Wine Quantity and Quality Variations in Relation to Climatic Factors in the Tokaj (Hungary) Winegrowing Region

László Makra,[1]* Béla Vitányi,[2] András Gál,[2] János Mika,[3,4]
István Matyasovszky,[5] and Tamás Hirsch[3]

Abstract: The effects of climatic elements on wine quantity and quality for the winegrowing region of Tokaj-Hegyalja, Hungary, were analyzed. By applying the Makra test, significant breaks were detected in both the wine quantity and wine quality data sets. Analysis of the relationship between climatic variables and wine quantity and quality, using different statistical methods, confirmed that the most important factors of wine quantity are hours of sunshine in May, June, July, and August and precipitation in September. For wine qualilty, mean temperature, precipitation, and hours of sunshine in May and September play key roles in addition to precipitation in July and hours of sunshine in August. The role of climate in September is most important, since aszú (*Botrytis*) formation, as an important component of wine quality, depends largely on conditions during September. Results indicate that the significant variables obtained by factor analysis better explain linear relationships between climate and wine quantity and quality than those obtained by the χ^2 test. Seven objective vintage climate types were defined using the methods of factor and cluster analysis. Results show that the classification of vintage climate types is more effective in explaining variations in wine quantity than variations in wine quality. Overall, the research identifies the characteristics and importance of the climatic variables with significant relationships with wine quantity and quality in the region. The results are useful in applying quantity and quality assessment strategies for wine production in the region.

Key words: Tokaj, Hungary, wine quality and production, climate, vintage climate type classification

Tokaj is the northernmost of prominent wine-producing regions in Europe (Supplemental Figure 1). The official name of the wine region is Tokaj-Hegyalja (*Hegyalja* translates as "foothills"), which is situated 200 km east of Budapest in northeastern Hungary. The Tokaj region consists of 28 villages and 6,000 ha of classified vineyards, of which an estimated 5,000 ha are currently planted (according to 2005 data from the Wine-Growing Research Institute of Tarcal [WGRIT 2005]). Tokaj is unique among wine regions in having been declared a UNESCO World Heritage Site in June 2002. In addition, the region is home to the world's oldest *Botrytis* wine: Tokaji aszú.

[1]Department of Climatology and Landscape Ecology, University of Szeged, H-6701 Szeged, P.O.B. 653, Hungary; [2]István Bocskai Secondary and Technical School, H-3900 Szerencs, Ondi út 1, Hungary; [3]Hungarian Meteorological Service, H-1525 Budapest, P.O.B. 38, Hungary; [4]Department of Geography, Eszterházy Károly College, H-3300 Eger, Leányka u. 6, Hungary; and [5]Department of Meteorology, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter u. 1/A, Hungary.

*Corresponding author (email: makra@geo.u-szeged.hu)

Acknowledgments: The authors thank the Wine-Growing Research Institute of Tarcal for climatic and wine quantity and quality data; Tibor Kovács (land manager, Hétszőlő RT Stock Company, Tokaj) for valuable advice and suggestions; and Zoltán Sümeghy for digital mapping.

Supplementary data is freely available with the online version of this article at www.ajevonline.org.

Manuscript submitted Sept 2008, revised Jan, Mar 2009, accepted May 2009. Publication costs of this article defrayed in part by page fees.

Aszú is a specific Tokaj wine style. The original meaning of *aszú* was "dried," but the term came to be associated with the type of wine made with botrytized (i.e., nobly rotten) grapes. Important preconditions of the formation of aszú are the late harvest, which traditionally starts in the region near the end of October, and the occurrence of the *Botrytis cinerea* fungus, which causes the berries to partially evaporate and shrivel. Sugar content can be as high as 780 g/L. Various wines are made from the aszú essence of the shriveled berries.

The optimum climate conditions for vinegrowing in Tokaj region are abundant precipitation at the end of the preceding year and in the spring of the current year; plentiful irradiance and warm air at bloom (abundant precipitation is harmful at this time); precipitation during berry development; and a long, dry and warm autumn during berry ripening (Gál 2004, Vitányi 2004). The climate in Tokaj is continental with dry, hot summers and cold winters. Frequent fog at the foot of the mountains (called the Tokaj "skirt") results in low irradiance and air temperature in spring, which slows budding. Regular spring frosts in April and May also limit the yield of the skirt regions. The autumn is fairly long and dry, and a period of sunny, warm weather in autumn (called "Indian summer") is typical (Vitányi 2004). Usually occurring after the first frost, Indian summer begins at the beginning of October and lasts for approximately 2 to 3 weeks. This weather makes the berries prone to attack by *Botrytis cinerea*, giving them high sugar, mineral, and organic acid content.

The role of soil is also important. The bedrock in the region is largely comprised of volcanic tufa, with slight variations from vineyard to vineyard, but consistently rich in minerals. The top soil is primarily loess in the south and primarily clay in the north (Gál 2004).

The relatively high ratio of vine on the skirts or even on the lowland, which has the most climatic risks (e.g., fog, frost, saturated soils), is the key factor that can adversely affect the crop yield and quality. Other factors, although not climatic ones, are inadequate technology and aging vineyards.

There are numerous studies on the climatic factors that influence wine quantity and quality. Temperature, irradiance, humidity, and rainfall have major impacts on crop yield (Steel and Greer 2008). The relationship between crop yield of winegrapes (among other major crops) and three climatic variables (minimum temperature, maximum temperature, and precipitation) has been analyzed (Lobell et al. 2007). Strong relationships between meteorological conditions and wine quality have been highlighted (Grifoni et al. 2006), with higher-quality wines obtained in the years characterized by a reduction in rainfall and warmer temperatures. A recent study found that with the increase of mean annual temperature and potential evapotranspiration, harvest dates advanced and sugar concentrations at harvest times increased the potential alcohol content (Laget et al. 2008). According to one analysis (Soar et al. 2008), higher frequencies above certain temperatures were associated with significantly better wine quality. However, the issue of how climate elements affect grape production and wine quality is more complex and depends on the values of the climate variables in the phenological intervals. Caprio and Quamme (2002) revealed both the unfavorable and beneficial effects of temperature and precipitation on grape production for different seasons of the year. Jones and Davis (2000) studied the relationships between climate and total production and quality. They found that during the last two decades of the 20th century, the increase in the number of warm days during flowering and veraison and a reduction in precipitation during maturation amounted to a general increase in vintage ratings and greater potential wine quality. In the second half of the 20th century, the majority of the world's highest quality wine-producing regions have experienced growing season warming trends (Jones et al. 2005).

