Open Journal of Statistics, 2026, 16(4), 247-258
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Robustness of Fitted Behavioural Relationships from Limited Geotechnical Datasets: Application to Sand Castle Test Data

Hans Rönnqvist

RQV Teknik AB, Hudiksvall, Sweden
Email: hans.ronnqvist@rqv.se

How to cite this paper: Rönnqvist, H. (2026) Robustness of Fitted Behavioural Relationships from Limited Geotechnical Datasets: Application to Sand Castle Test Data. *Open Journal of Statistics*, **16**, 247-258.
https://doi.org/10.4236/ojs.2026.164011

Received: June 21, 2026
Accepted: July 26, 2026
Published: July 29, 2026

Abstract

Engineering decisions are often based on fitted relationships derived from limited datasets, where individual observations may exert disproportionate influence on interpretation. This paper presents a practical framework for assessing the robustness of such relationships using a simple leave-one-out influence analysis. The approach is demonstrated using Sand Castle Test (SCT) data, in which relative compaction (RC) is related to collapse time (T) through a logarithmic relationship governing collapsibility and crack-stopping behaviour. The fitted slope is treated as a sensitivity descriptor, while the compaction window ΔRC (15 → 60) provides a physically meaningful measure of behavioural transition. Two case studies illustrate both robust and point-sensitive datasets. A normalised slope influence index is introduced to quantify dependence on individual observations and to identify data points controlling interpretation. Results show that apparently well-defined relationships may be sensitive to single observations, and that uncertainty in slope propagates directly into uncertainty in the inferred transition between collapsible and crack-holding behaviour. The proposed framework links statistical robustness to engineering interpretation by enabling identification of critical data points and supporting more defensible assessment of behaviour. Although demonstrated for SCT data, the approach is broadly applicable to geotechnical problems where fitted relationships are used to define project-specific criteria from limited datasets.

Keywords

Sand Castle Tests, Compaction-Collapse Behaviour, Robustness Analysis, Geotechnical Data Interpretation, Leave-One-Out Influence Analysis

1. Introduction

In many areas of engineering, behavioural interpretation relies on fitted relationships derived from relatively small experimental datasets. This is particularly common in geotechnical engineering, where laboratory testing is time-consuming and datasets are often sparse. Under such conditions, regression relationships may appear well defined while in reality being strongly influenced by individual observations.

In embankment dam engineering, filters are required both to retain migrating core material and to remain sufficiently collapsible to close developing cracks rather than maintain an open flow path [1] [2]. Crack-stopping behaviour and controlled collapsibility are therefore closely linked aspects of performance. The Sand Castle Test (SCT) is commonly used to assess this behaviour. In this test, a compacted sand column is subjected to controlled wetting, and the time to collapse provides a measure of apparent cohesion and stability. In practice, results are interpreted using collapse-time markers (notably $T = 15$ min and $T = 60$ min) as introduced by Soroush *et al.* [3], together with a logarithmic relation between relative compaction (RC) and collapse time (T). The resulting RC-log(T) representation forms the basis for interpretation, from which quantities such as RC@15 and ΔRC (15 → 60) are derived.

From a statistical perspective, inference based on small datasets is inherently sensitive to individual observations, and fitted parameters may vary under resampling or leave-one-out procedures [4]. While formal influence diagnostics exist (e.g. [5]), their results are not always directly interpretable in engineering terms. There is therefore a need for simple and transparent methods that allow practitioners to assess the robustness of fitted relationships and their implications for behavioural interpretation.

This need is consistent with long-recognised uncertainties in embankment dam engineering. As emphasised by Milligan [6], the ability to perform detailed analytical modelling often exceeds the ability to characterise material behaviour and construction conditions with comparable certainty. Peck [7] similarly noted that failures of earth dams are seldom attributable to deficiencies in analytical methods, but rather to misjudgement of conditions not readily captured by calculation.

The present study evaluates the robustness of fitted SCT compaction-collapse relationships using a leave-one-out approach. The analysed datasets are drawn from the SCT framework established by Rönnqvist [8]. While SCT data provide the case study, the methodology is general and applicable wherever behavioural interpretation relies on fitted relationships derived from limited datasets.

