



A comparative study of stress integration methods for the Barcelona Basic Model

Wojciech T. Sołowski^{a,*}, Matthias Hofmann^b, Günter Hofstetter^c, Daichao Sheng^a, Scott W. Sloan^a

^aARC Centre of Excellence for Geotechnical Science and Engineering, the University of Newcastle, 2308 NSW, Australia

^bALPINE BeMo Tunnelling GmbH, Innsbruck, Austria

^cUnit for Strength of Materials and Structural Analysis, University of Innsbruck, Austria

ARTICLE INFO

Article history:

Received 6 May 2011

Revised in revised form 13 November 2011

Accepted 12 March 2012

Keywords:

Unsaturated soils

Elasto-plasticity

Implicit stress integration

Explicit stress integration

Barcelona Basic Model

ABSTRACT

The paper compares the accuracy and stability of implicit and explicit stress integration schemes applied to the Barcelona Basic Model. In addition, the effect of the integration scheme on the convergence of Newton–Raphson algorithm is studied. By running a Newton–Raphson algorithm for a single stress point, the number of iterations required to reach convergence gives some insight on the convergence of the finite element solution at the global level. The explicit algorithms tested incorporate substepping with error control and are based on Runge–Kutta methods of different orders or on the Richardson extrapolation method. The implicit return-mapping algorithms follow the procedures of Simo and Hughes [1] and Hofmann [2].

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

In a finite element analysis the stress increments are computed at each integration point for known increments of strain and matric suction. To this end, different stress integration algorithms can be employed. Depending on the solution strategy, they are broadly classified as explicit or implicit algorithms. Currently, explicit algorithms are generally coded with automated substepping and error control, i.e. the strain and matric suction increments are subdivided automatically into smaller subincrements and the stress is computed for those subincrements to maintain a desired accuracy. The implicit algorithms are usually coded without substepping and error control (Hofmann [2]); however, they may be enhanced with the said features (e.g. Perez-Foguet et al. [3]).

The stress integration algorithm adopted affects not only the accuracy of the computed stresses, but also the overall (global) convergence of the finite element solution and, thus, the stability of the code. Despite the fact that choice of the stress integration strategy will affect analysis time, there is a dearth of research concerning this subject (some results are given by Potts and Ganendra [4]).

Usually, in a finite element analysis, the load is applied in incremental steps. After attaining equilibrium for the current load step, the subsequent load step is applied. During calculations for a particular load step, the stress integration algorithm is called at each integration point; non-convergence of the stress integration

algorithm at a single integration point will cause failure of the solution and will require the analysis to be restarted from the previous load step with smaller load step size. This is best to be avoided as it leads to a significant increase of the computation time.

In a usual implementation of the finite element method for non-linear problems the Newton–Raphson method is used for determining the solution for a particular load step in an iterative manner. The Newton–Raphson algorithm may cause divergence when: (a) the stress integration algorithm fails to provide the stresses at a single integration point or (b) the Newton–Raphson iterative process fails to converge at the structural level. Additionally, when the employed stress integration algorithm provides inaccurate results, the computed structural response may deviate from the actual one. Maintaining predictable accuracy during stress integration is therefore desirable.

It has been shown that the explicit stress integration algorithms with substepping and error control can be generally regarded as stable and provide results within a specified accuracy [5–10]. The explicit algorithms, however, generally do not offer a quadratic rate of asymptotic convergence at the global level, which potentially slows down the FE analysis. The influence of the explicit algorithms on the stability of the global solution is not well assessed.

First order implicit schemes can be linearized exactly yielding the so called consistent elasto-plastic tangent matrix. Using such a matrix, the Newton–Raphson method offers a quadratic rate of convergence at the global level once the solution is sufficiently close to the correct solution [1,3,11–13]. However, implicit algorithms, when coded without substepping and error control,

* Corresponding author.

E-mail address: wojciech.solowski@newcastle.edu.au (W.T. Sołowski).

may be inaccurate or even diverge for large strain increments (some initial research is presented in Sołowski et al. [14]). Explicit schemes can generally be of high order which is beneficial when high accuracy is required; high order schemes may, however, be inefficient for small increments when no substepping is required. For more details on the theoretical comparison of the implicit and explicit schemes see e.g. Deuffhard and Bornemann [15].

It is the aim of the present paper to assess explicit and implicit methods of stress integration in terms of convergence and accuracy and to offer some insights of the impact of the chosen stress integration algorithm on the global convergence of the finite element solution. In addition, a comparison of the convergence rate of implicit and explicit stress integration algorithms for a single Gauss point is given. The results have been obtained for the Barcelona Basic Model (BBM, see Alonso et al. [16]), which is probably the most well-known constitutive model for partially saturated soils. It is likely that they are also relevant for other advanced critical state elasto-plastic constitutive models. The non-convexity of the BBM, which occurs at the transition between positive and negative pore water pressures (see Wheeler et al. [17]; Sheng [18]), is not dealt with in this paper. As such, the change of the pore water pressure is strictly limited to the negative (suction) zone in all numerical examples studied in this paper. For readers not familiar with the Barcelona Basic Model it is briefly described in Appendix A.

2. Compared stress integration algorithms

Explicit stress integration algorithms used in the present comparison are based on Runge–Kutta schemes of various order and on the Richardson extrapolation method. The algorithms are coded with automatic substepping and error control. The algorithms based on Runge–Kutta schemes are similar to those proposed in [5–8]. The explicit algorithms considered are described in detail in Sołowski and Gallipoli [9], who also studied their accuracy and efficiency in Sołowski and Gallipoli [10] (brief description of the algorithms is given in Appendix B). The explicit algorithms used in this comparison include (i) the second order Modified Euler scheme, (ii) the third order Nystrom scheme and (iii) the extrapolation scheme. Higher order Runge–Kutta schemes have also been tested but, since they are only of benefit for very high accuracy analyses, it has been decided not to include them in this paper. The behaviour of the high order schemes is very similar to that of the presented algorithms. For more information on the high order schemes and when it becomes beneficial to use them the interested reader may refer to Sołowski and Gallipoli [9,10].

The general return mapping scheme follows that described by Simo and Hughes [1] and Hofmann [2] (who extended the procedure to handle unsaturated constitutive models). The general return mapping algorithm has been coded in a standard way. It requires the solution of a system of nonlinear equations at the integration point level, which consist of the consistency parameter, the net stress tensor and the hardening variable. However, when applied to the Barcelona Basic Model, it has proven to diverge for large strain increments. Thus, an improved algorithm specifically designed and optimized for the Barcelona Basic Model has been developed. A short description of the improved algorithm is given in Appendix C, while details of its implementation can be found in [2].

3. Comparison of stress integration algorithms accuracy

The accuracy of stress integration algorithms has been compared for three different types of soil. The sets of model parameters and the initial stress state for each type of soil are given in Tables 1

and 2, respectively. The initial stress states are also given schematically in Fig. 1. The volumetric and shear strain increments were in the range of zero to three percent. A series of unique stress integrations was performed. All of them start at the same initial stress state, but progressively larger increments of volumetric and shear strains are considered while the suction is kept constant. This yielded the iso-error maps given in Figs. 4–8.

To create an iso-error map, an “exact” solution is required against which all the approximate solutions can be compared. Such a solution has been obtained with a fifth order Runge–Kutta method. The solution was considered to be “exact” when doubling the number of subincrements led to changes only in the 15th digit of the computed stress increments. The relative error in each integration step has been computed as

$$E_p = \left| \frac{p - p^{exact}}{p^{exact} - p^{initial}} \right| \quad \text{and} \quad E_q = \left| \frac{q - q^{exact}}{q^{exact} - q^{initial}} \right| \quad (1)$$

with p and q denoting the hydrostatic and deviatoric stress, respectively. Note that for elastic material behaviour, the stress integration is exact for both the implicit and explicit schemes. Thus, only the stresses computed for elasto-plastic material behaviour contribute to the variations from the exact solution.

The explicit schemes with substepping and error control can furnish solutions of any desired accuracy [5–10]. To achieve an accuracy that is comparable to that of the implicit schemes, and to have better parity in computation time, the integration tolerance (which roughly corresponds to the expected integration error, see [9,10] for details) was set to 10%. Typically, in practice, the integration tolerance would be set to less than or equal to 1%.

