

Extension of crop model outputs over the land surface by the application of statistical and neural network techniques to topographical and satellite data

M. Bindi^{1,*}, F. Maselli²

¹Dept of Agronomy and Land Management (DISAT), University of Florence, 18 Piazzale delle Cascine, 51044 Florence, Italy

²Institute of Agrometeorology and Environmental Analysis for Agriculture, National Research Council (IATA-CNR), 18 Piazzale delle Cascine, 51044 Florence, Italy

ABSTRACT: The use of crop simulation models to evaluate cultivars and cropping practices has developed greatly in the last few years. These tools can provide unique advantages in several situations, for example, allowing a quick response when new needs arise or to extrapolate results of field experiments in different environmental (climate, soil) and agronomic (cultivars, cropping systems) situations. The operational utilisation of the results of models is however bounded by the problem of extrapolating then to all points on the land surface, which is not always a trivial task in topographically complex regions. The present work investigates the use of different methodologies for the extension of the outputs of a grapevine model in a rugged region of central Italy, Tuscany. In particular, 2 approaches were considered, the first based on statistical assumptions and the second on neural network reasoning. These techniques were applied using, as input parameters, topographical information layers and low-resolution satellite data related to vegetation development. The results obtained show that, in general, the neural network approach produced higher accuracy levels than the statistical approach, but the latter was more capable of merging information coming from different sources. Moreover, the estimates derived from the 2 methods show different spatial patterns and ranges, which must be taken into account when considering these approaches for possible operational uses.

KEY WORDS: Crop simulation models · Regional scale · Spatial analysis · Neural network · Remote sensing · Grapevine *Vitis vinifera* L.

—Resale or republication not permitted without written consent of the publisher—

1. INTRODUCTION

Well-tested computer simulation models of the growth, development, and yield of agricultural crops (perennial and annual) are increasingly being used at various scales from plot to field to identify production problems and to optimise management practices (Booltinlk & Verhagen 1997); simulation extends to the regional level to estimate resource use and crop production (Thornton et al. 1997). In particular, as field trials are too expensive and their results too slow in coming, the characterisation of regional production using crop models seems to provide enormous advan-

tages for planning agricultural investments with respect to any future scenario, from climate change to population increase.

The use of a site crop model at the regional scale has, however, raised questions about scaling-up procedures. The main question is whether to spatially interpolate inputs and run the model for every interpolated point or run the model for points where there are inputs and then spatially interpolate model output variables (Braga & Jones 1999). The quality of the results obtained using these 2 approaches is mainly dependent on the degree of nonlinearity of the model, as well as the spatial structure of the inputs.

Conventional techniques that can be used to explore the spatial distribution of the model outputs and ex-

*E-mail: bindi@sunserver.iata.fi.cnr.it

tend them on the land surface are semi-variance analysis and kriging, which are fundamental instruments of geostatistics (Davis 1973). When information on influential topographic factors is available, these techniques can be complemented by trend analysis, which is particularly effective in rugged terrain (Davis 1973). Recent works have also demonstrated that data taken by the AVHRR (advanced very high resolution radiometer) sensor mounted on board NOAA (National Oceanic and Atmospheric Administration) satellites are useful for eco-climatic classification, thanks to the abundant environmental information contained in the normalised difference vegetation index (NDVI) and thermal infrared images (Benedetti et al. 1994, Maselli et al. 1996). Accordingly, it may be hypothesised that NOAA-AVHRR imagery can provide information on the spatial distribution of the crop simulation model outputs.

All the methods mentioned for the extraction of the information contained in environmental factors and satellite images are based on statistical principles. Actually, a different approach to the same objective, which is based on the use of artificial neural networks (Richards 1993) exists. As will be seen later in more detail, neural network techniques are capable of learning relationships from the patterns of information and can be applied to data layers with completely different properties. The extent to which statistical and neural network approaches can be used to estimate crop productivity derived from plant simulation models in complex areas is, however, mostly unknown. In the present work statistical and neural network methodologies are therefore described and then applied to extend the main outputs from a grapevine model (i.e., duration of phenological stages, fruit dry matter, etc.) over the land surface of a topographically complex region in Central Italy (Tuscany). The results are evaluated visually and statistically against independent estimates of the model outputs.

2. STUDY REGION

Tuscany is situated in Central Italy (9–12° E, 44–42° N; Fig. 1). From an environmental point of view, the region is relatively unique, with its extremely heterogeneous topographical and climatic features. The topography ranges from flat areas near the coastline and along the principal river valleys, to hilly and mountainous zones towards the Appennine chain. Approximately two-thirds of the region is covered by hilly areas, one-fifth by mountains and only one-tenth by plains and valleys (Fig. 2).

From a climatic viewpoint, Tuscany is influenced by its complex orographic structure and by the direction

