Australian Journal of Crop Science AJCS
AJCS 5(2):162-168 (2011) ISSN:1835-2707

Application of supervised feature selection methods to define the most important traits affecting maximum kernel water content in maize

A. Shekoofa[1], Y. Emam[1], M. Ebrahimi[2], E. Ebrahimie

[1]Department of Crop Production and Plant Breeding, College of Agriculture, Shiraz University, Shiraz, Iran
[2]Bioinformatics Research Group, Green Research Center, Qom University, Qom, Iran

***Corresponding author: ebrahimie@shirazu.ac.ir**

Abstract

This study presents the results of applying supervised feature selection algorithms in the selection of the most important traits contributing to the maximum kernel water content (MKWC) as a major yield component. Data were obtained from a field experiment conducted during 2008 growing season, at the Experimental Farm of the College of Agriculture, Shiraz University, and from the literature. Experiments on the subject of sink/source relationships in maize were collected from twelve fields (as records) of different parts of the world, differing in 23 characteristics (features). The feature selection algorithm demonstrated that 15 features including: planting date (days), countries (Iran, Argentina, India, USA, Canada), hybrid types, Phosphorous fertilizer applied (kg ha^{-1}), final kernel weight (mg), soil type, season duration (days), days to silking, leaf dry weight (g plant^{-1}), mean kernel weight (mg), cob dry weight (g plant^{-1}), kernel number per ear, grain yield (g m^{-2}), nitrogen applied (kg ha^{-1}), and duration of the grain filling period (^{0}C day) were the most effective traits in determining maximum kernel water content. Among the effective traits (features), planting date (days) revealed to be the critical one. Hybrids and countries were the second most important affecting factors on the maize kernel water content. For the first time, our results showed that features classification by supervised feature selection algorithms can provide a comprehensive view on distinguishing the important traits which contribute to maize kernel water content and yield. This study opened a new vista in maize physiology using feature selection and data mining methods and would be beneficial to newcomers of this field.

Keywords: Data mining, Feature selection algorithms, *Zea mays* L., Maximum kernel water content.
Abbreviation: ANOVA, analysis of variance; KGR, kernel growth rate; K, potassium fertilizer applied; MKWC, maximum kernel water content; MC, moisture content; N, nitrogen fertilizer applied; P, Phosphorous fertilizer applied; RCBD, randomized complete block design.

Introduction

Data mining, using various methodologies, has been developed by both commercial and research centers. These techniques are used for industrial, commercial, and scientific purposes (Ebrahimi and Ebrahimie, 2010; Ebrahimi et al., 2010). Recently, agricultural and biological research studies have used various techniques of data mining for analyzing large data sets and establishing useful classification patterns in data sets. However, regarding its novelty and diverse branches, data mining methods are still supposed to bring more fruitful results (Matsumoto, 1998; Cunningham and Holmes, 1999; Hsiao et al., 2006; Amiri Chayjan, 2010; Ebrahimi and Mollazade, 2010). Data mining problems often involve hundreds, or even thousands, of variables. As a result, the majority of spent time and effort in the model-building process involves examining which variables to include in the model. Feature selection allows the variable set to be reduced in size, creating a more manageable set of attributes for modeling (Liu and Motoda, 2008). Feature selection has been an active research area in pattern recognition, statistics, and data mining communities.

The main idea of feature selection is to choose a subset of input variables by eliminating features with little or no predictive information (Handl and Knowles, 2006; Liu and Motoda, 2008; Bijanzadeh et al., 2010). The use of this method enables more complex data to be analyzed, compared to other methods (e.g. statistical techniques), particularly when the feature space is complex and all data do not follow the same distribution pattern (Drummond et al., 2002; Gautam et al., 2006). There are two types of feature selection algorithms: supervised and unsupervised. Supervised feature selection algorithms rely on measures that take into account the class information. A well-known measure is information gain, which is widely used in feature selection (Dash and Liu, 1997). For feature selection in unsupervised learning, learning algorithms are designed to find natural grouping of the examples in the feature space. Thus feature selection in unsupervised learning aims to find a good subset of features that forms high quality of clusters for a given number of clusters (Dy and Brodley, 2004; Liu and Motoda, 2008). Many factors such as planting date, soil type, fertilizer,