Our study was to analyze the effect of climatic elements on wine quantity and quality for the Tokaj winegrowing region. In addition, a key goal was to develop an objective, reliable classification system of vintage climate types prevailing over the Tokaj region through application of multivariate statistical methods. For each vintage climate type characterized by homogenous conditions of temperature, precipitation, and hours of sunshine, the values of wine quantity and wine quality were estimated.

Materials and Methods

The climatic data reported here is from the meteorological station operating at WGRIT, of the Ministry of Agriculture and Rural Development. The village of Tarcal is located in the southern part of the region studied (Supplemental Figure 1). Monthly data of three climatic variables were analyzed from April to September for the years 1901 to 2004: mean monthly temperature (T_{mean}, °C), monthly precipitation total (P, mm), and monthly hours of sunshine (S, hour).

Despite the widely accepted hypothesis that diurnal data generally contain more useful information on yield quantity and quality than monthly data, we have limited our study to monthly data because of the lack of quality data at diurnal resolution in the region. Tarcal is the only station with sunshine hour observations, and the complexity of the topography would not allow us to extend diurnal data of this single station to the entire region, particularly in the case of diurnal extremities. However, the Tokaj region has rarely experienced spring freeze events in the past, and the occurrence of such events may further decrease with global warming (Bartholy et al. 2008). However, spring freezes should not be taken out of consideration of our monthly approach (Gál 2004). Other weather extremes, such as a high number of rainy days or high air humidity, are explicitly or implicitly reflected by the monthly anomalies.

Ratings of annual wine production (in units of a thousand hectoliters) and wine quality consist of codes for the entire region. The original wine quality ranking was prepared and applied by WGRIT (Table 1). Our quality scores consist partly of wine quality characteristics and partly of the quantity measure of the so-called aszú berry production in the given year. Hence, the scores are comprised of both subjective (such as sensory quality ratings—aroma, flavor) and objective components (such as alcohol, sugar free extract, titrated acid, and citric acid content). Aszú berry production in a given year, as another part of the quality rating, is also an objective category.

Wine parameters are derived from the whole region of Tokaj; however, the monthly climatic variables only reflect the circumstances of one single station, namely Tarcal, which may weaken the representation of our analysis. Nevertheless, the small region surveyed (the longest distance

Table 1 Wine quality coding in the original system and the new ranking system (codes 5 and 8 are not represented in the time series because there were neither medium-quality vintages with a great deal of aszú nor vintages with a substantial amount of aszú).

Original system[a]	New system[b]
I	9
II	8
1+I	7
1+II	6
2+I	5
2+II	4
1	3
2	2
3	1

[a]I: aszú vintage with great deal of aszú; II: substantial amount of aszú; 1: high quality yield; 2: medium quality yield; 3: low quality yield.
[b]9: best quality; 1: weakest quality.

between two points is 52 km) may not exhibit spatially independent points within its area. The spatial correlation between two points of a plain area, located 50 km from each other, is over 0.97 in July–August, and over 0.99 in January–February (Czelnai et al. 1976) (the approximate distance between southwestern and northeastern ends of the Tokaj region). Similar correlations for monthly precipitation are 0.66 and 0.65. Hence, the spatial average temperature and precipitation over the region should vary more or less synchronously with the temperature and even precipitation variations of its single point, the Tarcal station. The surface-monitoring system could not adequately represent the topo-microclimatic characteristics of the relief, which could weaken this correlation between climate and wine parameters. However, that is the case when incorporating even substantially more stations.

The Makra test is a new interpretation of the classic two-sample test. The basic question is whether a significant difference can be found between the averages of an arbitrary subsample of a given time series and the whole sample (Makra et al. 2002). The test is evaluated at the 99% significance level.

There are several quantities for measuring the dependence between random variables. An optimal measure satisfies six criteria (Granger et al. 2004). A generalized correlation r_G based on mutual information is used here (Dionisio et al. 2004) and satisfies the above-mentioned criteria, except that r_G only takes values in the range [0,1]. A generalized correlation should be related to the correlation r via a simple functional form for bivariate normal random variables. Here, it is $r_G = |r|$. In general, $r_G > |r|$, therefore the difference $r_G - |r|$ provides information about how nonlinear the relationship is. In particular, a generalized correlation significantly different from zero indicates a certain nonlinearity when the correlation has to be taken as zero.

Factor analysis explains linear relationships among subsets of examined variables, which helps to reduce the dimensionality of the initial database without a substantial loss of information. First, factor analysis was applied to the initial data set consisting of 18 columns (climatic variables) and 104 rows (years) to reduce the 18 climatic variables to a smaller number of m. Factors can be viewed as main latent variables potentially influencing wine quantity and quality. The optimum number m of the retained factors is determined with the criterion of reaching the least percentage of the total variance (in our case 80%) in the original variables that needs to be achieved. After factor analysis, a special transformation of the retained factors was performed to determine to what degree the variables examined affected the resultant variable and to rank them according to their importance (Jolliffe 1990, 1993, Fischer and Roppert 1965, Jahn and Vahle 1968).

A cluster analysis was applied to the factor score time series that objectively grouped the years with similar climate conditions. Hierarchical cluster analysis using Ward's method (Ward 1963) was applied to the climatic variables of the 6-month (April–September) period over the 104 years examined. Ward's method attempts to minimize the sum of squares of elements within clusters forming at each step during the procedure. The procedure works with the Mahalanobis metric (Mahalanobis 1936), which is deemed better than the Euclidean metric and which takes into account the standard deviations of the components of the vectors to be clustered and the correlations among the components. We selected the number of clusters under possible cluster numbers from 3 to 10 to ensure nearly uniform occurrence frequencies of clusters. Intuitively, the final system of clusters delivers small variation of occurrence frequencies of clusters constrained on forming these clusters by Ward's method (Anderberg 1973, Hair et al. 1998).

Finally, one-way analysis of variance (ANOVA) was used to determine, for sample elements grouped into clusters, whether the intergroup variance was significantly higher than the intragroup variance. Post-hoc Tukey test was used to establish (after performing ANOVA on the averages of the groups examined) which groups differed significantly from each other (Tukey 1985, Makra et al. 2006). Linear trend analysis was used to check the significance of linear trends of arbitrary subsamples of the data set examined (Ezekiel and Fox 1970). Pearson's χ^2 test was applied to test the independence of row and column classifications in an unordered contingency table (Danielides et al. 2002, Bolla and Krámli 2005).

