J. Ind. Soc. Agril. Statist.
58(3), 2004 : 279-299

Statistical Validation of Yield Prediction Model Based on Satellite Data

F. Abdellani and J.F. Chamayou
Universite Paul Sabatier, 31062 Toulouse Cedex 4 – France

(Received : June, 2002)

SUMMARY

The aim of this paper is to present and describe the statistical validation of a yield prediction model based on remote sensing imagery. The main objectives of this validation task are

- the development of a model based on available data
- the definition of a methodology adapted to the available data
- an evaluation of the intermediate results corresponding to different components of the analysis and
- assessment of the overall accuracy of final yield estimates.

The validation task is essentially concentrated on the semi-empirical model applied to the estimation of corn yield and more precisely on the two main variables of this model : the computation of NDVI (Normalized Difference Vegetation Index) from remote sensing imagery and evaluation of the total dry matter of the crop.

Key words : Random homographical transformation, Cauchy attraction domain, VOIGT functions, MONTEITH model, Vegetation Index (NDVI), Brightness Index.

1. Introduction

1.1 General Frame of the Project

Agriculture is a major user of data from satellite remote sensing. But until now, most applications are restricted to a more or less descriptive analysis and remotely-sensed data for assessing the yield component utilisation is based on a more qualitative than quantitative use of satellite derived information. Quantitative analysis requires more complex methodologies, including the coupling of satellite data and crop production models to estimate yields.

Monitoring crop growth (and calibrating crop models) requires information at an adequate spatial resolution and adequate time frequency. Unfortunately high ground resolution and high time revisit do not exist with current satellite data. Combining high resolution data (with low time frequency) and high time

frequency data (with low ground resolution) from different satellites has some drawbacks, such as lack of radiometric compatibility and geometric accuracy. SPOT4, operating on the same platform a high time frequency sensor (VEGETATION) and a high ground resolution one (HRVIR : Haute Resolution Visible et Infra Rouge) with common bandwidths and common geometry, is expected to overcome these drawbacks. Under the VEGETATION Preparatory Program, the International User Committee (IUC) entrusted a consortium leaded by SCOT (Services et Conception de systèmes en Observation de la Terre) to investigate the combined use of VEGETATION and HRVIR to improve plant yield estimates (Husson (1998)).

The main goal of this project is the improvement of the methodology for integrating remote sensing data into crop models and VEGETATION data. The aim of this project is to implement crop forecasting systems based on agro-meteorological models for operational use. Its long term objective is to provide not only public crop analysis but also to give the private agri-business sector accurate information that could help better characterize the current crops of interest and detect changes that are usually only quantifiable at the end of the season.

1.2 Objectives of the Validation Task

We begin with a general description of the subject and then focus on the approximate models, as they will provide an analysis at a reduced cost.

These approximate models will be justified using field data extracted from SPOT-XS images corresponding to the areas of Orthez and Chartres, France and acquired in the frame of activity B of the MARS (Monitoring Agriculture with Remote Sensing) project. A statistical approach has been undertaken to account for the variability of the image data and of the ground experimental measurements.

This variability will be quantified using models with probability densities, distribution functions and confidence intervals of random variables representing the physical data. These models will use the available field observations.

2. Vegetation Index

2.1 Statistical Model

Varied agro-meteorological models using remote sensing information for crop yield prediction exist. These models can be split into two main categories : deterministic models and empirical ones. However, whatever the type of the selected model, the integration of remote sensing imagery often involves the use of the Normalized Difference Vegetation Index (NDVI). This value derived from the combination of Red and near Infra Red reflectances is widely used for the analysis of vegetation and must be considered as a major source of information to monitor crop development and to identify stress.

We define NDVI in remote sensing as

$$NDVI = \frac{\rho_{IR} - \rho_R}{\rho_{IR} + \rho_R} \tag{1}$$

where ρ_{IR} is the near Infra Red Reflectance and ρ_R the Red Reflectance.

The use of the two spectral bands ρ_R and ρ_{IR} on a multispectral satellite image allows the computation of a pseudo spectral band where the pixels have been transformed according to (1).

It is assumed that the observations (pixels in the Red and Infra Red spectral bands) are the result of a random trial. Let's suppose they are distributed according to a bivariate normal distribution $(\rho_R, \rho_{IR}) \sim N((\mu_1, \mu_2), \Sigma)$ with $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 r \\ \sigma_1\sigma_2 r & \sigma_2^2 \end{bmatrix}$. Further assume, each normal component (the Red and the Infra Red reflectance) had parameters respectively : mean (μ_1, μ_2), standard deviation (σ_1, σ_2), correlation coefficient (r), which characterize the vegetation class to which the pixels belong.

Relation (1) will be written under the form of a random homographical transformation

$$Z = \frac{1 - T}{1 + T}, \text{ where } T = \frac{\rho_R}{\rho_{IR}} \tag{2}$$

The problem now is to determine the distribution of Z from the distributions of ρ_R and ρ_{IR}. As the ratio of normal variates, Z distribution belongs to the attraction domain of the Cauchy distribution. This fact has the drawback of using a distribution without moments. The simplest case where ρ_R and ρ_{IR} are *mean zero* correlated normals with normal joint distribution gives a Cauchy distribution for Z. The more delicate case where ρ_R and ρ_{IR} are independent *not centered* normals (which is the application case) has been solved by Geary (1930) using an asymptotic result. When $\mu_1 >> \sigma_1, (\mu_2 >> \sigma_2)$, where $\mu_1, \mu_2, \sigma_1, \sigma_2$ are respectively, the means and variances of ρ_R and ρ_{IR}. We find that

$$X = \frac{\mu_2 - \mu_1 T}{\sqrt{\sigma_2^2 + \sigma_1^2 T^2}} \tag{3}$$

is distributed as a normal N (0, 1).

