

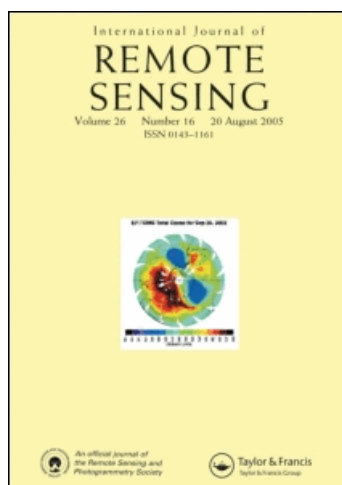
This article was downloaded by: [EBSCOHost EJS Content Distribution]

On: 15 September 2008

Access details: Access Details: [subscription number 902156990]

Publisher Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Remote Sensing

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713722504>

An agrometeorological-spectral model to estimate soybean yield, applied to southern Brazil

R. W. De Melo ^a; D. C. Fontana ^a; M. A. Berlato ^a; J. R. Ducati ^a

^a UFRGS - Faculdade de Agronomia - DPFA, Porto Alegre, RS, Brazil

Online Publication Date: 01 January 2008

To cite this Article De Melo, R. W., Fontana, D. C., Berlato, M. A. and Ducati, J. R. (2008) 'An agrometeorological-spectral model to estimate soybean yield, applied to southern Brazil', *International Journal of Remote Sensing*, 29:14, 4013 — 4028

To link to this Article: DOI: 10.1080/01431160701881905

URL: <http://dx.doi.org/10.1080/01431160701881905>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

An agrometeorological–spectral model to estimate soybean yield, applied to southern Brazil

R. W. DE MELO*, D. C. FONTANA, M. A. BERLATO and J. R. DUCATI
UFRGS – Faculdade de Agronomia – DPFA, Av. Bento Gonçalves, 7712, CX. Postal
15100, CEP 91501-970, Porto Alegre, RS, Brazil

(Received 18 September 2005; in final form 21 December 2007)

Soybean yield is modelled from data gathered from crops in Rio Grande do Sul State, Brazil. The model comprises an agrometeorological term, obtained by adjusting the multiplicative model of Jensen, modified by Berlato, and a spectral term, obtained from National Oceanic and Atmospheric Administration (NOAA) satellite images of the maximum Normalized Difference Vegetation Index (NDVI). The weather data used to calculate the relative evapotranspiration (ET_r/ET_0) cover the period from 1975 to 2000, and the NDVI/NOAA images were obtained from 1982 to 2000. Application of the agrometeorological–spectral model produced better yield estimates (of about 5%) than Jensen's model, allowing the further generation of yield maps for the most significant soybean production regions within the Rio Grande do Sul State.

1. Introduction: estimating crop yields

The accurate estimation of crop yield with adequate time prior to the harvesting period, either on a regional or a national scale, provides valuable information. Indeed, a knowledge of the future availability of agricultural commodities is crucial in an organized economy, and strategic to a nation's resources management. Conventional methods include polls conducted close to farmers, producers, cooperatives, suppliers of seeds and fertilizers, and other links of the production chain. These methods, being partially based on subjective information, are not very accurate and may even be subject to manipulation aimed at influencing market expectations. More objective approaches can be developed based on the assumption that harvest yields are strongly dependent on Nature itself, provided that standard farming practices are adopted. Such modern methods using direct environmental information are alternative, or complementary, to conventional methods of predicting harvests. Here we apply such a development to soybean crops in Rio Grande do Sul State, southern Brazil.

As the southernmost state in Brazil, Rio Grande do Sul (approximately 53°W, 30°S) harvests about 20% of the country's soybean crop. The production area is concentrated primarily in the northwest of the state, where 3 million hectares (8.3 million acres) produced 7 million metric tons in more than 140 000 farms, with sizes ranging from small to large according to 2000–2001 statistics. Local environmental conditions for this culture are adequate in some years, but, in most

*Corresponding author. Email: melo_rw@yahoo.com.br

cycles, frequency and intensity of rainfall at crucial times do not allow plants to attain their full productive potential (Matzenauer *et al.* 2002).

The relationship between water input, plant growth and grain formation is well known, and can be used to model crop yield. The so-called agrometeorological models have been proposed to estimate crop yields, with soil water availability as the sole independent variable. This type of modelling has its origins in the studies performed by Shantz and Piemeisel (1927), where a high correlation between transpiration and plant yield was revealed, explained by the strong association between photosynthesis and transpiration. Later, De Witt (1958) proposed estimating yield with the inclusion of the atmosphere's evaporative demand, that is using the relationship between transpiration and potential evaporation. Because of the practical problem of determination of transpiration, Jensen (1968) proposed the substitution of transpiration by the maximum evapotranspiration, linking the relative crop yield with the relative evapotranspiration (ET_r/ET_m , where ET_r is the real evapotranspiration and ET_m is the maximum evapotranspiration), and giving different weights for the different plant development phases. Since then, this model has been used extensively. Berlato (1987) obtained a very good fit with soybean data, making adjustments and validations on the agrometeorological model of Jensen, by replacing the ET_m parameter by ET_0 (reference evapotranspiration) and used data from experimental plots. This change was introduced to simplify the model for regional applications as the crop coefficient (K_c) is not used. In this model, each phenological stage has its own weight, according to plant sensitivity to water deficit. From these models it has been shown that the subperiod from flowering to grain filling is the most dependent on water deficit, having more influence on the definition of grain yield. Studies on sunflowers and maize in the Rio Grande do Sul region have been performed by Barni *et al.* (1996) and Matzenauer *et al.* (1995), respectively.

Fontana *et al.* (2001), using official data released by the Instituto Brasileiro de Geografia e Estatística (IBGE), extended the Jensen/Berlato agrometeorological model for soybeans to field conditions. The results were very promising, indicating that the use of agrometeorological models in harvest forecast presents advantages in terms of simplicity, objectivity and costs, compared with conventional methods.

In addition to agrometeorological models, it is possible to introduce specific information on plant conditions using modern techniques of remote sensing. This is based on the fact that vegetation growth and development are related to the absorbed solar energy, and may be associated with the reflected energy and collected by remote sensors under proper calibrations. Much of the investigation in this field uses vegetation indexes expressed by the ratio between the reflectances in the visible and infrared spectra, as responses in these spectral regions have different behaviours as the plant grows (Baret and Guyot 1991). The phenological development results in structural alterations in plants, leading to progressive changes in reflectances that define a spectral profile, characteristic of each vegetal species. Furthermore, alterations in a species spectral profile can be associated with its health condition (Justice *et al.* 1991; Eidenshink and Hass 1992). In general, it was observed that vegetation indexes correlated well with agronomical parameters such as yield because they express biomass evolution.

