Field soil aggregate stability kit for soil quality and rangeland health evaluations

J.E. Herrick [a,*], W.G. Whitford [b], A.G. de Soyza [a], J.W. Van Zee [a], K.M. Havstad [a], C.A. Seybold [c], M. Walton [a]

[a] *USDA–ARS Jornada Experimental Range, MSC 3JER, NMSU, Box 30003, Las Cruces, NM 88003, USA*
[b] *US EPA, ORD, NERL, Environmental Sciences Division-Las Vegas, P.O. Box 93478, Las Vegas, NV 81913 USA*
[c] *USDA–NRCS, Soil Quality Institute, Soil Science, ALS Bldg., Rm. 3017, Oregon State University, Corvallis, OR 97331 USA*

Abstract

Soil aggregate stability is widely recognized as a key indicator of soil quality and rangeland health. However, few standard methods exist for quantifying soil stability in the field. A stability kit is described which can be inexpensively and easily assembled with minimal tools. It permits up to 18 samples to be evaluated in less than 10 min and eliminates the need for transportation, minimizing damage to soil structure. The kit consists of two $21 \times 10.5 \times 3.5$ cm plastic boxes divided into eighteen 3.5×3.5 cm sections, eighteen 2.5-cm diameter sieves with 1.5-mm distance openings and a small spatula used for soil sampling. Soil samples are rated on a scale from one to six based on a combination of ocular observations of slaking during the first 5 min following immersion in distilled water, and the percent remaining on a 1.5-mm sieve after five dipping cycles at the end of the 5-min period. A laboratory comparison yielded a correlation between the stability class and percent aggregate stability based on oven dry weight remaining after treatment using a mechanical sieve. We have applied the method in a wide variety of agricultural and natural ecosystems throughout western North America, including northern Mexico, and have found that it is highly sensitive to differences in management and plant community composition. Although the field kit cannot replace the careful laboratory-based measurements of soil aggregate stability, it can clearly provide valuable information when these more intensive procedures are not possible. © 2001 Elsevier Science B.V. All rights reserved.

---

* Corresponding author. Tel.: +1-505-646-5194; fax: +1-505-646-5889.
E-mail address: jherrick@nmsu.edu (J.E. Herrick).

1. Introduction

Soil aggregate stability is widely recognized as a key indicator of soil (Karlen and Stott, 1994; Arshad et al., 1996) and rangeland health (reviewed in Herrick et al., 1999). It is closely related to a number of ecosystem properties, processes and functions, including the quantity and composition of soil organic matter (Tisdall, 1996), soil biotic activity (Wander et al., 1994), infiltration capacity (Pierson et al., 1994) and resistance to erosion (Blackburn and Pierson, 1994). The stability of larger macroaggregates, in particular, is largely a function of active soil organic matter fractions (Bethlenfalvay and Barea, 1994; Degens et al., 1994; Tisdall, 1996). These fractions have high turnover rates and are sensitive to management (Wander et al., 1994).

In spite of these positive attributes, aggregate stability is rarely included in soil quality and rangeland health evaluations. This is because the methods for evaluating aggregate stability suffer from a number of cost- and sampling-related limitations. Most existing methods are laboratory based (Beare and Bruce, 1993; also see Cammeraat, this volume). Laboratory facilities are frequently not available (but see simple kit described by Seybold and Herrick in this volume), and even when they are the cost is prohibitive. Measurement cost is further inflated by the need for high levels of replication, especially in rangeland settings. Aggregate stability varies widely across a variety of scales (Pierson et al., 1994) and soil textures. This problem is commonly dealt with by compositing. This can destroy the integrity of the aggregates. Furthermore, the variability itself may frequently be a good indicator of ecosystem condition (Herrick and Whitford, 1995).

Sampling and sample transport to laboratories are also major limitations. In non-cultivated arid and semi-arid soils, the most important aggregates for many ecosystem processes are located in the top 2–5 mm. In cultivated soil, the top 10–20 cm is generally sampled for changes in soil function. It is nearly impossible to sample and transport these materials, especially the thin surface samples, without modifying aggregate structure.