The central question addressed is therefore: how robust are fitted relationships used to interpret engineering behaviour, and to what extent do inferred thresholds depend on specific data points? To address this, a practical framework for robustness assessment is introduced based on leave-one-out influence analysis. The framework links statistical sensitivity to engineering interpretation by distinguishing four aspects: 1) presence of a behavioural threshold, 2) sensitivity of behaviour to the governing variable, 3) transition interval between behavioural states, and 4) robustness of the fitted relation.

2. Methodology and Conceptual Basis

The conceptual basis of the fitted relation and its engineering projection is illustrated in Figure 1, which defines the slope b and its projection as the compaction window ΔRC (15 → 60).

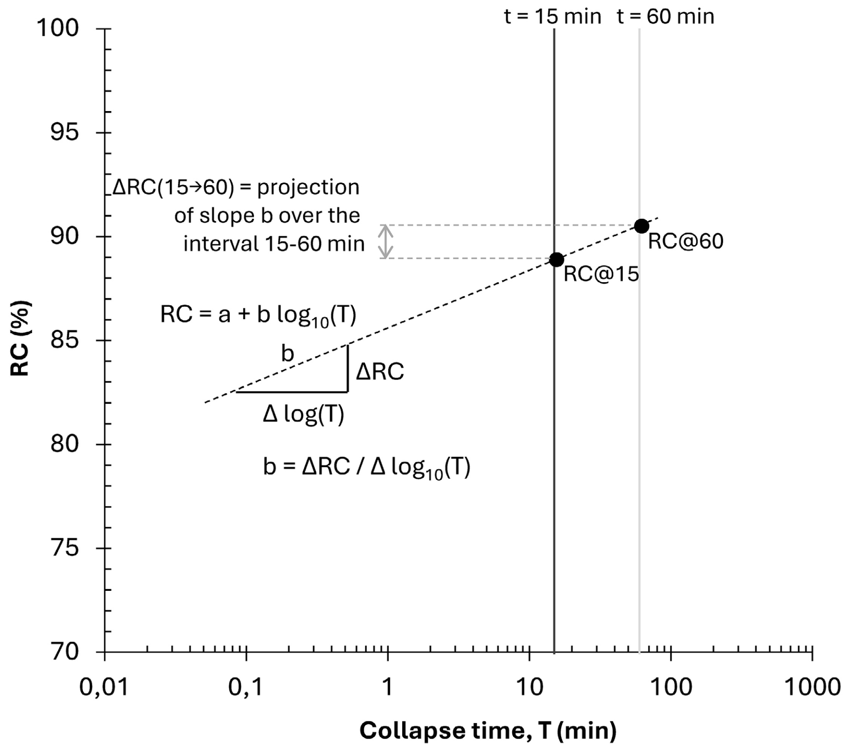

Figure 1. Conceptual illustration of the fitted relationship between relative compaction (RC) and collapse time (T). The fitted relation follows the form RC = a + $b\log_{10}(T)$, where b represents the slope of the RC-log10(T) relationship. The intersections with T = 15 min and T = 60 min define RC@15 and RC@60, and the vertical difference between these values defines the compaction window ΔRC (15 → 60).

2.1. Fitted Relationship

The fitted relation between relative compaction and collapse time is expressed as

$$\mathrm{RC} = a + b\log_{10}(T) \tag{1}$$

where a is the intercept and b is the slope of the RC-$\log_{10} T$ relation. The slope represents the rate of change of RC with respect to logarithmic collapse time.

2.2. General Slope Relation and Engineering Projection

The slope may be obtained from any two points on the fitted line:

$$b = \frac{\Delta \mathrm{RC}}{\Delta \log_{10}(T)} \tag{2}$$

An engineering transition interval between two collapse times T_1 and T_2 is therefore:

$$\Delta\mathrm{RC}\left(T_1 \to T_2\right) = b\log_{10}\left(\frac{T_2}{T_1}\right) \tag{3}$$

For the practical SCT anchors $T = 15$ min and $T = 60$ min:

$$\Delta\mathrm{RC}\left(15 \to 60\right) = b\log_{10}\left(\frac{60}{15}\right) \approx 0.60206\,b \tag{4}$$

Thus, ΔRC (15 → 60), as defined in Equation (4), is a projection of the fitted slope over a specified interval rather than an independent descriptor. Logarithms are expressed in base 10 throughout. The choice of logarithmic base affects only the numerical scaling of the fitted slope and does not influence interpretation.