Remote sensing products including satellite images that carry information in several spectral bands are suited to monitor vegetation growth conditions along its cycle and also to estimate yield (Reynolds *et al.* 2000; Liu and Kogan 2002;

Kalubarme *et al.* 2003; Ferencz *et al.* 2004). This is best performed by satellites with a short repeat cycle, such as the National Oceanic and Atmospheric Administration (NOAA) spacecrafts, which acquire images at a rate of several times a day. Justice *et al.* (1985), Batista *et al.* (1993), Motta *et al.* (2003) and, more recently, Liang *et al.* (2005) have shown that the use of Normalized Difference Vegetation Index (NDVI) images is a suitable tool for monitoring the evolution of canopy. Fontana *et al.* (1999) tested some methodologies of monitoring and forecasting harvests in Rio Grande do Sul State and observed that the time evolution of the NDVI is associated with the density of biomass over the land. The relationship between biomass and yield is logical in the sense that high yields are only attained when high levels of biomass have accrued; however, the inverse is not always true because high biomass is not a guarantee of high yields, especially in certain critical periods, such as water deficit in the reproductive phase. A connection linking biomass and yield may be made and was suggested by Zhong-Hu and Rajaram (1993).

Including spectral information in modelling yields adds non-negligible gains. In fact, the agrometeorological model expresses conditions of input of solar radiation, temperature, air humidity and available water. Additionally, the spectral component, besides these factors, expresses constraints and stresses, which are not considered in the agrometeorological component, such as differences in farming practices, varieties, and root depth (Rudorf and Batista 1990). The combination of agrometeorological models with data from remote sensing leads to models currently referred to as agrometeorological–spectral.

This model is additive, even if the spectral term is not independent of the factors that determine the value assumed by the agrometeorological term. This means that the effect of the meteorological conditions in crop growing, especially the water availability, is also present in the spectral term. However, this last term contains other factors that are not present in the agrometeorological term but will have an influence on the final yield. In fact, the spectral information is strongly linked to water availability in an earlier period (December and January) with respect to the onset of conditions that are more important to the agrometeorological term; these conditions prevail from January to March.

2. Methodology

The agrometeorological–spectral model for yield estimation is of the form:

$$Y = a + bA + cS \quad (1)$$

where Y is the estimated mean yield, A and S are the agrometeorological and spectral terms, respectively, and a , b and c are the model parameters. The agrometeorological term, A , is the product of a term containing the relative evapotranspiration from the multiplicative model of Jensen (1968), modified by Berlatto (1987), and a correction factor, F , in the form:

$$A = FY_A \quad (2)$$

The two terms forming A (F and Y_A) are obtained as follows. Y_A is given, for each year, by:

$$Y_A = Y_m \prod_{i=1}^n (E)_i^{\lambda_i} \quad (3)$$

where Y_m is the maximum yield observed in the series of observed crops, E is the

relative evapotranspiration (the ratio between the real and the reference evapotranspirations, ET_r/ET_0) in each period i in the crop cycle, and λ_i is the plant relative sensibility to water stress in period i .

The i periods are the months of January, February and March, where the correlations between the relative evapotranspiration and yield are more significant. The λ_i exponents were determined by multiple regression fitting of the logarithm transforms of the relative yield (Y/Y_m) and the relative evapotranspiration (ET_r/ET_0), with the zero-point passing through the coordinates origin, Y being the mean observed yield in each year of the series. From the soybean yield database of IBGE, which spans a 26-year period (1975–2000), 21 years were chosen for the model fitting, and 5 years (1982, 1985, 1990, 1991 and 1997) were taken randomly to validate the model. This random choice is not critical to the estimated yield, as discussed in section 5.

The adopted value for Y_m was the maximum mean yield observed during the period of study (2088 kg ha⁻¹ in 1998); it must be stressed that yields in that period did not present a significant temporal tendency.

Meteorological data for the model fitting came from stations installed in the soybean region (at Cruz Alta, Erechim, Iraí, Júlio de Castilhos, Passo Fundo, Santa Rosa and São Luiz Gonzaga counties), operated by either the Instituto Nacional de Meteorologia (INMET) or the Fundação Estadual de Pesquisa Agropecuária (FEPAGRO). The distribution of these seven stations over the production region is shown in figure 1. The region's climate is the *Cfa* type from Köppen's classification (1948), and is characterized as wet subtropical, with a mean annual temperature of 18.7°C and oscillations of about $\pm 5^\circ\text{C}$ over seasons. Rains are well distributed throughout the year, with a mean annual precipitation of 1680 mm.

Data on the maximum, minimum and mean air temperature, relative humidity, wind speed, rainfall, and sunlight in 10-day periods were used to calculate ET_0 , according to the Penman (1956) method. ET_r was determined from the water meteorological balance of Thornthwaite and Mather (1955), taking 75 mm as the available water capacity. Monthly ET_r/ET_0 values derived for each meteorological station were extended to the whole production region by spatial interpolation using the Kriging method. A grid of relative evapotranspiration values was generated, and these were transformed into images of ET_r/ET_0 with 9 km $\times$ 9 km pixels.

The estimated annual soybean yield, Y_A , from the modified Jensen model was derived from equation (3) and provided information on a pixel level. These images were generated only for the 1982–2000 period because satellite images used in the model's spectral component were not available prior to 1982.

The correction factor, F , in equation (2) is necessary because the Y used in the λ_i derivation is the mean observed annual yield. The resulting Y_A for each pixel will potentially contain super- or underestimates because each pixel in the area covered has its own yield. The information on the observed yield, Y , at the pixel level, comes from official data at the county level, which were treated to produce a raster file with the same 9 km spatial resolution of the Y_A image. The correction factor for each pixel was derived from the mean, calculated using all years of the fitting series of the ratios in each year between the estimated yield, Y_A , and observed yields, Y . This factor expresses local conditions, at the pixel level, leading to mean pixel yields that are different from the average yield for the whole region.

