

Genetic Programming for Agricultural Purposes

Clément Chion

LIVIA
École de Technologie Supérieure
Montréal (Québec), Canada
clement@livia.etsmtl.ca

Luis E. Da Costa

LIVIA
École de Technologie Supérieure
Montréal (Québec), Canada
costa@livia.etsmtl.ca

Jacques-André Landry

LIVIA
École de Technologie Supérieure
Montréal (Québec), Canada
jacques.landry@etsmtl.ca

ABSTRACT

Nitrogen is one of the most important chemical intakes to ensure the healthy growth of agricultural crops. However, some environmental concerns emerge (soil and water pollution) when a farmer applies nitrogen in excess. In this study, we propose a new method called GP-SVI to search for the best descriptive model of nitrogen content in a cornfield (*Zea mays*), thanks to airborne hyperspectral data and ground truth nitrogen measurements. Coupling the output of this descriptive model with variable-rate technologies (VRT) would allow farmers to practice site-specific management ensuring them economical savings and ecological benefits. GP-SVI is a parallel search of the best spectral vegetation index (SVI) describing a crop biophysical variable, derived from Genetic Programming (GP). Compared to statistical regression methods on our datasets, GP-SVI improves results obtained with classical approaches, in term of explained-variance and generalization error. We also show that the spectral bands selected by GP-SVI match those selected by Partial Least Square regression optimized by Genetic Algorithms (GA-PLS) as proposed by Leardi in “Application of genetic algorithm-PLS for feature extraction in spectral data sets”, in *Journal of Chemometrics*.

Categories and Subject Descriptors

I.2.10 [Computing Methodologies]: Artificial Intelligence – vision and scene understanding – modeling and recovery of physical attributes.

General Terms

Algorithms, Management.

Keywords

Genetic Programming (GP), hyperspectral imagery, remote sensing, precision farming, site-specific management, spectral vegetation indices (SVI), crop nitrogen content.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO'06, July 8–12, 2006, Seattle, Washington, USA.
Copyright 2006 ACM 1-59593-186-4/06/0007...\$5.00.

1. INTRODUCTION AND BACKGROUND

New sensor technologies have been appearing in agriculture over the last few years, giving birth to what’s called Precision Farming or Site-Specific Management, which is the application area of the presented study. The democratization of hyperspectral airborne sensor utilization lead to the acquisition of large volumes of data [13]. In agriculture, this means of data acquisition offers two main advantages namely: it offers a non-destructive approach to data acquisition, and it allows rapid data acquisition over large geographic areas under investigation; after an appropriate analysis, coupling the extracted information with variable-rate technologies (VRT) makes it possible for a farmer to adapt his fertilizer application to the spatially varying crop needs. These two characteristics have given rise to great hopes in agriculture in countries like Canada where fields are currently stretched over several hundreds of hectares. Indeed, over such large areas, an infield sample collection is impossible to obtain while still providing a precise and exhaustive evaluation of the variability of any biophysical variable, such as, for example, crop nitrogen content or *leaf area index* (LAI). However, more investigations are still necessary to improve the reliability of information extracted from hyperspectral data in order to justify the elevated cost associated with this acquisition.

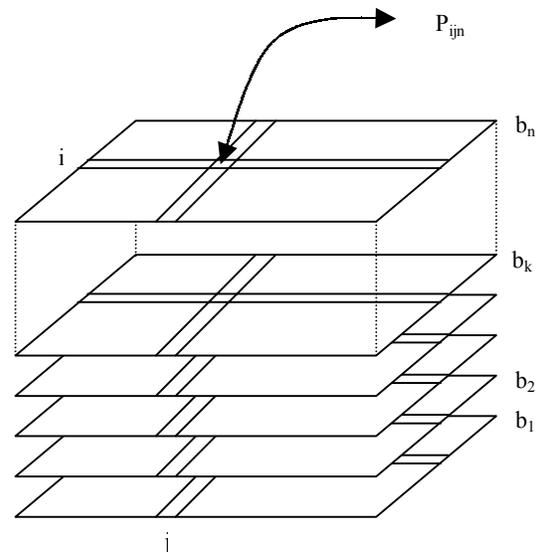


Figure 1. Hyperspectral data structure composed by n spectral bands (i.e. images) b_k