The investigated volumetric and deviatoric strain increments ranged from 0% to 3% for Compacted Kaolin and Barcelona Sandy Silt and 0.1% to 3% for Lower Cromer Till. The spacing of the strain increments for the iso-error maps is 0.1% (Compacted Kaolin and Barcelona Sandy Silt) and 0.029% (Lower Cromer Till). Therefore a total of $31 \times 31 = 961$ (Compacted Kaolin, Barcelona Sandy Silt) or $100 \times 100 = 10,000$ (Lower Cromer Till) individual stress computations are performed to create each error map. The computations for each strain increment are independent but start at the same initial state given in Table 1, and constitute a separate case of stress integration. The average relative solution error for the mean net stress and shear stress is presented in Fig. 2. This error is computed as an average of errors (calculated as in Eq. (1)) in all the increments computed for given constitutive model. One can see that the average relative error in the stresses from explicit schemes is generally well below the target of 10%. The implicit algorithms generally are characterized by larger errors than the explicit algorithms. Additionally, in the case of Lower Cromer Till, the general return mapping algorithm exhibits convergence problems for strain increments exceeding 3% (see Fig. 3). These results were not included in the error calculations presented in Fig. 2.

The iso-error maps obtained for the explicit schemes (see Figs. 4–6) resemble those compiled previously and reported in [10]. The error control used in all the tests resulted in the average relative error being below the required value of 10%. However, as the error estimate is approximate, the relative error in cases of single strain increments may exceed the prescribed tolerance. The only scheme tested that virtually guarantees that all the stresses are integrated with the required accuracy is the extrapolation method. However, when this method is used, the average error in stresses is two to three orders of magnitude smaller than the prescribed tolerance with the calculations requiring a significantly longer computing time. These issues are discussed in more detail in [10], while the algorithm is briefly described in Appendix B.

The implicit schemes (see Figs. 7 and 8) exhibit errors that increase gradually with the size of the strain increments. Thus, for

Table 1
Parameters of the Barcelona Basic Model, see also Appendix A, [16,9].

Parameter	Symbol	Compacted Kaolin	Barcelona Sandy Silt	Lower Cromer Till
Specific volume for zero suction at reference pressure	$N(0)$	1.9474	1.89	1.68523
Shear modulus	G (MPa)	3.3	3.0	7.0
Elastic stiffness parameter (elastic swelling index)	κ	0.015	0.01068	0.0077
Slope of NCL ^a at zero suction	$\lambda(0)$	0.14	0.0656	0.066
Critical line slope	M	0.82	1.003	1.2
Reference pressure	P_{ref} (kPa)	43	1	12
Atmospheric pressure	p_{atm} (kPa)	100	100	100
Swelling index for suction	κ_s	0.01	0.001	0.001
Parameter controlling cohesion increase with suction	k	1.24	0.6	0.8
Parameter controlling ratio of NCL ^a slopes at $S \rightarrow \infty$ and $s = 0$	r	0.26	0.659	0.25
Parameter controlling variation of NCL ^a slope with suction	β (kPa) ⁻¹	0.0164	0.00592	0.02

^a NCL – Normal Compression Line.

Table 2
Initial stress state and initial hardening parameter.

Parameter	Symbol	Compacted Kaolin	Barcelona Sandy Silt	Lower Cromer Till
Mean net stress	p (kPa)	45	500	6.6
Shear stress	q (kPa)	0	0	2.4
Hardening parameter	p_0^* (kPa)	55	54.94	20
Preconsolidation stress	p_0 (kPa)	90.3	828.2	20.9
Suction	s (kPa)	100	800	5

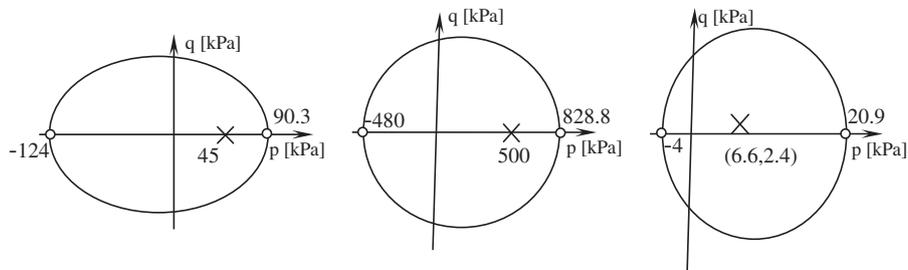


Fig. 1. Initial stress states for Compacted Kaolin (left), Barcelona Sandy Silt (middle) and Lower Cromer Till (right) in the BBM p-q yield locus intersection.

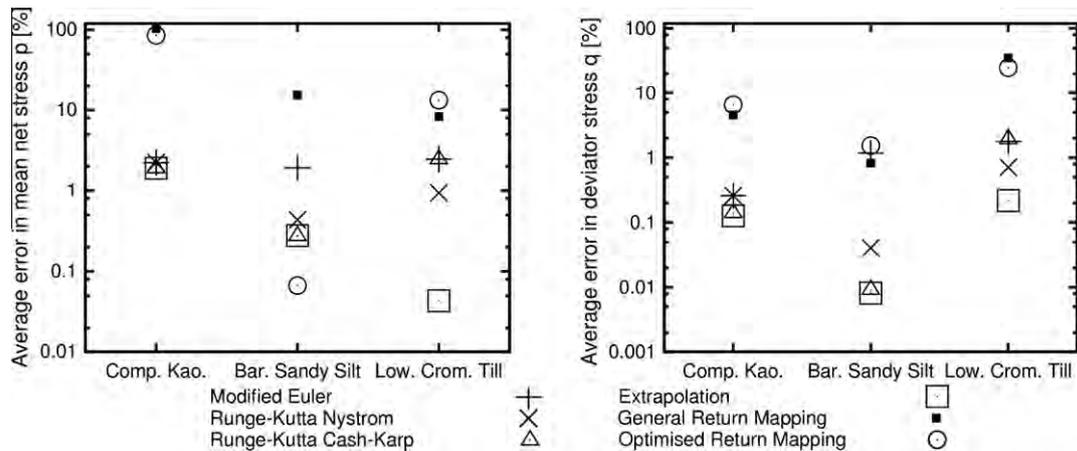


Fig. 2. Average error in mean net stress (left) and deviator stress (right) for three sets of constitutive model parameters.

small strain increments the accuracy is very high. The optimized return mapping scheme also offers very good stability, being able to accurately integrate large strain increments. Note that the apparently large relative errors in some cases are not that significant when the absolute values of stress are considered; This is especially true for the cases when the increase in the stress is relatively small when compared to the strain increment. The general

return mapping algorithm does have convergence problems in some cases (see Fig. 3). Hence, its use is not advisable for strain increments larger than 1%.

It appears that all the Runge–Kutta schemes show similar values of the average error in each test case (Fig. 2) and the resulting stresses are of roughly similar accuracy. Thus, as the efficiency of Runge–Kutta schemes depends on the required accuracy, the order

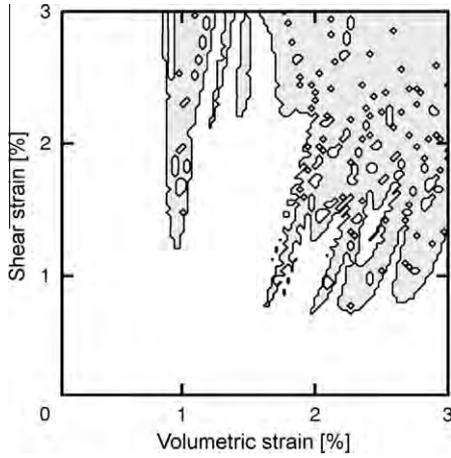


Fig. 3. Strain increments for which the general return mapping algorithm diverges for Lower Cromer Till.

of the method should be chosen accordingly. It appears that the third order Nystrom or second order Modified Euler method would offer the best performance for relatively small strain increments and tolerances exceeding 0.01% (see [9,10]).

From the limited tests performed it also seems that the optimized return mapping algorithm of Hofmann [2] is beneficial for the Barcelona Basic Model. The optimized algorithm converges for all the strain increments and is of an order of magnitude faster than the general return mapping algorithm with similar accuracy.

4. Impact of the stress integration algorithm on global convergence of the Newton–Raphson iteration

The various algorithms were not tested in a finite element environment, where global convergence for some typical boundary value problems could be measured. Instead, the convergence was assessed based on the Newton–Raphson iterations for a single stress point. The results obtained will not always translate directly to a finite element boundary value problem, but some results suggest that the single stress point test may give an acceptable approximation of the algorithm behaviour in a boundary value problem.

For implicit stress integration, a consistent elasto-plastic tangent matrix can be obtained for the Newton–Raphson iteration which leads to a quadratic rate of convergence provided that the starting point is sufficiently close to the solution (e.g. see Borja and Lee [11]). However, for the explicit schemes, the continuum elasto-plastic tangent matrix is usually used resulting in a linear rate of convergence (some research on improving the convergence of explicit schemes is available in Fellin and Ostermann [19] and Sloan et al. [20]). For the convergence study the parameters of the Barcelona Basic Model, given in Table 1, and the initial stress states, given in Table 3, are used.

