International Environmental Modelling and Software Society (iEMSs)
2010 International Congress on Environmental Modelling and Software
Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada
David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.)
http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

Modelling the seasonal climate effects on grapevine yield at different spatial and unconventional temporal scales

Subana Shanmuganathan, Philip Sallis and Ajit Narayanan
Geoinformatics Research Centre (GRC), (www.geo-informatics.org)
School of Computing and Mathematical Sciences, Auckland University of Technology, New Zealand. subana.shanmuganathan@aut.ac.nz

Abstract: The paper briefly outlines recent conventional approaches to modelling/ predicting seasonal climate effects on grapevine phenology and wine quality using weather data from national meteorological institutions and yield/ vintage ratings provided by sommeliers. The seasonal variability in climatic conditions can cause shifts in grapevine growth stages, phenological events, which in turn affect the formation and ratio of grape berry components, such as sugar, and pro-phenols, that give the colour, aroma and flavour attributes to the vintage relating to its wine style. Although winemaker ability is considered to be the major determinant of the quality of wine, the excellence of any vintage could still be enhanced considerably with grapes ripened under ideal weather conditions; this is evidenced by better vintage ratings and price hikes associated with better weather conditions in the past. Hence, viticulturists and enologists continuously strive to further scientific understanding on climate effects to increase yield and wine quality. Recent studies reveal that conventional rigorous statistical data analysis methodologies used long term data on crop yield/ wine quality and weather conditions for studying the associations between the variables and at regional scales. This data requirement impedes any such meaningful study on vineyards established recently, at a micro scale. The Geoinformatics Research Centre (GRC) approaches investigated to overcome this dilemma with data covering only a decade and at a vineyard scale are discussed. Climate data, such as monthly maximum, minimum and average temperature, monthly total rainfall, occurrence of frost days and growing degree days (GDD) (base 10^{o}C) along with yield is analysed using data mining techniques, such as clustering, then with regression and discriminant methods. The results show potential for predicting future yield/ wine quality under current weather conditions that could enhance winegrower ability to improve practices for better outcome from the vineyard in terms of yield quality and quantity.

***Keywords**:* Year-to-year climate variability, grapevine phenology and wine vintage

1. BACKGROUND

Modelling the seasonal climate effects on grapevine phenology and wine quality has agricultural as well as economic significance. The seasonal changes in climatic conditions can cause shifts in grapevine growth stages that are observable in terms of phenological events, such as budburst, *floraison*, *veraison*, harvest and then in yield as well. In fact, the seasonal changes to a greater extent can influence the formation and ratio (at favourable levels) of sugar and pro-phenols in grapes that in turn contribute towards wine alcohol content, colour, aroma and flavour of the vintage produced from the grapes of that season, of course this is in addition to the grapevine varietal genetics or '*cultiva*' factors. Despite the fact that the ultimate quality of the vintage is determined by winemaker talent and experience, many studies have proven that ideal climatic conditions lead to better wine quality (with higher wine ratings by sommeliers) and hence price hikes as discussed in Jones [2007]. The literature reviewed for this research reveals that in all such recent studies based on conventional approaches to analysing climate and wine rating data, analysts had used data at regional scales and over three decades in time span. This data requirement impedes the use of such methods with vineyard yield data covering shorter time span. To overcome this issue GRC researchers investigated into using novel approaches based on data mining techniques, to analyse data with shorter time span and the initial results of the research showed potential. The results arrived at quantify the monthly average temperature range (maximum, minimum and average temperatures), rainfall and occurrence of frost days, GDD (base 10^{o}C) that are known to be associated with grapevine yield of a vineyard in northern New Zealand. Hence, the approach being investigated is considered to be useful for analysing data sets that are at different spatial and at unconventional temporal scales and this paper details the methodology as well as results arrived at from this research.

2. SEASONAL CLIMATE EFFECTS ON VITICULTURE AND ENOLOGY

2.1 Literature

The literature reviewed for this study in modelling the seasonal climate effects on viticulture and enology under different climate regimes, reveals that the recent trend in this regard has been on analysing the effects of the phenomenon (natural or manmade), on yield / wine quality at macro scales and with date over three decades. The section looks at some of the major modelling approaches that have been successful at this scale.