Table 1. Authors information, locations and type of manipulative treatments of many literature data that were used for feature selection model of MKWC in maize

Type of treatment	Country	Authors
Defoliation, plant densities, hybrids	Iran	PhD thesis project-Avat Shekoofa et al. (2008-2009)
Defoliation, Restricted pollination	Argentina	Sala et al. (2007)
Hybrids	Argentina	Gambín et al. (2006)
Plant densities, hybrids	Argentina	Gambín et al. (2007)
Hybrids	India	Khanna-Chopra and Maheswari (1998)
Plant densities, Restricted pollination, hybrids	USA	Borrás et al. (2003)
Hybrids, nitrogen levels	Argentina	Melchiori and Caviglia (2008)
Defoliation, kernel removal	USA	Jones and Simmons (1982)
Hybrids	Canada	Subedi and Ma (2005)
Plant densities, Restricted pollination, hybrids	USA	Borrás and Westgate (2006)
Shading, thinning, hybrids	Argentina	Gambín et al. (2008)
Hybrids	USA	Brown et al. (1996)

Table 2. Feature selection classification on maximum kernel water content dataset and estimated important and unimportant inputs. Build setting of feature selection included as: use partitioned data: *false*, screen variability: *true*, min coefficient of variation: *0.0*, max percentage of records in a single category: *90.0*, screen for missing values: *true* , max percentage of missing values: *75.0*, max number of categories as a percentage of records: *95.0*, screen standard deviation: *true*, min standard deviation: *0.0*, base the p-value (importance) for categorical predictors on: *Pearson*, low importance from 0.0 and up to: *0.9*, medium importance above and up to: *0.95*

	Rank	Field	Type	Importance	Value	N°articles
True	1	Planting date (days)	Set	Important	1.0	10
True	2	Hybrids	Set	Important	1.0	12
True	3	Countries	Set	Important	1.0	12
True	4	P applied (kg/ha)	Range	Important	1.0	3
True	5	Final kernel weight (mg)	Range	Important	1.0	10
True	6	Soil type	Set	Important	1.0	5
True	7	Season duration (days)	Range	Important	1.0	4
True	8	Days to silking	Range	Important	1.0	4
True	9	Leaf dry weight (g/plant)	Range	Important	1.0	3
True	10	Mean kernel weight (mg)	Range	Important	1.0	8
True	11	Cob dry weight (g/plant)	Range	Important	1.0	3
True	12	Kernel number per ear	Range	Important	0.999	11
True	13	Grain yield (g/m^2)	Range	Important	0.996	6
True	14	N applied (kg/ha)	Range	Important	0.98	11
True	15	Duration of the grain filling period (^{0}C day)	Range	Important	0.962	9
False	16	Genotype type	Set	Unimportant	0.868	7
False	17	Density (plant/ha)	Range	Unimportant	0.848	10
False	18	Kernel dry weight(mg)	Range	Unimportant	0.702	3
False	19	Kernel growth rate (mg ^{0}C day^{-1})	Range	Unimportant	0.651	6
False	20	stem dry weight (g/plant)	Range	Unimportant	0.444	3
False	21	Defoliation	Set	Unimportant	0.413	6
False	22	Soil pH	Range	Unimportant	0.130	3
False	23	K applied (kg/ha)	Range	Unimportant	0.113	3

location, hybrid, season duration, etc. are all affecting yield and yield components of a grain crop. In this investigation, supervised feature selection algorithms were applied to select the most important traits (features) contributing to maximum kernel water content (MKWC) as a major yield component in maize (among a large number traits). The maize kernel undergoes large changes in water content during its development. Final kernel weight is closely related to MKWC (Saini and Westgate, 2000; Borrás et al., 2003). The mechanisms by which kernel water relations might regulate kernel development have not yet been established, however, a close coordination between dry matter accumulation, water content development, assimilate supply and crop husbandry has been reported (Borrás and Westgate, 2006; Melchiori and Caviglia, 2008). Kernel biomass accumulation stops when kernels reach a critical percent moisture content (MC) value (around 36% for maize), indicating the importance of maintaining a kernel water status above this critical one for increasing the duration of grain filling (Egli, 1990). Moreover, the kernel developmental stage during grain filling can be estimated by measuring kernel percent MC (Borrás and Westgate, 2006). Before invention of data mining methods, analyzing the results from one or two experiments was a common practice to achieve conclusions. In contrast, feature selection provided the opportunity to study the wide range of growing conditions by analyzing data from literature. The aim of our study was to shed light on the relationships between kernel water content, maize physiological process, and different field conditions using supervised feature selection method. Understanding these relationships play a critical role in predicting kernel weight responses when favorable and unfavorable growth conditions occur during the reproductive growth of maize crops. We thus expect to build an intelligent agricultural information system to assist the experts and to help an improvement on agricultural technologies.

