

**Universidade de São Paulo
Instituto de Matemática e Estatística**

Centro de Estatística Aplicada

Relatório de Análise Estatística

RAE-CEA–19P01

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

“Detecção e monitoramento de escorregamentos em encostas naturais a partir da utilização integrada de imagens espectrais e modelos de elevação”

**Lucas Bortolucci
Thiago Henrique Martinoni
Profa. Júlia Pavan Soler**

São Paulo, junho de 2019

CENTRO DE ESTATÍSTICA APLICADA - CEA – USP

TÍTULO: Relatório de Análise Estatística sobre o Projeto: “Detecção e monitoramento de escorregamentos em encostas naturais a partir da utilização integrada de imagens espectrais e modelos de elevação”.

PESQUISADOR: Guilherme Pereira Bento Garcia

ORIENTADOR: Prof. Dr. Carlos Henrique Grohmann de Carvalho

INSTITUIÇÃO: Instituto de Geociências – IGc usp

FINALIDADE DO PROJETO: Doutorado

RESPONSÁVEIS PELA ANÁLISE: Lucas Bortolucci

Thiago Henrique Martinoni

Júlia Pavan Soler

REFERÊNCIA DESTE TRABALHO: SOLER, J.M.P.; BORTOLUCCI, L.; MARTINONI, T.H. **Relatório de análise estatística sobre o projeto: “Detecção e monitoramento de escorregamentos em encostas naturais a partir da utilização integrada de imagens espectrais e modelos de elevação de alta resolução.”.** São Paulo, IME-USP, 2019. (RAE–CEA-19P01)

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

BUSSAB, W.O.; MORETTIN, P.A. (2017). **Estatística Básica**. 9ª ed. São Paulo: Saraiva.
JOHNSON, R. A.; WICHERN, D. W. (1992). **Applied Multivariate Statistical Analysis**. 3.ed. Upper Saddle River, NJ: Prentice Hall.

EVERITT, B.; HOTHORN, T. (2004) **An Introduction to Applied Multivariate Analysis with R**. New York: Springer.

GARCIA, G.P.B (2019). **Detecção e monitoramento de escorregamentos em encostas naturais a partir da utilização integrada de imagens multiespectrais e modelos de elevação de alta resolução**. São Paulo. 112p. Qualificação (Doutorado). Instituto de Geociências - USP.

PAULA, G.A. (2013). **Modelos de Regressão com apoio computacional**. São Paulo: Instituto de Matemática e Estatística - USP.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Microsoft Word Online

Microsoft Excel for Windows (versão 2016)

Rstudio (versão 1.2.1335)

R (versão 3.5.0)

Python (versão 3.7.1)

TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise Descritiva Multidimensional (03:020)

Análise de Componentes Principais (06:070)

Regressão Logística (07:090)

Regressão Não Linear (07:140)

ÁREA DE APLICAÇÃO

Ciências Físicas e Geoestatística (14:010)

Resumo

Com o objetivo de identificar escorregamentos de terra localizados em encostas naturais localizadas na Serra do Mar, o presente estudo busca estabelecer uma metodologia para a classificação de pixels vinculados a imagens feitas por satélites como pertencentes ou não a áreas de deslizamento. As informações utilizadas para essa classificação estão ligadas a outros estudos na área do geoprocessamento, sendo decorrentes de métodos bem estabelecidos dentro deste campo de pesquisa. A análise descritiva dos dados indicou que os índices de vegetação possuem uma correlação forte entre si, porém quase não se mostram correlacionados com as variáveis de terreno. Dentre os vários modelos paramétricos de classificação avaliados o que se mostrou mais eficiente na predição de áreas de escorregamento foi o modelo linear generalizado com ligação log-log complementar para a média. Além disso, um modelo de redes neurais usado no aprendizado de interesse também se mostrou promissor.

Sumário

1. Introdução	7
2. Objetivos	7
3. Descrição do estudo	8
4. Descrição das variáveis	8
5. Análise descritiva	11
6. Análise de regressão	15
7. Conclusão	17
8. Apêndice A	19
9. Apêndice B	25

1. Introdução

“A ocorrência de movimentos gravitacionais de massa, em encostas naturais ou em taludes, pode ser entendida como o resultado da interação de complexos fatores ambientais” (Garcia, 2019). É por este motivo que, para a identificação remota e em massa das consequências destes movimentos, é necessário o uso de diversas técnicas desenvolvidas ao longo do tempo, que possibilitam o reconhecimento de alguns dos fatores ambientais envolvidos no processo de recuperação da área estudada.

O plano inicial deste projeto propõe estabelecer uma metodologia para identificação e caracterização de cicatrizes de escorregamentos em encostas naturais localizadas na Serra do Mar densamente vegetadas, com o uso de informações extraídas a partir de imagens escaneadas da área de interesse feitas por satélites usados no contexto do sensoriamento remoto.

2. Objetivos

Os objetivos do projeto são:

- Identificação e caracterização de cicatrizes de escorregamentos em encostas naturais localizadas na Serra do Mar densamente vegetadas.
- Estudar a capacidade que os satélites têm de identificar esses escorregamentos quando capturam imagens dessas áreas.
- Analisar a influência de cada um de 16 índices de vegetação e de variáveis de classificação do terreno na identificação desses escorregamentos, assim como a relação entre eles.

3. Descrição do estudo

O estudo foi realizado a partir de uma imagem da região da Serra do Mar feita pelo satélite Landsat 8 (acessada pelo Google Earth em 15/01/2019), compreendendo grande parte do município de Caraguatatuba, SP, sendo selecionadas três áreas de

interesse para o estudo, uma de Escorregamento e as outras duas áreas de Não-escorregamento. Cada pixel foi considerado como uma unidade amostral.

Em uma das áreas, denominada Escorregamento, foram selecionados apenas os pixels pertencentes às regiões em que se sabia que algum escorregamento havia ocorrido. Já nas outras duas, todos os pixels foram incluídos no estudo, sendo, neste caso, conhecida a informação de que não ocorreu nenhum escorregamento, fazendo com que todos os pixels nessas duas áreas fossem incluídos no grupo de unidades amostrais classificadas como Não-Escorregamento.

Os valores de espectros de luz obtidos pelas imagens foram utilizados em cada pixel para calcular 16 índices de vegetação que foram utilizados como variáveis de interesse neste estudo.

Depois foram adicionadas 4 variáveis que classificam o terreno de cada pixel, para tentar diferenciar melhor os pontos classificados como Escorregamento.

4. Descrição das variáveis

No estudo temos dois tipos de variáveis.

Variável de grupo:

- Grupo: Escorregamento, Não Escorregamento. Classifica se houve escorregamento a partir da década de 60 no pixel compreendido pela área em questão.