IPHEN project as described by Mariani, et al., (2007) was initiated with an aim to produce and broadcast phonological maps for Italian wine varieties initially and then to extend the mapping to other fruits with agro-meteorological data. The maps were developed using a base simulation model originally developed to produce fortnightly maps of phonological phases for two widely grown Italian grapevine varieties, namely, the early *Chardonnay* and the late *Cabernet sauvignon.* The base simulation model uses summed values of heating normal hours instead of classical GDD and a detailed Digital Elevation Model (DEM) with pixels of 2 x 2 km. Thermal fields of maximum and minimum temperatures created by applying geostatistical procedures to National Agrometeorological Network were then calibrated using bibliographic data obtained from different wine regions of Italy. The model is corrected on a regular basis before the results are broadcast via the Internet (www.ucea.it).

Meanwhile, Tonietto & Carbonneau, (2004) talked about a multicriteria climate classification system, a methodology developed to describe the climate of vineyards on a macroclimate scale for the wine regions in the world using three synthetic climate indices relating to viticulture. The three indices used being 1) dryness index (DI) that in general used as an indicator of the potential water balance of the soil of Rio's index, but in the study used as a measure of the level of presence-absence of dryness, 2) heliothermal index for Huglin's heliothermal index, a coefficient of the thermal component that expresses the mean day length in relation to the latitude and 3) cool night index to describe night temperature during berry maturation. The three indices were portrayed to be viticulture descriptors and complementing each other. The authors concluded that the multicriteria climate system called as *Geoviticulture,* to be a complete one and could be used to represent the variability of the viticulture climates in all the different wine regions in the world, as the system was built with elements that could represent differences in grapevine varieties, vintage quality (sugar, colour and aroma) and styles/ appellations in wine. The system was initially presented for 97 grape-growing regions in 29 counties, as a research tool for classifying grape-growing regions and wine making that could be applied to mostly at larger scales i.e., world's wine regions or intra-annual variability within a wine region.

Lobell, et al., (2007) studied the yield-climate relationships for forecasting annual crop production and also for projecting the impact of future climate changes on various crops. In this study, authors used 1980-2003 data on annual yield of 12 major Californian crops and climate (minimum, maximum temperature and precipitation) to model the climate effects on the crop yields. The crops studied were; wine grapes, lettuce, almonds, strawberries, table grapes, hay, oranges, cotton, tomatoes, walnuts, avocados and pistachios. Regressions to find the correlations between yield and climate data for each of the crop were performed and from which fairly simple equations were developed that explained more than two third of the observed yield variance with only 2-3 climate variables, that were selected by explorative data analysis methodologies and described as most important factors.

Ashenfelter, (1995) established the correlations between the price of different vintages and the season's weather data that produced the vintages with an example set of French red *Bordeaux* wines (as judged by the prices of mature wines). The logarithm of vintage price was considered as surrogate for yield / quality of vintages. The age of vintage was regressed against some selected weather variables, such as temperature, (during growing season i.e., April-September), rain in September and August, rain in the months preceding the vintage i.e., October-March, average temperature in September R_2 (root mean squared error) for vintages of 1952-1980 (excluding 1954 and 1956, as these wines were rare, the two vintages being considered as the poorest in the decade).

Jones, (2007) discussed of climate and global wine quality factors and elaborated upon a study on a year-to-year comparison over a ten year period. The study included a description of wine quality factors in juxtaposition with prices and vintage ratings. Citing many earlier studies the author of this work pointed out that the analysis of the relationships between climate variables and wine prices to be based on an underlying hypothesis that beneficial climate conditions would invariably improve the wine quality and that in the past these had in turn led to short term price hikes. The paper also reflected the fact that the unavailability of consistent price data for multiple regions and with different styles over many years to be a shortcoming for any complete analysis/ study on long term effects. Furthermore, argued that the vintage ratings to be a strong determinant of the annual economic success of a wine region based on the work in Nemani, (2001) but then went on to say that the ratings could be determinants of wine quality not necessarily a predictor based on Ashenfelter, (2000) where ratings were described to be reflective of wine somewhat in an indirect way i.e., they had the same weather factors documented to be the determinants of the same wine quality.