2.3. Interpretation Variables

The fitted relation provides three complementary descriptors:

- RC@15: indicator of a behavioural threshold.
- b: sensitivity of RC to logarithmic time.
- ΔRC (15 → 60): projected transition interval.

2.4. Robustness Assessment

Robustness is evaluated using leave-one-out regression. Slopes b_{-i} are computed after removing each observation in turn.

$$\Delta b_{\text{spread}} = \max\left(b_{-i}\right) - \min\left(b_{-i}\right) \tag{5}$$

$$I_b = \frac{\Delta b_{\text{spread}}}{|b|} \tag{6}$$

The index I_b defined in Equation (6) quantifies overall sensitivity to individual observations, while Δb_i (Equation (7)) identifies specific influential points:

$$\Delta b_i = \frac{b_{-i} - b}{b} \tag{7}$$

The index I_b quantifies overall sensitivity of the fitted slope to individual observations, while Δb_i identifies specific influential points. Low values of I_b indicate a stable fitted relationship, whereas higher values indicate increasing dependence on individual observations. In the present examples, values of I_b on the order of 0.2 distinguish relatively stable from point-sensitive datasets and are therefore proposed as an indicative threshold for interpretation. $I_b = 0.2$ corresponds approximately to a situation where the total leave-one-out slope spread approaches 20% of the fitted slope magnitude. Further validation across broader datasets is required before general application.

Beyond this level, changes in the interpreted compaction window become sufficiently large that engineering interpretation may begin to depend noticeably on individual observations rather than the overall material response. The threshold should therefore be viewed as practical engineering guidance rather than a strict statistical limit. For elevated point sensitivity, the practitioner should examine the individual influence values Δb_i to identify which observations control the fitted slope. If a single point strongly steepens the regression, it may be appropriate to assess the alternative, flatter relation obtained when that point is removed and to consider whether this represents a more conservative interpretation for engineering purposes.

2.5. Preferred Basis of Interpretation

The slope b should be determined from the fitted regression rather than from individual point pairs. This approach uses the full dataset, reduces the effect of local scatter and provides a consistent basis for comparison between materials. It also allows interpretation even when $T = 60$ min is not strongly constrained by the achieved data.

2.6. Reference Dependence and Invariance

Relative compaction RC depends on the chosen compaction reference (for example Modified Proctor or Standard Proctor). Changing reference shifts absolute RC values, and therefore RC@15, but does not affect b or ΔRC (15 → 60), which depend only on differences along the fitted relation. This distinction is important when comparing datasets obtained using different compaction standards.

3. Case Study Dataset

The data analysed in this study originate from the SCT database presented by Rönnqvist [8], comprising natural and crushed filter materials tested over a range of relative compaction levels. For each specimen, collapse time is measured under controlled wetting and fitted logarithmic compaction-collapse relationships are derived. Method A defines compaction through density, expressed as relative compaction (RC), and measures the time required for collapse of a constructed sand column under controlled wetting.

In this note, RC is referenced to the maximum dry density obtained from Modified Proctor compaction [9]. If Standard Proctor compaction [10] were used instead, the absolute RC values would shift, but the slope b and the derived ΔRC (15 → 60) would remain unchanged because they depend only on differences in RC along the fitted relation. Consequently, quantities such as RC@15 shift with the compaction reference, whereas b and ΔRC (15 → 60) do not.

4. Results

The fitted relationships for the two representative materials are shown in **Figure 2**, which illustrates contrasting slopes and compaction windows for Materials A and D, while **Figure 3** shows the corresponding leave-one-out influence analysis and variation in slope under omission of individual observations. A quantitative comparison is summarised in **Table 1**, which presents slope, compaction window and robustness metrics for the two materials.

