**Original Scientific paper**
10.7251/AGREN2403039J
UDC 633.1:631.41(64)

# DEVELOPMENT OF A PARTICLE SIZE IMBALANCE INDEX TO DETERMINE THE SUITABILITY OF FEZ AND MEKNES SOILS FOR CEREAL CROPS

Ahmed JELLOUL[1], Lotfi KHIARI[1*], Saad DRISSI[2]

[1]Department of Soils and Food Engineering, Faculty of Agriculture and Food Sciences, Laval University, Quebec, Canada
[2]Department of Agronomy and Plant Breeding, National School of Agriculture, Meknes, Morocco
*Corresponding author: lotfi.khiari@fsaa.ulaval.ca

## ABSTRACT

To optimize cereal cultivation in the agro-systems of Fez-Meknes, Morocco, this study emphasizes the importance of soil texture and organic matter for cereal productivity, particularly in arid and semi-arid climates. The objective is to develop a comprehensive index based on five parameters of particle size properties (percentage of clay, fine silt, coarse silt, sand, and organic matter) to predict the yield potential of wheat (*Triticum aestivum*). Utilizing 298 independent experimental sites in the agricultural region of Fez-Meknes, wheat yields were analyzed in relation to soil texture. The analysis involved a complete compositional simplex ($S^5$), a log-centered transformation, and the calculation of five cumulative variance functions of the components, resulting in five cubic models. By identifying inflection points in these models linking cumulative functions to wheat yields, a critical yield of 24 quintals per hectare was determined, distinguishing between more and less productive wheat populations. The developed global particle size imbalance index, named $r^2_{granulo}$, exhibited a theoretical value of 3.1 and a value validated by the Cate-Nelson partitioning method of 5.2. These results indicate that the $r^2_{granulo}$ index can serve as a reliable indicator for evaluating soil suitability for cereal production in arid and semi-arid regions. An index value below 5.2 suggests favorable soil conditions for cereal cultivation. This index offers a practical tool for farmers and agronomists to optimize cereal yields by effectively assessing and managing soil properties.

**Keywords:** *Soil texture, compositional analysis, cereal soil, Simplex of composition, Morocco.*

## INTRODUCTION

In Morocco, approximately 69% of the national territory consists of agricultural land, although only 18% is arable. Cereals, mainly wheat and barley, dominate the national agriculture, covering 5.3 million hectares, representing 59% of the country's 8.8 million hectares of cultivated land (Harbouze *et al.*, 2019). Cereals are essential for the agricultural economy and fundamental in the diet of both rural and urban populations. Therefore, it is crucial to promote the sustainability of these agri-food systems by evaluating and improving soil quality. This concept, traditionally referred to as "soil quality," has evolved to include notions of "soil health" and "soil security." These contemporary terms reflect the soil's capacity to support food and fiber production while acting as a vital interface with the environment, ensuring its long-term functionality. Key indicators of this quality include soil texture and organic matter content, factors that directly influence other crucial properties for soil sustainability, health, and security (Lehmann *et al.*, 2020). Determining an organic component, represented by the percentage of organic matter (MO), and conducting a particle size analysis of the mineral fraction to identify four other components: clay (C), fine silt (FS), coarse silt (CS), and sand (S). Together, these elements constitute a complete compositional simplex $S^5$ = {OM, C, FS, CS, S}. Treating these data as compositional data is crucial to avoid methodological biases. Indeed, these five interdependent components define the quality of soils. Any change in one proportion necessarily affects the others (Xu *et al.*, 2017).

This research aims to develop an indicator that encompasses these five components of the simplex $S^5$ using unbiased compositional data analysis tools. The goal is to correlate this indicator with wheat yields in the Fez-Meknes region of Morocco, which is known for its cereal production. It is, therefore, hypothesized that the compositional transformation of the five components into a global index, noted $r^2_{granulo}$ (Khiari *et al*. 2001), is relevant if this global index can be related to cereal yields. This $r^2_{granulo}$ index could help the scoring process assess soil health and cereal suitability.