Hyperspectral sensors currently scatter the light energy reflected by a scene, called *reflectance*, over hundreds of electromagnetic spectral portions called spectral bands and noted b_i , each one represented by an image (see Figure 1); this provides the analyst with a large number of descriptors to describe biophysical variables. A sampling campaign in an agricultural field provides what is called *ground truth* data. These data are often expensive, which does limit their quantity. Coupling this *punctual* source of information with a “binding agent”, such as hyperspectral images, offers the opportunity to extend the punctual knowledge to the whole area under investigation. An important condition for success is to be able to pick up relevant sampling information during field campaigns. In fact, overall field variability (regarding the variable observed) must be represented by ground truth data and this can be accomplished thanks to farmer’s historical knowledge of his crop characteristics. Many approaches exist to deal with high dimensional hyperspectral datasets and derive information from them. Subsequently, we only mention three of them that appear to be the most widely spread in precision farming applications.

For each pixel in a scene, classification techniques [2, 4, 7, 17] consider hyperspectral data as an n -dimensional vector of characteristics P_{ij} , whose components are the greyscale values at location (i,j) , for each one of the n bands b_k . The aim is to attribute a label (supervised) or a cluster membership (unsupervised) to each pixel P_{ij} of the discretized scene (see Figure 1); consequently, these approaches fit with qualitative problems. In the case of datasets with very large dimensions (currently n is about several hundreds), the threat of *the curse of dimensionality*, first described by Bellman [1], forces the analyst to reduce the dimension of vector P_{ij} , using techniques of feature extraction [11, 20] or feature selection [12]. In the case of supervised classification, after a learning step (processed on a training subset of pixels, whose labels are known thanks to *a priori* knowledge), a program called a classifier is obtained and, after a validation step, it can be used to classify the rest of the scene. On the other hand, the case of unsupervised classification is used when no ground knowledge is available. In this case, a program produces a desired number of clusters of pixels, gathered together according to a criterion of resemblance (often statistical). Additional *a priori* information is then necessary to identify and aggregate similar clusters and to assign them a label.

To the same end, statistical methods are also used to create descriptive or predictive models of crop biophysical variables. As with the classification approach, these methods can be considered as “blind” in the sense that no knowledge on plant spectral properties is necessary to elaborate a model but only statistical characteristics of the data are taken into consideration. For example, many regression techniques are used to link quantitative crop variables to spectral data. Amongst them, one can mention multiple regressions (MR), partial least square regression (PLSR) [8, 19], stepwise regression [3] and some other derived ones. The goal of these methods is to find a weighted sum of specific wavelength reflectances that maximizes the explained variance of the variable samples under investigation. While MR methods consider the whole range of hyperspectral bands in the model, PLSR attempts to retain only a subset of the relevant ones.

Lastly, and probably the most widespread approach in precision farming, is the study of *spectral vegetation indices* (SVI). SVI are

defined by arithmetic operations between spectral bands b_i , extending from a simple bands ratio like the simple ratio index SR to more complex combinations like $MSAVI2$ [16]:

$$SR = \frac{NIR}{R}$$

and

$$MSAVI2 = (NIR + 1) - 0.5\sqrt{(2 \cdot NIR + 1)^2 - 8 \cdot (NIR - R)}$$

where NIR and R represent the scene reflectance in the near infrared and the red portion of the electromagnetic spectrum respectively.

The idea behind SVI is to promote the predominant effects of an observed variable by reducing the undesirable influence of others. The classical method in this process is to link spectral curve singularities (such as peaks or gaps of reflectance associated with the presence of a chemical components) to variables, thanks to laboratory experiments and theoretical knowledge [14, 15]; the next step is to compute an appropriate arithmetic combination of the selected bands to isolate the observed variable effects. Once the specific SVI is obtained, regressions can be performed between it and the biophysical variable in order to generate a descriptive model. Unfortunately, although this approach is often relevant at the leaf scale, in a laboratory context, it is not applicable at the crop scale because of many perturbing effects inherent to aerial acquisition of spectral data [9] (sensor resolution, sensor measure geometry, air humidity, influence of vegetation density, influence of correction steps, etc.). All these unstable parameters introduce some noise in hyperspectral data, which prevents the possibility of detecting the subtle pvariation in spectral variance caused by the biophysical variable under study. This partly explains why the development of universal SVI that would precisely quantify a specific biophysical variable in any configurations is still unrealistic.