Variáveis advindas de técnicas de geoprocessamento:

Foram escolhidos 16 índices de vegetação pelo pesquisador, sendo que mais detalhes sobre estes podem ser encontrados em Garcia (2019).

- ARVI (*Atmospherically Resistant Vegetation Index*)
Esse índice foi elaborado a partir do entendimento que os efeitos da atmosfera são mais significativos na banda R que na banda NIR, dessa forma

a diferença de radiância entre as bandas do vermelho (R) e do azul (B) cria uma nova banda RB que minimiza os efeitos de espalhamento causados por aerossóis na banda do vermelho.

- AVI (*Agricultural Vegetation Index*)
- Mensuração da vegetação verde/sadia em fase de crescimento. CIgreen (*Chlorophyll Index Green*)

Avaliação da relação entre o conteúdo de clorofila nas folhas e a resposta espectral em diferentes bandas, no caso do CIgreen utiliza-se a banda verde.

- IAF (Índice de área foliar ou LAI - *Leaf Area Index*)

O Índice de Área Foliar (IAF) é definido pela razão entre a área foliar de toda a vegetação por unidade de área utilizado por essa vegetação. Normalmente utilizado para caracterizar o crescimento ou vitalidade da cobertura vegetal.

- CVI (*Chlorophyll Vegetation Index*)

O índice CVI foi desenvolvido na tentativa de obter um índice de vegetação capaz de reunir a informação espectral da banda verde com sua sensibilidade real, para o conteúdo de clorofila e que fosse insensível a variação do IAF (Índice de Área Foliar).

- EVI (*Enhanced Vegetation Index*) Foi elaborado como um mecanismo de retorno que corrige simultaneamente os efeitos do solo e da atmosfera, uma vez que após analisados diversos tipos de solo e índices de vegetação melhorados, notou-se que a interação entre o solo e a atmosfera era inversamente proporcional. MSI (*Moisture Stress Index*)

MSI é um índice que apresenta grande correlação com o conteúdo de água das folhas e tenta relacionar as variações de reflectância com medidas fisiológicas significantes de estresse da folha.

- MVI5 e MVI7 (*Moisture Vegetation Index 5/7*)

Apresentam alta correlação com as medidas estruturais da floresta que foram coletadas em campo pelos pesquisadores e também são índices eficientes para estudos de florestas densas e úmidas.

- NDVI (*Normalized Difference Vegetation Index*)

O índice NDVI (*Normalized Difference Vegetation Index*) é o índice de vegetação mais amplamente utilizado. Muito utilizado em pesquisas de avaliação da vegetação global e regional, também demonstrou estar relacionado com a fotossíntese do dossel e não apenas com sua estrutura e IAF (Índice de área foliar).

- NPCI (*Normalized Pigment Chlorophyll Index*)

Foi elaborado na tentativa de estimar a taxa total de pigmentos de clorofila presente nas folhas, e está relacionado com a condição fisiológica das plantas.

- PSRI (*Plant Senescing Reflectance Index*)

Pode ser utilizado como medida quantitativa da senescência das plantas e amadurecimento das frutas.

- SAVI (*Soil-Adjusted Vegetation Index*)

A originalidade deste índice está em estabelecer um único modelo que permite descrever adequadamente o sistema solo-vegetação.

- MSAVI (*Modified Soil-Adjusted Vegetation Index*)

Para minimizar os efeitos do solo no SAVI, foi proposto o MSAVI (*Modified Soil Adjusted Vegetation Index*), substituindo um fator de ajuste por uma variável, o que reduz a influência do solo exposto.

- SR (*Simple Ratio*)

O índice SR é amplamente utilizado para o monitoramento e estimação de biomassa, especificamente em locais com grande densidade de cobertura vegetal, uma vez que esse índice é muito sensível à vegetação e possui boa correlação com a biomassa.

- VLg (*Visible Green Index*)

É um bom indicador do conteúdo de clorofila em vegetação com tons variando de amarelo até verde escuro.

Além desses índices, foram adicionadas mais 4 variáveis ao estudo, as quais caracterizam o terreno do pixel da imagem:

- Declividade, em graus
- Aspecto, em graus

- Curvatura Horizontal, adimensional
- Curvatura Vertical, adimensional

5. Análise descritiva

Nesta seção são apresentados resultados da análise descritiva dos dados. Detalhes sobre medidas resumo calculadas (como média, desvio padrão, quantis, coeficientes de correlação), bem como gráficos construídos (*box plot*), podem ser encontrados em Bussab e Moretim (2017). No Apêndice A estão apresentadas tabelas e, no Apêndice B, figuras resultantes da análise descritiva dos dados.

O tamanho amostral considerado no estudo é de 46.432 pixels, sendo 14.756 classificados no Grupo Escorregamento e 31.676 no Não-Escorregamento.

Na Tabela A.1 estão apresentadas estatísticas descritivas para as 20 variáveis explicativas de acordo com o grupo Escorregamento. Como pode ser visto, a escala das variáveis é bastante heterogênea, sendo os desvios padrões do índice AVI e da variável aspecto extremamente altos em relação às demais variáveis. Neste caso, para garantir que todas as variáveis contribuam de forma comparativa na análise e que estas possam ser analisadas conjuntamente, é recomendada a padronização das variáveis.

Na Tabela A.2 estão representadas as estatísticas descritivas para o grupo Não Escorregamento. Podemos observar em comparação com o grupo Escorregamento que apesar das médias serem parecidas, os desvios padrões deste grupo são sempre menores em comparação ao primeiro, mesmo tendo este mais que o dobro de observações que o grupo Escorregamento.

Na Figura B.1 é apresentado o gráfico com os valores dos coeficientes de correlação de Pearson entre as variáveis do grupo Escorregamento. Podemos verificar que algumas variáveis como AVI, EVI e PSRI estão pouco correlacionadas com as outras variáveis, porém há grupos de variáveis como IAF, MSAVI e NDVI, com alta correlação entre si. A variável MSI está negativamente correlacionada com todas as outras variáveis (com valores de correlação de moderada a alta), exceto com PSRI, AVI e EVI (com correlações muito próximas de zero). Também podemos observar nesse gráfico que as quatro variáveis de terreno (declividade, aspecto, curvatura horizontal e

curvatura vertical) são pouco correlacionadas com os índices, e entre elas apenas a curvatura horizontal e curvatura vertical parecem ter uma correlação significativa entre si.

Em comparação, no gráfico de correlação para o grupo Não Escorregamento (Figura B.2), podemos observar que os índices parecem ser menos correlacionados entre si, principalmente para a variável AVI, do que no grupo Escorregamento, com exceção da variável SR que apresenta uma correlação maior com as variáveis IAF, MSAVI, NDVI e SAVI neste grupo. Com relação as variáveis de terreno elas parecem ser um pouco mais correlacionadas com os índices nesse grupo.