Results

Linear trends of the entire 104-year period along with those in every possible (3–103 years) subperiod were computed and tested. The whole (104 element) data set for wine quantity revealed a significant trend at the 95% probability level. Trend analysis performed on all subperiods of the entire period having 3, 4, …103 years, as subsample elements, indicated very few significant trends. All these significant trends corresponded to subperiods of three or four elements, and these trends were found sporadically in the data set. The significant trend found for the whole wine quantity data set was subtracted from the original wine quantity data in each year, and further statistical analysis of wine quantity was performed on this trend-free data. But neither the data set of wine quality nor that of the 18 climatic variables showed any significant trends.

With the application of the Makra test, the wine quantity data set resulted in a significant positive break with a 31-year period (1916–1946) and a significant negative break with a 25-year period (1947–1971) (Figure 1). These intervals clearly fit the breaks of the data set, which precede the higher year-to-year variation of the record after 1971. On the other hand, the wine quality data set had only one significant break—a significant negative subperiod with a 45-year period (1938–1982) (Figure 2). The correlation coefficient between the trend-free wine quantity and wine quality data was 0.144, which was not significant at either 99% or 95% probability levels.

A further task was to determine the relative influence and importance of the 18 climate variables in determining wine quantity variations. The same analysis was performed for wine quality. First, factor analysis of wine quantity and the 18 climatic variables was conducted, after which a special transformation was applied. Ten factors were retained, accounting for 81.5% of the total variance of the original variables. Wine quantity was defined as the resultant variable, and the climatic variables as the influencing variables. In order to rank the effect of the influencing variables on the resultant variable, loadings of factors 2, 3, … 9 and 10 were transformed to factor 1 (Fischer and Roppert 1965, Jahn and Vahle 1968). Just three climatic variables had a significant connection (at least at 95% probability level) with wine quantity as the resultant variable: mean temperature in August, hours of sunshine in June, and hours of sunshine in August (by rank of importance). High values of these variables increased wine quantity, while corresponding low values decreased it. Accordingly, warm and sunny (dry) summers favored higher crop yield, since this kind of weather, with soil of adequate humidity, promotes undisturbed growth of berries. The effect of the other climatic variables could not be measured on their factor loadings because of insignificant relationships (Table 2).

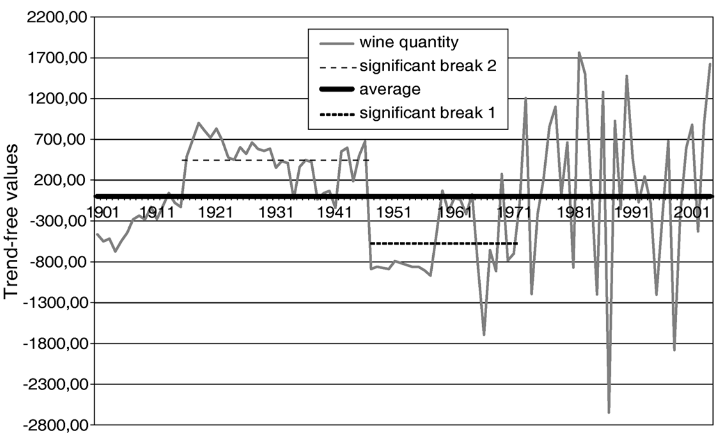

Figure 1 Significant breaks in the wine quantity data set, according to the Makra test. (The significant trend of the 104-year data series has been previously subtracted.)

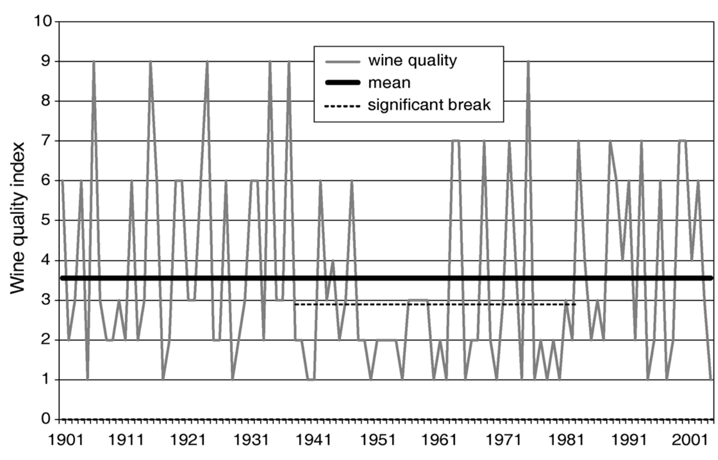

Figure 2 Significant breaks in the wine quality data set, based on the Makra test.

Second, factor analysis was performed on wine quality and the 18 climatic variables. Ten factors were retained, explaining 82.0% of the total variance of the original variables. Applying the special transformation, wine quality, as the resultant variable, had a significant relation at 99% probability level ($x_{0.01} = 0.254$) with four climatic variables: hours of sunshine in May, precipitation in June, mean temperature in May, and precipitation in September (ranked by importance). Furthermore, wine quality had a significant connection with precipitation in May at 95% probability level ($x_{0.05} = 0.195$). High values of hours of sunshine and high mean temperature in May, high precipitation in September, low precipitation in May and June, and low hours of sunshine in August improved wine quality. High values of hours of sunshine and high mean temperature in May favored the development of shoots and ensured an undisturbed period for blossoming. High precipitation in May and June revealed adverse consequences as it is detrimental to blooming and favorable for pests. Higher quality wines occur in dry years with high temperatures (Grifoni et al. 2006). Warm, dry summers result in high sugar and low acid levels at harvest, leading to higher quality wines (Jones and Storchmann 2001). High-quality winegrapes depend on the ability of maintaining mild to moderate levels of water stress in the crop during the growing season (Moller et al. 2007). While the results show that a warm and sunny August is important for quality, potentially more important is a wet September by which *Botrytis cinerea* can spread more effectively, causing berry evaporation, shrivel, and the high sugar content needed for the aszú wine production. The role of the remaining climatic variables could not be determined based on their factor loadings because of low significance (Table 3).

In order to detect the strength of relation between the climatic variables and wine quantity and quality, χ^2 test independence analysis was also performed. There was a significant connection between, on the one hand, hours of sunshine in May, June, and July and, on the other hand, wine quantity. A significant connection between wine quality and mean temperature and hours of sunshine in May was also detected (Table 4).

