Use of two hybrid algorithms in the investigation and prediction of the values of five quantitative traits of guar beans in different deficit irrigation methods

Seyed Hassan Mirhashemi

Received: 14 July 2022 / Accepted: 17 October 2022 / Published online: 2 November 2022
© The Author(s) 2022

Abstract
Nowadays, deficit irrigation is of particular importance in areas facing the water shortage and drought. This study focused on the investigation and prediction of the values of five quantitative traits of guar beans under different deficit irrigation methods. Deficit irrigation methods were carried out at the initial, development, mid, and late plant growth stages. The experiment was carried out in 25 treatments each with four replications in 2018 and 2019. Initially, the values of five quantitative traits of guar beans were divided into three categories, the values of which were clustered using the K-means algorithm. Then, clusters were predicted using a combination of K-means and CART algorithms. Finally, the relationship between different deficit irrigation methods and clusters was investigated by a combination of K-means and Apriori algorithms. The results of two hybrid algorithms determined that the amount of irrigation in the mid-stage of plant growth significantly affected the five quantitative traits of guar beans. After the mid-stage of the plant growth, the amount of irrigation in the development, initial, and late growth stages had the greatest effect on the quantitative traits of guar beans. Among the deficit irrigation methods, irrigation rates of 60% in the primary stage, 80% in the development stage, 100% in the mid-stage, and 40% in the late stage of the plant growth were the best deficit irrigation methods in the four stages of growth.

Keywords Apriori algorithm · CART algorithm · Dry areas · K-means algorithm · Plant growth stages

Introduction

Today, some parts of the world are facing the water shortage. In recent years, the water shortage has increased in most plains of Iran. Since drought affects the production of crops every year, the use of proper crop management methods, especially deficit irrigation methods, can pave the ground to achieve acceptable yields under water shortage.

Guar (Cyamopsis tetragonoloba L.) is an annual crop and a summer grain with high resistance to the drought (Whistler and Hymowitz 1979). It is a cereal grain crop of spring and summer with resistance to drought (Ashraf et al. 2002, 2005) that is usually cultivated in semiarid regions such as India, Pakistan, the USA, and South Africa. This crop can also be produced in other regions of the world with warm weather and water shortage (Sortino and Gresta 2007; Gresta et al. 2013). Guar can be successfully cultivated in areas where the temperature is > 30 °C in summer (Gresta et al. 2016, 2018; Santonoceto et al. 2019).

The dicotyledonous seed of guar is composed of three main parts: the hull, endosperm, and germ. The germ and hull are used as a meal in the animal and poultry feed, and the endosperm is the site of gum (Lee Jason et al. 2004). Guar seeds contain galactomannan gum that has various industrial applications and form a viscous gel even in cold water (Senapati et al. 2006; Pathak et al. 2010).

Guar was first used as an industrial product and then as fodder due to its high protein (Whistler and Hymowitz 1979). Guar is used as a thickener, gel, viscosifier, and stabilizer in industrial applications owing to its chemical reaction, high flexibility, and good solubility in cold water (Wang and Zhang 2009). It is also used in drilling, color, oil operations, cosmetics, pharmaceuticals, agriculture, civil engineering, foodstuffs, and textiles (Zambrano et al. 2004; Vaughna et al. 2011; Lubbe and Verpoorte 2011). Guar has widely been used in various industrial applications in recent

Seyed Hassan Mirhashemi
Hassan.mirhashemi@yahoo.com

Department of Water Engineering, Faculty of Water and Soil, University of Zabol, Zabol, Iran
years (Das et al. 2011). Most of the guar seeds are produced in India (80%) and Pakistan (15%). Therefore, there is a very good opportunity for the production of guar in the rest of arid and semiarid regions (Pathak and Roy 2015).

In this research, data mining algorithms were used to predict the effect of different methods of deficit irrigation on the quantitative traits of guar beans. Data mining is a process that automatically finds valuable information from data stored in large warehouses (Tan and Yu 2006). Some researchers now use data mining in different fields such as agriculture, healthcare, banking, sports, and hydrology. The tree data mining algorithm was used previously to predict variations in the depth of the aquifer (Mirhashemi et al. 2020).

In addition, some researchers have used data mining algorithm to extract association rules, which were used in India to link floods and droughts to climatic data and yielded strong relationships (Dhanya and Nagesh Kumar 2009). In a study by Mirhashemi and Mirzaei (2021), the Apriori algorithm was used to extract association rules, which were then employed to study the relationship between changes in the aquifer depth and the affecting factors.

In most plains of Iran, rainfall is usually lower in spring and summer than in the other seasons; thus, plants grown in these two seasons may be more exposed to water deficit. In this study, therefore, quantitative traits of guar were investigated in different deficit irrigation methods in spring and summer. This research aimed to predict quantitative traits of guar in different deficit irrigation methods and to investigate the relationships between different deficit irrigation methods and quantitative traits of guar by hybrid algorithms.

**Materials and methods**

The effects of different methods of deficit irrigation were investigated and predicted on grain yield, the number of pods m$^{-2}$, thousand-seed weight (TSW), harvest index (HI), and biomass of guar. The deficit irrigation methods were carried out in four initial, development, mid, and late stages of plant growth. The experiment was conducted in a farm in Alborz province (Iran) in spring and summer 2018 and 2019. The geographic coordinates of the experimental site are 35.8400° N, 50.9391° E with an altitude of 1.300 m from the sea level. The soil chemical and physical properties in the experimental site are shown in Table 1. According to the richness of the soil in terms of phosphorus and potassium, only 100 kg ha$^{-1}$ of nitrogen (from a source of urea) was added to the soil in the development and mid-stages of plant growth. The environmental conditions of farm in four stages of growth are summarized in Table 2. There was no effective rainfall during the experiment.

