Identificação de Áreas Prioritárias para Recuperação Florestal através da rede neural *Self-Organizing Maps* (SOM)

Gustavo Felipe Balué Arcoverde

Análise Espacial de Dados Geográficos (SER 301)

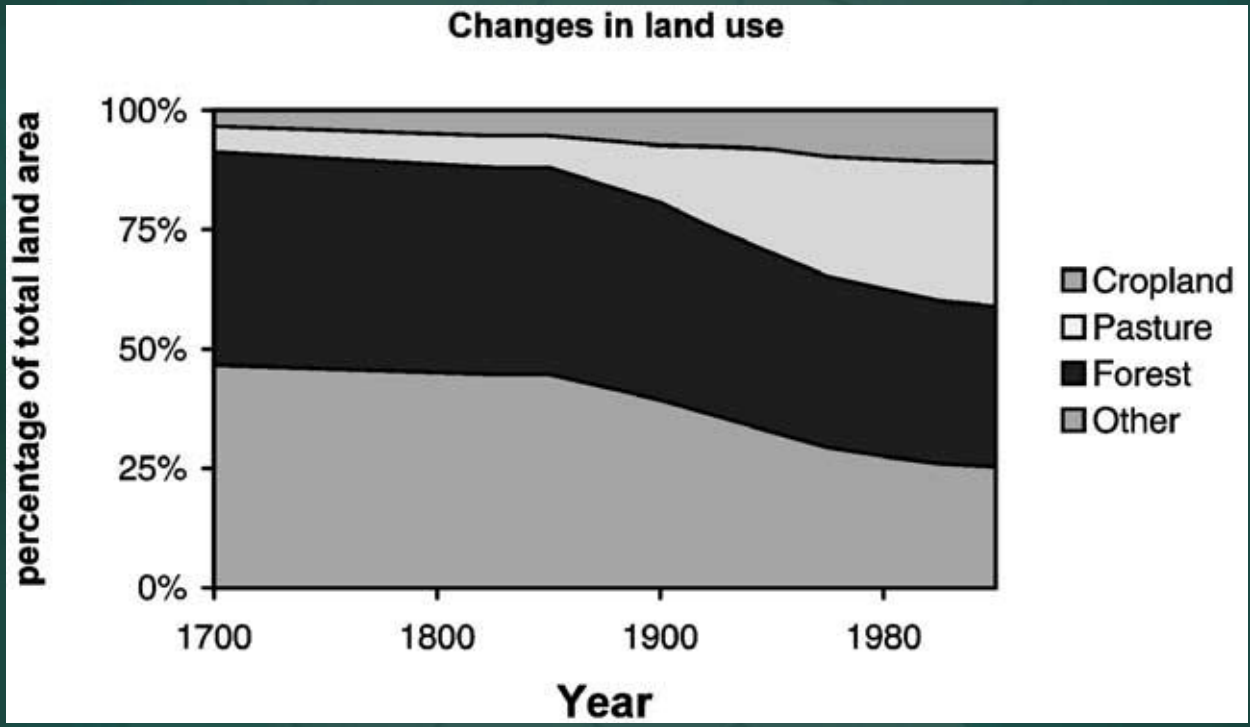
3/2009

Roteiro

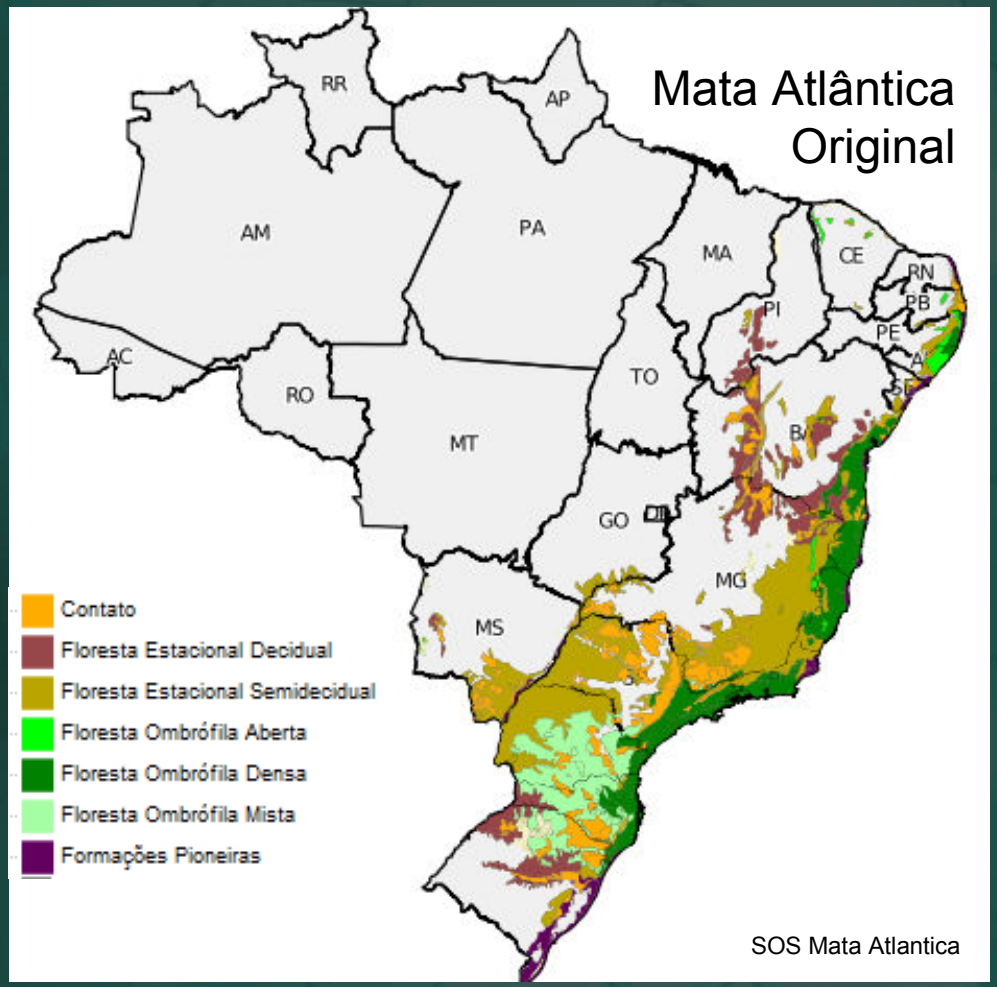
1. Contextualização
    1. 1 Problematização
2. Objetivo
3. Revisão: SOM
4. Metodologia
    4.1 Área de Estudo
    4.2 Preparo e Análise
5. Resultados
6. Considerações finais

1. Contextualização

Mudança de Uso e Cobertura da Terra é a principal alteração do planeta pelo homem, e tem ocorrido sem precedentes nas últimas décadas (Lambin, 2001)



1. Contextualização

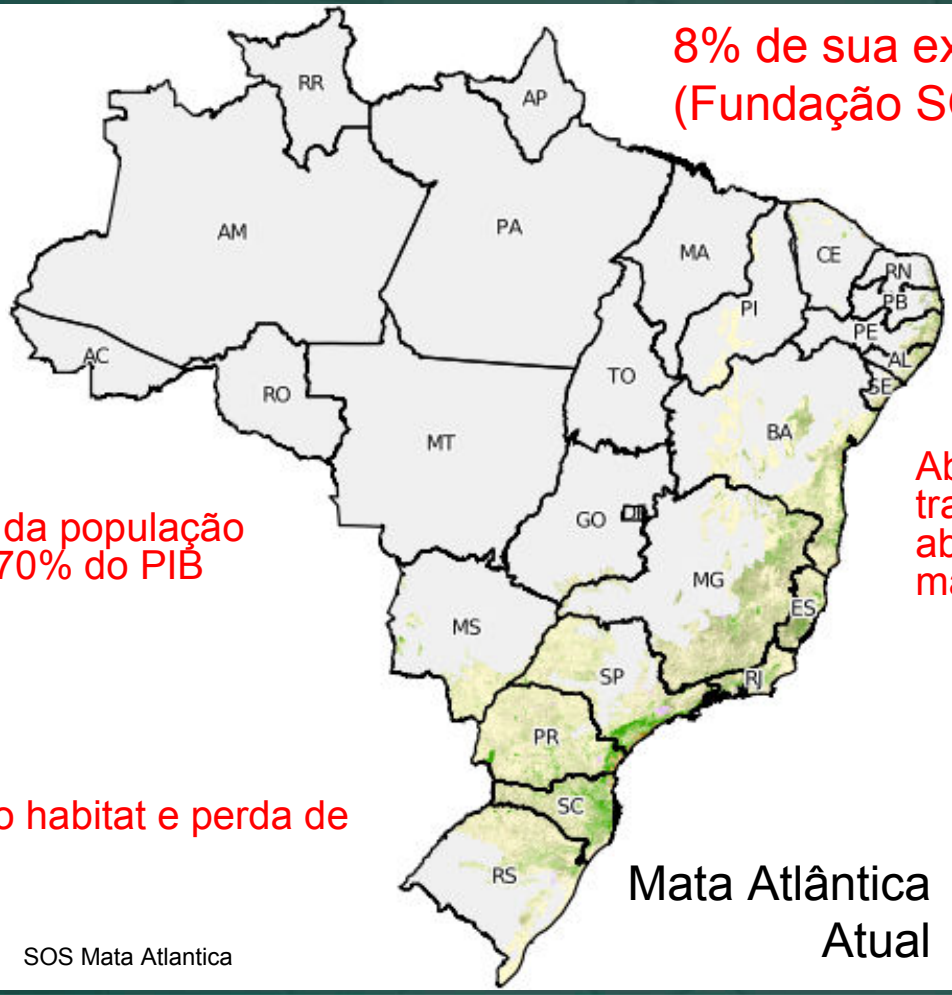


1. Contextualização

Um dos ecossistemas mais ricos do planeta e integra a lista dos 25 biomas de alta diversidade mais ameaçados (Mittermeier et al. 1999)

Abriga mais de 60% da população brasileira e mais de 70% do PIB nacional

Alta fragmentação do habitat e perda de sua biodiversidade



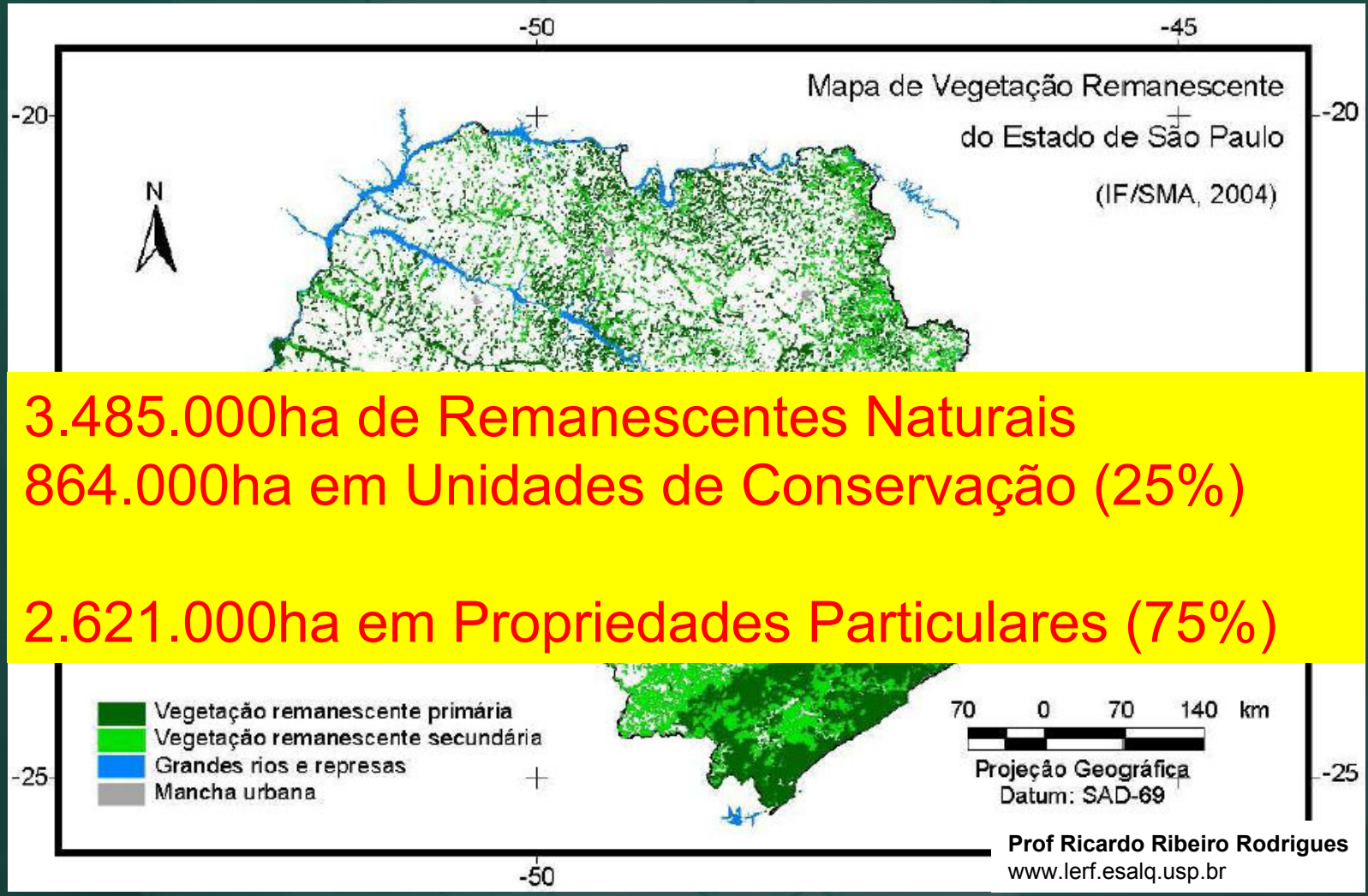
SOS Mata Atlantica

Mata Atlântica Atual

8% de sua extensão original (Fundação SOS Mata Atlântica e INPE)

Abrigo para várias populações tradicionais e garantia de abastecimento de água para mais de 110 milhões de pessoas

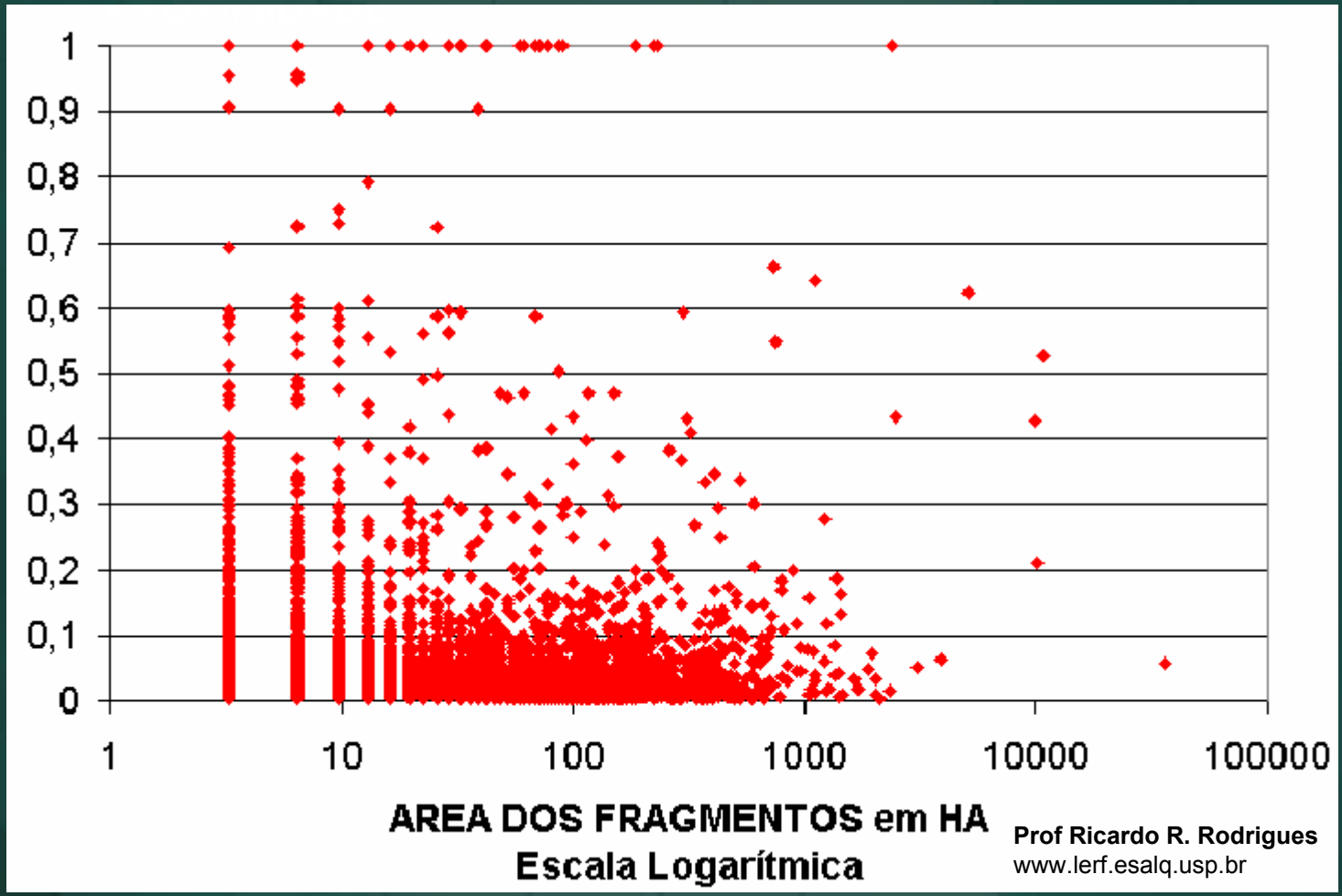
1. Contextualização



3.485.000ha de Remanescentes Naturais
864.000ha em Unidades de Conservação (25%)

2.621.000ha em Propriedades Particulares (75%)