The test aims at recovering the strains for given stresses in order to investigate the impact of the material tangent matrix on the convergence rate. For each initial stress state (given in Table 2), tests have been run for 51 increments of mean net stress and deviator stress (giving a total of 2601 stress increment combinations).

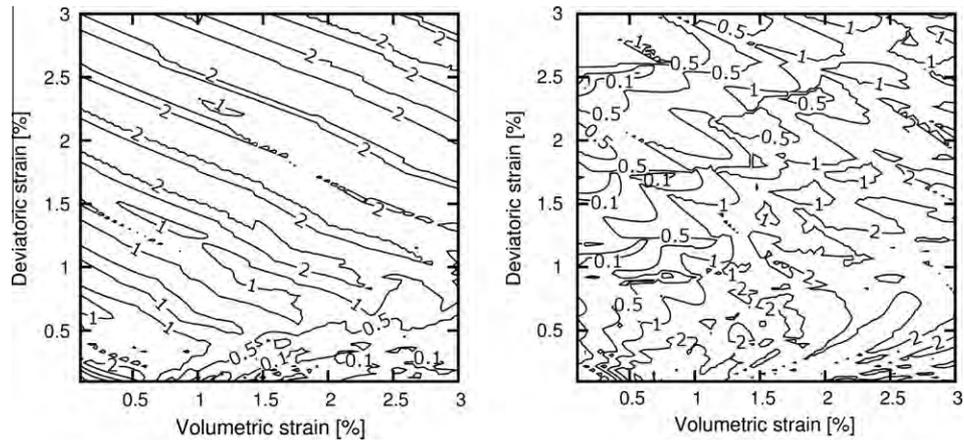


Fig. 4. Percentage error in mean net stress (left) and deviator stress (right) of Runge–Kutta Modified Euler integration scheme for Lower Cromer Till.

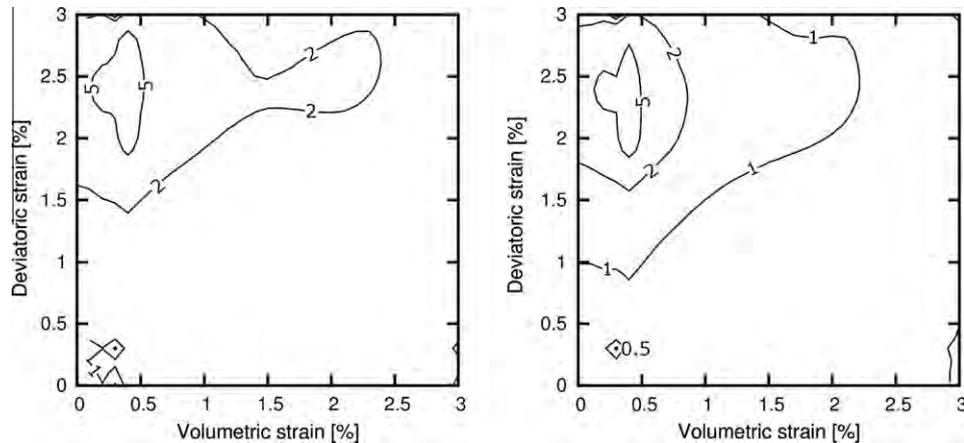


Fig. 5. Percentage error in mean net stress (left) and deviator stress (right) of Runge–Kutta Modified Euler integration scheme for Barcelona Sandy Silt.

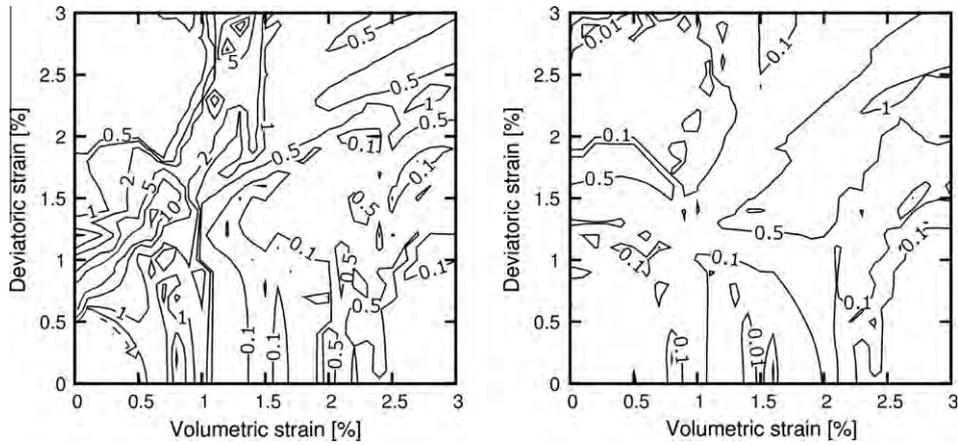


Fig. 6. Percentage error in mean net stress (left) and deviator stress (right) of Runge–Kutta Nystrom integration scheme for Compacted Kaolin.

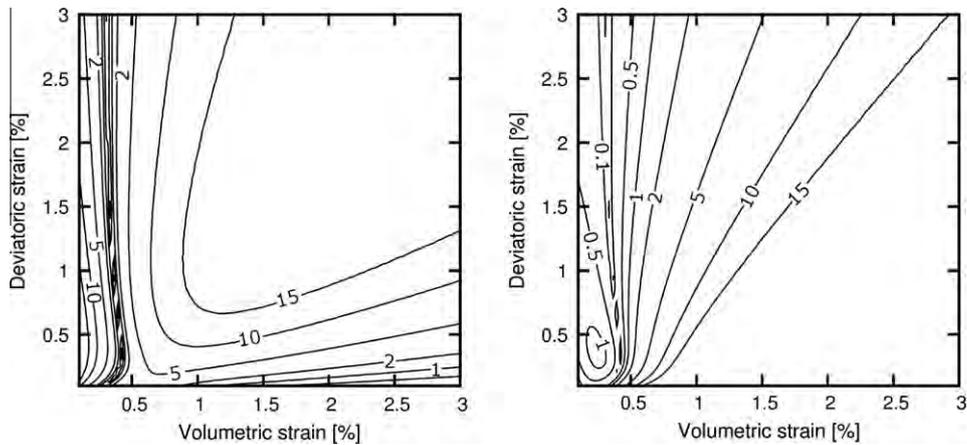


Fig. 7. Percentage error in mean net stress (left) and deviator stress (right) of optimized return mapping scheme for Lower Cromer Till.

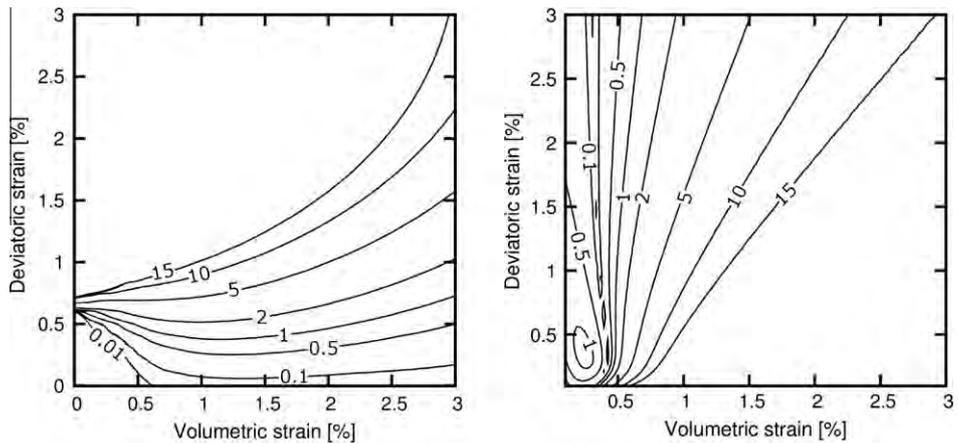


Fig. 8. Percentage error in mean net stress (left) and deviator stress (right) integrated with optimized return mapping algorithm for Compacted Kaolin.

Recording the number of iteration steps for each of the stress increments allows for the creation of an iso-map depicting the convergence rate from each starting point. For each soil type the suction increment was constant (as indicated in Table 3) and independent of the stress increment (so 2601 pairs $\Delta p_i - \Delta q_j - \Delta s$, $i, j = 0, \dots, 50$ were considered).