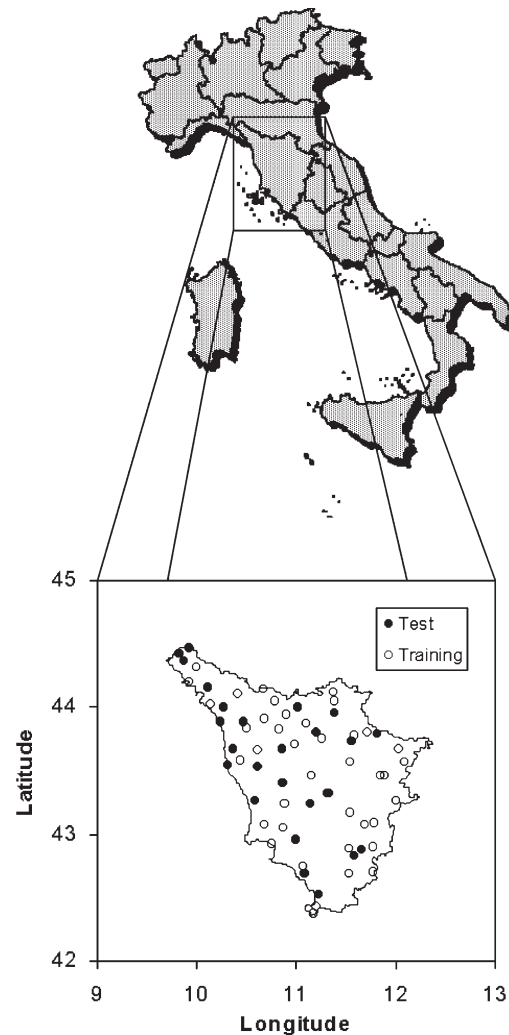


Fig. 1. Geographical position of Tuscany and of the meteorological stations within the region used in this study

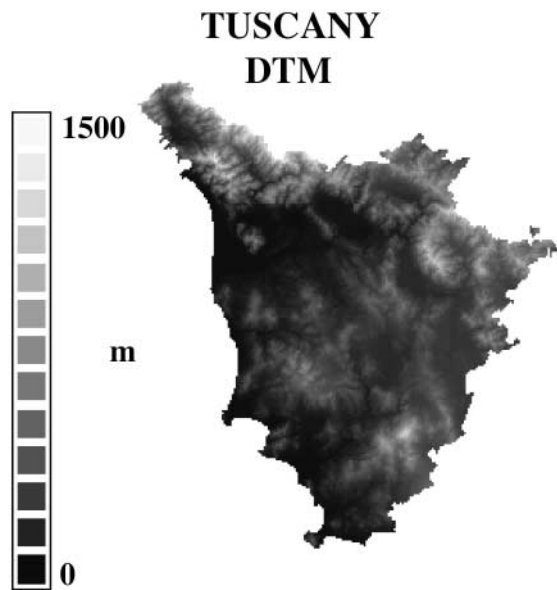
of the prevalent air flows (from west–northwest). As a result, the climate ranges from typically Mediterranean to temperate warm or cool, according to the altitudinal and latitudinal gradients and the distance from the sea (Rapetti & Vittorini 1995). The land use is predominantly agricultural where the land is flat and mixed agricultural and forestry in the hilly and mountainous areas. The main agricultural cover types are cereal crops in the plains and olive groves and vineyards on the hills. The upper mountain zones are almost completely covered by pastures and forests.

3. MATERIALS AND METHODS

3.1. Data. 3.1.1. Climate data: A set of historical climatic data was first constructed for the whole Tuscan

region. In particular, observed daily climatic data were collected for 67 stations evenly distributed over the region (Fig. 1) during the period from 1961 to 1990. The data set consisted of 5 variables: minimum, maximum and mean air temperature, total precipitation and global radiation¹. All these data were first error checked and transferred to a PC in a consistent format. Next, since the data series were incomplete, the Long Ashton Research Station (LARS) stochastic weather

¹Global radiation was estimated using the method proposed by Bristow & Campbell (1984)



generator (Racsko et al. 1991) was used to produce a 30 yr homogeneous set of synthetic data starting from the observed meteorological measurements. The 67 stations then were randomly divided into 2 groups. The first 40 stations were used for training the study methodologies and the second 27 stations for testing them. The spatial distribution of the training and test stations in Tuscany is illustrated in Fig. 1.

3.1.2. Ancillary data: First, digital topographic and geographic data were collected as a basis for subsequent trend and spatial analysis. A digital elevation model (DEM) of the region with a pixel size of 1.1×1.1 km was generated by digitising and processing the contour lines for every 100 m derived from 1:100 000 geographic maps of the Istituto Geografico Militare Italiano (IGMI) (Fig. 2a). Two digital images with north-south gradient and distance from the Tyrrhenian Sea were also produced by appropriate FORTRAN 77 programs.

3.1.3. Satellite data: The NOAA-AVHRR images used in the current study were derived from the '1-KM AVHRR Global Land Data Set', which was produced through international collaboration between NASA (National Aeronautic and Space Administration), NOAA, ESA (European Space Agency), Australia's CSIRO (Commonwealth Scientific and Industrial Research Organisation), China's SMC (Satellite Meteorological Center) and the USGS (United States Geological Survey). The original data set consisted of global 10 d NDVI and thermal infrared (TI) maximum value composite (MVC) images for 1992, 1993 and 1995, which are freely distributed on the web (<http://edcwww.cr.usgs.gov/landdaac/1KM/1kmhomepage>).

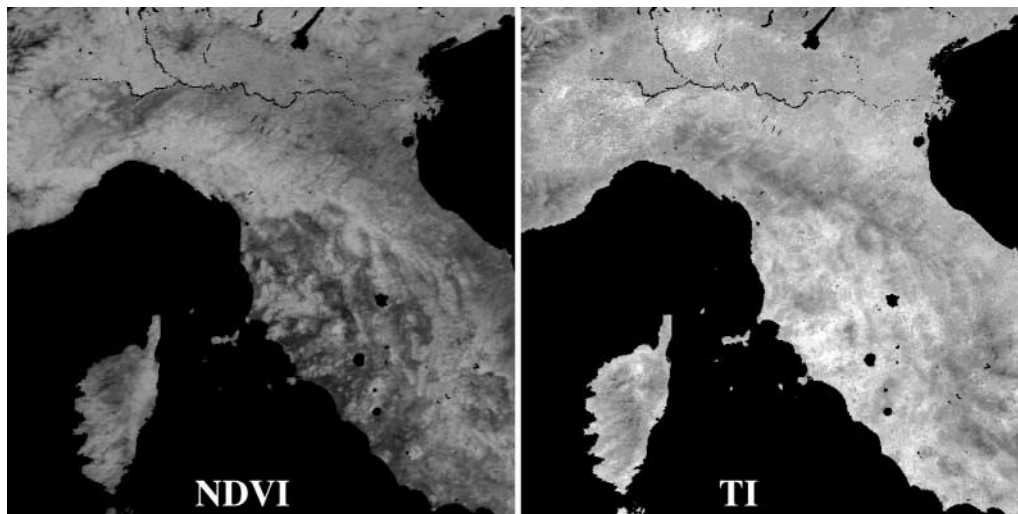


Fig. 2. Examples of data layers used for the study: (a) digital elevation model and (b) August 1995 NOAA-AVHRR NDVI (normalized difference vegetation index) and TI (thermal infrared) images

html). The pre-processing methodology applied to produce these data included the following steps:

