The place of expansive clays in the framework of unsaturated soil mechanics

Stephen Fityus, Olivier Buzzi *

Centre for Geotechnical and Materials Modelling, School of Engineering, University of Newcastle, NSW 2308, Australia

Abstract

This paper reviews and re-examines the place of expansive clay soils within the framework of unsaturated soil mechanics. Direct and indirect physical evidence is presented and discussed, which recognizes that natural expansive clays remain perpetually saturated (i.e. degree of saturation is one) as they gain or lose water over the suction range that is of interest to engineers, and that is likely to be encountered under normal field conditions. As a consequence, expansive soils do not have a unique saturated water content. Perpetual saturation is a consequence of the small particle size of smectitic clays, which in a structured clay soil, leads to very small pores and a high air entry value.

The behaviour of structural soil elements is distinguished from the behaviour of the total soil mass, to consider expansive clay soils as composite soil materials on a macro scale, composed of saturated, structured soil peds, separated by air-filled cracks. It is considered that peds of expansive clay soils in natural environments evolve, under cycles of wetting and drying, to attain an equilibrium micro-structure which allows them to shrink and swell by significant amounts in a completely reversible way. As peds remain saturated during wetting and drying, volume change behaviour is simplified, because the volume of water gained or lost from the soil equates directly to the volume change of the soil peds, and to the change in crack volume in the soil mass. This behaviour is discussed in the context of the normal shrinkage concept as used in soil science. It is suggested that constitutive models formulated in a continuum mechanics approach may be unsuited to the modeling of desiccating expansive clay soils even though they are commonly used.

© 2008 Elsevier B.V. All rights reserved.

Contents

1. Introduction ............................................................... 0
2. Perpetual saturation of expansive clays .................................................. 0
3. Significance of pore size and of soil structure ............................................... 0
4. The significance for the volume change .................................................. 0
5. Relevance of the unsaturated soils framework ............................................... 0
6. Conclusions ............................................................... 0
References ................................................................ 0

1. Introduction

The early study of unsaturated soils in engineering was strongly influenced by the need to understand the hydrodynamics and hydrostatics of soil–water systems and to manage the readily apparent problem of wetting- and drying-induced movements in swelling soils (e.g. Philip, 1969; 1972, Parcher and Liu, 1965; Miller, 1975). As shrinking clays are usually associated with water deficient situations, their status as unsaturated soils was established.

The first dedicated forum for the consideration of unsaturated soils to be significantly recognized by geotechnical engineering researchers was the ‘Pore Pressure and Suction in Soils’ conference (Butterworths, 1961). It was followed by a series of 7 dedicated expansive soils conferences that took place, between 1965 and 1992. In the beginning, what was known about suction and unsaturated soils borrowed heavily from the field of soil science. As geotechnical engineers came to terms with the topic, the study of unsaturated soils became more closely aligned with mainstream soil mechanics, and it became increasingly apparent that swelling clays were simply a subset of...
engineered soils, as partial saturation is an important condition in all soils. Accordingly, consideration widened to become the generalized field of unsaturated soil mechanics, and the new series of International Unsaturated Soils Conferences began in Paris in 1995.

Gradually we have become increasingly fascinated with the challenge of unsaturated soils and the quest for a unifying constitutive model that allows all unsaturated soils to be modeled in a continuum mechanics framework. However, along the way we seem to have lost touch with the special problem we started with. In some respects, truly expansive soils are now on the fringe of unsaturated soils research, with much of the current research focusing only on compacted swelling clays in specific applications, in particular, the use of bentonite clays in liners for nuclear waste disposal (e.g. Delage et al., 1998; Gens et al., 2002; Komine, 2004). Due to the development of a theoretical framework for unsaturated soils and of numerous hydromechanical models, expansive clay soils are treated as if they were unsaturated soils although the fundamental mechanisms of their volume change behaviour are different.