After performing χ^2 test independence analysis and special transformation for the factors retained, in some cases different climatic variables revealed significant connections with wine quantity and quality (Table 4). However, if a statistical problem is analyzed by different methods, certain difference of the results are expected. In the case of factor analysis, the factors are linear combinations of the original variables; hence, they display a linear relation between the variables. After performing the special transformation, factors 2–10 are transformed to factor 1 with their factor loadings that belong to the climatic variables and wine quantity and quality. In this way, the effect of climatic variables on wine quantity and quality as well as the rank of the climatic variables are presented (Table 2 and Table 3). To determine the reliability of the results, special transformation on factor 2 and factor 3 was also

Table 2 Special transformation. Effect of the climatic variables (mean temperature, precipitation, and sunshine hours) on wine quantity as resultant variable and the rank of the climatic variables on their factor loadings transformed to factor 1.

Climatic variables	Factor 1'[a]	Factor 2'[b]	Factor 3'[c]	Rank on factor 1'[a]
Wine quantity	0.884	-0.879	-0.877	–
Mean temp, April, $T_{mean, Apr}$	-0.161	0.172	0.175	8
Precipitation, April, P_{Apr}	-0.178	0.184	0.186	7
Sunshine hours, April, S_{Apr}	-0.002	0.023	0.023	18
Mean temp, May, $T_{mean, May}$	0.088	-0.078	-0.073	12
Precipitation, May, P_{May}	-0.086	0.083	0.082	13
Sunshine hours, May, S_{May}	**0.457**[d]	**-0.472**[d]	**-0.476**[d]	**1**[d]
Mean temp, June, $T_{mean, Jun}$	-0.106	0.194	0.193	10
Precipitation, June, P_{Jun}	**-0.314**	**0.262**[d]	**0.264**[d]	**3**[d]
Sunshine hours, June, S_{Jun}	*0.242*[d]	-0.150	-0.151	*4*[d]
Mean temp, July, $T_{mean, Jul}$	-0.143	0.158	0.158	9
Precipitation, July, P_{Jul}	0.017	-0.011	-0.009	16
Sunshine hours, July, S_{Jul}	0.025	-0.015	-0.020	15
Mean temp, Aug, $T_{mean, Aug}$	0.178	-0.164	-0.168	6
Precipitation, Aug, P_{Aug}	-0.046	0.051	0.051	14
Sunshine hours, Aug, S_{Aug}	**0.315**[d]	**-0.309**[d]	**-0.318**[d]	**2**[d]
Mean temp, Sept, $T_{mean, Sep}$	-0.016	0.028	-0.014	17
Precipitation, Sept, P_{Sep}	*0.238*[d]	*-0.230*[d]	-0.183	*5*[d]
Sunshine hours, Sept, S_{Sep}	0.094	-0.086	-0.140	11

[a]Factor 1': generated as consecutive transformations of factor loadings of the remaining nine factors to factor 1.
[b]Factor 2': omitting factor 1; generated as consecutive transformations of factor loadings of the remaining eight factors to factor 2.
[c]Factor 3': omitting factors 1 and 2; generated as consecutive transformations of factor loadings of the remaining seven factors to factor 3.
[d]Threshold of significance: *italic:* $x_{0.05} = 0.195$; **bold**: $x_{0.01} = 0.254$.

Table 3 Special transformation. Effect of the climatic variables (mean temperature, precipitation, and sunshine hours) on wine quality as resultant variable and the rank of the climatic variables on their factor loadings transformed to factor 1.

Climatic variables	Factor 1'[a]	Factor 2'[b]	Factor 3'[c]	Rank on factor 1'[a]
Wine quality	0.891	0.879	-0.746	–
Mean temp, April, $T_{mean, Apr}$	0.151	0.088	0.074	8
Precipitation, April, P_{Apr}	-0.090	-0.072	-0.069	12
Sunshine hours, April, S_{Apr}	-0.075	-0.151	**0.320**[d]	14
Mean temp, May, $T_{mean, May}$	**0.329**[d]	**0.261**[d]	0.009	**3**[d]
Precipitation, May, P_{May}	*-0.243*[d]	*-0.221*[d]	-0.026	*5*[d]
Sunshine hours, May, S_{May}	**0.392**[d]	**0.336**[d]	**-0.309**[d]	**1**[d]
Mean temp, June, $T_{mean, Jun}$	0.049	-0.042	0.154	15
Precipitation, June, P_{Jun}	**-0.373**[d]	**-0.286**[d]	*0.245*[d]	**2**[d]
Sunshine hours, June, S_{Jun}	0.031	-0.055	0.123	16
Mean temp, July, $T_{mean, Jul}$	0.012	-0.084	0.056	18
Precipitation, July, P_{Jul}	-0.193	-0.114	0.177	*7*[d]
Sunshine hours, July, S_{Jul}	-0.139	-0.203	0.063	10
Mean temp, Aug, $T_{mean, Aug}$	-0.124	*-0.238*[d]	*0.195*[d]	11
Precipitation, Aug, P_{Aug}	0.142	*0.206*[d]	-0.096	9
Sunshine hours, Aug, S_{Aug}	*-0.206*[d]	**-0.271**[d]	0.121	*6*[d]
Mean temp, Sept, $T_{mean, Sep}$	0.022	-0.037	-0.293	17
Precipitation, Sept, P_{Sep}	**0.302**[d]	**0.335**[d]	0.002	**4**[d]
Sunshine hours, Sept, S_{Sep}	0.084	0.058	**-0.389**[d]	13

[a]Factor 1': generated as consecutive transformations of factor loadings of the remaining nine factors to factor 1.
[b]Factor 2': omitting factor 1; generated as consecutive transformations of factor loadings of the remaining eight factors to factor 2.
[c]Factor 3': omitting factors 1 and 2; generated as consecutive transformations of factor loadings of the remaining seven factors to factor 3.
[d]Threshold of significance: *italic:* $x_{0.05} = 0.195$; **bold**: $x_{0.01} = 0.254$.

Table 4 Parameters of linear regressions for climatic variables, which significantly influence wine quantity and quality, obtained using χ^2 test independence analysis and factor analysis special transformation on factor loadings (transformed to factors 1, 2, and 3): coefficients, constants, and goodness of fit.