![Image](image-url)

**Table 1** Physical and chemical properties of the soil at experimental site

| Properties       | Unit   | Value |
|------------------|--------|-------|
| **Chemical**     |        |       |
| P                | mg kg$^{-1}$ | 0.15  |
| K                | mg kg$^{-1}$ | 80    |
| Total N          | %      | 420   |
| Organic carbon   | %      | 1.32  |
| Organic matter   | %      | 2.25  |
| **Physical**     |        |       |
| EC               | dS m$^{-1}$ | 1.17  |
| Silt             | %      | 28    |
| Clay             | %      | 12    |
| Sand             | %      | 60    |
| pH               |        | 7.40  |

**Table 2** The average amount of environmental conditions in the four stages of plant growth

| Environmental conditions | Initial | Development | Mid | Late |
|--------------------------|---------|-------------|-----|------|
| Air temperature (°C)     | 28      | 30          | 30  | 28   |
| Humidity (%)             | 20      | 18          | 15  | 20   |
| Sunny Hours (h)          | 10      | 11.5        | 11.5| 10   |

The amount of irrigation water ($dj$) was obtained for the plant by obtaining the field capacity (F.C.) for the soil and soil moisture at the time of irrigation (Pv) using a time-domain TDR (time-domain reflectometry) installed in soil farm and obtaining the root development depth ($z$) in different growth stages. Accordingly, irrigation water levels of 80, 60, and 40% were defined for the other irrigation levels in normal conditions. The TDR was also used to determine suitable irrigation intervals. In this method, the device was installed in different depths of root development at each stage of plant growth. By obtaining F.C., permanent wilting point (P.W.P), and the readily available coefficient for guar beans ($f$), the percentage of available water was calculated to obtain the irrigation point. Irrigation operations were done based on defined treatments when soil moisture reached the irrigation point according to the TDR measurements. The surface drip system was used for irrigation.

$$dj = (F.C - P.v) \times z$$

$$P.v = (F.C - P.W.P) \times f$$

Each experimental plot consisting of four planting lines with a length of three meters and a distance of 50 cm planting rows was selected in four replications. The distance between plants was 10 cm (Pathak and Roy 2015).
cultivation began on the first day of June (06/01), and the harvest was started in September 20 (09/20). Approximate intervals of 20, 30, 40, and 20 days were considered for the initial, development, mid, and late growth stages of guar beans, respectively.

The conditions of irrigation were different in each treatment (except the control treatment). Table 3 represents the naming of irrigation values in all four stages of plant growth.

Table 3 The naming of the irrigation values at the different growth stages of the guar bean plant

| The naming of irrigation amount | Initial stage | Development stage | Mid-stage | Late stage |
|--------------------------------|---------------|-------------------|-----------|-----------|
| 40% of full irrigation        | A             | B                 | C         | D         |
| 60% of full irrigation        |               |                   |           |           |
| 80% of full irrigation        |               |                   |           |           |
| Full irrigation               |               |                   |           |           |

The naming of irrigation in four stages of growth

| Initial stage | Development stage | Mid-stage | Late stage |
|---------------|-------------------|-----------|-----------|
| I1            | I2                | I3        | I4        |

Table 4 The names of different treatments related to the amounts of irrigation in the four stages of plant growth

| The names of different treatments |
|----------------------------------|
| ABCD                             |
| ADCB                             |
| BDAC                             |
| CBDA                             |
| DBAC                             |
| ABDC                             |
| BACD                             |
| BDCA                             |
| CDAB                             |
| DCBA                             |
| ACBD                             |
| BADC                             |
| CDBA                             |
| DCAB                             |
| ADDB                             |
| BCAD                             |
| CBDA                             |
| DACB                             |
| DDDD                             |

According to the irrigation treatments (Table 4), irrigation was different in each treatment (except the control treatment) in each stage of plant growth. In the treatment of ABCD, for instance, irrigation amounts of A, B, C, and D were done at the initial, development, mid, and late stages of plant growth, respectively.

The total dry weight of the plant was measured during the harvest phase of the guar plant. To measure the biomass of guar, the aerial parts were bottom cut and incubated in an oven at 80 °C for 48 h. After drying, the total dry weight of the plant was determined by a precision scale. To measure the number of pods per plant after physiological ripeness, the total number of pods per plant and the mean of pods per plant were counted in 20 randomly selected plants. To measure the number of seeds per pod, the seeds were separated from pods, all seeds were counted, and their mean was calculated as the number of seeds per pod. The grains obtained from the grain yield were randomly weighed by a precise scale over four replication of 1000 grains and the average was calculated as the TSW for each treatment. The HI for each treatment was obtained via dividing the grain yield by the biological yield.

The K-means algorithm was used to cluster five quantitative traits of beans. To better understand the results of the K-means algorithm, all relevant data about the amount of five traits of guar beans were further categorized into three low, medium, and high categories (Table 5). The optimal number of clusters was determined using the Silhouette index method.

### Silhouette validity index (SVI)

The SVI method is defined based on the average distance of each sample of a cluster with all samples in the same cluster and the average distance of the sample in other clusters with a given cluster (Eq. 3). Based on this view, the maximum amount of dispersion and correlation of data is determined for each cluster, and the maximum values of this index are used to determine the optimal number of clusters (Rousseeuw 1987).

\[ S(i) = \frac{(b(i) - a(i))}{\max\{a(i), b(i)\}} \]  

where \(a(i)\) is the mean distance between the ith sample and all the samples in \(X_i\), \(b(i)\) is the minimum mean distance between the sample and all clustered samples in \(X_k\) (\(k = 1, ..., c: k \neq j\)). The value of SVI varies between −1 and 1. An SVI value of 1 represents that clustering is done properly. If the SVI value is close to zero, it means that the sample can be attributed to a closer cluster and the sample position is equally distant from two clusters. An SVI value equal to −1 means an inappropriate clustering.

Table 5 Classification of five traits of guar beans

|                | Low          | Medium       | Much         |
|----------------|--------------|--------------|--------------|
| Grain yield (kg/ha) | 1300 ≤ … < 1500 | 1500 ≤ … < 1700 | 1700 ≤ … < 1900 |
| Pods number     | 300 ≤ … < 400 | 400 ≤ … < 500 | 500 ≤ … < 600 |
| 1000 seed weight (g) | 20 ≤ … < 25  | 25 ≤ … < 30  | 30 ≤ … < 35  |
| Harvest index   | 15 ≤ … < 20  | 20 ≤ … < 25  | 25 ≤ … < 30  |
| Biomass (kg/ha) | 2000 ≤ … < 4000 | 4000 ≤ … < 6000 | 6000 ≤ … < 8000 |
**Decision tree algorithm**

In addition to quantitative variables, the decision tree algorithm can predict qualitative variables. This method was first introduced by Breiman et al. (1984). The decision tree algorithm tries to minimize the diversity or variety (in terms of the target attribute) in the nodes. This lack of uniformity in the nodes is measurable using the impurity measures, and the Gini index is the most important and widely used one (Yoneyama et al. 2002). The difference of various kinds of decision trees is in the impurity measure, the splitting method, and pruning of tree nodes. In this study, the decision tree model was used to investigate and predict the five quantitative traits of guar beans. Therefore, the results of the three types of tree algorithms (CART, CHAID, and C5.0) were evaluated to predict the results of cluster classification of quantitative traits of guar beans.