## MATERIALS AND METHODS

Soil composite samples were collected from 298 locations, from the Fez-Meknes region located in northern Morocco, during the 2021/2022 autumn season. The sampling units within each plot were randomly selected. Soil samples were taken prior to plowing, reaching a depth of 30 cm for conventional tillage. These soil samples were air-dried after manually crushed and sieved through a 2 mm mesh. Organic C was determined by the Walkley-Black procedure (Nelson and Sommers, 1982). Soil texture was determined by the hydrometer method (Day, 1965).

The cutoff values for distinguishing between high and low yields populations were determined using the cumulative variance ratio functions $F_i^C(V_X)$ developed by Khiari et *al.* (2001a). A cubic adjustment was applied to model the evolution of these $F_i^C(V_X)$ in relation to yields. These cubic models identify the inflection point, which serves as the cutoff yield separating the productive sub-population (Pop+) from the (Pop$^-$), the less productive sub-population (Khiari *et al.*, 2001a). The global and individual granulometric indices (including organic matter) were computed using the Excel package (Microsoft, 2013) based on the theory outlined below (Equs1–6). The critical thresholds for these indices were determined using the Cate-Nelson binary partition method (Mathieu *et al.,* 2003). This involved an iterative process to derive sums of square values across multiple divisions at different levels of the independent variable $r^2_{granulo}$. The critical $r^2_{granulo}$ level corresponds to where the sum of squares reaches its maximum. At the same time, the dependent variable Y, is determined where the number of points in the error quadrants (false positive and false negative) is minimized. These binary partition calculations were conducted using R software (R 3.6.2) (Team, R.C, 2013).

## RATIONALE

The organic and inorganic components and the size distribution of the inorganic particles in the soil constitute a complete compositional simplex, $S^5$ = {OM, C, FS, CS, S}, as described by Aitchison (1982). This Simplex consists of five positive components OM> 0, C>0, FS> 0, CS> 0, S> 0), which together total 100%. This structure ensures that each component is proportionally represented, reflecting the comprehensive nature of the soil composition:

$$OM+C+FS+CS+S=100\% \quad (1)$$

Where: $OM$ represents the percentage of organic matter, $C$ represents the percentage of clay, $FS$ represents the percentage of fine silt, $CS$ represents the percentage of coarse silt, $S$ represents the percentage of sand. Each component is expressed as a percentage of the total soil composition.

The particle size fractions become invariant on the scale when divided by their geometric mean (G) defined as follows (Eq. 2):

$$G = [OM \times C \times FS \times CS \times S]^{1/5} \quad (2)$$

$V_{OM}$, $V_C$, $V_{FS}$, $V_{CS}$, and $V_S$ are the log-centered ratios of these five particle size components on their geometric means, respectively (Eq. 3).

$$V_{OM} = \frac{OM(\%)}{G};\ V_C = \frac{C(\%)}{G};\ V_{LS} = \frac{LS(\%)}{G};\ V_{CS} = \frac{CS(\%)}{G};\ V_S = \frac{S(\%)}{G} \quad (3)$$

By definition, the sum of the five particle size components is equal to 100% (Eq.1), and thesum of the log-centered ratios must be zero (Eq.4).

$$V_{OM} + V_C + V_{FS} + V_{CS} + V_S = 0 \quad (4)$$

The first standards for particle size analysis with the aim of integrating them into a global index are the averages ($V^*_{OM}$, $V^*_C$, $V^*_{FS}$, $V^*_{CS}$, and $V^*_S$) and standard deviations ($\sigma_{OM}$, $\sigma_C$, $\sigma_{FS}$, $\sigma_{CS}$, and $\sigma_S$) of the log-centered ratios of the five particle size components. The five log-centered ratios are normalized as follows:

$$I_{OM} = \frac{V_{OM}-V_{OM}^*}{\sigma_{OM}};\ I_C = \frac{V_C-V_C^*}{\sigma_C};\ I_{FS} = \frac{V_{FS}-V_{FS}^*}{\sigma_{FS}};\ I_{CS} = \frac{V_{CS}-V_{CS}^*}{\sigma_{CS}};\ I_S = \frac{V_S-V_S^*}{\sigma_S} \quad (5)$$