The exact probability density for T is given by Agard (1961) for $\mu_1 = \mu_2$. The generalization to $\mu_1 \neq \mu_2$ results in the probability density f (t) for T as

$$f(t) = Q(t)\left[1 + \sqrt{\pi}k(t)e^{k^2(t)} + \text{erf}(k(t))\right], t \in IR \tag{5}$$

where

$$Q(t) = \exp\left\{-\frac{\mu_1^2 + \mu_2^2 - 2r\mu_1\mu_2}{2(1-r^2)}\right\}\frac{\sqrt{1-r^2}}{\pi(1+t^2-2rt)} \tag{4}$$

$$k(t) = \frac{[(\mu_1 - r\mu_2) + t(\mu_2 - r\mu_1)]}{\sqrt{2(1-r^2)(1+t^2-2rt)}} \tag{6}$$

erf is the error function (see Spanier (1987)) and r is the correlation coefficient between ρ_R and ρ_{IR}.

We assume $\sigma_1 = \sigma_2 = 1$ which is not a restriction, since we just have to transform T by multiplying it with the ratio of standard deviations (σ_2/σ_1).

Using the complementary error function erfc, we can write the density as

$$f(t) = Q(t)\left\{[1 - \sqrt{\pi}k(t)e^{k^2(t)}\text{erfc}(k(t))] + \sqrt{\pi}\, k(t)\, e^{k^2(t)}\right\} \tag{7}$$

Where the term in brackets is

$$\frac{\sqrt{\pi}}{2}\left(\frac{\partial L(x,y)}{\partial x}\right)_{x=0} = 1 - \sqrt{\pi}ye^{y^2}\text{erfc}(y) \tag{8}$$

with L(x, y) the second kind Voigt function (see Haubold (1979)) , i.e.

$$L(x,y) = \frac{1}{\sqrt{\pi}}\int_0^\infty \exp\left(-yr - \frac{r^2}{2}\right)\sin(xr)dr \tag{9}$$

The asymptotic behaviour of L(x, y) is given by

$$L(x,y) = \frac{1}{\sqrt{\pi}}\left(\frac{x}{y^2} - \left(\frac{3}{2} + x^2\right)\frac{x}{y^4} + 0\left(\frac{1}{y^6}\right)\right) \tag{10}$$

After differentiation, we neglect the terms $0\left(\frac{1}{k^2(t)}\right)$ and we get for the approximate density $\tilde{f}(t)$

$$\tilde{f}(t) = Q(t)\sqrt{\pi}k(t)e^{k^2(t)} \tag{11}$$

This is the density of the random variable

$$M = \frac{\mu_2 - \mu_1 T}{\sqrt{1 + T^2 - 2rT}} \tag{12}$$

which is a normal N (0, 1).

This generalization of the result given by Geary to correlated variables allows us to deduce the Z confidence interval from the confidence interval of a reduced normal variate N at level α. Simply compute the roots (Z_α^- and Z_α^+) of the equation

$$n_\alpha = \frac{\mu_2 - \mu_1 \left(\frac{1 - Z_\alpha}{1 + Z_\alpha}\right)}{\sqrt{1 + \left(\frac{1 - Z_\alpha}{1 + Z_\alpha}\right)^2 - 2r\left(\frac{1 - Z_\alpha}{1 + Z_\alpha}\right)}} \tag{13}$$

where n_α is such that $P_r(N > n_\alpha) = \alpha$

It is necessary to justify the approximate confidence interval, by comparison of exact and approximate densities and distribution functions. Let us call this approximation the "Second order approximation". We distinguish it from a crude first order approximation which is based on the hypothesis : The pixels extracted from the image belonging to a given class have a NDVI distributed according to a normal distribution (see Agard (1961)) with mean $\frac{\mu_1 - \mu_2}{\mu_1 + \mu_2}$ and standard deviation $\frac{\sqrt{2(1-r)}}{\mu_1 + \mu_2}$.

A better approximation is obtained in the following manner.

Let $n_\alpha = 1$ in (13) and take the standard deviation of Z as $\xi = \frac{1}{2}(Z_\alpha^+ - Z_\alpha^-)$ i.e.

$$\xi = \frac{\{(\mu_2^2 - \mu_1^2)^2 - [(\mu_1 + \mu_2)^2 - 2(1 + r)][(\mu_2 - \mu_1)^2 - 2(1 - r)]\}^{1/2}}{[(\mu_1 + \mu_2)^2 - 2(1 + r)]} \tag{14}$$

from which the standard deviation $\left(\frac{\sqrt{2(1-r)}}{\mu_1 + \mu_2}\right)$ of the NDVI can be deduced leting μ_1 and $\mu_2 \to \infty$.

Let this standard deviatin be $\tilde{\xi}$. Then we have

$$N\left(\frac{\mu_1 - \mu_2}{\mu_1 + \mu_2}, \tilde{\xi}\right)$$

which must also be checked against the exact result.

2.2 Model Application

From the X S2 and X S3 spectral bands of SPOT multispectral HRV, on a lot of corn in the region of Orthez, the NDVI pseudo spectral band has been computed, applying a homographical transformation to all pixels.

Fig. 1 and 2 display the histograms in this parcel for X S2 and X S3 respectively as observed on 15/05/96.

In Fig. 3, one can see the NDVI pseudo spectral band and the parcel studied on 15/05/96 consisting of 170 pixels.

Fig. 4 shows the NDVI histograms of the parcel at each date of the corn growth for which we had a SPOT image of the region (7 dates) (the dots represent the exact distribution and the squares : the corresponding first order approximation), except on the three last dates for which the discrepancy is explained by a partial harvest of the parcels.

In Fig. 5, the cumulative distribution of the NDVI exact and first order approximation are plotted in the same graph. The second order approximation agrees with the experimental data using the Kolmogorov-Smirnov test and the Cramer-Von Mises test (see tab.7 and tab.7 for the results).

The image processing software used for these calculations is Multiscope (Fleximage) and the statistical software is S-PLUS. Specific routines in C have been developed to compute the model and its two approximations (this extra software is available on request).

Fig. 7 shows the boxplot of the NDVI as a function of time.

