SYNERGETIC USE OF REMOTE SENSING AND SOILBORNE DATA FOR REGIONAL YIELD PREDICTIONS OF MALTING BARLEY *(HORDEUM VULGARE L.)*

Christof J. Weissteiner [1], *Matthias Braun* [1] *and Walter Kühbauch* [2]

1. University of Bonn, Centre for Remote Sensing of Land Surfaces, 53115 Bonn, Germany {c.weissteiner/matthias.braun}(at)uni-bonn.de
2. University of Bonn, Department of Crop Science and Plant Breeding, 53115 Bonn, Germany lap(at)uni-bonn.de

ABSTRACT

Yield forecasts are of high interest to the malting and brewing industry in order to allow the most convenient organisation of the respective purchasing policy of raw materials. Within this investigation, malting barley yield forecasts (*Hordeum vulgare L.*), in Germany mostly grown as spring barley, are performed for typical growing regions in South-Western Germany. Multitemporal remote sensing data on the one hand and ancillary data such as meteorological, agrostatistical, topographical and pedological data on the other hand are used as input data for two versions of prediction models, which were based on an empirical-statistical modelling approach.

Since spring barley production is dependent on acreage and on the yield per area, *classification* is needed, which was performed by a supervised multitemporal classification algorithm, utilizing optical remote sensing data (LANDSAT TM/ETM+). The classification algorithm considers spectral data, topographical data (Digital Elevation Model) and expert knowledge input. The latter is important with regard to the particular phenological development of the observed crop, an expertise which was used to distinguish it from similar crops.

The basic version of the *yield estimation* model was conducted by means of linear correlation of remote sensing data (NOAA-AVHRR NDVI Maximum Value Composites), CORINE land cover data and agrostatistical data. In an extended version meteorological data (temperature, evapotranspiration) and soil data were incorporated. Both basic and extended prediction systems led to feasible results depending on the selection of the time span for NDVI accumulation. For NDVI accumulation across the grain-filling period, the mean deviation of the reported yield from the simulated one was 7.0% and 6.4% for the basic and extended yield estimation model, respectively.

Keywords: statistical model of yield forecast, NDVI yield prediction, malting barley

INTRODUCTION

The yearly variation of acreage and quality of agricultural crops determines the formation of prices. Food processing industry is therefore highly interested in yield forecasts in order to obtain an early knowledge about expected yields and quality of raw materials in order to optimize their purchasing policy.

Malting barley in Germany is mostly grown as spring barley (*Hordeum vulgare L.*). It has, however, to be considered that a percentage of spring barley, mainly lower quality spring barley, is also utilized as forage for livestock farming.

Since yield represents the major objective in the marketing of malting barley, the present study focused on yield quantity [tons ha^{-1}] and regional production, which is determined by the combination of both acreage and tons ha^{-1}. For acreage determination optical satellite imagery (LANDSAT TM and ETM+) was used. Pixel-based as well as object-oriented classification was carried out, a comparison of both was conducted later on. It is, however, still a challenge to classify individual crops and not crop groups like cereals as a whole, in particular for the given conditions (spatial resolution, land use structure). Similar works, which have been conducted in the past, achieved classification accuracies of 40 - 70% (1,2).

As described in the literature, a variety of yield prediction models are using remote sensing data (3). The used model for yield per area was empiric statistical, based on linear correlation of agrostatistical data and NDVI maximum value composites (MVC) between 1995 – 2002. Additionally, ancillary data (CORINE land cover, meteorological, phenological, pedological, topographical data) was incorporated. Linear correlation for yield estimation is propagated by several authors (4,5,6). NDVI, an indirect measure for vegetation vigour, is used to determine expected yields; the relationship between NDVI and yield performance is derived from historical agrostatistical and NDVI data (1995 – 2002). As a prerequisite land cover should be mainly dominated by the investigated crop or crop family. For investigations on yield quantity (yield ha^{-1}) optical low resolution imagery from NOAA-AVHRR (1.1 km) was utilized.

METHODS

Image analysis and classification

A pixel-based and an object-oriented strategy have been applied for the supervised multitemporal classification in order to enable comparisons between both methodologies.

Supervised pixel-based classification was done by ERDAS Imagine. Water bodies, forested areas and impervious areas were masked out in order to reduce noise in the image. The remaining agricultural areas of the available images were classified as agricultural crops by a Maximum-Likelihood algorithm (7). A DEM and knowledge-based rules were incorporated. The latter ones were interactively defined by NDVI thresholds and accounted for the expulsion of similar crops.

