IBGE
Instituto Brasileiro de Geografia e Estatística

Seminário de Metodologia do IBGE SMI 2016

CNEFE
Cadastro Nacional de Endereços para Fins Estatísticos

Integração e automação de processamentos espaciais nas atividades do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE)

9 de Novembro de 2016

Objetivo

Apresentar uma metodologia de trabalho e dois estudos de caso de como a utilização de ferramentas computacionais abertas para integração e automação de processamentos de dados espaciais ajudou a equipe do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) em seus trabalhos de planejamento e acompanhamento de atividades.

Breve histórico do CNEFE

- **Pré-2000**: Folhas de coleta do Censo Demográfico 1991 e Contagem da População 1996 são compiladas e validadas para utilização no Censo Demográfico 2000;
- **Pós-2000**: Folhas de coleta do Censo Demográfico 2000 são digitalizadas, formando um embrião do que viria a se tornar o CNEFE;
- **2004-2007**: O CNEFE é previsto nas metas do Plano Estratégico do IBGE desse triênio;
- **2010**: Constitui-se em ferramenta imprescindível a operação de coleta no Censo Demográfico 2010;
- **2011-Presente**: Principal repositório de endereços da instituição e para seleção de amostras com vistas às pesquisas domiciliares, com especial destaque para a Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua).

Aspectos evolutivos

- Grande evolução metodológica e tecnológica nas atividades e sistemas do CNEFE e também nas atividades que o utilizam enquanto insumo e/ou que possuam forte interface com o mesmo;
- Interação cada vez maior entre atividades afins dentro da instituição

  (Ex.: CNEFE x Base Territorial x PNAD Contínua);
- Significativa evolução das bases cartográficas operacionais censitárias a partir da adoção do conceito de Base Territorial contínua após 2010;
- Atualizações empreendidas pelo CNEFE e pela Base Territorial são atualmente concomitantes, paralelas e complementares entre si;
- Tornou-se extremamente necessário que uma série de procedimentos fossem adotados para garantir a integridade dos dados e a correta associação geográfica dos endereços do cadastro à Base Territorial.