3. Brightness Index

3.1 Statistical Model

The brightness index is given by the following formula

$$B = \sqrt{B_1^2 + B_2^2} \tag{15}$$

where B_1 and B_2 are the pixels radiometries in the two spectral bands. We assume that the distribution of the pixels radiometries is normal. The brightness index follows a Raleigh-Rice distribution, if B_1 and B_2 have the same variance and are independent, i.e. the density of B is given by

$$f(b) = \frac{b}{\sigma^2} \exp\left(-\frac{b^2 + m^2}{2\sigma^2} \right) I_0\left(\frac{bm}{\sigma^2} \right), b \in IR^* \tag{16}$$

where $m = \sqrt{\mu_1^2 + \mu_2^2}$, μ_1 and μ_2 are respectively the means of B_1 and B_2 and I_0 is a modified Bessel function of order zero. The distribution is close to normality when $m >> \sigma$. If B_1 and B_2 are correlated with ρ as correlation coefficient, then

$$f(b) = \frac{b}{\sigma^2\sqrt{1-\rho^2}} \exp\left(-\frac{b^2 + m^2 - 2\rho\mu_1\mu_2}{2\sigma^2(1-\rho^2)} \right) I(p, q, s; b), b \in IR^* \tag{17}$$

where the integral

$$I(p, q, s; b) = \frac{1}{2\pi} \int_0^{2\pi} \exp\left[b(p \cos\theta + q \sin\theta + \frac{s}{2} \sin(2\theta)) \right] d\theta \tag{18}$$

with

$$p = \frac{\mu_1 - \rho\mu_2}{\sigma^2(1-\rho^2)}, q = \frac{\mu_2 - \rho\mu_1}{\sigma^2(1-\rho^2)} \text{ and } s = \frac{b\rho}{\sigma^2(1-\rho^2)}$$

This integral can be represented as an infinite sum of products of modified Bessel functions (see Walsiljeff (1969)).

The asymptotic method of Laplace (see De Bruijn (1958)) can be applied to avoid numerical integration, with an error less than 10% on f(b). Let

$$I(b) \simeq \frac{\exp(-bh(\theta_0))}{\sqrt{2\pi b h''(\theta_0)}} \tag{19}$$

where

$$h(\theta) = -(p \cos\theta + q \sin\theta + \frac{s}{2} \sin(2\theta)) \tag{20}$$

Then the extremum is obtained by sloving

$$h'(\theta_0) = p \sin\theta_0 - q \cos\theta_0 - s \cos(2\theta_0) = 0 \tag{21}$$

The Newton-Raphson method gives

$$I(p, q, s; b) = \frac{1}{2\pi} \frac{\exp\left[b(p \cos\theta_0 + q \sin\theta_0 + \frac{s}{2} \sin(2\theta_0)) \right]}{\sqrt{b(p \cos\theta_0 + q \sin\theta_0 + 2s \sin(2\theta_0))}} \tag{22}$$

3.2 Model Application

Fig. 6 shows the comparison between the theoretical probability density of the brightness index and the computed histogram (15/05/96 Red, Infra Red

spectral bands). The χ^2 test is not significant at 5% level for 19 degrees of freedom.

4. Total Dry Matter

4.1 The Statistical Model

For several years, the idea that has been retained for crop productivity estimation is to combine a Dry Matter production model from intercepted radiation and radiation interaction with the plant canopy model in which remote sensing data are introduced. The present research trend is oriented towards model determination, because direct relationships between precise radiometric data and Dry Matter are empirical and therefore difficult to write a function of time and space. The model of Dry Matter calculation considered in this study is based on the Monteith model (1972), simplified by Varlet-Granchet (1982) and expressed as

$$M = \int_{t_0}^{t_n} \varepsilon_c \varepsilon_b \varepsilon_i(t) Rg(t) dt \tag{23}$$

Where M is the Total Dry Matter on stalks, Rg(t) is the global incident radiation (meteorological data), ε_c is the climate efficiency (ratio of the radiation energy available for the photosynthesis to the total radiation energy) which is a constant independent of time, ε_b is the conversion efficiency (ratio of the chemical energy accumulated in the Dry Matter to the radiation energy used for the photosynthesis) which is a constant slowly varying with time, and $\varepsilon_i(t)$ is the interception efficiency (ratio of the radiation energy used for the photosynthesis to the radiation energy available for the photosynthesis). Then

$$\varepsilon_i(t) = a\,(NDVI(t) - NDVI_{ground}) = a\,\widetilde{NDVI}(t)$$

Where a is a scale factor (invariant). $NDVI_{ground}$ is a deterministic calibration parameter corresponding to the NDVI for a bare soil (independent of time). NDVI (t) is a time varying known by multi-spectral satellite images sampled with the satellite observation frequency on the area, but with data losses due to the clouds cover (more than one week between samples). The distribution of NDVI is Gaussian, if the model of Section 3 is true. Let t_0 be the begining of the plant growth and write $t_0 = \max_t |\, NDVI(t) - NDVI_{ground} \,| \neq 0$; let t_n be the plant degeneration.

Rg is densely sampled and uncorrelated due to the weekly measurement intervals of the NDVI. Rg is considered to be deterministic in first approximation and uncorrelated from the NDVI.

Thus the total Dry Matter M is given by a stochastic integral (23) but only known under the form of a discrete sum from the weekly data collected

$$M = a\varepsilon_c\varepsilon_b \sum_{j=1}^{n} \widetilde{NDVI}(j)\, Rg(j) \tag{24}$$

This is a sum of correlated Gaussian random variates weighted by Rg(j). Thus M is also Gaussian with mathematical expectation

$$E(M) = \bar{a}\,\bar{\varepsilon}_c\bar{\varepsilon}_b \sum_{j=1}^{n} R_g(j) E(\widetilde{NDVI}(j)) \tag{25}$$

The second order moment is

$$E(M^2) = \bar{a}^2\bar{\varepsilon}_c^2\bar{\varepsilon}_b^2 \sum_{j=1}^{n}\sum_{k=1}^{n} Rg(j)\, Rg(k) E[\tilde{N}(j)\tilde{N}(k)] \tag{26}$$

Formula (25) is obtained from means, variances and correlation coefficients of the Red and Near Infra Red images at the dates j = 1,... , n using the Gaussian model.

Formula (26) needs the computation of all the

$$E(\widetilde{NDVI}(j).\widetilde{NDVI}(k)) \quad \text{for} \quad j = 1, \ldots, n \text{ and } k = 1, \ldots, n \tag{27}$$

The autocorrelation will decrease as $|j - k|$ increases.