The spectral term, S , was derived using an image database compiled by Clark Laboratories from data collected by the Advanced Very High Resolution

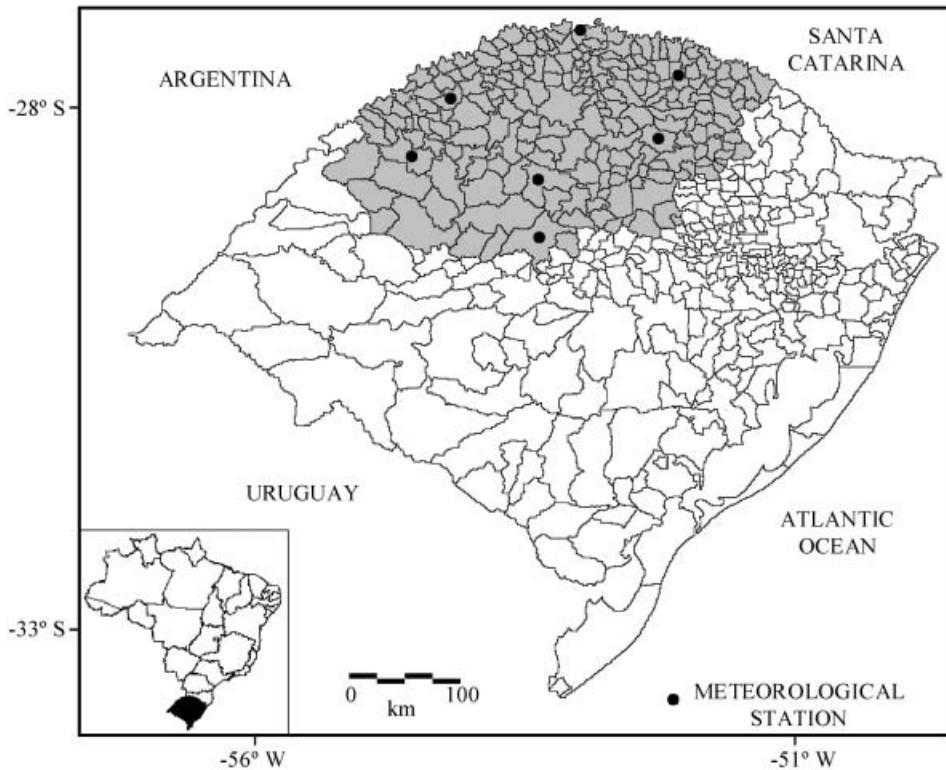


Figure 1. Distribution of the meteorological stations that provided data for this study. Also shown is the county division in Rio Grande do Sul State. The grey colour represents the significant soybean production region (Berlato and Fontana 1999).

Radiometer (AVHRR)/NOAA sensor, with radiometric and geometric corrections. In this database, the original pixels in an NOAA image, with 1100 m resolution at nadir, were resampled to 9 km $\times$ 9 km pixels over the entire image. Information is expressed in digital counts. These images spanning the 1982–2000 period were cut to generate 72 line–86 column subimages covering the Rio Grande do Sul State, with 9 km resolution. The NDVI was derived using the expression

$$\text{NDVI} = [(0.93/253) \times N_{\text{img}}] - 0.2 \quad (4)$$

where N_{img} is the NDVI value in the image, expressed in digital counts at the pixel level. This covers the (0, 253) range because two bits of the original image were reserved to express non-land targets or lack of data. S is defined as the average of the maximum monthly NDVI for the months (December and January) where the correlation between the NDVI and yield was significant within a 5% probability. Thus:

$$S = (N_{\text{Dec}} + N_{\text{Jan}})/2 \quad (5)$$

where N_{Dec} and N_{Jan} are the NDVI values in December and January, respectively. The addition of this spectral term led to a new yield estimate (equation (1)). The 1982–2000 period was used for the model fitting, with the exception of 1995 because of lack of satellite images. Model validation was achieved with the same years used

to validate the agrometeorological term. With these results in a pixel basis, it is possible to calculate the differences between estimated and observed yields for every pixel at every year.

The last part of the study was the evaluation of model stability and accuracy of estimates. It is an important reminder that the model was performed over a random choice of initial conditions (fitting and validation years). To investigate the model dependency on this choice, the model was run for 31 other initial conditions as follows. Again, year 1995 was not included in the series because of lack of orbital data:

- (a) fitting performed with the 13 years of maximum yield, and another 5 years for validation (model Max);
- (b) fitting similarly performed with the 13 years of minimum yield (model Min);
- (c) fitted years are the first 13 years in the 1982–2000 series (model 13a);
- (d) fitted years are 1983–1996, 1984–1997, and so on up to 1987–2000 (models 13b to 13f);
- (e) 14 years are taken for the fitting, starting with the period from 1982–1997, up to 1986–2000 (models 14a to 14e);
- (f) 15 fitting years, sets 1982–1997 to 1985–2000 (models 15a to 15d);
- (g) 16 fitting years, sets 1982–1998 to 1984–2000 (models 16a to 16c);
- (h) 17 fitting years, sets 1982–1999 and 1983–2000 (models 17a and 17b);
- (i) 18 fitting years, no validation years (model 18a);
- (j) finally, a random extraction of sets of fitting data (13 years) and validation data (5 years). This was performed for an additional eight models (13g to 13n), and was not coincident with models 13a to 13f, or with the original computation leading to equation (6).

3. Results

Determination of the agrometeorological term (equation (3)) was executed with data on relative yields Y/Y_m and is presented in table 1, with 1998 as the reference year (maximum yield in the period: 2088 kg ha⁻¹). In table 1, data on relative evapotranspiration are presented for months (January, February and March) where ET_r/ET_0 showed the best correlation coefficients with soybean yield (0.736, 0.514 and 0.609, respectively). Table 1 reveals the high interannual variability of soil water availability in Rio Grande do Sul State, 1998 being the sole year where ET_r/ET_0 exceeded 0.95 in all three months. The resulting λ_i values are presented in table 2, where it can be seen that yields are mostly defined in March, the stage of grain filling. Explanations for the relatively low value of λ in February call for further studies in the context of the phenology and critical periods of soybean.

Estimated yields through this modified Jensen model across the entire study area are shown in figure 2, both for those years used for model fitting and for the validation years. The correction factor F appearing in equation (2) is stable over time but varies spatially. This is presented in figure 3, where lighter grey shades represent regions where the modified Jensen model overestimates yields, a reducing correction being mediated by applying the F factor; darker grey shades express yield underestimates, prone to similar corrections. Finally, table 3 shows the mean values of the agrometeorological term (yields in kg ha⁻¹) for the region covered in this study, noting that this information is given only for the 1982–2000 period, as

Table 1. Relative soybean yields and values of relative evapotranspiration (ET_r/ET_0) for soybean, in January, February and March, in the region of significant production in Rio Grande do Sul State.