Object-oriented supervised classification was carried out by eCognition software (8) with previous segmentation and the implemented Standard Nearest Neighbor algorithm, which is in fact, to some extent, also a fuzzy classification. Like in the pixel-based classification, knowledge-based rules were incorporated, based on thresholds of an interactively gathered fuzzy rule set.

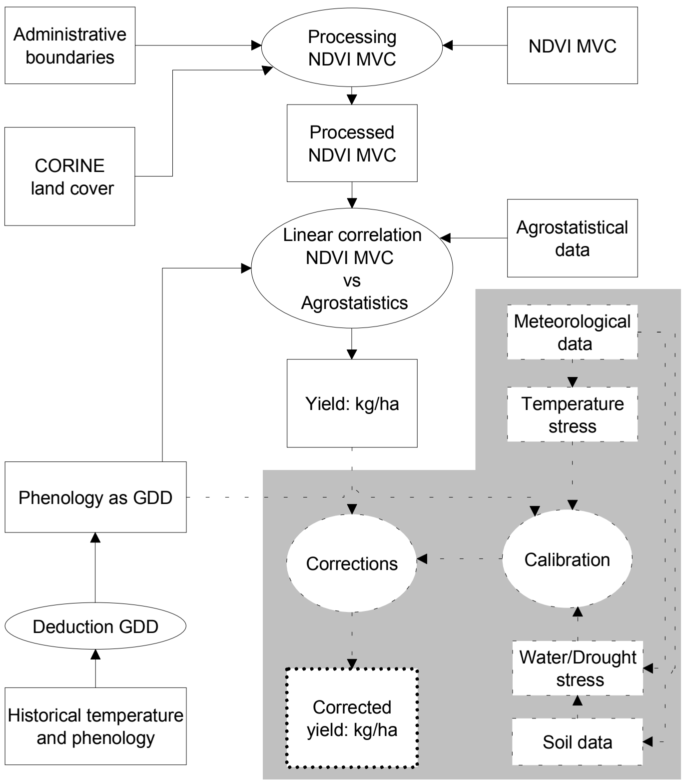
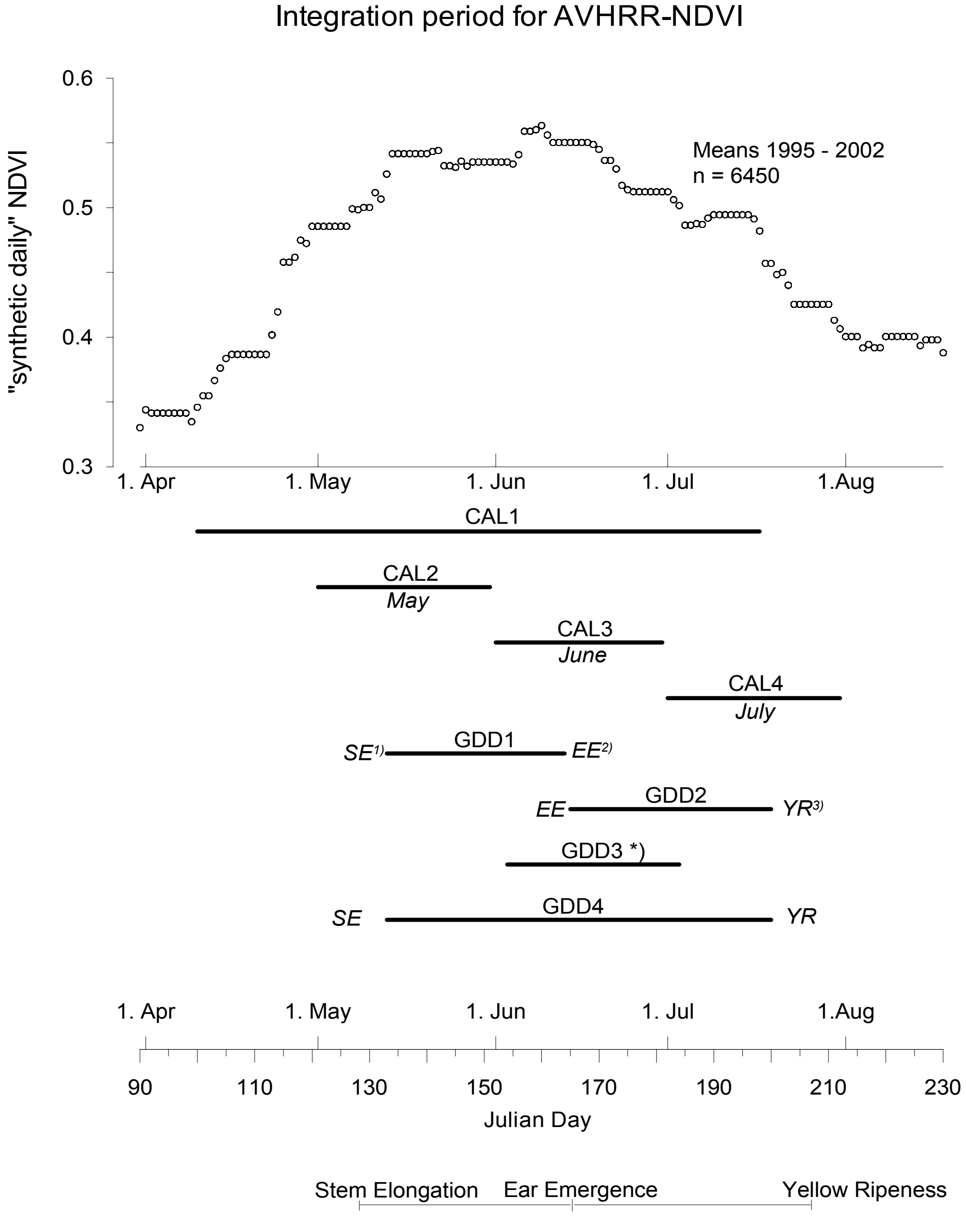