	Wine quantity				Wine quality			
		Factor analysis, transformed factor loadings[a]				Factor analysis, transformed factor loadings[a]		
Climatic variables	**χ^2 test**	**1***	**2***	**3***	**χ^2 test**	**1***	**2***	**3***
Mean temp, April, $T_{mean, Apr}$								
Precipitation, April, P_{Apr}								
Sunshine hours, April, S_{Apr}								-0.005
Mean temp, May, $T_{mean, May}$					0.260	0.126	0.222	0.238
Precipitation, May, P_{May}						-0.011	-0.008	-0.009
Sunshine hours, May, S_{May}	2.746	2.245	2.417	2.417	0.006	0.005	0.006	0.005
Mean temp, June, $T_{mean, Jun}$			-78.828	-78.828				
Precipitation, June, P_{Jun}		-1.188	-3.112	-3.112		-0.009	-0.010	-0.010
Sunshine hours, June, S_{Jun}	3.307	2.278						
Mean temp, July, $T_{mean, Jul}$								
Precipitation, July, P_{Jul}						-0.002	-0.006	-0.008
Sunshine hours, July, S_{Jul}	-0.080						-0.004	-0.005
Mean temp, Aug, $T_{mean, Aug}$			36.010	36.010			-0.312	-0.304
Precipitation, Aug, P_{Aug}							0.005	-0.004
Sunshine hours, Aug, S_{Aug}		4.543				-0.013	-0.002	-0.002
Mean temp, Sept, $T_{mean, Sep}$								0.080
Precipitation, Sept, P_{Sep}		2.125	1.886	1.886		0.010	0.008	0.013
Sunshine hours, Sept, S_{Sep}								0.006
Constant	-1336.954	-2163.397	1268.471	447.870	-1.870	4.573	7.748	6.058
Goodness of fit, using χ^2 test[b]	0.079				0.074			
Goodness of fit, factor 1[b]		0.096				0.116		
Goodness of fit, factors 1 and 2[b]			0.119				0.127	
Goodness of fit, factors 1, 2, and 3[b]				0.074				0.117

[a]Coefficients of the linear regression on the significant factor loadings transformed to: 1*, factor 1; 2*, factors 1 and 2; 3*, factors 1, 2, and 3.
[b]Goodness of fit (adjusted R square) with significant climatic variables (R square: relative explained variance).

applied. These factors also had relatively high information content, although not as high as factor 1. Comparison of factor loadings of different climatic variables for factors 1 to 3 showed that factor 2 complemented results obtained with factor 1. Consideration of factor 3 is, however, unnecessary for wine quantity (Table 4; see significant factor loadings of climatic variables for factors 2 and 3). Similar conclusion can be drawn for wine quality; hence, factor 3 is neglected when calculating best regression (see adjusted R squares in Table 4).

The goodness of the results (goodness of linear approximations calculated using the above two methods) can be compared if linear regression is performed for the wine parameters as the resultant variables, with the significant climatic variables obtained on the one hand by χ^2 test and, on the other, by applying special transformation and then comparing the adjusted R squares of the approximations. Based on the above, when performing linear regression, significant climatic variables belonging to factor 3 were omitted from further consideration both for wine quantity and wine quality. Only significant climatic variables belonging to factors 1 and 2 were considered for both resultant variables. According to the calculations, the adjusted R square of the linear approximation for both wine quantity and quality was higher for the significant climatic variables obtained from factor analysis special transformation (Table 4; final three rows). In other words, significant variables obtained by factor analysis better explain linear relationships between them and the resultant variables than those obtained by the χ^2 test. However, rejecting independence (χ^2 test) refers to both linear and curvilinear relationships. Relative explained variance (R square) resulted in low values for the approximations using both methods (Table 4).

Generalized correlations (r_G) are much higher than linear correlations, particularly for wine quantity (Table 5), but it is difficult to establish any correspondence between magnitudes of generalized correlations and linear correlations. A statistical test (Dionisio et al. 2004) was applied to determine whether an r_G differed significantly from zero. Generalized correlations clearly show relationships between climatic variables and wine quantity and quality.

These relationships, however, have complex forms weakly reproduced by linear or other simple functions.

On the basis of Pearson's χ^2 test, special transformation, and generalized correlation, climatic variables that were important in influencing wine quantity and quality were determined (Table 6). Hours of sunshine in May and June for wine quantity and mean temperature and hours of sunshine in May for wine quality are the only common variables that significantly influenced the resultant variables according to all three methods. For wine quantity, sunshine hours in July and August and precipitation in September are common variables for two of the methods. For wine quality, precipitation in July and sunshine hours in August are common variables for two of the methods.

Factor analysis of the 18 climatic variables resulted in nine factors explaining 80.2% of the total variance. A cluster analysis was applied to the 9-factor factor score time series to classify years objectively into groups of similar climate variables. Seven clusters (vintage climate types) were retained for the 6-month periods of the 104 years examined. The mean values of the climatic variables were then examined and those of wine quantity and quality were calculated for the seven characteristic vintage climate types (Table 7).

Pearson's χ^2 test was applied to determine whether wine quantity and quality depend on vintage climate types (Table 8). If the null hypothesis of independence is fulfilled, then neither wine quantity nor wine quality depend on vintage climate types; while, in the reverse case, there is relation between them. As a result, the likelihood of independence between wine quantity and vintage climate types is extremely low, below 0.07, whereas between wine quality and vintage climate types it is very high, above 0.34 (Table 8). Hence, we may conclude that wine quality was independent of vintage climate types. On the other hand, wine quantity was closely related to vintage climate types (at 90% probability level, but not at 99% or 95% levels).

Analysis of vintage climate type dependencies revealed that the highest wine quantity occurred during years dominated by vintage climate type 5. During the period examined, 14.4% of the years belonged to type 5—the most characteristic vintage climate type as it was dominated by three climatic variables that played important roles in wine quantity (Table 2, Table 7). The lowest wine quantity was associated with vintage climate type 1 (24.0%), with only one climatic variable having substantial influence on wine quantity (Table 2,

Table 5 Generalized and linear correlations of wine quantity and quality with climatic variables.