Where $I_{OM}$, $I_C$, $I_{FS}$, $I_{CS}$, and $I_S$ are the particle size indices.
The particle size indices defined by Eq.5 are standardized and linearized variables in a five-dimensional space (Aitchison, 1982). A diagnosed sample's single particle size imbalance index is the $r^2_{granulo}$ calculated according to Eq.6.

$$r^2_{granulo} = I^2_{OM} + I^2_C + I^2_{FS} + I^2_{CS} + I^2_S \quad (6)$$

The square sum of these five standardized and independent particle size indices provides a new variable that follows the chi-square ($\chi^2$) distribution law with 5 degrees of freedom (Ross, 1987). As described in Equations (5) and (6), the closer the particle size indices are to zero (as calculated by $r^2_{granulo}$ and the chi-square), the greater the likelihood of achieving a good yield.

## RESULTS AND DISCUSSION

**Step 1.** Selection of Cutoff Yield

The five cumulative variance ratio functions $F_i^C(V_X)$, depending on the yields, all showed sigmoidal forms (Figure 1), which fit the cubic polynomial model in a very highly significant way ($p < 0.01$) with a correlation coefficient $R^2$ of 0.99. The inflection points of these cubic equations, calculated using the formula −b/3a, range from 2 to 24 quintals per hectare. The highest yield of 24 quintals per hectare is used to distinguish between the "pop$^+$" sub-population (considered acceptable yields) and the "pop$^-$" sub-population (considered less acceptable yields). This critical yield of 24 quintals per hectare is significantly higher than the national average of 15 quintals. Notably, wheat production in Morocco for the 2022/2023 agricultural season is estimated at 4.1 million tons. This figure represents a decrease of about 17% compared to the average of the last five years. Still, it is a significant increase of over 60% compared to the 2022 production, which had been severely affected by drought, according to the FAO 2023. Morphological, physiological, and environmental factors all influence wheat production variability, both directly and indirectly.

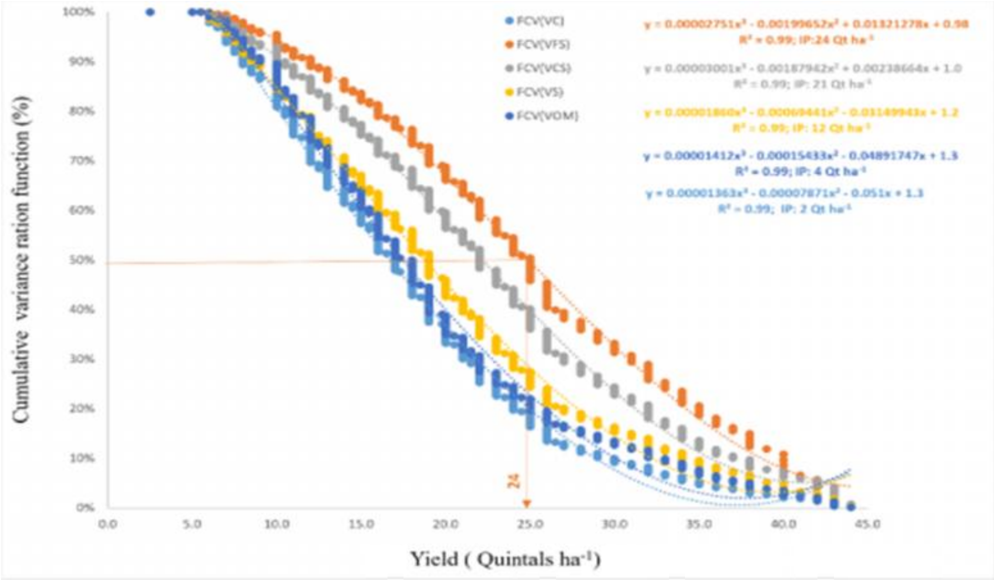


Figure 1. Relationship between wheat yield (Y) and the cumulative variance ratio functions $F_i^C(V_{OM})$, $F_i^C(V_C)$, $F_i^C(V_{FS})$, $F_i^C(V_{CS})$, and $F_i^C(V_S)$ fitted with cubic models of the form: $F_i^C(V_X) = aY^3+bY^2+cY+d$. IP: inflection point. $F_i^C(V_X)$: cumulative variance ratio function.