Figure 1: Schematic model for the basic (white background) and extended (grey background) NDVI linear correlation model for yield (kg ha^{-1}).

Yield prediction models

As shown in Figure 1, two different models, a basic and an extended model have been applied for yield prediction in terms of yield per area. However, both are based on a linear correlation model with NDVI MVC and agrostatistical data as main inputs. Both models utilize CORINE land cover data (9). Administrative units are defined by vector layers, which are used for zonal averaging over districts on the so-called European NUTS-3 level (10). Out of historical meteorological and phenological data Growing Degree Days (GDD) were generated and utilized for determination of starting and ending days of phenological stages for NDVI accumulation (Figure 2). An alternative modelling approach utilized calendar days for determination of NDVI accumulation periods.



*) GDD3: Comprises the period of the expected maximum photosynthetic activity
1) SE ... Stem elongation
2) EE ... Ear emergence
3) YR ... Yellow ripeness

Figure 2: NDVI means and time spans for NDVI integration. Integration periods were based on calendar units (CAL1 – CAL4) or on Growing Degree Days (GDD1 – GDD4) and thus based on phenology. GDD time spans are just roughly depicted since defined as local temperature sums.

For the extended yield prediction model two stress factors have been implemented, focusing on sensitive stages of plant development. Temperature stress was regarded in 3 stages (emergence, stem elongation, heading), water stress in two stages (stem elongation and heading). Due to high temperatures in the stages of stem elongation and heading losses in grain yield may occur. This is described by Wheeler et al. (11) for wheat and for barley, which reacts in a similar way, by Eagles et al. (12) and Chmielewski (13). Wardlaw et al. (14) estimate yield losses of 3 - 4 % for each degree exceeding 15°C during grain filling. On the other hand, this implies that low temperatures during yield formation may foster higher yields. At early phenological stages, like during emergence, high temperatures are required in order to allow a rapid and strong growth for young plants. Investigations by multiple linear regression analysis revealed yield losses at low temperatures (15) in early stages. Revaluation and devaluation of modelled yields had a dimension of 3.5% $°C^{-1}$ during heading and 2% $°C^{-1}$ during stem elongation and emergence. The calculation of temperature deviations was performed by a comparison of the long-term average (median 1990 - 2002) temperature and the actual temperature in a given phenological period and district.

Water stress was considered during stem elongation and heading. Especially the latter stage is highly sensitive to drought (16). A quantification of yield losses is estimable by the Crop Water Stress Index – CWSI (17), a relationship between ETa, the actual evapotranspiration and ETp the potential evapotranspiration.