Our objective was to design a system which would meet the criteria of being rapid (under 1 min per replicate including sampling), repeatable, inexpensive, easy to assemble from readily available materials, permit field evaluations of relatively undisturbed samples, and reflect management-induced changes in soil structure.

2. Design and construction

The method is based on a combination of observations following immersion of a soil fragment in water (Emerson, 1967; Tongway and Hindley, 1995) and percent soil remaining following wet sieving (Kemper and Roseanau, 1986). The kit consists of two $21 \times 10.5 \times 3.5$ cm plastic boxes divided into eighteen 3.5×3.5 cm sections (Fig. 1a), eighteen 2.5-cm diameter sieves with 1.5-mm openings (Fig. 1b) and a small metal spatula used for sampling (Fig. 1c). One box is used for sample collection; the other is used for sample testing. The boxes are commonly sold as fishing tackle boxes in sporting goods stores and as "parts" boxes in electronic and hobby stores. A water-
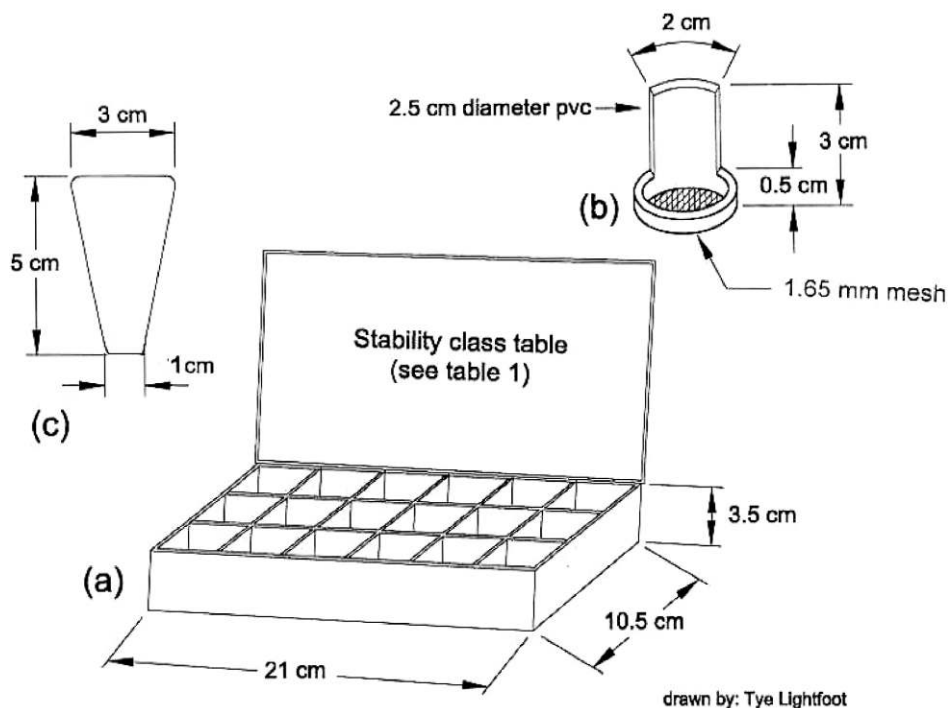
Fig. 1. Design for soil stability kit. (a) Box used for collecting and testing crust fragments. (b) One of 18 sieves included in the kit. (c) Sampling spatula.

proofed copy of the "Stability Class Table" (Table 1) is glued or taped to the inside of the test box. Construct each sieve from a 3-cm long section of 2.5-cm outer diameter

Table 1
Criteria for the assignment of crust fragments to stability classes

Stability class	Criteria for assignment to stability class (for Standard Characterization)[a]
0	Soil too unstable to sample (falls through sieve)[b]
1	**50%** of structural integrity lost within **5 s** of insertion in water
2	**50%** of structural integrity lost **5–30 s** after insertion
3	**50%** of structural integrity lost **30–300 s** fter insertion **or**
	< 10% of soil remains on sieve after five dipping cycles
4	**10–25%** of soil remains on sieve after five dipping cycles
5	**25–75%** of soil remains on sieve after five dipping cycles
6	**75–100%** of soil remains on sieve after five dipping cycles

Always sieve soils (even if rated ≤ 3) to verify class.

[a]Criteria may be modified to suit local requirements and conditions.