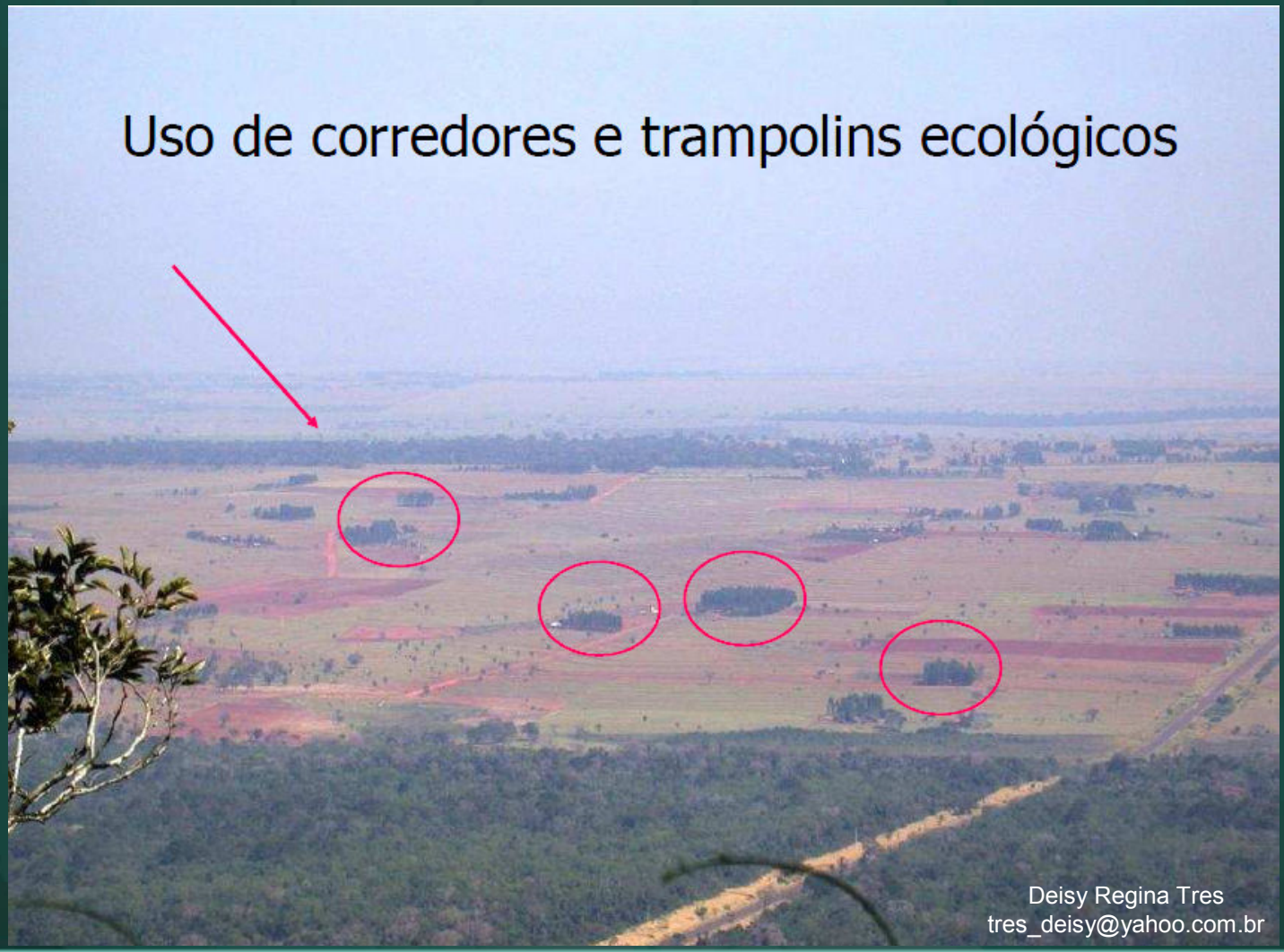
1. Contextualização



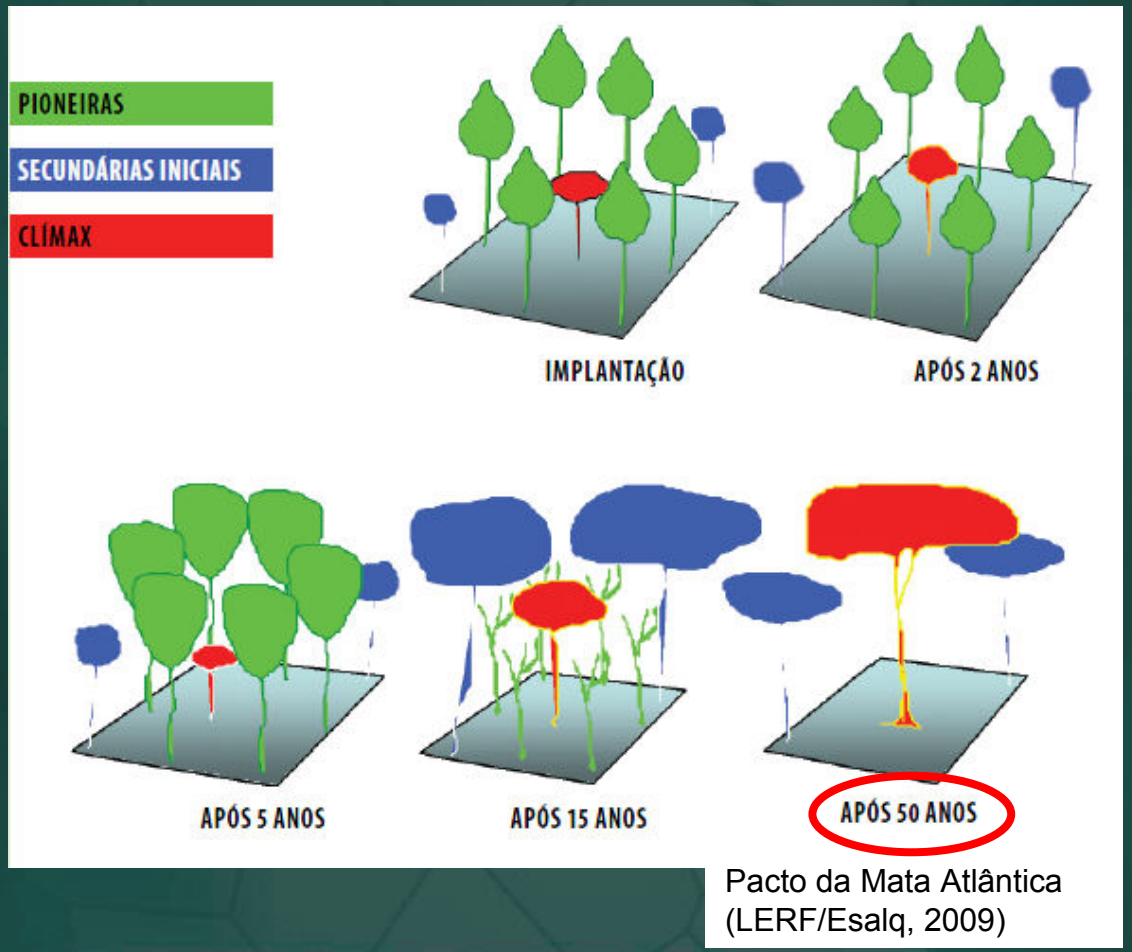
1. Contextualização

- Programas de Diagnóstico / Monitoramento / Recuperação / Prognósticos em SP:

1. Contextualização



1. Contextualização

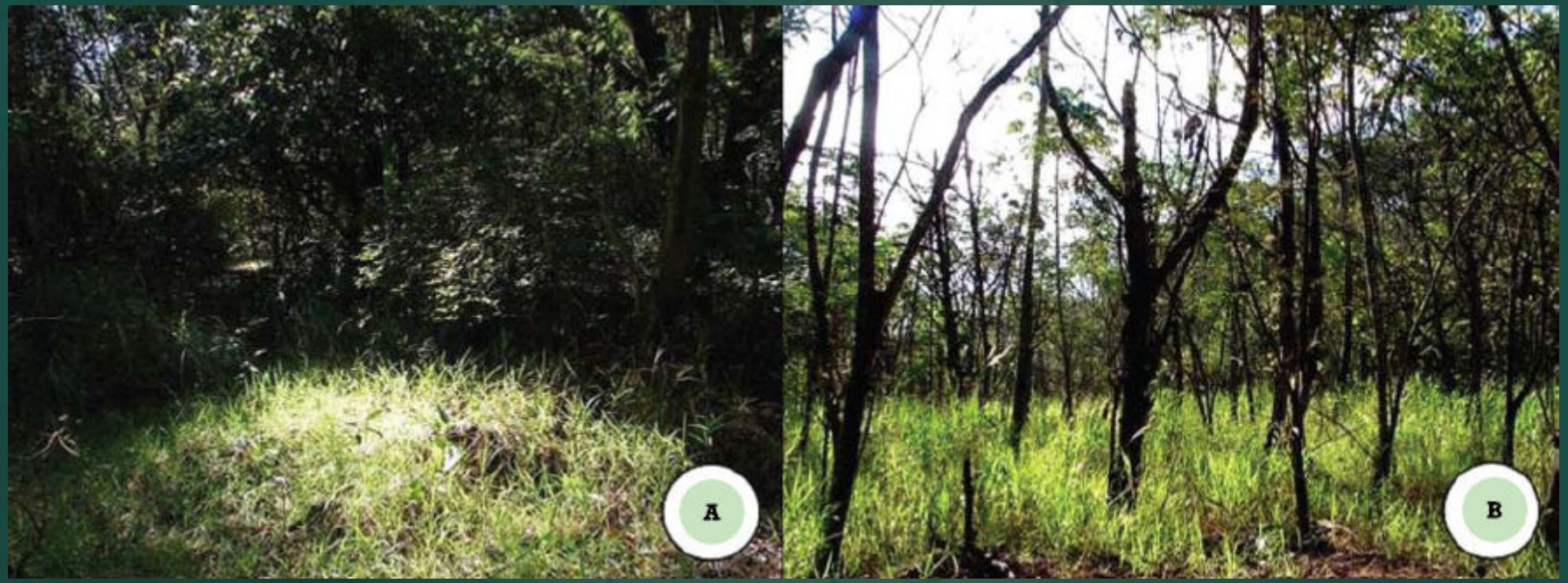


Pacto da Mata Atlântica
(LERF/Esalq, 2009)

1. Contextualização

Figura 4 – Mata ciliar restaurada há cerca de 50 anos

1. Contextualização



Pacto da Mata Atlântica
(LERF/Esalq, 2009)

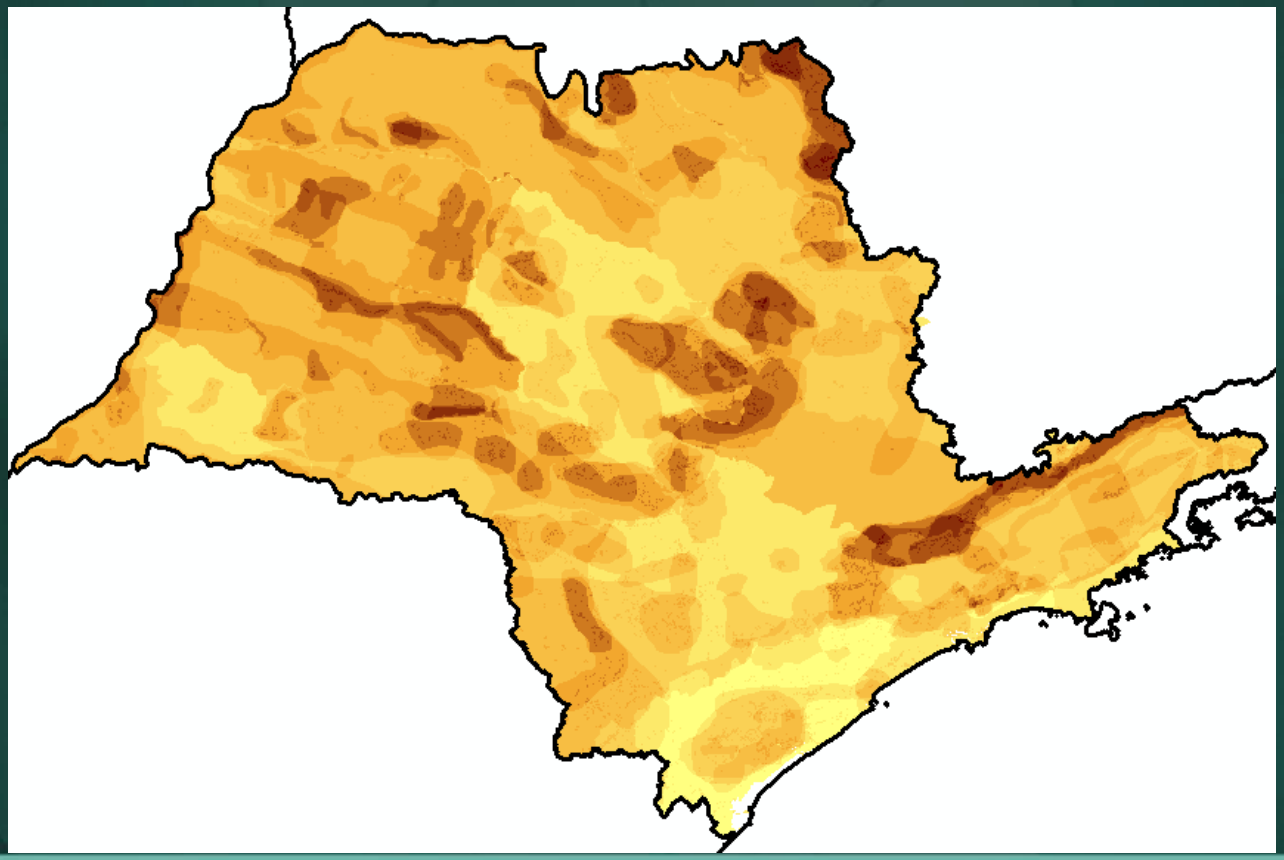
1. Contextualização

1. Contextualização

1. Contextualização

1. Contextualização

Mapa de Conectividade em SP (Biota/Fapesp): Áreas prioritárias para a Conservação Ambiental, auxiliando nas tomadas de decisão e compensação ambiental (Res. SMA 15/2008);



1. Contextualização
1.1 Problematização

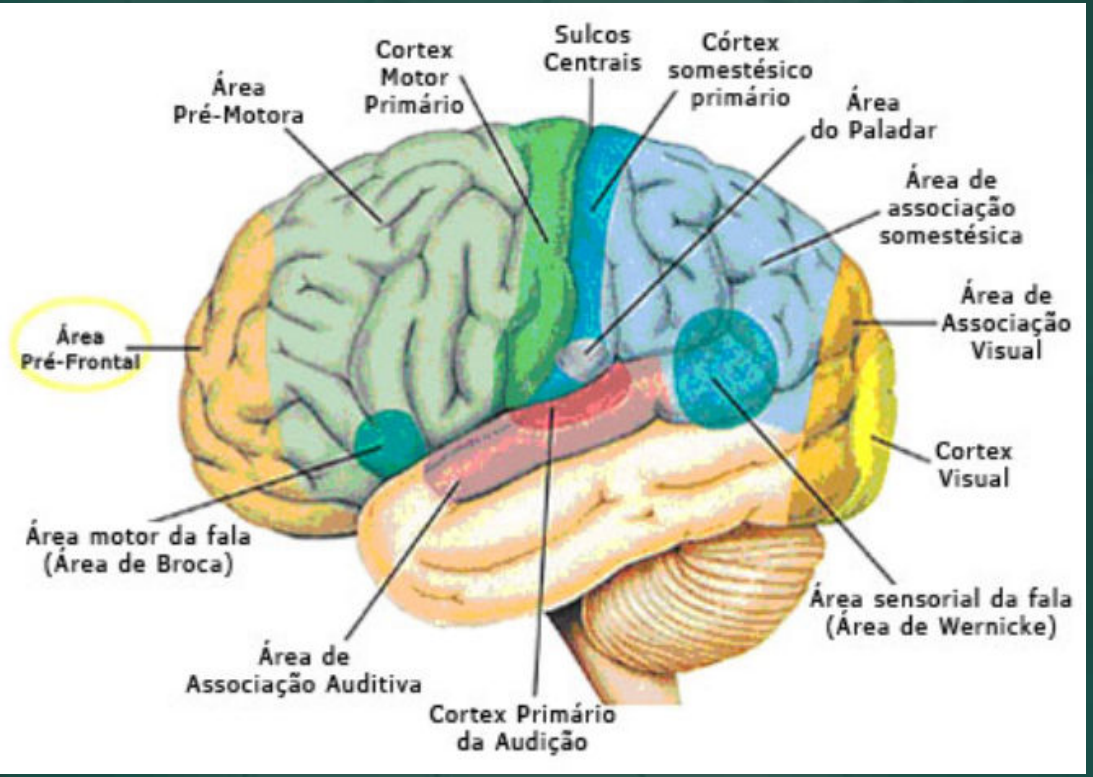
- O Mapa de Conectividade não tem o propósito de apontar áreas para recuperação florestal a nível de APP de curso d' água;
- A localização de áreas de Programas de Recuperação/Restauração Florestal possui muito mais um foco político do que técnico/científico;
- O uso do sensoriamento remoto e geoprocessamento pode dinamizar o processo de monitoramento e diagnóstico de áreas prioritárias para recuperação florestal;
- Classificação de imagens de satélite acaba sendo uma caixa- preta para usuários menos treinados;
- A análise das variáveis físico-espacias deve elencar elementos visuais, possibilitando a identificação de feições propícias para a recuperação florestal. O uso do SOM poderia ser uma alternativa para esta análise.

2.Objetivo

- Demonstrar a potencialidade do SOM para análise de variáveis;

- Identificar áreas prioritárias para recuperação florestal através do SOM a nível de Bacia hidrográfica e sub-bacia hidrográfica.

3. SOM


Área Pré-Motora
Cortex Motor Primário
Sulcos Centrais
Córtex somestésico primário
Área do Paladar
Área de associação somestésica
Área de Associação Visual
Área Pré-Frontal
Cortex Visual
Área motor da fala (Área de Broca)
Área sensorial da fala (Área de Wernicke)
Área de Associação Auditiva
Cortex Primário da Audição

3. SOM

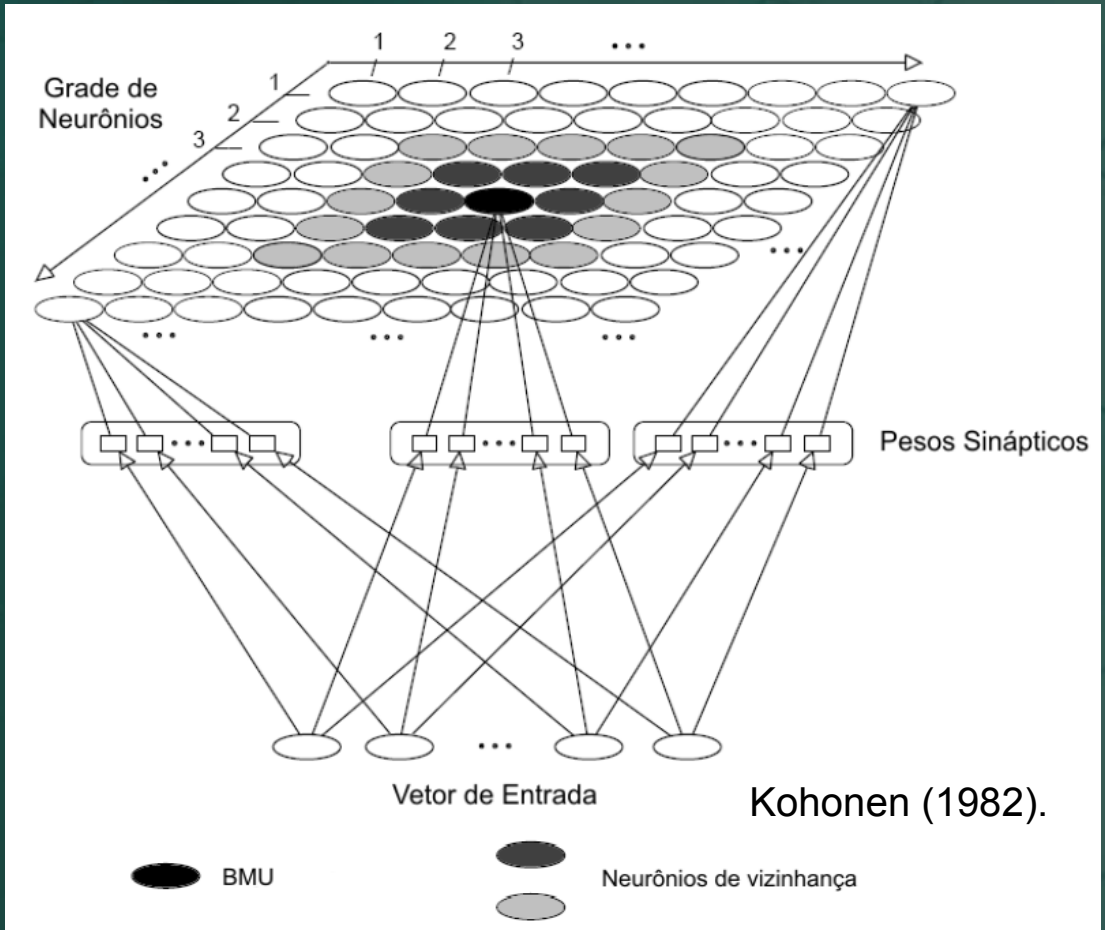
- Passos para o SOM (Haykin, 2001):

  - Inicialização dos dados: normalização e atribuição de pesos no arquitetura de neurônios do SOM – Mapa de Kohonen);

  - Treinamento (seqüêncial ou lote) (Fase de ordenação e convergência): competição, cooperação e adaptação sináptica;

  - Erros Associados.

3. SOM



Kohonen (1982).

Competição

$$d_j = \arg\min\left\{\sum_i^k \left\|(x_i^n - w_{ji}^n)\right\|\right\}$$

Cooperação

$$\beta_{ij'} = EXP\left(-\frac{d_{ij}^2}{2\sigma^2}\right)$$

Adaptação Sináptica

$$w_{ji}^{n+1} = w_{ji}^n + \alpha^n . \beta_{ij'} . \left(x_i^n - w_{ji}^n\right)$$

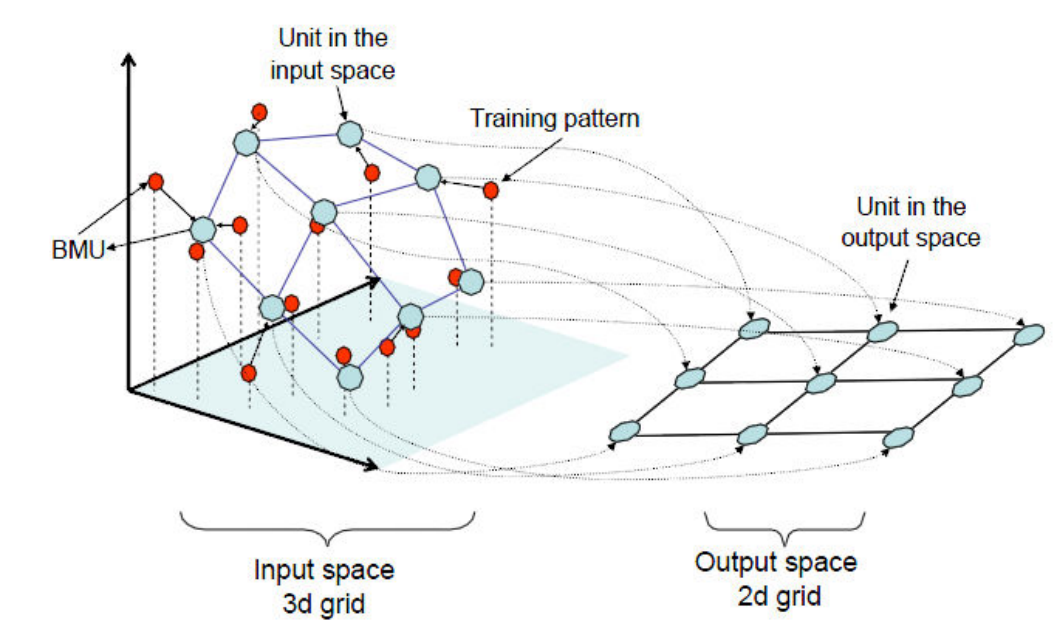
Erro de Quantização médio

$$qE = \frac{\sum_{k=1}^n \left\|x_k - w_{bmu}\right\|}{n}$$

Erro Topológico

$$tE = \frac{1}{n}\sum_{k=1}^n u(x_k)$$

3. SOM



Henriques, 2005

Competição

$$d_j = \arg\min\left\{\sum_i^k \left\|(x_i^n - w_{ji}^n)\right\|\right\}$$

Cooperação

$$\beta_{ij'} = EXP\left(-\frac{d_{ij}^2}{2\sigma^2}\right)$$

Adaptação Sináptica

$$w_{ji}^{n+1} = w_{ji}^n + \alpha^n . \beta_{ij'} . \left(x_i^n - w_{ji}^n\right)$$

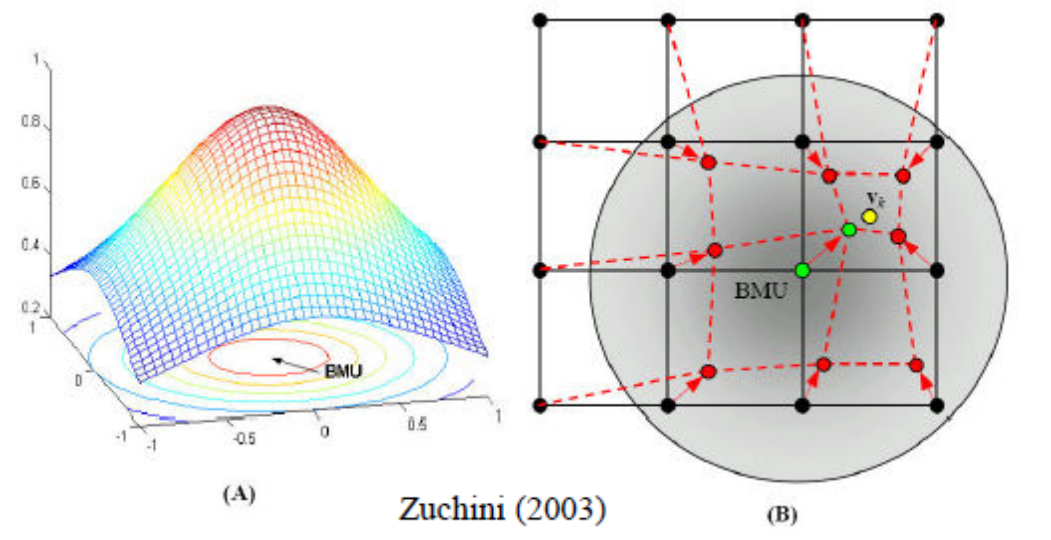
Erro de Quantização médio

$$qE = \frac{\sum_{k=1}^n \left\|x_k - w_{bmu}\right\|}{n}$$

Erro Topológico

$$tE = \frac{1}{n}\sum_{k=1}^n u(x_k)$$

3. SOM



Competição

$$d_j = \arg\min\left\{\sum_i^k \left\|(x_i^n - w_{ji}^n)\right\|\right\}$$

Cooperação

$$\beta_{ij'} = EXP\left(-\frac{d_{ij}^2}{2\sigma^2}\right)$$

Adaptação Sináptica

$$w_{ji}^{n+1} = w_{ji}^n + \alpha^n . \beta_{ij'} . \left(x_i^n - w_{ji}^n\right)$$

Erro de Quantização médio

$$qE = \frac{\sum_{k=1}^n \left\|x_k - w_{bmu}\right\|}{n}$$

Erro Topológico

$$tE = \frac{1}{n}\sum_{k=1}^n u(x_k)$$

3. SOM

SOMToolbox no Matlab
(http://www.cis.hut.fi/somtoolbox/)

Tabela 2.1- Determinação de tamanhos do Mapa de Kohonen pelo comando som_make do SOM toolbox. dlen = número de amostras do dado de entrada.