In this study, out of 298 specimens analyzed, 91 specimens originated from fields with an acceptable yield of 24 quintals per hectare or more. This significant representation underscores the correlation between yield outcomes and the quality of the environmental and genetic factors.

**Step 2.** A theoretical, critical value of the particle size imbalance index
The analysis of the total population of 298 specimens reveals that 91 specimens exceed the yield threshold of 24 quintals per hectare. At the same time, 207 falls below, as established in the first phase of this diagnostic system (Figure 1). Consequently, the proportion of yields considered less acceptable (pop-) is 69% (207/298). According to the Chi-square statistical table or Figure 2, this distribution corresponds to a theoretical critical value 3.1. This figure represents the theoretical maximum of the global particle size index beyond which soils are deemed unsuitable for cereal cultivation because it predicts a yield categorized in the less productive sub-population, pop-. This theoretical value needs validation to diagnose the cereal-growing potential of soils based on their organic matter content and proportions of clay, fine silt, coarse silt, and sand. Integrating these textural and organic components into a single global index (Eq. 7) could be relevant for diagnosing soil suitability for cereal cultivation, especially if this index correlates with actual yields obtained.

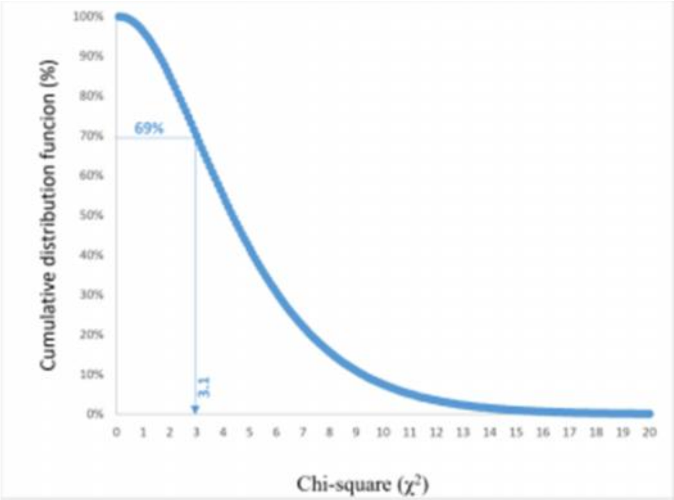


Figure 2. Cumulative Chi-Square distribution function for determining the critical theoretical value of the global particle size indicator with 5 degrees of freedom.

**Step 3.** Validation of the Threshold Particle Size Imbalance Index

This step involves establishing granulometric composition standards (Table 1) by first calculating the average centered logarithmic values for the 91 productive sub-population (pop+) specimens with 24 quintals per hectare or higher yields. Subsequently, the standard deviations of these same centered logarithmic values are determined, but this time considering the entire sampled population of 298 specimens. This approach helps to understand the variations in granulometric composition in relation to observed productivity.

Table 1: Granulometric compositional norms: means and standard deviations of centered logarithmic values of organic matter ($V^*_{OM}$), Clay ($V^*_C$), fine silt ($V^*_{FS}$), coarse silt ($V^*_{CS}$), and sand ($V^*_S$).

| | $V^*_{OM}$ | $V^*_C$ | $V^*_{FS}$ | $V^*_{CS}$ | $V^*_{FS}$ |
|---|---|---|---|---|---|
| **Mean** | - 1.661 | 1.152 | 0.285 | - 0.224 | 0.447 |
| **Standard deviation ($\sigma$)** | 0.266 | 0.676 | 0.284 | 0.273 | 0.803 |