This paper revisits some of the fundamentals of swelling clay behaviour in an attempt to reconcile what has long been appreciated about expansive clays, with the state-of-the-art in unsaturated soil mechanics research. The ultimate objective of the paper is to show that the framework of unsaturated soil does not apply properly to expansive soils even though it is commonly assumed so. Further work is needed to model the behaviour of truly expansive clays in a constitutive sense.

2. Perpetual saturation of expansive clays

Expansive clays are dominated by clay minerals with the potential for crystalline swelling, such as minerals of the smectite group. These are recognized as having very small particles, even among the clay minerals (Meunier, 2006). In an expansive clay soil with small particles and small pores, it is commonly observed that substantial matrix suctions can be sustained in the saturated soil (Marcial et al., 2002; Peron et al., 2007; Buzzi et al., 2007), before air entry occurs. This is shown by the data in Fig. 1. It suggests that the soil sustains significant changes in water content without desaturating. The basic structural units of such soils are, in fact, saturated soils over a very wide range of suction, and in truly expansive clay soils, this corresponds to most or all of the suction change range experienced under field condition (Nelson and Miller, 1992).

In a saturated element of expansive soil, the only air–water interfaces are on the outer surfaces of the element. Hence, a situation can be envisaged where, as water is evaporated from the soil element, the radii of the water phase menisci at the soil element surface become smaller, drawing deeper into the outer soil pores, exerting tension in the continuous internal soil water. This tension tries to draw water from the hydration layers of the counter-ions of the clay crystals, and when great enough, layers are lost and the interlayer spaces contract (Delage, 2007), causing the clay particles to shrink and the soil volume to reduce over a wide range of suction, without corresponding air entry (as shown in Fig. 1(c) where $e$ reduces from 1 to 0.5 without air entry). Volume changes (shrinking and swelling) in expansive clays are associated with changes in matric suction, driven by osmotic phenomena related to the hydration potential of counter-ions (Marshall and Holmes, 1988). The significance of this behaviour has profound implications for the behaviour of expansive clay soils. This will be discussed in a subsequent section.

3. Significance of pore size and of soil structure

Perpetual saturation of expansive clay soils can only be realized if the pore spaces are small enough to prevent air entry as suction increases. In any granular soil, the pore size distribution is a function of the particle size distribution. In a poorly graded soil, the size of the pores will reflect the size of the grains. In a well graded soil, the size of the pores will be strongly influenced by the size of the finest fraction(s), whilst the size of the largest pores and the pore size distribution, determine the air entry value and the slope of the soil water characteristic curve, respectively (Assouline et al., 1998; Marinho, 2005).

The structure of a swelling clay is far more complex and consideration cannot usually be limited to discrete soil particles. It is well recognized that natural clay soils and many reconstituted clay soils display a hierarchy of structures. These are variously referred to as particles (Quirk and Murray, 1991), lamellae (Oades and Waters, 1991) or micelles at the finest level; grains, crystals or quasi-crystals (Quirk and Murray, 1991) at the particle level; micro-aggregates (Oades and Waters, 1991) and clusters at the microscopic level; and peds (Thomasson, 1978), macro-aggregates (Oades and Waters, 1991) or prisms (Cabidoche and Ruy, 2001) on the macro scale. This proliferation of terminology is an artifact of soil science, as the tillage of soils in agricultural production can create soils with a wide variety of transient structural elements that are not all necessarily in

---

Please cite this article as: Fityus, S., Buzzi, O., The place of expansive clays in the framework of unsaturated soil mechanics, Applied Clay Science (2008), doi:10.1016/j.clay.2008.08.005

---

![Fig. 1. Natural expansive clay soils with high air entry values. a) Saturation degree vs suction for Maryland clay (adapted from Buzzi et al., 2007). b, c) Changes in saturation degree and void ratio with increasing suction (vapor equilibrium technique) in natural Romainville clay (after Cui et al., 2006).](image-url)
equilibrium with the prevailing natural environment. The existence of lamellae (the basic unit layers of a clay crystal) and quasi-crystals (stacks of lamellae) are well established as fundamental structures in clay soils. It is reasonable, on the basis of observations made in desiccated clay soils, to simplify larger structural orders by grouping aggregates and clusters into a single structural element group which will be referred to as peds. For the purposes of this discussion, a ped is a naturally occurring, structured soil element within a ripened (Pons and Van der Molen, 1973) heavy clay soil, that is bounded by discontinuities (typically cracks) which separate it from adjacent elements of similar form. The ped is the basic unit of a natural heavy clay soil at a macro scale.