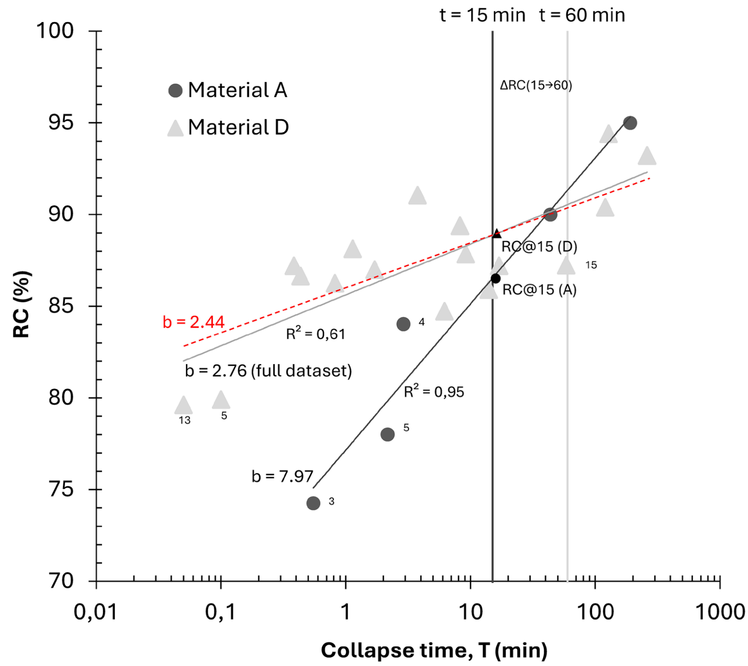

Figure 2. Fitted compaction-collapse relationships showing relative compaction (RC) as a function of collapse time (T, log scale) for Materials A and D. Material A exhibits a stable fitted slope and broad compaction window, whereas Material D shows a flatter slope and narrower transition interval. For Material D, the red dashed line shows the fitted relationship obtained after removal of the most influential observation (Point 5), resulting in a reduction of slope from $b \approx 2.76$ to $b \approx 2.44$. This illustrates how individual observations may influence the fitted relationship and the derived compaction window ΔRC (15 → 60).

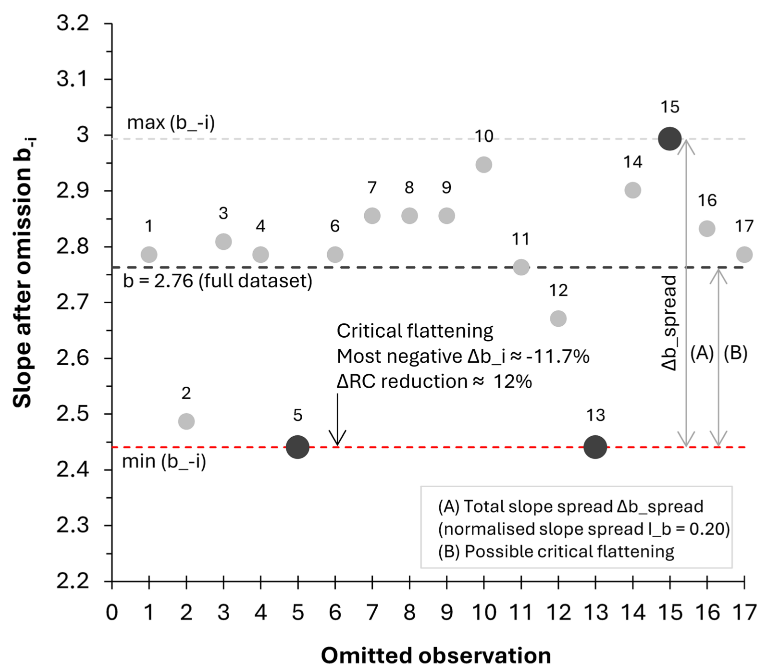

Figure 3. Leave-one-out slope analysis showing variation in fitted slope b under omission of individual observations based on sensitive case D. The resulting spread defines the robustness of the fitted relationship and forms the basis for the slope influence index I_b.

Table 1. Comparison of fitted relationship parameters for Materials A and D.

Material	b	ΔRC (15 → 60) (%)	R^2	I_b	Interpretation
A	7.97	4.7	0.95	0.11	Strong trend, high robustness
D	2.76	1.7	0.61	0.20	Lower correlation, elevated sensitivity

The fitted relations describe how achieved compaction governs collapse behaviour in the SCT. The slope b (Equation (2)) and the derived compaction window ΔRC (15 → 60) (Equation (4)) therefore directly control the interpretation of whether a material remains collapsible or transitions towards crack-holding behaviour. Errors in the fitted slope translate directly into errors in the inferred transition interval and may therefore affect engineering judgement regarding filter performance.