Número de neurônios (padrão)	$5 * \sqrt{dlen}$
Número de neurônios (pequeno)	$\frac{5 * \sqrt{dlen}}{4}$
Número de neurônios (grande)	$4 * \left(5 * \sqrt{dlen}\right)$

Fonte: Vesanto et al. (2000)

3. SOM

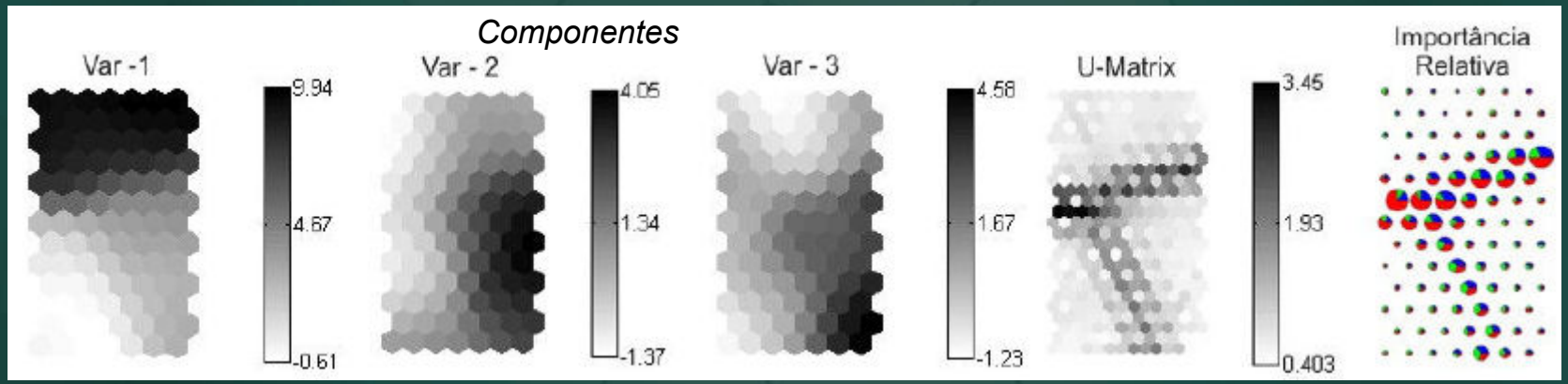
SOMToolbox no Matlab

Tabela 2.2 - Parâmetros padrões de treinamento utilizados pelo comando som_make do Som toolbox. max = ao maior lado do mapa de Kohonen; mpd = número de neurônios dividido por dlen.

	SEQÜÊNCIAL		LOTE	
	Ordenação	Convergência	Ordenação	Convergência
Taxa de treinamento inicial	0.5	0.05	-	-
Raio inicial (inicialização de aleatória)	$\frac{max}{4}$	1	$\frac{max}{4}$	1
Raio inicial (inicialização por autovetores)	$\frac{max}{8}$	1	$\frac{max}{8}$	1
Raio final	1	1	1	1
Número de épocas (padrão)	10*mpd	40*mpd	10*mpd	40*mpd
Número de épocas (curto)	2,5*mpd	10*mpd	2,5*mpd	10*mpd
Número de épocas (longo)	40*mpd	160*mpd	40*mpd	160*mpd

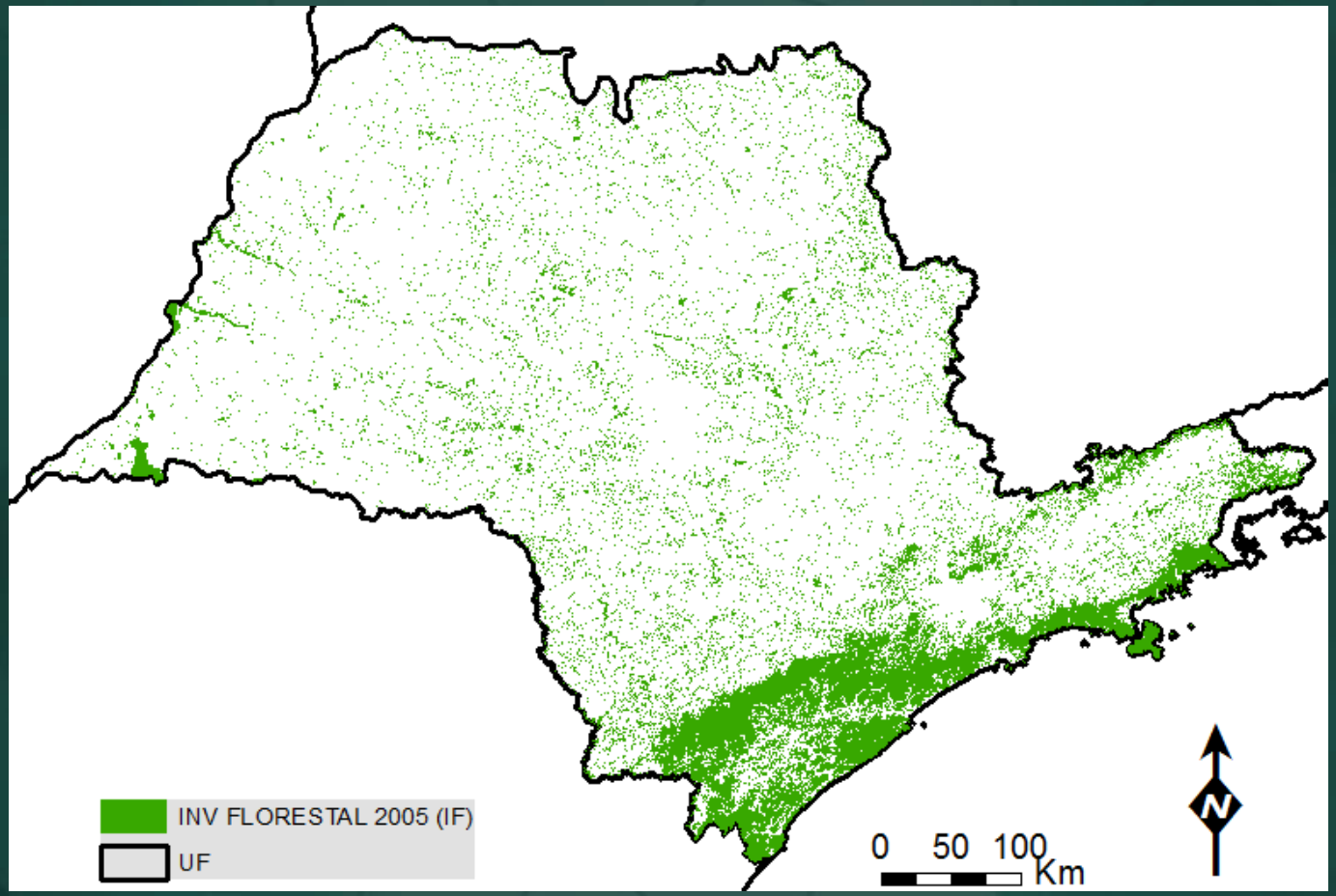
Fonte: Vesanto et al. (2000)

3. SOM


Componentes
Var -1
9.94
4.67
-0.61
Var - 2
4.05
1.34
-1.37
Var - 3
4.58
1.67
-1.23
U-Matrix
3.45
1.93
0.403
Importância Relativa

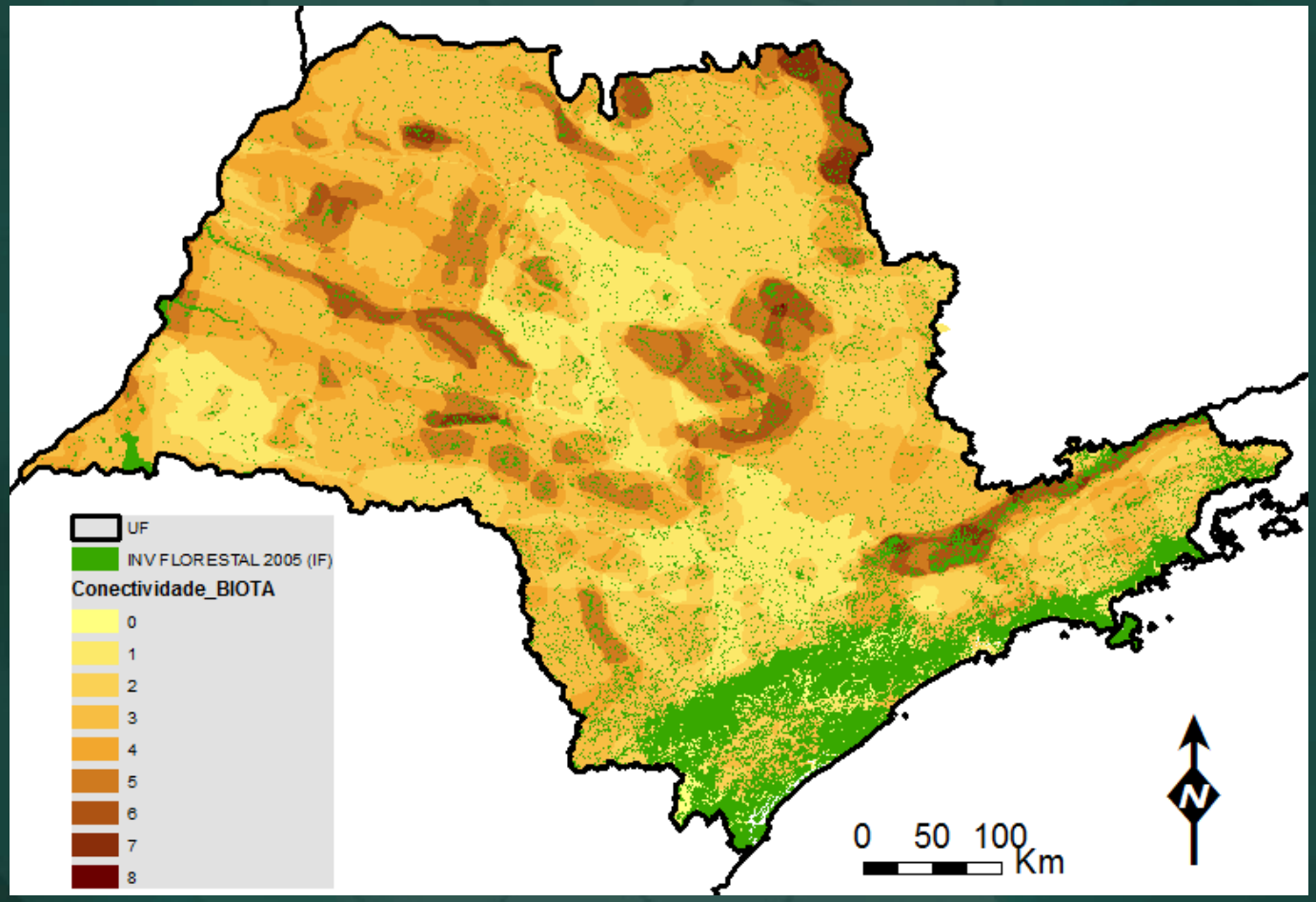
4. Material e Métodos

4.1 Área de Estudo



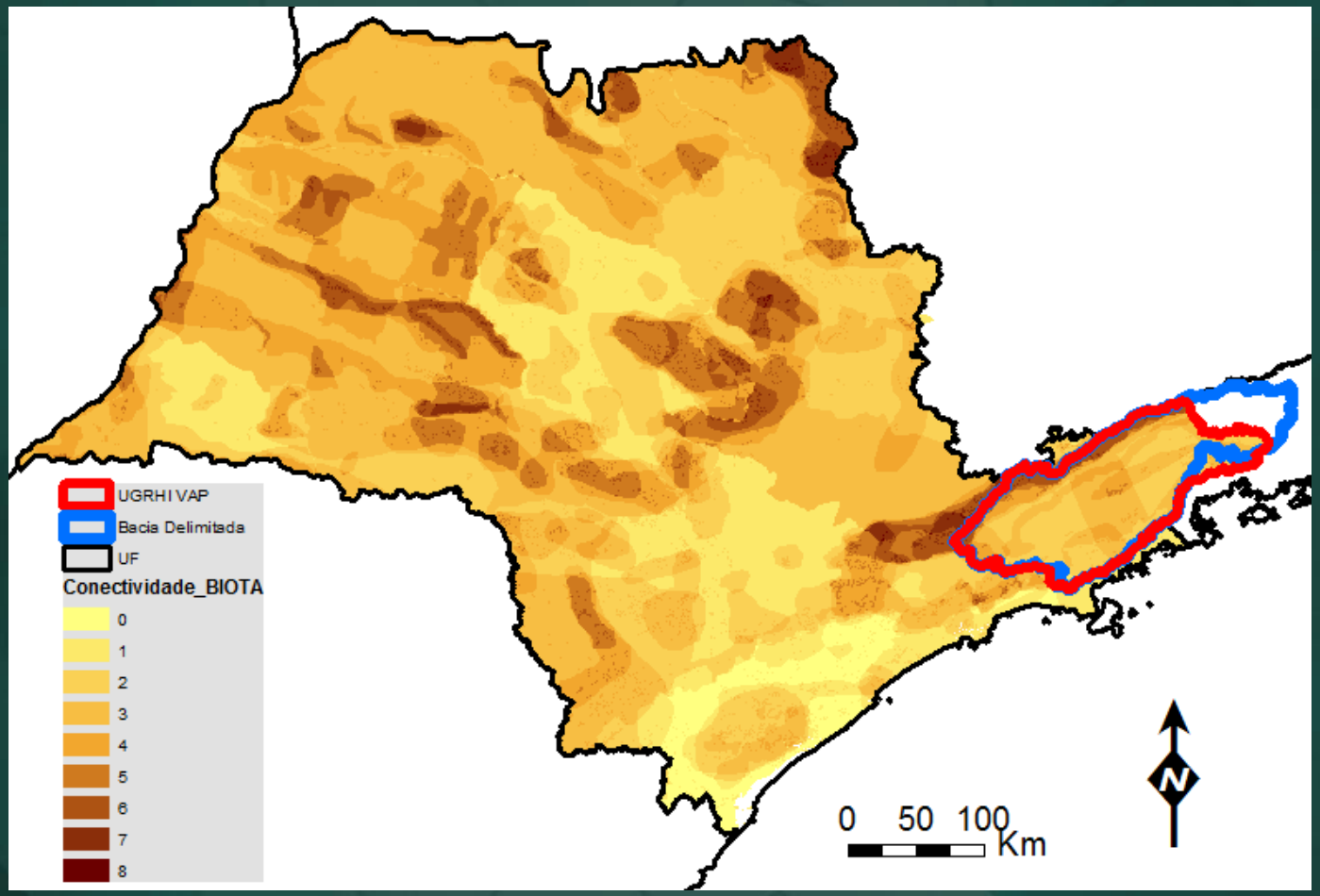
4. Material e Métodos

4.1 Área de Estudo



4. Material e Métodos

4.1 Área de Estudo



4. Metodologia

4.2 Material

4.2.1 UGRHI do rio Paraíba do Sul Paulista:

- Dado SRTM para divisão da bacia em sub-bacias (*Automated Geospatial Watershed Assessment tool* - AGWA) (Miller et
- al., 2007);
- Inventário Florestal (IF, 2005);
- Mapa de Conectividade (BIOTA/Fapesp, 2007).

4. Metodologia
4.2 Material
4.2.2 Sub-bacia delimitada:

- Inventário Florestal (IF, 2005);
- Mapa de Conectividade (BIOTA/Fapesp, 2007);
- Modelo Digital de Elevação (Vetores 1:50.000);
- Mapa de Solos (CEIVAP);
- Dados pluviométricos (ANA, DAEE);
- Localização de estradas pavimentadas (DEER), zona urbana consolidada (TM/Landsat 08/09/04), construções no meio rural em forma de pontos (DAEE);
- NDVI (TM/Landsat 08/09/04);

4. Metodologia

4.3 Preparo e Análise

4.3.1 UGRHI do rio Paraíba do Sul Paulista:

- Extração de variáveis por sub-bacias geradas:
  - Conectividade;
  - Taxa de remanescente florestal por APP de curso d´água;
  - Latitude e Longitude (coisas mais perto são mais parecidas, Tobler).

4. Metodologia
4.3 Preparo e Análise

- SOM:
  - Adequação da polarização (+ -) das variáveis. Valores crescentes indicariam prioridade para recuperação, valores decrescentes indicariam rejeição à recuperação.
  - Inicialização e Treinamento sequencial padrão do SOMtoolbox;
  - Análise Visual dos Dados;
  - Definição dos Agrupamentos por k-médias e Davies-Bouldin;

4. Metodologia

4.3 Preparo e Análise

4.3.2 Sub-Bacia Delimitada:

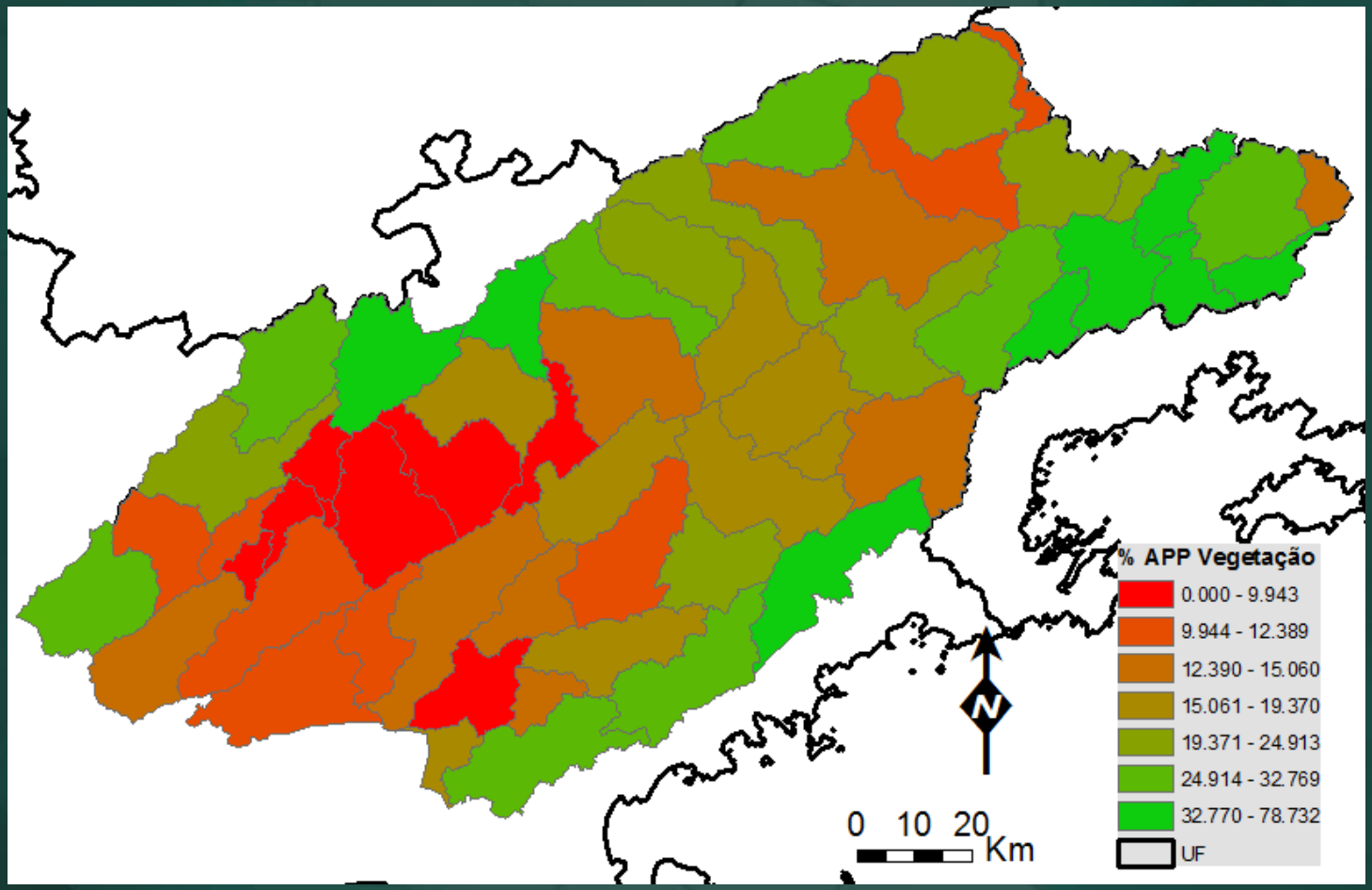
- Extração de variáveis por polígonos nas áreas de APP sem vegetação segundo Inv. Florestal (IF, 2005) (até 10 km de buffer):
  - Distância a fragmentos florestais;
  - Distância a fragmentos florestais maiores e coesos (métrica shape);
  - Distância a manchas urbanas consolidadas e rodovias pavimentadas;
  - Distância a construções rurais;
  - NDVI;
  - Potencial Natural de Erosão Natural = LS x K(solo) x R(chuva);
  - Latitude e Longitude (coisas mais perto são mais parecidas, Tobler).