The variances for the case j = k are known from the Gaussian model. It seems difficult to get an analytic form of (27) using means, standard deviations and the correlation of the Red and Infra Red images at the dates j and k since this mathematical expectation is a quadruple integral in X_j, Y_j, X_k, Y_k where these variables are correlated (correlation between dates and spectral bands). However a description of the correlation against time can be provided for the Red pairs (k, j), the Infra Red pairs (k, j), the Red (k, j)/(Infra Red) (k, j) and $\widetilde{NDVI}(k)/\widetilde{NDVI}(j)$.

4.2 Model Application

4.2.1 Correlograms

(a) Autocorrelograms

The correlograms for the red, the Infra Red and the NDVI spectral bands are plotted in Fig. 8, 9 and 10 where the model is a first order autoregressive process. The figures specify the correlation values and the dates from which the interval of dates is taken into account. The parameter ϕ has been fitted from the

observed results, i.e.

$$\rho_{ij} = \phi^{|\text{date (i)} - \text{date (j)}|} \tag{28}$$

where $|\text{date (i)} - \text{date(j)}|$ represent the differences of the m dates and

$$\phi_{ij} = \exp\left(\frac{1}{|\text{date (i)} - \text{date(j)}|} \ln|\rho_{ij}|\right) \qquad i, j = 1, \ldots, m \tag{29}$$

The sample is censored for negative correlation and

$$\phi_m = \frac{1}{\nu}\sum_i\sum_{j<i}\phi_{ij} \tag{30}$$

where ν is the number of dates for which the positivity condition is fulfilled. A confidence interval for the estimation of the correlations can be deduced from the classical Fisher theory (see Kendall and Stuart (1966) (p. 390 Chap. 16.33)).

Z_{ij} is computed from

$$Z_{ij} = \frac{1}{2}\ln\left(\frac{1+\phi_m^{|\text{date (i)} - \text{date(j)}|}}{1-\phi_m^{|\text{date(i)} - \text{date (j)}|}}\right) \tag{31}$$

but corrections are necessary for small sample sizes.

The two envelopes (surrounding the mean) are deduced from

$$b_{ij}^{\pm} = \tanh\left(Z_{ij} \pm t_\alpha \frac{1}{\sqrt{n-3}}\right) \tag{32}$$

where t_α is the value of the upper quantile for probability α.

(b) Intercorrelograms

The intercorrelogram between Red and Infra Red spectral bands is plotted in Fig. 11 together with the following formulation

$$\rho_{ij} = \phi^{|\text{date(i)} - \text{date(j)}|}\rho_0 \tag{33}$$

where

$$\rho_0 = \frac{1}{\nu}\sum_{i=1}^{\nu}\rho_i \tag{34}$$

The sample is censored for negative correlations and the corresponding standard deviation is computed to derive the ρ_0 confidence interval.

The estimation of ϕ, ϕ_m, is computed from ρ_{ij}/ρ_0 in the same way as before with censorship if ρ_{ij}/ρ_0 exceeds 1.

4.2.2 The Results

Applying the random Monteith model to the 8 test parcels for which ground measurements were performed, we get the results shown in Table 3. However, since the parcels have different areas, their contributions then have different weights and it is necessary to take this fact into account to extract the final result.

4.2.3 Harmonic Processing

The results are fractions (tons/hectare) of yield and an estimate of the expectation is calculated from the total dry matter produced by the whole set of parcels and their total area (which is given by the number of pixels by parcel). Let

$$\overline{X}_h = E(TDM) = \frac{n}{\sum_i \frac{n_i}{X_i}} = h(X_1, X_2, \ldots X_n) \tag{38}$$

where n is the total number of pixels, n_i is the number of pixels in parcel i and X_i is the total dry matter in parcel i.

The results processed harmonically of each parcel using the Monteith model for the satellite images are summarized in Tables 4, 5 and 6.

5. Conclusion

A statistical model with economical computation time has been provided for the Vegetation Index and also for the Total Dry Matter calculation. This model allows the calculation of the confidence intervals for these two quantities of agronomical interest and then the calculation of the confidence intervals for the yield prediction using the Monteith model on several test parcels harmonically averaged.

6. Appendix : Confidence Interval for the Harmonic Processing

According to the general theory of estimation (Kendall and Stuart (1966))

$$E[(\overline{X}_h - \theta)^2] = \left(\frac{Dh}{D\underline{\theta}}\right)_\theta \text{Var}[X] \left(\frac{Dh}{D\underline{\theta}}\right)_\theta^T$$

where the gradient in the harmonic case is

$$\left(\frac{Dh}{D\underline{\theta}}\right)_{\hat{\theta}} = \frac{n}{\left(\frac{n_1}{x_1} + \ldots \frac{n_k}{x_k}\right)^2} \left[\frac{n_1}{x_1^2}, \ldots \frac{n_k}{x_k^2}\right]^T$$

then the confidence interval is

$$\overline{X}_h \pm t_\alpha \sqrt{E[(\hat{\overline{X}}_h - \theta)^2]}$$

REFERENCES

Agard, J. (1961). Melange de deux populations normales et étude de quelques fonctions de variables normales. *Revue de Statistique Appliquee*, **9 4,** 53-70.

De Bruijin, N.G. (1958). *Asymptotic Methods in Analysis.* North Holland, Amsterdam.

Geary, R.C. (1930). The frequency distribution of the quotient of two normal variates. *Ann. Statist.*, 442-446.

Haubold, H.J., and John, B.W. (1979). New results concerning the analysis of VOIGT functions. *Astro Space Science*, **65,** 477-491.

Husson, A. (1998). Integration of VEGETATION and HRVIR data into yield estimation. Vegetation Preparatory Program, CNES-Toulouse.

Kendall, M. and Stuart, A. (1996). *Advanced Theory of Statistics* Oxford University Press.

Monteith, J.L. (1972). Solar radiation and productivity in tropical ecosystems. *J. Appl. Ecolo.*, **9,** 747-766.

Spanier, J. and Oldham, K.B. (1987). *An Atlas of Functions.* Hemisphere Publication, New York.

Varlet-Granchet (1982). Efficience de la conversion de l'énergie solaire par un couvert vegetal. *Acta Oecologica, Oecol. Plant.*, **17,** 3-26.