The allowable error (i.e. accuracy of the recovered stresses) is 1%. This means that if the applied stress increment is $\Delta p = 50$ kPa and $\Delta q = 30$ kPa, corresponding to $\Delta \sigma_{11} = 70$ kPa and $\Delta \sigma_{22} = \Delta \sigma_{33} = 40$ kPa, then the required accuracy is:

$$\begin{aligned} \text{Required Accuracy} &= Tol \sqrt{\Delta \sigma_{11}^2 + \Delta \sigma_{22}^2 + \Delta \sigma_{33}^2} = 0.01 \sqrt{8100} \\ &= 0.9 \text{ [kPa]} \end{aligned}$$

Hence, the iteration process is stopped when the calculated stress increments $\Delta \sigma_{11}^i, \Delta \sigma_{22}^i, \Delta \sigma_{33}^i$ in the i -th Newton–Raphson iteration step are such that:

$$\sqrt{(\Delta \sigma_{11} - \Delta \sigma_{11}^i)^2 + (\Delta \sigma_{22} - \Delta \sigma_{22}^i)^2 + (\Delta \sigma_{33} - \Delta \sigma_{33}^i)^2} < 0.9.$$

Table 3
Stress and suction increments for the Newton–Raphson convergence tests.

Initial value of a parameter		Compacted Kaolin	Barcelona Sandy Silt	Lower Cromer Till
Initial stress state 1, mean net stress	p (kPa)	88	820	20
Initial stress state 1, shear stress	q (kPa)	0	0	0
Initial stress state 2, mean net stress	p (kPa)	70	650	16
Initial stress state 2, shear stress	q (kPa)	48	440	9
Initial stress state 3, mean net stress	p (kPa)	50	500	12
Initial stress state 3, shear stress	q (kPa)	66	560	12
Mean net stress increment	Δp (kPa)	15	200	5
Shear stress increment	Δq (kPa)	15	200	5
Suction increment	(kPa)	–12	–200	–1

The described procedure is employed for testing the convergence at the level of a single integration point. It is similar for both explicit and implicit schemes; the only difference is that in the case of the explicit stress integration schemes, for each stress pair (Δp , Δq), the continuum elasto-plastic tangent matrix for the current stress state has been used while for the implicit scheme a consistent elasto-plastic tangent matrix has been adopted.

4.1. Results for explicit stress integration algorithms

It has been found that in the case of explicit schemes, the choice and accuracy of the stress integration procedure does not affect the convergence rate significantly as long as: (a) no drift correction scheme is used and (b) the tolerance of the integration algorithm is tighter than the tolerance for the Newton–Raphson iterations (thus the computed results are for an integration tolerance scheme of 0.1%) [5–10]. In some rare cases, the drift correction scheme (as in [9,21]) has been found to be detrimental to the global convergence rate. Therefore, to reduce the impact of drift correction (as the subtleties in coding of the drift correction algorithm may be the reason for the irregular behaviour observed) this study has been conducted without any drift correction.

The results depend significantly on the model parameters. The set of material parameters for the Lower Cromer Till caused the

most difficulties for the algorithm and resulted in a much larger number of iterations before reaching convergence.

Of the three initial stress states tested, the algorithms were found to be the most stable and converge in the lowest number of Newton–Raphson iterations for the initial isotropic stress state (i.e. $p \neq 0$, $q = 0$). The Barcelona Sandy Silt and Compacted Kaolin have similar convergence rates (requiring four and three iteration steps, respectively), whereas the convergence rate for the Lower Cromer Till is strongly increment dependent (Fig. 9, note that in Figs. 9–13, the initial stress state corresponds to the lower left corner of the figure).

The initial stress state close to the tip of the critical state ellipse (initial stress state 3) required the most iteration steps before reaching convergence. The results are shown in Figs. 10 and 11. Note that in the case of Lower Cromer Till, there are some areas where the Newton–Raphson scheme diverges (i.e. requires more than 250 iterations).

The rate of convergence in the case of explicit integration, although not quadratic, is sufficient for many engineering applications, especially for the model parameters corresponding to Barcelona Sandy Silt and Compacted Kaolin. The Lower Cromer Till presented some challenges due to the relatively large stress increments and suction decrease imposed. It has been chosen as an example to show that there are some cases for which the higher convergence rate provided by an implicit scheme is beneficial (though in FE analysis it is likely that a smart automatic load stepping algorithm would prevent such large strain increments).

4.2. Results for implicit stress integration algorithms

Generally, the number of Newton–Raphson iterations needed for the implicit stress integration using the consistent tangent matrix is smaller in all cases. The number of iteration steps for the Lower Cromer Till is relatively high, yet much lower than for the explicit stress integration (Fig. 12). However, the difference is usually not significant, especially for the initial isotropic stress state where both methods tend to give similar patterns of convergence for the Compacted Kaolin and Barcelona Sandy Silt (where the integration methods require 3 iteration steps to converge except for the explicit integration for Compacted Kaolin for which 3 or 4 iteration steps are needed, see Fig. 9 left). To sum up, in some cases, especially when the initial stress state is close to the tip of the ellipse, the implicit stress integration has substantial benefits, especially for larger strain increments and may result in a significant reduction in computations. This may be more important for finite

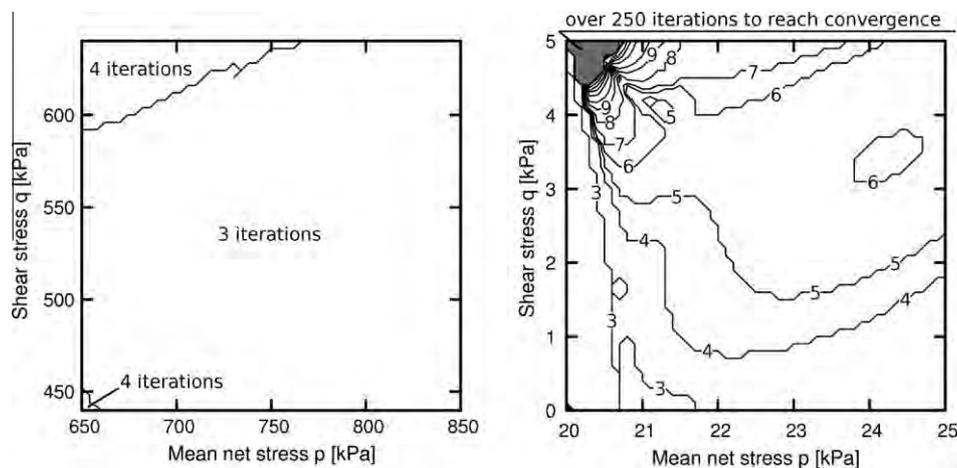


Fig. 9. Number of iterations until reaching convergence starting from the initial stress state 2 for Barcelona Silty Sand (left) and starting from the initial stress state 1 for Lower Cromer Till (right). Integration performed with Modified Euler scheme.

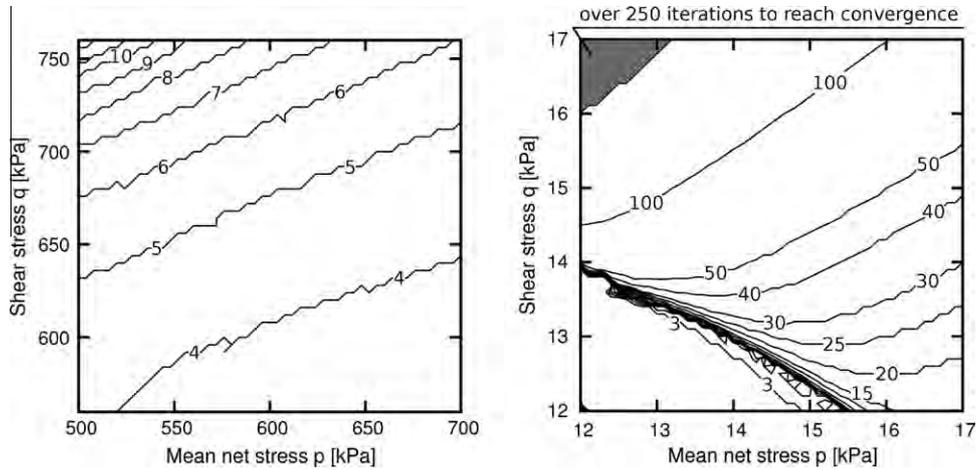


Fig. 10. Number of iterations until reaching convergence starting from the initial stress state 3 for Barcelona Sandy Silt (left) and Lower Cromer Till (right). Integration performed with Modified Euler scheme.