4. Metodologia
4.3 Preparo e Análise

- SOM:
  - Adequação da polarização (+ -) das variáveis. Valores crescentes indicariam prioridade para recuperação, valores decrescentes indicariam rejeição à recuperação.
  - Inicialização e Treinamento padrão do SOMtoolbox;
  - Análise Visual dos Dados;
  - Definição dos Agrupamentos por k-médias e Davies-Bouldin;
  - Histograma da soma das componentes normalizadas: escolha do agrupamento mais propício para a recuperação florestal.
  - Análise da frequencia dos neurônios vencedores sob o agrupamento definido

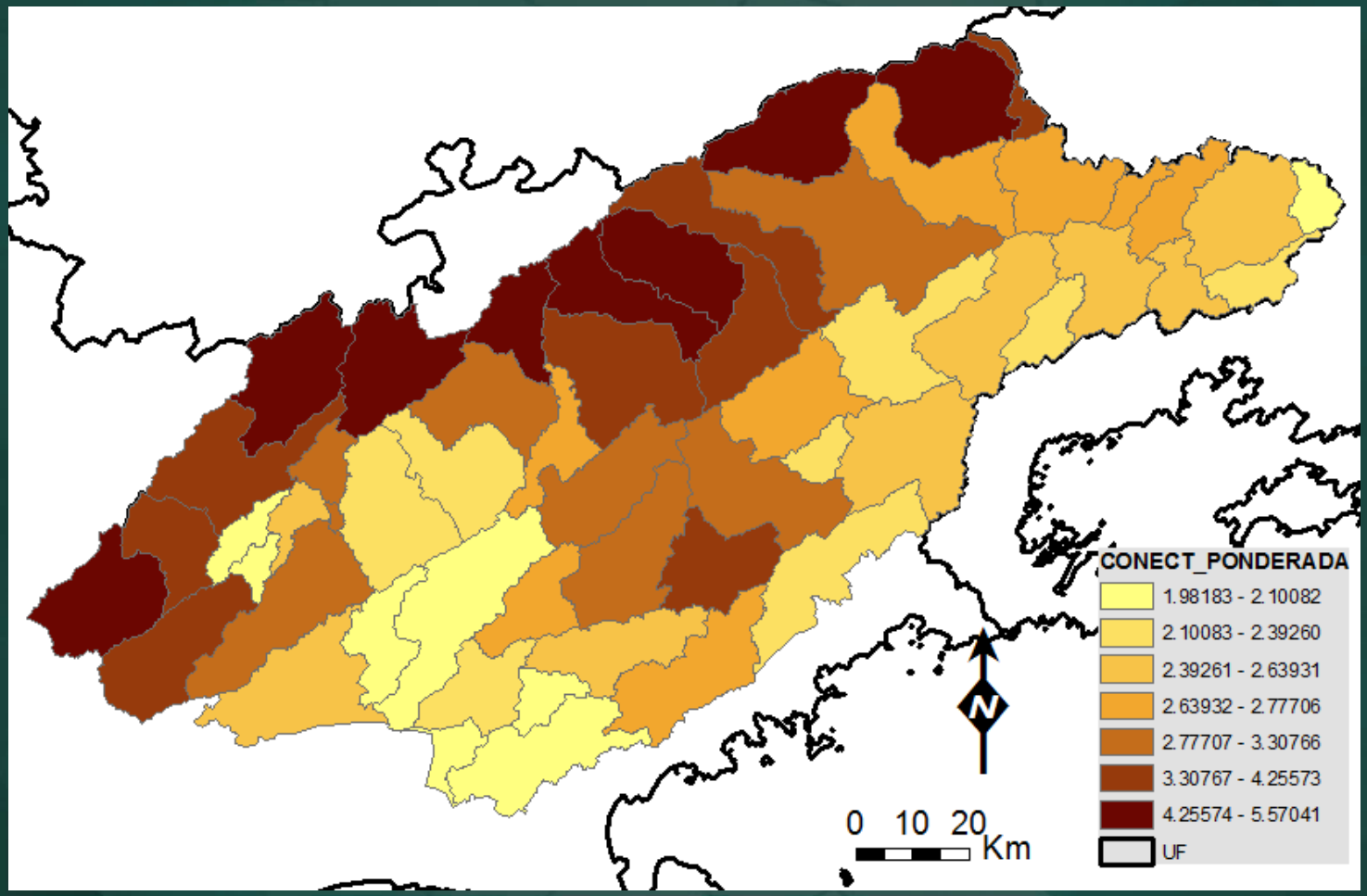
5. Resultados

5.1 UGRHI do rio Paraíba do Sul Paulista



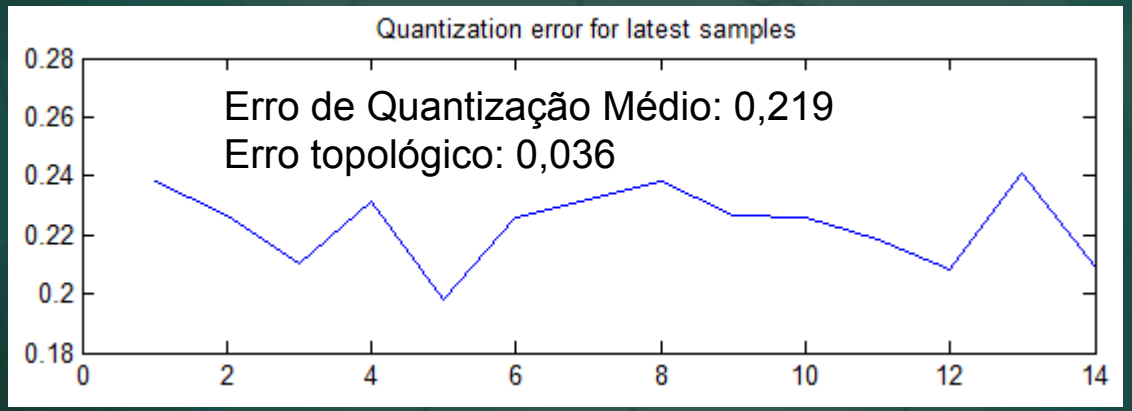
5. Resultados

5.1 UGRHI do rio Paraíba do Sul Paulista



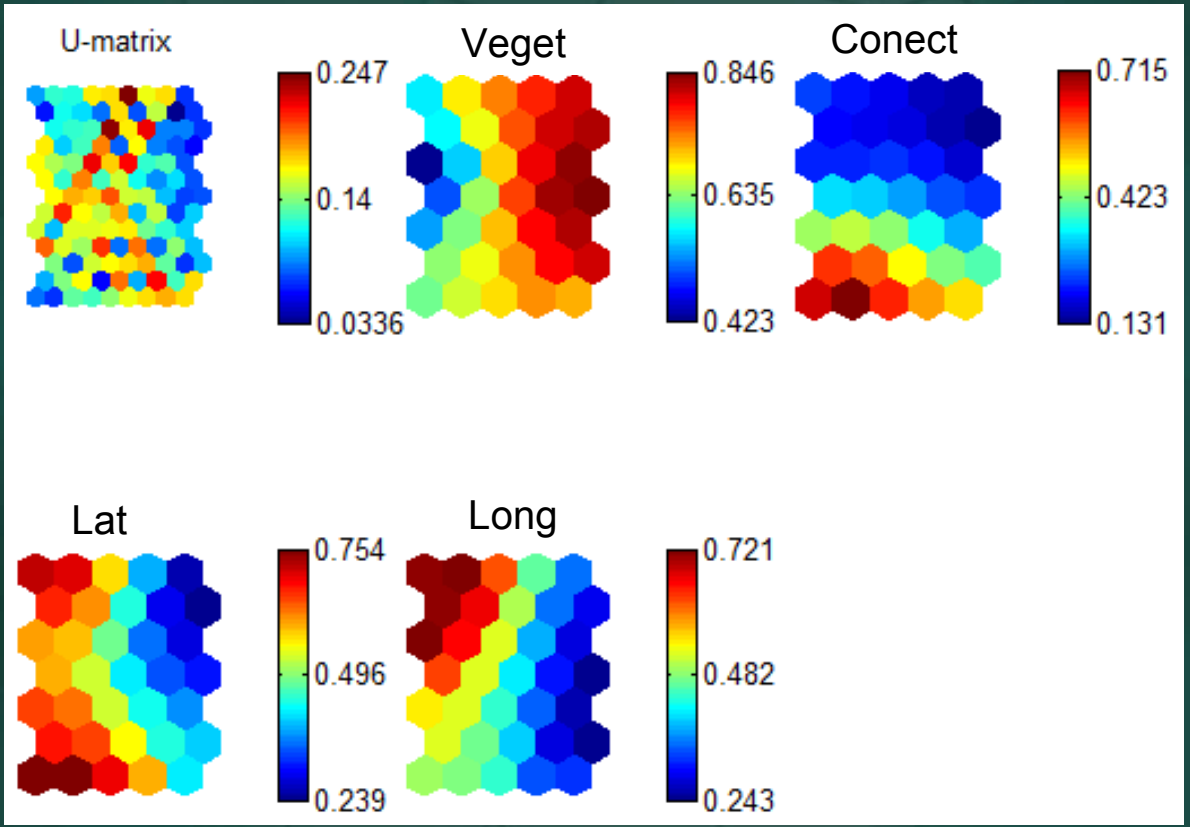
5. Resultados

5.1 UGRHI do rio Paraíba do Sul Paulista



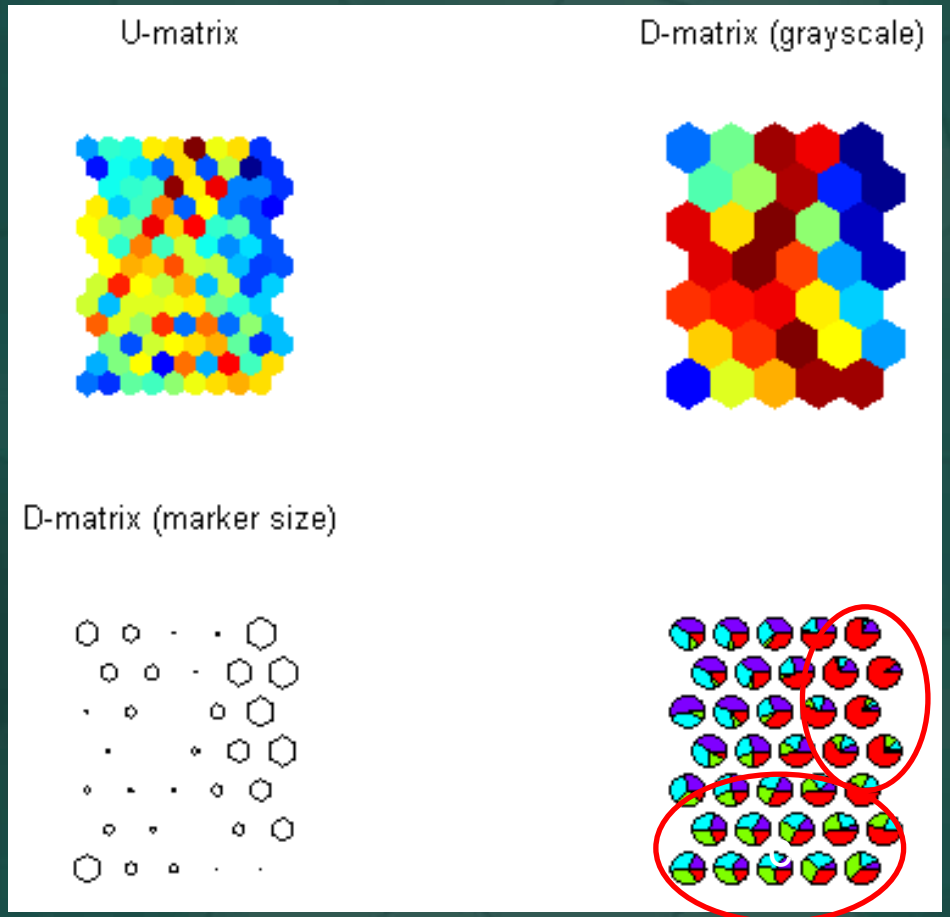
5. Resultados

5.1 UGRHI do rio Paraíba do Sul Paulista



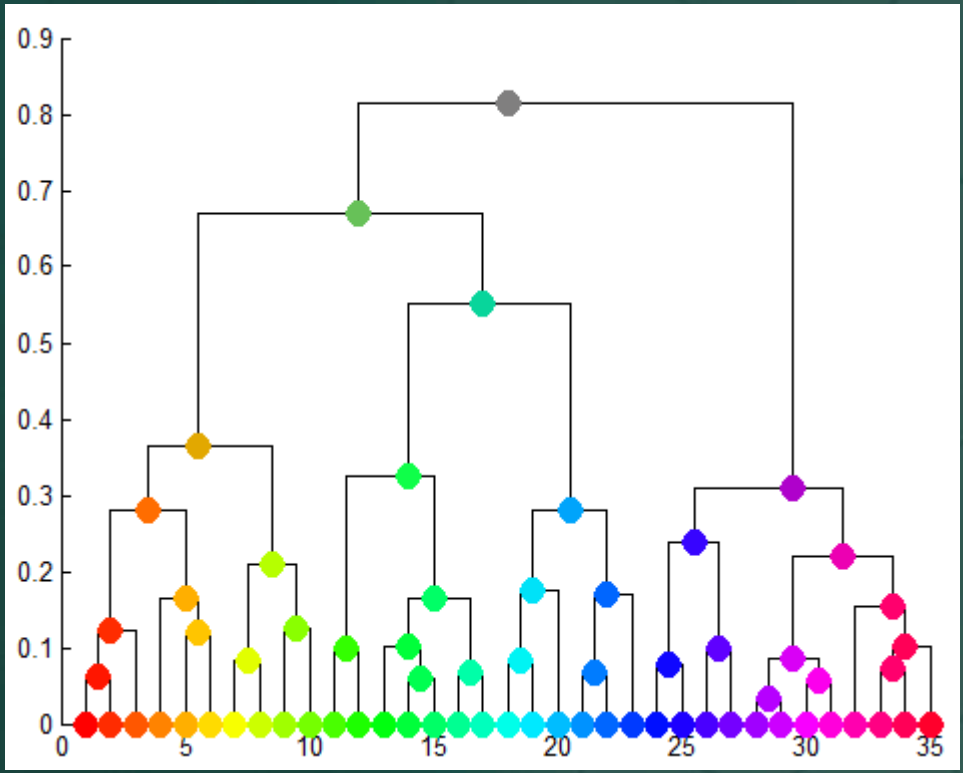
5. Resultados

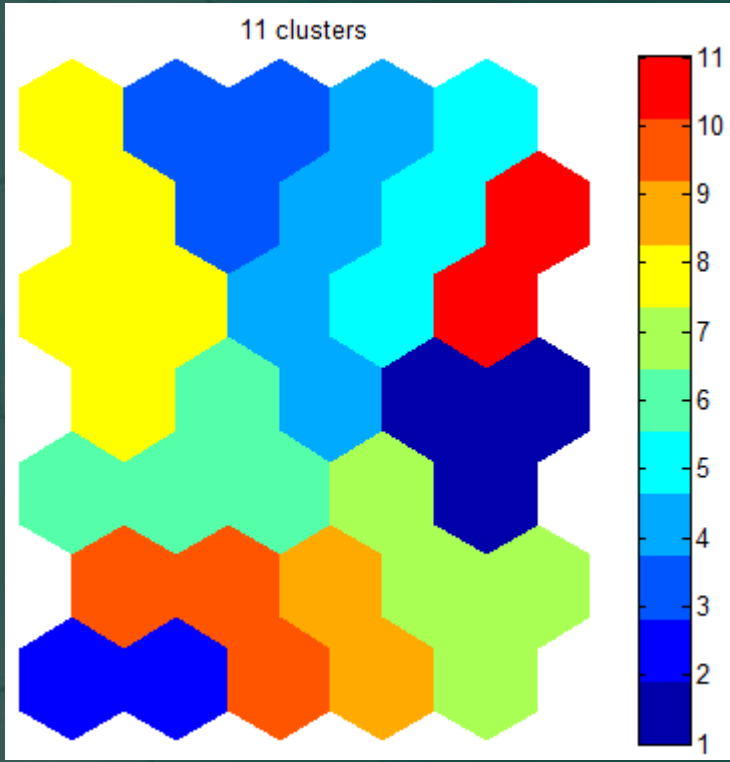
5.1 UGRHI do rio Paraíba do Sul Paulista



5. Resultados

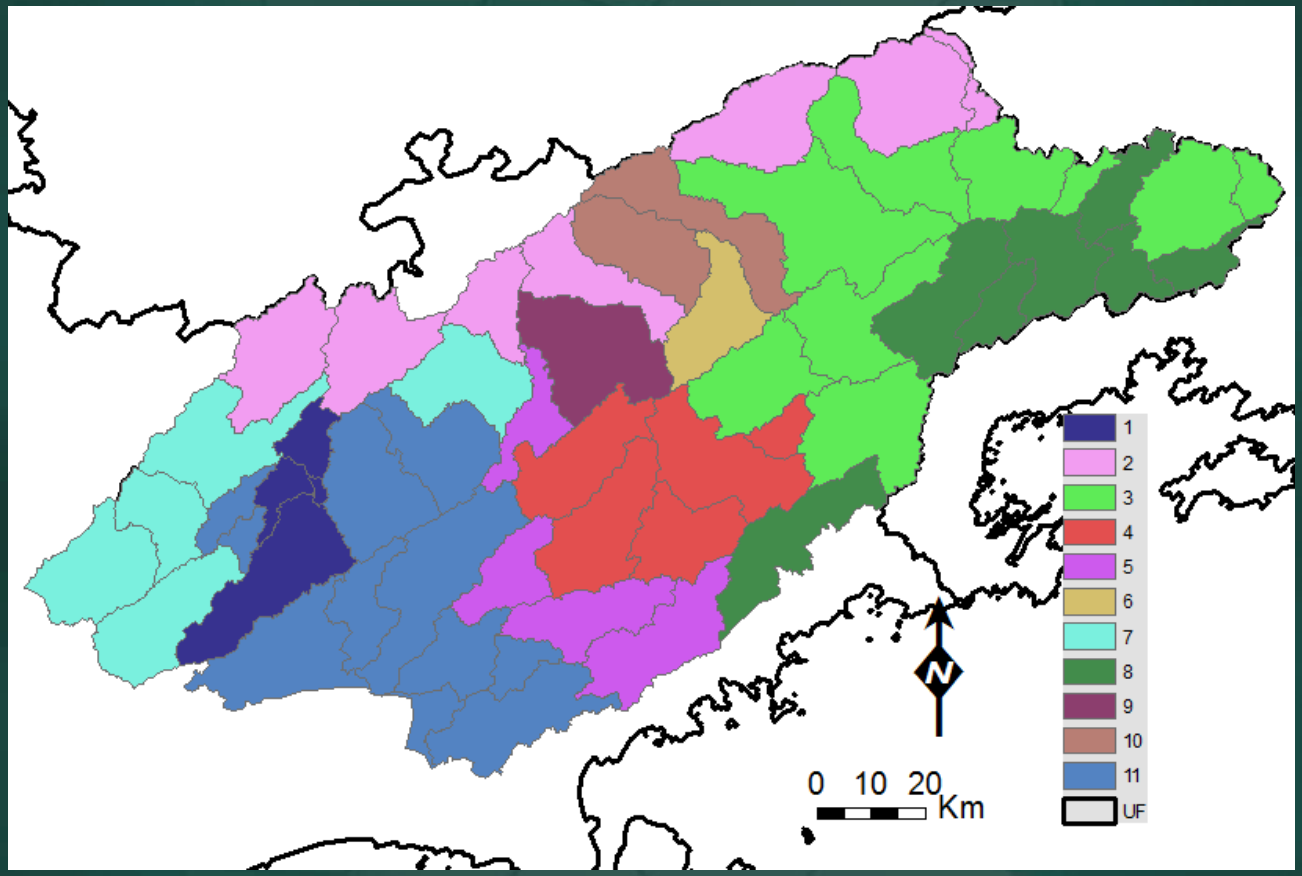
5.1 UGRHI do rio Paraíba do Sul Paulista