(1) Radiometric calibration of channels 1 and 2, described by Teillet & Holben (1993); (2) atmospheric correction following Teillet (1991); (3) computation of daily NDVI images and scaling of the NDVI from -1.0 to 1.0 as 0 to 200 ; (4) geometric registration and conversion into Interrupted Goode Homolosine projection (Goode 1925, Steinwand 1994); and (5) compositing over periods of approximately 10 d using the MVC procedure (Holben 1986). All available NDVI and TI images were windowed over central Italy. Next, only the images from 1995 were retained, since this was the only year which was completely covered from January to December (Fig. 2b).

3.2. Model description. The grapevine model was based on the simulation of the main processes regulating development, growth, and quality components of grapevine (Fig. 3) (Bindi et al. 1997a,b, 2000a). Crop development is divided into 2 periods: vegetative development and fruit growth. The duration of vegetative development, between budbreak and bloom, is calculated on the assumption that bloom occurs when 17 leaves have appeared on the main shoot. The duration of the fruit growth, between bloom and maturity, is assumed to be temperature-dependent and is calculated using cumulative degree days (1440 degree days with a base temperature = 10°C). Leaf area is estimated from the total number of actively growing shoots per unit area, the rate of leaf appearance and leaf expansion. Leaf area growth is estimated as a function of the total number of leaves that have appeared. Leaf area is used to calculate the amount of solar radiation intercepted by the leaf canopy so that crop biomass accumulation can be calculated on the basis of crop radiation-use efficiency (RUE). Daily fruit growth rate is calculated assuming that the fruit biomass index (i.e., ratio of fruit to current year's total biomass) increases linearly during fruit growth. Grape quality (i.e., sugar and acid content) is simulated as a function of intercepted solar radiation.

The grapevine model has already been validated extensively for Sangiovese and Cabernet cultivars within the Tuscany environment in 2 previous studies. Firstly, by comparison with data collected in model oriented field experiments (Bindi et al. 1997b), and, secondly, by comparison with observed data obtained from the Agrometeorological Service of Tuscany,

Institutes of Agricultural Ministry and private consortia of the most important viticultural areas of Tuscany (Bindi et al. 2000b). Both validation studies confirmed the ability of the model to reproduce the variability of grapevine yields for the study area.

3.3. Interpolation methodologies. The grapevine model was first run for each of the 67 selected stations using the daily synthetic weather data for 30 yr (1961 to 1990) produced by the LARS-WG stochastic weather generator and the parameters defining crop characteristics fixed for cv. Sangiovese in Bindi et al. (1997b). Next, the averages of the main model output variables (bud-break and maturity dates, grape yield, and acid and sugar grape concentrations) were computed for all stations. Finally, 2 different methodologies for extending site model output variables were evaluated and used to produce scenarios of viticultural production at a regional scale. Specifically, the capacity to reproduce the spatial variation of the examined parameters was evaluated by applying: (1) a statistical approach that combined and optimised the results obtained using a trend analysis, kriging and fuzzy classification of remotely sensed data; and (2) a neural network approach that could separately or jointly use topographical and/or remotely sensed data. The basic principles of these 2 approaches are

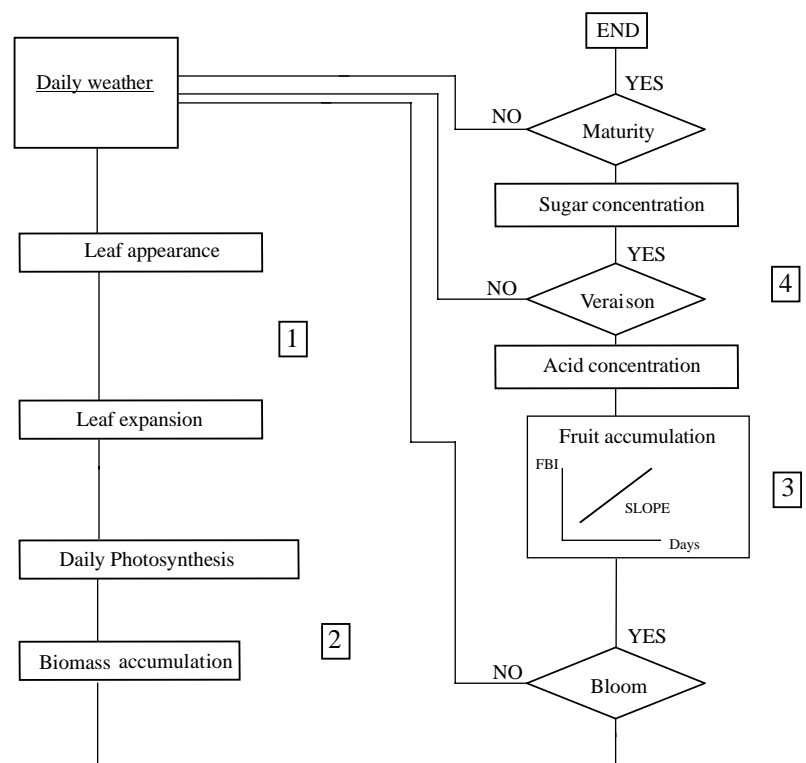


Fig. 3. Flow diagram of the grapevine model with the main simulated processes: (1) ontogeny; (2) biomass accumulation; (3) fruit growth; (4) sugar and acid accumulation (see Bindi et al. 1998a for details)

briefly reviewed, followed by a description of their application.