Due to the problems with SVI mentioned above, it appears clearly that analysts yet can’t extract accurate descriptive models from spectral datasets without ground truth data. In this study, as an alternative to pure statistical techniques, we propose the GP-SVI method for the assessment of crop biophysical variables thanks to the combination of hyperspectral and ground truth data. Considering the non-repeatability of hyperspectral sensor, GP-SVI aims at searching for specific SVI, in a case-by-case approach. This means that for each particular configuration (weather, nature of the underlying soil, acquisition time of day...), we look for new SVI to describe crop characteristics. GP-SVI is an evolutionary approach for finding SVI, inspired by Genetic Programming (GP). In our study, the biophysical variable under investigation is the nitrogen content in a cornfield. In other words, we aim at finding a descriptive model for nitrogen variability throughout a cornfield using airborne- hyperspectral data and nitrogen infield measurements. We first present the data of the study. Secondly, the GP-SVI method is presented in detail and the associated choices are explained and justified. Finally, we show, discuss and compare the performances of GP-SVI for the generation of descriptive models of cornfield nitrogen content.

2. Materials and Methods

2.1 Data Description

Data used to test and validate the proposed method was obtained from an intensive field campaign (GEOIDE 2000 project). During the growing season of a test cornfield (*Zea mays*) stretching from sowing to harvest, several hyperspectral images and ground measurements were collected on a field located on the McGill University Research Farm at Macdonald Campus, Ste-Anne de Bellevue, Quebec, Canada. The experimental field (see Figure 2) was randomised as plots with different combinations of weed and nitrogen treatments, resulting in a strong spatial heterogeneity.

During the growing season, three flights were made with a Compact Airborne Spectrographic Imager (CASI sensor) to obtain hyperspectral images at key stages of plants' growth. Data acquired are composed of 72 spectral bands ranging from 409 to 947 nm with $\Delta\lambda = 7.47\text{nm}$ and a 2-meter pixel resolution. At the time of each image acquisition, observations on canopy and soil biophysical variables were collected within the field to allow analyses of relations between remote spectral data and ground truth. For an exhaustive description of experimental plan, readers can refer to [6]. For the validation step of the new method presented in this paper, only SPAD (chlorophyll meter) data and LAI data were available as ground truth reference; these data are briefly discussed in the sections thereafter.

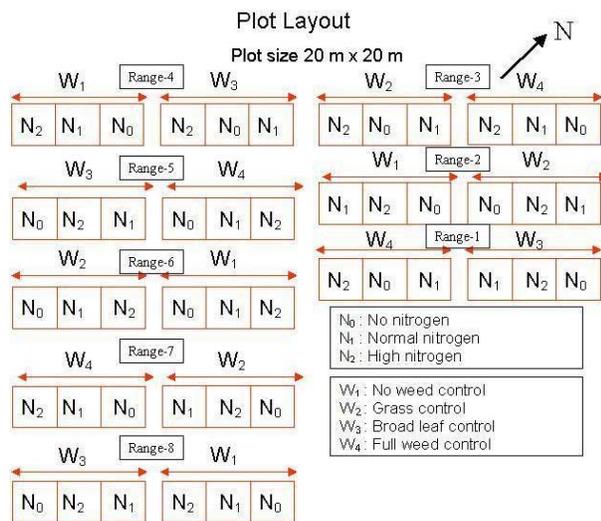


Figure 2. Experimental field plot layout

2.1.1 SPAD: Ground Truth Data

In this study, focus is made on canopy nitrogen content at the beginning of the reproductive stage with data taken early in August. During this period, 88 observations of plant nitrogen content are exploitable for analysis, taking into account that two types of measures were needed. Firstly, chlorophyll meter SPAD measurements were transformed into plant nitrogen concentration values, thanks to Dwyer's relation found in [5]. Secondly, these

data were weighted with Leaf Area Index value (LAI) to obtain a final value proportional to total nitrogen content in a pixel [18].