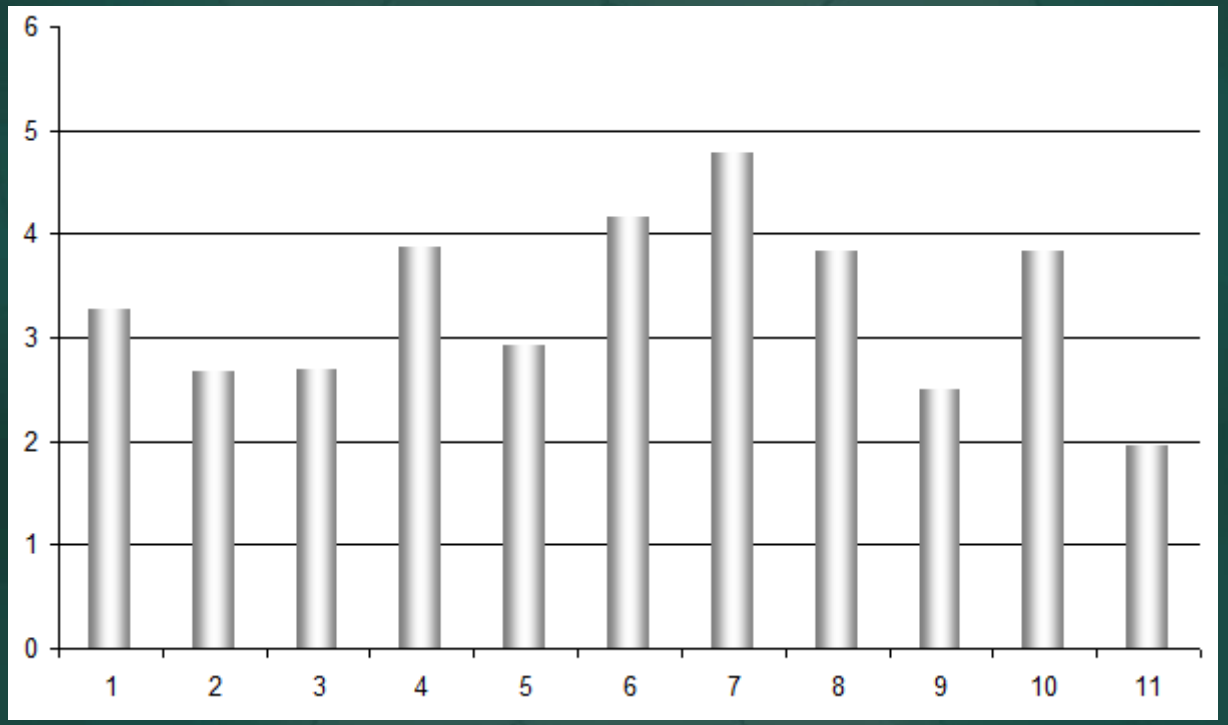
5. Resultados

5.1 UGRHI do rio Paraíba do Sul Paulista



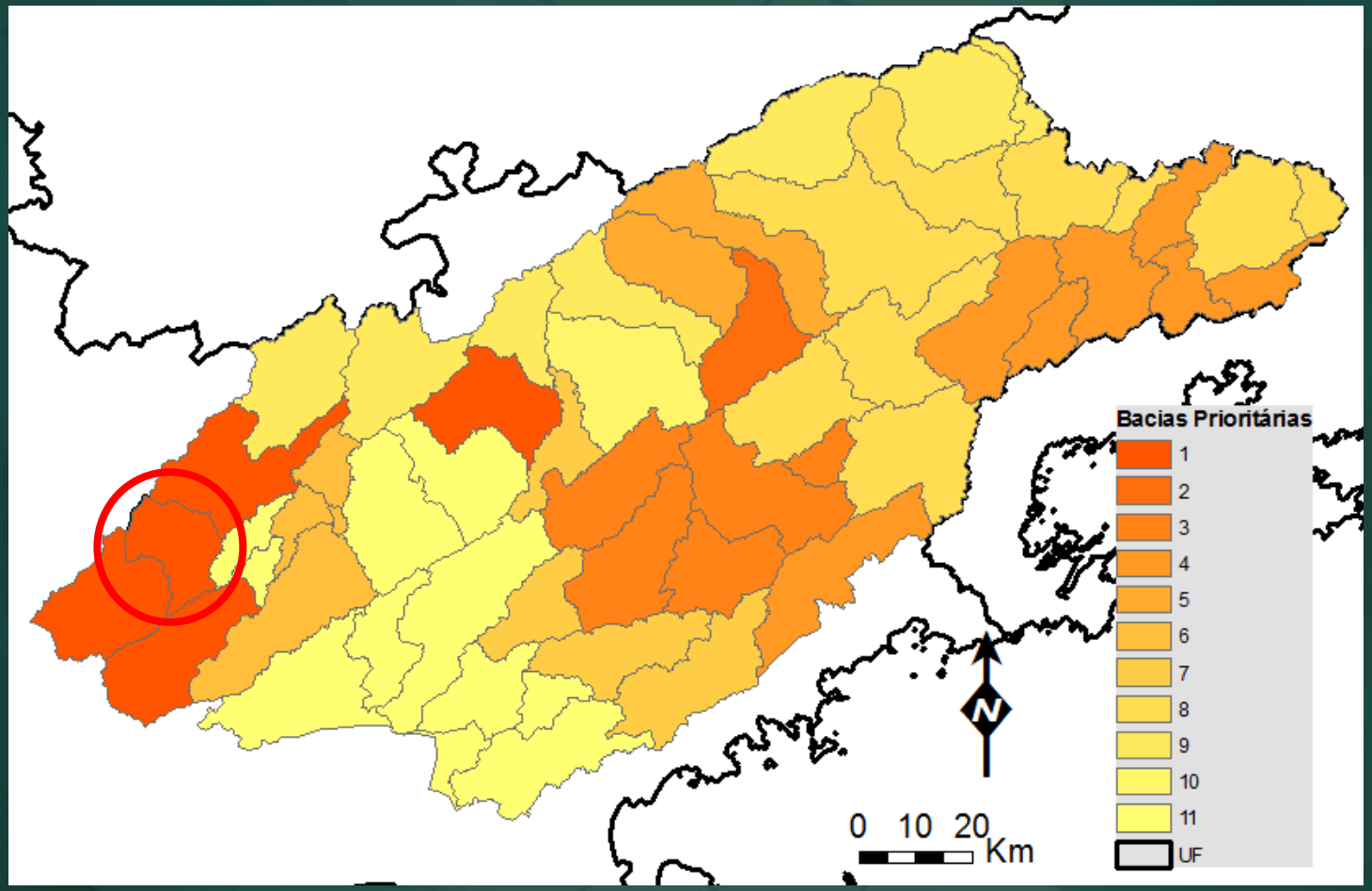
5. Resultados

5.1 UGRHI do rio Paraíba do Sul Paulista



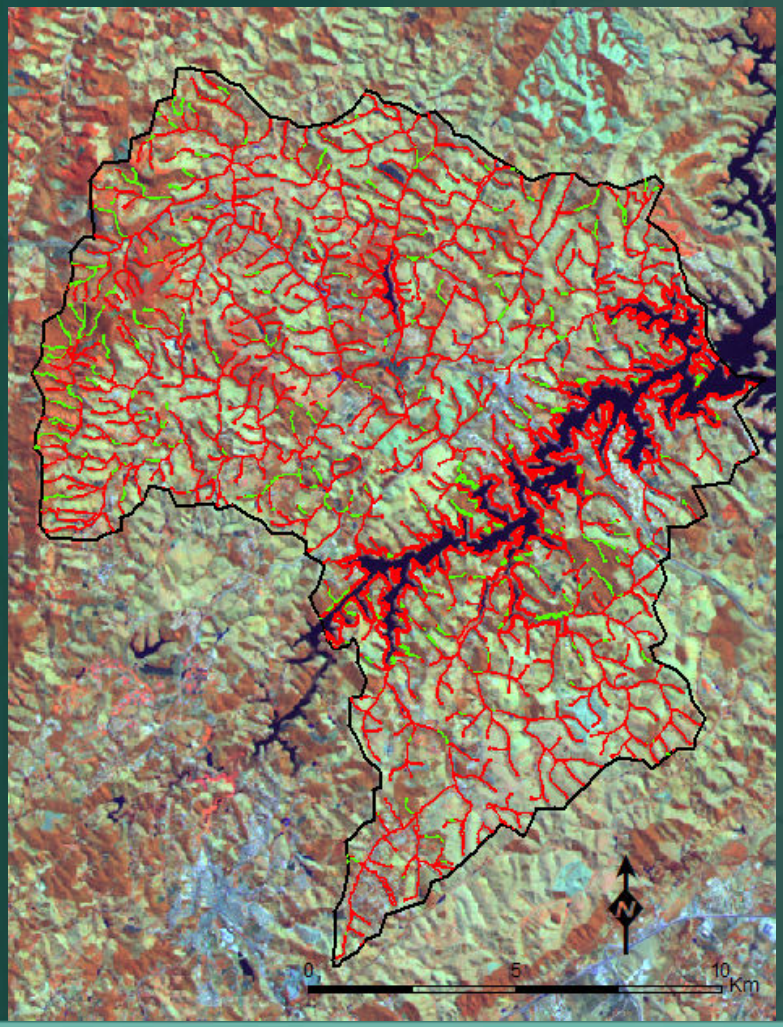
5. Resultados

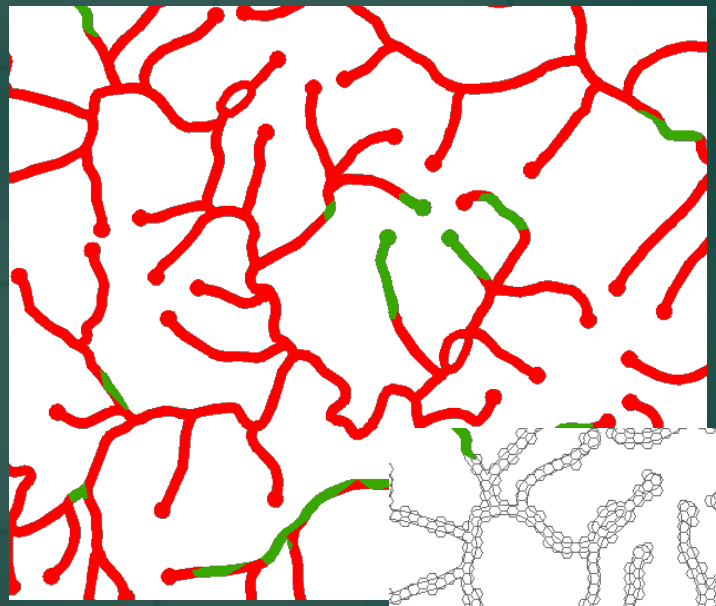
5.1 UGRHI do rio Paraíba do Sul Paulista