Year	Relative soybean yield	ET_r/ET_0			Year	Relative soybean yield	ET_r/ET_0		
		January	February	March			January	February	March
1975	0.757	0.919	0.880	0.962	1988	0.538	0.729	0.684	0.483
1976	0.787	0.940	0.854	0.912	1989	0.879	0.928	0.763	0.905
1977	0.811	0.908	0.863	0.759	1990*	0.874	0.888	0.800	0.905
1978	0.606	0.591	0.452	0.578	1991*	0.325	0.473	0.421	0.511
1979	0.407	0.178	0.694	0.724	1992	0.954	0.747	0.974	0.995
1980	0.724	0.626	0.446	0.926	1993	0.958	0.986	0.778	0.948
1981	0.793	0.848	0.896	0.629	1994	0.831	0.599	0.992	0.830
1982*	0.569	0.308	0.887	0.615	1995	0.953	0.863	0.807	0.833
1983	0.786	0.772	0.887	0.851	1996	0.816	0.835	0.949	0.861
1984	0.743	0.951	0.881	0.786	1997*	0.775	0.841	0.940	0.662
1985*	0.779	0.374	0.916	0.874	1998	1.000	0.951	0.968	0.965
1986	0.499	0.592	0.737	0.840	1999	0.699	0.627	0.902	0.674
1987	0.791	0.857	0.899	0.689	2000	0.787	0.761	0.749	0.858

*Years used for validation.

Table 2. Statistics for derived exponents of the modified Jensen model for soybean yield estimates, Rio Grande do Sul State, during 1975–2000.

Month	Exponent (λ)	Probability (p)
January	0.401	0.0001
February	0.192	0.1536
March	0.475	0.0025

explained in section 2. The yield for 1995 is not provided because satellite images were not available.

Regarding the spectral term, figure 4 shows the mean time evolution of the NDVI along the time course. The maximum NDVI in February and March coincides with the flowering and grain filling periods, where the cultures present the highest biomass density. However, NDVI correlations with yield (figure 5) are highest for December (0.424) and January (0.517), which correspond to the beginning of flowering. An even higher correlation (0.77) was found for the same months by Fontana (1995) between soybean yield and Global Vegetation Index (GVI) from a smaller (4-year) database. Nevertheless, both February and March are months that largely influence yield, with significant correlations expected between NDVI and yield. The lack of correlation may result from a lack of sensitivity of the NDVI to detect small variations in biomass when the Leaf Area Index (LAI) is high. This was, in fact, suggested by other investigators (Gamon *et al.* 1995, Fonseca *et al.* 2002).

Table 3 presents the values for the spectral term used for the fitting and validation of the agrometeorological/spectral model. It is apparent that the mean values of the spectral term do not vary significantly along the time course, even if these are means for the whole region. Greater spatial and temporal variations inside the area do exist.

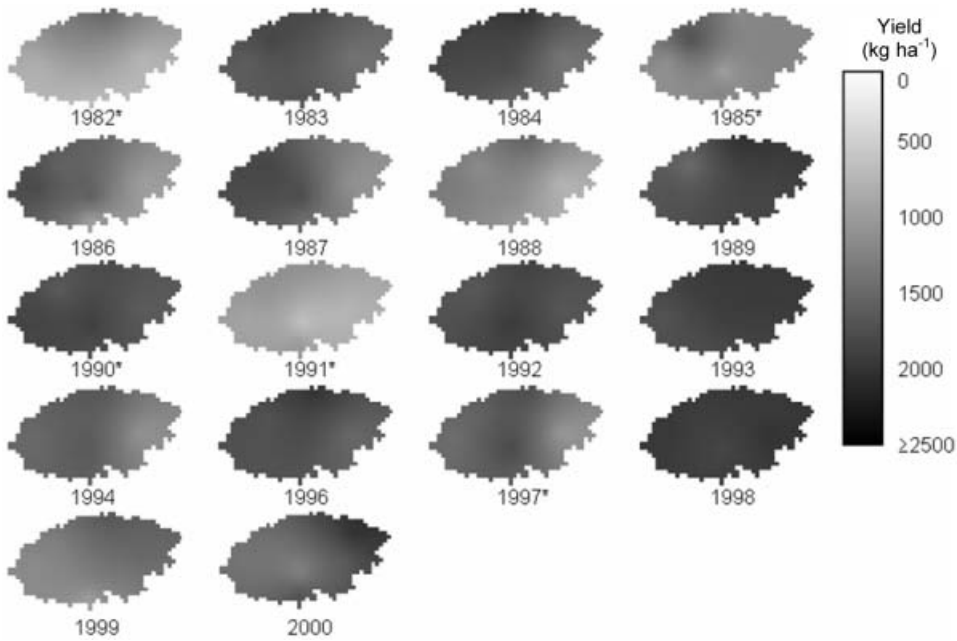


Figure 2. Soybean yields (kg ha^{-1}) estimated by the modified Jensen model, as a function of the observed yields, considering fitting years and validation years. Validation years are marked by an asterisk.



Figure 3. Values of the correction factor derived for the agrometeorological term over the study area.

The above results, obtained through multiple linear regression, led to the parameters of the model, which are presented in table 4. The agrometeorological–spectral model used to estimate the soybean yield in Rio Grande do Sul State is

$$Y = -2.63487 + 1084A + 4.63420S \tag{6}$$

Table 3. Values of the agrometeorological term (A) and of the spectral term (S) of the agrometeorological–spectral model for soybean yield estimation in Rio Grande do Sul State.

Year	A	S	Year	A	S	Year	A	S
1982*	937	0.580	1989	1780	0.549	1995		
1983	1594	0.553	1990*	1726	0.565	1996	1688	0.514
1984	1656	0.553	1991*	870	0.549	1997*	1459	0.564
1985*	1221	0.579	1992	1758	0.557	1998	1913	0.576
1986	1327	0.501	1993	1866	0.587	1999	1337	0.564
1987	1467	0.566	1994	1427	0.589	2000	1567	0.561
1988	1093	0.564						

*Years used for validation.