**CART (classification and regression tree) algorithm**

The CART is a binary recursive tree that exactly splits the parent nodes into two offspring groups and continues recursive splitting until another branch cannot be built. In recursive splitting of data into subgroups, it supports only two subgroups and is a fast algorithm with backward pruning (Chattamvelli 2011).

**CHAID (Chi-squared automatic interaction detector) algorithm**

The CHAID algorithm was presented by Morgan and Sonquist (1963). The algorithm was promoted to an algorithm for analyzing nominal-scale dependent variables by Morgan and Messenger (1973) and was converted to the CHAID algorithm by Kass (1980).

**C 5.0 algorithm**

This algorithm is a single-variate decision tree and the improved form of the C4.5 algorithm. Similar to CART, this algorithm first creates an almost full tree, but its pruning strategy is quite different. It performs the classification by splitting data into subsets containing more uniform records than their parents (Chattamvelli 2011). In this study, five clusters as dependent variables (objective function) and deficit irrigation methods were introduced as independent variables to tree algorithms.

**Association rule algorithm**

Agrawal et al. (1993) used the association rules to determine the important and hidden affiliations of a database. These association rules predict the occurrence of an object based on the occurrence of other objects. Therefore, the algorithms of finding association rules find all possible association rules within a database. A set of “If–then” rules, which expresses the connections between the simultaneous occurrence of a set of objects, is an important outcome of this approach. This algorithm expands the set of rules obtained from data and extracts the rules with the highest level of information.

**Apriori model**

This algorithm was first introduced by Agrawal and Srikant (1994) and is one of the most important findings in the mining of association dependency rules (Zhao and Bhownick 2003). This algorithm requires the classification of all input and output fields as it is optimized for this type of data (Cheung et al. 1996).

**Introduction of three basic terms of association rules**

**Confidence**

This coefficient determines the accuracy of a rule and suggests that if the exchange controls the antecedent conditions, then it will also obtain the consequent conditions. This measure was introduced by Agrawal et al. (1993).

**Support**

This parameter expresses the possible coexistence of A and B in the A → B rule. This measure was also introduced by Agrawal et al. (1993).

**Lift**

This is a criterion for evaluating the correlation value of two data sets. If the amount of lift is lower than one, the two sets have a negative correlation; if it is greater than one, the two data sets have a positive correlation; and if it is equal to one, the two sets have no correlation and are independent. This is one of the most commonly used measures of fitting rules introduced by Brin et al. (1997).

In this study, deficit irrigation methods and the types of clusters were selected as the antecedent and consequent sets.

**Results and discussion**

First, the data about five quantitative traits of the guar plant in both 2018 and 2019 were analyzed by IBM SPSS Statistics. According to Tables 6 and 7, the quantitative traits of guar beans in different deficit irrigation methods were significant at a 1% level. Figures 1, 2 and 3 show the average
Table 6 The ANOVA results of five quantitative traits of guar beans in various deficit irrigation methods in 2018

| S.O.V    | df | Biomass of guar | Harvest index | 1000 seed weight | Pods number | Seed yield |
|----------|----|-----------------|---------------|------------------|-------------|------------|
| Repetition | 3  | 1,044,694.5     | 36.65         | 5.78             | 1024.33     | 7.30       |
| Irrigation | 24 | 7,387,975**     | 25.48**       | 26.47**          | 50,673.4**  | 85,535.3** |
| Error     | 72 | 21,378.67       | 0.166         | 0.026            | 11.3        | 15.3       |
| CV %      | –  | 23.7            | 11.5          | 8.6              | 20          | 8.7        |

*Significant at the probability level of 0.01

Table 7 The ANOVA results of five quantitative traits of guar beans in different deficit irrigation methods in 2019

| S.O.V    | df | Biomass of guar | Harvest index | 1000 seed weight | Pods number | Seed yield |
|----------|----|-----------------|---------------|------------------|-------------|------------|
| Repetition | 3  | 1,048,152.7     | 31.95         | 5.55             | 1102.86     | 11.66      |
| Irrigation | 24 | 7,380,512.9**   | 26.4**        | 25.15**          | 50,042.8**  | 85,997.6** |
| Error     | 72 | 21,723.829      | 0.510         | 0.353            | 22.846      | 22.927     |
| CV %      | –  | 24.2            | 11.8          | 8.9              | 20          | 9          |

*Significant at the probability level of 0.01

Fig. 1 Mean values of changes in the number of pods m⁻² and seed yield under different deficit irrigation methods in 2018 and 2019

Fig. 2 Mean values of changes in the harvest index and the thousand-seed weight in various deficit irrigation methods in 2018 and 2019
amount of changes in five quantitative traits of guar beans in the different deficit irrigation methods in both 2018 and 2019.

Figures 1, 2 and 3 show the values of quantitative traits of guar beans under various deficit irrigation methods.

According to these figures and the coefficient of variation (CV) in Tables 6 and 7, the biomass of guar beans and the TSW of guar beans had the highest and the lowest CVs, respectively, relative to the other quantitative traits of guar beans in different deficit irrigation methods. The K-means algorithm was used to cluster the data related to quantitative traits of guar beans. At each run of the K-means algorithm, the clustering results were evaluated by the SVI. In the end, five clusters were achieved as the most optimal number of clustering with the highest SVI value of 0.90 (Fig. 4).

Five clusters were formed for five quantitative traits of guar beans using the K-means algorithm. The highest contribution percentage of quantitative traits classified in cluster formation is shown in Fig. 5. In the first cluster, for example, the highest contribution of the five quantitative traits belongs to the C category where the contribution of TSW, biomass of guar, and the number of pods was equal to 100%, with 92 and 72% for the HI and the yield, respectively. In the first cluster, the C category showed the most contribution of quantitative traits of guar beans, suggesting that this cluster possesses the greatest values of quantitative traits compared to the other clusters.