So far the section looked at the approaches to modelling climate change effects on viticulture and wine quality with data covering over three decades. As far as New Zealand is concerned data available on wine production covers only short periods of time, i.e., less than a decade. The main reason for this being the chequered past of the New Zealand wine industry; it is since the last decade, certain wine appellations, such as *Sauvignon Blanc* of Marlborough, *Pinot Noir* of Central Otago, *Chardonnay* of North Island, became famous as stated by Cooper (2008) hence the lack of long term data impedes modelling climate effects with conventional approaches. The paper details on the novel approaches being investigated to overcome this impediment, as part of the GRC's overarching project called *Eno-Humanas* that is aimed at building models to studying the effects of independent factors, such as environmental and climate conditions on dependent factors in viticulture and its products, such as shifts in phenological events, berry ripening process, yield and wine quality. The major issue in the *Eno-Humnas* project is establishing the links between the independent factors that consist of more precise data and the dependent factors with less precise data, such as wine quality-which is arguably a subjective matter relating to human sensory perception described by sommeliers about wine taste, mainly about wine colour, aroma, mouth feel and aftertaste. Further details on GRS's projects relating to wine sensory perception are discussed in (Shanmuganathan, et al., 2009 a & b)

2.2 The methodology

The vineyard studied in this research was established in 1996 hence, with this ten year yield data, recent popular conventional methods that require long-term grapevine phenology could not be performed. To overcome this issue, this yield data is analysed initially using self-organising map[1] based data clustering (explorative) and then with statistical methods, (regression and discriminant). Weather data used in the study inlude concurrent monthly climate average values, calculated by National Institute of Water and

1 . A self-organising map is an artificial neural network (ANN) based on an unsupervised algorithmic learning. SOM based data clustering and mapping are useful in projecting multidimensional data sets onto low (1 or 2-D) displays enabling the discovery of new knowledge in the form of patterns/relationships in the maps.

Atmosphere (NIWA) using one of its meteorological station's daily weather recordings, extracted via the institution's web portal (NIWA, 27 September 2009).

2.3 The data

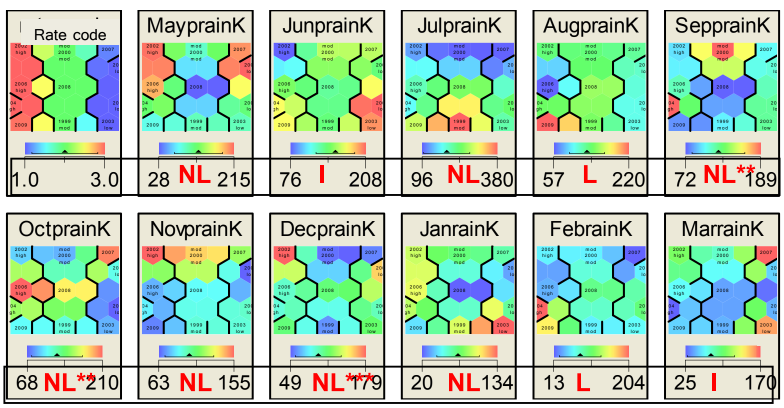
Yield data being modelled is from a vineyard in northern New Zealand and covers from 1997-2006. The vineyard data was classified into *low, moderate* and *high* yield years. Data on monthly maximum, minimum and average temperature, total rainfall, occurrence of frost days, GDD (base 10^{o}C) for this time period, was extracted from NIWA's web portal. This monthly data was calculated using daily weather data logged at Henderson River Pk, (36.85539S, 174.62383E) near the vineyard where the yield data comes from.

3. RESULTS AND DISCUSSION

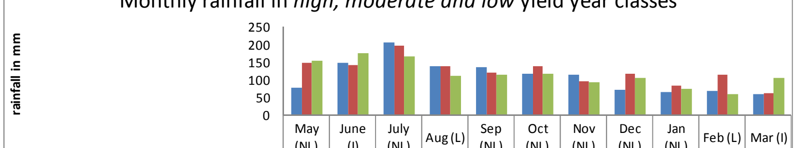
The section elaborates upon the results of data mining and statistical methodologies.