Considerando as variáveis resposta padronizadas, foram construídos *box plots* separando-se os grupos Não-Escorregamento (0) e Escorregamento (1). É possível observar na Figura B.3 a presença de muitas observações atípicas (*outliers*) segundo o critério adotado no *box plot*. As centralidades dos dados (medianas) parecem estar próximas, independentemente da variável e do grupo. Já a variabilidade central dos dados (entre os quartis Q1 e Q3) mostra-se maior no grupo 1 (Escorregamento) relativamente ao grupo 0 (Não Escorregamento), para todas as variáveis.

Com a padronização das variáveis, podemos observar na Figura B.4, pelo dendograma do *heatmap* dos índices padronizados do grupo Escorregamento, que se pode dividir essas variáveis em 3 grupos. O primeiro sendo composto pelas variáveis CVI, SR, Clgreen, NPCI, AVI e EVI, o segundo grupo composto pelas variáveis MSAVI, SAVI, IAF, NDVI, ARVI, MVI5 E MVI7 e Aspecto e o terceiro grupo composto pela curvatura vertical, curvatura horizontal, declividade, PSRI e MSI, que parece ser o grupo em que as variáveis são mais distantes entre si.

Por esse mesmo gráfico (Figura B.4), considerando o dendograma das unidades amostrais, podemos classificar os pixels (unidades amostrais) do grupo Escorregamento em 4 ou 5 grupos, dependendo do nível de corte escolhido. De maneira geral, nota-se que o primeiro grupo de pixels, cujo agrupamento está definido na parte superior da figura, caracteriza-se pela alta correlação entre as variáveis PSRI e

MSI, bem como AVI e EVI. Já o último agrupamento de pixels, definido na parte inferior da figura, mostra as mais baixas correlações entre estas mesmas variáveis.

Para descrever melhor o padrão de correlação entre os índices de vegetação no grupo Escorregamento, foi realizada uma Análise de Componentes Principais, buscando revelar como cada variável contribui para a variância total dos dados. Detalhes desta técnica podem ser encontrados em (Jonhson e Wichern, 1992; Everitt, 2004). Considerando a diferença nas escalas desses índices, a análise foi realizada com as variáveis padronizadas, o que corresponde a obter os componentes principais diretamente da matriz de correlação entre as variáveis em questão.

Na Análise de Componentes Principais, podemos ver pela Figura B.5 que os cinco primeiros componentes explicam uma porcentagem considerável da variância total dos dados padronizados (87%, segundo a Tabela A.3).

Observando os coeficientes (cargas) das variáveis respostas dessas cinco componentes (Tabela A.4), no primeiro componente, os índices MSI e PSRI receberam valores negativos, sendo que os demais tiveram influência por meio de valores positivos. Neste componente, as variáveis de terreno tiveram os seus pesos praticamente nulos. Já o segundo componente recebe maior influência das variáveis AVI, EVI e PSRI.

As variáveis de terreno acabam tendo maior influência nos 3 próximos componentes, sendo que o terceiro componente tem um peso grande nas variáveis curvatura horizontal e curvatura vertical. Já nos componentes 4 e 5 as variáveis que possuem o maior peso são, respectivamente, declividade e aspecto.

A Figura B.6 mostra a dispersão dos pixels do grupo Escorregamento de acordo com seus escores nos dois primeiros componentes principais. Já na Figura B.7, é apresentado o *biplo*t, que representa simultaneamente as unidades amostrais e os coeficientes das variáveis nestes dois primeiros componentes. É possível visualizar um padrão homogêneo na distribuição dos pixels com a maioria estando dentro do intervalo [-10,5] para o primeiro componente e entre [-5,5] para o segundo componente. Porém é interessante notar que, apesar de apenas duas variáveis terem peso negativo no

primeiro componente, grande parte das observações tem seus escores negativos para este, além de que há vários *outliers* para o lado esquerdo do eixo das abscissas.

6. Análise de regressão

Para a classificação dos pixels como pertencentes ou não a áreas de escorregamentos, os 16 índices de vegetação foram utilizados, juntamente com as quatro variáveis de terreno, no contexto de ajustes de regressão com uso de modelos de regressão logística. Detalhes sobre tais ajustes podem ser consultados em Paula, G.A. (2013).

No ajuste de modelos de regressão para finalidade de classificação é importante estabelecer quais serão os critérios que permitirão a escolha de um modelo final. Neste caso, além da significância dos efeitos das variáveis ajustadas pelo modelo e dos gráficos de diagnóstico, foi utilizada uma base de dados externa para avaliação da eficiência da predição. Essa base externa corresponde à imagem feita pelo satélite de uma parte do litoral de Caraguatatuba, SP, da qual não se possui informações a respeito de todos os pixels com relação ao pertencimento destes a áreas de escorregamento. Todavia, foi possível extrair valores de todos os pixels para as 20 variáveis explicativas (16 índices e 4 características de terreno). Isso torna viável a predição de observações de acordo com os modelos utilizados e a estimação da probabilidade dos pixels pertencentes à imagem estarem contidos em zonas de escorregamento, permitindo, portanto, suas classificações.

Os resultados para cada modelo ajustado foram observados e sua adequabilidade aos objetivos inicialmente estabelecidos foram avaliados com auxílio e supervisão dos pesquisadores. Nesse sentido, modelos de regressão foram pesquisados sendo que o seguinte modelo c-log-log (log-log complementar) foi o modelo final selecionado para a classificação.

É importante ressaltar duas características a respeito do modelo ajustado: que o índice de vegetação SAVI é combinação linear de outros índices dentro da base de dados, por isso foi descartado no ajuste, o que resultou no uso de 19 variáveis das 20

propostas inicialmente; que uma nova coluna foi adicionada à matriz de planejamento do modelo resultante da inclusão não somente do termo linear ligado à declividade, mas também do termo quadrático. A inclusão desse termo no ajuste, além de ter se mostrado estatisticamente significativa, foi decisivo para que o modelo tivesse eficácia em identificar diferenças nas probabilidades de escorregamentos para áreas com características terrenas geologicamente distintas, como planícies e costas.

$$\log(-\log(1 - \pi_i)) = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots + \beta_{19} x_{19,i} + \beta_{20} x_{19,i}^2 \quad (6.1)$$

em que, π_i é a probabilidade do i -ésimo pixel pertencer a uma área de escorregamento, $x_{j,i}$ é o valor da variável j ($j=1,2,\dots,19$) para o i -ésimo pixel, α é o valor esperado do $\log(-\log(1 - \pi_i))$ para pixels que têm todos os valores nulos para os índices de vegetação e variáveis de terreno, β_j ($j=1, 2, \dots, 19$) é o acréscimo esperado em α para cada mudança de uma unidade de medida da j -ésima variável explicativa considerada inicialmente no estudo (com exceção da declividade e do índice SAVI), fixando-se todas as demais variáveis, sendo β_{19} e β_{20} associados à variável declividade.