In the third cluster, the A category has the utmost contribution of quantitative traits of guar beans, indicating that this cluster contains the least values of quantitative traits relative to the other clusters. All quantitative traits of the second cluster, except the grain yield, are in the B category, implying that this cluster is at a medium level in terms of quantitative traits. The fourth cluster is superior to the fifth cluster in terms of the values of traits and includes B and C clusters. In this research, the five formed clusters were arbitrarily named based on the quantitative traits. Accordingly, the first, fourth, second, fifth, and third clusters have very good, good, medium, bad, and very bad values of quantitative traits, respectively. According to the size of five clusters in Fig. 6, the highest and the least sizes belong to the first (36%) and fourth (7%) clusters, respectively, among the other ones.

In the next step, the clusters in different deficit irrigation methods were predicted using a combination of tree and clustering algorithms. In this way, the results of the clustering algorithm and deficit irrigation methods as input data of the tree algorithm. Three tree algorithms, including CHAID, CART, and c5, were used for this stage. The results of performance evaluation of tree algorithms in the prediction of clusters (Table 8) indicate that the CART algorithm has the best performance in the prediction of clusters.

The CART algorithm predicted five clusters obtained from the results of the K-means algorithm (Fig. 7). The data about each branch were discussed separately to better understand the tree diagram. The tree diagram has two main branches. The third cluster and the other clusters are
predicted by the first main branch and the second main branch, respectively. At the beginning of each branch, irrigation is located at the mid-stage of plant growth (I3), which indicates the great importance of irrigation in the third stage to predict clusters.

**First main branch**

**Deficit irrigation (A) in the mid-stage of plant growth (I3)**

The amount of irrigation was as high as 40% of full irrigation (A). The third cluster is the predicted one.

To further understand this branch, we used the data of frequencies and surrogates, which are not shown in the tree diagram (it is in the lower part of the software) (Table 9). The lower part of the software contains the surrogates, frequencies,
and the history of each branch. Frequency is active in discrete target data and represents the percentage of each discrete value of the objective function in the target branch so that the dominant frequency is introduced as the estimated value of that branch. The surrogates represent other rules governing the branches of interest depending on their priority. This priority indicates the ability of the rule to make purity in sub-branches.

Frequencies refer to more data about the leaves (results), and surrogates are about data of the branches (rules).

According to Table 9, the first main branch predicts the fifth cluster, in addition to the third cluster, but with a very lower likelihood than the third cluster. Table 9 also reveals that in addition to irrigation (A) at the stage (I3), the C and B irrigation in the stages of I1 and I2 also affects the prediction of the third cluster. The amount of irrigation decreases from the I1 to I3 stages, and the plant experiences more water stress at each stage than the previous one. The third and fifth clusters have the lowest values of quantitative traits (very bad and bad values) compared to the other clusters, which demonstrate the large negative effect of deficit irrigation (A) of the quantitative traits in the mid-stage of plant growth.

The second main branch

Deficit irrigation (B) in the mid-stage of plant growth (I3)

In the second main branches, the second and fifth clusters are predicted by the tree diagram when deficit irrigation is done in the I3 stage. So when the deficit irrigation B and A, respectively, occur in the I3 of I1 stages, the second cluster, but if irrigation D is done at the I1 stage, the fifth cluster will be predicted by the tree diagram. As there were no surrogates, only the frequencies data were used in these two branches. According to frequencies data, the second and fifth clusters in these two branches are predicted by different percentages.

The second cluster has a better condition with medium quantitative traits than the fifth cluster with slightly bad traits. However, the branch that predicted the second cluster has less irrigation level in the initial growth stage than the branch that predicted the fifth cluster. Considering that four
different irrigation methods were carried out in different stages of plant growth, more irrigation in the development or the late stage of plant growth would likely have a greater positive effect on the quantitative traits of guar beans than the initial growth stage (Table 10).

The third deficit irrigation (C) in the mid-stage of plant growth (I3)

In the next branch, when deficit irrigation C was done in the I3 phase, clusters 1, 2, and 4 were predicted by a tree diagram.

In the first case, when deficit irrigation occurs in C and A cases in the mid- and development stages of plant growth,

respectively, the second cluster is predicted by the tree diagram. Frequencies and surrogates data were used for further information about this branch.

Given that full irrigation (D) was done in the I1 stage, it had no effect on the selection of very good and good (first and fifth) clusters, and deficit irrigation A at the I2 stage had a greater effect on the reduced values of quantitative traits and the choice of the second cluster (medium values of quantitative traits) (Table 11).

In the second case, when deficit irrigation C and B are applied in the I3 and I2 stages, respectively, the fourth cluster is predicted by the tree diagram. The irrigation in the I2 stage in this branch has increased compared to the previous branch and is placed in case B. The predicted cluster is also in a better position than the previous branch cluster. According to the rules of Table 12, the variation of irrigation amounts A and D in the stages of I1 and I4 does not affect the changes in cluster prediction.

In the third case, when the deficit irrigation of C and D is performed in the I3 and I2 stages, respectively, the first and fourth clusters are predicted at 60 and 40%, respectively, by the tree diagram. The very good and good conditions of the first and the fourth clusters, respectively, indicate the positive effect of additional irrigation in the I2 stage compared

| Cluster name | Prediction percentage | Primary | I3 in D |
|--------------|-----------------------|---------|---------|
| 1            | 90                    | 1       | I2 in C |
| 4            | 10                    | 2       | I1 in B |
|              |                       | 3       | I4 in A |