3.3.1. Statistical approach: The application of statistical methods generally requires knowledge of the distributions of the input and output variables considered and of the relationships which link the former to the latter. Since topographic factors and satellite images have different properties and relationships with vegetation variables, different procedures had to be applied for their treatment in the present context, with the merging of the relevant outcomes postponed to a second phase. In particular, trend analysis and kriging were first applied to extract the topographic and spatial dependencies from the topographical information layers. Next, a fuzzy classification method, recently proposed by one of the authors (Maselli 2001), was used to extract the spectral information from the NOAA-AVHRR images. Finally, the outcomes of the 2 methods were integrated using an error minimisation criterion.

3.3.1.1. Trend analysis and kriging: Every natural phenomenon and its manifestation has a spatial structure. According to geostatistics (Davis 1973), the spatial variability of most environmental parameters can be approximated by a stochastic surface. Kriging was developed as a stochastic estimation method based on local weighted averaging, where the weights can be derived from the experimental semi-variogram. This is an estimate of a theoretical function calculated from a finite number of points which correspond to all pairs of observations separated by distance h . In practice, the kriging estimator is given by the formula:

$$Z_k(x) = \sum_{i=1}^N \lambda_i Z(x_i) \quad (1)$$

where $Z_k(x)$ is the study random function at point x and λ_i and $Z(x_i)$ are the kriging coefficients and the measurements of the function in the N observations, respectively.

In the present approach, the altitudes of the station pixels were first used in a conventional linear regression model with respect to the study parameters in order to remove systematic trends. Next, kriging was applied to interpolate the residuals of the study parameters for all pixels using an exponential semi-variogram model. Finally, the study parameters for each pixel were obtained, reintroducing the systematic trend due to altitude. As described in Davis (1973), estimates of error variance can be also obtained by kriging for each pixel, and these were currently retained for subsequent use.

3.3.1.2. Fuzzy classification of NOAA-NDVI data: Data taken by AVHRR mounted on board NOAA satellites have already been demonstrated to be useful for

eco-climatic classifications, thanks to the abundant environmental information contained in NDVI and TI data (Maselli et al. 1996, 1998). However, NDVI and partly also TI data are related to crop growth dynamics in a complex, nonlinear manner, since vegetation conditions respond to environmental factors in a complex way. Usual regression techniques would therefore be ineffective for the processing of these data, due to the relevant rigid statistical constraints. A suitable alternative could be based on the identification of spectral similarities among pixels in a more flexible way. Following these considerations, a fuzzy logic approach was considered to be most appropriate. Such an approach, recently developed for the estimation of forest parameters from Landsat TM (thematic mapper) data (Maselli et al. 1995) and for the estimation of agroclimatic parameters from NOAA-NDVI data (Maselli et al. 1998), was used in this study.

Specifically, in this approach the image pixels were the fuzzy elements and each training entity (corresponding to the pixels around a meteorological station) was treated as an individual class. Fuzzy probabilities of attribution of all pixels to each entity were consequently found out by a softened maximum likelihood classification of the remotely sensed data (Wang 1990a,b). These probabilities were then used to estimate the study parameters for each pixel through a weighted average strategy as follows:

$$Z_t = \sum_{i=1}^N Pr_i Z_i \quad (2)$$

where Z_t = estimated parameter for the pixel considered; N = number of classes (stations); Pr_i = pixel fuzzy probability of attribution to class (station) i ; and Z_i = measured parameter of class (station) i . Some modifications to the original method have been recently proposed by the same author (Maselli 2001) in order to improve the consideration of more informative channels and to obtain per-pixel estimates of error variances as in kriging (Davis 1973).

For the application of this new version, the 36 NDVI and TI MVCs of 1995 were first further composited over monthly periods, in order to reduce residual atmospheric disturbances. Next, the spectral signatures of the training stations were extracted from 3×3 pixel windows around each of them. The most informative bands were then selected from the 12 NDVI plus 12 TI MVCs by examining the correlations between the spectral distances among the training stations and the corresponding parameter's differences (Maselli 2000). In this way, the 10 satellite bands listed in Table 1 were selected for the application of the method described over the whole study region.

3.3.1.3. Statistical merging of the estimates from the two sources: As the estimates of the model output vari-

Table 1. Topographical and satellite data used. All data listed were utilised with the neural network approach, whilst the layers related to spatial dependencies (latitude, longitude, distance from the sea) were substituted by the use of kriging within the statistical approach

Data layer	Parameters
Topographical data	Elevation Latitude, longitude Distance from the sea
Satellite data (1995)	NDVI (March, April, August) TI (February, March, April, May, June, August, October)

ables were derived from 2 sources which could be considered independent, they could be combined to produce optimal estimates. According to Gelb (1974), if 2 independent measurements of the same variable Z are available, then the combined optimal estimate Z_0 can be derived as:

$$Z_0 = \left(\frac{1}{Var_1} + \frac{1}{Var_2} \right)^{-1} \left(\frac{Z_1}{Var_1} + \frac{Z_2}{Var_2} \right) \quad (3)$$

where Z_1 and Z_2 = independent measurements of the variable (kriging and fuzzy); and Var_1 and Var_2 = error variances of the 2 measurements, which were produced by both the kriging and fuzzy classification procedures (Davis 1973, Maselli et al. 1998)

