

BAYESIAN NETWORK MODEL TO PREDICT AREAS FOR SUGARCANE EXPANSION IN BRAZILIAN CERRADO

Redes Bayesianas na Previsão de Áreas para Expansão da Cana-de-açúcar no Cerrado Brasileiro

Alexsandro Cândido de Oliveira Silva, Leila Maria Garcia Fonseca & Thales Sehn Körting

**Instituto Nacional de Pesquisas Espaciais – INPE
Divisão de Processamento de Imagens – DPI**

Caixa Postal 515-12227-010 – São José dos Campos – SP, Brasil
{alexsandro.silva, leila.fonseca, thales.korting}@inpe.br

*Received on January 18, 2017/ Accepted on April 21, 2017
Recebido em 18 de Janeiro, 2017/ Aceito em 21 de Abril, 2017*

ABSTRACT

The growing demand for ethanol has powered the shift of the sugarcane frontier into the Brazilian Cerrado, mainly in the states of Goiás and Mato Grosso do Sul. Therefore, this study aims to propose Bayesian Network models for identifying potential areas for sugarcane expansion in Goiás and Mato Grosso do Sul states. The models take into account constraint factors in relation to the sugarcane expansion such as topography, soil aptitude, climate conditions, and available infrastructures. Results showed that Bayesian Network models proposed in this study were able to represent the tendency of sugarcane expansion.

Keywords: Bayesian Networks, Probabilities, Suitable Areas, Sugarcane Expansion, Cerrado.

RESUMO

A crescente demanda por etanol tem impulsionado o deslocamento da fronteira da cana-de-açúcar para o Cerrado Brasileiro, principalmente nos estados de Goiás e Mato Grosso do Sul. Portanto, este estudo tem como objetivo propor modelos de Rede Bayesiana para identificar áreas potenciais para a expansão da cana-de-açúcar nos estados de Goiás e Mato Grosso do Sul. Os modelos levam em consideração fatores de restrição em relação à expansão da cana-de-açúcar, como topografia, aptidão do solo, condições climáticas e infraestruturas disponíveis. Os resultados mostraram que os modelos propostos neste estudo foram capazes de representar a tendência de expansão da cana-de-açúcar.

Palavras-chave: Redes Bayesianas, Probabilidades, Áreas Adequadas, Expansão da Cana-de-açúcar, Cerrado.

1. INTRODUCTION

The intense demand for ethanol in recent years has stimulated the sugarcane expansion to new areas, which intensify the competition with other agricultural areas, increase the land prices and reduce the options for further expansions (ADAMI *et al.*, 2012; GRANCO *et al.*, 2015). The saturation of traditional producing areas such as in the state of São Paulo (CASTRO *et al.*, 2010; SHIKIDA, 2013) along with higher production costs for the industry has motivated producers to search potential new areas for production of sugarcane.

The ethanol industry identified in the Brazilian Cerrado opportunities for investment, especially in the states of Goiás and Mato Grosso do Sul – the Cerrado Biome covers 97% of Goiás and 61% of Mato Grosso do Sul (BRASIL, 2009). However, even though climate and soil conditions are favorable to sugarcane cultivation (SHIKIDA, 2013) and land prices are affordable, few mills were operating in the region. Until 2005, only 22 mills had established in Goiás and Mato Grosso do Sul, which was a limiting factor for farmers to start planting sugarcane, since the crop needs to be promptly processed after harvesting (GRANCO *et al.*, 2015).

To enhance the attractiveness for the sugarcane industry, the government provided support through fiscal incentives, credit lines and investments in transportation infrastructure. As the result, the number of mills and areas planted with sugarcane in Goiás and Mato Grosso do Sul increased by approximately three times from 2005 to 2015 (GRANCO *et al.*, 2015). Consequently, these states, which previously had an economy centered on cattle ranching and grains (soy-bean and corn), witnessed a strong sugarcane expansion (SHIKIDA, 2013).

The rapid sugarcane expansion and the eventual land cover changes in the Cerrado led the Brazilian government to implement the Sugarcane Agroecological Zoning in 2009 to regulate the expansion and sustainable sugarcane production in Brazil (MANZATTO *et al.*, 2009). To identify potential areas for sugarcane crop in the Cerrado Biome, Ribeiro *et al.* (2015) carried out Boolean spatial analysis. However, Boolean analysis results only two classes (favorable and non-favorable), which do not adequately represent

the spatial phenomena. Other inference method for spatial data integration, such as Bayesian inference, can numerically express the potential of sugarcane areas from 0 to 1, which allows a decision surface (CÂMARA *et al.*, 2004).

Within this context, this study aims to propose a Bayesian inference model for identifying potential areas for sugarcane expansion in the states of Goiás and Mato Grosso do Sul. We used a Bayesian Networks approach: the enhanced Bayesian Network for Raster Data (e-BayNeRD) method (SILVA *et al.*, 2014), which is an enhanced version of the BayNeRD algorithm (MELLO *et al.*, 2013). The e-BayNeRD method is a probabilistic approach based on raster data observations and it is able to incorporate experts' knowledge for analysis. In the next section, we present a brief description about the theory of Bayesian Networks and the e-BayNeRD method employed in this study.

This paper is based on SILVA *et al.* (2016), previously presented in XVII Brazilian Symposium on Geoinformatics (GEOINFO 2016) (<http://www.geoinfo.info/geoinfo2016>).

2. BAYESIAN NETWORK MODEL

Bayesian Networks (BN) are defined in terms of two components: (i) qualitative component – a Directed Acyclic Graph (DAG), in which nodes represent the variables in the model and the statistical dependence between pair wise variables is indicated by directed arrows that start in a parent node and end in a child node, as illustrate on Figure 1; and (ii) quantitative component – probability functions associated to each variable denoting the strengths of the links in the BN model (AGUILERA *et al.*, 2011; LANDUYT *et al.*, 2013).