Como pode ser observado na Tabela A.5, alguns dos efeitos mostraram-se não significantes (valores- p maiores que 5%). Apesar disso, a inclusão das variáveis associadas a estes efeitos e o uso da ligação para π_i vista em 6.1 foram determinantes para uma classificação da base de dados externa que fosse satisfatória, principalmente para a correta inclusão da zona do oceano no grupo de pixels classificados como pertencentes ao grupo de não-escorregamento.

Ainda, na Tabela A.5, pode-se observar que os valores de alguns coeficientes se mostram em escala muito diferente dos demais. Isso está ligado tanto à diferença de escalas das próprias variáveis envolvidas no estudo, como à alta correlação entre os índices de vegetação, o que pode dificultar a interpretação dos coeficientes que cada variável tem no modelo.

O modelo de classificação adotado foi validado também por meio da Curva ROC. Neste caso, utilizando 0,4 como um ponto de corte arbitrariamente escolhido pela análise da curva ROC (Figura B.8), temos que a sensibilidade e a especificidade do modelo para a base de dados usada para o ajuste, assim como os gráficos de diagnóstico (Figura B.9 e B.10), parecem indicar que o modelo foi bem ajustado.

A classificação da imagem completa de parte do litoral de Caraguatatuba pode ser vista na Figura B.11, onde cada pixel foi pintado de acordo com as cores posicionadas na escala ao lado direito da imagem conforme as probabilidades ajustadas pelo modelo. Os resultados indicam que o modelo escolhido viabilizou uma maior “seletividade” ao classificar um pixel como pertencente a uma área de escorregamento, isolando de uma maneira mais próxima da realidade essas zonas.

É importante ressaltar que, apesar do uso de resultados advindos da classificação de uma base de dados externa (imagem do litoral de Caraguatatuba), não se possui nenhuma informação a respeito da maior parte dos pixels quanto ao pertencimento ou não das unidades amostrais a áreas de escorregamento e, por isso, informações relevantes presentes nessas áreas não são captadas pelo modelo por não estarem incluídas na base de dados do ajuste. Isso sugere que mais análises devem ser feitas futuramente com a inclusão de novos dados.

7. Conclusão

O principal objetivo do presente estudo é encontrar uma metodologia para a identificação de cicatrizes em encostas naturais. Para esse fim, apesar da seleção de variáveis desempenhar um papel importante, nesta aplicação, de maneira geral, teve pouco impacto em melhorar a eficiência predita do modelo. Deste modo, entende-se que a maior parte dos índices de vegetação, assim como as variáveis de terreno, mostraram-se importantes na classificação. Contudo, sob uma análise supervisionada, destaca-se a inclusão do termo quadrático associado ao efeito da variável declividade, o qual teve o maior impacto na comparação entre modelos.

Apoiando-se nas tentativas de ajustes de diversos modelos no processo de modelagem, pôde-se deduzir que não há informação necessária a respeito da distribuição das variáveis explicativas em áreas geográficas como o Oceano ou mais próximas da praia para que os modelos detectem apropriadamente a presença de pixels contidos em zonas de escorregamento, limitando assim a flexibilidade dos ajustes quanto às funções de ligação e seleção de variáveis.

Os resultados obtidos com o modelo final escolhido (modelo de regressão com função de ligação log-log complementar) pareceram satisfatórios levando em conta a base de dados fornecida, pois conseguiram isolar áreas que possivelmente pertencem a zonas de escorregamento de uma forma convincente, ao mesmo tempo que indicou um ajuste adequado à base de dados.

APÊNDICE A

Tabelas

Tabela A.1 Medidas descritivas dos índices de vegetação do Grupo Escorregamento

	Média	Mediana	Desvio padrão	Mínimo	Q1	Q3	Máximo
ARVI	0,01	0,01	0,00	0,00	0,01	0,01	0,01
AVI	5130	5140	1600	385	3900	6300	10340
Clgreen	6,93	7,17	1,48	0,20	6,22	7,93	11,45
CVI	5,02	5,14	0,82	1,29	4,61	5,56	9,20
EVI	0,57	0,56	0,23	-0,08	0,40	0,73	2,70
IAF	1,07	1,11	0,18	-1,15	1,05	1,15	1,27
MSAVI	0,91	0,92	0,05	0,09	0,91	0,93	0,96
MSI	0,46	0,43	0,11	0,28	0,41	0,46	2,05
MVI5	0,38	0,40	0,08	-0,34	0,37	0,42	0,57
MVI7	0,68	0,70	0,09	-0,24	0,68	0,72	0,79
NDVI	0,84	0,86	0,06	0,05	0,84	0,87	0,91
NPCI	0,75	0,77	0,07	-0,05	0,73	0,80	0,86
PSRI	-0,12	-0,12	0,18	-1,52	-0,22	-0,02	0,56
SAVI	1,26	1,29	0,10	0,07	1,25	1,31	1,37
SR	12,61	13,01	2,83	1,10	11,28	14,4	22,08
VLg	0,22	0,23	0,05	-0,18	0,21	0,25	0,37
Declividade	22,60	22,94	8,63	0,43	16,50	28,71	51,17
Aspecto	217,32	255,69	105,13	0,01	129,54	300,05	359,98
CurvaturaV	0,00	0,00	0,002	-0,02	-0,01	0,00	0,02
CurvaturaH	0,00	0,00	0,003	-0,02	-0,01	0,00	0,02

Tabela A.2 Medidas descritivas dos índices de vegetação do Grupo Não-Escorregamento

	Média	Mediana	Desvio padrão	Mínimo	Q1	Q3	Máximo
ARVI	0,01	0,01	0,00	0,00	0,01	0,01	0,01
AVI	4610	4670	900	760	3990	5200	8370
Cigreen	6,78	6,77	0,53	3,13	6,43	7,12	9,53
CVI	5,09	5,08	0,40	1,74	4,82	5,34	9,207,49
EVI	0,52	0,52	0,12	-0,03	0,43	0,60	1,02
IAF	1,07	1,07	0,03	0,66	1,05	1,09	1,21
MSAVI	0,91	0,92	0,01	0,82	0,91	0,92	0,94
MSI	0,46	0,46	0,02	0,34	0,45	0,48	0,77
MVI5	0,36	0,36	0,02	0,13	0,35	0,37	0,49
MVI7	0,68	0,68	0,01	0,40	0,67	0,69	0,77
NDVI	0,84	0,84	0,01	0,70	0,83	0,85	0,89
NPCI	0,76	0,77	0,03	0,26	0,75	0,79	0,84
PSRI	-0,02	-0,02	0,07	-1,13	-0,07	0,02	0,21
SAVI	1,26	1,26	0,02	1,04	1,26	1,28	1,34
SR	11,92	11,88	0,96	5,57	11,30	12,49	17,69
VLg	0,21	0,21	0,02	0,05	0,19	0,22	0,40
Declividade	13,43	12,43	7,55	0,02	7,53	18,38	45,28
Aspecto	199,72	217,85	103,58	0,00	111,06	290,96	359,97
CurvaturaV	0,00	0,00	0,002	-0,01	0,00	0,00	0,01
CurvaturaH	0,00	0,00	0,002	-0,01	0,00	0,00	0,01

Tabela A.3 Valores das variâncias e porcentagem da variância total acumulada para as primeiras 5 primeiros componentes principais para a matriz de correlação do grupo Escorregamento.