For each stress factor an additional variable was introduced to control the magnitude of correction, which could theoretically range between 0 and 100 %. This variable allowed the model to be optimized using the Simplex method (18). Also interactions between the stress factors could be indirectly accounted for by these variables.

RESULTS

Land use classification

Since field plot size in South-West Germany is rather small, classification results for both pixel- and object-oriented classification algorithms are given separately for subsets of different field plot extents (table 1 and table 2). Obviously, the larger the field plots, the better the classification. Pixel-based and object-oriented classification results were better in the region of interest of BW. Generally, object-oriented classification led to better results in the smaller field plots in the regions of BW and RLP compared to the pixel-based approach, whereas results for larger field plot sizes did not differ much.

Table 1: Results of pixel-based classification of spring barley for both regions of interest (BW and RLP). Results are expressed as percentages of classified area.

	Rhineland-Palatinate (RLP)		Baden-Württemberg (BW)	
Statistics applied on:	n	Classification accuracy	n	Classification accuracy
All field plots	92	46 %	29	58 %
Field plots > 1,0 ha	60	49 %	20	61 %
Field plots > 2,0ha	33	57 %	11	69 %
Field plots > 2,5 ha	22	61 %	7	72 %

Table 2: Results of object-oriented classification of spring barley for both regions of interest (BW and RLP). Results are expressed as percentages of classified area.

	Rhineland-Palatinate (RLP)		Baden-Württemberg (BW)	
Statistics applied on:	n	Classification accuracy	n	Classification accuracy
All field plots	92	48 %	29	65 %
Field plots > 1,0 ha	60	51 %	20	69 %
Field plots > 2,0 ha	33	57 %	11	71 %
Field plots > 2,5 ha	22	60 %	7	73 %

Yield predictions

Results of the yield prediction by the basic linear correlation model with accumulated NDVIs by calendar units and GDDs are shown in Figure 3. Generally, GDD-based accumulation led to better results. Both strategies show relatively low mean deviations in respect to the reported yields (on an average < 11.5%). However, correlation coefficients of the worse results indicated low correlations and hence low stability of prognosis. The accumulation period of July seems to be inappropriate, because it is too close to maturity, whereas earlier time spans represent higher correlations and hence lower deviations. GDD-based accumulation gave best results for the period from ear emergence to yellow ripeness (grain filling period). At the end of this period yield predictions were possible with an average deviation of 7.0%. Similar results are possible by accumulating NDVIs in the assumed period of maximum photosynthetic activity, which represents a slightly earlier period, even though correlation was low in one observed district. Due to its low correlation the early period from stem elongation to ear emergence (GDD1) is not recommended for prognosis, whereas an extension to yellow ripeness (GDD4) stabilizes the relation.

The last column of prediction deviations of Figure 3 stands for the extended yield prediction results. The deviations are slightly improved regarding the basic yield prediction model (GDD2). The distribution of the deviations reveals the highest compactness with an average deviation of 6.4%.