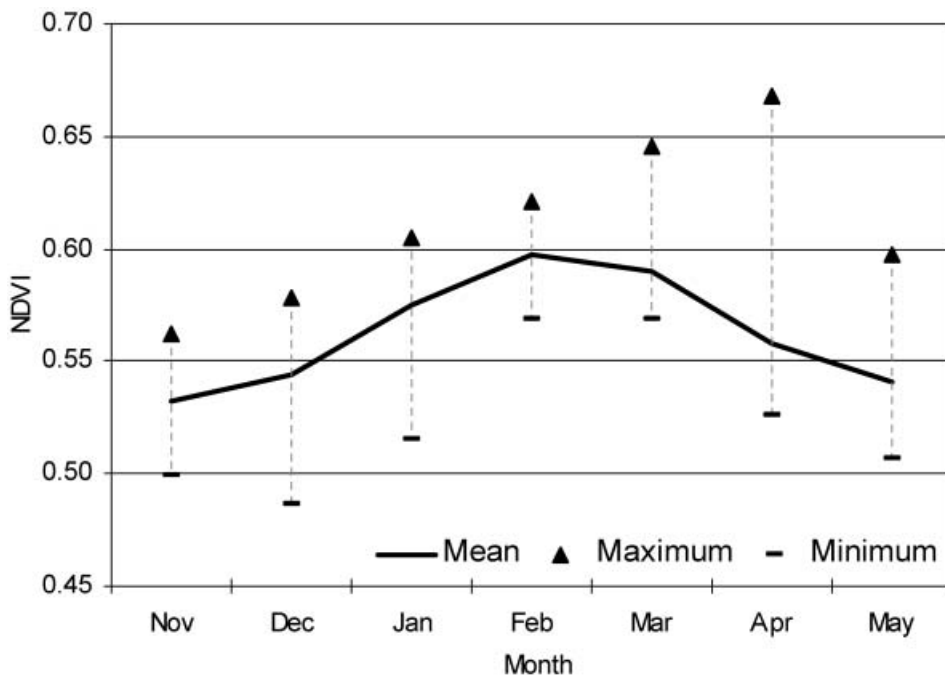


Figure 4. Time evolution of the NDVI in the soybean production region in Rio Grande do Sul State (1998–2000).

Figure 6 presents the images of yield estimates and shows the interannual and spatial yield variability inside the region. Table 5 presents the mean soybean yields obtained by the agrometeorological–spectral model.

As discussed in section 2, model stability may be estimated by running different sets of initial conditions. Thirty-one different sets of data were used and they are given in table 6.

4. Discussion

The comparison between estimated and observed yields is presented in figure 7, and the correlation coefficients (R) for the fitting and validation sets are 0.96 and 0.94, respectively. The correlation between predicted and observed yields for the complete

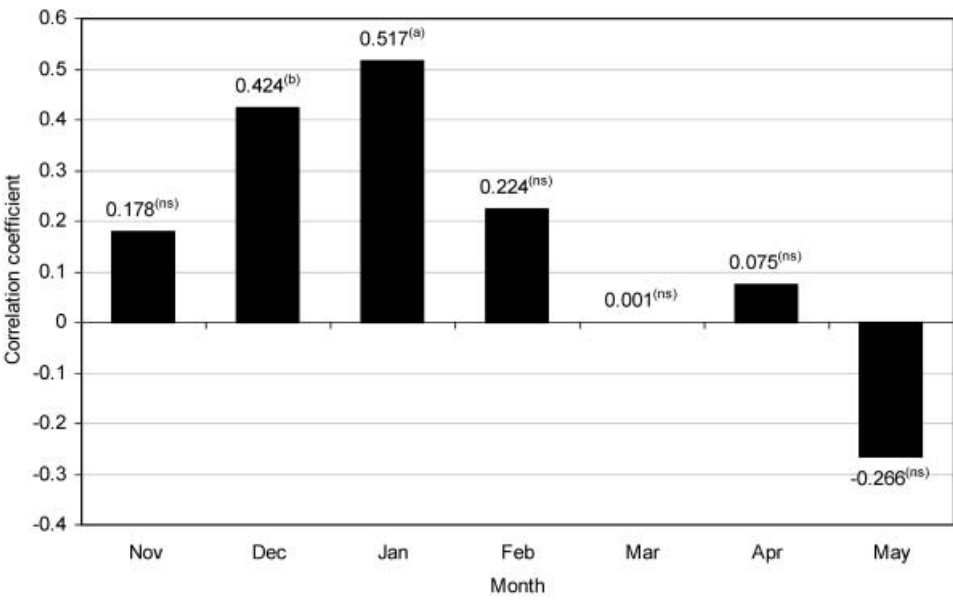


Figure 5. Correlation coefficients between NDVI values and average soybean yields, 1982–2000. Coefficients significant to 5% (a), 10% (b) and non-significant (ns) are shown.

Table 4. Statistics of parameters of the agrometeorological–spectral model for soybean yield estimation, for Rio Grande do Sul State, during 1982–2000.

Term	Parameter	Probability (<i>p</i>)
Intercept	–2634.87	0.002313
Agrometeorological	1.08	0.000005
Spectral	4634.20	0.002669

series of data (fitting and validation years) is 0.95. These results are very promising when compared with those obtained by other investigators. Rudorff and Batista (1990) found coefficients varying from 0.50 to 0.69, working with sugarcane cultures in São Paulo State. Fontana and and Berlato (1998), studying soybean in Rio Grande do Sul State, verified that the agrometeorological–spectral model was able to explain 55.2% of the variations in yield for this culture; these results, however, were preliminary because of the smaller database. Liu and Kogan (2002), using AVHRR/NOAA images to estimate soybean yield in several regions in Brazil, found a correlation coefficient of 0.26 for Rio Grande do Sul State.

It must be stressed that this model is predictive, as both terms use information collected up to the month of March, leading to a yield estimate well over a month before harvest.

Figure 8 shows the behaviour of yield estimates from the agrometeorological term and from the agrometeorological–spectral model compared with the observed yields in the period of study. It is noted that the estimated yields follow the high interannual variability of the observed yields in Rio Grande do Sul State. The results from the agrometeorological term can be compared with those from the agrometeorological–spectral model. The spectral term in fact reduces deviations in

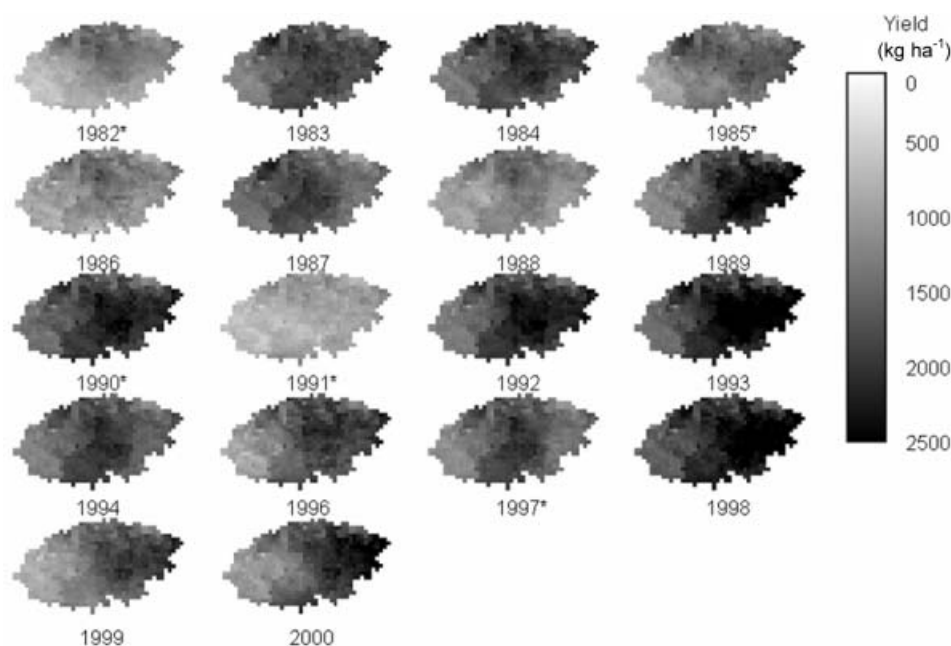


Figure 6. Soybean yields (kg ha^{-1}) estimated by the agrometeorological-spectral model, in the region of significant production in Rio Grande do Sul State, from 1982–2000. Validation years are marked by an asterisk.