[b]Improper sampling can result in "false zeros". To verify, moisten the soil surface with a fine mist, then remove a sample (moisture will improve cohesion). Allow to air-dry, then perform the test. This should result in 0. If it does not, remove samples more carefully, or use this technique for all samples. For any site with zeros recorded, means should be calculated both with and without zeros included.

PVC tubing (thin wall). Remove three-quarters of the circumference from the top 2.5 cm of the tube, leaving a 2 (width) × 2.5 (length) cm tab. Glue standard aluminum window screen (1.5-mm opening) on one end of the 3-cm long PVC section. The spatula can be cut from any piece of thin metal such as a tin or aluminum can.

3. Methods

For most systems, we recommend a minimum of nine to eighteen samples (1/2 to 1 box) from two depths: the surface and the depth to which disturbance normally occurs. Where a physical or biological crust is present, use the top 2–4 mm for the surface sample. If a crust is not present, simply use the aggregate nearest to the sampling point. Each sample consists of a 6- to 8-mm diameter soil fragment. Smaller fragments may result in an underestimate, while larger fragments may lead to an overestimate. The fragment sampling points should be randomly selected. Dry soils have a tendency to shatter. This problem can be avoided by excavating a 1-cm deep hole the width of the spatula beside the sampling point, then carefully lifting up on the fragment. In extreme cases, the soil may be moistened with a light mist.

Place each sample in a separate sieve. All samples should be at a uniform moisture content. Tests are normally applied to air-dry samples. Although there are good arguments for pre-wetting all samples (Kemper and Roseanau, 1986; Beare and Bruce, 1993), we have omitted this step in most cases in order to save time. Instead, samples are air-dried by leaving them in the sieves for up to 1 h. Drying can be accelerated by placing the open kit on the dashboard of a vehicle or, on cloudy days, on top of the engine block. In systems that are normally moist, all samples may be pre-wetted using a third box. The third box is filled with fine sand and covered with gauze. Sufficient de-ionized water should be added to generate a 1-cm tension at the sand surface.

After collecting and equilibrating all 18 samples, fill the second (test) tray with de-ionized water to a depth of 2 cm and begin carefully immersing the samples in the cells at a rate of one every 15 s (Table 2). During the first 5 min, the samples are rated (Table 1, Classes 1–3) based on time to slaking. During this period, the test is nearly identical to the Emerson slake test. In our work with land managers and government agency personnel in both the US and Mexico, we have found that the word "melt" (derretir in Spanish) is more useful than "slake". After 5 min, sieve each sample by lifting it completely out of the water and returning it to the bottom five times at a rate of one complete cycle for every 2 s. After five cycles, rate the sample again. All samples, including those previously rated 3 or below, should be sieved. Ratings may be increased (e.g., from 2 or 3 to 4) depending on the results of the sieving. For example, a sample that appears to slake within 30 s but ends the test with 40% remaining on the sieve would be rated 5 after an initial rating of 2. This situation has been relatively uncommon (< 5%) in our experience. For management comparisons in areas with high stability, the number of dipping cycles can be increased. Samples that are hydrophobic will generally float. This should be noted and the sample gently pushed under the water until it is sufficiently wetted to remain immersed.

Table 2
Mean soil stability classes for bare and vegetated microsites at three transects on a grazing allotment in the Gila National Forest, NM, USA

	Surface (0–3 mm)			Sub-surface (20–25 mm)		
	Bare[a]	Vegetated	All	Bare[a]	Vegetated	All
Transect 1						
Average(S.E.)	2.5(n/a)	5.4(0.2)	4.8(0.2)	1.5(n/a)	4.7(0.2)	4.0(0.2)
Maximum	4.0	6.0	6.0	2.0	6.0	6.0
Minimum	1.0	3.0	1.0	1.0	3.0	1.0
n	2.0	7.0	9.0	2.0	7.0	9.0
Transect 2						
Average(S.E.)	5.0(n/a)	5.7(0.1)	5.6(0.1)	5.5(n/a)	3.3(0.2)	3.8(0.2)
Maximum	5.0	6.0	6.0	6.0	5.0	6.0
Minimum	5.0	5.0	5.0	5.0	1.0	1.0
S.E.	0.0	0.1	0.1	0.4	0.2	0.2
n	2.0	7.0	9.0	2.0	7.0	9.0
Transect 3						
Average(S.E.)	3.5(n/a)	4.3(0.2)	4.1(0.2)	2.0(n/a)	2.9(0.2)	2.7(0.1)
Maximum	6.0	6.0	6.0	2.0	5.0	5.0
Minimum	1.0	2.0	1.0	2.0	1.0	1.0
S.E.	1.8	0.2	0.2	0.0	0.2	0.1
n	2.0	7.0	9.0	2.0	7.0	9.0