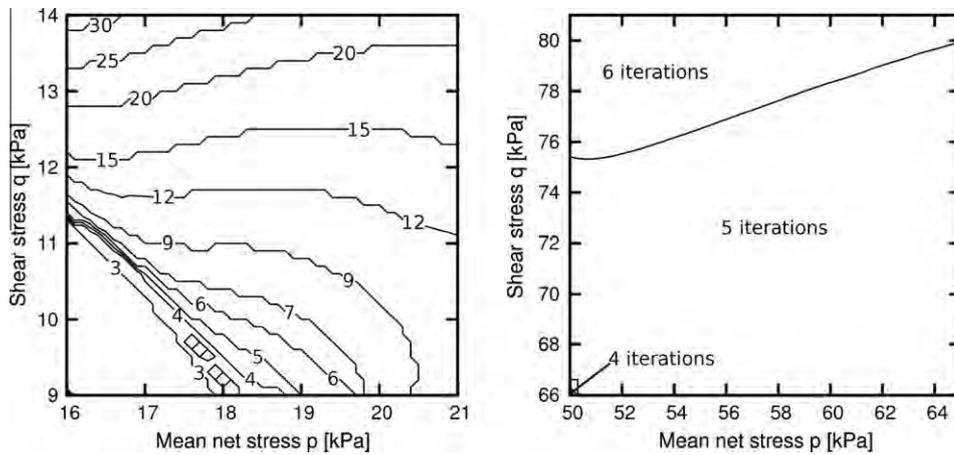


Fig. 11. Number of iterations until reaching convergence starting from the initial stress state 2 for Lower Cromer Till (left) and starting from the initial stress state 3 (right) for Compacted Kaolin. Integration performed with Modified Euler scheme.

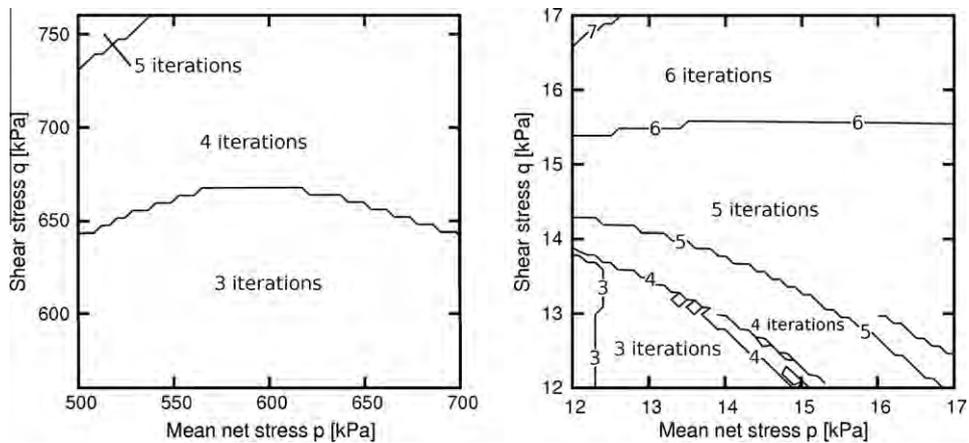


Fig. 12. Number of iteration steps using the optimized return mapping scheme until reaching convergence starting from the initial stress point 3 for Barcelona Sandy Silt (left) and Lower Cromer Till (right).

element analysis where a broad variety of stress states is present. This benefit comes at a cost of accuracy, as shown in subsequent paragraphs.

The quadratic rate of convergence of the optimized return mapping algorithm is confirmed in Table 4 where a typical iteration process is shown. The increment in Table 4 has been calculated

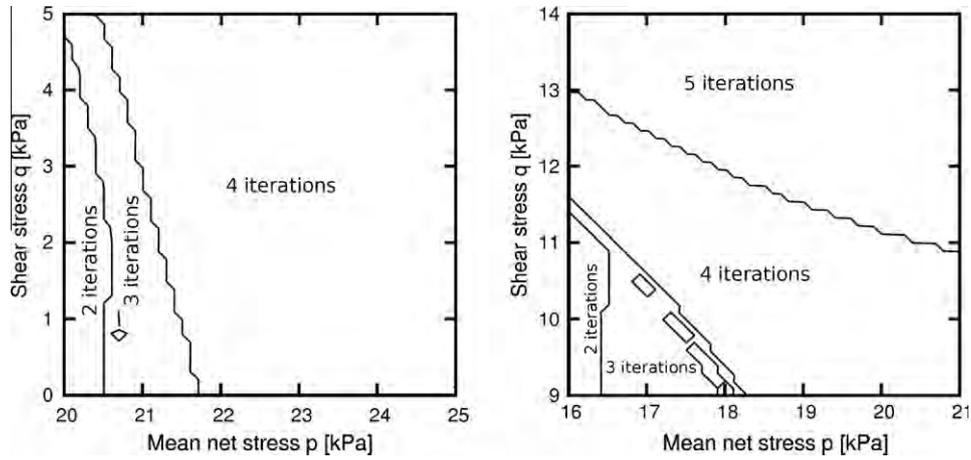


Fig. 13. Number of iteration steps for the optimized return mapping scheme starting from the initial stress state 1 (left) and 2 (right) for Lower Cromer Till.

Table 4
Iteration process for stress increment for Barcelona Sandy Silt at point 3. Implicit stress integration (Improved return mapping scheme).

Iteration	$\Delta\epsilon_{11}$ (%)	$\Delta\epsilon_{22}, \Delta\epsilon_{33}$ (%)	$\Delta\sigma_{11}$ (kPa)	$\Delta\sigma_{22}, \Delta\sigma_{33}$ (kPa)	Error (kPa)	σ_{11} (kPa)	$\sigma_{22} = \sigma_{33}$ (kPa)
1	3.8200	-1.4027	221.75	39.659	153.73	1095.08	352.99
2	4.6212	-1.8604	151.92	-38.115	24.315	1025.25	275.22
3	5.1660	-2.1296	146.30	-52.521	3.183	1019.63	260.81
4	5.2537	-2.1731	145.36	-54.620	0.070	1018.69	258.71
5	5.2556	-2.1741	145.34	-54.666	0.003	1018.67	258.67

Table 5
Iteration process for stress increment for Barcelona Sandy Silt at point 3. Explicit stress integration scheme (Modified Euler).

Iteration	$\Delta\epsilon_{11}$ (%)	$\Delta\epsilon_{22}, \Delta\epsilon_{33}$ (%)	$\Delta\sigma_{11}$ (kPa)	$\Delta\sigma_{22}, \Delta\sigma_{33}$ (kPa)	Error (kPa)	σ_{11} (kPa)	$\sigma_{22} = \sigma_{33}$ (kPa)
1	2.2222	-1.1058	-157.32	-237.112	397.70	716.02	76.22
2	4.8569	-2.2179	19.974	-155.008	189.35	893.31	158.33
3	5.4324	-2.4300	98.377	-116.007	98.641	971.71	197.33
4	5.1914	-2.2878	124.70	-99.737	66.994	998.04	213.60
5	4.7478	-2.0626	132.85	-90.199	51.779	1006.18	223.13
6	4.3327	-1.8567	136.62	-80.927	38.147	1009.95	232.41
7	4.0231	-1.7046	139.45	-71.767	24.889	1012.79	241.57
8	3.8296	-1.6099	141.68	-64.578	14.486	1015.01	248.76
9	3.7239	-1.5581	143.24	-59.975	7.793	1016.58	253.36
10	3.6702	-1.5317	144.07	-57.499	4.201	1017.40	255.83
11	3.6445	-1.5189	144.79	-56.038	2.013	1018.13	257.30
12	3.6308	-1.5122	145.07	-55.333	0.977	1018.41	258.00

for Barcelona Sandy Silt taking point 3 as the initial state; the stress increment was $\Delta p = 12$ and $\Delta q = 200$ kPa (which corresponds to increment of $\Delta\sigma_{11} = 145.33$ kPa and $\Delta\sigma_{22} = \Delta\sigma_{33} = -54.67$ kPa and final total stresses $\sigma_{11} = 1018.67$ and $\sigma_{22} = \sigma_{33} = 258.67$) and the required accuracy was 1% (thus the corresponding acceptable error was 1.6462 kPa). By contrast, in this case the convergence rate for the explicit scheme is relatively low (see Table 5, same increment and initial conditions as in Table 4). It needs to be mentioned, however, that the stress increment is large and in the majority of the cases tested three or four iteration steps were enough to achieve convergence for the explicit stress integration with a tolerance of one percent. If stricter tolerances are required the benefit of using the consistent tangent matrix would be more pronounced.