Table 5. Mean yields for soybean, at the region of significant production in Rio Grande do Sul State, estimated by the agrometeorological-spectral model.

Year	Yield (kg ha^{-1})	Year	Yield (kg ha^{-1})	Year	Yield (kg ha^{-1})
1982*	1067	1989	1840	1995	
1983	1655	1990*	1853	1996	1577
1984	1725	1991*	853	1997*	1559
1985*	1373	1992	1853	1998	2107
1986	1125	1993	2105	1999	1429
1987	1581	1994	1643	2000	1664
1988	1164				

*Years used for validation.

estimates. Modelling using only the agrometeorological term compares to observed yields with an accuracy of 93%; the introduction of the spectral term leads to an accuracy of 98%. In two periods, 1987 to 1992 and 1997 to 2000, the yield tracings are very similar, with differences being less than 200 kg ha^{-1} .

Inspection of table 6 reveals that the agrometeorological-spectral model is fairly stable with respect to choices of fitting and validation years. Table 6 is also useful in helping to assess the uncertainties associated with the estimated yields. Column 6 of table 6 gives the mean percentage deviation of estimated yield for each simulation; values are concentrated around 6.31%, or about 88 kg ha^{-1} . These errors contain a component coming from the drifting of the choices for fitting years, the value from this component is of the order of 1%, which is the mean oscillation of values in column 6.

Table 6. Simulations made to test for the model stability. Each simulation/model uses different sets of fitting and validation years.

Model	Average yield (kg ha ⁻¹)	Mean deviation	Correlation coefficient (<i>R</i>)	Absolute mean deviation (kg ha ⁻¹)	Mean deviation (%)	Parameters		
						<i>a</i>	<i>b</i>	<i>c</i>
Obs.l	1578	365						
AS	1565	346	0.952	87	6.34	-2634.87	1.084	4634.20
13g	1592	370	0.955	87	6.04	-3446.01	1.129	6011.20
13h	1567	342	0.949	82	6.01	-3854.97	0.995	7053.44
13i	1539	376	0.949	94	6.41	-2689.28	1.188	4409.66
13j	1536	372	0.955	90	6.03	-3538.31	1.136	6057.20
13k	1579	381	0.951	88	6.04	-4297.56	1.120	7534.79
13l	1601	336	0.954	83	6.36	-3174.88	1.015	5845.01
13m	1568	307	0.943	88	6.82	-3606.90	0.869	6944.08
13n	1554	355	0.952	87	6.18	-2733.73	1.112	4715.77
Max	1647	238	0.954	116	10.37	-1536.17	0.730	3754.23
Min	1599	381	0.954	88	5.98	-3839.14	1.152	6665.99
13a	1564	356	0.951	84	6.01	-3947.91	1.046	7079.55
13b	1573	364	0.954	86	6.03	-3188.50	1.123	5532.22
13c	1578	365	0.954	87	6.08	-3203.17	1.124	5566.75
13d	1587	364	0.954	89	6.31	-3053.76	1.128	5304.63
13e	1566	379	0.949	95	6.48	-2724.33	1.198	4494.01
13f	1572	371	0.941	99	7.02	-2111.62	1.189	3434.26
14a	1580	354	0.955	83	5.99	-3268.25	1.080	5803.33
14b	1576	364	0.954	86	6.05	-3208.28	1.122	5578.93
14c	1574	358	0.954	85	6.07	-3116.27	1.104	5455.72
14d	1588	364	0.954	89	6.32	-3056.01	1.127	5314.35
14e	1564	379	0.949	95	6.47	-2718.24	1.197	4482.05
15a	1583	354	0.955	83	6.01	-3281.10	1.080	5830.51
15b	1573	358	0.954	85	6.05	-3122.90	1.103	5468.60
15c	1575	358	0.954	85	6.08	-3120.58	1.101	5473.70
15d	1585	363	0.954	89	6.28	-3050.51	1.124	5305.56
16a	1580	349	0.955	83	6.01	-3202.40	1.065	5723.49
16b	1574	357	0.954	85	6.06	-3127.78	1.100	5488.20
16c	1573	357	0.954	85	6.06	-3115.85	1.100	5465.42
17a	1580	349	0.955	83	6.02	-3203.24	1.065	5727.68
17b	1572	357	0.954	85	6.04	-3122.43	1.099	5478.63
18a	1578	348	0.955	82	5.99	-3198.47	1.063	5720.70

This model is based heavily on use of observational data, which, by definition, carry errors or uncertainties. Errors of an external nature come from the sources of data used to evaluate both *S* and *A* terms, and the methods of data handling. The agrometeorological term, *A*, is derived from meteorological information, which itself carries observational and instrumental errors, collected from a finite number of ground stations. This spatially discrete distribution of data is made continuous by interpolation techniques to produce maps on a pixel basis and this certainly introduces additional errors, which are difficult to assess. Even if the meteorological data came from a relatively sparse network of ground stations, the results indicate that the interpolation procedure would lead to satisfactory estimates, suggesting that the model is valid and will yield improved figures if more stations become available. The *A*, *S* and *Y* images generated all had a 9 km resolution, compatible with the NDVI images. The yield images, *Y*, cover large territories, and yield information can be extracted for smaller areas, at the county scale. Uncertainties of