[a] Sample sizes were insufficient to report standard errors for the bare microsites.

The depth of the sub-surface sample should be standardized within a monitoring program to reflect the depth(s) from which soil is brought to the soil surface by the current disturbance regime. In rangeland systems, 2–3 cm is generally appropriate as this represents the depth of soil that is most likely to be exposed at the surface by grazing disturbances. In cultivated systems, maximum tillage depth is commonly used. Other depths may be used according to the known or expected disturbance regime. For example, deeper samples should be collected in areas with high levels of burrowing by small mammals, or in forested systems in which whole-tree harvesting is planned.

4. Comparison with percent by weight

A comparison between qualitative evaluations of stability class and quantitative measurements of stability were completed using a Glendale clay loam (fine-silty, mixed (calcareous), thermic Typic Torrifluvent), a Belen sandy clay loam (montmorillonitic (calcareous), thermic Vertic Torrifluvent) and a Bluepoint loamy sand (mixed thermic, Typic Torripsamment). A total of twenty-four 5–8-mm diameter crust fragments were collected from a variety of microsites in order to generate a range of values. The air-dry fragments were weighed and placed in individual 1.5-mm opening sieves and placed on a motorized platform (described in Kemper and Rosenau, 1986), which generates 1.5 cm of vertical movement at a rate of one cycle every 2 s. This is approximately equivalent
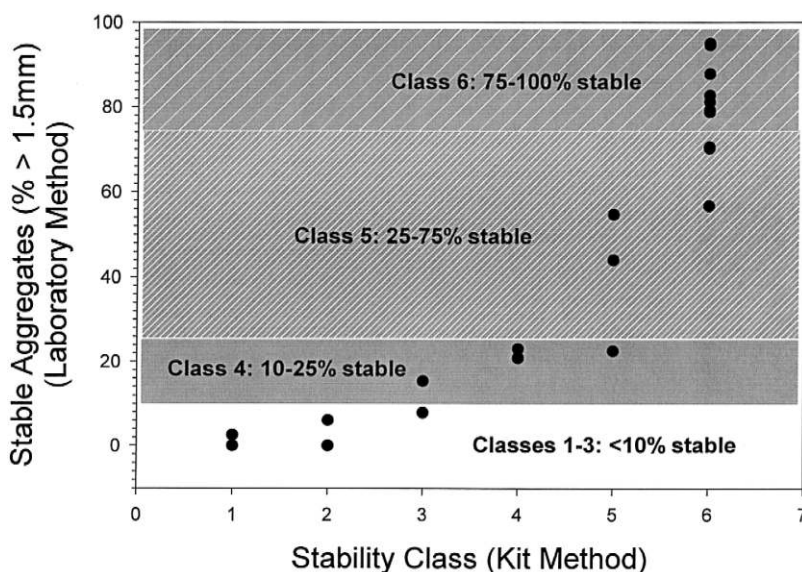
Fig. 2. Relationship between estimated stability class (Table 1) and laboratory determination of aggregate stability in which the same disruptive forces applied by hand in the field were generated mechanically in the lab and oven-dry weight replaced ocular estimation. Shaded areas represent expected percent stable aggregates for each stability class.

to the recommended procedure for the stability kit. The procedures described above for immersing and rating the fragments were followed using the motorized platform. At the end of the test, the fragments were oven-dried at 60°C and re-weighed. Aggregate stability was calculated as the soil remaining as a percent by weight of the original soil fragment. The stability kit values were correlated with percent aggregate stability (Fig. 2). Seventy-six percent of the values fell within the expected range based on the stability class, and all samples were correctly estimated within one class.