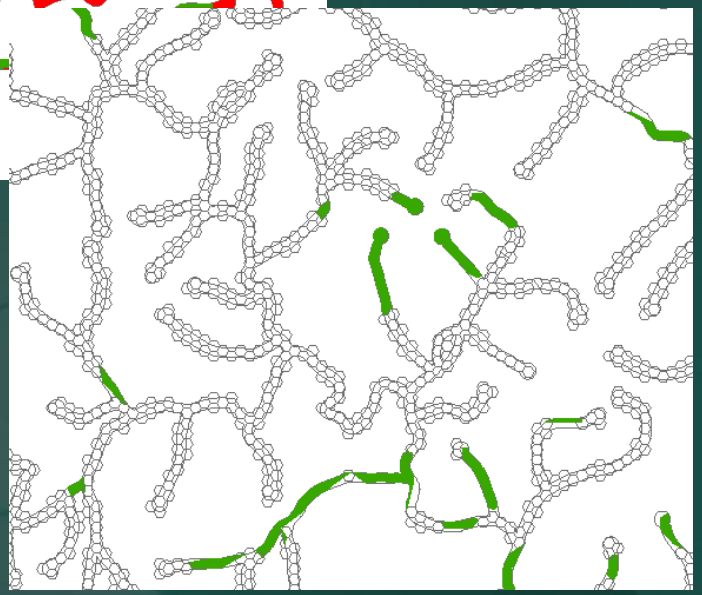


5. Resultados

5.2 Sub-Bacia

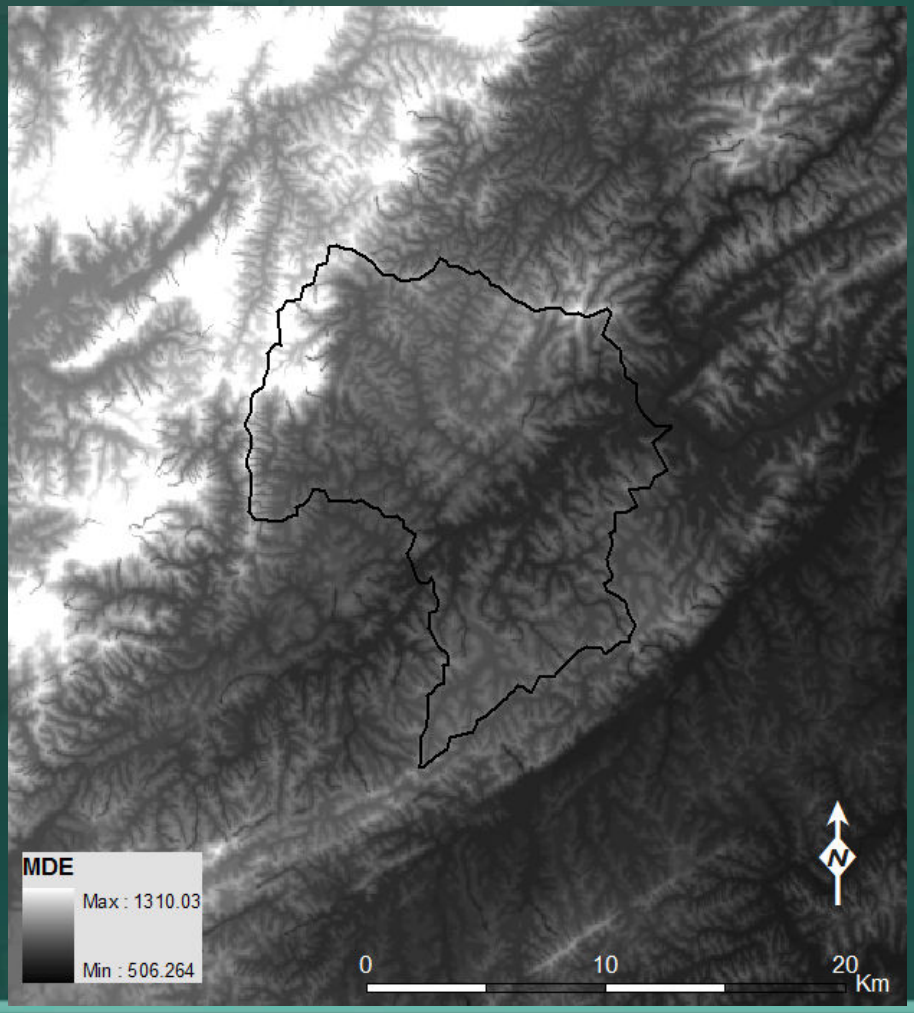






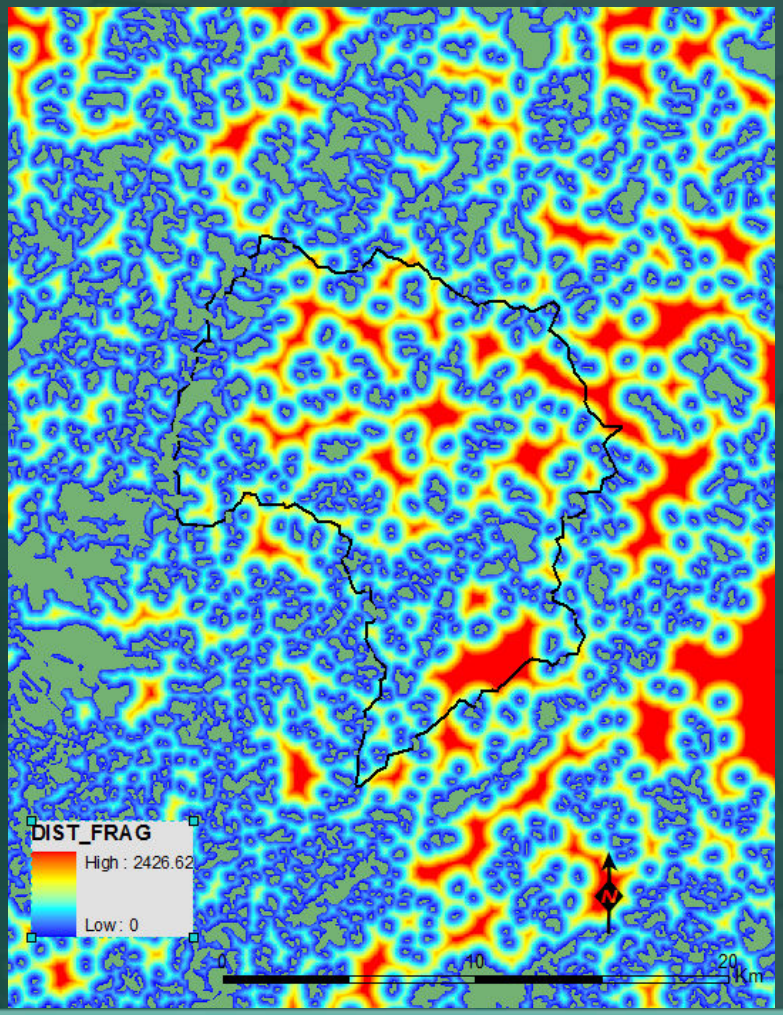
5. Resultados

5.2 Sub-Bacia



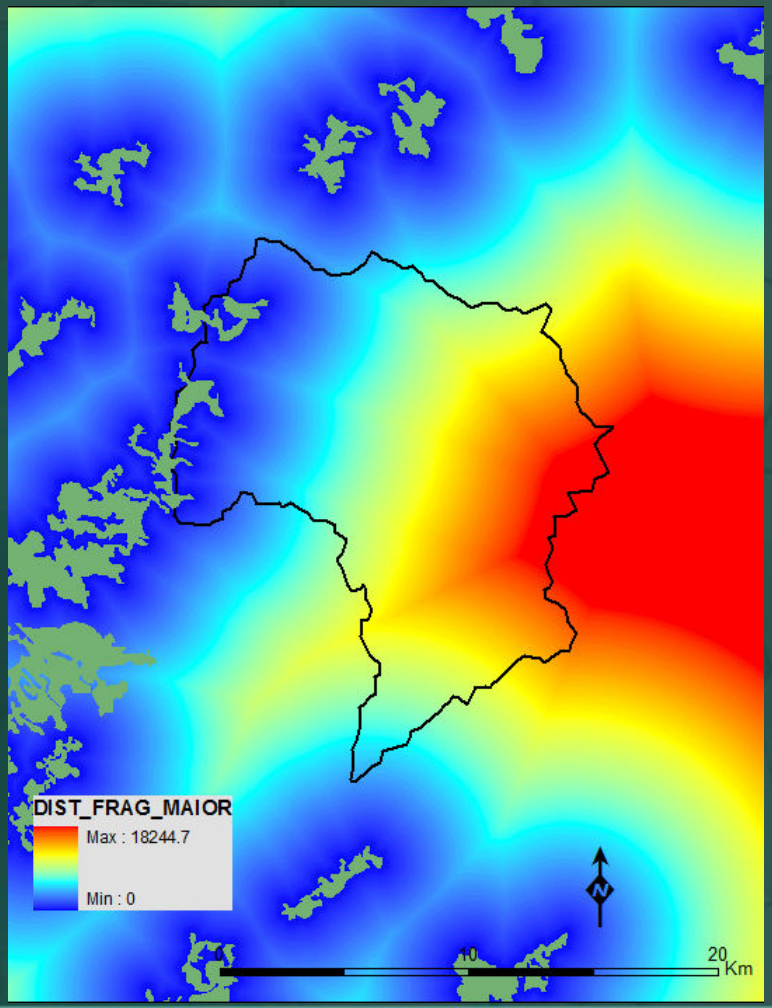
5. Resultados

5.2 Sub-Bacia



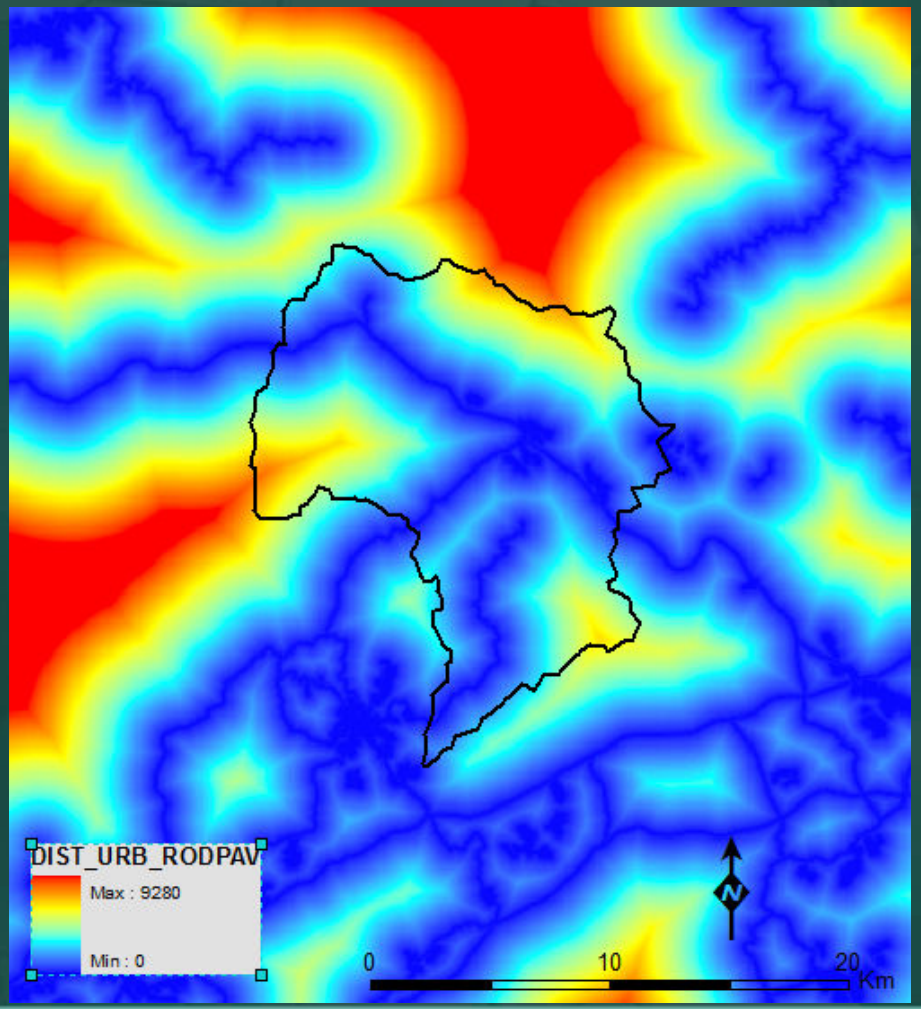
5. Resultados

5.2 Sub-Bacia



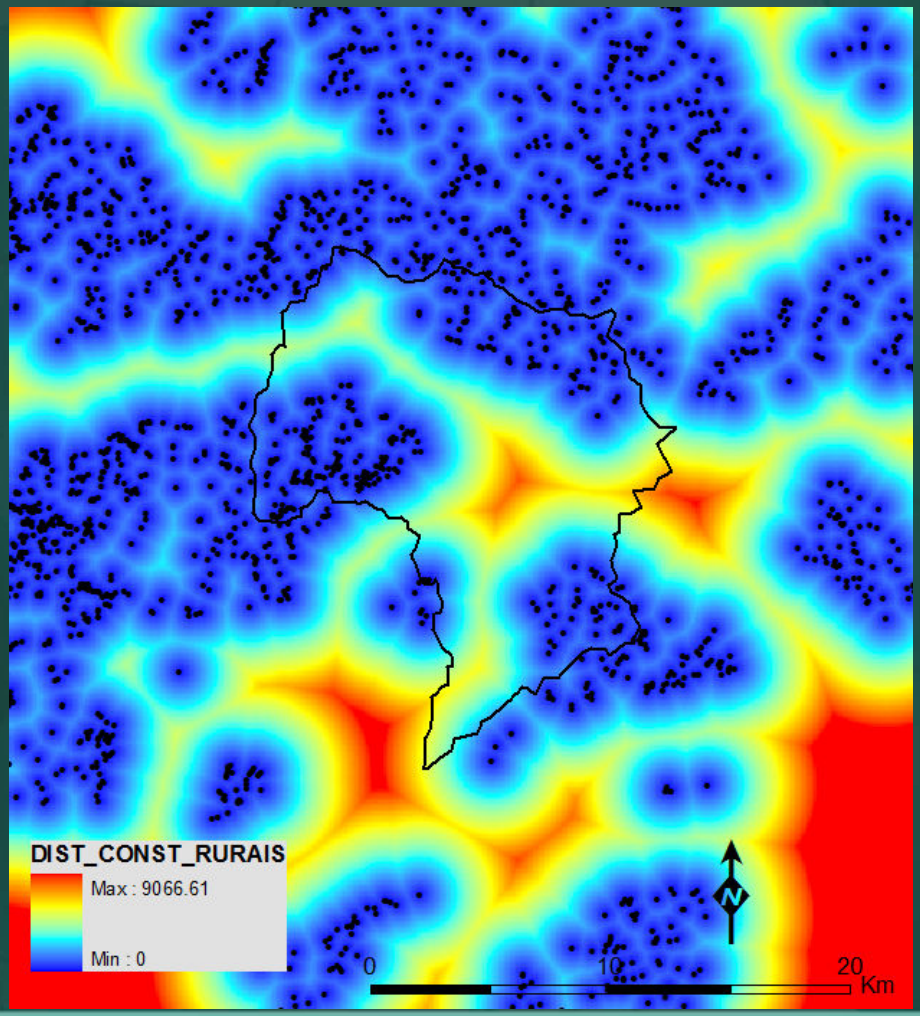
5. Resultados

5.2 Sub-Bacia



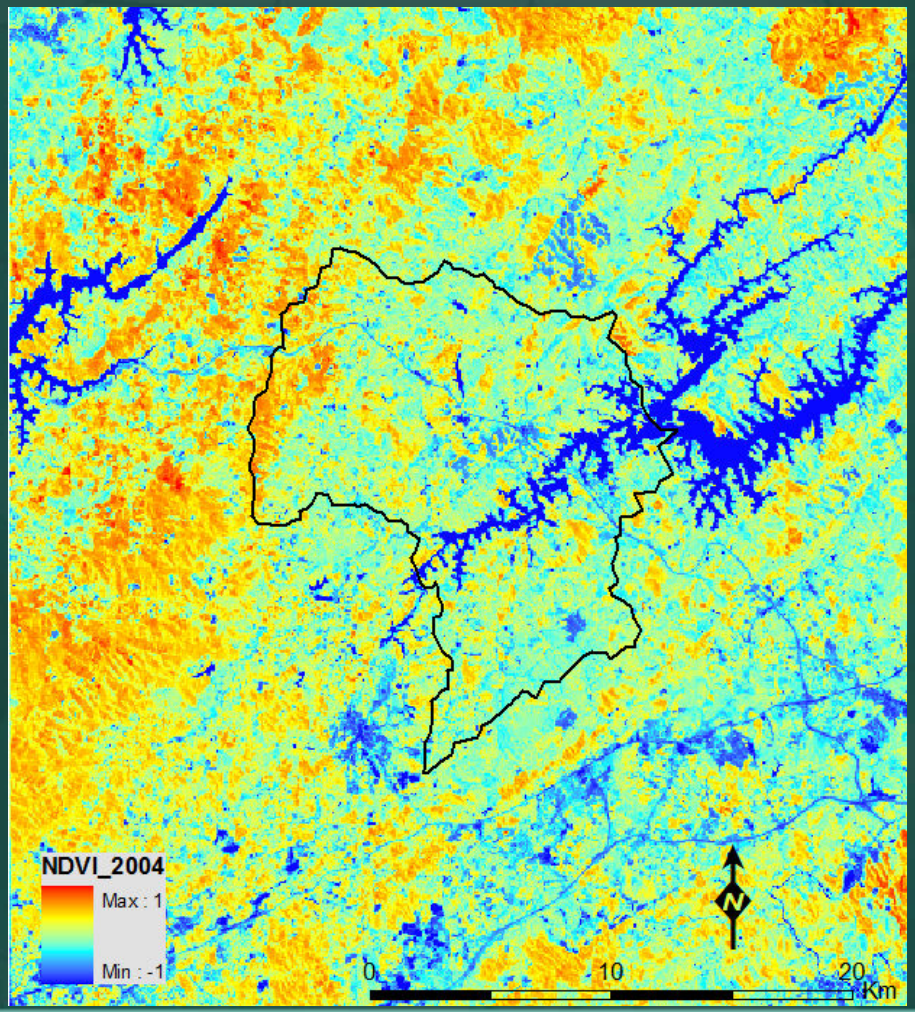
5. Resultados

5.2 Sub-Bacia



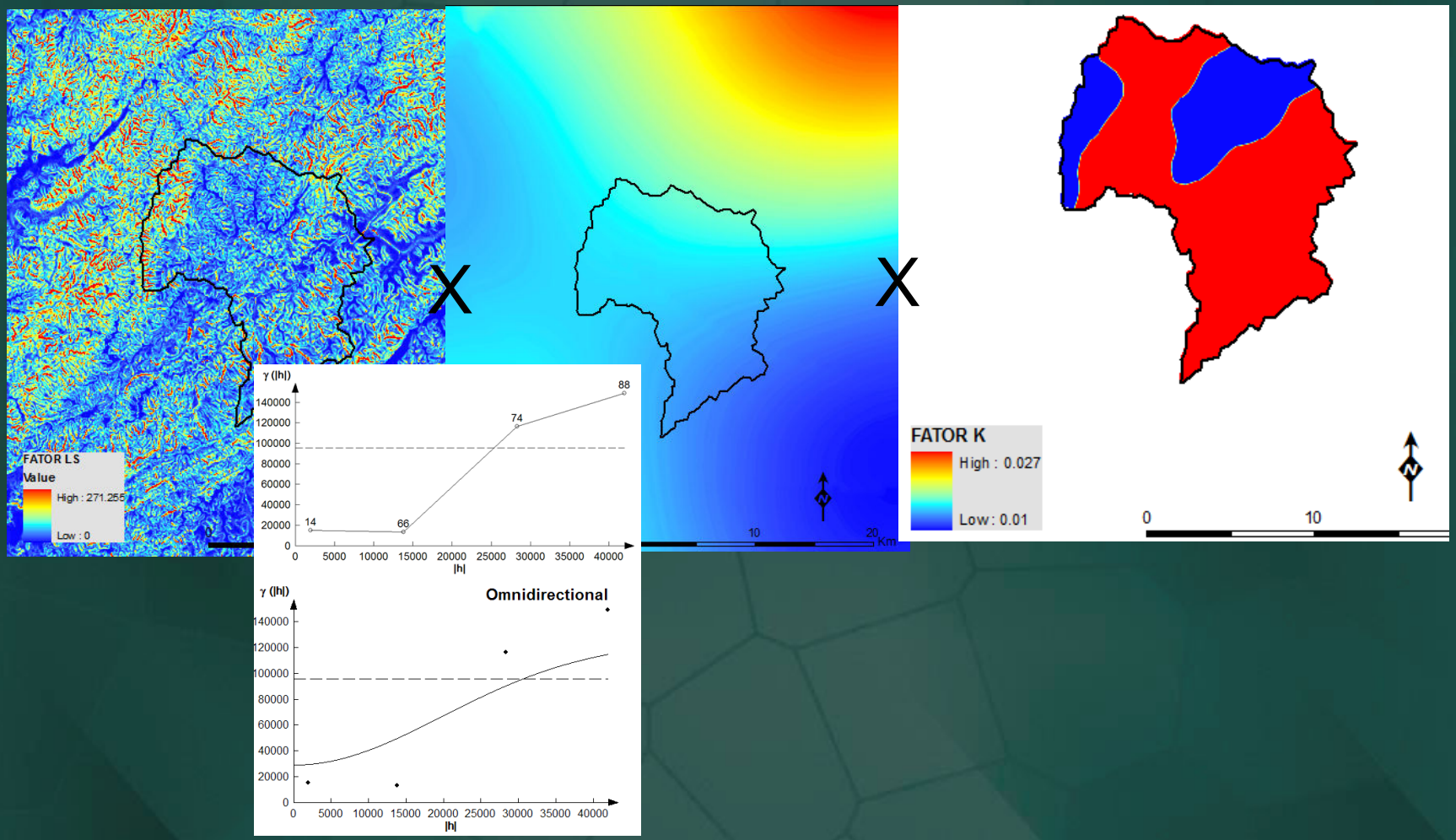
5. Resultados

5.2 Sub-Bacia



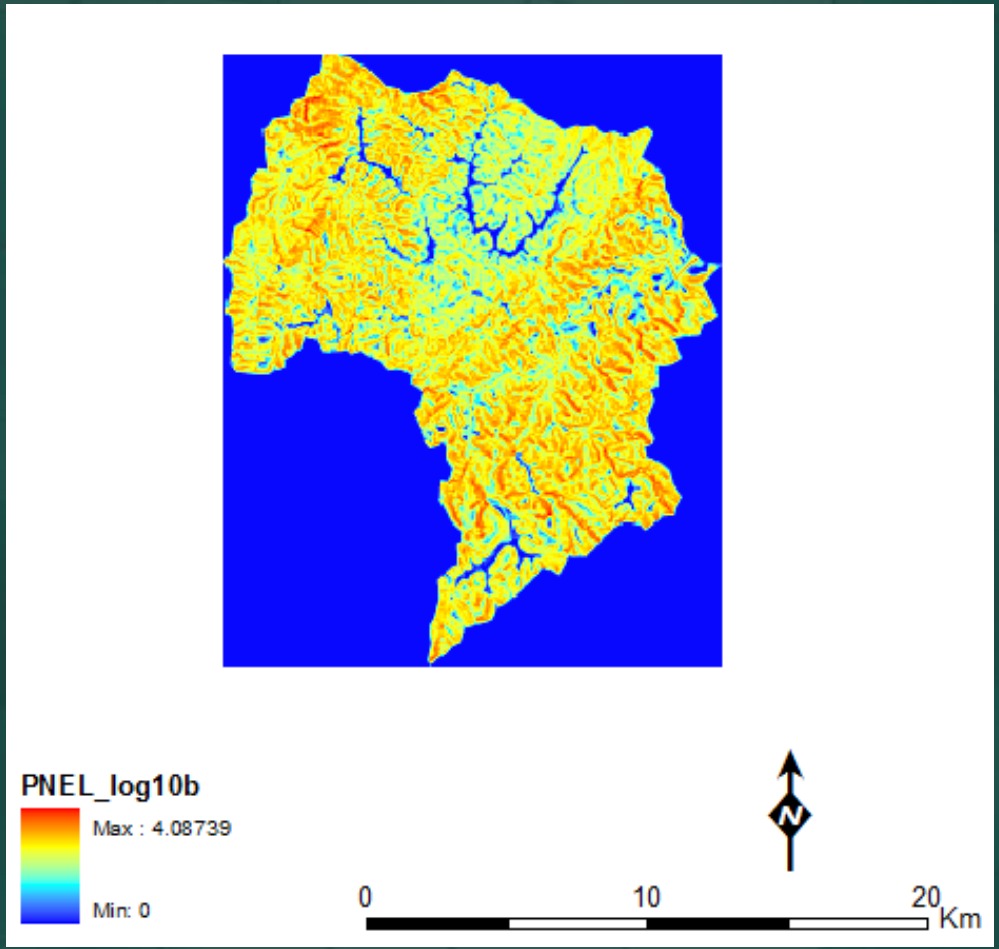
5. Resultados

5.2 Sub-Bacia



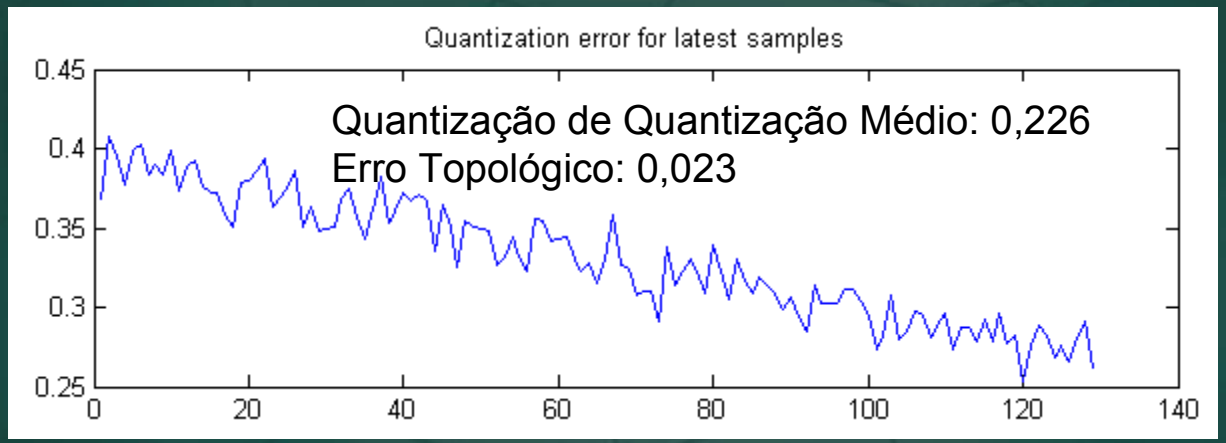
5. Resultados

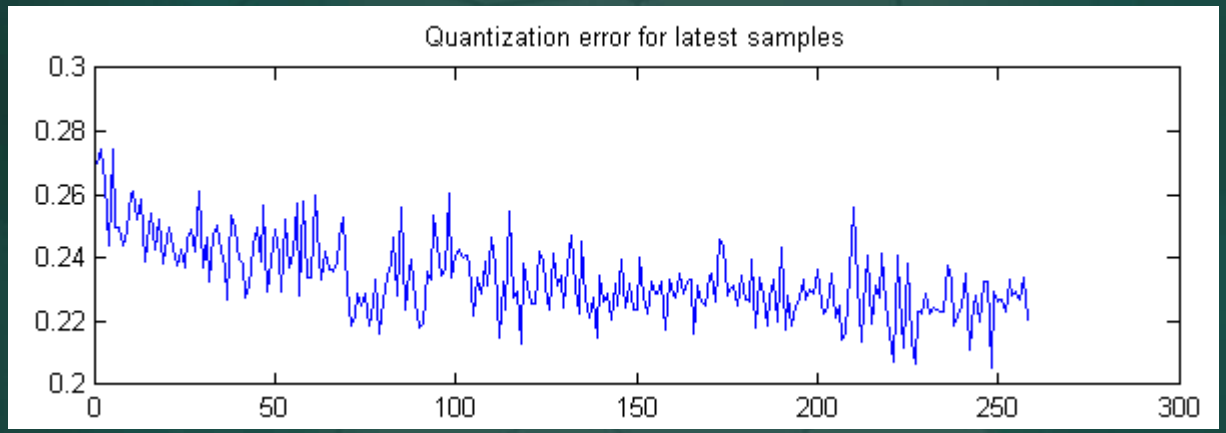
5.2 Sub-Bacia



5. Resultados

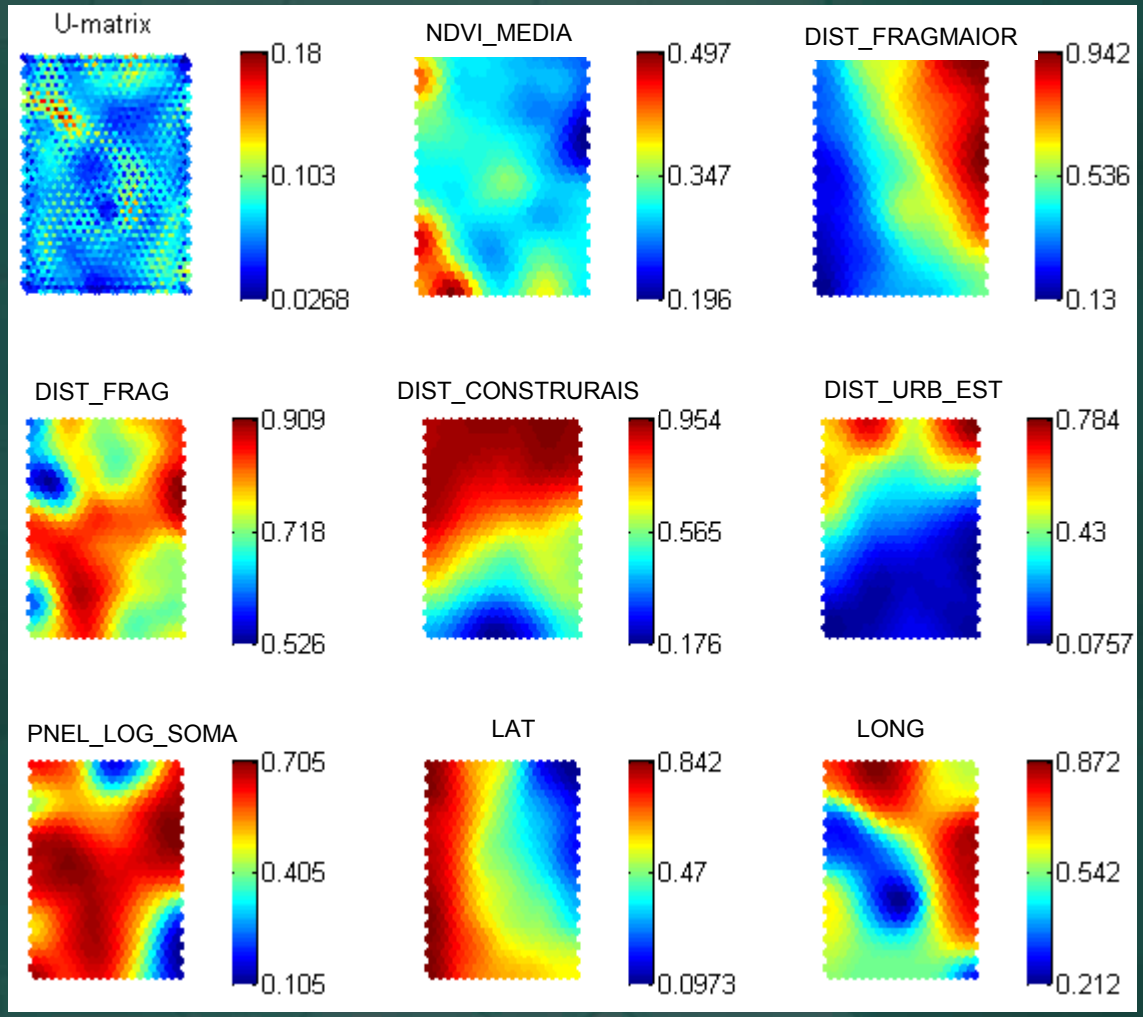
5.2 Sub-Bacia





5. Resultados

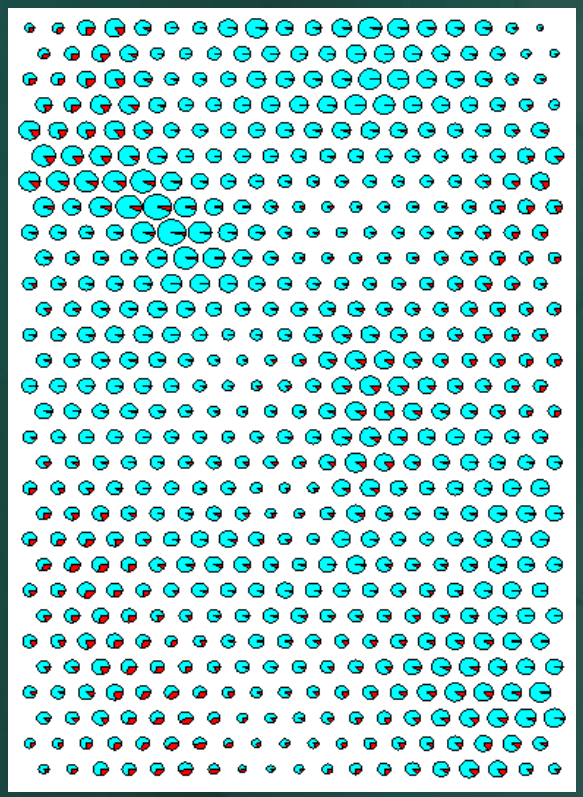
5.2 Sub-Bacia