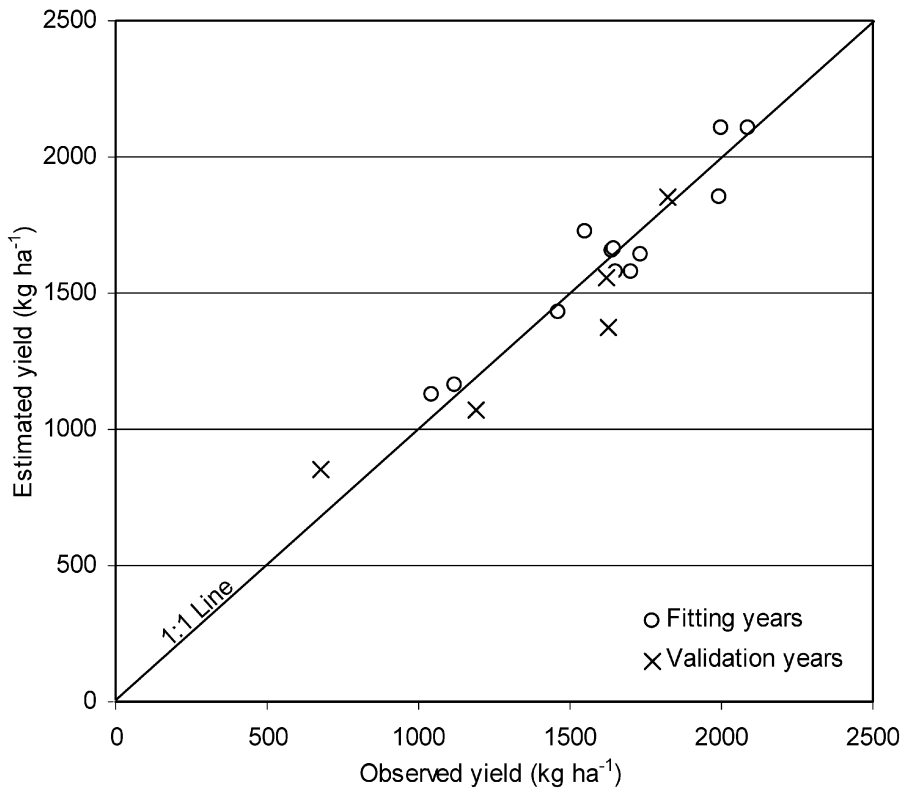


Figure 7. Soybean yields estimated by the agrometeorological-spectral model, as a function of observed yields, considering fitting ($R=0.96$) and validation ($R=0.94$) years (1982–2000).

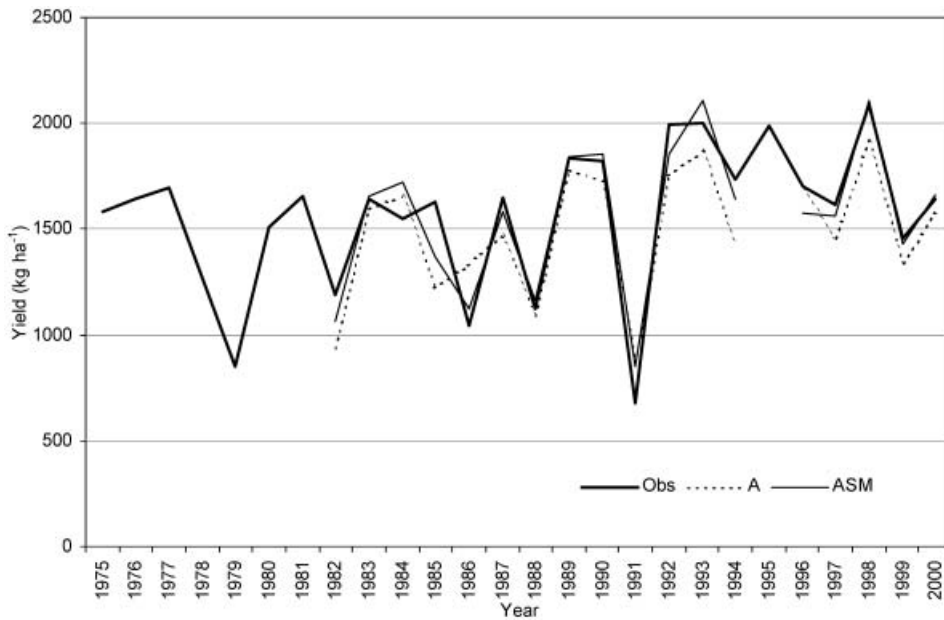


Figure 8. Soybean yield estimates in Rio Grande do Sul State, obtained from the agrometeorological term (A), and by the agrometeorological-spectral model (ASM), compared with the observed yield (Obs), along the observed time course.

the model also arise from the spectral term, which carries the errors contained in the calibration given in equation (4). The way those uncertainties accumulate, when put together, is complex, and the more straightforward way to reach an estimate on the model's quality is a direct comparison involving estimated and observed yields. This is presented in table 6 and it is apparent from its column 2 that the observed yield (1578 kg ha^{-1}) is well within the range of estimated values.

5. Conclusions

The agrometeorological-spectral model is a promising estimator of soybean yields and deserves further development, including the addition of more recent data to the database. The introduction of the spectral term, even if of a smaller impact when compared to the agrometeorological term, resulted in a significant (about 5%) improvement in yield estimates.

References

- BARET, F. and GUYOT, G., 1991, Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, **35**, pp. 161–173.
- BARNI, N.A., BERLATO, M.A., BERGAMASCHI, H. and RIBOLDI, J., 1996, Modelo grometeorológico de predição do rendimento do girassol. I. Relação entre rendimento e índice hídrico. (Agrometeorological model of sunflower yield estimative. I. Relation between yield and water index.) *Pesquisa Agropecuária Gaúcha*, Porto Alegre, **2**, pp. 7–17.
- BATISTA, G.T., SHIMABUKURO, Y.E. and LAWRENCE, W.T., 1993, Monitoramento da cobertura vegetal através de índices de vegetação do NOAA-AVHRR. (Monitoring of vegetation through vegetation indices of the NOAA-AVHRR.) In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 7, Curitiba. *Anais...* (São José dos Campos: INPE), pp. 30–37.
- BERLATO, M.A., 1987, *Modelo de relação entre o rendimento de grãos de soja e deficiência hídrica para o Estado do Rio Grande do Sul* (Model of the relationship between of the grain soybeans yield and water deficit for the state of Rio Grande do Sul). 103f. Thesis (Doctorate in Meteorology) - Pós Graduação Meteorologia, Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- DE WITT, C.T., 1958, *Transpiration and crop yields* (Wageningen: Institute of Biological and Chemical Research on Field Crop and Herbage) erse-Landbowk, Onder Z.n°64.6-S Gravenhage.
- EIDENSHINK, J.C. and HAAS, R.H., 1992, Analyzing vegetation dynamics of land systems with satellite data. *Geocarto International*, **1**, pp. 53–61.
- FERENCZ, Cs., BOGNÁR, P., LICHTENBERGER, J., HAMAR, D., TARCSAI, Gy., TIMAR, G., MOLNÁR, G., PÁSZTOR, Sz., STEINBACH, P., SZÉKELY, B., E. FERENCZ, O. and FERENCZ-&AACUTE;RKOS, I., 2004, Crop yield estimation by satellite remote sensing. *International Journal of Remote Sensing*, **25**, pp. 4113–4149.
- FONSECA, E.L., ROSA, L.M.G. and FONTANA, D.C., 2002, Caracterização espectral de *Paspalum notatum* em diferentes níveis de adubação nitrogenada. (Spectral characterization of *Paspalum notatum* under different nitrogen fertilization) *Pesquisa Agropecuária Brasileira*, **37**, pp. 365–371.
- FONTANA, D.C., 1995, *Índice de vegetação global para o monitoramento da vegetação e sua correlação com elementos agrometeorológicos e rendimento de grãos de soja* (Global vegetation index for the monitoring of the vegetation and its correlation with agrometeorological elements and soybean grain yield). Thesis (Doctorate in Plant Sciences) - Programa de Pós-Graduação em Fitotecnia, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre.