The absolute scale of peds varies, depending on many factors. As suggested above, several orders of ped sizes may exist within a soil mass. However, it is the internal structure of peds that is of particular interest to the present discussion. Clay particles are small, and so, the inter-particle pores are small. The particle sizes of natural montmorillonites are well distributed between 50 and 1600 nm but with as much as 20% smaller than 50 nm (Roberson et al., 1998). Spherical particles 50 nm in diameter would produce pores with air entry values of around 5 MPa, based on Jurin's law (Fredlund and Rahardjo, 1993). However, basic smectite clay quasi-crystals are sheet-like, with length/thickness ratios of the order of 100 (Oades and Waters, 1991), having crystal thicknesses of only 2 to 100 nm and lamellae thicknesses of around 1 nm. The particle packing arrangement of sheet-like particles is difficult to characterize, and the pore size distribution is so small that it cannot readily be observed by ESEM or CT scanning, or even inferred by techniques such as mercury porosimetry. It is likely that the size of pores within peds of expansive clay are much smaller than the measured clay crystal size, but greater than the crystal thickness. Quirk and Murray (1991) suggest that inter-particle pores exist as slit- and wedge-shaped pores between the lamellae and/or clay crystals in a soil and that these are less than 10 nm wide. A similar model is proposed by Meunier (2006). Oades and Waters (1991) suggest that in fine, highly expansive clays such as the montmorillonites, the smallest of these pores could be as small as 3–4 nm, which means that they could theoretically resist air entry up to suctions as great as 80 to 100 MPa. Such high values of air entry values have been found in numerous studies of expansive clays reconstituted and consolidated from slurries (Aylmore and Quirk, 1962; Marcial et al., 2002; Villar, 2000). However, despite the scarcity of soil water characteristic curve measurements on undisturbed heavy clays, values of around 10–20 MPa can be found (Cui et al., 2006; Holmes, 1955; Buzzi et al., 2007).

Experimental research in soil mechanics typically considers soil in one of three conditions: undisturbed (the natural or field condition), remolded (by the consolidation from a soil slurry) and compacted (compact from some unsaturated, crumbed state). Quite clearly, soils prepared by remolding or compaction will have a structure very different to natural soils, and so, different arrangements and size distributions of pores. These differences are greatest for clay soils. Clay soils prepared from a slurry are usually saturated, and will have a relatively uniform pore size distribution with the largest pores determined by the magnitude of the imposed consolidation load. On the contrary, the compaction of granulated clays results in an artificially large pore size network as mentioned in Lloret et al. (2003) or highlighted by Jayawickrama and Lytton (1993). Although there are still likely to be micro-aggregates (granules) of clay that remain saturated (or unsaturated beyond their limit of shrinkage) in the compacted soil, they are sufficiently small so that the soil approximates an unsaturated granular soil.

Natural soils have usually experienced numerous wetting/drying cycles of different suction amplitudes, during which their structure has evolved. The idea of a “ripened” structure for heavy clay soils is not new in the soil science literature, although it has not received detailed attention. Pons and Van der Molen (1973) quote two Dutch publications where the concept of a physical ripening process was developed. More recently, Monroy et al. (2007) have examined the evolution of pore size distribution when subjecting clay to wetting and loading. Their results suggest that the ripening process tends to reduce the size of inter aggregate pores. This leads to higher air entry values and resulting in perpetual saturation. Alonso et al. (1999) have also shown that remolded clays exhibit significant permanent strains when subjected to cyclic drying and wetting but also, that the volume change behaviour becomes increasingly reversible as the number of cycles increases, as if an equilibrium state is being achieved.