This procedure is similar to that of Kemper and Roseanau, except that we begin with a single, large (6–8 mm) fragment, a 1.5-mm screen is used instead of a 250u, we allow the sample to slake for 5 min before sieving instead of initiating sieving immediately upon immersion, the sieving time is only 10 s instead of 30 min, and the sample complete clears the surface of the water during sieving. Many of these differences appear in other modifications of Kemper and Roseanau's procedure (Beare and Bruce, 1993).

5. Field tests

We have applied the method in a wide variety of agricultural and natural ecosystems throughout western North America, including northern Mexico, and have found that it is highly sensitive to differences in management and plant community composition. It has also been evaluated and adopted for use in a citizen-soil quality monitoring program for

agricultural soils in the state of Illinois, USA. The pre-adoption evaluation included a wide range of soil textures, from clay loams to sands. The only soils in which the test has not been found to be useful are wetland and extremely sandy soils in which there is little formation of aggregates larger than 1.5 mm in diametereven under good conditions. Some of the most useful comparisons in arid systems have been within sites: ratios between average values under plants and in plant interspaces (Herrick and Whitford, 1995).

We have found that most users are able to easily complete 18 tests in 10 min by following the schedule included in the data form (Fig. 3). Allowing for 10 min for sampling and up to 10 min for setup and cleanup, we have found that a site can be adequately characterized in less than 30 min. We have successfully used it as an evaluation, monitoring and educational tool (e.g., de Soyza et al., 1997).

The results of one field test completed on September 24, 1998 are presented in Tables 2 and 3. Samples were removed from randomly selected points along each of three permanent 30-m vegetation transects located on fine sandy loam soils within a 5900-ha grazing allotment within the Gila National Forest in New Mexico, USA. Earlier evaluations based only on species composition data indicated that the allotment was in fair to poor condition. All three transects are currently dominated by blue grama (*Bouteloua gracilis*).

Soil stability values ranged from 1 to 6. This high level of variability is typical for rangeland settings and reflect spatial variability in organic matter inputs and aggregation and degradation processes. Values were higher at the surface, where cyanobacterial filaments were observed, and generally higher under vegetation, where organic matter tends to concentrate and raindrop impact is limited. Values were relatively lower (many below 4) in bare areas, especially in the sub-surface. Surface gravel was absent from all three transects and litter cover was generally confined to areas beneath plant canopies. This indicates that erosion susceptibility is likely to be increased by grazing disturbances, which tend to bring sub-surface soil to the surface in the plant interspaces. However, plant cover was quite high and the average distance between plant canopies

Surface (left side of box)

Immerse	Sieve	Class	Immerse	Sieve	Class	Immerse	Sieve	Class
0:00	5:00		0:45	5:45		1:30	6:30	
0:15	5:15		1:00	6:00		1:45	6:45	
0:30	5:30		1:15	6:15		2:00	7:00	

Sub-Surface (right side of box)

Immerse	Sieve	Class	Immerse	Sieve	Class	Immerse	Sieve	Class
2:15	7:15		3:00	8:00		3:45	8:45	
2:30	7:30		3:15	8:15		4:00	9:00	
2:45	7:45		3:30	8:30		4:15	9:15	

Fig. 3. Data form and testing schedule, which permits a full box of 18 fragments to be tested in under 10 min. The table is designed so that it reflects the organization of the cells in the box, with the "Surface" samples located in the nine cells on the left side the box, and the "Sub-surface" samples located in the nine cells on the right side of the box. Times are in min:s.

Table 3
Vegetation characteristics for transects from which soil samples were collected on a grazing allotment in the Gila National Forest, NM, USA

Transect	Areal cover (%)	Basal cover (%)	Distance between plant canopies (cm)		Distance between plant bases (cm)	
			Average	Maximum	Average	Maximum
2	66.7	12.8	13	143	27	166
4	71.7	16.3	11	42	21	110
7	64.2	14.7	11	54	21	92

was quite low at all three locations (Table 3), reducing the probability that erosion cells will develop. This example illustrates the importance of interpreting soil stability kit values in the context of vegetation cover, spatial distribution and disturbance regime.