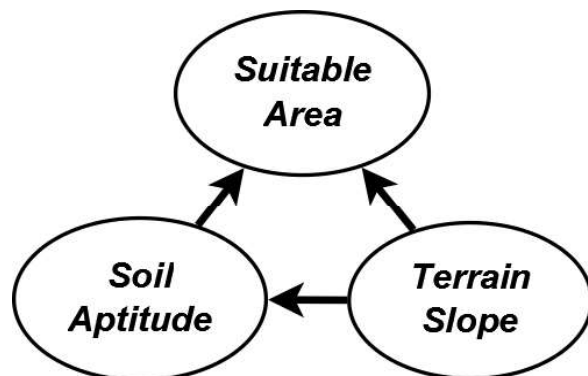


Fig. 1 - Example of Bayesian Network graphical model.

Figure 1 shows an example of BN graphical model, in which Suitable Area variable is statistically dependent of Soil Aptitude variable and both are statistically dependent of Terrain Slope variable. Prior probability is assigned to variable without parent (e.g.: Terrain Slope), whereas conditional probability is assigned to descendant ones (e.g.: Soil Aptitude and Suitable Area).

The prior knowledge of an event is updated taking into account new evidence through the Bayes' theorem (Eq. 1):

$$P(A = a | B = b) = \frac{P(B = b | A = a)P(A = a)}{P(B = b)} \quad (1)$$

in which $P(A=a)$ is the prior probability of the event A; $P(B=b | A=a)$ is the likelihood function and $P(A=a | B=b)$ is the posterior probability; and $P(B=b)$ is a normalizing constant (NEAPOLITAN, 2004). Upper-case letters denote the variables and the same but lower-case letters denote the state or value of the variable. This ability to compute posterior probabilities given new evidence is called inference.

Bayesian Networks Model has been used in many applications employing Bayesian inference to develop a plausible reasoning and to describe the occurrence probability of a phenomenon. Aguilera *et al.* (2011) argue that the BN have mainly been used as a technique for inference in Environmental Sciences. Dlamini (2010) presented a BN approach to estimate the probability of wildfire occurrence based on satellite-detected wildfires data and a set of social, physical, environmental and climatic factors.

McCloskey *et al.* (2011) proposed a BN to identify suitable sites for the economic development and landscape conservation, while Gonzalez-Redin *et al.* (2016) used BN approach to find areas for a sustainable timber production and biodiversity conservation. Mello *et al.* (2013) developed a BN model for raster data analysis to study soybean mapping based on remote sensing variables.

However, Aguilera *et al.* (2011) reported that few studies have used methods based on BN in agriculture applications. Therefore, we expected that this study contribute to show the potentiality of BN models, which are used to infer potential areas for sugarcane expansion in this study.

2.1 e-BayNeRD algorithm

To identify the appropriate areas for sugarcane expansion, we used the e-BayNeRD method (SILVA *et al.* 2014), which is briefly described below. Figure 2 shows the e-BayNeRD's workflow.

The method deals with raster data in GeoTiff format, in which each GeoTiff corresponds to a variable (node) in the BN model. The variable that represents the studied phenomenon is called target variable and its GeoTiff must contain the reference data for training. To illustrate, suppose that Suitable Area is the target variable in the BN model exemplified in Figure 1. The others variables in the model are called context variables, and they can have some relation with the target one and they can have some interrelation. After entering all the raster data, the user designs the BN graphical model (i.e., the DAG) by defining the relations among all variables.

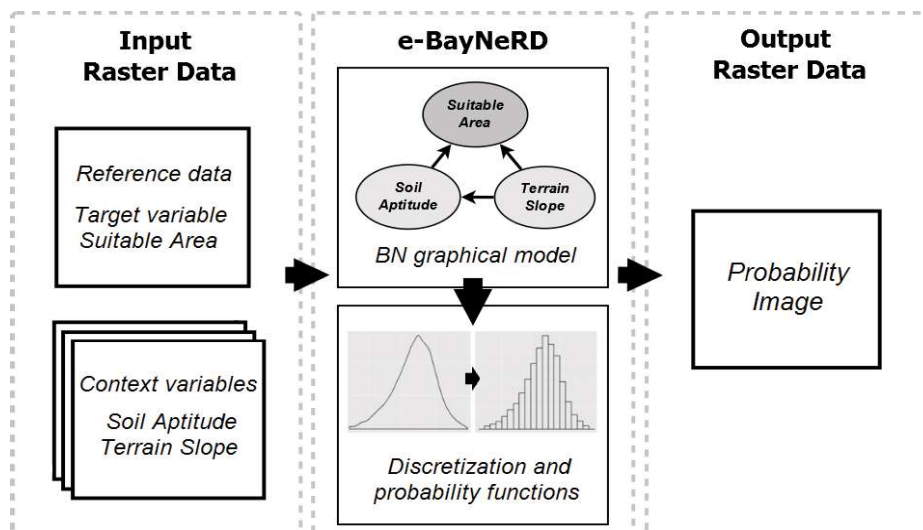


Fig. 2 - e-BayNeRD's Workflow.

Once the BN graphical model is defined, user needs to convert continuous context variables into categorical ones. In the discretization step, the range of observed values for each variable is divided into intervals according to the lower and upper limits chosen by the user. Each interval is one category; therefore, a variable will have n categories if the user discretizes it into n intervals. Figure 3 illustrates the Terrain Slope variable discretized in three categories: [0, 8), [8, 13), [13, max]. Values into the bars indicate the probability to find a pixel within the defined interval limits. The limits of the intervals for each variable should be appropriately chosen to describe the variable as best as possible according to the studied phenomenon.