Figure 3 illustrates a noticeable downward trend in wheat yield based on the global granulometric index, $r^2_{granulo}$ (Figure 3a). The Cate-Nelson method enabled partitioning the pop+ and pop- subpopulations using two distinct critical thresholds: 27 quintals per hectare yield and an $r^2_{granulo}$ index of 5.2 (Figure 3). These thresholds differ from the theoretical critical values previously identified, namely a yield of 24 quintals per hectare (Figure 1) and an $r^2_{granulo}$ index of 3.1 (Figure 2). The partitioning by the Cate-Nelson method proved to be more selective regarding yield and less selective concerning the $r^2_{granulo}$ index. The critical values of the $r^2_{granulo}$ index, which distinguish a balanced granulometric composition suitable for cereal cultivation from an unbalanced composition unsuitable for cereals, are set at 3.1 based on the theoretical foundations of this compositional approach as detailed by Khiari et al. (2001a, b), and at 5.2 when this index is actually measured. Based on these two critical thresholds, we can classify the areas studied in the Fez-Meknes region of Morocco into three groups: (i) a category with a very balanced texture and particularly conducive to cereals, characterized by an $r^2_{granulo}$ index less than 3.1, typically associated with good physical fertility and acceptable yields; (ii) a category with a balanced texture where the $r^2_{granulo}$ index ranges between 3.1 and 5.2; and (iii) a category with an unbalanced texture, where the $r^2_{granulo}$ index exceeds 5.2, indicating poor physical fertility and a granulometry inappropriate for cereals, leading to low yields. Among the five performance metrics of the Cate-Nelson partition — robustness, specificity, sensitivity, Negative Predictive Values (NPV), and Positive Predictive Values (PPV) — only the NPVs demonstrated a satisfactory reliability level of 76%. This ratio [TN/(TN + FN)] represents the likelihood that a wheat field is unsuitable for soil with an unbalanced texture when the $r^2_{granulo}$ index exceeds 5.2. Thus, for $r^2$granulo values in group iii (above 5.2), only one error was made for every four diagnoses. This result encourages the exclusion of soils unsuitable for wheat cultivation based on their granulometric and organic composition.However, for groups (i) and (ii), with $r^2_{granulo}$ indices below 3.1 and 5.2, the diagnostic based on a balanced granulometric composition is not as reliable. The probabilities associated with specificity, sensitivity, PPV, and robustness are below 50%. On the other hand, as the NPV approaches closer to 100% and moves away from 0%, it effectively discriminates between soils suitable and less suitable for cereal cultivation. In other words, the diagnostic based on this global granulometric index is only effective when values exceed the critical threshold of 5.2.

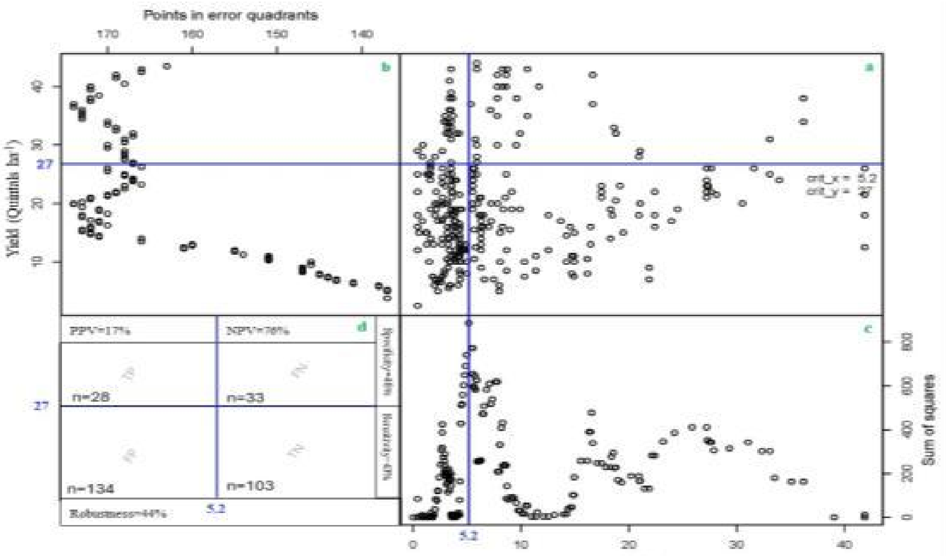


Figure 3. Binary partitioning results according to the Cate-Nelson method for wheat yield based on the global granulometric index $r^2_{granulo}$, showing (a) the distribution of points and identification of two thresholds for yield and $r^2_{granulo}$; (b) the variation in the number of points outside the model for determining the limit yield; (c) the sum of squares for determining the $r^2_{granulo}$ limit; and (d) a summary table (bottom left); performance indicators of the partition model include specificity, sensitivity, accuracy, negative predictive value (NPV), and positive predictive value (PPV). Moreover and according to Šimundic (2009), the NPV is considered a good criterion for discrimination because it falls between 70 and 80%, in a diagnostic performance system that rates results as excellent, very good, good, sufficient, poor, and useless, respectively in the following probability ranges: [90-100%], [80-90%], [70-80%], [60-70%], [50-60%], and [0-50%].