Table 14 Frequencies and surrogates data of D in I3

Table 15 The rules extracted from the first cluster as the consequent and the amount of irrigation in different stages of growth as the antecedent

| No. | Antecedent                  | Confidence % | Support % | Lift   |
|-----|------------------------------|--------------|-----------|--------|
| 1   | I2 = B and I3 = D           | 100          | 8         | 2.78   |
| 2   | I1 = C and I3 = D           | 100          | 8         | 2.78   |
| 3   | I4 = B and I3 = D           | 100          | 8         | 2.78   |
| 4   | I1 = A and I3 = D           | 100          | 8         | 2.78   |
| 5   | I4 = A and I3 = D           | 100          | 8         | 2.78   |
| 6   | I2 = B and I1 = A and I3 = D| 100          | 4         | 2.78   |
| 7   | I1 = C and I4 = B and I3 = D| 100          | 4         | 2.78   |
| 8   | I2 = C and I1 = B and I3 = D| 100          | 4         | 2.78   |
| 9   | I2 = A and I1 = C and I4 = B and I3 = D | 100 | 4 | 2.78 |
| 10  | I4 = C and I2 = B and I1 = A and I3 = D | 100 | 4 | 2.78 |
| 11  | I2 = B and I1 = C and I4 = A and I3 = D | 100 | 4 | 2.78 |
| 12  | I4 = B and I1 = A and I2 = C and I3 = D | 100 | 4 | 2.78 |
| 13  | I4 = A and I2 = C and I1 = B and I3 = D | 100 | 4 | 2.78 |
| 14  | I4 = D and I3 = D and I2 = D and I1 = D | 100 | 4 | 2.78 |
| 15  | I3 = D                     | 98           | 28        | 2.73   |
| 16  | I2 = A and I3 = D           | 94           | 8         | 2.60   |
| 17  | I4 = C and I3 = D           | 94           | 8         | 2.60   |
| 18  | I1 = B and I3 = D           | 94           | 8         | 2.60   |
| 19  | I4 = A and I1 = B           | 88           | 8         | 2.43   |
| 20  | I2 = A and I4 = C and I1 = B| 88           | 4         | 2.43   |
| 21  | I4 = B and I1 = A           | 81           | 8         | 2.26   |
| 22  | I2 = B and I4 = A           | 75           | 4         | 2.08   |
| 23  | I2 = B and I1 = A           | 63           | 8         | 1.73   |
| 24  | I4 = A                     | 54           | 24        | 1.50   |
| 25  | I2 = C and I1 = B           | 50           | 8         | 1.38   |
to I1 and I4 stages. Although full irrigation in the previous branch was done in the initial and late stages, the first cluster (very good) is not predicted (Table 13).

**Full irrigation (D) in the mid-stage of plant growth (I3)**

Finally, when the amount of irrigation D is at the I3 stage, the tree diagram predicts the first cluster. According to Table 14, full irrigation in the mid-stage and irrigation C at the development stage of plant growth had significant effects in the prediction of very good and good clusters. Table 14 also indicates that the best-deficit irrigation methods should be applied in the initial, development, mid, and late stages of plant growth in the order of B, C, D, and A, respectively, for large amounts of quantitative traits.

---

Table 16  Rules extracted between the second cluster as the consequent and the amount of irrigation in different stages of growth as the antecedent

| No. | Antecedent                                                                 | Confidence % | Support % | Lift |
|-----|---------------------------------------------------------------------------|--------------|-----------|------|
| 1   | I3 = C and I2 = A and I1 = B and I4 = D                                   | 88           | 4         | 3.98 |
| 2   | I3 = C and I2 = A                                                         | 81           | 8         | 3.69 |
| 3   | I3 = B and I4 = C and I1 = A                                              | 63           | 4         | 2.84 |
| 4   | I4 = C and I1 = A and I2 = D                                              | 63           | 4         | 2.84 |
| 5   | I1 = A and I2 = C and I4 = D                                              | 63           | 4         | 2.84 |
| 6   | I4 = A and I2 = C and I1 = D                                              | 63           | 4         | 2.84 |
| 7   | I3 = B and I1 = A and I2 = C and I4 = D                                   | 63           | 4         | 2.84 |
| 8   | I3 = B                                                                    | 58           | 24        | 2.65 |
| 9   | I3 = B and I2 = A and I1 = D                                              | 50           | 4         | 2.27 |
| 10  | I3 = B and I2 = A and I4 = C and I1 = D                                   | 50           | 4         | 2.27 |
| 11  | I1 = B and I4 = D                                                         | 44           | 8         | 1.99 |
| 12  | I2 = A                                                                    | 42           | 24        | 1.89 |
| 13  | I3 = C and I4 = B                                                         | 38           | 8         | 1.70 |
| 14  | I3 = C and I1 = D                                                         | 38           | 8         | 1.70 |
| 15  | I3 = B and I2 = A and I1 = C and I4 = D                                   | 38           | 4         | 1.70 |
| 16  | I3 = C and I2 = B and I1 = A and I4 = D                                   | 38           | 4         | 1.70 |
| 17  | I3 = C                                                                    | 33           | 24        | 1.52 |
| 18  | I4 = A and I1 = D                                                         | 31           | 8         | 1.42 |
| 19  | I2 = C and I4 = D                                                         | 31           | 8         | 1.42 |
| 20  | I2 = C and I1 = D                                                         | 31           | 8         | 1.42 |
| 21  | I1 = A                                                                    | 27           | 24        | 1.23 |
| 22  | I1 = D                                                                    | 27           | 28        | 1.22 |
| 23  | I2 = A and I4 = C                                                         | 25           | 8         | 1.14 |
| 24  | I4 = A                                                                    | 23           | 24        | 1.04 |
Table 17  Rules extracted between the third cluster as the consequent and irrigation amounts in different stages of growth as the antecedent

| No. | Antecedent                                                                 | Confidence % | Support % | Lift |
|-----|-----------------------------------------------------------------------------|--------------|-----------|------|
| 1   | I3 = A and I2 = B                                                           | 100          | 8         | 4.8  |
| 2   | I3 = A and I4 = D                                                           | 100          | 8         | 4.8  |
| 3   | I3 = A and I4 = C and I2 = B and I1 = D                                     | 100          | 4         | 4.8  |
| 4   | I3 = A and I2 = B and I1 = C and I4 = D                                     | 100          | 4         | 4.8  |
| 5   | I3 = A and I2 = C and I1 = B and I4 = D                                     | 100          | 4         | 4.8  |
| 6   | I3 = A and I2 = C                                                           | 94           | 8         | 4.5  |
| 7   | I3 = A and I1 = D                                                           | 94           | 8         | 4.5  |
| 8   | I3 = A and I4 = C                                                           | 88           | 8         | 4.2  |
| 9   | I3 = A and I1 = B                                                           | 88           | 8         | 4.2  |
| 10  | I3 = A and I4 = B and I2 = C and I1 = D                                     | 88           | 4         | 4.2  |
| 11  | I3 = A                                                                     | 85           | 24        | 4.1  |
| 12  | I3 = A and I1 = C                                                           | 75           | 8         | 3.6  |
| 13  | I3 = A and I4 = C and I1 = B and I2 = D                                     | 75           | 4         | 3.6  |
| 14  | I3 = A and I4 = B                                                           | 69           | 8         | 3.3  |
| 15  | I3 = A and I2 = D                                                           | 63           | 8         | 3.0  |
| 16  | I3 = A and I1 = C and I4 = B and I2 = D                                     | 50           | 4         | 2.4  |
| 17  | I4 = B and I2 = C                                                           | 44           | 8         | 2.1  |
| 18  | I4 = B and I1 = D                                                           | 44           | 8         | 2.1  |
| 19  | I2 = B                                                                     | 33           | 24        | 1.6  |
| 20  | I2 = C                                                                     | 31           | 24        | 1.5  |
| 21  | I4 = D                                                                     | 30           | 24        | 1.45 |
| 22  | I1 = C                                                                     | 27           | 24        | 1.3  |
| 23  | I1 = C and I2 = D                                                           | 25           | 8         | 1.2  |
| 24  | I4 = B and I2 = D                                                           | 25           | 8         | 1.2  |
| 25  | I4 = B                                                                     | 23           | 24        | 1.1  |