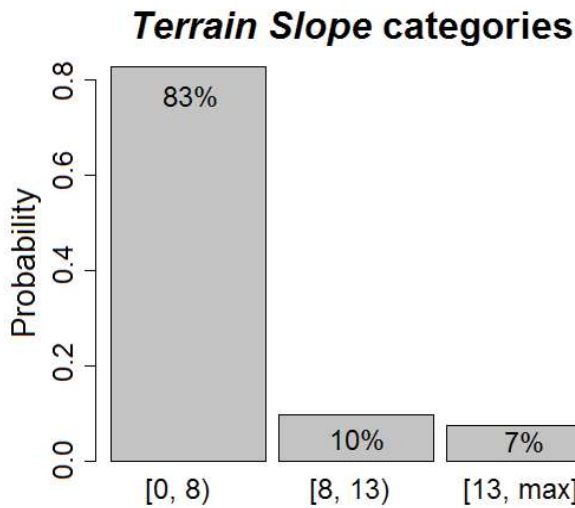


Fig. 3 – Example of the discretization step: Terrain slope variable discretized in three categories.

In the e-BayNeRD algorithm, probability functions are computed based on pixel counting according to the dependence relations among variables and their categories. Prior probability is assigned to those variables without parents (e.g.: Terrain Slope variable), whereas conditional probability is assigned to descendant ones. Table 1 shows the conditional probability table associate to the Soil Aptitude variable, in which High, Moderate and Low are categories for Soil Aptitude. Values in table are conditional probabilities $P(\text{Soil Aptitude} = i | \text{Terrain Slope})$, which indicate the probability of finding a pixel equal to some Soil Aptitude category given that this pixel is equal to some Terrain Slope category.

Table 1: Conditional probability Table of the soil aptitude variable given The Terrain slope variable

Soil Aptitude	Terrain Slope (in %)		
	[0, 8)	[8, 13)	[13, max]
High	0.713	0.524	0.497
Moderate	0.027	0.074	0.070
Low	0.260	0.402	0.433

After computing the probability functions associated with each variable, e-BayNeRD is able to calculate the probability of target presence given the values observed in the context variables. When the probability is computed for each pixel in the study area, the output, called Probability Image, is formed.

3. STUDY AREA

The study area covers Goiás and Mato Grosso do Sul states in the Brazilian Cerrado biome excluding some regions of no interest for sugarcane expansion, as showed in Figure 4. Some areas excluded in study are: urban areas, water bodies, areas under environmental protection laws (i.e. conservation units, indigenous lands, the Upper Paraguay River basin, Pantanal Biome), and areas where commercial sugarcane production is not allowed such as agrarian settlements or quilombo communities.

Figure 4 shows also cultivated sugarcane areas in Goiás and Mato Grosso do Sul for 2008/2009 crop year (dark green color). These regions were excluded from the study area because they were also excluded in the Sugarcane Agroecological Zoning, which was used as a reference data for training and assessing the model to infer potential areas for sugarcane expansion proposed in this work.

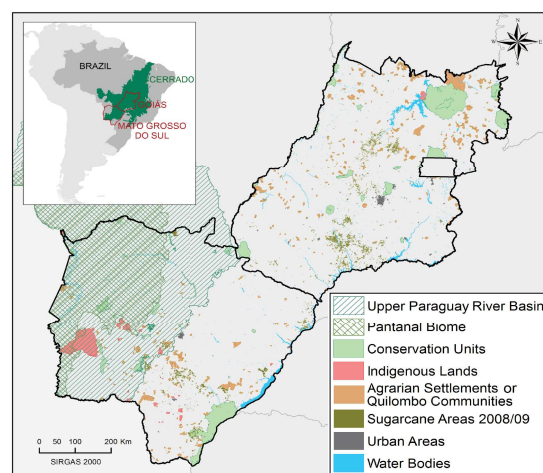


Fig. 4 - States of Goiás and Mato Grosso do Sul and masked out areas from the study.

4. SELECTED VARIABLES TO BUILD THE BN MODEL

This section presents the variables chosen to compose the BN model. The target variable represents appropriate areas for sugarcane expansion and the context variables are constraining factors that exhibit any kind of relation with the sugarcane expansion.

4.1 Target variable

Most of the suitable areas for sugarcane expansion in Brazil are located in Goiás and Mato Grosso do Sul states (MANZATTO *et al.*, 2009). Figure 5 shows the target variable Suitable Area for sugarcane expansion according to the Sugarcane Agroecological Zoning (Source: <http://geo.cnpma.embrapa.br/>). Suitable and non-suitable areas are denoted by green and yellow colors, respectively. About 70% of the pixels in each class were randomly selected to compose the reference data for training. The remaining 30% were used for accuracy assessment.

4.2 Context variables

Considering that Goiás and Mato Grosso do Sul states have topography that is adequate for sugarcane expansion (SHIKIDA, 2013), we selected Terrain Slope variable as a context variable in the model. Terrain Slope variable was settled as parent of Suitable Area variable in the BN model. (Terrain Slope variable source: <http://srtm.csi.cgiar.org/>).

Soil in our study area presents favorable conditions for sugarcane cultivation although there are some poor-nutrient areas as stated by (SHIKIDA, 2013). As soil conditions influence sugarcane cultivation, we set Soil Aptitude variable as parent of Suitable Area variable. Considering that terrain slope is an important factor to determine the aptitude to some crop cultivation, we set Soil Aptitude variable as descendant of Terrain Slope variable. (Soil Aptitude variable source: <http://mapas.mma.gov.br/i3geo/datadownload.htm>).

Furthermore, it is important to consider climate risks to cultivate sugarcane (BRUNINI *et al.*, 2008). In this sense, the climate data annual mean Temperature and annual Precipitation were also selected as context variables. As these

variables indicate suitable climate conditions regions for sugarcane cultivation, we set them as parent of Suitable Area variable in the BN model (Temperature and Precipitation variables source: <http://www.worldclim.org/>).

Sugarcane expansion in the South-central region of Brazil has occurred mainly on pasturelands and eventually on annual crops lands (ADAMI *et al.*, 2012; CASTRO *et al.*, 2010). Taking this into account, the land use maps produced by IBGE (12 classes) were rasterized to represent the Land Use variable. Land use classes were grouped in five classes: agriculture; planted pasture; forest; grassland; and natural pasture. Agriculture, planted and natural pasture represent 11%, 45% and 40% of the land use, respectively. Considering that each land use class sets limits to the sugarcane expansion, Land Use variable is considered as Suitable Area variable parent. We define Land Use variable as descendant of Terrain Slope and Soil Aptitude variables (Land Use variable source: <http://portaldemapas.ibge.gov.br/>).