	Wine quantity		Wine quality	
Climatic variables	**Generalized**	**Linear**	**Generalized**	**Linear**
Mean temp, April, $T_{mean, Apr}$	**0.643**[a]	0.101	*0.289*[a]	0.140
Precipitation, April, P_{Apr}	0.038	-0.005	0.044	-0.022
Sunshine hours, April, S_{Apr}	**0.912**[a]	0.108	0.159	-0.021
Mean temp, May, $T_{mean, May}$	*0.290*[a]	0.142	**0.449**[a]	**0.260**[b]
Precipitation, May, P_{May}	0.016	-0.100	0.255	*-0.227*[b]
Sunshine hours, May, S_{May}	**0.772**[a]	0.120	**0.371**[a]	*0.228*[b]
Mean temp, June, $T_{mean, Jun}$	0.231	0.039	0.271	0.074
Precipitation, June, P_{Jun}	0.078	-0.060	0.249	*-0.207*[b]
Sunshine hours, June, S_{Jun}	**0.710**[a]	*0.209*[b]	*0.294*[a]	0.041
Mean temp, July, $T_{mean, Jul}$	**0.648**[a]	-0.001	0.229	0.041
Precipitation, July, P_{Jul}	0.085	0.069	0.271	-0.110
Sunshine hours, July, S_{Jul}	**0.707**[a]	-0.068	*0.326*[a]	-0.044
Mean temp, Aug, $T_{mean, Aug}$	**0.364**[a]	*0.250*[b]	0.164	-0.094
Precipitation, Aug, P_{Aug}	0.112	-0.084	0.095	0.089
Sunshine hours, Aug, S_{Aug}	**0.758**[a]	*0.199*[b]	*0.309*[a]	-0.110
Mean temp, Sept, $T_{mean, Sep}$	0.222	-0.029	**0.383**[a]	-0.031
Precipitation, Sept, P_{Sep}	*0.335*[a]	0.011	0.192	0.182
Sunshine hours, Sept, S_{Sep}	**0.661**[a]	0.052	0.228	0.018

[a]Generalized correlation, thresholds of significance: *italic*: $x_{0.05} = 0.273$; **bold**: $x_{0.01} = 0.351$.
[b]Linear correlation, thresholds of significance: *italic*: $x_{0.05} = 0.195$; **bold**: $x_{0.01} = 0.254$.

Table 6 Comparison of climatic variables that significantly influence wine quantity and quality on the basis of Pearson's χ^2 test, special transformation, and generalized correlation. (•: significant relation at least at 99% probability level.)

	Wine quantity			Wine quality		
Climatic variables	**χ^2 test**	**Special transf.**	**Generalized correlation**	**χ^2 test**	**Special transf.**	**Generalized correlation**
Mean temp, April, $T_{mean, Apr}$			•			
Precipitation, April, P_{Apr}						
Sunshine hours, April, S_{Apr}			•			
Mean temp, May, $T_{mean, May}$			•	•	•	•
Precipitation, May, P_{May}					•	•
Sunshine hours, May, S_{May}	•	•	•	•	•	•
Mean temp, June, $T_{mean, Jun}$						•
Precipitation, June, P_{Jun}		•			•	
Sunshine hours, June, S_{Jun}	•	•	•			
Mean temp, July, $T_{mean, Jul}$			•			
Precipitation, July, P_{Jul}					•	•
Sunshine hours, July, S_{Jul}	•		•			•
Mean temp, August, $T_{mean, Aug}$			•			
Precipitation, August, P_{Aug}						
Sunshine hours, August, S_{Aug}		•	•		•	•
Mean temp, Sept, $T_{mean, Sep}$						•
Precipitation, Sept, P_{Sep}		•	•		•	
Sunshine hours, Sept, S_{Sep}			•			

Table 7). The highest wine quality occurred in years with a higher frequency of vintage climate type 3 (19.2%), with only precipitation in September playing an important role in wine quality (Table 3, Table 7). The lowest wine quality was associated with vintage climate type 2 (13.5%), where two of the six climatic variables that characterized this type (Table 7) were relevant to wine quality (Table 3).

In order to determine the effect of the individual vintage climate types on wine quantity and quality, ANOVA was performed for these two resultant variables. Results show that the means of wine quantity differed with 85.9% probability, while those of wine quality differed with 75.2% probability between the individual vintage climate types. The analysis of variance did not reveal any significant difference, at least at 95% probability level, in the mean values of wine quantity and quality between the individual vintage climate types. The post-hoc Tukey test shows a low vintage climate type difference, with only types 1 and 5 showing differences at the 93% level. These two vintage climate types were presented above, since the highest wine quantity belongs to type 5, while the lowest one belongs to type 1.

Discussion

The relationship between climatic variables and wine quantity and quality was studied using different statistical procedures. In order to assess the effect of different vintage climate types on wine quantity and quality, objective multivariate statistical methods were applied to the climatic,

Table 7 Mean values of climatic parameters and of wine quantity and quality for the years of the individual vintage climate types (clusters) (**bold:** maximum; *italic:* minimum).

Cluster	1	2	3	4	5	6	7
Number of cases (years)	**25**	14	20	*7*	15	10	13
Relative frequency (%)	**24.0**	13.5	19.2	*6.7*	14.4	9.6	12.5
Mean temp, April, $T_{mean, Apr}$ (°C)	*10.1*	10.8	11.1	12.8	11.9	10.6	**12.8**
Precipitation, April, P_{Apr} (mm)	43.9	**61.5**	35.3	30.5	37.5	*25.2*	49.6
Sunshine hours, April, S_{Apr} (hr)	159.9	*143.8*	175.6	187.2	188.3	181.7	**198.7**
Mean temp, May, $T_{mean, May}$ (°C)	**17.6**	*14.6*	17.1	17.3	16.3	15.8	17.2
Precipitation, May, P_{May} (mm)	*42.5*	**81.6**	47.9	53.3	61.9	*57.0*	75.0
Sunshine hours, May, S_{May} (hr)	206.6	166.9	247.8	231.1	**251.0**	*116.7*	180.8
Mean temp, June, $T_{mean, Jun}$ (°C)	**20.4**	18.9	19.0	19.3	20.3	*18.3*	20.2
Precipitation, June, P_{Jun} (mm)	67.5	85.3	79.6	107.2	63.5	**121.8**	*48.6*
Sunshine hours, June, S_{Jun} (hr)	236.0	234.8	223.4	228.2	**280.7**	*207.3*	260.1
Mean temp, July, $T_{mean, Jul}$ (°C)	21.7	*20.4*	20.9	21.4	21.6	21.0	**23.6**
Precipitation, July, P_{Jul} (mm)	60.8	78.6	58.0	**129.4**	67.1	84.3	*38.0*
Sunshine hours, July, S_{Jul} (hr)	261.8	238.8	256.6	*235.7*	278.5	275.6	**303.5**
Mean temp, Aug, $T_{mean, Aug}$ (°C)	20.9	*19.7*	20.7	20.0	**22.5**	20.0	21.6
Precipitation, Aug, P_{Aug} (mm)	58.2	61.4	56.8	**137.8**	*46.1*	60.5	48.7
Sunshine hours, Aug, S_{Aug} (hr)	240.7	248.8	255.2	*230.2*	**288.5**	252.4	248.0
Mean temp, Sept, $T_{mean, Sep}$ (°C)	17.0	16.1	*14.9*	17.1	**17.7**	16.5	17.2
Precipitation, Sept, P_{Sep} (mm)	*29.2*	32.1	**82.0**	56.7	41.2	30.0	51.4
Sunshine hours, Sept, S_{Sep} (hr)	190.0	178.3	*158.6*	194.8	**220.2**	198.2	173.8
Wine quantity (L/ha)	*2232.7*	2454.1	2441.4	2532.7	**3250.2**	2748.1	2544.7
Wine quality (index)	3.8	*2.6*	**4.6**	3.7	3.0	3.2	3.3