Figure 8 shows the importance of irrigation in the four stages of plant growth in predicting the five clusters. Accordingly, irrigation in the I3 stage with a relatively high difference from the other irrigation phases has the greatest importance in predicting clusters.

Results of the Apriori and K-means hybrid algorithm

To extract the association rules, it is necessary to introduce the antecedent and consequent data to the Apriori algorithm. For this purpose, clusters and the amount of irrigation in four stages of plant growth were introduced as the consequent and the antecedent, respectively, to the Apriori algorithm. The rules with the highest validity relative to the other rules were used among the extracted rules (Tables 15, 16, 17, 18, 19). The rules are sorted in descending order of confidence percentage, decreasing from the first to the last rule. In Table 15, the rules extracted by the Apriori algorithm are related to the first cluster.

Most of the rules in Table 15 have a 100% confidence interval that specifies the high validity of the rules. For example, the first rule can be defined in such a way that the values of quantitative traits of guar beans belonged to the first cluster in 100 cases where irrigation amounts B and D were applied in I2 and I3, respectively.

Furthermore, a support coefficient of 8% reveals that irrigation amounts B and D in the I2 and I3 stages, respectively, occurred at the same time with the first cluster in 8% of the total irrigation factors influencing the values of the first cluster. Given that the support value is the percentage of simultaneous occurrence of consequent and antecedent factors, lower values of antecedent factors result in a higher percentage of support.

The positive value of lift also marks a positive correlation between antecedent and consequent factors. Therefore, there is a positive correlation between consequent and antecedent factors in all rules in Table 15.

In rule 14, full irrigation (D) is done in all stages of the plant growth; hence, it has 100% confidence associated with the first cluster (very good values of quantitative traits). The values of the three assessment indices in rules 9–13 are equal to those of rule 14. Therefore, the irrigation rules of
9–13 can be applied to achieve very good values of quantitative traits as in rule 14.

Rule 15 consists of an antecedent factor (the amount of irrigation D in the I3 stage) and the maximum support value (28%). In addition to rule 15, the amount of irrigation D is carried out at the I3 stage in all rules of Table 2, which further confirms the great correlation of irrigation D in the I3 stage with very good values of quantitative traits.

In Table 16, the confidence level of the first rule is greater than the other rules. In this rule, irrigation amounts B, A, C, and D were done in the stages of I1, I2, I3, and I4, respectively. Considering that the second cluster has moderate values of quantitative traits and the duration of irrigation A is 30 days in the I2 stage, it is appropriate in terms of water saving and the increase in the values of quantitative traits.

The support value is 24% in all rules involving an antecedent factor, except for rule 23 with the maximum support of 28%. In this rule, irrigation D was done in the I1 stage. The support coefficient states that irrigation D in the I1 stage occurred simultaneously with the second cluster in 28% of the total set of irrigation values affecting the second cluster.

In Table 16, the I3 stage only includes the irrigation amounts B and C, further demonstrating the correlation of almost medium (B and C) irrigation in the I3 phase with the second cluster (medium).

As shown in Table 17, irrigation A was carried out in the I3 phase in five rules with a 100% confidence value. Given that the values of quantitative traits in the third cluster are in a very bad condition, irrigation A in the mid-stage had a significant effect on reducing the values of quantitative traits in guar beans. Although full irrigation (D) was carried out in I1 and I4 stages in these five rules, it did not affect the increase in the values of quantitative traits. Therefore, irrigation A should not be performed at the I3 stage due to its high reducing effect on the quantitative traits of guar beans.

In addition, among the rules in which antecedent factors contain one element, rules 11 and 19, respectively, have the highest percentages of confidence and lift. In these rules, the antecedent factor of irrigation amounts A and B occurred at the I3 and I2 stages, respectively. These rules demonstrate the importance of irrigation amounts A and B in the I3 and I2 stages, respectively, in the values of quantitative traits in guar beans.

In Table 18, the highest percentages of confidence and lift belong to the first rule. Given that the fourth cluster has good values of quantitative traits, the first rule irrigation