Another context variable is the Distance to the nearest Water Body, which was computed using the drainage network. This variable was selected to take into account conservation of natural vegetation around water bodies/rivers. This suggests that sugarcane cultivation does not occur in areas close to water bodies. Therefore, Distance to the nearest Water Body variable is set as parent of Suitable Area variable. It is noteworthy that small distances to water bodies are generally associated with higher slope. Therefore, Distance to the nearest Water Body variable is also descendant of Terrain Slope variable (Drainage network source: <http://portaldemapas.ibge.gov.br/>).

Installation of new mills was also a driver for sugarcane expansion in Goiás and Mato Grosso do Sul states, usually located close to roads to facilitate the production transport (GRANCO *et al.*, 2015). Hence, the Distance to the nearest Mill and Distance to the nearest Road were selected as context variables, and they were built from highways network and mills location, respectively. For logistical and economic reasons, it is expected that the sugarcane expansion occurs near mills and roads. Thus, the Distance to the nearest Mil and

Distance to the nearest Road variables can be defined as parents of the Suitable Area variable in BN model (Highways network source: <http://portal demapas.ibge.gov.br/> and points of mills location source: <http://ctbe.cnpem.br/pesquisa/producaobiomassa/cana-info/>).

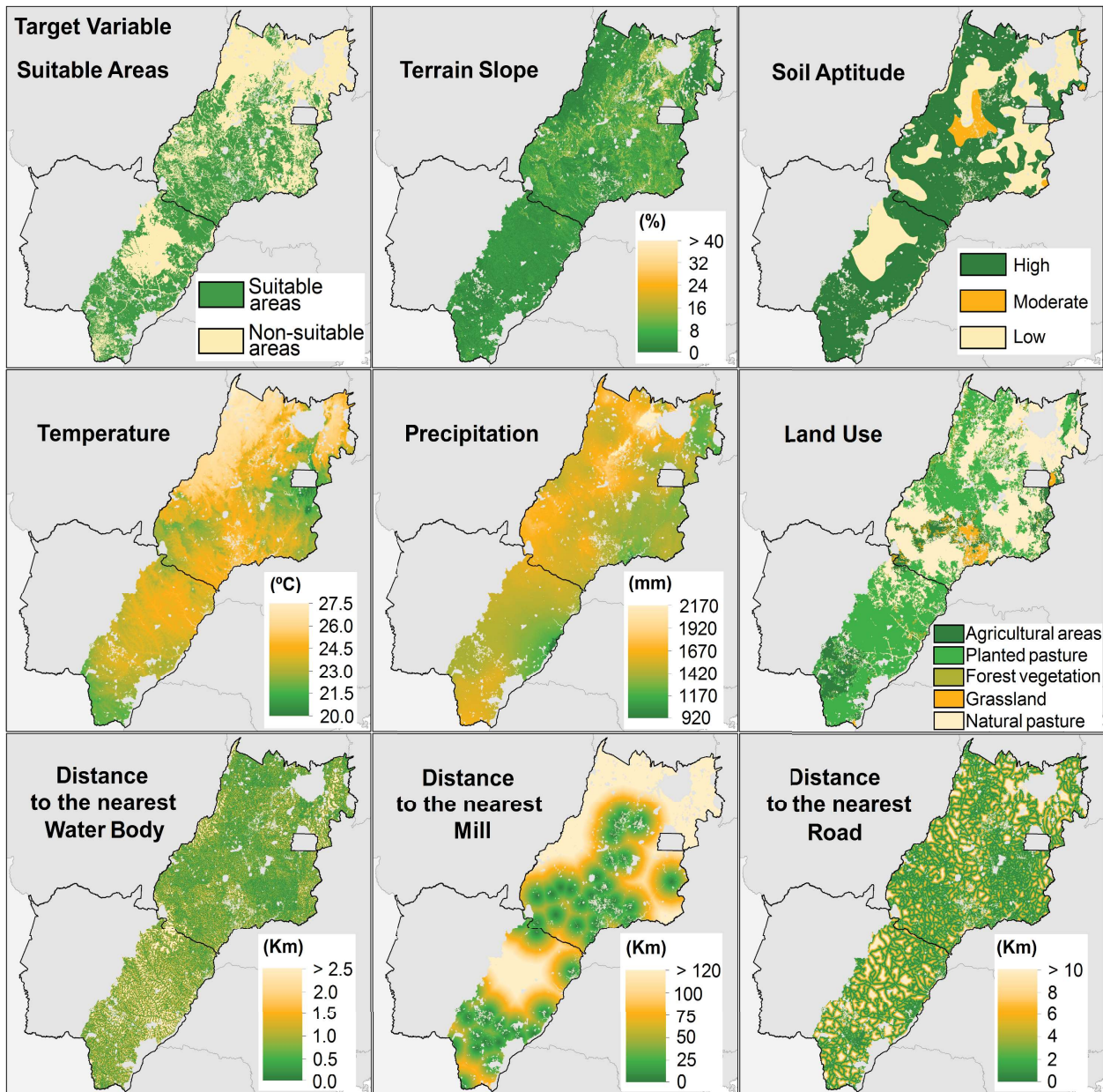


Fig. 5 - GeoTiff data of target and context variables.

4.3 BN graphical models and discretization phase

Two BN models were created. The first one identifies sustainable areas for sugarcane expansion taking into account the variables Terrain Slope, Soil Aptitude, Precipitation, Temperature, Land Use and Distance to the nearest Water Body. The second one specifies logistically appropriate areas as well, considering the variables Distance to the nearest Mill and Distance to the nearest Road. Additionally,

the Probability Image produced by the first BN model was incorporated into the second BN model as a context variable. Figure 6 shows the first and second BN models.

The limits of the intervals should be appropriately chosen to describe as best as possible the context variable according to the target variable Suitable Area. Table 2 presents the interval limits and the categories defined for context variables, which were chosen based on our knowledge and experience.