3.1 Data mining

A self-organising map (SOM) was created using a commercial software package called Viscovery[2], with the vineyard yield class and monthly climate data to look for relationships between the two sets of variables, (dependent and independent). In the SOM, the year yield class was given higher priority to favour clustering on this factor (*low, moderate and high*) so that the final mapping enhances the visualisation of any difference/s between the yield year classes and various climate variables being studied. Hence, in the SOM component plane of yield (in the top left corner of fig 1 a), *high, moderate (mod) and low* yield years with rate codes (3, 2 and 1) are clustered and mapped in the left, centre and right respectively. The corresponding monthly rainfall component planes (MayprainK-MarrainK) show the distribution of each of these variables in relation to the yield year classes. This monthly rainfall distribution is also displayed in a bar chart graph (fig 1b).



Monthly rainfall in *high, moderate and low* yield year classes



	May (NL)	June (I)	July (NL)	Aug (L)	Sep (NL)	Oct (NL)	Nov (NL)	Dec (NL)	Jan (NL)	Feb (L)	Mar (I)
C1: 1998 mod 2000 mod, 2008, 1999 mod R^2 = 0.8903	76.2	147.3	206.6	139.5	136	117.1	114.2	69.8	65.2	69	57.6
C2: 2002 high, 2006 high, 2004 high, 2009 R^2 = 0.7383	148.9	143	196.1	138.1	121	139.6	95.7	116.7	83.1	113.8	61
C3: 2007, 2005 low, 2001 low, 2003 low R^2 = 0.8057	155.2	175.7	166.6	110.9	113.3	116.7	91.3	105.5	74.4	58.1	105

L: Linear (increase low to high); NL: Non-linear and I: Inverse correlation

Figure 1 a: SOM component planes, each showing the range for that variable in each map unit/cluster. For example, February rainfall 58.1mm, 69mm and 113.8mm relate to *low,*

[2] Viscovery SOMine is a commercial software package that is very useful in explorative data mining, visual cluster analysis, statistical profiling, segmentation and classification. The data mapping is based on Kohone self-organizing maps (SOMs) and classical statistics in an intuitive workflow environment (Kohonen, 2001)

moderate and *high* yield year classes respectively b: graph showing the monthly rainfall distribution in *high, moderate* and *low* yield. ** regression and *** discriminant predictors.

International Environmental Modelling and Software Society (iEMSs)
2010 International Congress on Environmental Modelling and Software
Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada
David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.)
http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