2.1.2 CASI hyperspectral data

Data were pre-processed by John Miller's research team from the Centre for Research in Earth and Space Science (CRESS) of York University in Toronto, Ontario, Canada. This pre-processing step comprises atmospheric correction, radiometric calibration and geometric corrections, which are mandatory operations prior to the creation of a geographically referenced database. All these geographic data were imported into a Geographic Information System (GIS) software, namely PCI GEOMATICA V9.1.6. A geographically referenced database was then created to allow easier data manipulations.

Once the pre-processing stage executed, a 3x3 mean filter was applied upon images for two reasons. Firstly, SPAD measurements were collected in a 1-meter perimeter around a flag, which implies a possible influence of neighbour pixels due to resolution. Secondly, accuracy of geometric correction is about 0.5 pixel [6] (about 1 meter because of the 2-meter image resolution). If one considers the error coming from a GPS for assessment of flags position, the use of the mean filter is then fully justified.

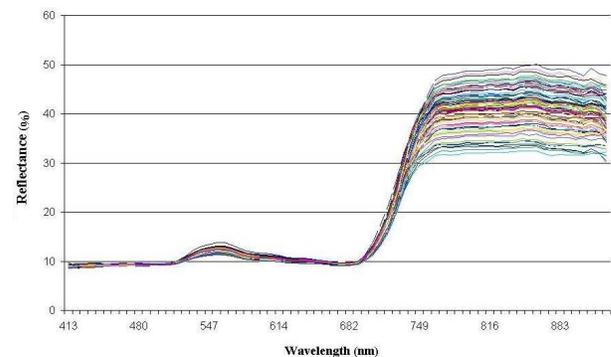


Figure 3. Spectral reflectance curves at the SPAD sample location

Once the initial processing steps performed, relevant data were extracted from the GIS environment to be treated by the GP-SVI algorithm developed using the Microsoft Visual C++ V6.0 environment. The CASI spectral curves of the 88 sampling points (points where nitrogen canopy is known) are shown in the Figure 3; it shows that data variability is mainly located in the green and in the near infrared areas of the electromagnetic spectrum. The following subsection present with the GP-SVI method. The choices and adjustments made are presented and justified.

2.2 GP-SVI Method Description

2.2.1 General Concept

The GP-SVI method is inspired by the Genetic Programming approach as proposed by Koza [10]. This method aims at

searching for the best SVI among evolving populations of potential solutions, regarding their ability to describe the observed crop biophysical variable. To do so, we evaluate the strength of the correlation between each individual and the ground truth samples; in fact, since an SVI can be considered as an image (an arithmetic combination of images is still an image), regression can be computed between ground truth data and SVI's grayscale values at sample locations. Four types of regressions were investigated: linear, logarithmic, exponential and power. The general framework of GP-SVI is described in Figure 4.