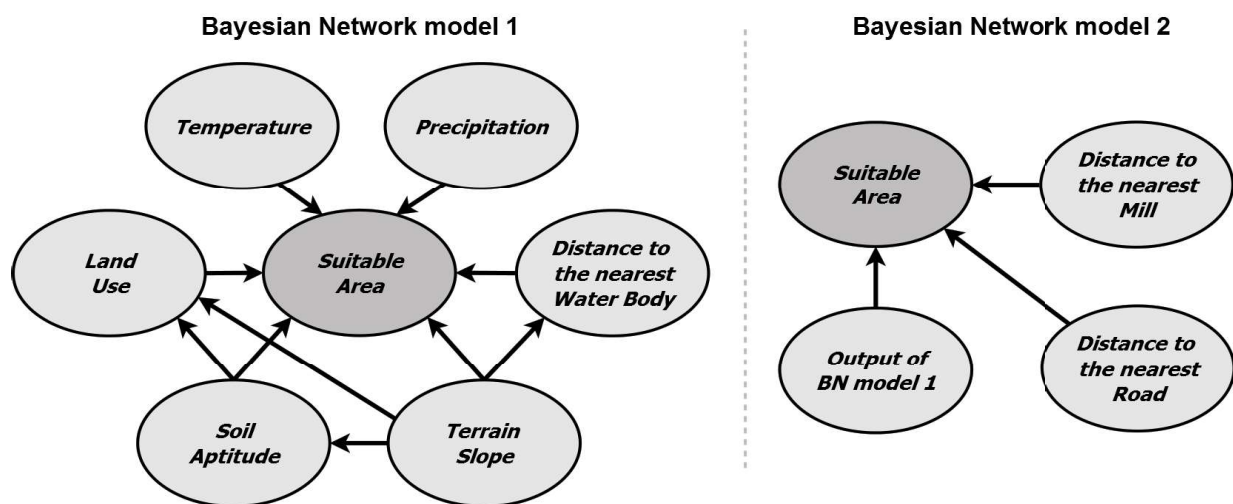


Fig. 6 - BN model 1 (left): identify sustainable areas; BN model 2 (right): identify sustainable and logistically appropriate areas.

Table 2 - Intervals and categories defined for context variables

Context variable	Terrain Slope (%)	Soil Aptitude	Land Use	Temperature (°C)
	[0, 8)	High	Agriculture	[20, 23.6)
	[8, 13)	Moderate	Planted pasture	[23.6, 24.9)
	[13, max]	Low	Forest	[24.9, max]
	-	-	Grassland	-
	-	-	Natural pasture	-
Context variable	Precipitation (mm)	Distance to the nearest Water Body (km)	Distance to the nearest Mill (km)	Distance to the nearest Road (km)
	[920, 1600)	[0, 0.25)	[0, 25)	[0, 1)
	[1600, max]	[0.25, 0.5)	[25, 50)	[1, 2)
	-	[0.5, 1)	[50, max]	[2, 5)
	-	[1, max]	-	[5, max]

5. RESULTS AND DISCUSSION

Probability Image is the main e-BayNeRD outcome, in which each pixel value represents the probability of such area to be a sugarcane growing area given the observed context variables values. The Probability Images produced by both BN models are shown in Figures 7. Green colored pixels represent areas with high probability for sugarcane expansion. High probability values were achieved when context variables exhibited good conditions for sugarcane plantation: Terrain Slope < 8%; high Soil Aptitude; Temperature < 24.9 °C; Distance to the nearest Water Body > 1km; agriculture or grassland as Land Use; Distance to the nearest Mill < 50km; Distance to the nearest Road < 5km; and Precipitation had no much influence. This result is similar to those obtained by Ribeiro *et al.* (2015). This result can be explained by two factors: agricultural aptitude and logistical issues.

According to Castro *et al.* (2010) and Silva & Miziara (2011), there is sugarcane expansion tendency in the south-central Goiás state. Areas in the south of Goiás present appropriate conditions for sugarcane cultivation. The Sugarcane Agroecological Zoning (MANZATTO *et al.*, 2009) pointed out that most areas in Goiás have conditions for sugarcane expansion. Our BN models showed high probabilities for sugarcane expansion in the south-central Goiás. Silva & Miziara (2011) argue that sugarcane spatial dynamics is directly related to the installation of new mills. Indeed, our results showed a similar spatial distribution between mills locations (context variable Distance to the nearest Mill in Figure 5) and potential areas for sugarcane expansion. This distribution is known as the south-central expansion axis (CASTRO *et al.*, 2010), which can be observed in our results showed in Figure 7.