It is interesting to note that the implicit and explicit schemes give quite different results. The difference reaches 30% - and this is due to the implicit integration scheme. Thus, although the use of a consistent tangent matrix speeds up the convergence process, it may lead to larger errors in the stress update and to larger errors in a FE analysis. Again, this is unlikely to happen when the load

steps are small, which will be the case if a smart load stepping scheme is used, see e.g. Abbo and Sloan [22] and Sheng et al. [23], or if subincrementation with error control is also employed for the implicit schemes.

5. Conclusions

The paper compares the performance of explicit and implicit stress integration algorithms for the Barcelona Basic Model. It appears that the main advantages of explicit schemes include less complex implementation (as only first order derivatives are needed) and error control which allows for user-specified accuracy. The explicit algorithms are also relatively robust, though some convergence problems may occur at the global level due to using the continuum elasto-plastic tangent matrix, especially when large load increments are considered. The main benefit of the implicit schemes is that they offer faster convergence at the global level, especially in the case of large load increments, due to the use of the consistent elasto-plastic tangent matrix.

The standard implementation of implicit schemes may, however, lead in some cases to inaccurate computations unless they are combined with substepping and error control. It may also be that the implicit algorithms are more expensive in terms of computation time. Nevertheless, under some circumstances, in a finite element analysis employing implicit stress integration the solution may be computed significantly faster than using explicit stress integration (compare [3]). This, however, may come at a cost of some stress states being computed with surprisingly large errors, especially when relatively large strain increments occur in a single load step. Such inaccuracy at a single integration point may lead to erroneous results in finite element analysis of a boundary value problem.

When small strain increments are considered (which is often the case when a smart automatic load stepping algorithm is used), it appears that there is no significant difference between the implicit and explicit approach for stress integration. Both approaches offer good accuracy and decent convergence rates. The better global convergence rate may favour slightly the implicit schemes, whereas the ease of coding and adaptive error control may favour the explicit approach to stress integration.

Appendix A. Brief description of Barcelona Basic Model

The Barcelona Basic Model (BBM) is a constitutive model, based on plasticity theory, for describing unsaturated soil behaviour. The BBM is formulated in terms of the net stress tensor σ_{ij} and suction s . The former is the total stress in excess of the pore air pressure, the latter is the difference between the pore air pressure and the pore water pressure.

For stress states located within the elastic domain, enclosed by the yield surface, the rates of the elastic volumetric strain $\dot{\epsilon}_v^e$ and deviatoric strain $\dot{\epsilon}_{ij}^e$ are given as

$$\dot{\epsilon}_v^e = \frac{\kappa}{1+e} \frac{\dot{p}}{p} + \frac{\kappa_s}{1+e} \frac{\dot{s}}{s+p_{atm}}, \quad \dot{\epsilon}_{ij}^e = \frac{\dot{s}_{ij}}{2G} \quad (\text{A.1})$$

with the material parameters κ and κ_s , representing the elastic stiffness for rates of the mean net stress $p = (\sigma_{ij}\delta_{ij})/3$ and suction s , respectively; e , p_{atm} , $\dot{s}_{ij} = \dot{\sigma}_{ij} - \dot{p}\delta_{ij}$ and G denote the void ratio, the atmospheric air pressure, the deviatoric stress rate and the shear modulus, respectively. It follows from (A.1)₁ that the elastic volumetric strain rate $\dot{\epsilon}_v^e$ depends on both the mean net stress p and suction s .

The stress point (p, e) lies on the unloading–reloading line (URL) with the slope κ within the elastic domain and on the normal (isotropic) compression line (NCL) with the suction-dependent slope

$$\lambda(s) = \lambda(0)[(1-r)e^{-\beta s} + r] \quad (\text{A.2})$$

for isotropic plastic loading. $\lambda(s)$ describes the soil stiffness during plastic loading in a hydrostatic test for given suction in terms of the respective stiffness $\lambda(0)$ at saturated conditions and the material parameters r and β .

The intersection point of the URL and the NCL is the preconsolidation stress p_0 (Fig. A1). The NCL is defined by the slope $\lambda(s)$ and the void ratio $N(s) - 1$ at $p = 1$ with $N(s)$ denoting the respective specific volume. From geometric considerations in the diagram for the volumetric behaviour, showing the void ratio e in terms of $\ln p$, follows

$$e = N(s) - 1 - \lambda(s) \ln p_0 + \kappa \ln \left(\frac{p_0}{p} \right) \quad (\text{A.3})$$

where $N(s)$ is equal to

$$N(s) = N(0) - \kappa_s \ln \left(\frac{s+p_{atm}}{p_{atm}} \right) + \lambda(s) \ln \left(\frac{p_{ref}}{1} \right) \quad (\text{A.4})$$

and p_{ref} is the reference pressure. The $N(s)$ is different to that defined in Alonso et al. [16], but the resulting model is identical. From (A.3) one obtains

$$p_0 = \exp \left(\frac{-N(s) + (1+e)}{\kappa - \lambda(s)} \right) p^{\frac{\kappa}{\kappa - \lambda(s)}} \quad (\text{A.5})$$

The yield surface is defined as (Fig. A2)

$$f = J_2 - \frac{M^2}{3}(p + p_s)(p_0 - p) \quad (\text{A.6})$$

with the second invariant of the deviatoric stress tensor $J_2 = s_{ij}s_{ij}/2$ and

$$p_s = ks \quad p_0 = p_{ref} \left(\frac{p_0^*}{p_{ref}} \right)^{\frac{\lambda(0) - \kappa}{\lambda(s) - \kappa}} \quad (\text{A.7})$$

In (A.6) M defines the slope of the critical state line. p_s and p_0 both depend on suction according to (A.7). For negative values of p the intersection with the plane $J_2 = 0$ is given by p_s according to (A.7)₁ with the material parameter k describing the increase in cohesion due to suction (Fig. A2). The preconsolidation pressure p_0 and the one for saturated conditions p_0^* are located on the so called loading collapse yield curve (LC curve) according to (A.7)₂. This curve is the intersection of the yield surface with the plane $J_2 = 0$ for positive values of p (Fig. A2). p_{ref} serves as a reference pressure such that for $p_0^* = p_{ref}$ (A.6)₂ degenerates to $p_0 = p_{ref} = \text{const.}$

The plastic strain rate is determined from the non-associated flow rule

$$\dot{\epsilon}_{ij}^p = \dot{\gamma} \frac{\partial g}{\partial \sigma_{ij}} \quad (\text{A.8})$$

with the flow potential

$$g = \alpha J_2 - \frac{M^2}{3}(p + p_s)(p_0 - p) \quad (\text{A.9})$$

where α is a constant. Note that because of using different stress parameters compared to [12], α in (A.9) and in [12] are not identical.

The hardening law relates the rate of the preconsolidation pressure at saturated conditions p_0^* , which serves as the hardening parameter, to the volumetric plastic strain rate $\dot{\epsilon}_v^p$ by

$$p_0^* = p_0^* \frac{1+e}{\lambda(0) - \kappa} \dot{\epsilon}_v^p \quad (\text{A.10})$$

(A.10) describes the evolution of the yield surface. The latter is shown for two different values of p_0^* in Fig. A2.

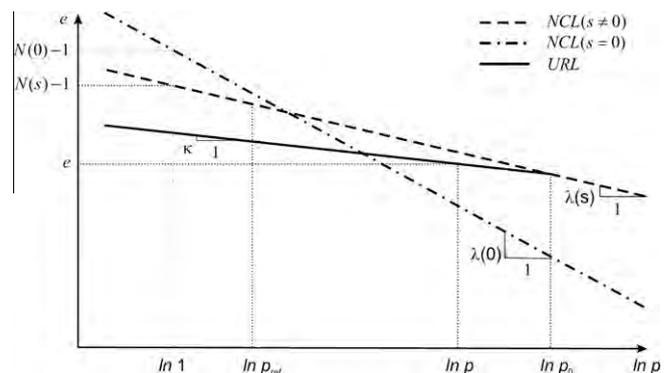


Fig. A1. Volumetric behaviour of the BBM.

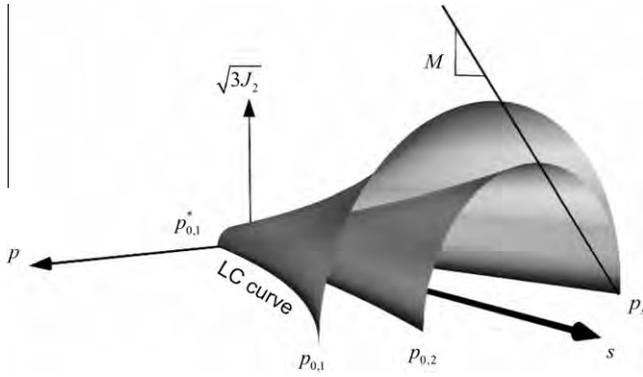


Fig. A2. Yield surface of the BBM for different values of the preconsolidation pressure.