	Comp. 1	Comp. 2	Comp.3	Comp. 4	Comp. 5
Desvio Padrão	3,25143	1,86548	1,19021	1,03389	0,93793
% Var.	0,52859	0,17400	0,07083	0,05345	0,04399
% Acum. da Var.	0,52859	0,70259	0.92737	0,82687	0,87085

Tabela A.4 Coeficientes (cargas) das 5 primeiras componentes para os Índices de vegetação padronizados do grupo Escorregamento.

	Comp. 1	Comp.2	Comp.3	Comp.4	Comp.5
ARVI	-0,28516	0,17546	0,01136	0,04397	0,02318
AVI	0,11371	-0,46990	-0,05039	0,08744	0,01060
Cigreen	0,26704	-0,13967	0,14082	-0,20766	-0,16698
CVI	-0,20416	-0.18625	0,21168	-0,31775	-0,31407
EVI	-0,01705	-0,51211	-0,06318	0,07707	-0,00573
IAF	-0,30130	0,02037	-0,02313	0,06860	-0,02468
MSAVI	-0,28709	0,03718	-0,04477	0,12562	-0,03383
MSI	-0.27769	1-0,10699	0,07443	-0,00186	-0,13153
MVI5	0,28045	0,11004	-0,04847	-0,04348	0,10928
MVI7	-0,28804	0,08989	-0,08410	-0,00755	0,10658
NDVI	-0,30128	0,02081	-0,02309	0,06864	-0,02463
NPCI	-0.21061	-0,35732	-0,07549	0,05745	-0,08632
PSRI	-0,12328	-0,42912	-0,12548	0,12125	-0,10290

SAVI	-0,30130	0,02037	-0,02313	0,06860	-0,02468
SR	0,27637	-0,10804	0,08609	-0,11987	-0,06108
VLg	0,24942	0,09529	-0,12032	0,21178	0,22620
Declividade	-0,00878	0,10737	-0,02925	0,79105	-0,46169
Aspecto	0,00986	0,22406	0,08984	-0,22844	-0,71250
CurvaturaV	-0,00609	0,04948	-0,65046	-0,18265	-0,13776
CurvaturaH	-0,02506	0,03010	-0,65938	-0,14144	-0,12841

Tabela A.5 Estimativas e testes para os coeficientes dos modelos.

	Estimativa	Desvio Padrão	Valor t	Pr(> t)
Intercepto	19038,30	615,16	30,95	< 0,001
ARVI	277,73	66,74	4,16	< 0,001
AVI	$9,87 \times 10^{-4}$	$1,62 \times 10^{-4}$	6,05	< 0,001
Clgreen	-1,34	0,55	-2,43	0,02
CVI	3,42	0,48	7,11	< 0,001
EVI	-2,21	1,23	-1,80	0,071
IAF	$1,47 \times 10^4$	$4,78 \times 10^2$	30,89	< 0,001
MSAVI	$2,05 \times 10^2$	12,47	16,44	< 0,001
MSI	-18,63	7,10	-2,62	0,01
MVI5	11,12	7,39	1,50	0,013
MVI7	-27,75	1,32	-21,00	< 0,001
NDVI	$-41,55 \times 10^4$	$1,33 \times 10^3$	-31,01	< 0,001

NPCI	-29,22	1,95	-14,97	< 0,001
PSRI	-10,78	0,49	-22,15	< 0,001
SR	0,70	0,18	3,83	< 0,001
VLg	51,73	4,69	11,02	< 0,001
Declividade	0,17	0,01	32,10	< 0,001
Declividade²	$-2,39 \times 10^{-3}$	$1,23 \times 10^{-4}$	-19,39	< 0,001
Aspecto	$2,23 \times 10^{-3}$	$1,17 \times 10^{-4}$	19,05	< 0,001
CurvaturaV	77,22	5,35	14,43	< 0,001
CurvaturaH	-4,36	4,46	-0,98	0,33

APÊNDICE B

Figuras

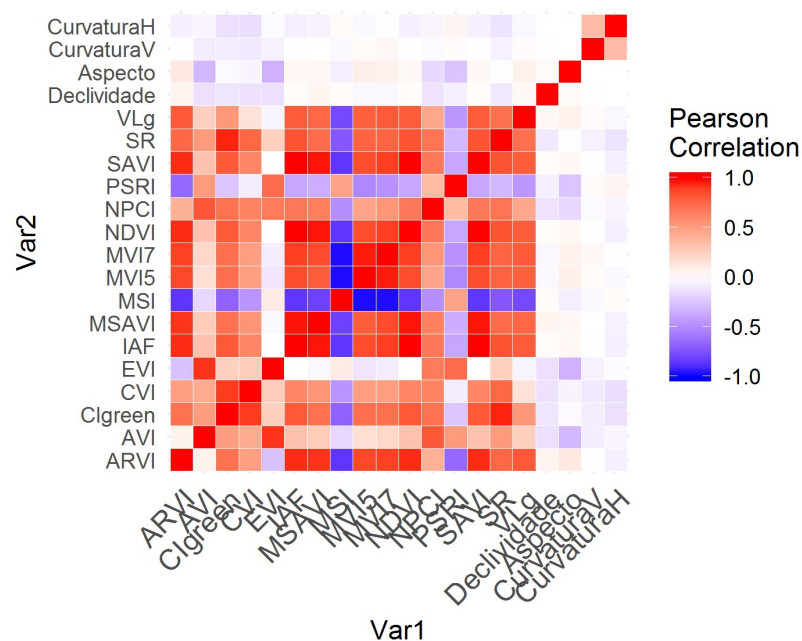


Figura B.1 Representação gráfica da matriz de correlação entre os Índices de vegetação para os pixels observados no grupo Escorregamento