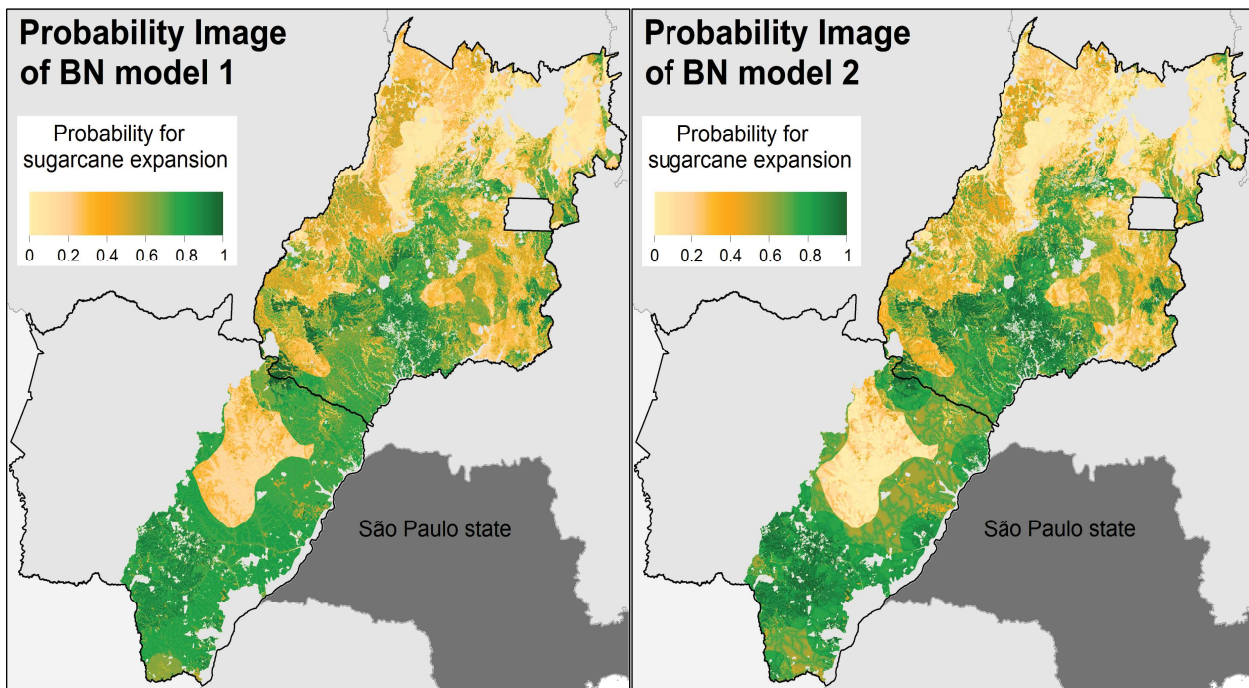


Fig. 7 - Probability Images of: BN model 1 (left) and BN model 2 (right). Green colored pixels represent areas with high probability for sugarcane expansion.

On other hand, sugarcane expansion is more restricted in the north of Goiás, where Terrain Slope (above 13%) and Soil Aptitude (low) are not favorable. Results also presented low probability values in the west and east of Goiás. In this case, the limiting factor in the modeling was the high annual Temperature in the west and the high Terrain Slope and the low Soil Aptitude in east.

Regarding the state of Mato Grosso do Sul, Probability Images in Figure 7 show that the most suitable regions for sugarcane expansion are located in the eastern and southern regions, which are near São Paulo state, the main domestic market for ethanol. This result is in accordance to those presented by Manzatto *et al.* (2009) and Ribeiro *et al.* (2015). Mato Grosso do Sul state presents favorable agricultural and environmental conditions for sugarcane cultivation. However, low probability values in the central region of the state occurred mainly due to low Soil Aptitude in this region. Pasture represents most land use in Mato Grosso do Sul, but the highest probability values in Mato Grosso do Sul were achieved in agriculture areas in the southern of the state.

In addition to the good agriculture and

climate conditions in the Mato Grosso do Sul state, it is important to observe that most mills are concentrated in the southern state (context variable Distance to the nearest Mill in Figure 5). It is possible to note that potential areas for sugarcane expansion stated more specific in the result of second BN model, which considered the Distance to the nearest Mill and Distance to the nearest Road context variables. Mills distribution in the southern Mato Grosso do Sul facilitates the logistics and transportation of ethanol, since this region has better road infrastructure connecting it with São Paulo state (GRANCO *et al.*, 2015).

5.1 BN models assessment

We analyzed the Probability Images (Figure 7) for sugarcane areas only in 2012/2013 crop year, which correspond to 955100 hectares approximately, according to the Canasat project. Sugarcane areas were monitored in the South-central region of Brazil by the Canasat Project from 2005 to 2013 (RUDORFF *et al.*, 2010).

Probability Images can be used to indicate the best regions for sugarcane expansion. Based on these probabilities, we sliced the range of probability values in aptitude categories for

sugarcane expansion, as illustrated in Figure 8. We verified that the expansion has mostly occurred in areas in which BN models considered as high probability of sugarcane cultivation. For BN model 1, about 75% of sugarcane areas have probability ≥ 0.7 ; for BN model 2, about

85% of sugarcane areas have probability ≥ 0.7 . These results indicate that the second BN model, which infer about sustainable and logistically appropriate areas for sugarcane expansion, is better than the first BN model.

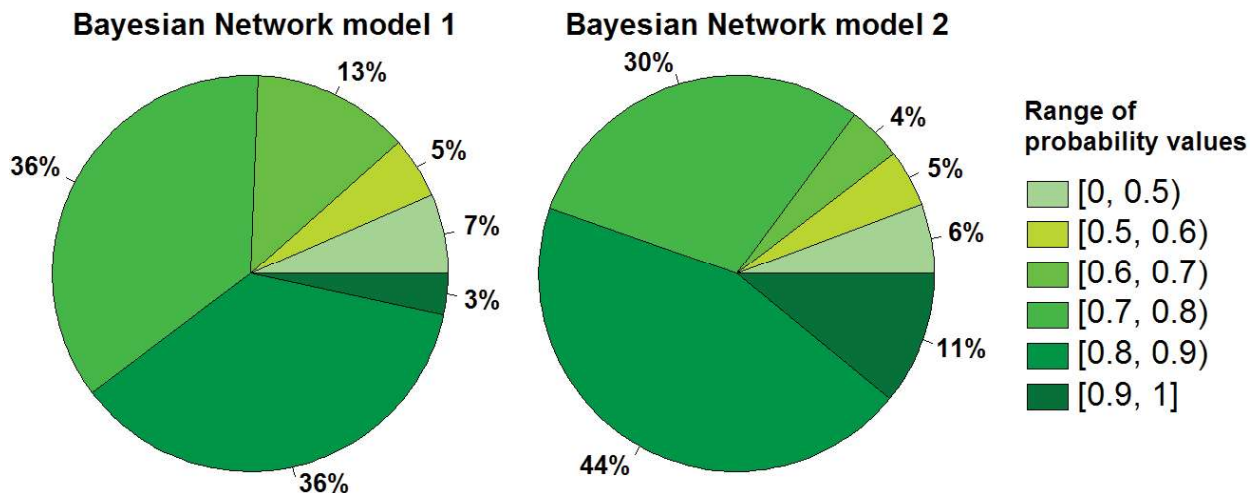


Fig. 8 - Percentage of sugarcane areas in the 2012/2013 crop year according to the range of probability values for each BN model.