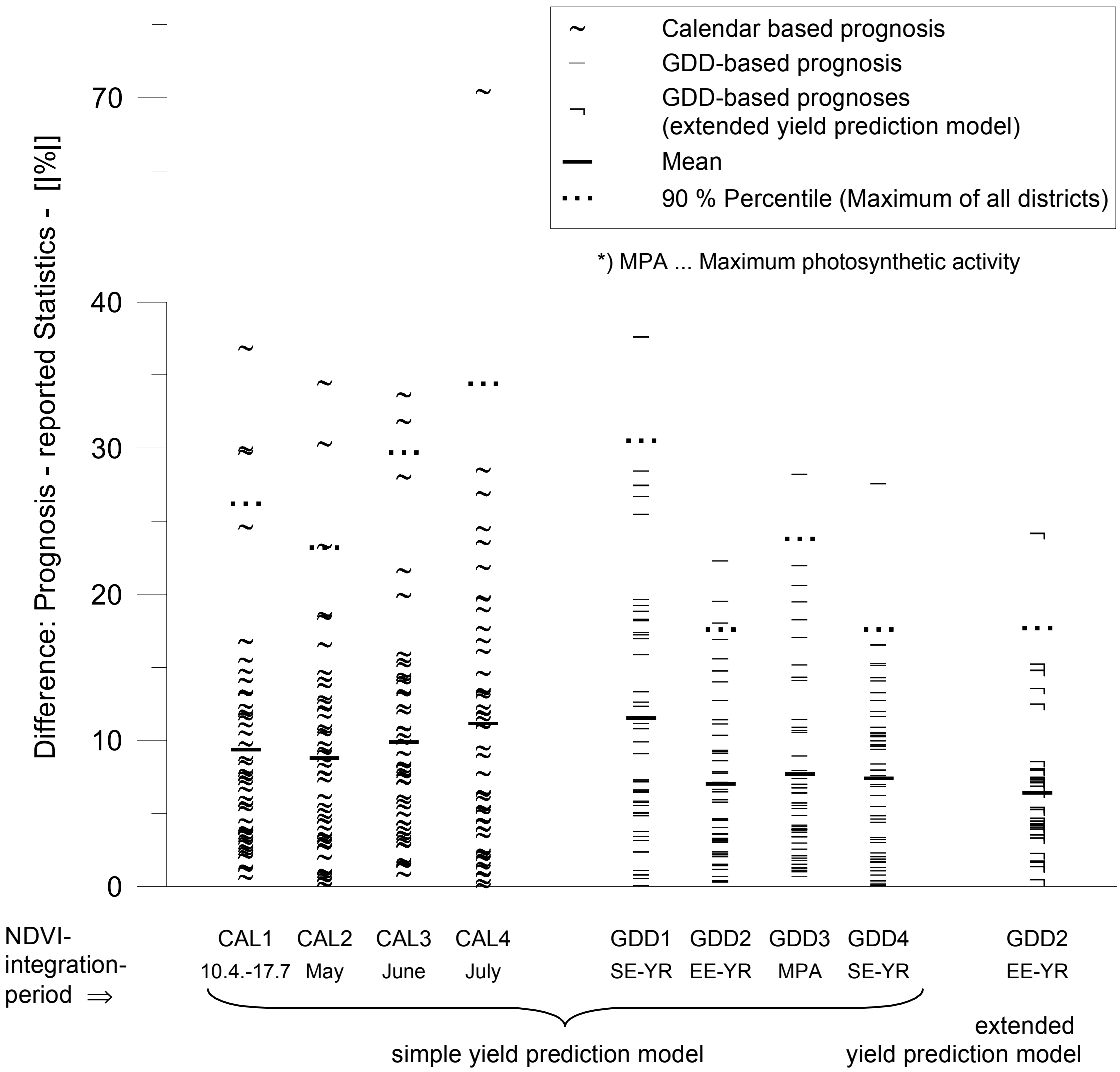


Figure 3: Prediction deviations for the simple (calendar and GDD based) and extended yield forecast model for different NDVI integration periods.

CONCLUSIONS

The results of the classification could clearly be improved when the field plot size was incremented. However, an area of 2 ha is desirable when 30 m resolution imagery is used in order to reach (producer's) classification accuracies of 60 % or more. The better results were generally achieved in the region of BW, which might be due to larger field plots and less variety in the cropping pattern. The object-oriented classification method led to an improved classification of small field plots in both regions BW and RLP. Although the classification accuracies were rather low, qualitative comparisons with regard to previous years allow the user to extract valuable information. It has to be noted that spring barley was classified and not, as in many land use classifications, a whole crop family (e. g. cereals) or barley comprising winter and spring barley. The high confusion associated with that and the fact that relatively coarse remote sensing data (in respect to the field plot size) was used led to the mentioned accuracy results.

Regarding yield predictions, the comparison between calendar-based and GDD-based accumulation periods for yield prediction is not possible in a consistent way, since periods of accumulation were not the same in both methodologies. Time spans GDD3 and CAL3 bear the highest resemblance to each other, the GDD-based accumulation performed clearly better in this case. This is plausible, since GDDs are indirectly taking into account also the phenology of crops.

Critical analysis of the resulting deviation of modelled yield from reported yield must consider also the robustness of relationships in terms of the correlation coefficient or the maximum deviation (herein expressed as 90 % percentile). This is essential, since also very weak relationships may lead to apparently reliable results. This requirement was best fulfilled by accumulation of NDVIs from ear emergence (EE) to yellow ripeness (YR), which comprises the heading stage and the grain filling period (see GDD2). Obviously, the earlier the accumulation period the worse the prediction, since events of interest which affect NDVI, occurring after this date, are not included. However, for some customers an early yield forecast might be of great value even though precision is lower. Accumulation of NDVIs in the assumed maximum photosynthetic activity (GDD3) focuses on this issue and allows a prediction at the end of June, while prediction results are still feasible.