6. Conclusions

The stability kit described here can be inexpensively and easily assembled with minimal tools. It permits a large number of samples to be evaluated quickly and eliminates the need to transport samples, minimizing damage to soil structure. Although it cannot replace careful laboratory-based measurements of soil aggregate stability, it can clearly provide valuable information when these more intensive procedures are not possible.

Acknowledgements

References

Arshad, M.A., Lowery, B., Grossman, B., 1996. Physical tests for monitoring soil quality. In: Doran, J.W., Jones, A.J. (Eds.), Methods for Assessing Soil Quality. SSSA Spec. Publ., vol. 49. Soil Sci. Soc. Am, Madison, WI, USA, pp. 123–141.

Beare, M.H., Bruce, R.R., 1993. A comparison of methods for measuring water-stable aggregates: implications for determining environmental effects of soil structure. Geoderma 56, 87–104.

Bethlenfalvay, G.J., Barea, J.-M., 1994. Mycorrhizae in sustainable agriculture: I. Effects on seed yield and soil aggregation. Am. J. Altern. Agric. 9, 157–161.

Blackburn, W.H., Pierson Jr., F.B., 1994. Sources of variation in interrill erosion on rangelands. In: Blackburn, W.H., Pierson, F.B. Jr., Schuman, G.E., Zartman, R. (Eds.), Variability in Rangeland Water Erosion Processes. Soil Sci. Soc. Am. J, Madison, WI, USA, pp. 1–10.

Degens, B.P., Sparling, G.P., Abbott, L.K., 1994. The contribution from hyphae, roots and organic carbon constituents to the aggregation of a sandy loam under long-term clover-based and grass pastures. Eur. J. Soil Sci. 45, 459–468.

de Soyza, A.G., Whitford, W.G., Herrick, J.E, 1997. Sensitivity testing of indicators of ecosystem health. Ecosyst. Health 3, 44–53.

Emerson, W.W., 1967. A classification of soil aggregates based on their coherence in water. Aust. J. Soil Res. 5, 47–57.

Herrick, J.E., Whitford, W.G., 1995. Assessing the quality of rangeland soils: challenges and opportunities. J. Soil Water Conserv. 50, 237–242.

Herrick, J.E., Weltz, M.A., Reeder, J.D., Schuman, G.E., Simanton, J.R., 1999. Rangeland soil erosion and soil quality: role of soil resistance, resilience, and disturbance regime. In: Lal, R. (Ed.), Soil Erosion and Soil Quality. CRC Press, Boca Raton, FL, USA, pp. 209–233.

Karlen, D.L., Stott, D.E., 1994. A framework for evaluating physical and chemical indicators of soil quality. In: Doran, J.W., Coleman, D.C., Bezdicek, D.F., Stewart, B.A. (Eds.), Methods for Assessing Soil Quality. SSSA Spec. Publ., vol. 35. Soil Sci. Soc. Am, Madison, WI, USA, pp. 53–72.

Kemper, W.D., Roseanau, R.C., 1986. Aggregate stability and size distribution. In: Klute, A. (Ed.), Methods of Soil Analysis: Part I. 2nd edn. Agron. Monogr., vol. 9. Am. Soc. Agron. and Soil Sci. Soc. Am, Madison, WI, USA, pp. 425–442.

Pierson, F.B., Blackburn, W.H., Van Vactor, S.S., Wood, J.C., 1994. Partitioning small scale spatial variability of runoff and erosion on sagebrush rangeland. Water Res. Bull. 30, 1081–1089.

Tisdall, J.M., 1996. Formation of soil aggregates and accumulation of soil organic matter. In: Carter, M.R., Stewart, B.A. (Eds.), Structure and Organic Matter Storage in Agricultural Soils. CRC Lewis, Boca Raton, FL, USA, pp. 57–96.

Tongway, D., Hindley, N., 1995. Manual for Soil Condition and Assessment of Tropical Grasslands. CSIRO, Canberra, Australia, 60 pp.

Wander, M.M., Traina, S.J., Stinner, R.B., Peters, S.E., 1994. The effects of organic and conventional management on biologically-active soil organic matter pools. Soil Sci. Soc. Am. J. 58, 1130–1139.