Two complementary indices are commonly used in binary classifications: sensitivity, which indicates true positives rate (target areas labeled as target), and specificity, which indicates true negatives rate (non-target areas labeled as non-target). As BN models output a predicted probability, several thematic maps were produced by varying the probability threshold from 0 to 1. For example, if threshold = 0, all pixels are classified as suitable area. If all pixels that are suitable for sugarcane cultivation are correctly classified, then the sensitivity value is equal to 100%. On the other hand, pixels which are not suitable but are misclassified lead to specificity value equals to 0%.

The Receiver Operating Characteristic (ROC) curve (HANLEY; MCNEIL, 1982) is a graph that shows performance of a binary classifier in terms of two indices. ROC curve graph is plotted with sensitivity (Y-axis) versus 1-specificity (X-axis). To generate the ROC curve it is necessary to plot sensitivity versus 1-specificity for all possible classification

thresholds. The upper left corner is the best point, where both indexes are equal to 100%.

Figure 9 shows the ROC curve for both BN models. In the ROC curve, points plotted above the diagonal represent a classification better than a random one (random guess). This means that BN models did a good job in separating suitable areas and non-suitable areas classes. In the Probability Image for BN model 1, threshold = 0.45 resulted in a point closest to the upper left corner with sensibility = 84% and specificity = 67%. In the BN model 2, threshold = 0.51 resulted in a point with sensibility = 85% and specificity = 72%.

Area under the curve (AUC) is used to quantify a BN model performance. AUC is the percentage of area under the ROC curve (FAWCETT, 2006). Therefore, points in the ROC curve closest as possible to the upper left corner represent the best accuracy. AUC for BN model 1 and BN model 2 was 82% and 84%, respectively. This indicates that the BN models are able to differentiate suitable area and non-suitable area classes.

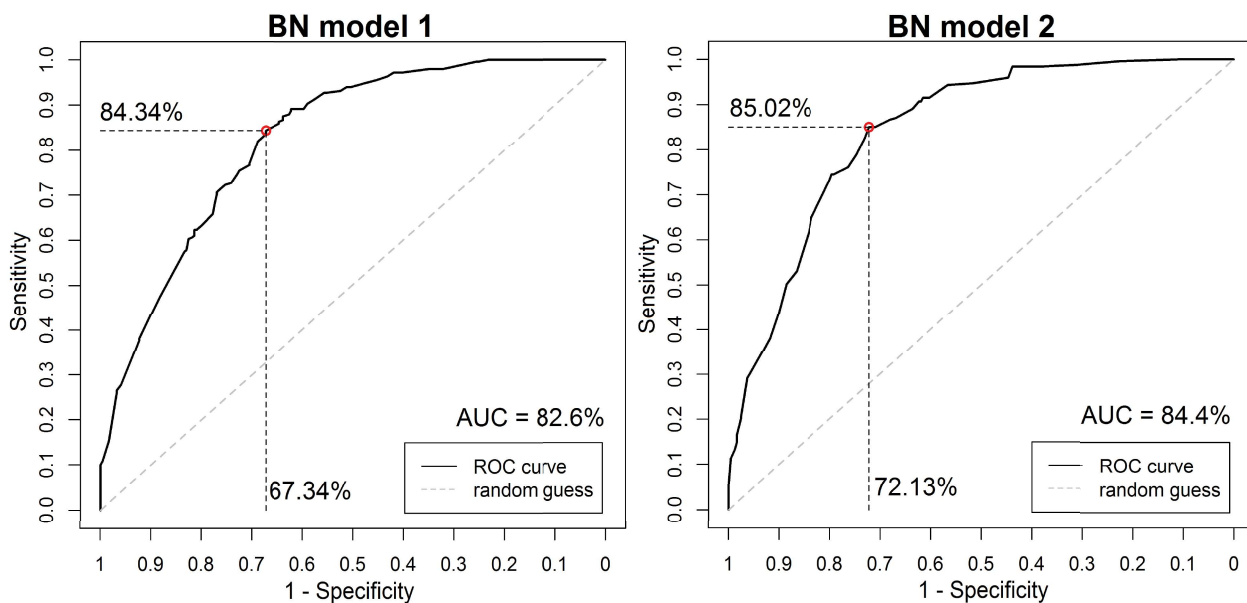


Fig. 9 - ROC curve of the first (left) and second (right) BN model.

6. CONCLUSION

This work used Bayesian Network method based on raster data observations to identify adequate areas for sugarcane expansion in the Brazilian Cerrado in the states of Goiás and Mato Grosso do Sul. Most of new sugarcane areas in recent crop years occurred in regions in which the model assigned high probability values ($> 70\%$) to them. This indicates that the Probability Image output of BN models can be used to direct the sugarcane expansion to most suitable areas. The Bayesian Network model was able to distinguish suitable and non-suitable areas for sugarcane expansion and it can be used as a planning decision tool.

In general, the study area presented favorable agro-environmental conditions for sugarcane plantation. Among all constraining factors, Terrain Slope and Soil Aptitude variables were the main drivers for sugarcane expansion. The highest probability values for expansion were achieved in agricultural areas and grassland. Future research should consider only pasture areas for sugarcane expansion, since there is a concern about competition with other agricultural areas; and also compare the e-BayNeRD method with others similar algorithms.

ACKNOWLEDGEMENTS

The authors thank to CAPES for financial support; and the reviewers for their valuable comments and inputs.