Table 8 Frequencies of the wine quantity and wine quality in the vintage climate types (clusters) obtained.

Vintage climate type	Wine quantity (L/ha)					Wine quality[a]						
	725/1663	1663/2600	2600/3538	3538/4476	4476/5413	1	2	3	4	6	7	9
1	7	10	5	2	1	2	8	7	0	3	3	2
2	3	6	2	3	0	6	3	2	0	3	0	0
3	7	7	3	1	2	0	5	4	3	4	1	3
4	2	2	3	0	0	1	3	0	0	2	1	0
5	1	3	6	1	4	3	5	3	1	2	1	0
6	0	6	2	2	0	2	4	1	1	1	0	1
7	1	7	5	0	0	4	3	2	0	1	3	0

[a]1: lowest quality, 9: highest quality.

wine quantity, and wine quality data sets. Objective vintage climate types prevailing over the Tokaj region were determined. Although the procedure is known and has been applied in the technical literature (e.g., Makra et al. 2006), this method can be regarded a new approach for analyzing and interpreting the interannual differences for the mean values of wine quantity and quality in this wine-producing region.

Vintage climate types of similar climate conditions, derived for the 104-year period between 1901–2004, were related to the annual values of wine quantity and quality. However, wine parameters are also affected by other important natural factors (e.g., soil type, slope degree and aspect, latitude), which cannot be directly analyzed via our methodology. These factors are implicitly incorporated in the area averages of wine parameters with fairly uniform spatial distribution in the 104-year period investigated, with the exception of the 1960s and 1970s (see below). Hence these factors may not substantially alter any climate effects.

Significant positive trends have been found in both the temperature of the growing season and the vintage quality ratings in the second half of the 20th century in the majority of the world's highest quality wine-producing regions (Jones et al. 2005). However, a significant increasing trend obtained in our wine production data set was not reflected in the climatic variables analyzed. Our results concern relations between climatic variables; furthermore, a general increase in vintage ratings is in accordance with other studies (Jones and Davis 2000). However, the climatic conditions of better wine quality over the Tokaj region differ from those of other regions (Grifoni et al. 2006, Laget et al. 2008, Soar et al. 2008). In Tokaj, the low mean temperature, hours of sunshine, and high precipitation that occur in September are most important in determining higher-quality wines, in addition to the high mean temperature, low precipitation, and high hours of sunshine that occur in May.

For wine quality, warm years are considered good years in the Tokaj region. In addition, good aszú can be formed from low-quality vines and vice versa. September is a decisive month; if it is relatively cool, cloudy, and rainy, then a good year of aszú is expected. On the other hand, if berries are small, then their skin is thick and *Botrytis cinerea* cannot attack the berries. Moreover, if this fungus infects the skin of the berries too early (the grape is not ripened), aszú will not develop. Hence, it is not easy to determine the precise meteorological conditions for aszú formation. But unlike our findings, Steel and Greer (2008) report on a high predisposition of grapes to non-*Botrytis* bunch rots, most notably bitter rot (*Greeneria uvicola*) and ripe rot (*Colletotrichum acutatum*) in Australia in the wetter eastern sites. In contrast, high light exposures coupled with extremely hot (>35°C) temperatures at inland vineyards cause severe skin damage. This sunburn damage does, however, increase the incidence of *Botrytis* latent infections in grapes. But in the Tokaj region, bunch rots and skin damage contribute to a loss of berry quality, hence detrimentally affecting wine quality.

Our study considered three monthly climate variables: temperature, precipitation, and sunshine duration. Relative humidity might also be considered as a possible factor, but it has a strong correlation with hours of sunshine in Hungary (Wantuch-Dobi 2002). The impact of climate on wine quantity and quality may also be investigated by studying large-scale atmospheric circulation factors and searching for connections between diurnal weather patterns or monthly circulation indices (e.g., NAO) and local wine characteristics. This approach is, however, beyond the scope of our study.

Conclusions

Analysis of the relationship between climatic variables and the resultant variables (wine quantity and quality) using different statistical methods confirmed that the most important factors of wine quantity in the Tokaj region are hours of sunshine in May, June, July, and August and precipitation in September. On the other hand, mean temperature, precipitation, and hours of sunshine in May and September play a basic role in wine quality, as does precipitation in July and hours of sunshine in August. The weather in September is very important for aszú wine production, whereby increased rainfall during this month leads to higher occurrence of *Botrytis cinerea* and more concentrated grape sugar and flavor levels. It was also found that significant climatic variables obtained by factor analysis better explained linear relationships between them and the resultant variables than those obtained by the χ^2 test. However, rejecting independence (χ^2 test) refers to a relationship of any kind of character (not only linear).

The classification of vintage climate types, as the most homogeneous groups of climatic factors, was more effective in separating the mean values of wine quantity than those of wine quality. The application of objective vintage climate types in classifying crop yields and wine quality may be recommended as a new tool for quantifying environmental cause-effect relationships.

Literature Cited

Anderberg, M.R. 1973. Cluster Analysis for Applications. Academic Press, New York.

Balassa, I. 1991. Tokaj-Hegyalja szőleje és bora (Vine and wine of Tokaj-Hegyalja; in Hungarian). Tokaj-Hegyaljai ÁG Borkombinát, Tokaj.