4.1. Example A: Robust Case

Material A represents a high-correlation case ($R^2 \approx 0.95$) with slope $b \approx 7.97$ and a comparatively wide compaction window ΔRC (15 → 60) ≈ 4.7%. This indicates relatively low sensitivity to densification across the tested range, and a stable interpretation of collapse behaviour with respect to compaction. Leave-one-out analysis yields a low influence index ($I_b \approx 0.11$), with the largest absolute point influence of about 5.8%, confirming that the fitted slope is highly stable with respect to removal of individual observations. The variation in slope is limited, indicating that the fitted relation is representative of the dataset rather than controlled by specific points.

4.2. Example D: Sensitive Case

Material D represents a contrasting case with lower correlation ($R^2 \approx 0.61$), slope $b \approx 2.76$, and a narrower compaction window ΔRC (15 → 60) ≈ 1.7%. This indicates higher sensitivity to densification. Leave-one-out analysis yields $I_b \approx 0.20$, indicating elevated point sensitivity. Removal of individual observations produces noticeable changes in slope, with the largest absolute point influence of about 11.7%, reflecting the greater dependence of the fitted relation on individual observations, and implying that the inferred transition between collapsible and crack-holding behaviour may shift depending on the dataset.

This effect is illustrated in **Figure 2**, where removal of the most influential observation (Point 5) results in a reduction of the fitted slope from $b \approx 2.76$ to $b \approx 2.44$. The corresponding flattening of the fitted relationship leads to a reduced compaction window ΔRC (15 → 60), demonstrating how interpretation of collapse behaviour may shift depending on individual observations. This highlights the practical importance of identifying influential data points when interpreting SCT results.

4.3. Comparison

The comparison shows that regression quality and robustness are related but not identical. The present examples show that moderate R^2 does not necessarily invalidate a fitted trend, while robustness must be evaluated independently, as high correlation alone does not guarantee low point sensitivity. Material A combines high correlation with low sensitivity, whereas Material D exhibits both weaker correlation and greater dependence on individual observations. This highlights the need to evaluate robustness explicitly rather than relying solely on regression metrics.

The quantitative comparison of slope, compaction window and influence index is provided in Table 1, while detailed variation in leave-one-out slopes is summarised in Table 2.

Table 2. Summary of leave-one-out slope variation for representative datasets.

Material	b	max(b_{-i})	min(b_{-i})	I_b	Largest Δb_i (%)
A	7.97	8.38	7.51	0.11	5.8
D	2.76	2.99	2.44	0.20	11.7

5. Discussion

5.1. Engineering Interpretation of Fitted Relationships

The interpretation of fitted relationships should be supported by both graphical representation (Figures 1-3) and quantitative measures (Tables 1-3). Figure 1 defines the compaction window concept, Figure 2 presents fitted behaviour, and Figure 3 quantifies robustness, while the tables provide corresponding numerical measures.

Table 3. Most influential observations for Materials A and D based on leave-one-out analysis.

Material	Omitted point	RC (%)	T (min)	b_{-i}	Δb_i
D	5	79.9	0.10	2.44	−0.117
D	13	79.6	0.05	2.44	−0.117
D	15	87.3	58.48	2.99	0.083
A	3	74.3	0.55	7.51	−0.058
A	4	84.0	2.90	8.38	0.052
A	5	78.0	2.17	7.64	−0.040

From an engineering perspective, the fitted relationship describes how achieved compaction governs collapse behaviour and therefore crack-stopping ability. The slope b (Equation (2)) defines sensitivity, while the compaction window ΔRC (15 → 60) (Equation (4)) expresses the rate of transition between practical collapse-time anchors. A narrow window implies that small variations in achieved density may shift a material from collapsible to crack-holding behaviour, whereas a broader window indicates a more gradual transition. The reliability of this interpretation therefore depends directly on the stability of the fitted slope.

