon_p^p, s\right) = \left(\frac{\bar{p}' - A_\pi}{\rho - 1}\right)^2 - \left(\frac{\bar{q}}{M_\pi}\right)^2 - A_\pi^2$$

where $M_\pi$ is the slope of the saturated soil critical state line (CSL), $\bar{p}' = \gamma A'_\pi$, $\bar{q}' = \gamma x M_\pi A_\pi$, $x = q/(Mp' + q_0)$, and $\gamma = \left(1 + (\rho - 1)\sqrt{1 + x^2 \rho (\rho - 2)}\right)/\left(1 + (\rho - 1)^2 x^2\right)$. $M_\pi$ and $A_\pi$ are assumed to be material parameters that are independent (in particular, of suction $s$). Also, $\rho$ is a material parameter. The bounding surface plasticity soil model has the following limitations and shortcomings (An, 2010): (1) more experimental data is needed to define the suction dependence of material parameters; and (2) an objective relation, defined by the retention curve, is needed between the degree of saturation and suction.

### 4. PARTICLE BASED AND MESHFREE METHODS

The finite element method is a widely accepted and used approach to the solution of engineering problems which can be modeled using a continuum approach. However, simulations of explosions, fragmentations, and inherently granular problems require the use of adaptive meshing techniques that can become computationally intensive (Belytschko, Liu, and Moran, 2000). Particle-based and meshfree methods offer engineers a new methodology with which they
may more accurately tackle highly discrete or granular problems. Particle based methods offer a number of advantages. The connectivity between nodes, or particles, is (re)computed at each time step and this allows for simulations of large deformations (Li and Liu, 2002). Fracture and other discontinuous behaviors are explicitly captured by particle-based methods. The following is a brief overview of three of the most commonly used particle based and mesh-free methods; the discrete element method, smoothed particle hydrodynamics, and reproducing kernel particle methods.

4.1 Discrete Element Method (DEM)

In the case of the finite element method, the material (soil in this study) is assumed to be a continuum. For the cases in which the granular behavior of soil is to be accurately modeled the discrete element method (DEM) is applied. The DEM was developed to simulate the dynamic behavior of granular material such as granular flow. In the DEM, the material is represented by an assembly of particles with simple shapes (circles and spheres), although there have been simulations in which non-circular rigid particles are used (Tutumluer et al., 2006; Reeves et al., 2010). The elastic and inelastic properties at the contact between the particles are introduced using springs with spring constants (elastic response) and dashpots with viscous damping constants. The contact forces between particles are calculated from the interpenetration between those particles using the spring constant and the viscous damping constant (Oida and Momozu, 2002). The displacements of the particles are obtained for a certain time interval by solving the governing kinetic equations of motion. This process is repeated for all particles in the analyzed region for very short time intervals until the end of the simulation time. Some of the shortcomings associated with the discrete element method are as follows: it can be computationally very inefficient for soil in which the granular effect can be approximated using a
continuum model (Oida and Momozu, 2002), it is difficult to accurately determine the spring and
damping constants that define the contact forces between the particles (Khulief and Shabana,
1987), and the representation of soil cohesion and adhesion properties is difficult to incorporate
within DEM analysis (Asaf et. al., 2006).

4.2 Smoothed Particle Hydrodynamics (SPH)

As one of the earliest meshfree methods, *smoothed particle hydrodynamics* has been widely
adopted and used to solve applied mechanics problems. In SPH, the idea is to discretize the
material into particles, with each particle having a unique neighborhood over which its properties
are "smoothed" by a localized interpolation field, called the kernel function (Li and Liu, 2002).
The neighborhood of each element defines the interaction distance between particles, often
referred to as the smoothing length. Smoothed particle hydrodynamics has been used to model
soil behavior. In particular, Bui *et al.* (2008) proposed a Drucker-Prager model for elastic-plastic
cohesive soils which showed good agreement to experimental results. However, the model
suffered from tensile instability which was overcome by using the tension cracking treatment,
artificial stress, and other methods. Recently, SPH methods have been used in conjunction with
FEM to produce tire-soil interaction models, but it was concluded that further validation would
be required to analyze the effects of SPH parameters on results (Lescoe et. al., 2010).