3.3.2. Neural network approach: A neural network (NN) can be viewed as a computer system that is made up of several simple and highly interconnected processing elements similar to the neuron structure found in the human brain (McClelland et al. 1986a,b). Problems which are normally not solvable by traditional algorithmic approaches can be solved with an NN approach (Davidson & Lee 1991). NNs are suitable for problems which require the interpretation of large data sets. In addition they can also be used to solve problems in which the inputs and the corresponding outputs are known, but the relationships between inputs and outputs are not well understood. These conditions are commonly found in many environmental and agricultural applications. Consequently NN models have successfully been used in biological applications to predict processes such as crop phenology (e.g., soybean), insect pest treatments thresholds, weather variability, optimum temperature in greenhouses, etc. (Seginer & Sher 1992, Elizondo et al. 1994, McClendon & Bachelor 1995).

Another property of NNs is that they require no assumption on the distributions of the data used, so that they can contemporaneously treat data with different statistical properties (Richards 1993). In the current case, this complete nonparametric approach allowed the application of the same procedure to both

topographical and satellite data, changing only the configuration of the network depending on the numbers of input variables used.

In practice, different 3-layer feedforward NN models were developed for extending simulated phenological stages and yield components over the regional land surface. A graphic representation of the typical NN model considered for extending the model outputs is presented in Fig. 4. As done previously, the mean model outputs computed for 30 yr at the training stations were used to train and test 3 NN models. The first NN model was trained and tested adopting as inputs the topographical factors (4 layers, Table 1); whereas the other 2 NN models used as input the satellite images (10 images), and all the input parameters (4 layers plus 10 images). The models were developed using QNET neural network software (NN Modelling, Vesta Services Inc.). For the first NN model, a sensitivity analysis was performed to determine the optimal number of hidden nodes (over a range of 5 to 25 with a 5 nodes step) and to tune the proper learning rate and momentum.

3.3.3. Accuracy evaluation of scaling-up methods:

The performance of the 2 methodologies (statistical and neural network) when applied to the different data sets (topographic factors, satellite images and both) were tested by comparing the actual model outputs for the test stations to the relevant extended estimates. Two statistics were taken into account to evaluate the comparison. The first was as an indicator of correlation between the 2 variables (correlation coefficient, r) and the second was as an indicator of mean difference between them (root mean square error, RMSE).

4. RESULTS

4.1. Method performances

4.1.1. Statistical approach

In general, the statistical approach provided good estimates of model output variables (Table 2). Specifically, it was found that the kriging method based on the 'topographical data' gave quite satisfactory accuracy in the extension of fruit dry matter and sugar content (Table 2). In both cases a correlation coefficient higher than 0.91 was reached with quite low RMSEs (0.95 t ha⁻¹ for fruit dry matter and 4.3% for sugar content). Similar accuracy was achieved for acid content ($r = 0.93$ and RMSE = 2.5 g l⁻¹) and slightly lower bud-break dates ($r = 0.84$ and RMSE = 16 d). Finally, a much lower accuracy was obtained for maturity dates ($r = 0.74$ and RMSE = 102 d), for which some explanation is needed. Specifically, the extension of this parameter

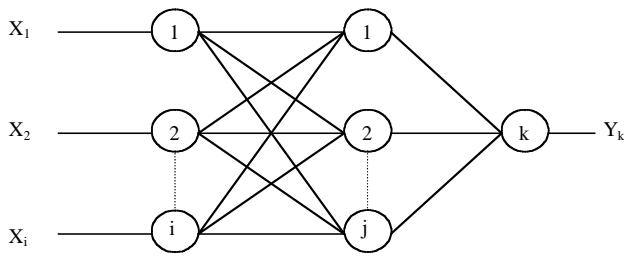


Fig. 4. Three-layer feedforward NN with i input nodes, j hidden nodes, and output k

was carried out keeping all model outputs, which covered a range of grapevine maturation dates much wider than that which really occurred in the region. This permitted the evaluation of the actual extension capacity of the different methodologies, but also increased the mean errors obtained.

Again from Table 2, it can be seen that the fuzzy estimation approach based on ‘satellite data’ produced results with accuracy levels similar or slightly lower than those of the kriging approach. From the statistical comparison with the test data (Table 2), it can be seen that the estimation of the budbreak and maturity dates were globally as accurate as in the previous case, while a slightly lower accuracy was obtained for fruit dry matter and acid and sugar contents.

The statistical merging of the 2 data sources produced optimal estimates of the model outputs (Table 2).

Table 2. Results of the statistical comparison between the model outputs for the 27 test stations and the corresponding estimates from the 2 approaches (statistical and NN) applied to the 3 data sets (topographical, satellite and merged). Bud = budbreak dates (DOY), Mat = maturity dates (DOY), Aci = acidity content (g l^{-1}), Sug = sugar content (%), Fru = fruit dry matter (t ha^{-1}) (*significant at $p < 0.01$). DOY = day of year

Data set	Model output	Approach			
		Statistical r	Statistical RMSE	Neural network r	Neural network RMSE
Topographical data	Bud	0.841*	16.22	0.760*	12.83
Satellite data	Bud	0.87*	15.31	0.867*	15.75
Topogr.+sat. data	Bud	0.89*	14.96	0.940*	9.58
Topographical data	Mat	0.737*	102.11	0.893*	99.20
Satellite data	Mat	0.747*	107.54	0.724*	112.18
Topogr.+sat. data	Mat	0.766*	102.72	0.839*	80.73
Topographical data	Aci	0.931*	2.47	0.953*	1.50
Satellite data	Aci	0.889*	2.34	0.830*	2.50
Topogr.+sat. data	Aci	0.936*	2.30	0.897*	1.94
Topographical data	Sug	0.919*	4.27	0.930*	3.02
Satellite data	Sug	0.895*	4.17	0.923*	3.09
Topogr.+sat. data	Sug	0.925*	4.09	0.960*	2.16
Topographical data	Fru	0.916*	0.95	0.921*	0.68
Satellite data	Fru	0.881*	0.93	0.906*	0.77
Topogr.+sat. data	Fru	0.926*	0.91	0.931*	0.64