The above discussion has not considered the effect of overburden stresses or external stresses, although it is widely appreciated that these are important factors for the swelling/shrinkage of the soil, (Philip, 1972; Groenewold and Bolt, 1972) and for the water retention curve (Stroosnijder and Bolt, 1984; Philip, 1969). Whilst the authors believe these can be easily accommodated by an extension of the concepts presented here, they will not be considered further in the present discussion, as their consideration is not required to achieve its conclusion.

4. The significance of the volume change

The persistence of saturation over a wide range of suction has a profound significance for the treatment of volume change models for expansive soils. To begin with, there is no unique saturated water content value as was recognized early by Terzaghi (1925). However, the most significant aspect ensuing from this situation is that soil peds undergo drying and shrinkage without air entry. This means that the volume decrease of a clay ped during shrinkage should be exactly equal to the volume of water lost from the ped. This affords great simplification in the modeling of volume changes in expansive soils at any scale, including the scale of an entire soil mass. If the reference soil volume is taken as the volume at field capacity, when no cracks are present, then the volume change which occurs in this soil mass due to drying is equal to the volume of water lost in going from the field capacity water content to the drier water content. This is easily calculated. It is not necessary to consider the effects of partial saturation: at any water content above the air entry value, the soil mass has two components only: cracks which are completely dry and peds which are completely saturated, and the volume lost by one is equal to the volume gained by the other.

The idea of a one-to-one relationship between volume change and water lost is not new in soil science or soil mechanics (Johnston and Hill, 1944; Croney and Coleman, 1953; Aitchison and Holmes, 1953; Newmann and Thomasson 1979; Chan, 1982; McGarry, 1988; Daniells, 1989; Cabidoche and Ruy, 2001; Cabidoche and Voltz, 1995; Taboada et al., 2001) but what has to be specified pointed out is that normal shrinkage extends to high values of suctions for expansive clays. For example, Holmes (1955) studied swelling of an illitic clay, noting that it remained saturated over a suction range from 10 kPa to 5 MPa. The significance of the one-to-one correspondence between water content change and soil volume change, and its subsequent implications for expansive soil modeling, were also appreciated by Swarbrick (1994), forming the basis of his “unit water balance” model for expansive clay soils.

These ideas can be considered in the context of the soil shrinkage model of McGarry and Malaffant (1987) for unconfined soil units, which proposed that shrinkage can be divided into a series of phases. Starting from field capacity, there is a phase where there is water loss with little or no volume change, as water is drained, under gravity, from blind cracks and isolated macro-voids/defects. Then there is a phase of normal shrinkage, where the structural elements of the soil reduce in volume, by an amount equal to the volume of water lost. Once the air entry suction value is reached, this is followed by the residual phase, where air enters the intergranular pores as water is lost from the pores, and the soil desaturates. It is important to note
that there is little or no corresponding shrinkage during the residual phase. McGarry and Malafant’s (1987) model, expressed in terms of specific volume and gravimetric water content, is presented in Fig. 2.

Whilst this soil shrinkage model is generally applicable to all soils, it is manifested differently in cohesive heavy clay soils and more free-draining, textured clay soils. In natural heavy clay soils, structural pores are predominantly present as cracks, and smaller non-crack pores are less common. Non-crack pores can be fully occluded within the peds, or connected to the crack network. Connected cracks can be considered as extensions of the crack network, and so, behave as the cracks do, as true structural pores, in which water movement is dominated by gravity. Occluded pores do not behave as structural pores, as they are completely surrounded by intact, low permeability clay, and so they cannot drain freely. In fact, because they are surrounded by clay with a very low air entry value, air is unable to enter these large pores until the air entry value of the entire ped is reached. Hence, water gain/loss from these pores is consistent with the mechanism of normal shrinkage, and not structural shrinkage.