4.3 Reproducing Kernel Particle Methods (RKPM)

RKPM improves the accuracy of the SPH method for finite domain problems (Chen et al., 1997).
In this method a modification of the kernel function, through the introduction of a correction
function to satisfy reproducing conditions, results in a kernel that reproduces polynomials to a
specific order. Unlike traditional SPH methods, the RKPM method can avoid the difficulties
resulting from finite domain effects and minimize the amplitude and phase errors through the use
of a correction function which allows for the fulfillment of the completeness requirement. While RKPM methods have not been used in vehicle-terrain interactions, they have been quite successfully applied to geotechnical applications. RKPM methods demonstrate promising potential for large deformation problems but require a systematic approach for the selection of appropriate dilation parameter in order to be made robust (Chen et al., 1997).

5. EXISTING SOFTWARE

There are several commercial and research computer programs that are used in soil modeling. Some of these programs are based on nonlinear finite element algorithms. In this section, a brief review of some of these programs is presented.

ABAQUS

ABAQUS is a popular FE analysis program that contains a wide variety of material models and is further configurable by user defined material models. The inelastic material models that are deployed with ABAQUS /Standard (as of version 6.8) include the following broad categories: metal plasticity, fabric materials, jointed materials, concrete, and permanent set in rubberlike materials. The plasticity models that are included in ABAQUS /Standard and are of most relevance include: extended Drucker-Prager models, Modified Drucker-Prager/Cap models, Mohr-Coulomb plasticity, Critical State (Clay) plasticity models, and crushable foam plasticity models. These inelastic material models are applicable only when the elastic region is linearly elastic. In addition to the built in material models, one may develop user defined material models for use with ABAQUS. The constitutive models of the user defined materials can be programmed in the user subroutine UMAT. Many user defined material models can be found as extensions for ABAQUS. For example, the finite elasto-plasticity material model
(FeFp material model) developed for use with nonlinear elastic finite strains and nonlinear plastic hardening is offered as an extension.

**ANSYS** ANSYS is another popular computer aided engineering (CAE) FE program capable of static, non-linear, thermal, modal, frequency response, and coupled field analysis and transient simulation. Its widespread popularity can be said to spur from its parametric language (ANSYS Parametric Design Language) which allows for the scripting of all commands necessary to perform pre-processing, solution, and post-processing of a problem. It includes a diverse set of material models. Among which are included the following: anisotropic elastic, plastic kinematic, Mooney-Rivlin rubber, three parameter Barlat plasticity, strain rate dependent plasticity, and geological Cap.

**MARC** The general purpose FE program MARC includes a modified critical state model in conjunction with a new nonlinear elastic law which gives a non-associative elastoplastic model for geomaterials within the regime of large strains. This model includes fully implicit integration and algorithmic tangent moduli, resulting in a quadratic rate of convergence in global Newton iterations. The essential features of this model are the satisfaction of the principle of conservation of energy, flexibility with respect to consideration of the evolution of Poisson’s ratio and the physically meaningful interpretations of all model parameters (Liu et al., 2000, 1996).

**OPENSEESPL** The standard incremental theory of elasto–plasticity is implemented in OPENSEESPL. Implementation within OpenSEES finite element platform allows for the use of existing material models and development of new elasto–plastic material models by simply combining yield functions, plastic flow directions (or plastic potential functions) and evolution laws into a working elastic–plastic models. An advantage of using OPENSEESPL is that one
may benefit from the use of the object oriented paradigm offered for the separation of elastic models, yield function, plastic flow directions, and evolution laws (hardening and/or softening).

**CRISP**  CRISP was developed at Cambridge University starting 1975 (Britto and Gunn, 1987) and further developed by the CRISP Consortium Ltd (Carter et al., 1982). In CRISP, a small strain formulation is implemented and, of the material models implemented in CRISP, the elastic perfectly-plastic Mohr–Coulomb model may be used as the basis for modeling certain soil behavior.

**TREMORKA and SHAKE91**  These programs are written for the analysis of earthquake phenomena. These linearized programs are based on a nonlinear model proposed by Schnabel et al. (1972), except for the method used to choose frequency-dependent moduli and damping (Hartzell et al., 2004).

**NOAHW and NOAH**  NOAHW is a second-order, staggered-grid finite difference code. NOAH is a multi-spring model for total stress analysis.

**LS-DYNA**  Is a scalable combined implicit/explicit solver for highly nonlinear transient problems. It offers a comprehensive materials library which includes the following broad categories: metals, plastics, glass, foams, elastomers, fabrics, concrete and soils, etc. The code can be augmented to include user defined material models. An example of a model incorporated into LS-DYNA is the two-invariant inviscid cap model developed for simulating soil behaviors under high strain rate based on a two-invariant, inviscid cap model proposed by (Simo et al., 1988) and augmented with the Perzyna’s viscoplastic formulation (Tong and Tuan, 2007).