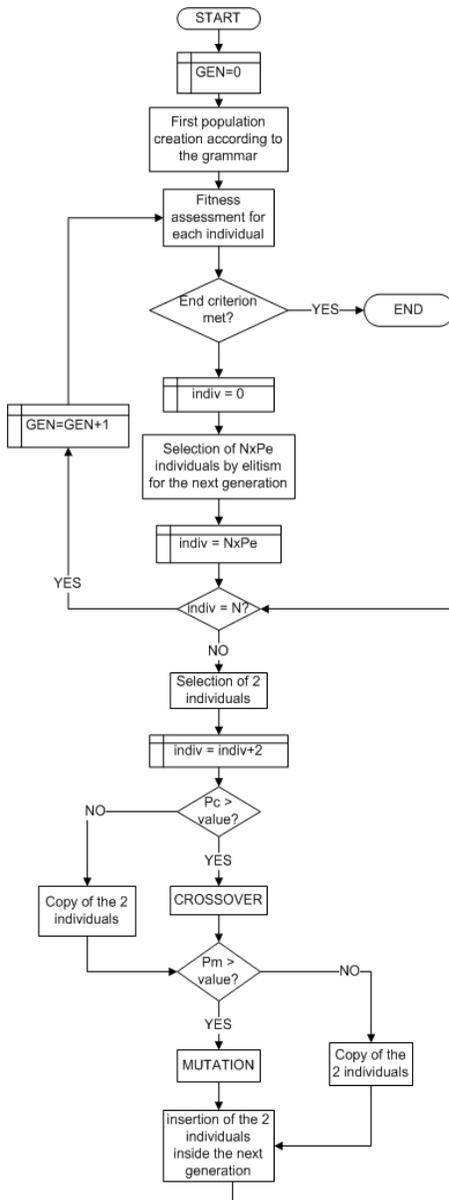


Figure 4. General framework of the GP-SVI method.

In the next subsection, we present the grammar and the representation mode of individuals.

2.2.2 Representation Mode and Grammar

The representation mode of population individuals is using a binary tree structure, which consists of a node holding a value which is followed by two nodes or zero; consequently, an individual is represented by a network of nodes in a “pyramidal structure” (i.e. thin at top and thick at bottom). This structure is chosen for its simplicity of implementation and its ability to describe the entire solution space; indeed, solutions are in the shape of arithmetic combinations of bands which can be assembled to build any type of arithmetic sequences. We also needed a reading convention to avoid ambiguity under non-commutative operator nodes (“/” and “-”); we choose to give the priority to the left branch of each node to settle this problem. Finally, reading is performed from the bottom to the top of the tree.

Now that a convention is chosen for the representation of individuals, we must define the grammar to construct these potential solutions.

The solutions that we are seeking have the shape of an SVI, which implies that each individual is a more or less complex sequence of arithmetic operations between spectral bands. Consequently, using Languages Theory formalism, we are able to propose the following context free grammar G, defined as a quadruple:

$$G = (N, T, P, S)$$

- N: set of variables (non terminal);
- T: set of constants (terminal);
- P: set of production rules;
- S: start symbols.

With the following definitions:

$$N = \{O, B, S\};$$

$$T = \{\alpha_i, b_j\}, \text{ with } 0 < i < 5 \text{ and } 0 < j < 71;$$

$$P = \{S \rightarrow O, O \rightarrow \alpha_i O O, O \rightarrow \alpha_i B O, O \rightarrow \alpha_i O B, O \rightarrow \alpha_i B B, B \rightarrow b_j\};$$

Where:

α_i represents the set of arithmetic operators, which means that $\alpha_i = \{“+”, “-”, “/”, “\times”\}$.

b_j represents the set of CASI spectral bands, which means that $b_j = \{b_1, b_2, \dots, b_{70}\}$.

Let’s note that a starting symbol S can’t be replaced with B because b_j is terminal and doing so would imply the presence in the initial population of “monoband individuals” (i.e. solution described by only one band); this would be a loss of diversity from the start.

These basic rules allow the creation of an initial generation and provide the basis for building the children resulting from genetic

operator applications. With these rules, we see the risk of individual's size bloat which brought us to introduce the parameter *max_prof* to limit individuals' size at the initialisation step; we also introduced a threshold parameter *max_node* to control individuals' size. Parameter values are given later in this section.

2.2.3 Fitness Function

As mentioned above, we computed regressions between each population SVI and ground measurements. This led to the determination of a Pearson coefficient R_k associated with each of the k SVI. Because of code bloat threat, and because we prefer a solution as short as possible (instinctively, we believe that a short solution is more likely to offer better generalization), we decided to introduce a means in the fitness function for penalizing long individuals. Consequently, we proposed the following fitness function $f(k)$ as given below:

$$f(k) = \frac{|R_k|}{\left[1 + \log\left(\frac{1 + L_k}{1 + L_{min}}\right) \right]}$$

Incidentally, the probability $P(k)$ for an individual k to be selected among the n -size population is given by:

$$P(k) = \frac{f(k)}{\sum_{i=1}^n f(i)}$$

The logarithm is introduced to smooth the effect of only dividing the Pearson coefficient by the length.