5.2. Statistical Robustness and Point Sensitivity

Established statistical influence measures, such as Cook's distance and DFFITS, as well as robust regression techniques, provide formal tools for identifying influential observations in fitted relationships. However, these approaches are not always directly interpretable in terms of engineering quantities. The present framework complements such methods by expressing sensitivity directly in terms of slope variation and the resulting change in interpreted behavioural thresholds, providing a transparent and practically applicable measure of robustness. From a statistical perspective, fitted relationships derived from limited datasets may appear stable while being influenced by individual observations. This sensitivity is quantified using leave-one-out analysis, where the robustness index I_b measures overall slope stability and Δb_i identifies specific influential points. It identifies vulnerable data points, quantifies how ΔRC (15 → 60) may shift due to those points, and supports conservative interpretation through alternative, flatter fits when sensitivity is elevated. In the case studies, Material A shows low Δb_i and stable ΔRC, indicating a reliable fit, whereas Material D exhibits influential points that materially alter the slope, requiring targeted scrutiny and cautious interpretation. The practical impact of point sensitivity is illustrated in **Figure 2**, where removal of a single influential observation produces a measurable change in slope and corresponding compaction window, demonstrating how statistical sensitivity translates directly into engineering interpretation. Influential observations should not be interpreted solely as statistical anomalies. In geotechnical testing, such points may reflect genuine transitions in material behaviour, for example associated with changes in moisture conditions or structural collapse mechanisms. The present framework therefore identifies sensitivity in the fitted relationship but does not prescribe removal of observations; rather, it highlights cases where interpretation should be supported by experimental judgement.

5.3. Implications for Threshold-Based Interpretation

This distinction is critical when fitted relationships are used to define behavioural thresholds. High correlation does not guarantee robustness, and moderate correlation does not invalidate a trend. For point-sensitive datasets, interpretation should therefore consider plausible alternative fits derived from removal of influential observations.

For datasets exhibiting both moderate regression quality and elevated point sensitivity, interpretation should be undertaken conservatively. A practical approach is to consider the most negative Δb_i, corresponding to a flatter fitted relation, as a lower-bound estimate of behaviour. Reductions in fitted slope on the order of 10% - 15% may produce comparable reductions in the derived compaction window ΔRC (15 → 60).

In practical terms, this may correspond to density differences on the order of tens of kg/m^3, which are within the range of normal construction variability. Consequently, materials exhibiting narrow compaction windows and elevated sensitivity may transition from collapsible to crack-holding behaviour under routine variations in achieved density. In such cases, apparent collapsibility may be overestimated if robustness is not explicitly considered.

5.4. Link to Proposed Framework

The combined use of graphical representation, fitted relationships, and robustness metrics provides a consistent framework for interpretation. Engineering behaviour (ΔRC), statistical stability (I_b and Δb_i), and practical consequence (density variation) are directly linked, allowing informed and transparent assessment of collapse behaviour from limited datasets.

5.5. Proposed Practical Framework for Robustness Assessment

The proposed framework is intentionally simple. Its novelty lies not in the statistical method itself, but in its direct application to engineering interpretation. By linking slope sensitivity to a physically meaningful descriptor (ΔRC) and to practical decision-making, the method provides a transparent tool for assessing confidence in fitted relationships derived from limited datasets.

Proposed procedure:

1) Evaluate robustness

Fit the RC-log(T) relationship and compute I_b and Δb_i using leave-one-out analysis. Classify the dataset as robust (low I_b) or point-sensitive (elevated I_b). In practical implementation, regression fitting in common software environments may use either base-10 or natural logarithms. The choice of logarithmic base does not affect interpretation, provided consistency is maintained. Where natural logarithms are used, the fitted slope should be converted to the equivalent base-10 slope to ensure consistency with the definition of ΔRC (15 → 60). This is achieved by dividing the slope obtained from ln(T) by ln(10), *i.e.* $b = b_{ln}/\ln(10)$, where b_{ln} is the slope obtained from the RC-ln(T) relation. This conversion ensures that derived quantities remain directly comparable and physically interpretable within the SCT framework.

2) Identify and verify influential observations

For point-sensitive datasets, examine Δb_i to identify controlling observations. Check data quality, test conditions, and representativeness. Where warranted, repeat laboratory tests or review field data to confirm or refute the influence of these points.