6. **COMPARISON OF SOIL MODELS**
The soil models described in Sections 2, 3, and 4 offer a broad overview of the various common methods used in soil modeling. The categories for terrain-vehicle interaction presented in this paper include: terramechanics, continuum mechanics based soil models (FEA), and particle (DEM) and mesh-free methods (SPH and RKPM). Vehicle interaction studies exist that include a combination of the aforementioned broad categories, as in Nakashima and Oida (2004) for example. Table 1 offers a comparison of some of the soil models presented in this paper based on the ability of the model to capture work hardening, fracture, and cyclic loading.

The analytical terramechanics approach remains among the most popular methods used for vehicle-terrain interaction studies in MBS simulations (Ding et. al., 2011). This popularity can be attributed to the efficiency of most implementations for analytical terramechanics. Of the FEA, DEM, and mesh-free methods; the DEM is unattractive due to its computational cost and inherent difficulty in capturing cohesive and adhesive tensile phenomena of soil (though it can capture soil rupture and other particle phenomenon quite easily). Mesh-free methods are computationally intensive and require further testing and validation (Lescoe et. al., 2010). While FEA and DEM methods are gaining popularity, the initial resistance towards the use of these methods was due to the computational intensity required for such terrain-vehicle interaction. Considerable progress in computational power of personal computing systems is making this avenue of analysis more appealing. FEA implementations of soil plasticity that employ highly efficient algorithms have been developed in recent decades, often with quadratic rates of convergence, for the solution of the plasticity equations. Leveraging these algorithms within a FEA/MBS environment will allow for the development of high fidelity vehicle-terrain interaction models.
Considering the continuum-based soil models category of Table 1, it can be seen that of the models presented many capture both work hardening as well as cyclic loading. Cam-Clay type soil models offer an attractive entry to the modeling of soils through FEA plasticity theory because of the sequential developments of these types of models and the general acceptance of such models in the geomechanics community. Cam-Clay models began with infinitesimal strain assumptions and have been developed to the case of finite strains. Furthermore, Cam-Clay models have been extended to capture the cyclic behavior of soils (Carter et al., 1982).

7. CONTINUUM SOIL PLASTICITY FE IMPLEMENTATION

The rate form of the constitutive equations can be used with other plasticity equations to define a set of differential equations that can be integrated using implicit integration methods or the return mapping algorithm. The plasticity equations presented in section 3.1 are typically solved in either an implicit or explicit fashion. The explicit solution, while easier, requires small time steps for stability. Implicit schemes are more computationally intensive, but are stable; small time steps may still be needed for accuracy.

7.1 Integration Algorithm

The constitutive equations that govern the behavior of the hyperelastic elastoplastic finite deformation Cam-Clay model (Borja et al., 1996) are summarized in this section for convenience. The yield function is defined by

\[ f = f(P, Q, P_c) = \frac{Q^2}{M^2} + P(P - P_c) \]  

(28)
where \( P, Q, P_c, \) and \( M \) are the finite deformation analogs of the parameters defined for the infinitesimal case using the Kirchhoff stress tensor. The hardening law expressed in terms of the plastic component of the volumetric strain is given by

\[
\frac{\dot{J}^p}{J^p} = -\Theta \dot{\varepsilon}_v^p, \quad \dot{\varepsilon}_v^p = \frac{\dot{J}^p}{J^p}, \quad \Theta = \frac{1}{\hat{\lambda} - \hat{\kappa}}
\]  

(29)

where \( J^p \) is the plastic component of the Jacobian. The parameters \( \hat{\lambda} \) and \( \hat{\kappa} \) can be calculated from the corresponding infinitesimal model analogs. The discrete flow rule at time \( t_{n+1} \) for implicit time integration in the space defined by the elastic Eulerian logarithmic stretches can be written as

\[
\varepsilon_{\varepsilon}^{n+1} = \varepsilon_{\varepsilon}^{n} - \Delta \phi_{\varepsilon}^{n+1} \frac{\partial g}{\partial \beta} \bigg|_{n+1}
\]

(30)

where \( \beta \) is the Kirchhoff stress tensor, and \( \Delta \phi \) is a plastic multiplier.