Appendix B. Brief description of explicit stress integration algorithm

At the beginning it is assumed that the initial stress state σ , strain state ε and hardening parameter p_0^* is known and the vector of the strain increment $\Delta\varepsilon$ consisting of the strain components $\Delta\varepsilon_{ij}$ and the suction increment Δs is given. To obtain the stress increment corresponding to the strain $\Delta\varepsilon$ with the explicit stress integration schemes, the purely elastic part of the strain increment $\Delta\varepsilon^e$ and corresponding stress $\Delta\sigma^e$ must be computed. With the elastic rule as described in (A.1) it is possible to analytically compute the stresses $\Delta\sigma^e$ for a known strain increment $\Delta\varepsilon^e$. However, the strain increment $\Delta\varepsilon^e$ is usually obtained numerically from the Pegasus algorithm [6,9]. The remaining part of the strain increment $\Delta\varepsilon^p = \Delta\varepsilon - \Delta\varepsilon^e$ is assumed to be elasto-plastic and is integrated with a Runge–Kutta method with adaptive substepping (subincrementation) and error control.

Integration of the elasto-plastic part of the increment requires the knowledge of the elasto-plastic matrix \mathbf{D}^{ep} . However, due to the fact that strain vector $\Delta\varepsilon$ has, in general, seven elements, the \mathbf{D}^{ep} has to be enhanced and has dimensions of 6×7 , with the last column corresponding to the suction increment. The derivation of the matrix is given in [9].

The explicit stress integration is equivalent to integration of equation

$$\dot{\sigma} = \mathbf{D}^{ep} \dot{\varepsilon}^p \quad (\text{B.1})$$

so the total stress increment is

$$\Delta\sigma = \Delta\sigma^e + \int_{\varepsilon+\Delta\varepsilon^e}^{\varepsilon+\Delta\varepsilon} \mathbf{D}^{ep} d\varepsilon^p \quad (\text{B.2})$$

As the matrix \mathbf{D}^{ep} depends on the current stresses, strains and the hardening parameter values, the stresses are integrated numerically using the Runge–Kutta algorithm with automatic substepping and error control (for the description of an alternative method, the Richardson extrapolation, see 9). In short, the idea is to divide the strain increment $\Delta\varepsilon^p$ into a number of subincrements $\delta\varepsilon_i$ such that error estimate obtained during integration of each subincrement is below a specified tolerance. After integration of a given subincrement $\delta\varepsilon_i$, the error is estimated. If the error is too large, the computed stress corresponding to $\delta\varepsilon_i$ is rejected, the subincrement size is decreased and the integration is repeated for the i -th (now smaller) subincrement. Otherwise, the integration for the subincrement $\delta\varepsilon_i$ is deemed successful and the size of the next subincrement $\delta\varepsilon_{i+1}$ is determined based on the obtained error estimate.

Assuming the initial conditions before computation of i -th subincrement as σ_{i-1} , ε_{i-1} and $p_{0(i-1)}^*$, for each subincrement $\delta\varepsilon_i$ the stress is computed as

$$\delta\sigma_i = \sum_{j=1}^{\text{NoS}} b^{(j)} \delta\sigma^{(j)} \quad (\text{B.3})$$

where NoS is the number of stages and $b^{(j)}$ are specified for a given Runge–Kutta method. The necessary stresses $\delta\sigma^{(j)}$ are computed as

$$\delta\sigma^{(j)} = \mathbf{D}_i^{ep} \left(\varepsilon_{i-1} + c^{(j)} \delta\varepsilon_i; \sigma_{i-1} + \sum_{k=1}^{j-1} a^{(jk)} \delta\sigma^{(k)}; p_{0(i-1)}^* + \sum_{k=1}^{j-1} a^{(jk)} \delta p_0^{*(k)} \right) \delta\varepsilon_i \quad (\text{B.4})$$

where $a^{(jk)}$ and $c^{(j)}$ are again specified for a given Runge–Kutta scheme. The increase in hardening parameter is computed in a similar manner. The error estimate is given as

$$\mathbf{E}(\delta\sigma) = \delta\sigma - \delta\sigma^{(low)} = \sum_{j=1}^{\text{NoS}} (b^{(j)} - d^{(j)}) \delta\sigma^{(j)} \quad (\text{B.5})$$

where $d^{(j)}$ are constants specified for a given Runge Kutta scheme. The subincrement $\delta\varepsilon_i$ is assumed to be integrated accurately enough when the error is smaller than the user specified tolerance Tol as shown below

$$\mathbf{R}(\delta\sigma_i) = \frac{|\mathbf{E}(\delta\sigma_i)|}{|\sigma_{i-1} + \delta\sigma_i|} < Tol \quad (\text{B.6})$$

The relative size of the next increment to integrate is computed as

$$\chi = \zeta \sqrt[m]{\frac{Tol}{R}} \quad (\text{B.7})$$

where $\zeta < 1$ is a constant and m is the order of the Runge–Kutta method. If the stresses are not accurate enough, the size of $\delta\varepsilon_i$ is reduced to $\delta\varepsilon_i = \chi \delta\varepsilon_i$ and the increment is reintegrated. Otherwise, the scheme integrates the subsequent strain increment $\delta\varepsilon_{i+1} = \chi \delta\varepsilon_i$. The scheme continues until the whole strain increment $\Delta\varepsilon_p$ is integrated with the required accuracy. After the computation of each $\delta\sigma_i$ the stresses may not satisfy the yield locus equation due to inaccuracies in the numerical solution. In such cases an additional drift correction is used, which restores the stresses to the yield surface. For more details see [4–6,9,10,21].

Appendix C. Brief description of implicit stress integration algorithm

From the known equilibrium state at time instant t_n and from the known increments of the total strain of the soil skeleton and suction, $\Delta\varepsilon_{ij}$ and Δs , for the current time step $\Delta t = t_{n+1} - t_n$, the unknown plastic strain, net stress and hardening variable are to be computed.

Since the BBM is characterized by (i) a single yield surface (A.6), (ii) a single plastic potential (A.9) only depending on the first and second invariant of the net stress tensor, $I_1 = 3p$ and J_2 , (iii) a single hardening parameter that can be expressed in terms of only the mean net stress using (A.5) and (A.7)₂ for the corrector step and (iv) a relation for the deviatoric stress rate (A.1)₂ depending on the deviatoric elastic strain rate through the shear modulus G , the general return mapping algorithm, proposed in [1], can be formulated in a very efficient manner in terms of a single scalar nonlinear equation for the mean net stress p at t_{n+1} .

To this end, the flow rule (A.8) is split into volumetric and deviatoric parts

$$\dot{\varepsilon}_v^p = \dot{\gamma} 3 \frac{\partial g}{\partial I_1}, \quad \dot{\varepsilon}_{ij}^p = \dot{\gamma} \frac{\partial g}{\partial J_2} s_{ij}. \quad (\text{C.1})$$

Backward Euler integration of (C.1) yields

$$\Delta \varepsilon_v^p = \gamma 3 \frac{\partial g}{\partial I_1} = \Delta \varepsilon_v - \varepsilon_v^e, \quad \Delta e_{ij}^p = \gamma \frac{\partial g}{\partial J_2} s_{ij} = \Delta e_{ij} - \Delta e_{ij}^e, \quad (C.2)$$

where $\gamma = \dot{\gamma} \Delta t$. Note that quantities with the subscript n refer to the converged values at t_n , whereas all other quantities refer to the current values at t_{n+1} . From (A.1)₂ follows in case of a constant value of G

$$s_{ij} = 2G e_{ij}^e = 2G[(e_{ij} - e_{ij,n}^p) - (e_{ij}^p - e_{ij,n}^p)] = s_{ij}^{trial} - 2G \Delta e_{ij}^p. \quad (C.3)$$

Use of (C.2) in (C.3) gives

$$\left(1 + \gamma \frac{\partial g}{\partial J_2} 2G\right) s_{ij} = s_{ij}^{trial}. \quad (C.4)$$

The term enclosed by the brackets is a scalar quantity, and, hence, s_{ij} and s_{ij}^{trial} differ only by a scalar factor. Thus, from (C.4) it follows

$$\left(1 + \gamma \frac{\partial g}{\partial J_2} 2G\right)^2 J_2 = J_2^{trial} \quad (C.5)$$