In the next subsection, we briefly describe the genetic operators chosen.

2.2.4 Genetic Operators

Exploration of the various types of genetic operators is beyond the scope of the present study. Subsequently, most of the chosen operators are classical ones.

As described in the GP-SVI framework in Figure 4, three genetic operators take place during the process. First, a proportion p_e of the solutions present in generation n is directly transferred into the next generation $n+1$. This transfer uses a strong elitism function, which is the simplest and the fastest to implement. Crossover operator occurs with probability p_c close to 1, considering that it is the core of the genetic process. A truncation point is randomly selected in each one of two selected solutions' structures and sub-trees are exchanged between these solutions to yield two new individuals; they systematically take place in the new population without consideration of their fitness values. Finally, the mutation operator occurs with a low probability p_m , on the output of crossover operator. This function randomly chooses a truncation

point in an individual's structure and replaces the above sub-tree with a randomly generated one.

Strategy for parameter values determination is described in the next subsection and the summary of these values is presented.

2.2.5 Parameters Values

Table 1. Simulation parameters

Parameter	Description	Value
Regression type	Nature of the regression	logarithmic
p_m	Probability of mutation occurrence	10%
p_c	Probability of crossover occurrence	98%
p_e	Proportion of individuals selected by elitism	10%
N_g	Number of generations	3000
<i>max_prof</i>	Maximum individual's depth in the first population	3
<i>max_node</i>	Maximum number of nodes for an individual	25
N_{ind}	Number of population individuals	300

To perform the simulations, we have set parameter values with the help of some sensitivity analysis as well as by trial and error approaches.

Here again, this step could be optimised using Genetic Algorithm or other techniques but it is also beyond the scope of this study. All chosen parameters are summarized in Table 1 below.

As regression types are concerned, observations made over 30 runs have shown a subtle superiority of the logarithmic model over 3 other types. A sensitivity analysis was conducted to set p_m , p_e and *max_prof* values whereas *max_node* and p_c were chosen more empirically. N_g and N_{ind} were set according to the performance offered by the simulation computer. Likewise, plotting generalization error (evaluated on a test data subset), we observed that overfitting occurred around the 2800th generation, which helped us setting $N_g = 3000$.

Now that parameters are chosen, the next section presents the preliminary results of GP-SVI and compare these with results obtain using other methods applied to the data described in the section II.

3. Results and Discussion

Table 2. Comparison of methods performance on a test data set

Method	RMSE% (best model)	RMSE% (mean on 30 runs)	Explained variance (best model)
GP-SVI	12.6%	13.9%	84.8%
GA-PLS [12]	15.7%	17.2%	73.9%
NDVI ₂ [8]	18.8%	-	70.2%

Preliminary results are presented and discussed in this section. We observed two main trends in our results while searching for a model explaining nitrogen concentration through a cornfield. Firstly, the GP-SVI method offered superior performances (on our dataset) over all other tested methods according to the two indicators mentioned above. Models derived using regression from all classical SVI and nitrogen measures were tested, and the best one (called NDVI₂) as proposed by Hansen and al. in [8], is presented in Table 2. Multiple regressions have been performed but results are not as good as with GA-PLS, which is why they are not presented in table 3. Secondly, concerning band selection, the results found are quite close to those obtained by GA-PLS and are presented in the Table 3.

Table 3. 10 most selected bands in the best models by GP-SVI and GA-PLS over 30 runs

Method	i th CASI spectral band									
	3	6	9	10	12	19	31	37	47	63
GP-SVI	3	6	9	10	12	19	31	37	47	63
GA-PLS [12]	9	10	11	18	19	20	26	31	43	62

Analysing Table 3, it is interesting to observe that GP-SVI and GA-PLS agree 40% of the time on the selected bands (b_9, b_{10}, b_{19} and b_{31}). Moreover, 20% of the other selected bands are just one band apart. This result is interesting considering that there are 70 spectral bands. We also noticed (not shown here) that b_{10} is largely the most selected CASI band over all runs; more than 50% of the best individuals contained b_{10} .