The above equations can be shown to lead to the following set of equations that can be used to define a scalar return mapping algorithm (Borja, 1998) in the invariants of the elastic logarithmic stretches

\[
\varepsilon_{\varepsilon}^{c} = \varepsilon_{\varepsilon}^{c_{\text{inf}}} - \Delta \phi \frac{\partial g}{\partial P}, \quad \varepsilon_{\varepsilon}^{c_{\text{inf}}} = \varepsilon_{\varepsilon}^{c_{\text{inf}}} - \Delta \phi \frac{\partial g}{\partial Q}
\]

\[
f \left( P, Q, P_c \right) \leq 0, \quad \Delta \phi \geq 0, \quad \Delta \phi f \left( P, Q, P_c \right) = 0
\]

(31)

An example implicit integration scheme for the finite deformation Cam-Clay plasticity soil model can be developed by considering Eq. 32 as a set of simultaneous nonlinear equations. An application of the Newton-Raphson method can be used to solve this set of nonlinear equations. To this end, the residual vector \( r \) and the vector of unknowns \( x \) are written as follows:
The Newton-Raphson solution procedure requires the iterative solution of the algebraic system
\[
\left( \frac{\partial r^p}{\partial x^p} \right) \Delta x^p = -r^p , \quad \text{where } \Delta x^p \text{ is the vector of Newton differences. A closed form expression for the consistent tangent operator } \left( \frac{\partial r^p}{\partial x^p} \right) \text{ can be found and the algorithm can be made more efficient by the application of the static condensation technique (Borja et al., 1996).}
\]

8. INTEGRATION OF SOIL PLASTICITY WITH ANCF/MBS ALGORITHMS

The FE implementation of the soil mechanics plasticity equations requires the use of an approach that allows employing general constitutive models. The vehicle/soil interaction can lead to a significant change in geometry that cannot be captured using finite elements that employ only translational displacement coordinates without significant refinement. In some soil applications, such a significant change in geometry may require the use of elements that employ gradients and accurately capture curvature changes. This requirement can be met using the FE absolute nodal coordinate formulation (ANCF).

8.1 Absolute Nodal Coordinate Formulation (ANCF)

ANCF finite elements do not employ infinitesimal or finite rotations as nodal coordinates; instead, absolute slopes and displacements at the nodal points are used as the element nodal coordinates. The position vector \( r^j \) of an arbitrary point on element \( j \) can be defined in a global coordinate system \( XYZ \) as
\[
r^j = S^j \left( x^j, y^j, z^j \right) e^j(t) .
\]
In this equation, \( x^j, y^j, \) and \( z^j \) are the
element spatial coordinates, $S^j$ is the shape function matrix, $\mathbf{e}'$ is the vector of element nodal coordinates, and $t$ is time. The nodal coordinate vector $\mathbf{e}^{jk}$ at node $k$ can be defined as follows:

$$
\mathbf{e}^{jk} = \begin{bmatrix}
\mathbf{r}^{jkT} & \left( \frac{\partial \mathbf{r}^{jk}}{\partial x^j} \right)^T & \left( \frac{\partial \mathbf{r}^{jk}}{\partial y^j} \right)^T & \left( \frac{\partial \mathbf{r}^{jk}}{\partial z^j} \right)^T
\end{bmatrix}^T
$$

(33)

Fully parameterized ANCF finite elements allow using a general continuum mechanics approach to define the Green-Lagrange strain tensor $\varepsilon = \left( J^T J - I \right) / 2$, where $J$ is the matrix of position vector gradients. In dynamic soil problems, ANCF leads to a constant inertia matrix and to zero Coriolis and centrifugal forces. The mass matrix obtained using ANCF finite elements can always be written as $M^j = \int_V \rho' S'S' dV^j$, where $\rho'$ and $V^j$ are, respectively, the mass density and volume of the finite element. ANCF finite elements allow for straightforward implementation of general constitutive models including the continuum mechanics-based soil models discussed in this paper.