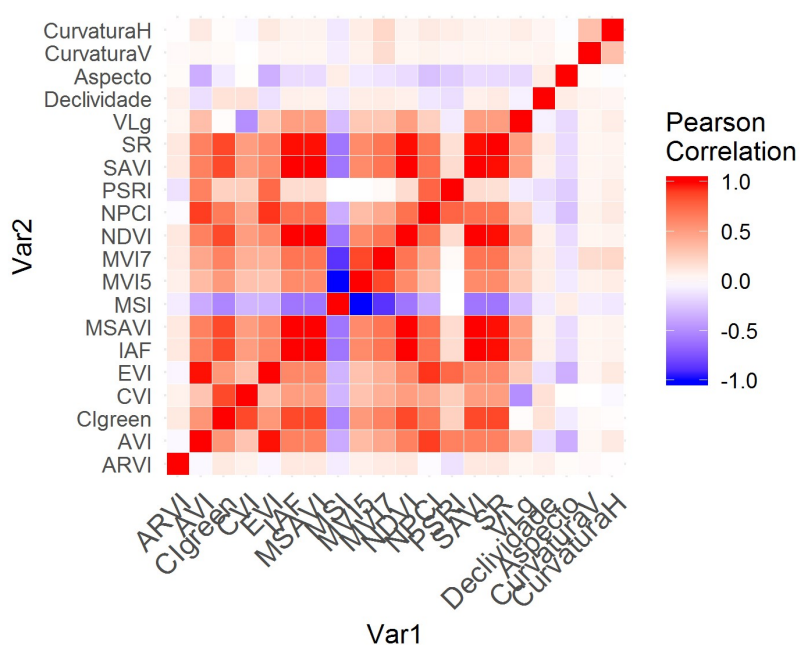


Figura B.2 Representação gráfica da matriz de correlação entre os Índices de vegetação para os pixels observados no grupo Não-Escorregamento

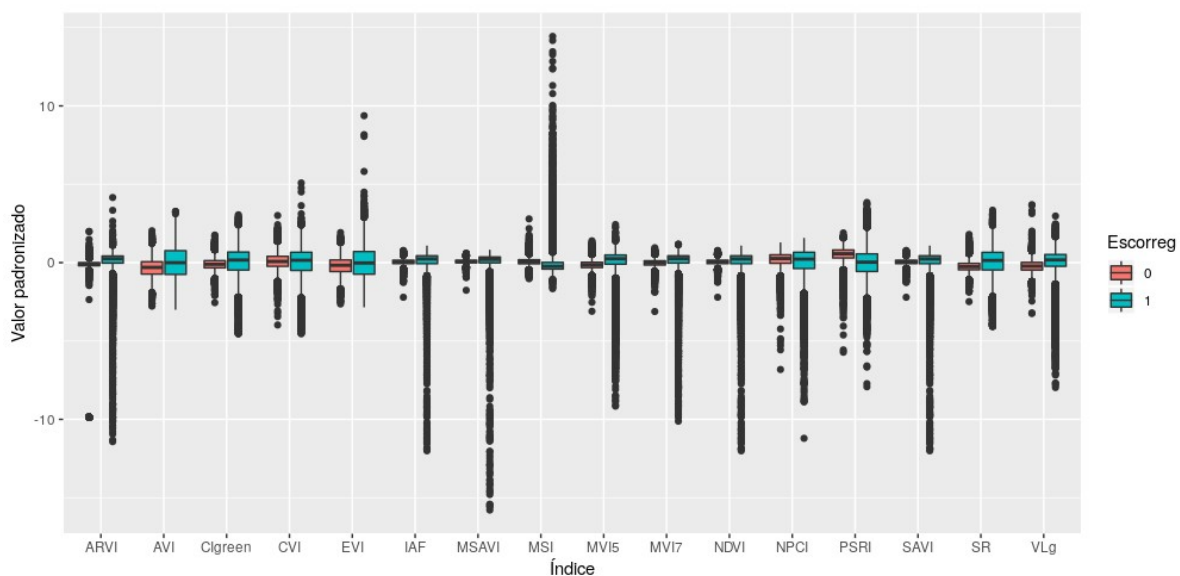


Figura B.3 *Boxplots* das variáveis padronizadas de acordo com os grupos, Não-Escorregamento (0) e Escorregamento (1).

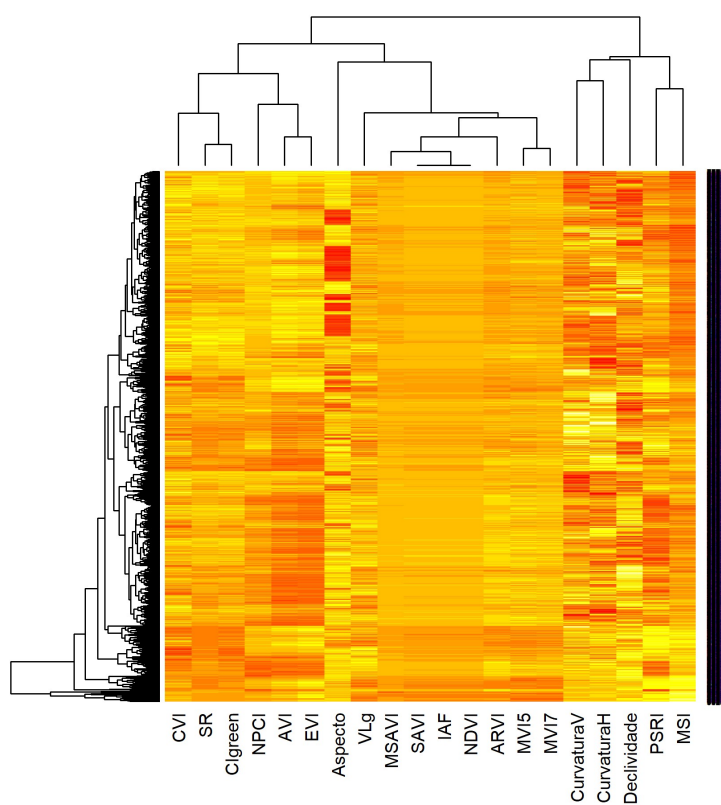


Figura B.4 *Heatmap* para as variáveis padronizadas em relação às observações pertencentes ao grupo Escorregamento.