## CONCLUSIONS

Determining soil quality in Morocco for appropriate crop allocation is crucial for optimizing agricultural productivity in a sustainable manner. The most commonly referenced soil quality indicators in international literature are particle size composition and organic matter content. These components form a complete compositional simplex $S^5$ = {OM, clay, fine silt, coarse silt, sand}, leading to the global granulometric index $r^2_{granulo}$. This index defines soil quality and shows a correlation with wheat yields. This study has developed an effective diagnostic system for identifying Fez-Meknes wheat fields unsuitable for cultivation when the $r^2_{granulo}$ index exceeds 5.2. This outcome, demonstrating the discrimination of soils unsuitable for cereal cultivation with a diagnostic performance of 76% based on their granulometric and organic composition, is promising for establishing tailored strategies based on the global $r^2_{granulo}$ index of soils.

**ACKNOWLEDGMENTS**

The authors thank representatives from the Moroccan Ministry of Agriculture and the National School of Agriculture in Meknes and the local farmers from the regions studied for their invaluable support and contributions. The Natural Sciences and Engineering Research Council of Canada (NSERC) covered the cost of training in compositional analysis and modeling techniques under grant number RDCPJ528053-18.

## REFERENCES

Aitchison, J. (1982). The statistical analysis of compositional data: Chapman Hall. J. R. Stat. Soc., 44, 139–177.

Aitchison, J.; Greenacre, M. (2002). Biplots of compositional data. J. R. Stat. Soc. Series C Appl. Stat., 51, 375–392.

Day, P.R. (1965). Particle fractionation and particle-size analysis. p.545-567. ln C.A. Black (ed.) Methods of soil analysis. Part 1. Physical and mineralogical properties. Agron. Monogr. 9. ASA, Madison, WI.

FAO, 2023 : GIEWS Country Brief: Morocco, Postedc26 Jul 2023, published 26 Jul 2023

Harbouze, R., Pellissier, J. P., Rolland, J. P., & Khechimi, W. (2019). Rapport de synthèse sur l'agriculture au Maroc (Doctoral dissertation, CIHEAM-IAMM). https://hal.science/hal-02137637v2

Khiari, L.; Parent, L.-E.; Tremblay, N. (2001a). Selecting the high-yield subpopulation for diagnosing nutrient imbalance in crops. Agron. J., 93, 802–808.

Khiari, L.; Parent, L.-É.; Tremblay, N. (2001b). Critical compositional nutrient indexes for sweet corn at early growth stage. Agron. J., 93, 809–814.

Lehmann J., Bossio DA., Kögel-Knabner I. et al. (2020). The concept and future prospects of soil health. Nat Rev Earth Environ.;1:544–53.

Mathieu, C.; Pieltain, F.; Jeanroy, E. (2003). Analyse Chimique Des Sols: Méthodes Choisies; Lavoisier: Paris, France,.

Nelson, L.; Anderson, R. (1977). Partitioning of soil test—Crop response probability. Soil Test. Correl. Interpret. Anal. Results, 29, 19–38.

Nelson, D.W., and Somners, L.E. (1982). Total carbon, organic carbon, and organic matter. p. 539-579. ln A.L. Page et al. (ed.) Methods.

Ross, S.M. 1987. Introduction to probability and statistics for engineers and scientists. John Wiley & Sons, New York.

Šimundi'c, A.-M. (2009). Measures of diagnostic accuracy: Basic definitions. Ejifcc, 19, 203.

Team, R.C. (2013). R: A Language And Environment For Statistical Computing; Foundation for Statistical Computing: Vienna, Austria,.

Xu, Y., Jimenez, M.A., Parent, S.-É., Leblanc, M., Ziadi, N., Parent, L.E. (2017). Compaction of Coarse-Textured Soils: Balance Models across Mineral and Organic Compositions. Front. Ecol. Evol., 5, 83.