8.2 Dynamic Equations

For a finite element or a deformable body, the principle of virtual work can be written using the reference configuration as

$$
\int_V \rho \ddot{\mathbf{r}}^T \delta \mathbf{r} dV + \int_V \mathbf{\sigma}_{p2} : \delta \mathbf{e} dV - \int_V \mathbf{f}_b^T \delta \mathbf{r} dV = 0
$$

(34)

In this equation, $V$ is the volume, $\rho$ is the mass density, $\mathbf{r}$ is the global position vector of an arbitrary point, $\mathbf{\sigma}_{p2}$ is the second Piola Kirchhoff stress tensor, $\mathbf{e}$ is the Green-Lagrange strain tensor, and $\mathbf{f}_b$ is the vector of body forces. The second term in the preceding equation can be recognized as the virtual work of the elastic forces, it can be rewritten to define the generalized elastic forces, that is
\[ \delta W_s = \int_V \sigma_{p_2} : \delta \epsilon dV = Q'e \delta e \]  \hspace{1cm} (35)

Where \( \delta e \) is the virtual change in the nodal coordinates associated with a particular ANCF finite element or a body, and \( Q'e \) is the vector of the generalized elastic forces. The vector of elastic forces often takes a fairly complicated form, especially in the case of plasticity formulations, and is obtained using numerical integration methods. The principle of virtual work leads to the following equations of motion:

\[ M\ddot{e} + Q_s - Q_e = 0 \]  \hspace{1cm} (36)

where \( M \) is the symmetric mass matrix, and \( Q_e \) is the vector of body applied nodal forces.

8.3 Integration with MBS Algorithms

The objective of this paper is to present a review of soil mechanics formulations that can be integrated with computational MBS algorithms used for the virtual prototyping of vehicle systems. These algorithms allow for modeling rigid, flexible, and very flexible bodies. The small deformation of flexible bodies in vehicle systems are often examined using the floating frame of reference (FFR) formulation. Therefore, efficient modeling of complex vehicle system dynamics requires the implementation of different formulations that can be used for rigid body, small deformation, and large and plastic deformation analyses. A Newton-Euler or Lagrangian formulation can be used to model rigid bodies, the FFR formulation that employs two sets of coordinates (reference and elastic) can be used to model small deformations, and ANCF finite elements can be used to model large and plastic deformations including soil deformations.

MBS algorithms are designed to exploit the sparse matrix structure of the resulting dynamic equations. Because ANCF finite elements lead to a constant inertia matrix, Cholesky coordinates can be used to obtain an identity generalized mass matrix, leading to an optimum
sparse matrix structure. Computational MBS algorithms are also designed to solve a system of differential and algebraic equations. The differential equations define the system equations of motion, while the algebraic equations define the joint constraints and specified motion trajectories. The nonlinear algebraic constraint equations can be written in a vector form as \( C(q,t) = 0 \), where \( q \) is the vector of the system generalized coordinates, and \( t \) is time. Using the constraint equations and the equations of motion, the augmented form of the equations of motion can be written as (Shabana, 2005):

\[
\begin{bmatrix}
M_{rr} & M_{rf} & 0 & C_{q_r}^T \\
M_{fr} & M_{ff} & 0 & C_{q_f}^T \\
0 & 0 & M_{aa} & C_{q_a}^T \\
C_{q_r} & C_{q_f} & C_{q_a} & 0
\end{bmatrix}
\begin{bmatrix}
\ddot{q}_r \\
\ddot{q}_f \\
\ddot{q}_a \\
\lambda
\end{bmatrix}
= \begin{bmatrix}
Q_r \\
Q_f \\
Q_a \\
Q_c
\end{bmatrix}
\] (37)

where subscripts \( r, f \) and \( a \) refer, respectively, to reference, elastic, and absolute nodal coordinates, \( M_{rr}, M_{rf}, M_{fr}, M_{ff} \) are the inertia sub-matrices that appear in the FFR formulation, \( M_{aa} \) is the ANCF constant symmetric mass matrix, \( C_q \) is the constraint Jacobian matrix, \( \lambda \) is the vector of Lagrange multipliers, \( Q_r, Q_f, \) and \( Q_a \) are the generalized forces associated with the reference, elastic, and absolute nodal coordinates, respectively, and \( Q_c \) is a quadratic velocity vector that results from the differentiation of the kinematic constraint equations twice with respect to time, that is \( C_q \ddot{q} = Q_c \). The generalized coordinates \( q_r \) and \( q_f \) are used in the FFR formulation to describe the motion of rigid and flexible bodies that experience small deformations. The vector \( q_a \) is the vector of absolute nodal coordinates used to describe the motion of flexible bodies that may undergo large displacement as well as large and plastic deformations as in the case of soils. The vector \( q_a \) includes the nodal coordinate vector
of all ANCF bodies, including the ANCF soil coordinates. Similarly, the mass matrix $\mathbf{M}_{\text{soil}}$ includes the soil inertia matrix as well as the inertia of the vehicle components modeled using ANCF finite elements. This mass matrix can be made into an identity mass matrix using Cholesky coordinates, leading to an optimum sparse matrix structure. The generalized force vector $\mathbf{Q}_s$ includes also the contributions of the forces $\mathbf{Q}_e$ and $\mathbf{Q}_s$ of Eq. 37. The vectors $\mathbf{Q}_e$ and $\mathbf{Q}_s$ account for the vehicle soil interaction forces.