| No. | Antecedent                  | Confidence % | Support % | Lift  |
|-----|-----------------------------|--------------|-----------|-------|
| 1   | I3 = C and I2 = A and I4 = A and I1 = D | 50           | 4         | 7.14  |
| 2   | I3 = C and I2 = B           | 38           | 8         | 5.36  |
| 3   | I3 = C and I4 = A           | 38           | 8         | 5.36  |
| 4   | I3 = C and I1 = D           | 38           | 8         | 5.36  |
| 5   | I3 = C and I4 = B and I1 = A and I2 = D | 38       | 4         | 5.36  |
| 6   | I3 = C and I4 = B           | 31           | 8         | 4.46  |
| 7   | I3 = C and I1 = A           | 31           | 8         | 4.46  |
| 8   | I3 = C and I2 = D           | 31           | 8         | 4.46  |
| 9   | I3 = C                      | 27           | 24        | 3.87  |
| 10  | I2 = B and I4 = A           | 25           | 8         | 3.57  |
| 11  | I2 = B and I1 = D           | 25           | 8         | 3.57  |
| 12  | I4 = A and I1 = D           | 25           | 8         | 3.57  |
| 13  | I3 = C and I2 = A and I4 = B and I1 = D | 25      | 4         | 3.57  |
| 14  | I3 = C and I2 = B and I1 = A and I4 = D | 25      | 4         | 3.57  |
| 15  | I3 = C and I4 = A and I1 = B and I2 = D | 25      | 4         | 3.57  |
| 16  | I4 = B and I1 = A           | 19           | 8         | 2.68  |
| 17  | I4 = B and I2 = D           | 19           | 8         | 2.68  |
| 18  | I1 = A and I2 = D           | 19           | 8         | 2.68  |
| 19  | I2 = B                      | 13           | 24        | 1.79  |
| 20  | I4 = A                      | 13           | 24        | 1.79  |
| 21  | I3 = C and I2 = A           | 13           | 8         | 1.79  |
| 22  | I3 = C and I1 = B           | 13           | 8         | 1.79  |
| 23  | I4 = A and I2 = D           | 13           | 8         | 1.79  |
| 24  | I1 = B and I2 = D           | 13           | 8         | 1.79  |
| 25  | I1 = A                      | 10           | 24        | 1.48  |
pattern is suggested for regions with deficient in water (when the goal is the fourth cluster). Irrigation D should be done because the water demand in the plant and the duration of irrigation are not high in the I1 stage. Irrigation B and C should, respectively, be done at the developmental and mid-stages of plant growth when the duration of the plant growth is longer than the other growth stages, and irrigation A is appropriate in the late stage of growth.

In all the rules of Table 18, irrigation C was carried out at the I3 stage, indicating a high correlation between irrigation C in the I3 phase and the fourth cluster. The lowest percentages of confidence and lift are related to the last rule; thus, irrigation A in the I1 stage is not highly correlated with the fourth cluster (good values of quantitative traits).

The values of the evaluation index are equal in the fourth and fifth rules, but irrigation amounts C and D are different in the I4 phase of the two rules. Thus, irrigation amounts C or D in the I4 stage and irrigation B at the I3 stage have the same relationship with the fifth cluster. In terms of water saving, therefore, it is best to do irrigation C instead of irrigation D in the fourth stage after applying irrigation B at the I3 stage (when the goal is the fifth cluster).

Irrigation B was done in all the rules where the antecedent elements involve the I3 phase. In this table, therefore, Irrigation B in the I3 stage has a high correlation with the fifth cluster (bad values of quantitative traits).

### Conclusions

Most plains in Iran have been plagued by water shortage and drought in the last few years. Accordingly, this study sought to investigate and predict five quantitative traits of guar beans in different deficit irrigation methods. For this purpose, various deficit irrigation methods were applied in the first four growth stages of the guar bean plant. First,

| No. | Antecedent | Confidence % | Support % | Lift |
|-----|------------|--------------|-----------|------|
| 1   | I3 = B and I2 = A | 50           | 8         | 3.57 |
| 2   | I3 = B and I2 = A and I4 = C and I1 = D | 50           | 4         | 3.57 |
| 3   | I3 = B and I2 = A and I1 = C and I4 = D | 50           | 4         | 3.57 |
| 4   | I3 = B and I4 = C | 44           | 8         | 3.13 |
| 5   | I3 = B and I4 = D | 44           | 8         | 3.13 |
| 6   | I3 = B and I1 = D | 44           | 8         | 3    |
| 7   | I3 = B          | 40           | 24        | 2.83 |
| 8   | I3 = B and I4 = C and I1 = A and I2 = D | 38           | 4         | 2.68 |
| 9   | I3 = B and I1 = A and I2 = C and I4 = D | 38           | 4         | 2.68 |
| 10  | I3 = B and I4 = A and I2 = C and I1 = D | 38           | 4         | 2.68 |
| 11  | I4 = C and I2 = D | 31           | 8         | 2.23 |
| 12  | I3 = B and I4 = A | 31           | 8         | 2.23 |
| 13  | I3 = B and I2 = D | 31           | 8         | 2.23 |
| 14  | I2 = A and I4 = D | 31           | 8         | 2.23 |
| 15  | I1 = A and I4 = D | 25           | 8         | 1.79 |
| 16  | I2 = C and I1 = D | 25           | 8         | 1.79 |
| 17  | I2 = A and I4 = C | 25           | 8         | 1.79 |
| 18  | I3 = B and I1 = C and I4 = A and I2 = D | 25           | 4         | 1.79 |
| 19  | I1 = C          | 21           | 24        | 1.49 |
| 20  | I2 = D          | 20           | 28        | 1.40 |
| 21  | I2 = A          | 19           | 24        | 1.34 |
| 22  | I4 = C and I1 = A | 19           | 8         | 1.34 |
| 23  | I1 = A and I2 = C | 19           | 8         | 1.34 |
| 24  | I4 = D          | 16           | 28        | 1.14 |
| 25  | I1 = A          | 15           | 24        | 1.04 |
five quantitative traits of guar beans were clustered by the K-means algorithm, the results of which consisted of five optimal clusters selected by the SVI. According to the results of clustering, the first and third clusters, respectively, had the highest and lowest values of quantitative traits in guar beans compared to the other clusters.

According to the results of the CART-K-means hybrid algorithm, the third cluster is predicted when irrigation A (40% of full irrigation) occurs at the mid-stage of the guar bean growth. The first cluster (with the maximum values of quantitative traits) was predicted in two modes. The first mode was when full irrigation was performed at the mid-stage of plant growth, and the second was when full irrigation and 80% of full irrigation (irrigation C) occur in the development and mid-stages of plant growth, respectively. The order of the importance of irrigation amounts in different stages of growth in the prediction of clusters was determined as the mid-, development, initial, and late stages of plant growth, respectively, by the CART-K-means hybrid algorithm.

Furthermore, the Apriori-K-means hybrid algorithm determined the relationship between different irrigation levels in the four stages of plant growth with the clusters. According to the results, the irrigation amount in the mid-stage of plant growth had the greatest relationship with the clusters. In the mid-stage of plant growth, the highest relationships were observed between irrigation amounts D (full irrigation) with the first cluster, B (60% of full irrigation) and C with the second cluster, A with the third cluster, C with the fourth cluster, and B with the fifth cluster, respectively.

According to the results of this study, the best method of irrigation in the initial, development, mid-, and late stages of plant growth is in order of B, C, D, and A, respectively, to increase the values of quantitative traits in the guar bean plant.