Vignolles, C. (1995). Modelisation de la production agricole a moyenne echelle : apport du couplage des mesures par satellite a hautes resolutions spatiales et a hautes frequences temporelles. Universite Paul Sabatier-Toulouse.

Walsiljeff, A. (1969). Verallgemeinerung einer Summenformel Von Neumann und Lommel. *Zeitschrift Angewandte Mathematik and Physik,* **20,** 389-402.

Table 1. Kolmogorov-Smirnov test

Shot date	Second order approximation	amp; Corresponding probability	First order approximation	Corresponding probability
28/3/96	0.05469	31%	0.06869	60%
15/5/96	0.10023	93%	0.09686	92%
30/5/96	0.08331	81%	0.08203	80%
15/6/96	0.0866	84%	0.09957	93%
17/7/96	0.15534	R	0.22969	R
15/8/96	0.0808	78%	0.23226	R
7/9/96	0.11208	97%	0.2017	R

Table 2. Cramer-von Mises test: values and probabilities for the approximation to NDVI histograms

Shot date	Second order approximation	Corresponding probability	First order approximation	Corresponding probability
28/3/96	0.09742	35%	0.16832	49%
15/5/96	0.2326	72%	0.20283	71%
30/5/96	0.30836	62%	0.23170	61%
15/6/96	0.18682	64%	0.29840	72%
17/7/96	0.89750	91%	3.11124	98%
15/8/96	0.33178	60%	3.97832	98%
7/9/96	0.26854	78%	2.53045	96%

Table 3. Correlation between reference and estimation : 0.75

Parcel	Ground reference	Estimation	Standard deviation	Hail	Irrigation	Confidence interval
1	26	23	4.4	no	no	[17.3, 28.6]
2	30.4	21.7	4.6	no	yes	[15.8, 27.8]
3	10.6	14.1	4.8	yes	no	[7.9, 19.8]
4	13.8	13	5.1	yes	yes	[6.8, 19.2]
5	17.8	11	5.8	no	no	[4.04, 17.9]
6	19.5	22.7	4.7	no	no	[17, 28.3]
7	15.6	14	5.1	yes	yes	[7.8, 20.1]
8	15.3	13.6	6.9	yes	yes	[5.3, 21.8]

Table 4. Corn results in Orthez : estimation for the 8 parcels with ground investigation

Harmonic mean : 17.2 tons/hectare
Harmonic standard deviation : 2.4 tons/hectare
95% Confidence interval : [12.6, 21.8]
90% Confidence interval : [13.2, 21.1]

Table 5. Corn results in Orthez : estimation of all parcels studied (39 parcels)

Harmonic mean : 15.7 tons/hectare
Harmonic standard deviation : 1.6 tons/hectare
95% Confidence interval : [12.6, 18.8]
90% Confidence interval : [13.1, 18.3]

Table 6. Wheat results in Chartres : estimation of all the parcels studied (41 parcels)

Harmonic mean : 21.15 tons/hectare
Harmonic standard deviation : 0.51 tons/hectare
95% confidence interval : [20.2 22.2]
90% Confidence interval : [20.3 22.0]

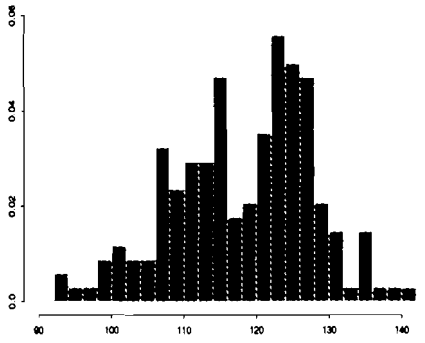

Fig. 1. Red reflectance normalized histogram

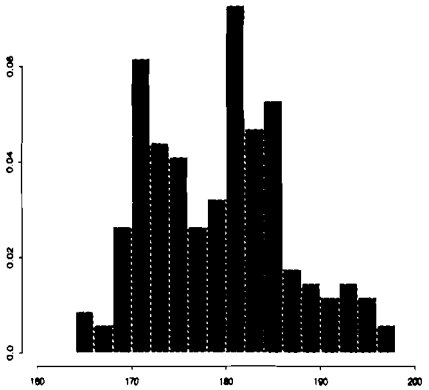

Fig. 2. Near Infra Red reflectance normalized histogram

Fig. 3. NDVI pseudo spectral band (450*550 pixels) and a parcel of 170 pixels (©SPOT Image Copyright 1996 CNES).