The solution of Eq. 38 defines the vector of accelerations and Lagrange multipliers. The independent accelerations can be integrated to determine the coordinates and velocities including those of the soil. The soil coordinates can be used to determine the total strain components that enter into the formulation of the soil constitutive equations. Knowing the strains, the soil properties, yield function, and the flow rule; the state of soil deformation (elastic or plastic) can be determined as previously discussed in this paper. Knowing the state of deformation, the constitutive model appropriate for this state can be used to determine the elastic force vector $\mathbf{Q}_e$. Therefore, the structure of Eq. 38 allows for systematically integrating soil models into MBS algorithms used in the virtual prototyping of complex vehicle systems.

9. SUMMARY
In this paper, soil mechanics formulations that can be integrated with FE/MBS algorithms to study vehicle dynamics are reviewed. Several simple models including analytical terramechanics models are discussed. Bekker’s model as well as other parametric and analytical terramechanics models have been used in the study of track/soil interaction and can be implemented in MBS algorithms using simple discrete force elements. These simple models, however, have serious limitations because they do not capture the distributed elasticity and plasticity of the soil. The
limitations of the discrete element method (DEM) are also discussed in this paper. More general continuum plasticity soil formulations are reviewed. Among the continuum soil plasticity formulations discussed in this paper are the Mohr-Coulomb, Drucker-Prager, modified Cam-Clay, Barcelona Basic, elasto-plastic cap model for partially saturated soil, viscoplastic cap, and bounding surface plasticity unsaturated models.

Existing computer codes that are used in the soil modeling are reviewed. The integration algorithm that is commonly used to solve the plasticity equations is discussed. This algorithm can be integrated within the absolute nodal coordinate formulation (ANCF) to develop a computational procedure that allows for the study of vehicle/soil interaction. The ANCF/soil model integration will be the subject of future investigations.

ACKNOWLEDGEMENTS

This research was supported by the US Army Tank-Automotive Research, Development, and Engineering Center (TARDEC) (Contract # W911NF-07-D-0001).
REFERENCES


<table>
<thead>
<tr>
<th>Category</th>
<th>Soil Model</th>
<th>Hardening or Work Hardening</th>
<th>Fracture</th>
<th>Cyclic Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Continuum Based Models</strong></td>
<td>Mohr Coulomb Elastic Perfectly-Plastic</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Drucker-Prager</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modified Clam-Clay</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>Elasto-Plastic Cap Model for Partially Saturated Soils</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visco-Plastic Cap Model</td>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bounding Surface Plasticity Un saturated Soil Model</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>Elasto-Plastic Barcelona Basic Model</td>
<td>O</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td><strong>Terramechanics Models</strong></td>
<td>Bekker's Soil Model</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>Modified Bekker Soil Model</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td><strong>Particle Based Models</strong></td>
<td>Discrete Element Method</td>
<td></td>
<td></td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>Smoothed Particle Hydrodynamics</td>
<td></td>
<td></td>
<td>O</td>
</tr>
</tbody>
</table>

Table 1 Summary of Soil Models
Figure 1 Response of soil with respect to shearing (Whitlow, 1995)

Figure 2 Stress at a point R units away from the point load (Wong, 2010)
Figure 3 Contact area under a circular loading area (Wong, 2010)

Figure 4 Stress at a point due to a rectangular loading area (Wong, 2010)
Figure 5 Idealized flexible track and terrain interaction (Wong, 2010)

(a) Mohr-Coulomb  
(b) Drucker-Prager  
(c) Strain-hardening cap model  
(d) Lade-Duncan

Figure 6 Failure surfaces in stress space (Brinkgreve, 2005)

Figure 7 The Modified Cam-Clay model (An, 2010)
Figure 8 Static yield surface for Cap Model (Kohler, 2007)

Figure 9 Yield surface of the extended cap model in terms of net stress and matric suction (Kohler, 2007)