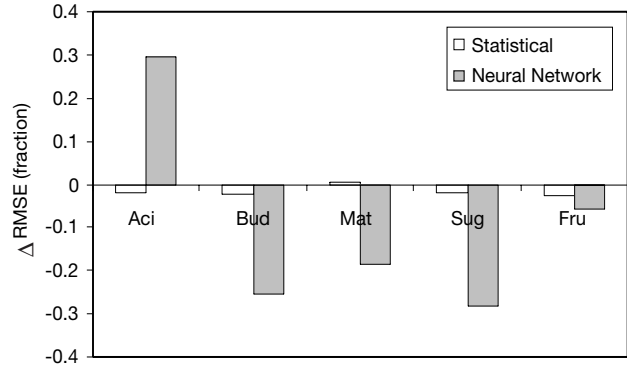
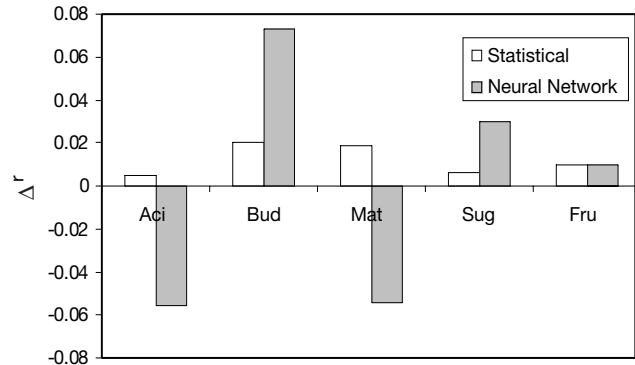


Fig. 5. Differences in r and RMSE obtained by the integration procedure with respect to the use of the better basic data layer (topographical or satellite data), expressed as fractions to make the values comparable for parameters with different scales. For explanations of abbreviations, see Table 2

In Fig. 5, the differences produced by the joint consideration of the 2 data layers (topographical and satellite data) are reported with respect to the better single layer. As can be seen, the estimation accuracy of almost all parameters was slightly increased in terms of both statistics. In particular, good results were obtained for budbreak and maturity dates, which were poorly estimated by both the kriging and fuzzy procedures. This was an indirect indication that the per-pixel error estimates produced by both the kriging and fuzzy methods were informative for the uncertainty of the estimated parameters and could be exploited by the error minimisation criterion given in Eq. (3).

4.1.2. Neural network approach

The NN approach was also quite satisfactory from the viewpoint of accuracy in the estimation of model output variables (Table 2). More specifically, the NN model

developed using topographical data gave a higher accuracy in the estimation of fruit dry matter and sugar and acid contents. In all 3 cases a correlation coefficient higher than 0.92 was reached with relatively low RMSEs (0.68 t ha⁻¹ for fruit dry matter, 3.0% and 1.5 g l⁻¹ for sugar and acid contents). A relatively lower accuracy was again achieved for maturity dates ($r = 0.89$ and RMSE = 99 d), for which the explanation has been previously provided. Also, a lower accuracy was obtained for budbreak dates ($r = 0.76$ and RMSE = 13 d).

The development of an NN model using the satellite data produced results with accuracy levels slightly lower than those obtained using topographical data. By developing the NN model with both data sources, only marginal and inconsistent improvements were obtained in the accuracy levels of the estimates (Table 2). In particular, as can be seen in Fig. 5, the correlation coefficients were increased and the RMSEs reduced for budbreak dates, sugar content and fruit dry matter, whereas there were opposite results for acid content (a decrease in r and increase in RMSE). In contrast, there was a reduction in r as well as a decrease in RMSE for maturity dates. A possible explanation of these results is that the AHVRR data contained information which was very variable in space and not solely related to vegetation development (Maselli et al. 1998). This probably caused overfitting problems in the application of the NN process, which, being completely nonparametric, was relatively sensitive to the presence of anomalous values in the training or test data.

4.2. Regional representations of the optimum extended estimates

On the basis of the results of the model performances, the statistical and NN approaches with the complete data sets (topographical plus satellite data) were selected for extending the model output variables over the entire region. The extended estimates of phenological development and grapevine yield and quality are commented on in the following sections.

While describing these images, it is worth making a preliminary remark about the variability of the estimates derived by the 2 methods. The statistical method, being based on weighted average strategies, tended to reduce the ranges of the parameters, which was not the case with the neural network procedure. This resulted in extended estimates from the 2 methods which, even if strongly correlated, have somewhat different appearances, as is clearly visible in Figs. 6 & 7.

4.2.1. Phenological development

The dates of budbreak and maturity as predicted by the grapevine model are shown in Fig. 6. Budbreak was predicted to occur mainly from late March to early April in areas along the coast (west and southwest Tuscany), from early to mid April in the centre of the region and from the mid to late April in most continental areas (east and northeast Tuscany). Physiological maturity was predicted to occur from mid to late August in southwestern Tuscany. Further inland, in most of central and northern areas grapevines were found to mature in September. In areas close to the northern and eastern limits of grapevine cultivation maturity was modelled to occur in October.