If these ideas are considered in the context of a mechanism involving swelling clay particles, it can be speculated that volume change during normal shrinkage is the sum of the volume lost through the collapse of clay quasi-crystals as interlayer water is extracted by increasing suction, and water expelled from inter-particle pores, as the shrinking particles cause the pores to become compressed (Popescu, 1986). Taking this idea further, a situation can be envisaged where the particle shrinkage that occurs at low suctions causes pore shrinkage, which in turn raises the air entry value, leading to more normal shrinkage, particle shrinkage and further pore size reduction, etc. That is, the resistance to air entry increases significantly as the soil water becomes harder to extract, thus maintaining the soil in a saturated state over a wide range of suction values. Such a concept was also suggested by Marinho (2005).

The onset of the residual phase corresponds to the removal of all of the readily extractable interlayer water, with the subsequent stiffening of the soil skeleton, and displacement of the residual pore water by air. The transition from normal shrinkage to residual shrinkage effectively marks the limit of shrinkage, although the precise relationship between air entry and the shrinkage limit is not straightforward and it cannot be explored further here in the space available.

5. Relevance of the unsaturated soils framework

Although cracked clay soil masses are partially saturated in an overall sense, on a microscopic scale, the peds are saturated up to high suction values. This idea is shown schematically in Fig. 3. Homogenization techniques are commonly used to describe a soil mass as a continuum (e.g. Raats, 2002) provided that the physics of the phenomena involved prevail (e.g. water flow). However, the porosity in an expansive soil exists in the form of macroscopic desiccation cracks, between large peds of saturated clay, giving the soil the macroscopic characteristics of a very coarse granular soil. This structure does not develop surface films and water bridges between peds in the same way as occurs between grains in partially saturated silts and sands. As a consequence, the generalized constitutive models for partially saturated soils developed in the continuum mechanics framework are not directly applicable to natural expansive clay soils because the mechanical models do not apply.

Most of the unsaturated soil constitutive models are based on the fact that suction increases with decreasing degree of saturation (Gens and Alonso, 1992; Alonso et al., 1990; Sheng et al., 2008; Wheeler et al., 2003; Alonso et al., 1999) and very few of them incorporate suction

![Fig. 2. The 3 phase shrinkage model of McGarry and Malafant (1987).](image)

![Fig. 3. Schematic representation of a desiccated clay soil showing the differentiation between the unsaturated soil mass and the saturated soil elements.](image)
change within the saturated range (e.g. Nuth and Lalou, 2007). Moreover, even when this behaviour is incorporated, it is over a range of water content far less significant than the water content range over which natural expansive soils remain saturated. The models cited above seem to be able to describe the behaviour of particular situations such as compacted bentonite (Lloret et al., 2003) only because, as mentioned previously, an artificially large pore size network has been imparted to the soil by the compaction process, leading to unsaturated soil specimens. Most of the unsaturated soil models are based on Terzaghi's principle in the saturated range (e.g. Sheng et al., 2008; Wheeler et al., 2003; Alonso et al., 1999). These latter could be adapted to describe behaviour in the saturated range, by transferring the suction to an equivalent net stress once the degree of saturation reaches one, thereby providing that the principle of effective stresses formulated by Terzaghi (1936) holds for expansive clays. The problem would then be treated as a classical saturated soil problem. However, it has been shown in the literature that electro-chemical forces within expansive clays are a significant component of the mechanical equilibrium so that the principle of effective stress, which does not account for them, does not always hold for expansive clays (Lambe and Whitman, 1959; Lambe, 1960; Sridharan and Venkatappa Rao, 1973; Hueckel, 1992). Consequently, the suction change in the saturated range cannot be simulated using an equivalent change in mean or net stress because the response of the soil might be different.