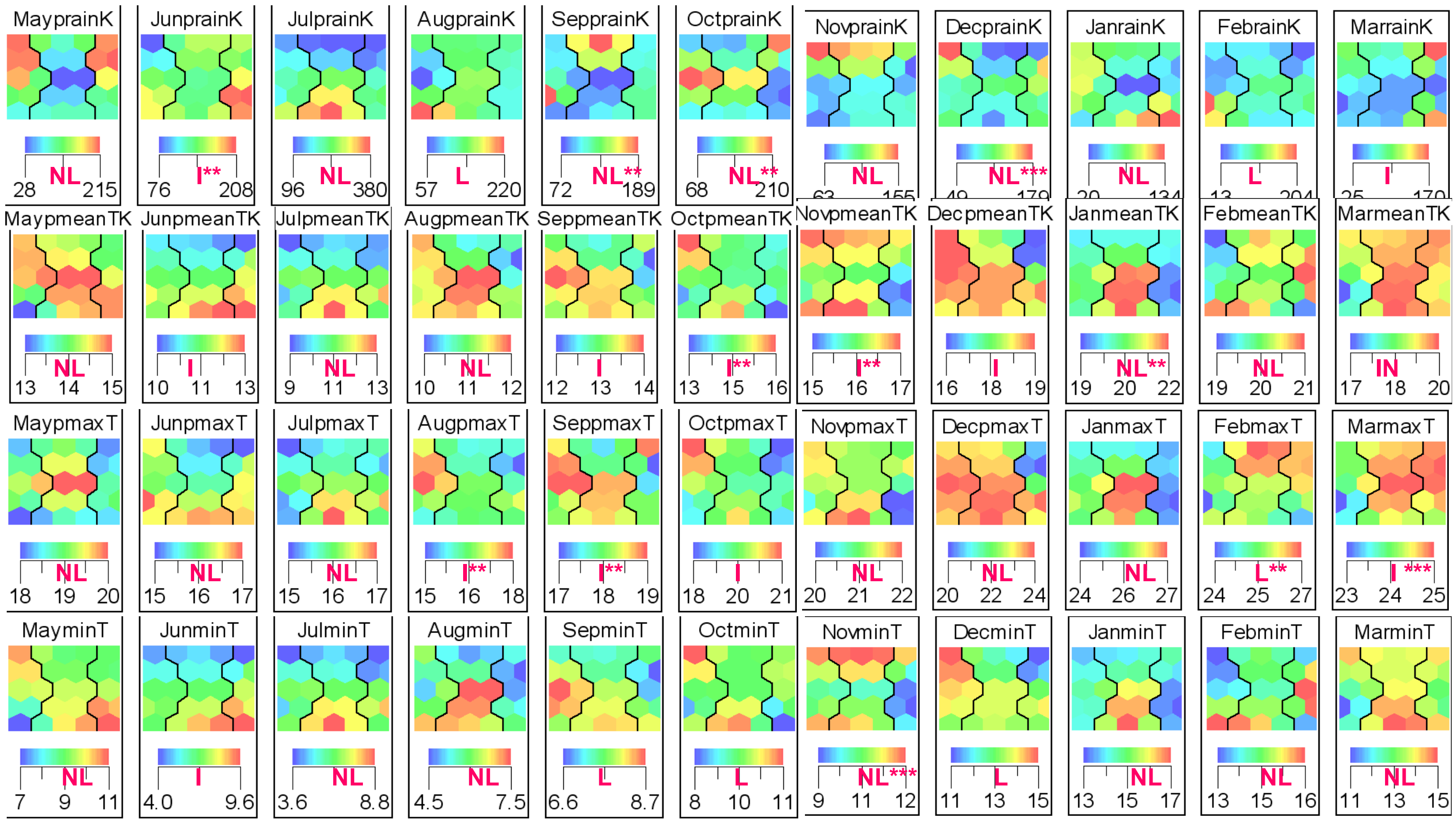
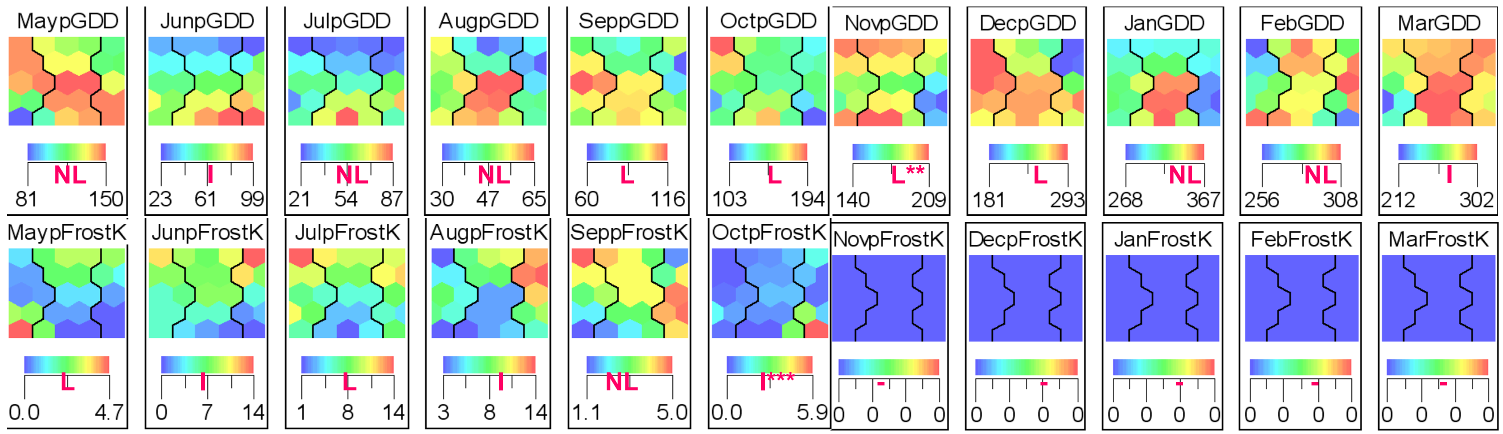
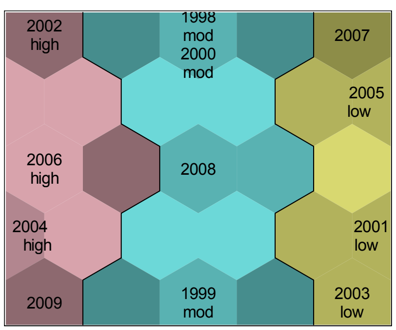