Materials and methods

Data collection

Data presented in this study was collected from the literature on the subject of sink/source relationships in maize (Table 1). In addition, data was also obtained from the experiment carried out without any discernible nutrient or water limitations during 2008 growing season, at the Experimental Farm of the College of Agriculture, Shiraz University, Badjgah, (29° 50´ N and 52° 46´ E; elevation: 1810 m above mean sea level). The experimental design was a randomized complete block design (RCBD) with three replicates and treatments designed in a split-split plot arrangement. Three hybrids (370, Maxima 524 and 704) were the main plots, the plant densities (75 000, 85 000 and 95 000 plants ha^{-1}) were allocated to subplots, and defoliation (control-without defoliation, 50% of defoliation at 25, and 35 days after silking) in the sub-subplots.

Target features

Consequently, twelve records with 23 features including Phosphorous (P) applied (kg ha^{-1}), final kernel weight (mg), grain yield (g m^{-2}), season duration (days), days to silking, leaf dry weight (g plant^{-1}), mean kernel weight (mg), cob dry weight (g plant^{-1}), kernel number per ear, nitrogen (N) applied (kg ha^{-1}), plant density (plant ha^{-1}), stem dry weight (g plant^{-1}), kernel dry weight (mg), duration of the grain filling period (^{0}C day), kernel growth rate (mg ^{0}C day^{-1}), soil pH, potassium (K) applied (kg ha^{-1}), location, hybrid name, hybrid type, defoliation, planting date, and soil type with the MKWC were recorded. The MKWC was set as output variable and the rest of variables as input variables.

Feature selection procedures

Here, we applied feature selection algorithm to recognize those attributes that have any strong correlation with MKWC. The algorithm considered one attribute at a time to see how well each predictor alone predicts the target variable (output). The general feature selection process is illustrated in Fig. 1. The importance value of each variable was then calculated as (1- p) where p was the p value of the appropriate test of association between the candidate predictor and the target variable. The association test for categorized output variables (e.g., MKWC) was different from the test for continuous ones. When the target value was continuous, p values based on the F statistic were used. If some predictors are continuous and some are categorical in the dataset, the criterion for continuous predictors is still based on the *p* value from a transformation and that for categorical predictors from the *F* statistic. Predictors are ranked by the following rules: (1) Sort predictors by *p* value in ascending order. (2) If ties occur, follow the rules for breaking ties among all categorical and all continuous predictors separately, then sort these two groups (categorical predictor group and continuous predictor group) by the data file order of their first predictors (Liu and Motoda, 2008; Clementine® 11.0 Algorithms Guide).

Statistical analysis

The idea was to perform a one-way ANOVA and F test for each predictor; otherwise, the p value was based on the asymptotic t distribution of a transformation on the Pearson correlation coefficient. When some-but not all-predictors are categorical and the target is also categorical, importance can be ranked based on either the Pearson or likelihood-ratio chi-square. The predictors were then labeled as 'important', 'marginal', and 'unimportant' with values above 0.95, between 0.95 and 0.90, and below 0.90, respectively. In spite of feature selection that was used to reduce the complexity of input parameters by eliminating marginally important and/or unimportant predictors, classification was used in an innovative way to reduce the complexity of output variable MKWC by converting it from a continuous variable with an unrestricted number of possible values to a flag variable with possible variables of 'True' or 'False'.