Both the basic and the extended yield forecast models are a combination of remote sensing data and ancillary data. Although the extended model uses more ancillary data, the gathered results were not improved in the same relation as data input flow increased. A reason for this might be the fact that the NDVIs, which represent a copy of vegetation vigour, already include this information. Therefore, the use of additional meteorological and pedological data sets is to a certain extent probably leading to a double correction. However, impacts of drought, heat or cold might be hardly detectable by NDVI when they occur in a late or very early stage of the growing period.

The achieved results encourage an operational use of the forecast model, especially the basic yield prediction model, which may be categorized as a low cost model, due to the limited data requirements.

ACKNOWLEDGEMENTS

This work was supported by BMBF and DLR (DLR-50EE0106) within the project "Yield & Quality". The authors would like to thank the assisting farmers for supporting field investigations.

REFERENCES

1 Hall-Könives K (1990) Crop monitoring in Sweden. International Journal of Remote Sensing, 11, 461-484

2 Jewell N (1989) An evaluation of multi-date SPOT data for agriculture and land use mapping in the United Kingdom. International Journal of Remote Sensing, 10: 939-951

3 Atzberger C (1997) Estimates of winter wheat production through remote sensing and crop growth modelling - A case study in the Camargue region, Akademische Abhandlungen zu den Geowissenschaften, Berlin, VWF Verlag für Wissenschaft und Forschung

4 Genovese G, C Vignolles, T Nègre & G Passera (2001) A methodology for a combined use of normalised difference vegetation index and CORINE land cover data for crop yield monitoring and forecasting. A case study in Spain. Agronomie, 21: 91-111

5 Illera P, J A Delgado, A Fernández Unzueta & A A Fernández Manso (2000) Integration of NOAA-AVHRR and meteorological data in a GIS - Application for vegetation monitoring in Castilla y León, Spain. In: Remote Sensing in the 21st century: Economic and Environmental Applications, edited by J L Casanova (Balkema, Rotterdam), 47-54

6 Bochenek Z (2000) Operation use of NOAA data for crop condition assessment in Poland. In: Remote Sensing in the 21st century: Economic and Environmental Applications, edited by J L Casanova (Balkema, Rotterdam), 387-392

7 Swain P H & S M Davis (1978) Remote Sensing: The quantitative approach. (McGraw-Hill, New York), 166-174

8 Baatz M (2001) eCognition - object oriented image analysis - User Guide (Munich)

9 EEA (2003) CORINE land cover. 12/2000; European Environment Agency http://dataservice.eea.eu.int/

10 EU (2004) Introduction to the NUTS and the Statistical regions of Europe. European Union, 18-2-2004 http://europa.eu.int/comm/eurostat/ramon/nuts/home_regions_de.html

11 Wheeler T R, T D Hong, R H Ellis, G R Batts & J I L H P Morison (1996) The duration and rate of grain growth, and harvest index, of wheat (Triticum aestivum L.) in response to temperature and CO2. Journal of Experimental Botany, 47: 623-630

12 Eagles H A, A G Bedggood, J F Panozzo & P J Martin (1996) Cultivar and environmental effects on malting quality in barley. Australian Journal of Agricultural Research, 46: 831-844

13 Chmielewski F-M (1998) Der Einfluß der Witterung auf die Ertragsbildung von Sommergerste. Mitteilungen der Gesellschaft für Pflanzenbauwissenschaft, 11: 97-98

14 Wardlaw I F, I A Dawson, P Munibi & R Fewster (1989) The tolerance of wheat to high temperatures during reproductive growth. 1. Survey procedures and general response patterns. Australian Journal of Agricultural Research, 40: 1-13

15 Hünting K, C J Weissteiner & W Kühbauch (2003) Yield prediction of malting barley based on meteorological data. Proceedings of 2003 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Toulouse, France, July 21-25, 2003. 383-385

16 Doorenbos J & A H Kasam (1979) Yield response to water, FAO irrigation and drainage paper, 33, FAO publications (FAO, Rome)

17 Jackson R D (1982) Canopy temperature and crop water stress. In: Advances in irrigation, edited by D Hillel (Academic, New York), 43-85

18 Nelder J A & R Mead (1965) A simplex method for function minimization. Computer Journal, 7: 307-313