Making use of $\gamma = (\Delta \varepsilon_v - \Delta \varepsilon_v^e) / (3 \partial g / \partial I_1)$, resulting from (C.2)₁, yields

$$\left(3 \frac{\partial g}{\partial I_1} - \frac{\partial g}{\partial J_2} 2G(\Delta \varepsilon_v^e - \Delta \varepsilon_v)\right)^2 J_2 - \left(3 \frac{\partial g}{\partial I_1}\right)^2 J_2^{trial} = 0 \quad (C.6)$$

In (C.6) the incremental volumetric strain $\Delta \varepsilon_v$ is known from the current estimate of Δe_{ij} at t_{n+1} and the converged values of the strain at t_n . J_2 and $\Delta \varepsilon_v^e$ in (C.6) can be replaced by

$$J_2 = \frac{M^2}{3} (p + p_s)(p_0 - p) \quad (C.7)$$

and

$$\Delta \varepsilon_v^e = \frac{\kappa}{1+e} \ln\left(\frac{p}{p_n}\right) + \frac{\kappa_s}{1+e} \ln\left(\frac{s + p_{atm}}{s_n + p_{atm}}\right) \quad (C.8)$$

following from (A.6) and (A.1)₁. Integration of the rate of the void ratio

$$\dot{e} = -(1+e)\dot{\varepsilon}_v \quad (C.9)$$

yields the value of the void ratio at t_{n+1}

$$e = (1 + e_n) \exp(-\Delta \varepsilon_v) - 1. \quad (C.10)$$

(C.6) together with (C.7), (C.8), (C.10), and (A.5) represents a nonlinear scalar equation for the unknown I_1 (or, equivalently, for $p = I_1/3$), which can be solved, e.g., by the Newton method. Once p has been determined from this equation, it is inserted into (A.5), yielding p_0 , and the latter into the recast form of (A.6)₂ yielding the hardening parameter

$$p_0^* = p_{ref} \left(\frac{p_0}{p_{ref}}\right)^{\frac{\lambda(s)-\kappa}{\lambda(0)-\kappa}} \quad (C.11)$$

The Newton iteration for solving (C.6) together with (C.7), (C.8), (C.10), and (A.5) can be started by the trial values p^{trial} and s_{ij}^{trial} , which are computed by assuming elastic material behaviour for the current time step. An improved trial value for the mean net stress p^{start} will be obtained by projecting p^{trial} on the isotropic compression line by

$$p^{start} = \exp\left(\frac{N(s) - (1+e)}{\lambda(s)}\right). \quad (C.12)$$

p^{start} is only used as an improved trial stress, when $p^{start} < p^{trial}$ holds.

For known net stress at t_{n+1} the rate of the latter in terms of the strain rate and suction rate is obtained as

$$\dot{\sigma}_{ij} = C_{ijkl}^{ep} \dot{\varepsilon}_{kl} + C_{ij}^s \dot{s} \quad (C.13)$$

with

$$C_{ijkl}^{ep} = \frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} \quad \text{and} \quad C_{ij}^s = \frac{\partial \sigma_{ij}}{\partial s} \quad (C.14)$$

denoting the consistent material tangent moduli with respect to the total strain ε_{kl} and the consistent tangent moduli with respect to suction s , respectively. They are determined by making use of the split of the net stresses into the mean net stress p and the deviatoric stress s_{ij} , introduced in Appendix A, as

$$\frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} = \frac{\partial p}{\partial \varepsilon_{kl}} \delta_{ij} + \frac{\partial s_{ij}}{\partial \varepsilon_{kl}} \quad \text{and} \quad \frac{\partial \sigma_{ij}}{\partial s} = \frac{\partial p}{\partial s} \delta_{ij} + \frac{\partial s_{ij}}{\partial s} \quad (C.15)$$

$\partial p / \partial \varepsilon_{kl}$ and $\partial p / \partial s$ are computed by implicit analytic differentiation of (C.6), $\partial s_{ij} / \partial \varepsilon_{kl}$ and $\partial s_{ij} / \partial s$ are computed by analytic differentiation of the relation (C.4) between the deviatoric stress and the deviatoric trial stress, respectively [2].

References

- [1] Simo J, Hughes T. Computational inelasticity. Interdisciplinary applied mathematics series. Springer; 1998.
- [2] Hofmann M. Integration algorithms and parameter identification of elastoplastic constitutive models for partially saturated soils, PhD Thesis, University of Innsbruck; 2009.
- [3] Perez-Foguet A, Rodriguez-Ferran A, Huerta A. Consistent tangent matrices for substepping schemes. *Comput Methods Appl Mech Eng* 2001;190(35–36):4627–47.
- [4] Potts D, Ganendra D. An evaluation of substepping and implicit stress point algorithms. *Comput Methods Appl Mech Eng* 1994;119:341–54.
- [5] Sloan SW. Substepping schemes for the numerical integration of elastoplastic stress-strain relations. *Int J Numer Methods Eng* 1987;24:893–911.
- [6] Sloan SW, Abbo AJ, Sheng DC. Refined explicit integration of elastoplastic models with automatic error control. *Engineering Computations* 2001;18(1/2):121–54. Erratum: *Engineering Computations* 2002;19(5/6):594–594.
- [7] Sheng DC, Sloan SW, Gens A, Smith DW. Finite element formulation and algorithms for unsaturated soils. Part 1: Theory. *International Journal for Numerical and Analytical Methods in Geomechanics* 2003;27(9):745–65.
- [8] Sheng DC, Smith DW, Sloan SW, Gens A. Finite element formulation and algorithms for unsaturated soils. Part 2: Verification and Applications. *International Journal for Numerical and Analytical Methods in Geomechanics* 2003;27(9):767–90.
- [9] Sotowski WT, Gallipoli D. Explicit stress integration with error control for the Barcelona Basic Model. Part I: Algorithms formulation. *Computers & Geotechnics* 2010;37(1–2):59–67. <http://dx.doi.org/10.1016/j.compgeo.2009.07.00>.
- [10] Sotowski WT, Gallipoli D. Explicit stress integration with error control for the Barcelona Basic Model. Part II: Algorithms efficiency and accuracy. *Computers & Geotechnics* 2010;37(1–2):68–81. <http://dx.doi.org/10.1016/j.compgeo.2009.07.00>.
- [11] Borja RI, Lee SR. Cam-Clay plasticity, part I: implicit integration of elastoplastic constitutive relations. *Comput Methods Appl Mech Eng* 1990;78:49–72.
- [12] Borja RI, Sama KM, Sanz PF. On the numerical integration of three-invariant elastoplastic constitutive models. *Comput Methods Appl Mech Eng* 2003;192:1227–58.
- [13] Hickman RJ, Gutierrez M. An internally consistent integration method for critical state models. *Int J Numer Anal Methods Geomech* 2005;29:227–48.
- [14] Sotowski WT, Hofmann M, Hofstetter G. Comparison of explicit and implicit integration schemes for the Barcelona Basic Model. In: Buzzi, Fityus, Sheng, editors. *Proceedings of the 4th Asia – Pacific Conference on Unsaturated Soils*. Taylor & Francis Group; 2010. p. 705–10.
- [15] Deufhard P, Bornemann F. *Scientific computing with ordinary differential equations*. Texts in applied mathematics, vol. 42. Springer; 2002.
- [16] Alonso EE, Gens A, Josa A. A constitutive model for partially saturated soils. *Géotechnique* 1990;40(3):405–30.
- [17] Wheeler SJ, Gallipoli D, Karstunen M. Comments on use of the Barcelona Basic Model for unsaturated soils. *Int J Numer Anal Methods Geomech* 2002;26:1561–71.
- [18] Sheng D. Non-convexity of the Barcelona Basic Model – Comment on SJ Wheeler, D Gallipoli & M Karstunen (2002; 26:1561–1571). *Int J Numer Anal Methods Geomech* 2003;27:879–81.
- [19] Fellin W, Ostermann A. Consistent tangent operators for constitutive rate equations. *Int J Numer Anal Methods Geomech* 2002;26:1213–33.
- [20] Sloan SW, Abbo AJ, Sheng DC. Accelerated convergence of Newton-Raphson method using least squares approximation to the consistent tangent matrix.

- In: Khalili and Oeser (Eds.), Proceedings of the 13th international conference of the IACMAG, 2011; vol. 1. p. 15–20.
- [21] Potts DM, Gens A. A critical assessment of methods of correcting for drift from the yield surface in elasto-plastic finite element analysis. *Int J Numer Anal Methods Geomech* 1985;9:149–59.
- [22] Abbo AJ, Sloan SW. An automatic load stepping algorithm with error control. *Int J Numer Methods Eng* 1996;39(10):1737–59.
- [23] Sheng DC, Sloan SW, Abbo AJ. An automatic Newton–Raphson scheme. *Int J Geomech* 2002;2(4):471–502.