3) Adopt a defensible interpretation

If influential points cannot be conclusively resolved, adopt a conservative interpretation by considering alternative fits (e.g. a flatter slope corresponding to the most negative Δb_i). Translate the resulting change in slope into ΔRC (15 → 60) using Equation (4) and reassess the implied transition between collapsible and crack-holding behaviour.

4) Update engineering decisions

Where reduced ΔRC indicates a narrow compaction window, adjust specifications and procedures accordingly (e.g. tighter compaction control, revised acceptance criteria, or additional verification testing) to account for construction variability and reduce the risk of unintended transition to crack-holding behaviour.

A similar procedure may be applied in other geotechnical contexts where fitted relationships are used to define project-specific criteria, allowing robustness assessment to be linked directly to governing parameters and acceptance thresholds relevant to the problem at hand.

6. Conclusions

- Fitted relationships derived from limited datasets may appear reliable but can be sensitive to individual observations; uncertainty in the fitted slope therefore propagates directly into uncertainty in the compaction window ΔRC (15 → 60) and the inferred transition between collapsible and crack-holding behaviour.
- Leave-one-out analysis provides a simple and effective diagnostic that quantifies slope stability and identifies influential data points controlling interpretation.
- The framework enables practical engineering actions, including verification of critical data points and adoption of conservative interpretations based on alternative, flatter fits when sensitivity is elevated.
- Application to the case studies demonstrates how the method distinguishes between robust and point-sensitive datasets, identifies influential observations, and supports more defensible interpretation of compaction-collapse behaviour.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Vaughan, P.H. and Soares, H.F. (1982) Design of Filters for Clay Cores of Dams. *Journal of the Geotechnical Engineering Division*, **108**, 17-31. https://doi.org/10.1061/ajgeb6.0001233

[2] FEMA (2011) Filters for Embankment Dams: Best Practices for Design and Construction. Federal Emergency Management Agency, Denver, CO, USA.

[3] Soroush, A., Tabatabaie Shourijeh, P., Aghajani, H.F., Mohammadinia, A. and Aminzadeh, A. (2012) A Review of the Sand Castle Test for Assessing Collapsibility of Filters in Dams. *Geotechnical Testing Journal*, **35**, 503-516. https://doi.org/10.1520/gtj103051

[4] Efron, B. and Tibshirani, R.J. (1993) An Introduction to the Bootstrap. Chapman &

Hall.

[5] Cook, R.D. (1977) Detection of Influential Observation in Linear Regression. *Technometrics*, **19**, 15-18. https://doi.org/10.1080/00401706.1977.10489493

[6] Milligan, V. (2003) Some Uncertainties in Embankment Dam Engineering. *Journal of Geotechnical and Geoenvironmental Engineering*, **129**, 785-797. https://doi.org/10.1061/(asce)1090-0241(2003)129:9(785)

[7] Peck, R.B. (1980) "Where Has All the Judgment Gone?" The Fifth Laurits Bjerrum Memorial Lecture. *Canadian Geotechnical Journal*, **17**, 584-590. https://doi.org/10.1139/t80-065

[8] Rönnqvist, H. (2026) Crack-Stopping in Natural and Crushed Filter Materials Assessed Using the Sand Castle Test. *Geotechnique Letters*. https://doi.org/10.1680/jgele.26.00030

[9] ASTM D1557 (2014) Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Modified Effort. ASTM International.

[10] ASTM D698 (2021) Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Standard Effort. ASTM International.

List of Notations

RC	Relative compaction (%)
SCT	Sand castle test
T	Collapse time (min)
T_1, T_2	General collapse times defining an interval (min)
a	Intercept of fitted relation ($\log_{10}$ base)
b	Slope of RC-$\log_{10}(T)$ relation
b_{-i}	Slope with observation i removed
Δb_i	Relative change in slope due to removal of observation i
Δb_{spread}	Absolute slope spread from leave-one-out analysis
I_b	Normalised slope influence index
ΔRC (15 → 60)	Compaction window between $T = 15$ min and $T = 60$ min
ΔRC ($T_1 \to T_2$)	General compaction window between two collapse times
RC@15	Relative compaction at $T = 15$ min
RC@60	Relative compaction at $T = 60$ min
R^2	Coefficient of determination