Bartholy, J., R. Pongracz, G. Gelybo, and P. Szabo. 2008. Analysis of expected climate change in the Carpathian Basin using the PRUDENCE results. Idojaras (special issue) 112(3-4):249-264.

Bolla, M., and A. Krámli. 2005. Theory of Statistical Conclusions (in Hungarian). Typotex Kft, Budapest.

Caprio, J.M., and H.A. Quamme. 2002. Weather conditions associated with grape production in the Okanagan Valley of British Columbia and potential impact of climate change. Can. J. Plant Sci. 82(4):755-763.

Czelnai, R., L.S. Gandin, and W.I. Zachariew. 1976. Statistische Struktur der Meteorologischen Felder (in Russian and German). 364 pp. OMSZ Hivatalos Kiadványai, Budapest.

Danielides, V., C.S. Nousia, D. Patrikakos, A. Bartzokas, C.J. Lolis, H.J. Milionis, and A. Skevas. 2002. Effect of meteorological parameters on acute laryngitis in adults. Acta Oto-Laryngol. 122(6):655-660.

Dionisio, A., R. Menezes, and D.A. Mendes. 2004. Mutual information: A measure of dependency for nonlinear time series. Phys. A-Stat. Mech. Appl. 344(1-2):326-329.

Ezekiel, M., and K.A. Fox. 1970. Correlation and Regression Analysis. Linear and Non-linear Methods (in Hungarian). Közgazdasági és Jogi Könyvkiadó, Budapest.

Fischer, G., and J. Roppert. 1965. Ein Verfahren der Transformationsanalyse faktorenanalytischer Ergebnisse. *In* Lineare Strukturen in Mathematik und Statistik unter besonderer Berücksichtigung der Faktoren- und Transformationsanalyse (in German). Verlag Physica, Wien-Würzburg, Austria-Germany.

Gál, A. 2004. Natural conditions of viniculture in Tokaj-Hegyalja. *In* Proceedings for the Third International Conference on Application of Natural, Technical and Economic Sciences. J. Puskás (ed.), p. 52. Berzsenyi Dániel College, Szombathely, Hungary.

Granger, C.W., E. Maasoumi, and J. Racine. 2004. A dependence metric for possibly nonlinear processes. J. Time Ser. Anal. 24:649-669.

Grifoni, D., M. Mancini, G. Maracchi, S. Orlandini, and G. Zipoli. 2006. Analysis of Italian wine quality using freely available meteorological information. Am. J. Enol. Vitic. 57:339-346.

Hair, J.F., R.E. Anderson, R.L. Tatham, and W.C. Black. 1998. Multivariate Data Analysis. Prentice Hall, New Jersey.

Jahn, W., and H. Vahle. 1968. Die Faktoranalyse und ihre Anwendung. Verlag die Wirtschaft, Berlin.

Jolliffe, I.T. 1986. Principal Component Analysis. Springer-Verlag, New York.

Jolliffe, I.T. 1990. Principal component analysis: A beginner's guide. I. Introduction and application. Weather 45:375-382.

Jolliffe, I.T. 1993. Principal component analysis: A beginner's guide. II. Pitfalls, myths and extensions. Weather 48:246-253.

Jones, G.V., and R.E. Davis. 2000. Climate influences on grapevine phenology, grape composition, and wine production and quality for Bordeaux, France. Am. J. Enol. Vitic. 51:249-261.

Jones, G.V., and K.H. Storchmann. 2001. Wine market prices and investment under uncertainty: An econometric model for Bordeaux Crus Classes. Agr. Econ. 26(2):115-133.

Jones, G.V., M.A. White, O.R. Cooper, and K. Storchmann. 2005. Climate change and global wine quality. Climatic Change 73(3):319-343.

Laget, F., J.L. Tondut, A. Deloire, and M.T. Kelly. 2008. Climate trends in a specific Mediterranean viticultural area between 1950 and 2006. J. Int. Sci. Vigne 42(3):113-123.

Lobell, D.B., K.N. Cahill, and C.B. Field. 2007. Historical effects of temperature and precipitation on California crop yields. Climatic Change 81(2):187-203.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. Proc. Nat. Inst. Sci. India 12:49-55.

Makra, L., S. Horváth, R. Pongrácz, and J. Mika. 2002. Long term climate deviations: An alternative approach and application on the Palmer drought severity index in Hungary. Phys. Chem. Earth. 27:1063-1071.

Makra, L., J. Mika, A. Bartzokas, R. Béczi, E. Borsos, and Z. Sümeghy. 2006. An objective classification system of air mass types for Szeged, Hungary with special interest to air pollution levels. Meteorol. Atmos. Phys. 92(1-2):115-137.

Moller, M., V. Alchanatis, Y. Cohen, M. Meron, J. Tsipris, A. Naor, V. Ostrovsky, M. Sprintsin, and S. Cohen. 2007. Use of thermal and visible imagery for estimating crop water status of irrigated grapevine. J. Exp. Bot. 58(4):827-838.

Soar, C.J., V.O. Sadras, and P.R. Petrie. 2008. Climate drivers of red wine quality in four contrasting Australian wine regions. Aust. J. Grape Wine Res. 14:78-90.

Steel, C.C., and D.H. Greer. 2008. Effect of climate on vine and bunch characteristics: Bunch rot disease susceptibility. *In* Proceedings of the International Symposium on Grape Production and Processing. Book series: Acta Hortic. 785:253-262.

Tukey, J.W. 1985. The problem of multiple comparisons (1953) (unpublished manuscript). *In* The Collected Works of John W. Tukey. Vol. II. Time Series, pp. 1965-1984. Wadsworth, Monterey, CA.

Vitányi, B. 2004. Energy balance of vine in Tokaj-Hegyalja. *In* Proceedings for the Third International Conference on Application of Natural, Technical and Economic Sciences. J. Puskás (ed.), p. 65. Berzsenyi Dániel College, Szombathely, Hungary.

Wantuch-Dobi, I. 2002. Multi-dimensional stochastic modeling of daily meteorological datasets (in Hungarian). Ph.D. thesis, Eötvös Loránd, University, Budapest.

Ward, J.H. 1963. Hierachical grouping to optimize an objective function. J. Am. Stat. Ass. 58:236-244.

WGRIT. 2005. Wine quantity and quality and acreage data for Tokaj-Hegyalja region (manuscript, in Hungarian). Wine-Growing Research Institute of Tarcal, Hungary.