Results and discussion

This work was motivated by the research on feature selection algorithm. Features classification (Table 2) indicated that among 23 tested features, 15 features were the most important features related to the MKWC pattern recognition (Table 2). These included planting date, countries, hybrids, P applied, final kernel weight, soil type, season duration, days to silking,

Table 3. Results of General liner models between maize planting date (days) and maximum kernel water content (MKWC) (mg). Statistics are reported for planting date that had a significant relationship (P<0.01) between differences of planting date and MKWC

Location	Planting date (days)		P value
Department of Plant Production at the University of Buenos Aires	1 October		
INTA-Balcarce-Buenos Aires-Argentina		15 October	0.0000
Department of Plant Production at the University of Buenos Aires		23 September	0.0522
Department of Plant Production at the University of Buenos Aires		8 October	0.0001
INTA-Parana-Argentina		late December	0.0000
INTA-Parana-Argentina		mid September	0.0000
INTA-Balcarce-Buenos Aires-Argentina	15 October		
Department of Plant Production at the University of Buenos Aires		23 September	0.0031
Department of Plant Production at the University of Buenos Aires		8 October	0.0129
INTA-Parana-Argentina		late December	0.0000
INTA-Parana-Argentina		mid September	0.0000
Department of Plant Production at the University of Buenos Aires	23 September		
Department of Plant Production at the University of Buenos Aires		8 October	0.0673
INTA-Parana-Argentina		late December	0.0000
INTA-Parana-Argentina		mid September	0.0000
Department of Plant Production at the University of Buenos Aires	8 October		
INTA-Parana-Argentina		late December	0.0000
INTA-Parana-Argentina		mid September	0.0000
INTA-Parana-Argentina	late December		
INTA-Parana-Argentina		mid September	0.9110
Ames-IA-USA	10 May		
Bruner-Iowa Stat University-Ames		11 May	0.9854
Ames-IA-USA		29 May	0.1483
Badjgah-Shiraz-Iran		5 June	0.1834
Bruner-Iowa Stat University-Ames	11 May		
Ames-IA-USA		29 May	0.3140
Badjgah-Shiraz-Iran		5 June	0.5208
Ames-IA-USA	29 May		
Badjgah-Shiraz-Iran		5 June	0.8036

leaf dry weight, mean kernel weight, cob dry weight with 1.0 value, and kernel number per ear (0.999 value), grain yield (0.996 value), N applied (0.98 value), and duration of the grain filling period (0.962 value). The rest of features included as hybrid type (0.868 value), plant density (0.848 value), kernel dry weight (0.702 value), kernel growth rate (0.651 value), stem dry weight (0.444 value), defoliation (0.413 value), soil pH, and K applied revealed to be unimportant features. We found that the classifier performance improved by eliminating redundant features (Table 2). In our study, redundant features were genotype, density (plant ha^{-1}), kernel dry weight (mg), kernel growth rate (mg ^{0}C day^{-1}), stem dry weight (g plant^{-1}), defoliation, soil pH, K applied (kg ha^{-1}). The results showed that planting date, hybrids, and countries (1.0 value) were the most important effective traits on MKWC (Table 2). The relationship between one important management decision, planting date, and yield potential has been previously documented by Lauer et al. 1999 and Nielsen et al. 2002. In spite of the fact that kernel number per unit land area is the most important yield component, MKWC related to kernel dry weight is also an important contributor to grain yield (Saini and Westgate, 2000; Gambín et al., 2008). It should be noted that when the source capacity of the crop was reduced during grain filling, the developmental pattern of kernel water content was not as seriously affected as was final kernel weight (Sala et al., 2007). In nine locations, there was a significant (P< 0.01) relationship between the planting dates and MKWC (Table 3). In many of these nine locations such as, INTA-Parana-Argentina, Ames-IA-USA, Bruner-Iowa Stat University-Ames, INTA-Balcarce-Buenos Aires-Argentina, Department of Plant Production at the University of Buenos Aires, and Badjgah-Shiraz-Iran, there were a significant relationship between early, average and late planting date (such as, 1 October, 8 October, 15 October, 23 September and late December) with MKWC (Table 3). According to the applied feature selection algorithms, the planting date, (with 1.0 value) was the most effective feature. The MKWC of maize kernel in early planting date (1 October) was (51.11%) more than the late planting date (late December). Nielsen et al., (2002) and Kucharik, (2008) reported that planting date and countries were strongly correlated to the yield and yield components of maize. This result showed the impact of planting date on grain yield. More studies are needed to quantify the specific effects of crop management, biophysical changes on yield and yield component, and chemical changes effects on both management decisions and productivity. The results of this research suggested that classifying by feature selection algorithm might