Figure 2: SOM components of (from row 1-4) monthly rainfall (mm), mean temperature (°C), maximum and minimum temperature (°C) each component clustered based on vineyard annual yield year class, left to right *high*, *moderate* and *low.* Monthly rainfall of May, July, September - January show nonlinear (NL) correlations with yield class. August, November and February show linear (L) correlations. Finally, March rainfall shows inverse (IN) correlation with *high* yield. NL: nonlinear, L: linear and I: inverse correlation. ** regression and *** discriminant predictors.





Classification Results[b,c]

		ratecode	Predicted Group Membership 1	2	3	Total
Original	Count	1	3	0	0	3
		2	0	3	0	3
		3	0	0	3	3
		Ungrouped cases	0	2	1	3
	%	1	100.0	.0	.0	100.0
		2	.0	100.0	.0	100.0
		3	.0	.0	100.0	100.0
		Ungrouped cases	.0	66.7	33.3	100.0
Cross-validated[a]	Count	1	3	0	0	3
		2	0	3	0	3
		3	0	0	3	3
	%	1	100.0	.0	.0	100.0
		2	.0	100.0	.0	100.0
		3	.0	.0	100.0	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 100.0% of original grouped cases correctly classified.

c. 100.0% of cross-validated grouped cases correctly classified.

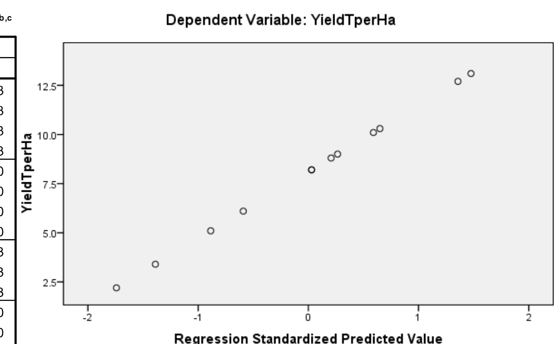


Figure 3 a: SOM components of monthly total GDD (base 10^{o}C) and total frost days clustered based on vineyard annual yield ratings, left to right *high*, *moderate*, low *yield*. L: linear, NL: nonlinear and I: inverse correlation to yield class. ** regression and *** discriminant predictors.
b: SOM of yield and monthly climate variables as in figures 1 and 2a. **c:** Results of regression analysis, yield rating against climate variables.

International Environmental Modelling and Software Society (iEMSs)
2010 International Congress on Environmental Modelling and Software
Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada
David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.)
http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

The SOM results displayed in figs 1 a, 2, 3 a -b show the patterns of correlations between yield year classes and the climate variables studied. The SOM findings are verified with regressions and discriminant analyses and the results are discussed here onwards.

3.2 Statistical data analysis results

Based on regression analyses performed on the same data set predictors for the dependent variable yield (tone per ha) were: (Constant), November GDD, September rain, February maximum temperature, October mean temperature, September maximum temperature, August maximum temperature, November mean temperature, October rain, January mean temperature and June rain (Table 4) for the vineyard being studied in this study.