4.2.2. Grapevine yield and quality

The distribution of potential fruit dry matter and acid and sugar content is shown in Fig. 7. Fruit dry matter was found to range from 7.2 up to 7.8 t ha⁻¹ for most of Tuscany. Fruit dry matter was high in western areas, where the season was longer and radiation receipts were higher, and gradually declined northwards and eastwards. Predictions of sugar content exhibited a different pattern to fruit dry matter, varying from 22.5% in the southern and eastern areas to less than 21% in some small inland areas. As for sugar, the highest acid contents were predicted in

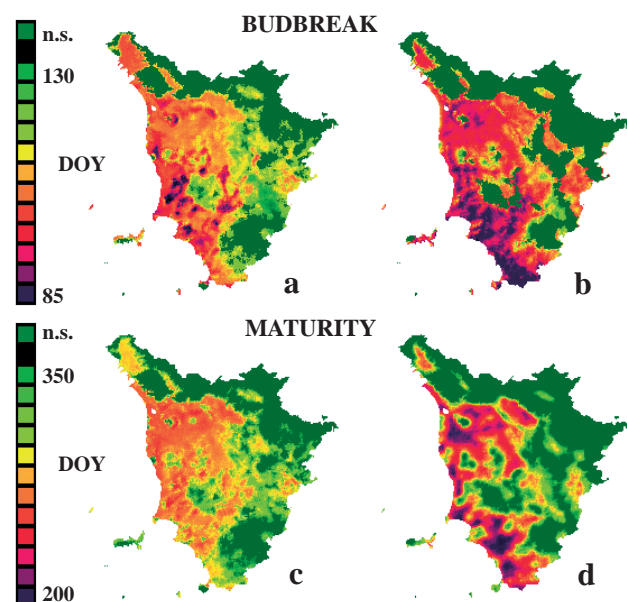


Fig. 6. Extended estimates of the phenological model outputs obtained by the (a,c) statistical and (b,d) neural network approaches. DOY = day of year; n.s. = nonsuitable areas (i.e., grapevine does not reach maturity stage)

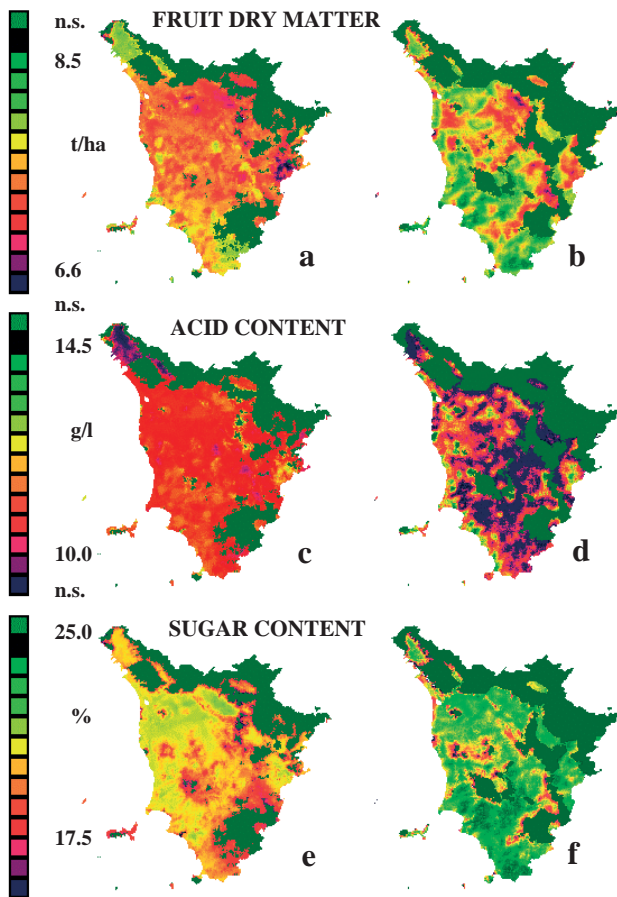


Fig. 7. Extended estimates of the productivity model outputs obtained by the (a,c,e) statistical and (b,d,f) neural network approaches (n.s. = nonsuitable areas).

southern Tuscany (up to 12.5 g l^{-1}) and they gradually declined northwards and eastwards.

5. DISCUSSION AND CONCLUSIONS

In the current work 2 methods based on different concepts and assumptions were applied to 2 information sources in order to extend the outputs of a grapevine simulation model over the surface of a complex region in central Italy. From the results obtained, the following conclusions can be drawn: (1) Both statistical and NN approaches are effective in extracting the topographical and spectral information of the layers considered. In general, however, the latter method produces higher accuracy levels. (2) The 2 information sources contain similar amounts of information on grapevine development and productivity, and part of this information is clearly redundant. The use of satellite data can, however, cause overfitting problems when using the NN approach. (3) The statistical

method is more effective than the NN procedure in combining the 2 information sources, thanks to the production of error estimates which are exploited by the error minimisation criterion applied. (4) As regards the extended estimates, the statistical and NN methods produce maps with similar information, but the former, unlike the latter, tends to reduce the variable ranges.

The different performances of the 2 methods and data sets must be obviously confirmed by further investigations. We feel, however, that these results are worth of dissemination as a first step towards the application of these relatively new techniques on an operational basis.

Acknowledgements. The authors wish to thank Luca Fibbi for technical assistance. This research was supported partly by the Commission of EU (Project CLIVARA no. ENV4-CT95-0154, Environment Programme) and partly by the Italian Space Agency (ASI).