Funding: The authors received no specific funding for this work.

Declarations

Conflict of interest: The corresponding author states that there is no conflict of interest.

Open Access: This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases. In ACM SIGMOD Rec 22(2):207–216
Agrawal R, Srikant R (1994) Fast algorithms for mining association rules. In: Bocca JB, Jarke M, Zaniolo C (eds) Proceedings of 20th International Conference on Very Large Databases, Santiago de Chile, Sept 12–15, 1994: 487–499
Ashraf MY, Akhtar K, Sarwar G, Ashraf M (2002) Evaluation of arid and semi-arid ecotypes of guar (Cyamopsis tetragonoloba L.) for salinity (NaCl) tolerance. J Arid Environ 52:473–482
Ashraf MY, Akhtar K, Sarwar G, Ashraf M (2005) Role of the rooting system in salt tolerance potential of different guar accessions. Agron Sustain Dev 25:243–249
Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression trees. Chapman & Hall/CRC Press, Boca Raton
Brin S, Motwani R, Silverstein C (1997) Beyond market baskets: generalizing association rules to correlations. ACM SIGMOD Rec 26(2):265–276
Chattamvelli R (2011) Data mining algorithm. Alpha Science, Oxford
Cheung DW, Han J, Ng VT, Wong CY (1996) Maintenance of discovered association rules in large databases: an incremental updating technique. In: Proceedings of the twelfth international conference on database engineering. IEEE, pp 106–114
Das D, Ara T, Dutta S, Mukherjee A (2011) New water resistant biobased biocide film based on guar gum. Bioresour Technol 102:5878–5883
Dhanya CT, Nagesh Kumar D (2009) Data mining for evolution of association rules for droughts and floods in India using climate inputs. J Geophys Res Atmos. https://doi.org/10.1029/2008JD010485
Gresta F, Sortino O, Santonoceto C, Issi L, Formantici C, Galante YM (2013) Effects of sowing times on quantitative and qualitative traits of four varieties of guar (Cyamopsis tetragonoloba) in a Mediterranean environment. Ind Crops Prod 41:46–52
Gresta F, Mercati F, Santonoceto C, Abenavoli MR, Ceravolo G, Araniti F, Anastasi U, Sunseri F (2016) Morpho-agronomic and AFLP characterization to explore guar (Cyamopsis tetragonoloba L.) genotypes for the Mediterranean environment. Ind Crops Prod 86:23–30
Gresta F, Cristaudo A, Tostile C, Anastasi U, Guaraccia P, Catara S, Onofri A (2018) Germination of guar (Cyamopsis tetragonoloba (L.) Taub.) genotypes with reduced temperature requirements. Aust J Crop Sci 12:954–960
Kass GV (1980) An exploratory technique for investigating large quantities of categorical data. Appl Stat 29(2):119–127
Lee Jason T, Connor-Appleton S, Haq Akram U, Bailey Christopher A, Cartwright Aubrey L (2004) Quantitative measurement of negligible trypsin inhibitors activity and nutrient analysis of guar mean fraction. J Agric Food Chem 52(21):6492–6495
Lubbe A, Verpoorte R (2011) Cultivation of medicinal and aromatic plants for specialty industrial materials. Ind Crops Prod 34:785–801
Mirhashemi SH, Mirzaei F (2021) Extracting association rules from changes in aquifer drawdown in irrigation areas of Qazvin plain, Iran. Groundw Sustain Dev 12:100495

 Springer
Mirhashemi SH, Haghighat Jou P, Mirzaei F, Panahi M (2020) The study of environmental and human factors affecting aquifer depth changes using tree algorithm. Int J Environ Sci Technol 17(3):1825–1834
Morgan J, Sonquist J (1963) Problems in the analysis of survey data, and a proposal. J Am Stat Assoc 58:415–434
Morgan JN, Messenger RC (1973) THAID, a sequential analysis program for the analysis of nominal scale dependent variables
Pathak R, Roy MM (2015) Climatic responses, environmental indices and interrelationships between qualitative and quantitative traits in clusterbean *Cyamopsis tetragonoloba* (L.) Taub. under arid conditions. Proc Natl Acad Sci India Sect B Biol Sci 85(1):147–154
Pathak R, Singh SK, Singh M, Henry A (2010) Molecular assessment of genetic diversity in clusterbean (*Cyamopsis tetragonoloba*) genotypes. J Genet 89(2):243–246
Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65
Santonoceto C, Mauceri A, Lupini A, Gresta F, Chiera E, Sunseri F, Mercati F, Anastasi U (2019) Morpho-agronomic characterization and genetic variability assessment of a guar germplasm collection by a novel SSR panel. Ind Crops Prod 138:111568
Senapati MK, Srinatha A, Pandit JK (2006) In vitro release characteristics of matrix tablets: study of karya gum and guar gum as release modulators. Indian J Pharma Sci 68(6):824–826
Sortino O, Gresta F (2007) Growth and yield performances of five guar cultivars in a Mediterranean environment. Ital J Agron 4:359–364
Tan KC, Yu Q (2006) A coevolutionary algorithm for rules discovery in data mining. Int J Syst Sci 12(37):835–864
Vaughna SF, Berhowa MA, Winkler-Mosera JK, Leeb E (2011) Formulation of a biodegradable, odor-reducing cat litter from solvent-extracted corn dried distillers grains. Ind Crops Prod 34:999–1002
Wang L, Zhang LM (2009) Viscoelastic characterization of a new guar gum derivative containing anionic carboxymethyl and cationic 2-hydroxy-3-(trimethylammonio) propyl substituents. Ind Crops Prod 29:524–529
Whistler RL, Hymowitz T (1979) Guar: agronomy, production, industrial use, and nutrition. Purdue University Press, West Lafayette
Yoneyama Y, Suzuki S, Sawa R, Yoneyama K, Power GG, Araki T (2002) Increased plasma adenosine concentrations and the severity of pre eclampsia. Obstet Gynecol 100(6):1266–1270
Zambrano F, Despinoy P, Ormenese RC, Faria EV (2004) The use of guar and xanthan gums in the production of light low fat cakes. Int J Food Sci Technol 39:956–966
Zhao Q, Bhowmick SS (2003) Association rule mining: a survey. Technical report, CAIS, Nanyang Technological University

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.