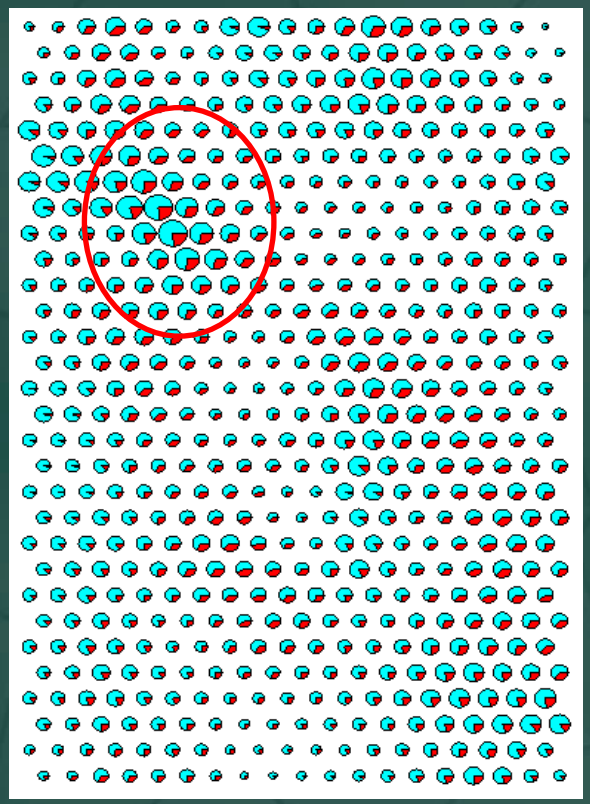
5. Resultados

5.2 Sub-Bacia

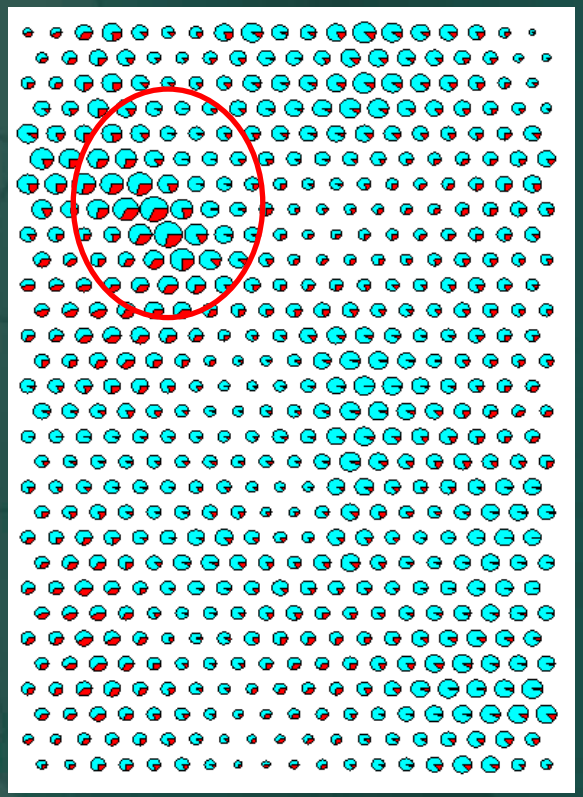
NDVI_MEDIA



DIST_FRAGMAIOR



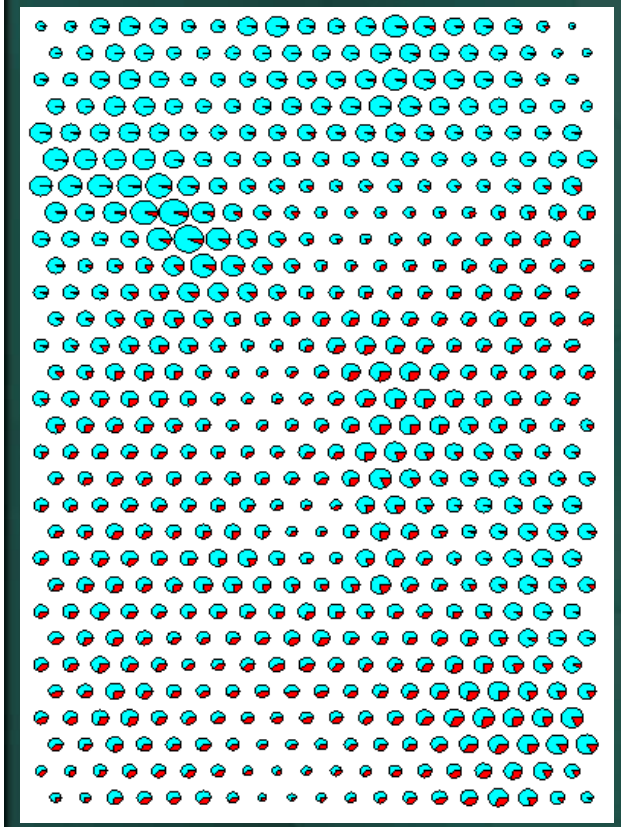
DIST_FRAG



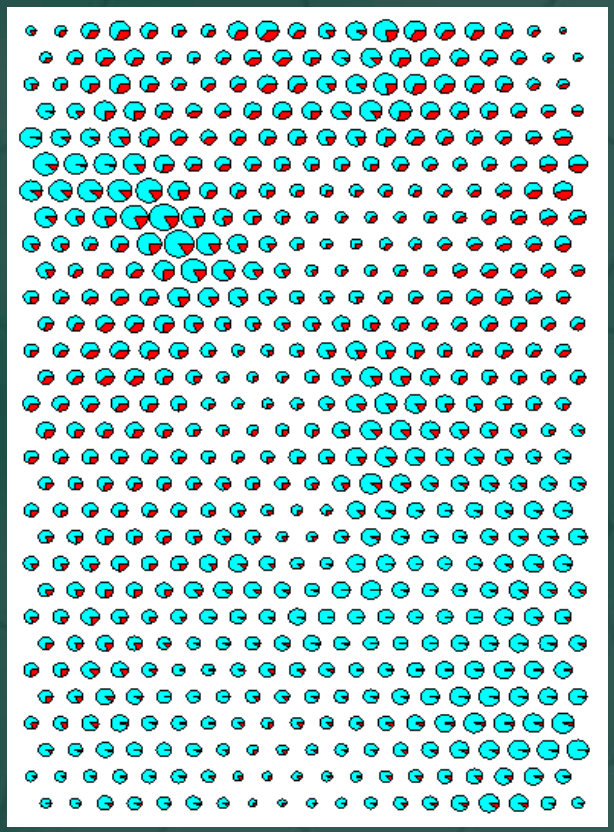
5. Resultados

5.2 Sub-Bacia

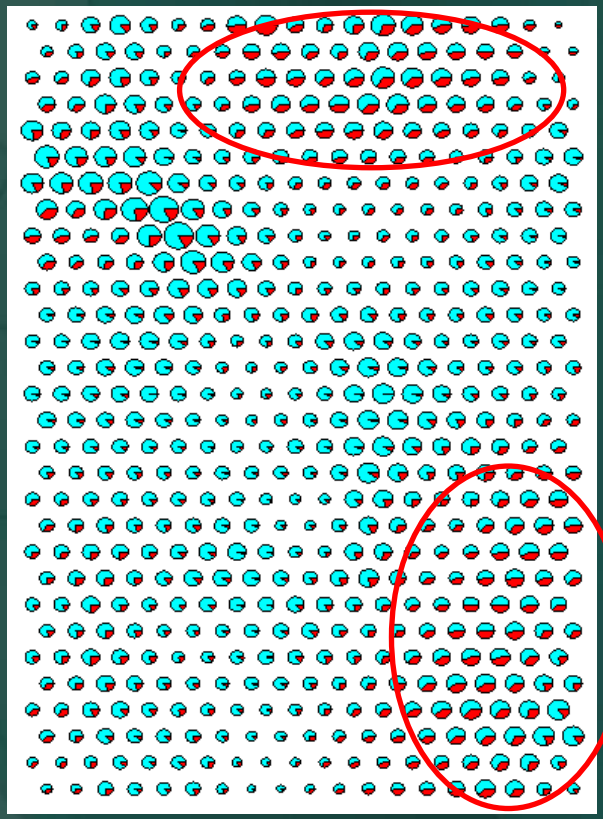
DIST_CONSTRURAIS



DIST_URB_EST



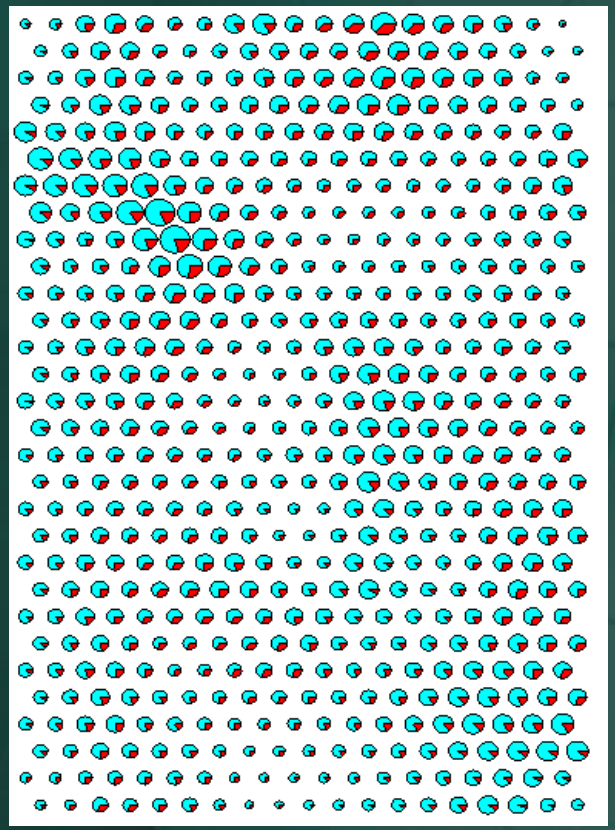
PNEL_LOG_SOMA



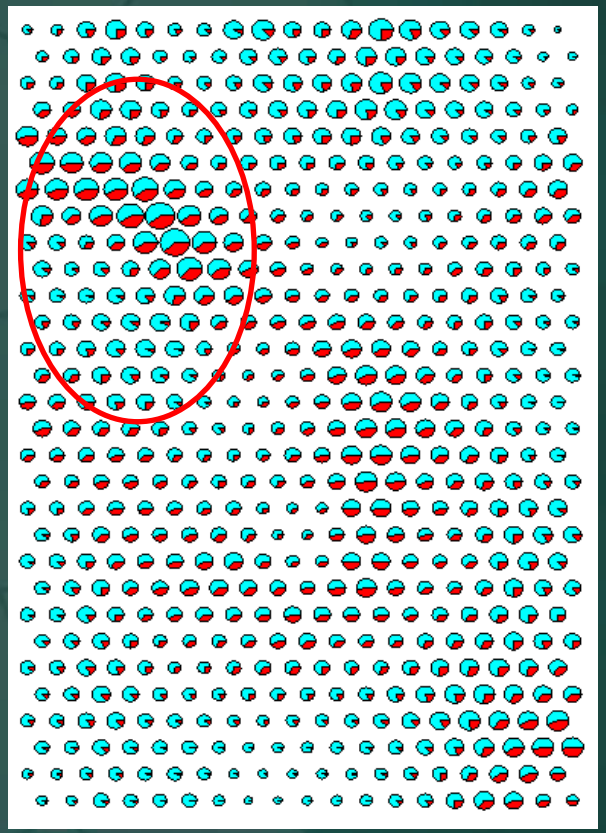
5. Resultados

5.2 Sub-Bacia

LAT

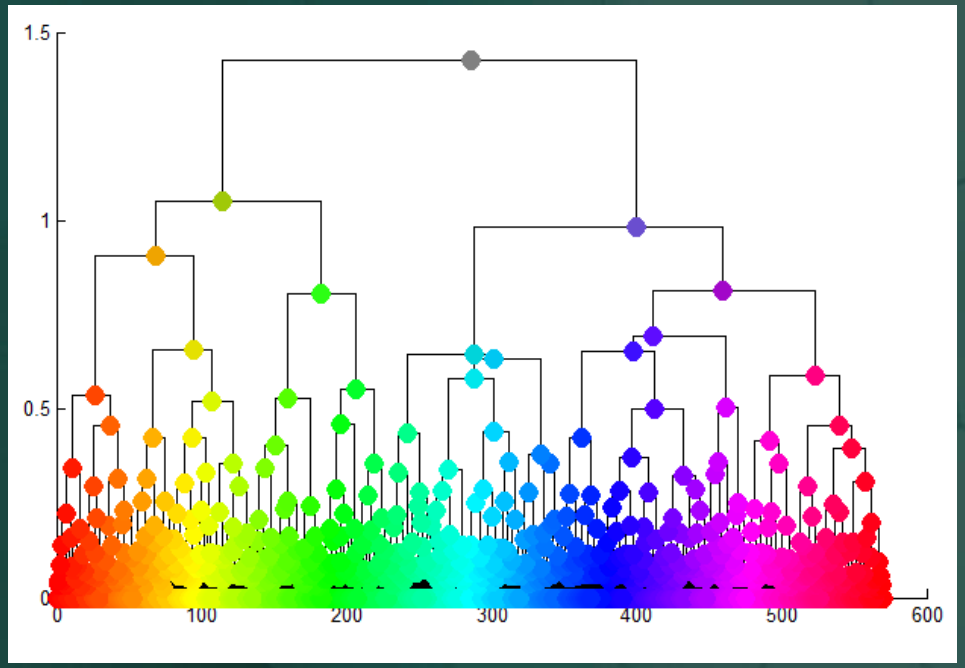


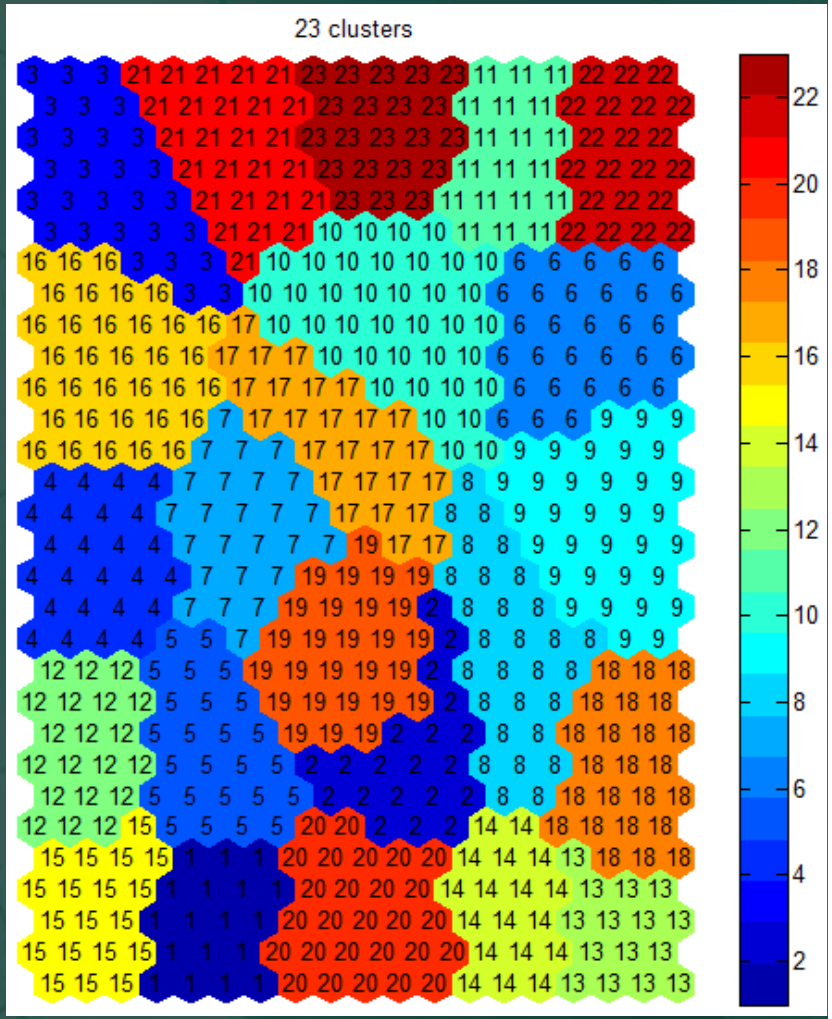
LONG



5. Resultados

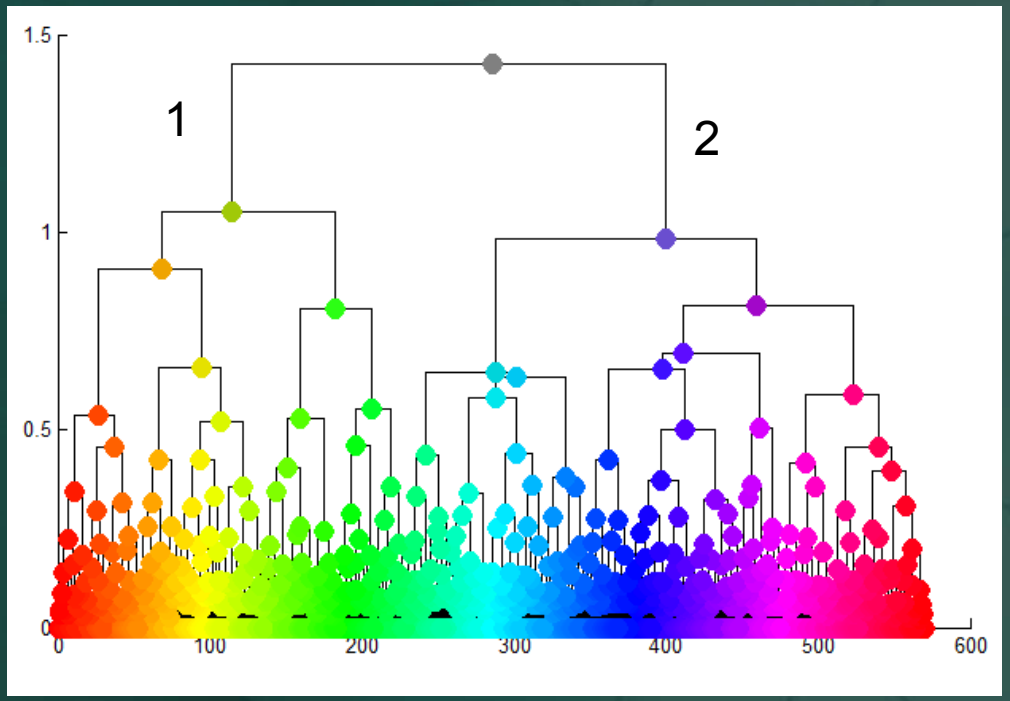
5.2 Sub-Bacia

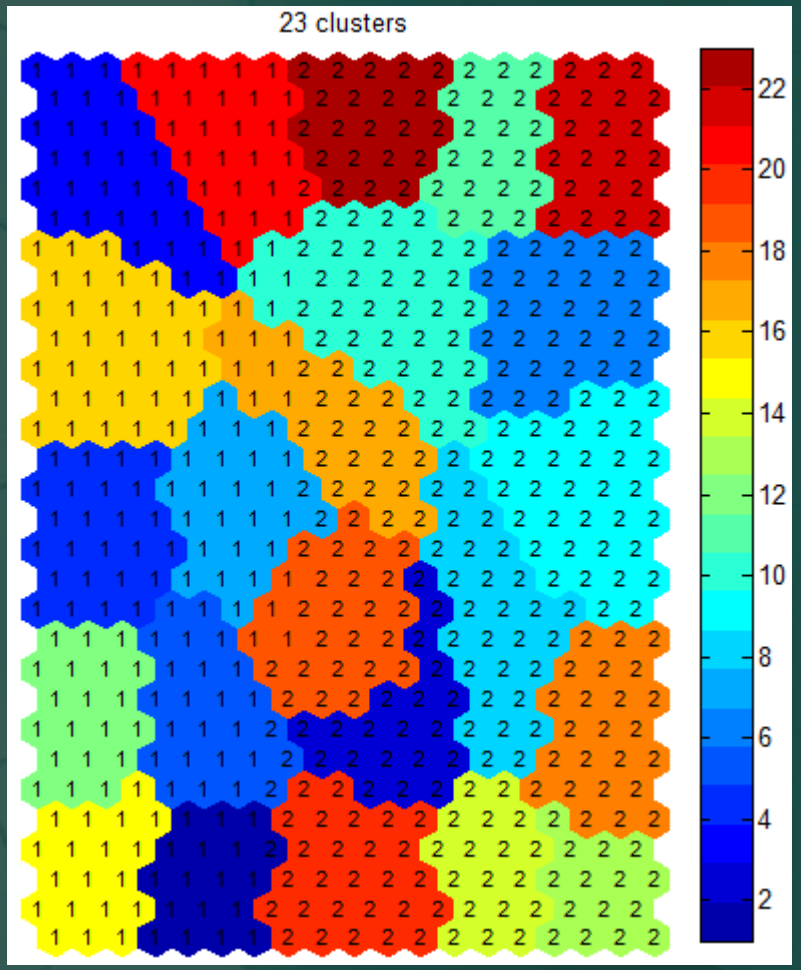




5. Resultados

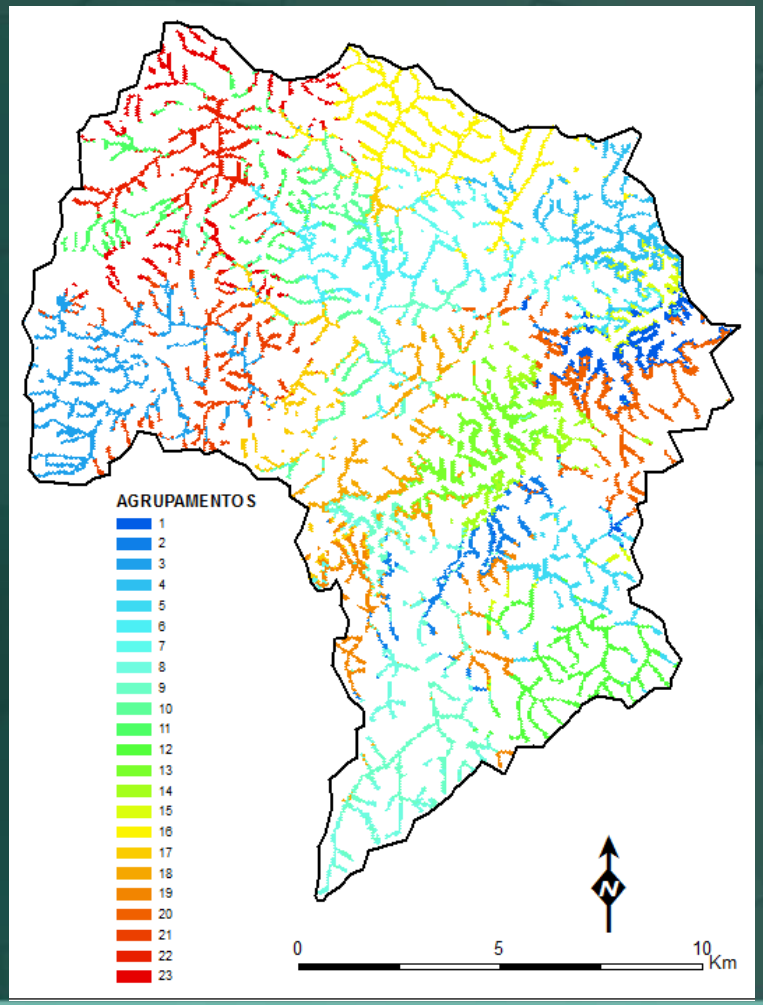
5.2 Sub-Bacia





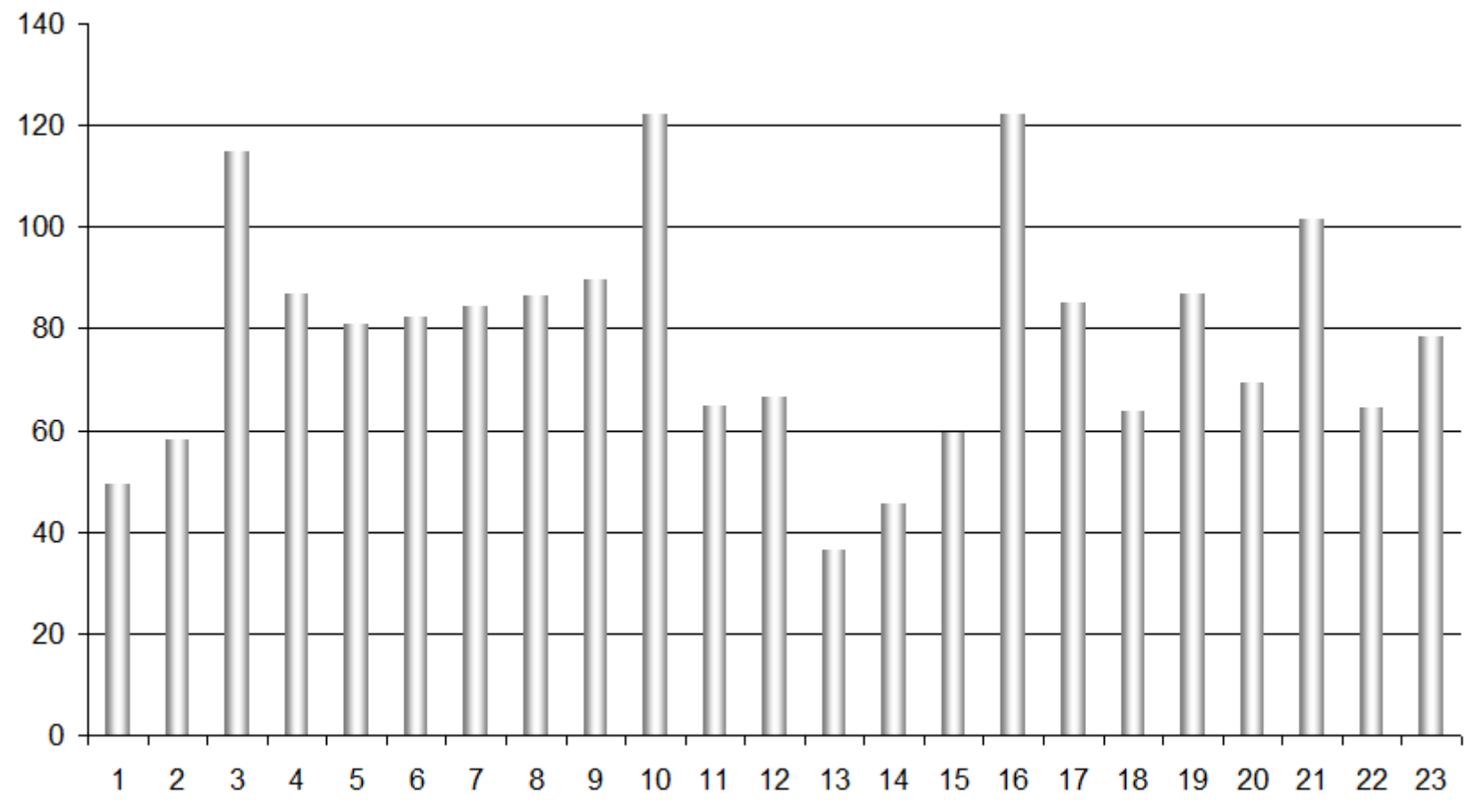
5. Resultados

5.2 Sub-Bacia



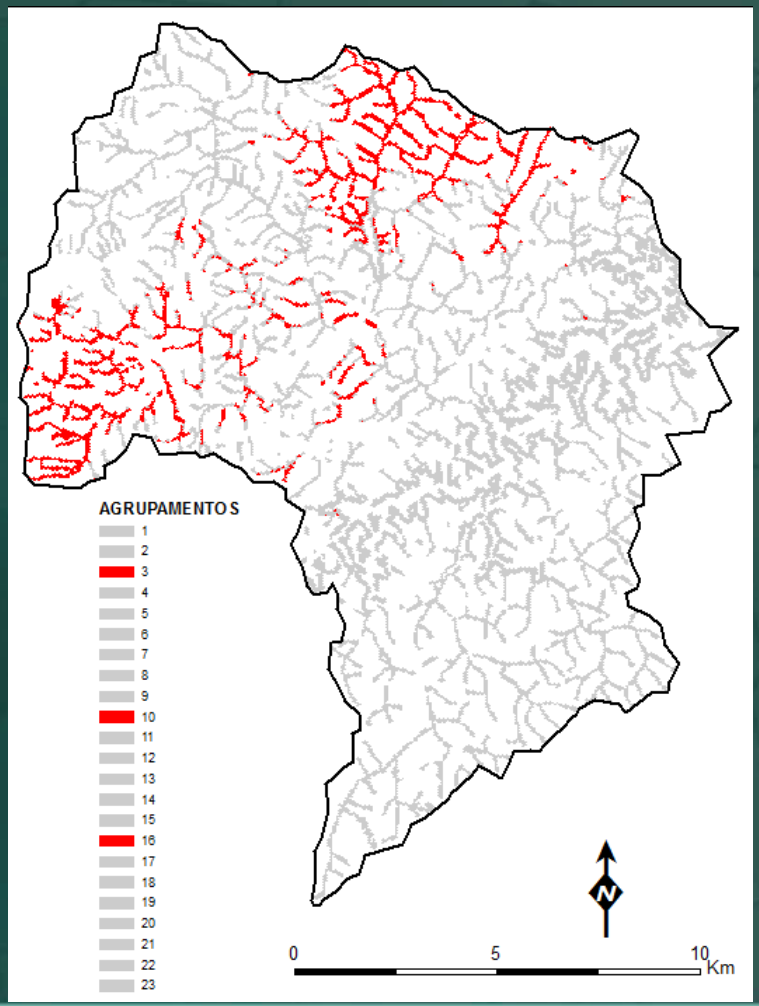
5. Resultados

5.2 Sub-Bacia