The best descriptive model is derived from logarithmic regression between the SVI called IV_1 and ground truth. This SVI is defined by the expression given hereafter and the derived descriptive model is presented in the Figure 5.

$$IV_1 = \frac{b_3 \times b_8 \times (b_{10})^3 \times b_{56}}{b_{31} \times (b_{37})^2 \times b_{70} \times (b_{34} - b_{63} + b_{47})}$$

The best predictive model shown in Figure 5 is defined by the relation $10.843 \times \ln(IV_1) - 19.987$ and is represented as a grayscale image. White values stand for high nitrogen content whereas black ones represent low nitrogen content. It is important to note that for each block of three adjacent $N_i W_j$ treatment combinations (see Figure 2), it is always possible to identify the N_1, N_2 and N_3 treatments.

We note that 3 bands are located in the blue part of the reflectance spectrum, 4 are in the near infrared, 2 in the red portion and finally 1 in the green. Figure 3 shows that data variance is mainly regrouped in the infrared part of the reflectance spectrum so that it seems normal to find good models composed of spectral bands in this area. Nevertheless, it may appear surprising to find several bands in the blue portion where data variance is low. We would also expect to find several bands in the green portion where data variance was quite important, and IV_1 has just one in this part. These remarks reveal that nitrogen content variance may be due to subtle variation in reflectance curves and this reinforces the fact that it is difficult to find SVI only by studying spectral curves.

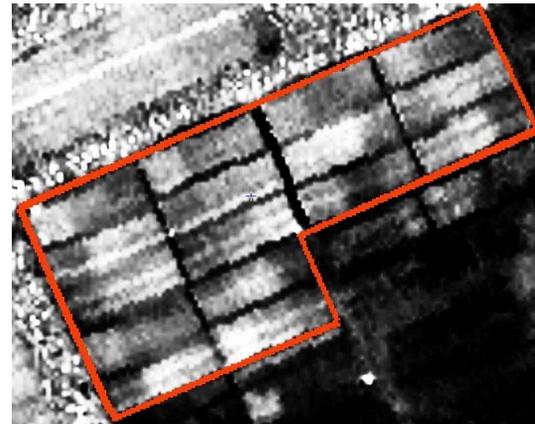


Figure 5. Best descriptive model found by GP-SVI over all performed runs.

Of course, it would be relevant to compare our results to other methods on a benchmark in order to assess rigorously GP-SVI performance. Nevertheless, preliminary results found are encouraging and leave the door open for further investigations and improvements.

4. CONCLUSION

Hyperspectral data contain large volume of information and extraction of relevant information for any specific application is a great challenge. Evolutionary computation (EC) approaches are less friendly than statistical techniques but they sometimes improve significantly the obtained results. Statistical methods are often chosen, partly because of their strong foundation and perhaps because they result in only one solution contrary to EC. Moreover, the great number of parameters to optimize and the threat of data overfitting or code bloat (for GP) also certainly tend to bridle their expansion. In the context of agricultural

applications, this study suggests that great improvements may be obtained by the GP-SVI method when compared to classical approaches considering the problem of crop nitrogen description. Although only preliminary results are presented here, it is important to notice that GP-SVI can be extended to any regression problem between a set of descriptive variables and a set of observations, and that not only for agricultural problems. On the nitrogen assessment problem, it is important to note that the CASI spectral range available for this study (409 nm to 947 nm) doesn't give any information about the middle infrared portion; however, new sensors like AVIRIS allow the acquisition of reflectance data at wavelengths up to 2500 nm, offering greater potential to elaborate accurate models.