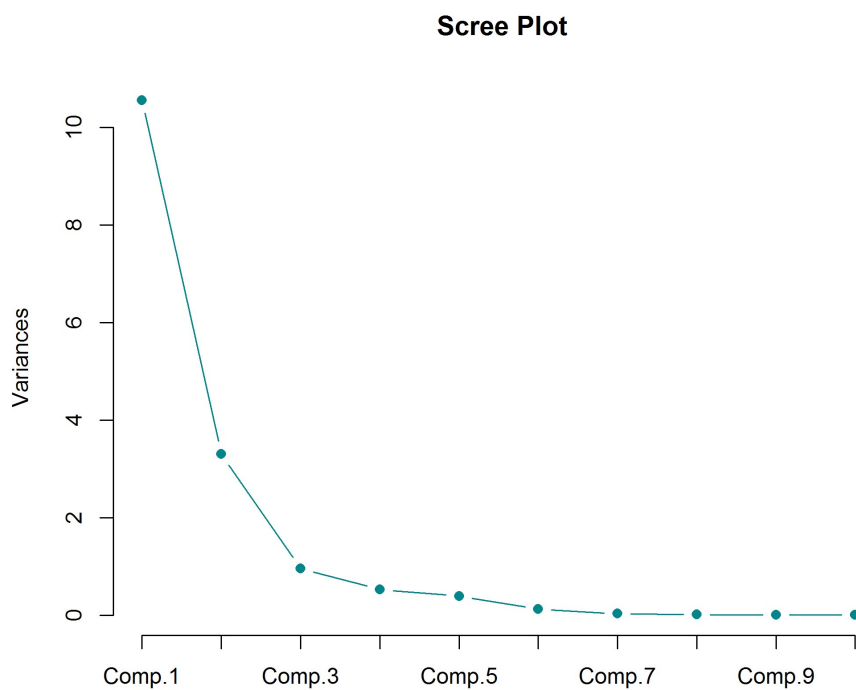


Figura B.5 *Screeplot* para a Análise de Componentes Principais das observações padronizadas pertencentes ao Grupo Escorregamento.

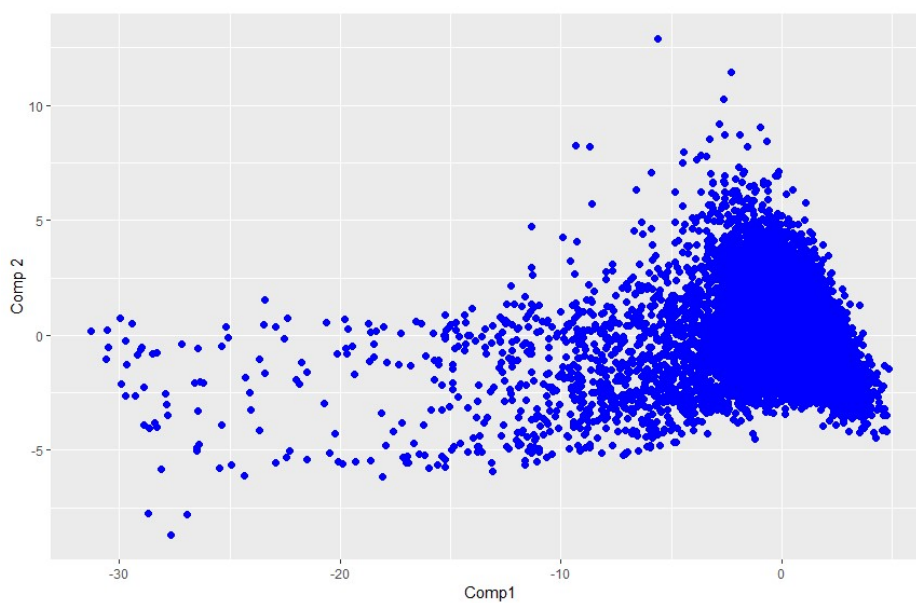


Figura B.6 Gráfico dos escores das unidades amostrais do Grupo Escorregamento para os dois primeiros componentes.

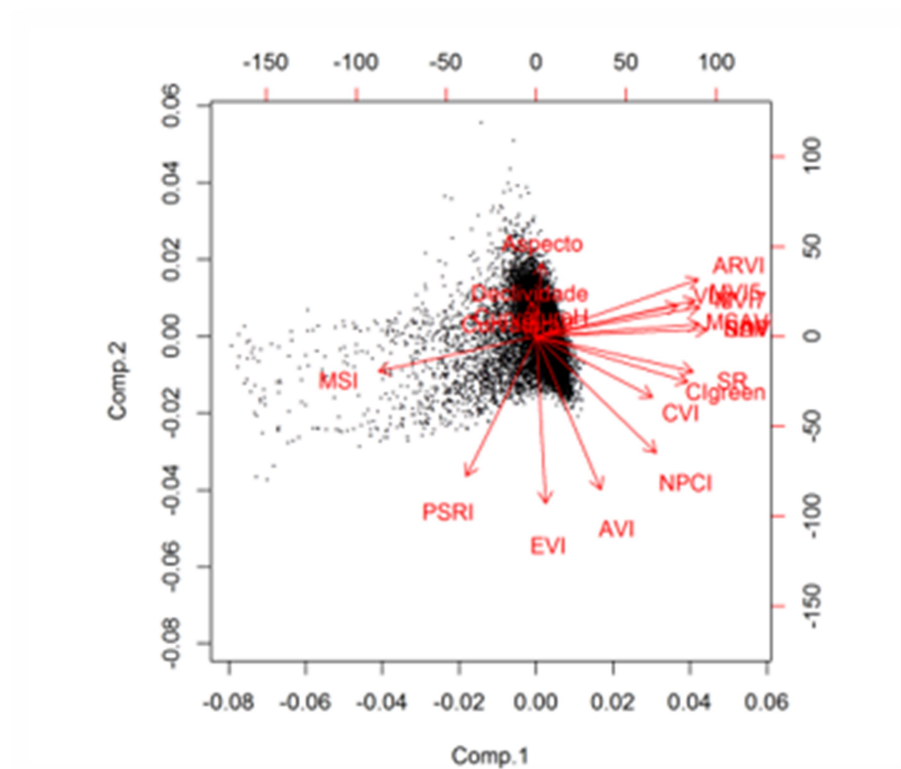


Figura B.7 *Biplot* dos primeiros dois componentes na Análise de Componentes Principais para as variáveis do Grupo Escorregamento

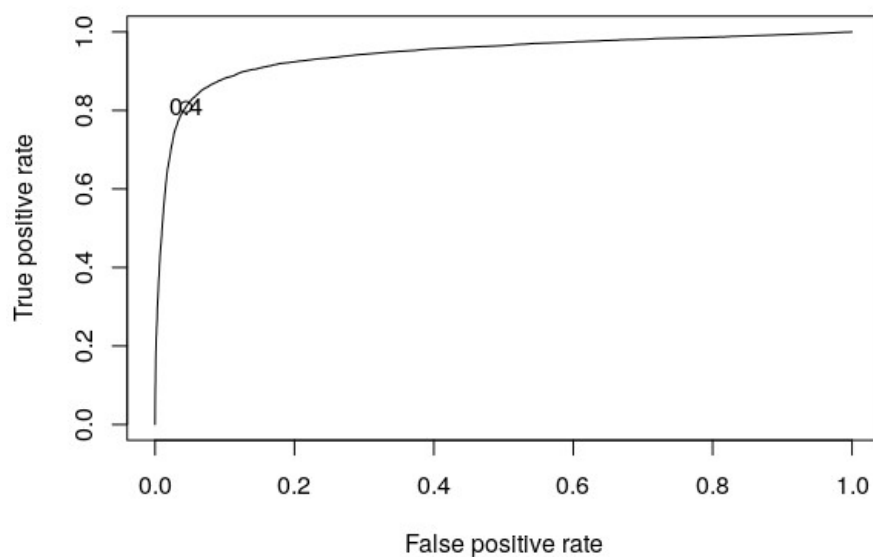


Figura B.8 Curva ROC do modelo final ajustado para a base de dados fornecida.