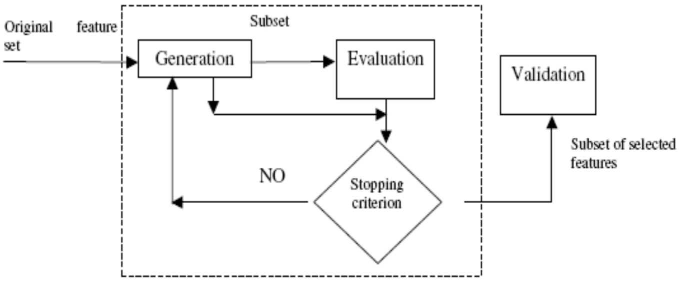

Fig 1. General procedure of feature selection (Dash and Liu, 1997)

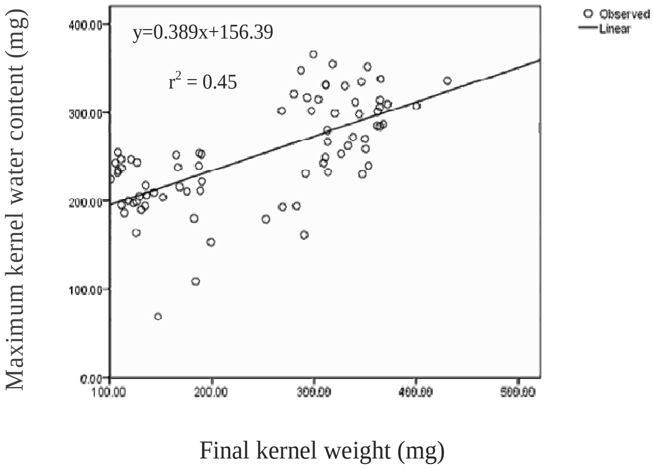

Fig 2. Scatter plot of maximum kernel water content (MKWC) and final kernel weight of maize in this work. Best-fit linear regression is plotted in case where the relationship was significant at P<0.01

be an effective tool for the comparison of features such as hybrids, countries, fertilizers, and yield components (mean kernel weight, final kernel weight, etc.) within the different study areas. Selection of the appropriate hybrids is an important cultural practice. In this study, the average MKWC of maize kernel showed a range from 137.33 (mg) in DK 752 hybrid to maximum 304.5 (mg) in Ax842 MG hybrid), 304.5 (mg). Some of the countries of this research, a significant relationship ($P<0.05$) existed between hybrids and MKWC (data not shown). Final kernel weight and kernel number per ear (KNPE) were two the most relevant traits with 1.0 value affecting severely on MKWC (Table 2). Our result showed that final kernel weight was correlated with maximum kernel water content ($R^2=0.45$, $P<0.0001$) (Fig. 2). This finding is consistent with previous reports where water content was used as an early estimator of final kernel weight (Saini and Westgate, 2000; Borrás et al., 2003). Cob dry weight, leaf dry weight, and MKWC have shown to be closely related (Table 2). The relationship between final kernel weight and grain yield has been demonstrated by some researchers e.g. Khanna-Chopra and Maheswari, (1998) and Subedi and Ma (2005). It is probable that kernels adjust their weight by receiving assimilates or remobilization from the cob and leaf in the end of growing season. Modifications in kernel weight can be explained by changes in kernel growth rate, which was closely correlated to maximum water content during rapid grain filling. Several studies of commercial maize germplasm have shown how kernel water accumulation can be used to normalize genotypic variation in kernel development and predict environmental effects on kernel growth (Borrás and Westgate, 2006; Gambín et al., 2007; Borrás et al., 2009). Because maize kernels increase water content before rapid accumulation of reserves, the maximum achieved water content in early development provides a fairly accurate estimate of potential kernel size and is closely related to the kernel growth rate (KGR) (Borrás et al., 2003; Borrás and Westgate, 2006; Borrás et al., 2009). Kernel water concentration has further been used with good success for estimating the percent of maximum kernel weight, achieved at any grain-filling stage in different species (Schnyder and Baum, 1992; Borrás et al., 2003). From the results presented in this research, it seems that a subset (15 features) of final kernel weight traits of the original traits (23 features) can be assumed as the relevant important traits of the MKWC samples. Understanding the metabolic factors determining how maximum kernel water content is achieved and regulated is essential to increase sink strength under favorable conditions in future works. Another interesting item rising from this research is the fact that different features can be found, providing the same percentage of right classification (Table 2). The features selected by the feature selection algorithm corresponding to one possible solution, will give information to agronomists as to what traits are more relevant for MKWC. In conclusion, a supervised feature selection has been used in this paper to reduce the time and cost of feature acquisition, as well as reducing classifier training and testing time providing more understandable results. The study also showed that feature selection algorithm provided better accuracies for predicting the more relevant features present in MKWC (maximum kernel water content) of maize yield components. Feature selection is also helpful in improving classifier accuracy, provided that noisy, irrelevant or redundant features are eliminated. Our results recommended that feature classification by supervised feature selection algorithm may be a suitable option for determining the important features such as planting date, location, hybrids, final kernel weight, grain yield, mean kernel weight, cob dry weight, etc. The aim of our complementary works is to identify a more efficient and computationally stable technique for determining values of important features among all inputs. The weak point of analyzing the result of just one experiment (one field) is that commonly, the outcome of this experiment is strong and reliable just for that specified condition. In fact, in many cases, it is not possible to extend the data of separate experiments providing conflicting results. The results of this study showed that feature classification by using supervised feature selection algorithms is a suitable option for determining the important features contributing to MKWC providing a comprehensive view about this trait. Determining critical key features is valuable for maize grain yield improvement in the maize-producing countries. This is the first work in identifying the most important factors on maize grain yield from many fields in the world by using supervised feature selection algorithm.