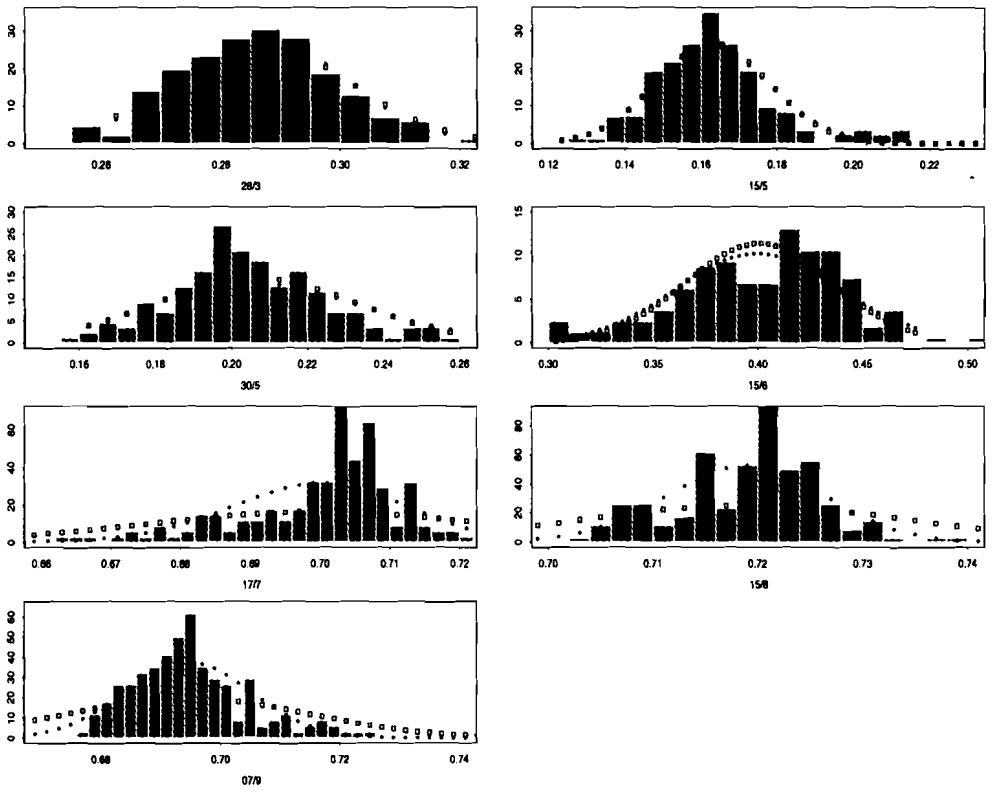

Fig. 4. NDVI histograms and approximations (Probability densities). From top and from left to right to bottom : 28/3, 15/5, 30/5, 15/6, 17/7, 15/8 and 7/9.

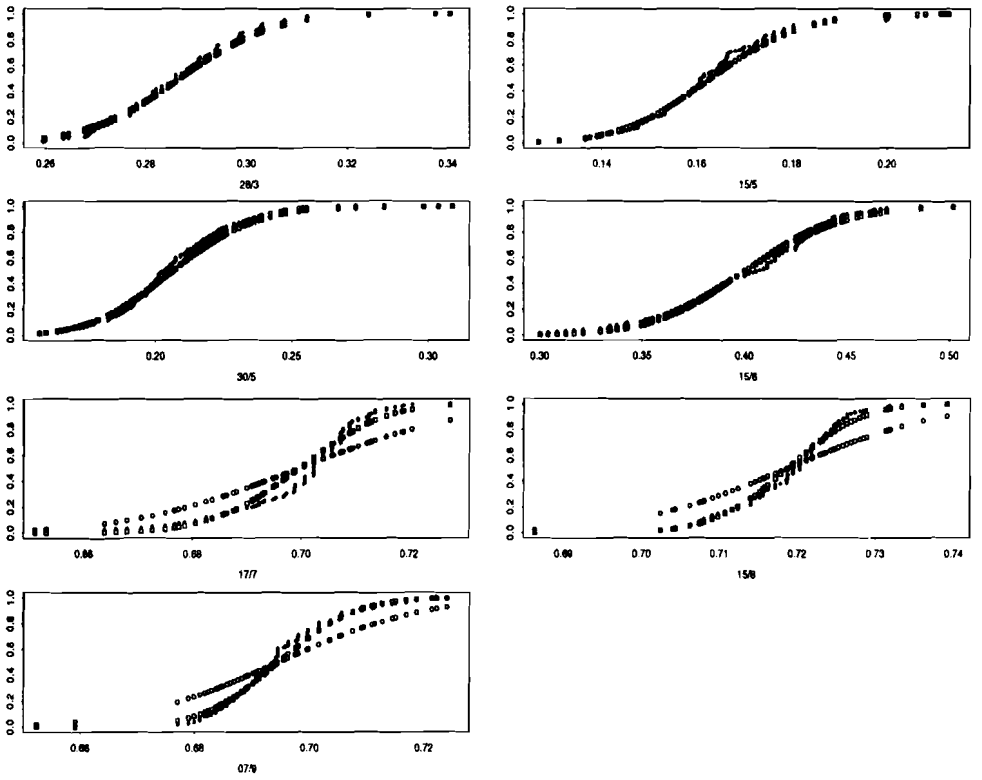

Fig. 5. NDVI Cumulative distribution and approximations (Repartition function). From top and from left to right to bottom : 28/3, 15/5, 30/5, 15/6, 17/7, 15/8, and 7/9.