LITERATURE CITED

- Benedetti R, Rossini P, Taddei R (1994) Vegetation classification in the middle Mediterranean area by satellite data. *Int J Remote Sensing* 15:583–596
- Bindi M, Miglietta F, Gozzini B, Orlandini S, Seghi L (1997a) A simple model for simulation of growth and development in grapevine (*Vitis vinifera* L.). I. Model description. *Vitis* 36:67–71
- Bindi M, Miglietta F, Gozzini B, Orlandini S, Seghi L (1997b) A simple model for simulation of growth and development in grapevine (*Vitis vinifera* L.). II Model validation. *Vitis* 36:73–76
- Bindi M, Fibbi L, Lanini M, Miglietta F (2000a) Effects of elevated carbon dioxide on grapevine. In: Butterfield RE, Lonsdale KG, Downing TE (eds) Climate change, climatic variability and agriculture in Europe. Final Report (1996–1998) to the Commission of the European Union, Contract ENV4-CT95-0154, Environmental Change Unit, Oxford, p 71–81
- Bindi M, Fibbi L, Maselli F, Miglietta F (2000b) Modelling climate change impacts on grapevine in Tuscany. In: Butterfield RE, Lonsdale KG, Downing TE (eds) Climate change, climatic variability and agriculture in Europe. Final Report (1996–1998) to the Commission of the European Union, Contract ENV4-CT95-0154, Environmental Change Unit, Oxford, p 191–216
- Booltink HWG, Verhagen J (1997) Using decision support systems to optimise barley management of spatially variable soil. In: Kropff MJ et al. (eds) Applications of systems approaches at the field level, Vol 2. Systems approaches for sustainable agriculture development 6. Kluwer, Dordrecht, p 219–233
- Braga RP, Jones JJ (1999) Predicting spatial pattern of simulated yield from point input data. In: Donatelli M, Stockle C, Villalobos F, Villar Mir JM (eds) Proc Int Symp Modelling Cropping Systems. ESA, University of Lleida, Lleida, p 149–150
- Bristow KL, Campbell GS (1984) On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric For Meteorol* 31:159–166
- Davidson CS, Lee RH (1991) Artificial neural networks for automated agriculture. In: Proc 1991 Symp Automated

- Agriculture for the 21st Century. American Society of Agricultural Engineers (ASAE), St. Joseph, MI, p 106–115
- Davis JC (1973) Statistics and data analysis in geology. John Wiley & Sons, New York
- Elizondo D, Hoogenboom, C, McClendon RW (1994) Development of a neural network model to predict daily solar radiation. *Agric For Meteorol* 71:115–132
- Gelb A (ed) (1974) Applied optimal estimation. The MIT Press, Cambridge, MA
- Goode JP (1925) The Homolosine projection: a new device for portraying the Earth's surface entire. *Annals Assoc Am Geogr* 15:119–125
- Holben BN (1986) Characteristics of maximum-value composite images from temporal AVHRR data. *Int J Remote Sensing* 7:1417–1434
- Maselli F (2001) Extension of environmental parameters over the land surface by improved fuzzy classification of remotely sensed data. *Int J Remote Sensing* (in press)
- Maselli F, Conese C, De Filippis T, Norcini S (1995) Estimation of forest parameters through fuzzy classification of TM data. *IEEE Trans Geosci Remote Sensing* 33:77–84
- Maselli F, Conese C, Petkov L, Maracchi G (1996) Eco-climatic classification of Tuscany through NOAA-AVHRR data. *Int J Remote Sensing* 17:2369–2384
- Maselli F, Petkov L, Maracchi G (1998) Extension of climate parameters over the land surface by the use of NOAA-AVHRR and ancillary data. *Photogram Eng Remote Sensing* 64:199–206
- McClelland JL, Rumelhart DE, the PDP Research Group (1986a) Parallel distribution processing: exploration in the microstructure of cognition. Vol I: Foundations. MIT Press, Cambridge, MA
- McClelland JL, Rumelhart DE, the PDP Research Group (1986b) Parallel distribution processing: exploration in the microstructure of cognition. Vol II: Psychological and biological models. MIT Press, Cambridge, MA
- McClendon RW, Batchelor WD (1995) An insect pest management neural network. ASAE Paper No. 95-3560, American Society of Agricultural Engineers (ASAE), St. Joseph, MI
- Racsco P, Szeidl L, Semenov MA (1991) A serial approach to local stochastic weather models. *Ecol Model* 57:27–41
- Rapetti F, Vittorini S (1995) Carta climatica della Toscana. Pacini Editore, Pisa
- Richards JA (1993) Remote sensing digital image analysis: an introduction, 2nd edn. Springer-Verlag, Heidelberg
- Seginer I, Sher A (1992) Neural-nets for greenhouse climate control. ASAE Paper No. 927013. American Society of Agricultural Engineers (ASAE), St. Joseph, MI
- Steinwand DR (1994) Mapping raster imagery to the interrupted Goode Homolosine Projection. *Int J Remote Sensing* 15:3463–3472
- Teillet PM (1991) Radiometric and atmospheric correction procedures for AVHRR preprocessing in the solar reflective channels. In: Proc 5th Int Colloquium on Spectral Signatures of Objects in Remote Sensing, Courchevel, p 511–516
- Teillet PM, Holben BN (1993) Towards operational radiometric calibration of NOAA AVHRR imagery in the visible and infrared channels. *Can J Remote Sensing* 20:1–10
- Thornton PK, Bowen WT, Ravelo AC, Wilkens PW, Farmer G, Brock J, Brink JE (1997) Estimating millet production for famine early warning: an application of crop simulation modelling using satellite and ground-based data in Burkina Faso. *Agric For Meteorol* 83:131–136
- Wang F (1990a) Fuzzy supervised classification of remote sensing images. *IEEE Trans Geosci Remote Sensing* 28:194–201
- Wang F (1990b) Improving remote sensing image analysis through fuzzy information representation. *Photogram Eng Remote Sensing* 56:1163–1169

Editorial responsibility: Gerd Esser, Gießen, Germany

*Submitted: December 6, 1999; Accepted: September 15, 2000
Proofs received from author(s): March 8, 2001*