As noted above, heavy clay soils with a ripened structure experience most of their volume change whilst remaining saturated. This makes the concept of using the degree of saturation as a variable (stress state or otherwise) meaningless and the water content should be preferred to describe the level of hydration of the soil. For instance, the common use of unsaturated permeability, such as Van Genuchten’s (1980) formulation, in unsaturated soil mechanics should be replaced by a saturated permeability depending on the water content (several examples given in Raats, 2002), on the porosity (e.g. Kozeny–Carman’s formulation (Kozeny, 1927) or a more appropriate derivative) or on void ratio (Gibson et al., 1967). Moreover, unlike the trend usually considered for unsaturated soils, the permeability of a cracked clay soil (dual permeability system with cracks permeability and bulk permeability) decreases with increasing water content. Indeed, the cracks tend to close as the soil gets wet significantly reducing their permeability. This cannot be fully compensated by the associated increase of bulk permeability with wetting resulting in an overall reduction of soil mass permeability whilst the water content increases.

Further, for most soils with a liquid limit of more than 25%, the SWCC plots as a straight line over the usual suction range of interest (up to 10,000 kPa) (Marinho and Chandler, 1993; McQueen and Miller, 1974). Beyond this, air enters the soil and after this, the rate of further shrinkage is greatly reduced. It is suggested that, small the volume changes experienced in this unsaturated, residual shrinkage phase are consistent with the behaviour of unsaturated soils in the generalized sense. In a non-clay granular soil, the mechanism of volume change is envisaged as skeletal compression in response to the net tension in the air–water interface films (“contractile skins” of Fredlund and Rahardjo, 1993). Only after an expansive clay ped desaturates, are such air–water-interface films present throughout the soil, and the tension they exert replaces particle shrinkage as the principal factor affecting soil volume. Hence, the transition from normal shrinkage to residual shrinkage marks a change in the principal mechanism of soil volume change, and a good constitutive model should incorporate this. As a general rule, it is suggested that the clays that are likely to cause real engineering problems in the classic expansive soil sense are those which have sufficient smectic soil content to fill the spaces between larger grains so that the maximum pore sizes are small enough to resist drying up to suctions of at least 10 MPa, which seems to be within the practical suction range in most real situations (McKeen, 1992). Considerable further work is needed, however, to better define appropriate ranges of applicable behavioural models for particular soils, and to fully understand the significance of the shrinkage limit in the context of these models.

6. Conclusions

The unique and important properties of expansive clay soils that were recognized by soil scientists and geotechnical pioneers of the past, have to some extent been forgotten or overlooked by unsaturated soils researchers of the present. There is a need to appreciate that truly expansive soils are not just the extreme case of the more generally defined unsaturated soils. Rather, they are a special case of saturated soils where large volumetric strains occur in response to suction changes over many orders of magnitude, corresponding to large changes in the saturated water content of the soil, at a degree of saturation that remains very close to one. They are able to remain saturated because of a significant fraction of very fine smectitic clay that limits the maximum pore size to very small values that preclude air entry, and which is able to adsorb and desorb water by varying the thickness of the counter-ion hydration layers, thereby accommodating large volumes of shrinkage and swelling of the soil particles themselves. The net result is a soil skeleton that responds to suction by collapsing both its soil particles and its interparticle pores, so that water is lost without a corresponding entry of air. As a consequence, there is theoretical equality between the volume of water gained or lost, the change in the volume of the saturated soil elements and the change in the volume of the desiccation cracks. This effectively means that in a cracked expansive soil mass, the volume of water lost from the soil elements is equal to the volume of air gained by the shrinkage crack network. This behaviour affords significant simplifications for the modeling of volume change in natural clay soils up to their pedal air entry value, which in truly expansive soils, is likely to be close to the highest suction likely to be encountered under field conditions. The framework of unsaturated soil mechanics as it is currently developed is not appropriate to describe the mechanical behaviour of expansive soils, considering the perpetual saturation, the irrelevance of degree of saturation as a soil parameter and the uncertainty about the validity of the effective stress principle. Beyond the air entry values, volume change behaviour of soil elements is consistent with the volume change behaviour of non-expansive soils. The boundary between expansive and non-expansive soils is not well defined, and more consideration is needed to better understand the behaviour of slightly expansive clay soils such as broadly graded soils with small smectite (and illite) contents and soil dominated by non-swelling clays such as kaolinite.

References