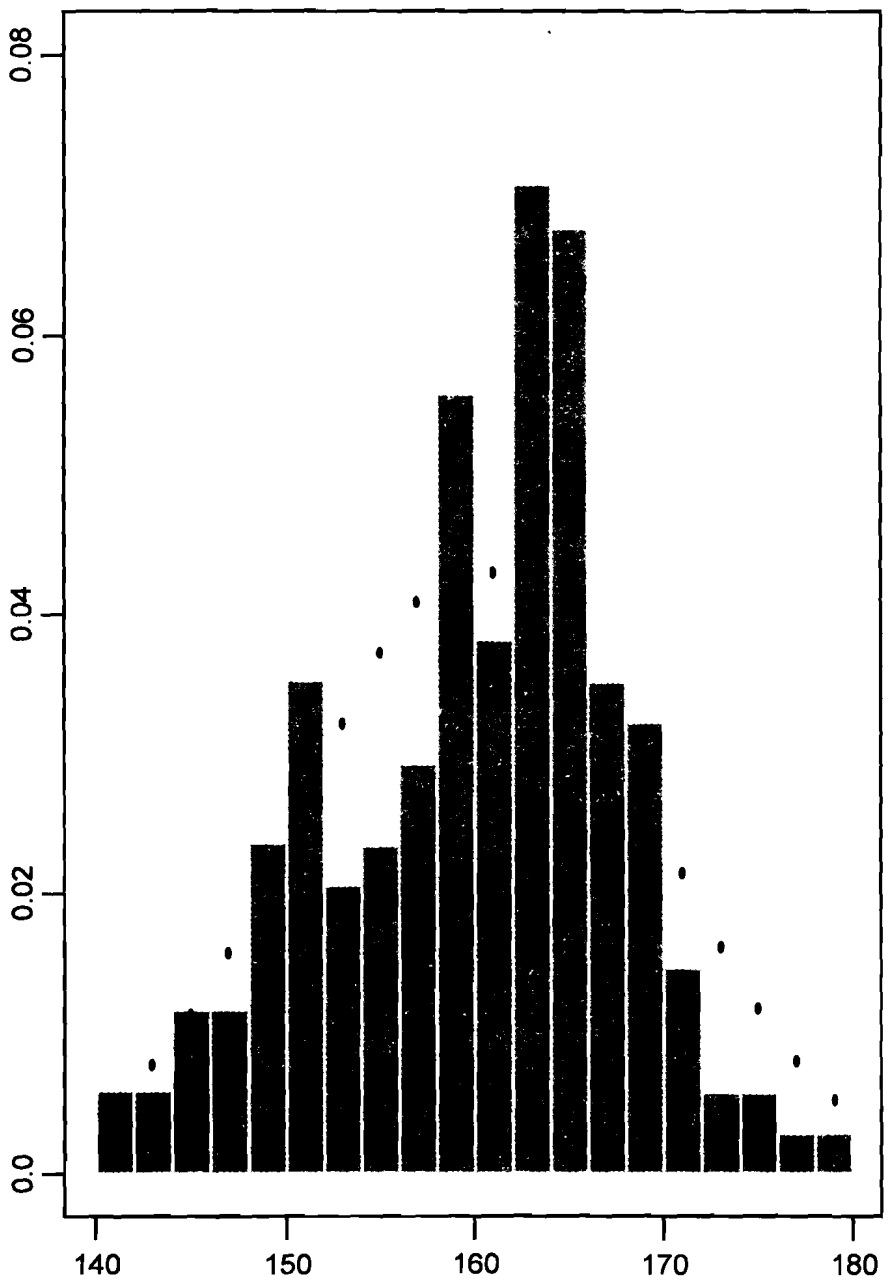

Fig. 6. The brightness index histogram and the corresponding theoretical probability (standard deviation = 0.7)

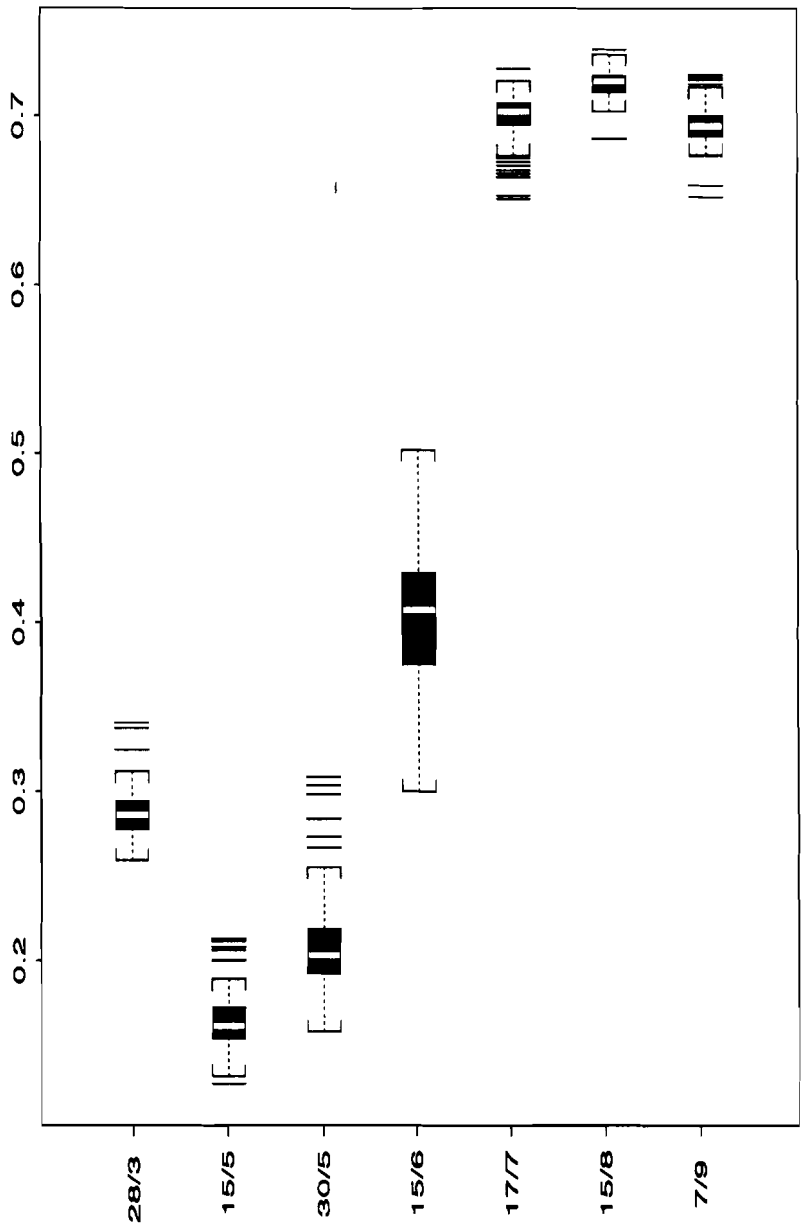

Fig. 7. NDVI dispersion characteristics versus date

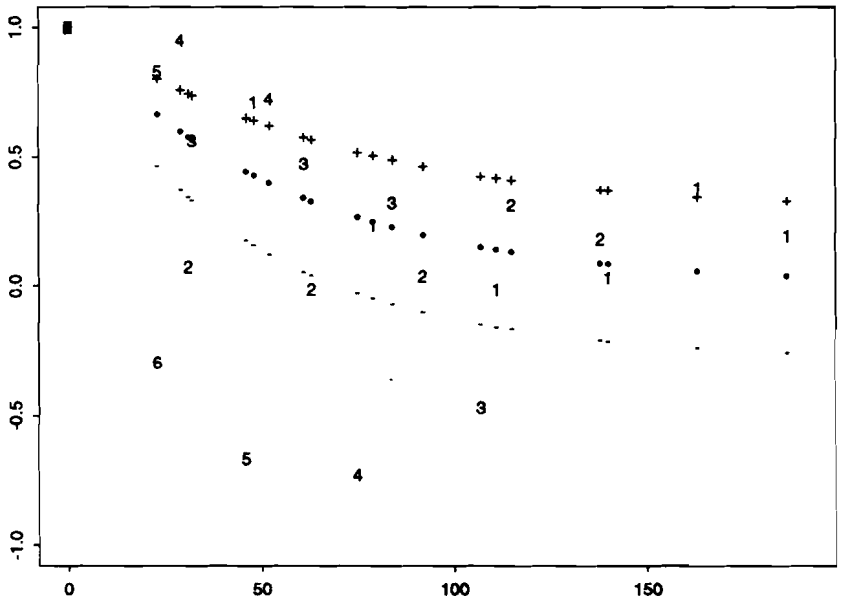

Fig. 8. Red correlogram versus date (in days)
1 : 28/3 2 : 15/5 3 : 30/5 4 : 15/6 5 : 17/7 6 : 15/8 7 : 7/9