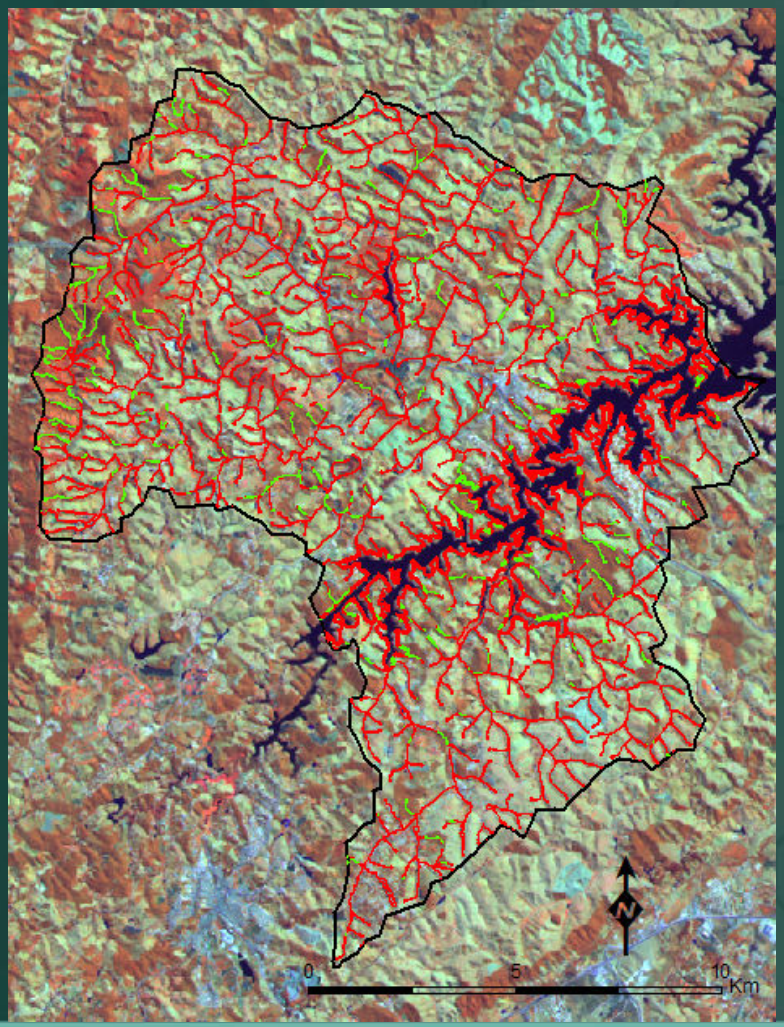
5. Resultados

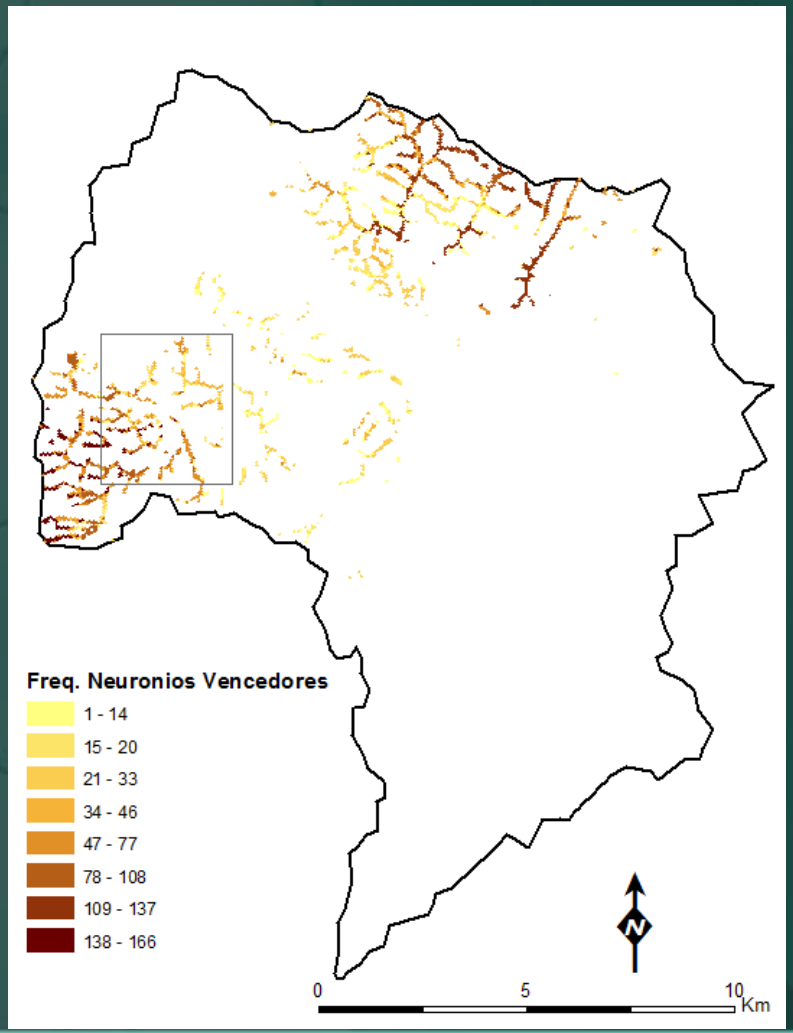
5.2 Sub-Bacia



5. Resultados

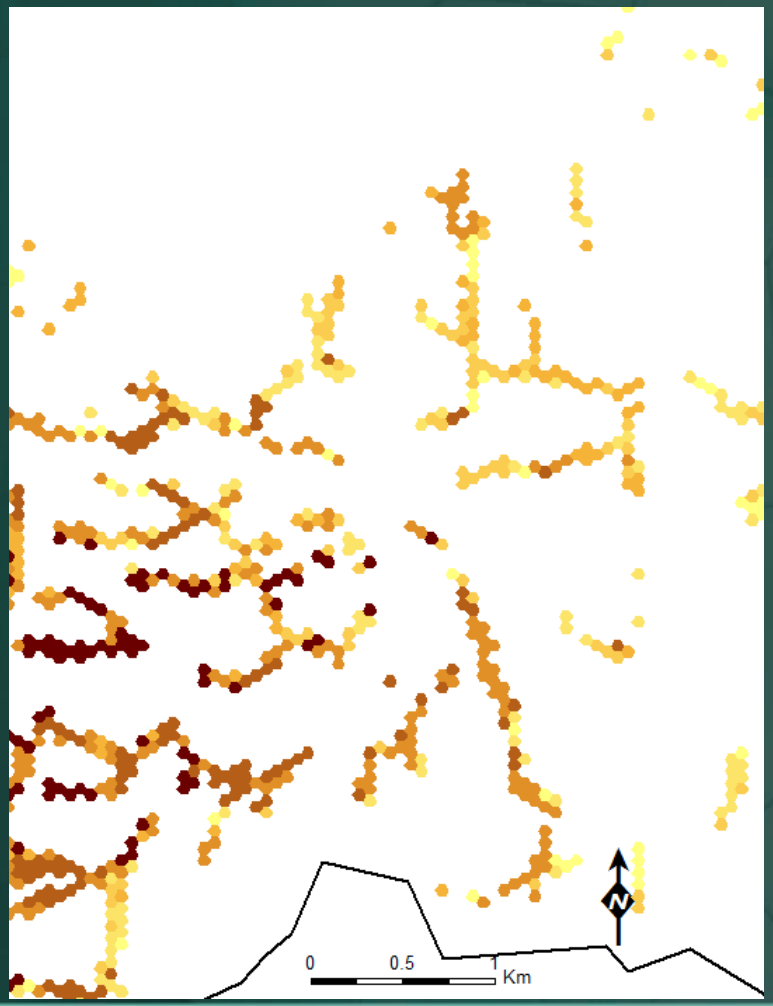
5.2 Sub-Bacia

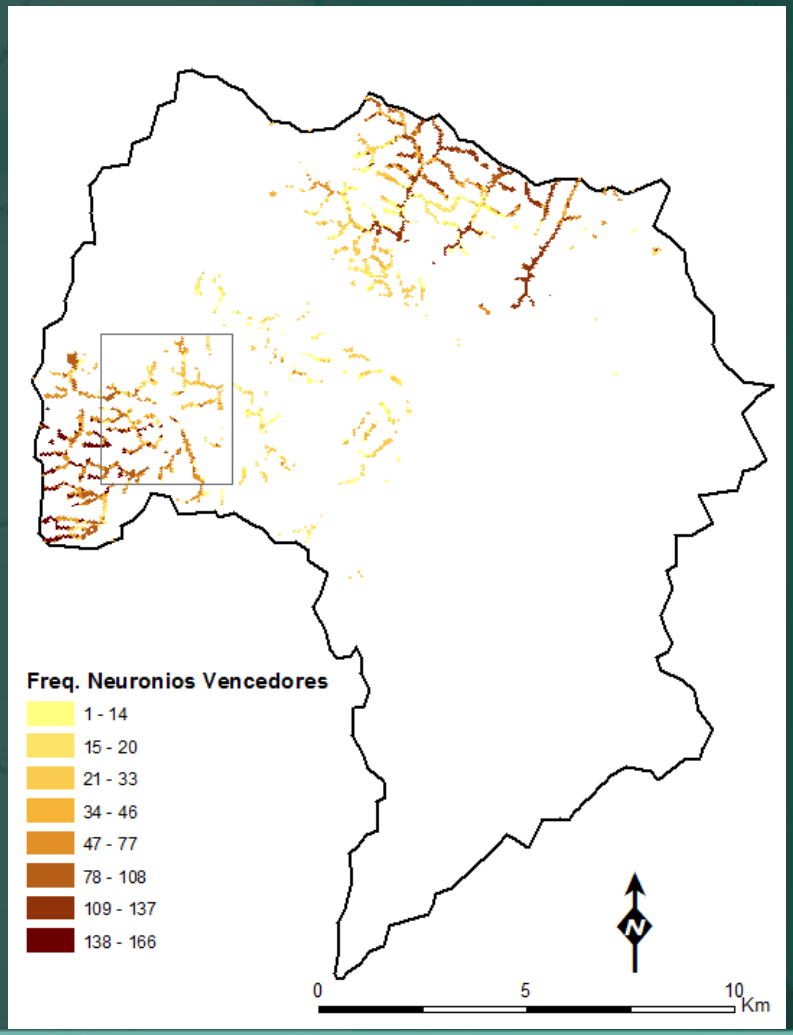




5. Resultados

5.2 Sub-Bacia





6. Considerações Finais

- As áreas identificadas para recuperação florestal são coerentes com a realidade;
- O SOM é uma ferramenta eficaz para a detecção de feições para a recuperação florestal;
- O Uso da componente espacial como variável é útil para visualizar correlação espacial;
- A frequencia de variáveis que tiveram o mesmo neurônio vencedor deve ser melhor investigada, mas há indicios de que a mesma retrata a “aderência” das variáveis utilizadas nos agrupamentos escolhidos;
- Estudos futuros se darão em relação à temporalidade das variáveis, e na quantificação e análises estatísticas dos dados gerados.

OBRIGADO!