Table 1: Table showing regression analysis results. Dependent variable vineyard yield (tons per ha) is regressed against monthly total of rain (mm), occurrence of frost days GDD, mean of average, maximum and minimum temperatures (^{o}C) of May to March grapevine growing cycle.

Model Summary[m]

del	R	R Square	ed R Square	. Error of the Estimate	Change Statistics: uare Change	F Change	df1	df2	F Change	-Watson
1	.873[a]	.762	.738	1.7323	.762	32.062	1	10	.000	
2	.914[b]	.835	.798	1.5220	.073	3.954	1	9	.078	
3	.946[c]	.894	.855	1.2906	.060	4.517	1	8	.066	
4	.969[d]	.938	.903	1.0567	.044	4.933	1	7	.062	
5	.982[e]	.965	.935	.8610	.027	4.543	1	6	.077	
6	.982[f]	.964	.944	.8006	.000	.052	1	6	.827	
7	.995[g]	.990	.982	.4587	.026	15.323	1	6	.008	
8	.998[h]	.996	.992	.3006	.006	8.973	1	5	.030	
9	.999[i]	.999	.996	.2087	.002	6.371	1	4	.065	
10	1.000[j]	1.000	.999	.0892	.001	18.905	1	3	.022	
11	1.000[k]	1.000	1.000	.0250	.000	36.211	1	2	.027	
12	1.000[l]	1.000	1.000	.0021	.000	280.124	1	1	.038	1.838

a-l. Predictors: (Constant), NovpGDD, SepprainK, FebmaxT, OctpmeanTK, SeppmaxT, AugpmaxT, NovpmeanTK, OctprainK, JanmeanTK, JunprainK m. Dependent Variable: YieldTperHa

Tables 2 a and b: Tables showing discriminant analysis results which indicate the predictors for this vineyard yield (tons per ha).

Step	Entered	Wilks' Lambda: Statistic	df1	df2	df3	Exact F: Statistic	df1	df2	Sig.
1	NovminT	.214	1	2	6.000	10.996	2	6.000	.010
2	DecprainK	.057	2	2	6.000	8.002	4	10.000	.004
3	MarmaxT	.017	3	2	6.000	8.883	6	8.000	.003
4	JanmaxT	.002	4	2	6.000	14.388	8	6.000	.002
5	OctpFrostK	.000	5	2	6.000	86.499	10	4.000	.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 132.

b. Minimum partial F to enter is 3.84.

c. Maximum partial F to remove is 2.71.

d. F level, tolerance, or VIN insufficient for further computation.

Variables in the Analysis

Step		Tolerance	F to Remove	Wilks' Lambda
1	NovminT	1.000	10.996	
2	NovminT	.469	19.936	.509
	DecprainK	.469	6.957	.214
3	NovminT	.428	17.657	.167
	DecprainK	.324	8.908	.093
	MarmaxT	.659	4.654	.057
4	NovminT	.086	63.100	.106
	DecprainK	.074	33.011	.056
	MarmaxT	.178	12.480	.023
	JanmaxT	.161	8.909	.017
5	NovminT	.002	2297.097	.049
	DecprainK	.002	899.719	.019
	MarmaxT	.057	21.483	.000
	JanmaxT	.006	166.719	.004
	OctpFrostK	.007	114.850	.002

Meanwhile, discriminant analysis ran on the same set of data produced November minimum temperature, December rainfall, March and January maximum temperature, and October frost as predictors of yield for this vineyard (Tables 2a and b).