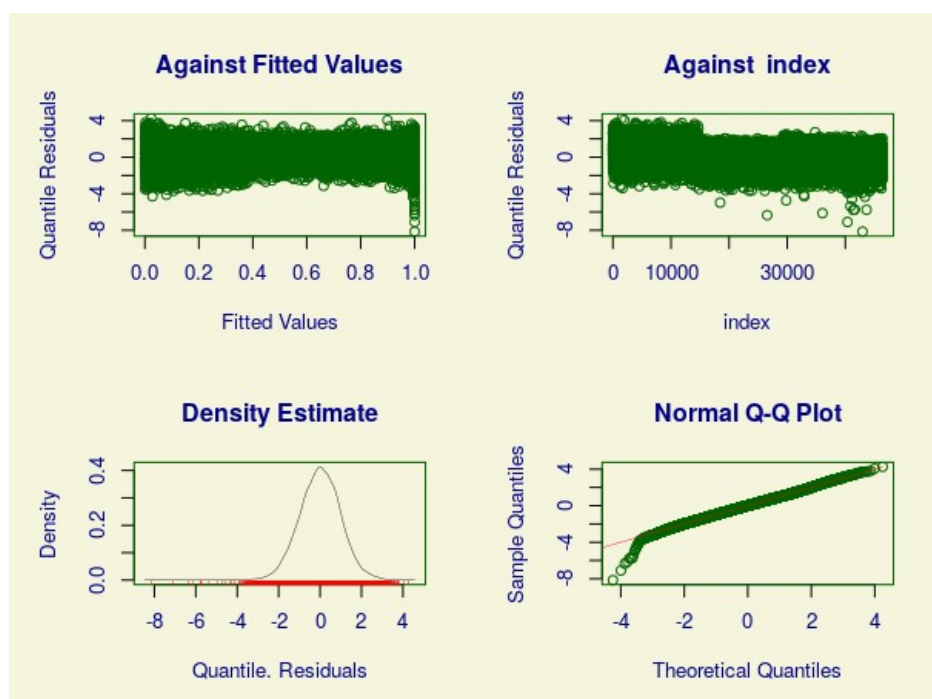


Figura B.9 Gráficos de diagnóstico do modelo final ajustado.

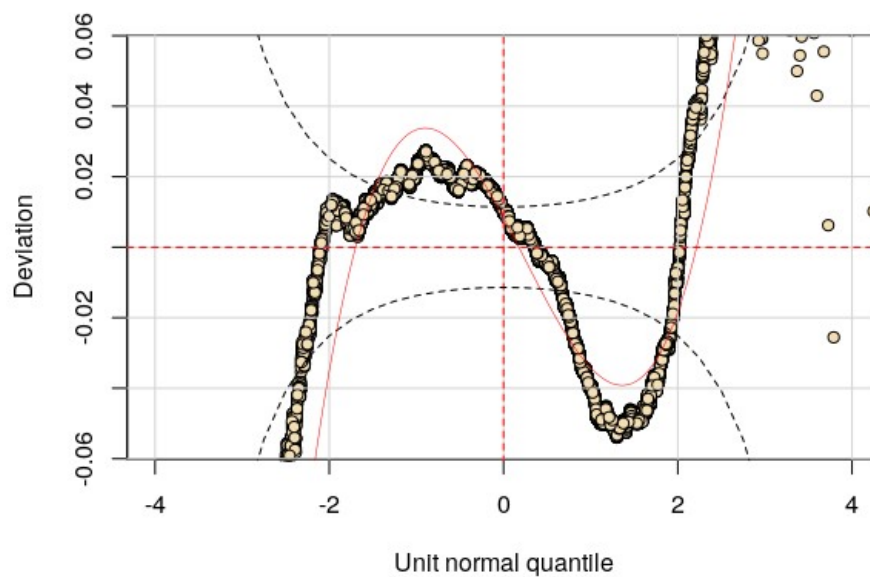


Figura B.10 Gráfico Wormplot utilizado no diagnóstico do modelo final.

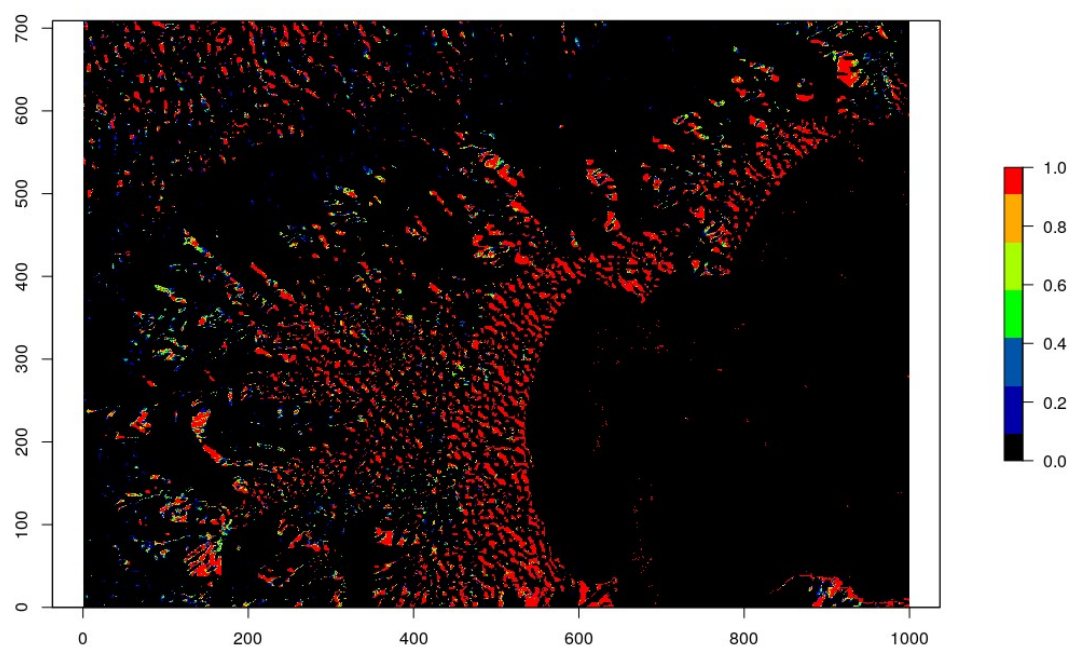


Figura B.11 Imagem com probabilidades ajustadas pelo modelo final para cada pixel dentro do litoral de Caraguatatuba, SP.