5. REFERENCES

- [1] Bellman, R. Adaptive Control Processes: A Guided Tour, 1961.
- [2] Chen, K.S., Tzeng, Y.C., Chen, C.F., Kao, W.L. and Ni, C.L., Classification of multispectral imagery using dynamic learning neural network. in Proceedings of IGARSS '93 - IEEE International Geoscience and Remote Sensing Symposium, 18-21 Aug. 1993, (Tokyo, Japan, 1993), IEEE, 896-898.
- [3] Curran, P.J., Kupiec, J.A. and Smith, G.M. Remote Sensing the Biochemical Composition of a Slash Pine Canopy. IEEE Transactions on Geoscience and Remote Sensing, 35 (2). 415-420.
- [4] Dubrovin, V.I., Subbotin, S.A., Morshchavka, S.V. and Piza, D.M., The plant recognition on remote sensing results by the feed-forward neural networks. in ANNIE 2000. Smart Engineering Systems Design Conference, 5-8 Nov. 2000, (St. Louis, MO, USA, 2000), ASME, 697-702.
- [5] Dwyer, L.M., Anderson, A.M., Ma, B.L., Stewart, D.W., Tollenaar, M. and Gregorich, E. Quantifying the nonlinearity in chlorophyll meter response to corn leaf nitrogen concentration. Canadian Journal of Plant Science, 75 (1). 179-182.
- [6] Goel, P.K., Prasher, S.O., Landry, J.A., Patel, R.M., Bonnel, R.B., Viau, A.A. and Miller, J.R. Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. Computers and electronics in agriculture, 38. 99-124.
- [7] Goel, P.K., Prasher, S.O., Patel, R.M., Landry, J.A., Bonnell, R.B. and Viau, A.A. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. Computers and Electronics in Agriculture, 39 (2). 67-93.
- [8] Hansen, P.M. and Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment, 86. 542-553.
- [9] Jacquemoud, S. Utilisation de la haute résolution spectrale pour l'étude des couverts végétaux: développement d'un modèle de réflectance spectrale Méthodes Physiques en Télédétection, Paris VII, Paris, 1992, 92.
- [10] Koza, J.R. Genetic Programming-On the programming of Computers by Means of Natural Selection. MIT Press, Cambridge, London, 1992.
- [11] Kumar, S., Ghosh, J., Crawford, M.M. and Member, I. Best-Bases Feature Extraction Algorithms for Classification of Hyperspectral Data. IEEE Transactions on Geoscience and Remote Sensing, 39 (NO.7). 1368-1379.
- [12] Leardi, R. Application of genetic algorithm-PLS for feature selection in spectral data sets. Journal of Chemometrics, 14. 643-655.
- [13] Moran, M.S., Inoue, Y. and Barnes, E.M. Opportunities and limitations for image-based remote sensing in precision crop management. Remote Sensing of Environment, 61 (3). 319-346.
- [14] Peñuelas, J. and Filella, I. Technical Focus: Visible and near-infrared reflectance techniques for diagnosing plant physiological status. Trends in plant science, 3 (1). 151-156.
- [15] Peñuelas, J., Piñol, J., Ogayar, R. and Filella, I. Estimation of plant water content by the reflectance Water Index WI (R900/R970). International Journal of Remote Sensing, 18 (13). 2869-2875.
- [16] Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H. and Sorooshian, S. A modified soil adjusted vegetation index. Remote Sensing of Environment, 48 (2). 119-126.
- [17] Shock, B.M., Carpenter, G.A., Gopal, S. and Woodcock, C.E., ARTMAP neural network classification of land use change. in Proceedings of the World Congress of Computers in Agriculture and Natural Resources, 19-21 Sept. 2001, (Iguacu Falls, Brazil, 2001), American Soc. Agric. Eng, 22-28.
- [18] Strachan, I.B., Pattey, E. and Boisvert, J.B. Impact of nitrogen environmental conditions on corn as detected by hyperspectral reflectance. Remote Sensing of Environment, 80 (2). 213-224.
- [19] Xiang, Z., Suhong, L. and Zhenkun, T., A Method for estimating Chlorophyll Content of Wheat from Reflectance Spectra. in Geoscience and Remote Sensing Symposium, IGARSS'04, (2004), 4504-4507.
- [20] Yao, H. and Tian, L. A genetic-algorithm-based selective principal component analysis (GA-SPCA) method for high-dimensional data feature extraction. IEEE Transactions on Geoscience and Remote Sensing, 41 (6 PART I). 1469-1478.