4. CONCLUSIONS

The paper looked at some major recent approaches to modelling seasonal climate effects on grapevine growth stages, phenology and yield. In general the trend is this regard seems to be more towards analysing yield data at larger spatial scales i.e., regional with data covering at least over three decades. This impedes any conventional approaches to modelling at a micro scale i.e., vineyard and with shorter time span. The paper described an approach consisting of data mining and conventional statistical data analysis methods namely, regression and discriminant, investigated for modelling the year-to-year variability in climatic conditions and its effects on viticulture using yield data from a vineyard in northern New Zealand and with monthly weather data extracted from NIWA's web portal recorded at a nearby metrological station. Despite the SOM results that showed many variables as in relation to yield classes (*low, moderate and high*), regression results produced monthly November GDD, September rainfall, February and August maximum temperature, October, November and January mean temperature, October and June rainfall as predictors of the dependent variable annual vineyard yield (ton/ ha). Interestingly, discriminate analysis results produced monthly November minimum temperature, December rainfall, March maximum temperature and October Forest as predictors of the annual yield in the particular vineyard. The former two independent variables influence *veraison*, March Maximum temperature affects berry ripening and October forest events immensely impact on budburst. Forest events if not handled properly could destroy a whole year's production. Hence, based on the results of data mining, regression and discriminate analysis approaches it is possible to gain scientific understanding on anecdotal evidence in viticulture in a quantitative manner that could be applied to improve viticulture practices whereby the quality and quantity of the yield could be enhanced in this case the grapes and the ultimate produce the wine.

ACKNOWLEDGMENTS

Michael Brajkovich of Kumeu River Wines is acknowledged for his advice on viticulture and winemaking, and sharing one of their vineyard yield data. The authors wish to thank Peter Sumich of Auckland University of Technology for his efforts in this regard.

REFERENCES

Ashenfelter O. and G.V.Jones, The demand for expert opinion: BordeauxWine. VDQS Annual Meeting, d'Ajaccio, Corsica, France. October, 1998. Published in Cahiers Scientifique from the Observatoire des Conjonctures Vinicoles Europeenes, Faculte des Sciences.

Ashenfelter, O., Ashmore, D., and Lalonde, R., Bordeaux, Wine Vintage Quality and the Weather - 1995. Chance vol 8 No. 41995: 7-14.

Cooper, M., (2008) Wine Atlas of New Zealand (2nd Ed) Hardback -Hodder Moa. ISBN: 1869710916 pp408

Jones, G. V., Climate change: observations, projections, and general implications for viticulture and wine production 2007. - Vols. Practical Winery & Vineyard: 44-64.

Jones G. V., White, M. A., Cooper, O. R., and K. Storchmann, Climate and Global Wine Quality [Journal]. - 2005. Climatic Change by Springer vol. 73:319–343.

Jones G.V., Making Wine in a Changing Climate 2004. Geotimes, Aug 2004, Vol. 0, No.7,22-27.

Kohonen, T., Self-Organizing Maps, ISBN 3-540-67921-9, ISSN 0720-678X Heidelberg, New York, 1995, 1997, 2001. Springer series in Information Sciences, Vol. 30, Springer, Berlin.Third Extended Edition, 2001 501 pp

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

Mariani L [et al.] IPHEN: a model for real time production of grapevine phenological maps (2007) . In proceedings of the Climate and Viticulture Congress, Zaragoza, 10-14 April, pp 272-278 See also www.ucea.it

Nemani R. R., White, M. A., Cayan, D. R., Jones, G. V., Running, S.W., and J. C. Coughlan, Asymmetric climatic warming improves California vintages 2001.Clim. Res. Vol 19:25-34

NIWA (National Institute of Water and Atmosphere) [Online]. - 27 September 2009. - http://cliflo.niwa.co.nz.

Shanmuganathan, S., Sallis, P., and Narayanan, A., Statistical data analysis incorporating web text mining to establish correlations between grape wine taster comments and wine ratings in proceeding of the 6th International Conference on Information Information Technology in Asia 2009 (CITA09) "Towards Human-Centered Computing" Hilton Kuching Hotel, Sarawak, 6 - 9th July 2009. Published by IEEE computer society, (2009a) pp 258-262.

Shanmuganathan, S., Sallis, P., and Narayanan, A., Modelling climate change effects on wine quality based on expert opinions expressed in free-text format: The WEBSOM Approach in ICONIP 2008, Part I, LNCS 5506, 2009. - M. Kppen et al. (Eds.), Springer-Verlag Berlin Heidelberg 2009b. pp. 915-925

Tonietto, J., and Carbonneau, A., (2004) A multicriteria climatic classification system for grape-growing regions worldwide Agricultural and Forest Meteorology 124 (2004) 81–97. www.sciencedirect.com

Viscovery [Online]. - 9 February 2010. - http://www.viscovery.net