Finally, the algorithm was applied to be modified to use for other applications and can use any number of parameters.

Conclusions

Although traditional statistical methods have been already applied in agricultural experiments, we expect recent supervised feature selection methods to bring still more fruitful results. The main advantages of feature selection methods are the reduction of the data processing time, decrement in the requirements of data storage space, decreasing in the cost of data acquirement and the most important, it allows to select a subset of the original features which contribute with the largest amount of information for a particular problem (reduction in the dimensionality of the input data). As such, Supervised Feature Selection Methods is an interesting tool available to define critical crop physiological traits. However, and as shown in this article, it was not able to provide outcomes physiologists are not able to detect with simple field experiments. The fact that planting date is a critical issue for KMWC is already known by physiologist and breeders, and current farming practices have already identified the optimum planting date for every production area.

Acknowledgment

We would like to thank the Department of Crop Production and Plant Breeding of Shiraz University for supporting this research. Lucas Borrás, Departamento de Produccion Vegetal, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, S2125ZAA Zavalla, Santa Fe, Argentina, commented very helpfully on a draft of this manuscript, authors wish to thank for useful comments.

References

Amiri Chayjan R (2010) Modeling of sesame seed dehydration energy requirements by a soft-computing approach. Aust J Crop Sci 4(3):180-184

Bijanzadeh E, Emam Y, Ebrahimie E (2010) Determining the most important features contributing to wheat grain yield using supervised feature selection model. Aust J Crop Sci 4 (6): 402-407

Borrás L, Westgate ME, Otegui ME (2003) Control of kernel weight and kernel water relations by post-flowering source–sink ratio in maize. Ann Bot 91: 857–867

Borrás L, Westgate ME (2006) Predicting maize kernel sink capacity early in development. Field Crop Res 95: 223–233.

Borrás L, Zinselmeier C, Lynn Senior M, Westgate ME, Muszynski MG (2009) Characterization of grain-filling patterns in diverse maize germplasm. Crop Sci 49: 999–1009

Brown G, Wuethrich K, Halter A (1996) Kernel dry weight loss during post-maturity drydown intervals in corn. Available online: http://www.agry.purdue.edu/ext/corn/research/rpt94-01.htm

Clementine® 11.0 Algorithms Guide. Copyright© (2007) by Integral Solutions Limited. All rights reserved. Printed in the United States of America. Pp. 250. http://www.spss.com/ worldwide/

Cunningham SJ, Holmes G (1999) Developing innovative applications in agriculture using data mining. In the Proceedings of the Southeast Asia Regional Computer Confederation Conference, Singapore.1999

Dash M, Liu H (1997) Feature selection for classification. Intell Data Ana 1(3): 131–156