REFERENCES

- ADAMI, M.; RUDORFF, B. F. T.; FREITAS, R. M.; AGUIAR, D. A.; SUGAWARA, L. M.; MELLO, M. P.. Remote Sensing Time Series to Evaluate Direct Land Use Change of Recent Expanded Sugarcane Crop in Brazil. *Sustainability*, v. 4, n. 12, p. 574–585, 2 abr. 2012.
- AGUILERA, P. A.; FERNÁNDEZ, A.; FERNÁNDEZ, R.; RUMÍB, R.; SALMERÓN, A.. Bayesian networks in environmental modelling. *Environmental Modelling & Software*, v. 26, n. 12, p. 1376–1388, dez. 2011.
- BRASIL, M. A. **Conheça os biomas brasileiros**. Available in: <<http://www.brasil.gov.br/meio-ambiente/2009/10/biomas-brasileiros>>. Access: aug 10 2016.
- BRUNINI, O.; CARVALHO, J.P.; PÁDUA JÚNIOR, A. L.; MORAES, J. F. L.; PRADO, H.; LANDELL, M. G. A.. **Zoneamento de culturas bioenergéticas no estado de São Paulo: aptidão edafoclimática da cultura da cana-de-açúcar**. Campinas, IAC, 31p. 2008.
- CASTRO, S. S.; ABDALA, K.; SILVA, A. A.; BÔRGES, V. M. S.. A expansão da cana-de-açúcar no Cerrado e no estado de Goiás: elementos para uma análise espacial do processo. *Boletim Goiano de Geografia*, v. 30, n. 1, p. 171–191, 31 ago. 2010.
- DLAMINI, W. M. A Bayesian belief network

- analysis of factors influencing wildfire occurrence in Swaziland. **Environmental Modelling & Software**, v. 25, n. 2, p. 199–208, fev. 2010.
- FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861–874, jun. 2006.
- GONZALEZ-REDIN, J.; LUQUE, S.; POGGIO, L.; SMITH, R.; GIMONA, A.. Spatial Bayesian belief networks as a planning decision tool for mapping ecosystem services trade-offs on forested landscapes. **Environmental Research**, v. 144, p. 15–26, 2016.
- GRANCO, G.; CALDAS, M. M.; BERGTOLD, J. S.; SANT'ANNA, A. C.. Exploring the policy and social factors fueling the expansion and shift of sugarcane production in the Brazilian Cerrado. **GeoJournal**, n. Goldemberg 2007, p. 1–18, 27 ago. 2015.
- HANLEY, J. A; MCNEIL, B. J. The Meaning and Use of the Area under a Receiver Operating (ROC) Curvel Characteristic. **Radiology**, v. 143, n. 1, p. 29–36, 1982.
- LANDUYT, D.; BROEKX, S.; D'HONDT, R.; ENGELN, G.; AERTSENS, J.; GOETHALS, P. L. M.. A review of Bayesian belief networks in ecosystem service modelling. **Environmental Modelling & Software**, v. 46, p. 1–11, ago. 2013.
- MANZATTO, C. V.; ASSAD, E. D.; BACCA, J. F. M.; ZARONI, M. J.; PEREIRA, S. E. M.. **Zoneamento agroecológico da cana-de-açúcar**. Rio de Janeiro, Embrapa Solos, 55p. 2009.
- MCCLOSKEY, J. T.; LILIEHOLM, R. J.; CRONAN, C. Using Bayesian belief networks to identify potential compatibilities and conflicts between development and landscape conservation. **Landscape and Urban Planning**, v. 101, n. 2, p. 190–203, 2011.
- MELLO, M.; RISSO, J. ATZBERGER, C.; APLIN, P.; PEBESMA, E.; VIEIRA, C.; RUDORFF, B.. Bayesian Networks for Raster Data (BayNeRD): Plausible Reasoning from Observations. **Remote Sensing**, v. 5, n. 11, p. 5999–6025, 15 nov. 2013.
- CÂMARA, G.; MOREIRA, F. R.; BARBOSA, C.; ALMEIDA FILHO, R.; BÖNISCH, S.. Inferência Geográfica e Suporte à Decisão. In: **Introdução à Ciência da Geoinformação**. São José dos Campos: INPE, 2004. cap 9, p. 241–288. Available in: < <http://mtc-m12.sid.inpe.br/col/sid.inpe.br/sergio/2004/04.22.07.43/doc/publicacao.pdf> >. Access: aug 10 2016.
- NEAPOLITAN, R. E. **Learning Bayesian Networks**. 2. ed. New Jersey: Person Prentice Hall, 693p. 2004.
- RIBEIRO, N. V.; FERREIRA, L. G.; FERREIRA, N. C. Padrões e impactos ambientais da expansão atual do cultivo da cana-de-açúcar : uma proposta para o seu ordenamento no bioma Cerrado. **Ateliê Geográfico**, v. 9, n. 2, p. 99–113, 2015.
- RUDORFF, B. F. T.; AGUIAR, D. A.; SILVA, W. F.; SUGAWARA, L. M.; ADAMI, M.; MOREIRA, M. A.. Studies on the Rapid Expansion of Sugarcane for Ethanol Production in São Paulo State (Brazil) Using Landsat Data. **Remote Sensing**, v. 2, n. 4, p. 1057–1076, 9 abr. 2010.
- SHIKIDA, P. F. A. Expansão canavieira no Centro-Oeste: limites e potencialidades. **Revista de Política Agrícola**, v. XXII, n. 2, p. 122–137, 2013.
- SILVA, A. A.; MIZIARA, F. Avanço do setor sucroalcooleiro e expansão da fronteira agrícola em Goiás. **Pesquisa Agropecuária Tropical**, v. 41, n. 3, p. 399–407, 6 jul. 2011.
- SILVA, A. C. O.; FONSECA, L. M. G.; KÖRTING, T. S. **Bayesian network model to predict areas for sugarcane expansion in Brazilian Cerrado**. In: XVII Brazilian Symposium on GeoInformatics, 2016. **Proceedings of...** Campos do Jordão. National Institute for Space Research, 2016. p. 216–227.
- SILVA, A. C. O.; MELLO, M. P.; FONSECA, L. M. G. **Enhancements to the Bayesian Network for Raster Data (BayNeRD)**. In: XV Brazilian Symposium on Geoinformatics, 2014. **Proceedings of...** Campos do Jordão. National Institute for Space Research, 2014. p. 73–82.