- FONTANA, D.C. and BERLATO, M.A., 1998, Modelo agrometeorológico-espectral para a estimativa do rendimento da soja no Estado do Rio Grande do Sul: um estudo preliminar. (Agrometeorological-spectral model to estimate the soybean yield in the state of Rio Grande do Sul: a preliminary study). In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, **10**, Santos. *Anais...* Santos: [s.n.] 1 CD-ROM.
- FONTANA, D.C., DUCATI, J.R., WEBER, E., FIGUEIREDO, D.C., BERLATO, M.A. and BERGAMASCHI, H., 1999, *Desenvolvimento e teste de metodologia em monitoramento e previsão de safras no Rio Grande do Sul* (Development and test of a methodology in crop monitoring and estimation in Rio Grande do Sul). Porto Alegre: CEPSSRM/UFRGS. Não paginado. (Série D: Relatório Técnico 003/99).
- FONTANA, D.C., BERLATO, M.A., LAUSCHNER, M.H. and MELO, R.W., 2001, Modelo de estimativa de rendimento de soja no Estado do Rio Grande do Sul. (Estimation model for soybean yield in the State of Rio Grande do Sul, Brazil) *Pesquisa Agropecuária Brasileira*, Brasília, **36**, pp. 399–403.
- GAMON, J.A., FIELD, C.B., GOULDEN, M.L., GRIFIN, K.L., HARTLEY, A.E., JOEL, G., PENUELAS, P. and VALENTINI, R., Relationships between NDVI, canopy structure and photosynthesis in three californian vegetation types. *Ecological Applications*, **5**, pp. 28–41.
- JENSEN, M.E., 1968, Water consumptions by agricultural plants. In: *Water deficits and plant growth*, T.T. Kozlowsky (Ed.), (New York: Academic Press), **2**, pp. 1–22.
- JUSTICE, C.O., TOWNSHED, R.G., HOLBEN, B.N. and TUCKER, C.J., 1985, Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing*, **6**, pp. 999–1318.
- JUSTICE, C.O., TOWNSHED, J.R.G. and KALB, V.L., 1991, Representation of vegetation by continental data sets derived from NOAA-AVHRR data. *International Journal of Remote Sensing*, **12**, pp. 999–1021.
- KALUBARME, M.H., POTDAR, M.B., MANJUNATH, K.R., MAHEY, R.K. and SIDDHU, S.S., 2003, Growth profile based crop yield models: a case study of large area wheat yield modeling and its extendibility using atmospheric corrected NOAA AVHRR data. *International Journal of Remote Sensing*, **24**, pp. 2037–2054.
- KÖPPEN, W., 1948, *Climatología* (México, DF: Fondo de Cultura Económica), p. 71.
- LIANG, E.Y., SHAO, X.M. and HE, J.C., 2005, Relationships between tree growth and NDVI of grassland in the semiarid grassland of north China. *International Journal of Remote Sensing*, **26**, pp. 2901–2908.
- LIU, W.K. and KOGAN, F., 2002, Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices. *International Journal of Remote Sensing*, **23**, pp. 1161–1179.
- MATZENAUER, R., BERGAMASCHI, H., BERLATO, M.A. and RIBOLDI, J., 1995, Modelos agrometeorológicos para estimativa do rendimento de milho em função da disponibilidade hídrica no Estado do Rio Grande do Sul. (Agrometeorological models to estimate the corn yield in function of the water availability in the state of Rio Grande do Sul) *Pesquisa Agropecuária Gaúcha*, Porto Alegre, **1**, pp. 225–241.
- MATZENAUER, R., BERGAMASCHI, H., BERLATO, M.A., MALUF, J.R.T., BARNI, N.A., BUENO, A.C., DIDONÉ, I.A., ANJOS, C.A., MACHADO, F.A. and SAMPAIO, M.R., 2002, *Consumo de água e disponibilidade hídrica para milho e soja no Rio Grande do Sul* (Water consumption and water availability for corn and soybean in Rio Grande do Sul) (Porto Alegre: FEPAGRO), p. 105 (Boletim FEPAGRO,10).
- MOTTA, J.L.G., FONTANA, D.C. and WEBER, E., 2003, Evolução temporal do NDVI/NOAA em áreas cobertas por pixels com proporções variáveis de soja. (Time evolution of NDVI/NOAA in areas covered by pixels with variable proportions of soybean). *Revista Brasileira de Agrometeorologia*, Santa Maria, **11**, pp. 356–360.
- PENMAN, H.L., 1956, Evaporation: an introductory survey. *Netherland Journal of Agricultural Science*, **4**, pp. 9–29.

- REYNOLDS, C.A., REYNOLDS, C.A., YITAYEW, M., SLACK, D.C., HUTCHINSON, C.F., HUETE, A. and PETERSON, M.S., 2000, Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data. *International Journal of Remote Sensing*, **21**, pp. 3487–3508.
- RUDORFF, B.F.T. and BATISTA, G.T., 1990, Yield estimation of sugarcane based on agrometeorological-spectral models. *Remote Sensing of Environment*, **33**, pp. 183–192.
- SHANTZ, H.J. and PIEMEISEL, L.N., 1927, The water requirement of plants at Akron, Colorado. *Journal of Agriculture Research*, **34**, pp. 1093–1190.
- THORNTHWAITE, C.W. and MATHER, J.R., 1955, The water budget and its use in irrigation. In. *THE YEARBOOK OF AGRICULTURE: Water*. (Washington, D.C.: Department of Agriculture), pp. 346–358.
- ZHONG-HU, H. and RAJARAM, S., 1993, Differential responses of bread wheat characters to high temperature. *Euphytica*, **72**, pp. 197–203.