Drummond ST, Sudduth KA, Joshi A, Birrell SJ, Kitchen NR (2002) Statistical and neural network methods for site-specific yield prediction. Trans Am Soc Agric Eng 47(1): 5–14

Dy J, Brodley C (2004) Feature selection for unsupervised learning. J Mach Lear Res 5: 845–889

Ebrahimi M, Ebrahimie E (2010) Sequence-Based Prediction of Enzyme Thermostability Through Bioinformatics Algorithms. Curr Bioinform. 5: 195-203

Ebrahimi M, Ebrahimie E, Shamabadi N, Ebrahimi M (2010) Are there any differences between features of proteins expressed in malignant and benign breast cancers? Journal of Research in Medical Sciences. 15(6): 299-309

Ebrahimi E, Mollazade K (2010) Integrating fuzzy data mining and impulse acoustic techniques for almond nuts sorting. Aust J Crop Sci 4(5):353-358

Egli DB (1990) Seed water relations and the regulation of the duration of seed growth in soybean. J Exp Bot 41: 243–248

Gambín BL, Borrás L, Otegui ME (2006) Source–sink relations and kernel weight differences in maize temperate hybrids. Field Crop Res 95: 316– 326

Gambín BL, Borrás L, Otegui ME (2007) Kernel water relations and duration of grain filling in maize temperate hybrids. Field Crop Res 101: 1-9

Gambín BL, Borrás L, Otegui ME (2008) Kernel weight dependence upon plant growth at different grain-filling stages in maize and sorghum. Aust J of Agric Res 59: 280-290

Gautam R, Panigrahi S, Franzen D (2006) Neural network optimisation of remotely sensed maize leaf nitrogen with a genetic algorithm and linear programming using five performance parameters. Biosyst Eng 95 (3): 359–370

Handl J, Knowles J (2006) Feature subset selection in unsupervised learning via multiobjective optimization. Inter J of Comput Intell Res 2(3): 217–238

Hsiao HW, Tasi MS, Wang SC (2006) Spatial data mining of colocation patterns for decision support in agriculture. Asian Journal of Health Info Sci 1: 61-72

Jones RJ, Simmons SR (1983) Effect of altered source-sink ratio on growth of maize kernels. Crop Sci 23: 129-134

Khanna-Chopra R, Maheswari M (1998) Effect of altering source availability on expression of sink capacity in a maize hybrid and its parents. Eur J Agron 9: 101-107

Kucharik CJ (2008) Contribution of planting date trends to increased maize yields in the Central United States. Agron J 100: 328-336

Lauer JG, Carter PR, Wood TM, Diezel G, Wiersma DW, Rand RE, Mlynarek MJ (1999) Corn hybrid response to planting date in the northern Corn Belt. Agron J 91: 834-839

Liu H, Motoda H (Eds) (2008) Computational Methods of Feature Selection. Chapman & Hall/CRC Taylor & Francis Group 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487‑2742 © 2008 by Taylor & Francis Group, LLC Chapman & Hall/CRC is an imprint of Taylor & Francis Group, an Informa business. <http://www.crcpress.com>

Matsumoto K (1998) An Experimental Agricultural Data Mining System. Arikawa, S., and Motoda, H., (Eds.): DS'98, LNAI 1532, pp. 439–440

Melchiori RJM, Caviglia OP (2008) Maize kernel growth and kernel water relations as affected by nitrogen supply. Field Crop Res 108: 198-205

Nielsen RL, Thomison PR, Brown GA, Halter AL, Wells J, Wuethrich KL (2002) Delayed planting effects on flowering and grain maturation of dent corn. Agron J 94: 549-558

Sala RG, Westgate ME, Andrade FH (2007) Source/sink ratio and the relationship between maximum water content, maximum volume and final dry weight of maize kernels. Field Crop Res 101: 19–25

Saini HS, Westgate ME (2000) Reproductive development in grain crops during drought. Adv Agron 68: 59–96

Schnyder H, Baum U (1992) Growth of the grain of wheat (*Triticum aestivum* L.). The relationship between water content and dry matter accumulation. Eur J Agron 2: 51–57

Subedi KD, Ma BL (2005) Ear position, leaf area, and contribution of individual leaves to grain yield in conventional and leafy maize hybrides. Crop Sci 23: 2246-2257