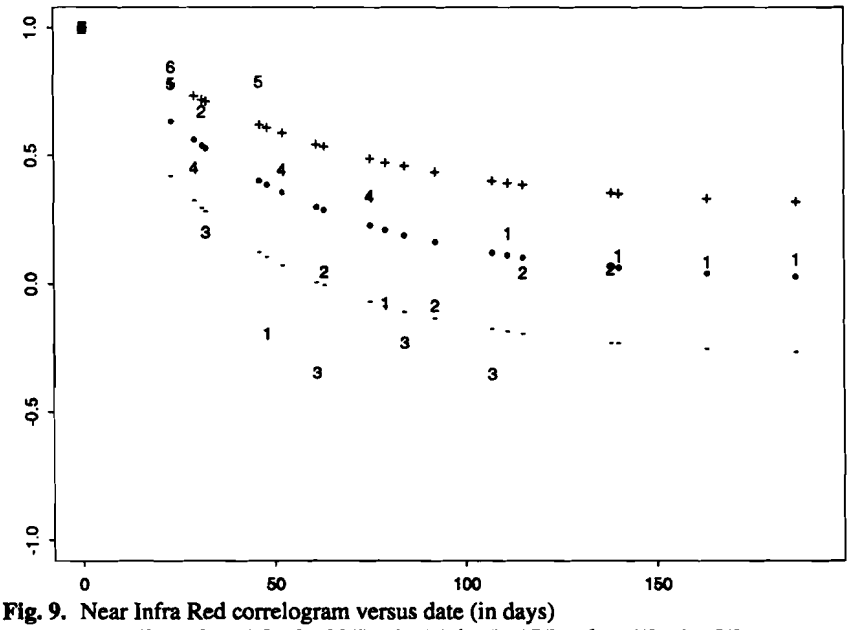

Fig. 9. Near Infra Red correlogram versus date (in days)
1 : 28/3 2 : 15/5 3 : 30/5 4 : 15/6 5 : 17/7 6 : 15/8 7 : 7/9

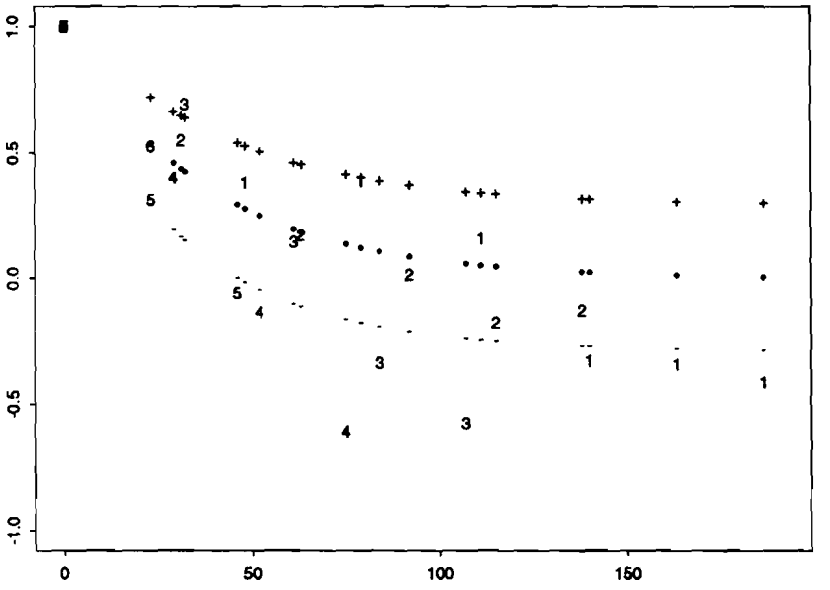

Fig. 10. NDVI correlogram versus date (in days)
1 : 28/3 2 : 15/5 3 : 30/5 4 : 15/6 5: 17/7 6 : 15/8 7 : 7/9

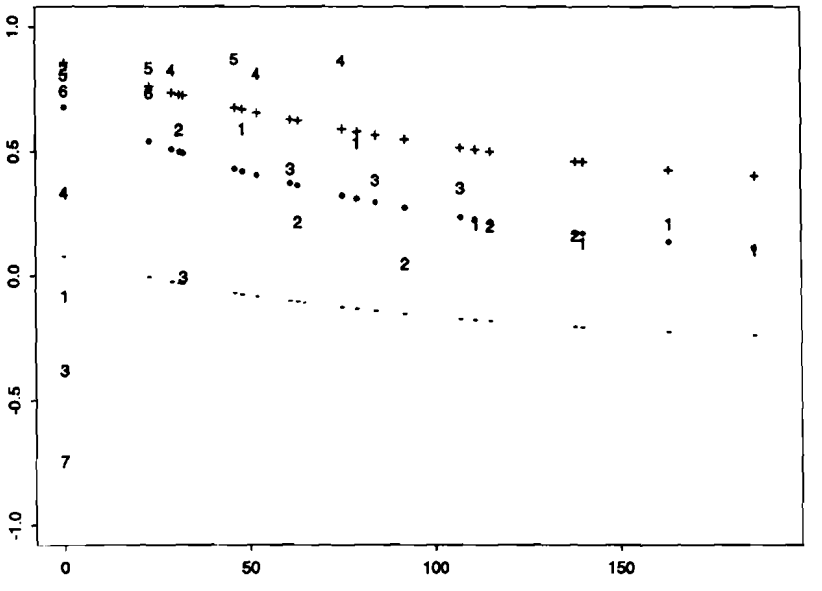

Fig. 11. Red-Near infra red correlogram versus date (in days)
1 : 28/3 2 : 15/5 3: 30/5 4: 15/6 5 : 17/7 6 : 15/8 7: 7/9