↓

Integração de bases de dados e procedimentos

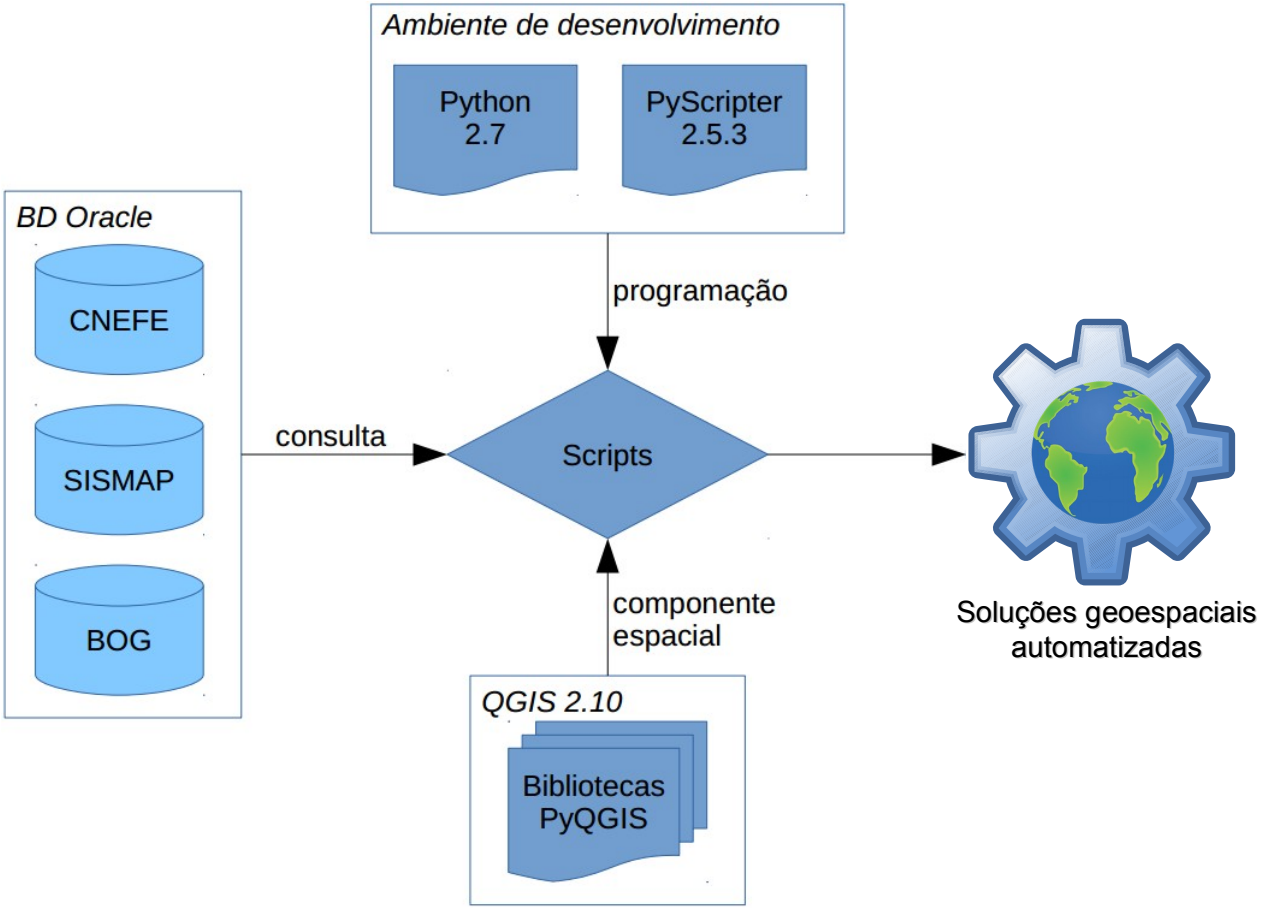
Metodologia

Insumos e materiais

- Bases de dados Oracle do CNEFE Universo (PR/COC), SISMAP (DGC/CETE) e BOG (DGC/CETE);
- Python 2.7 - linguagem de programação de alto nível e orientada a objetos que permite a integração de fontes de dados e a automação de procedimentos através de scripts;
- PyScripter 2.5.3.0 - ambiente de desenvolvimento integrado (IDE) utilizado;
- Bibliotecas PyQGIS (QGIS 2.10) - proporcionam inteligência geográfica a linguagem, permitindo que processamentos e aplicações sejam realizados através da linha de comando na interface gráfica do QGIS e também em modo *standalone* (fora do ambiente gráfico da aplicação), proporcionando uma liberdade na proposição de aplicações.

Metodologia

Modelo conceitual de automação de processos no CNEFE



Estudos de caso

- Transferência de endereços entre setores;
- Registro do trajeto do atualizador em campo.

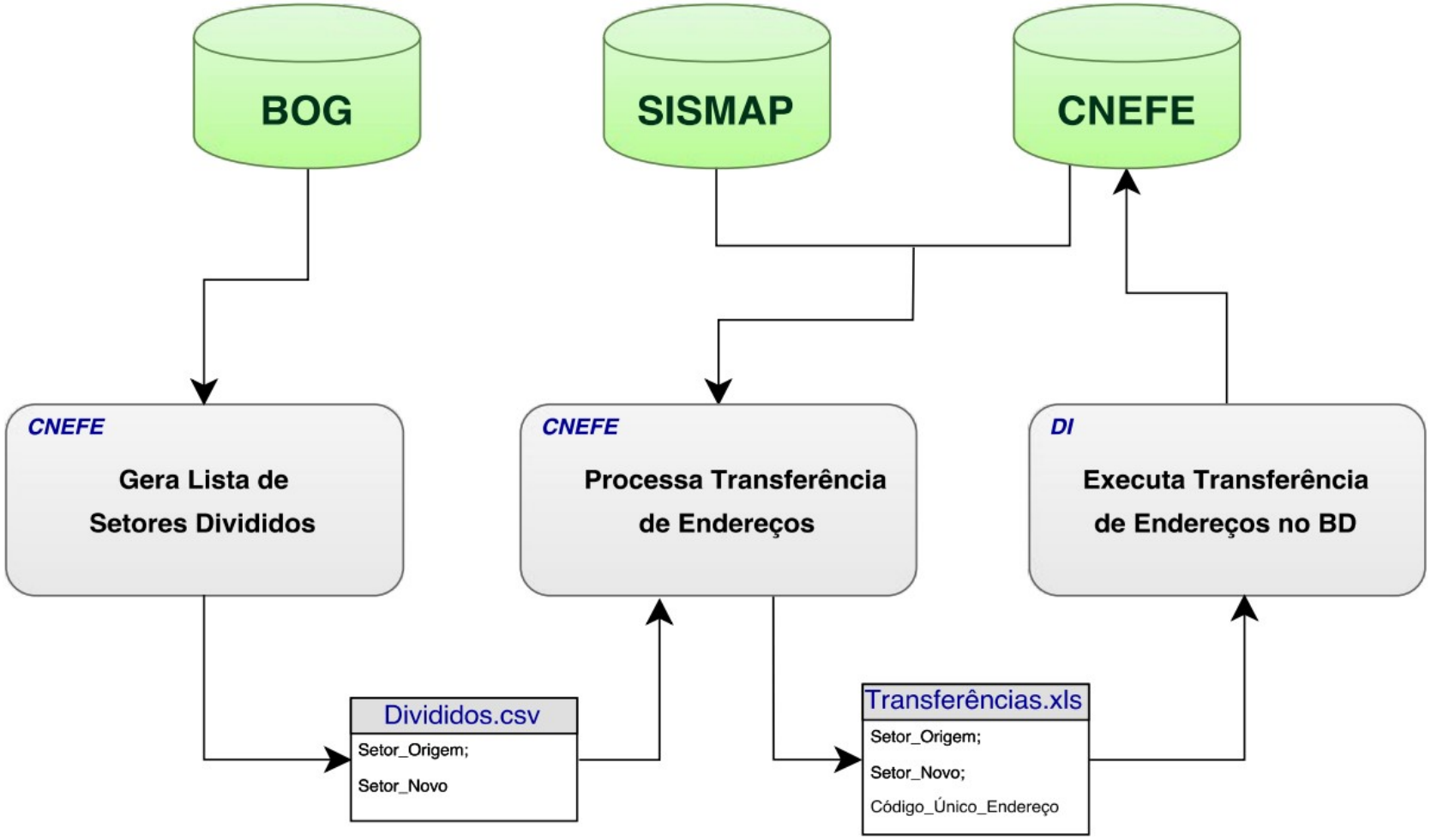
Estudos de caso

Transferência de endereços entre setores

- Cenário bastante comum atualmente;
- Visa a garantia da correta associação espacial do CNEFE, pois as geometrias das faces de quadra encontram-se no banco do SISMAP;
- Crescente número de divisões de setores oriundas da adoção de um planejamento contínuo da Base Territorial a partir de 2014;
- Processamento automático: mais eficiente frente a crescente demanda por transferência de endereços para suporte às atividades do CNEFE para a PNAD Contínua.

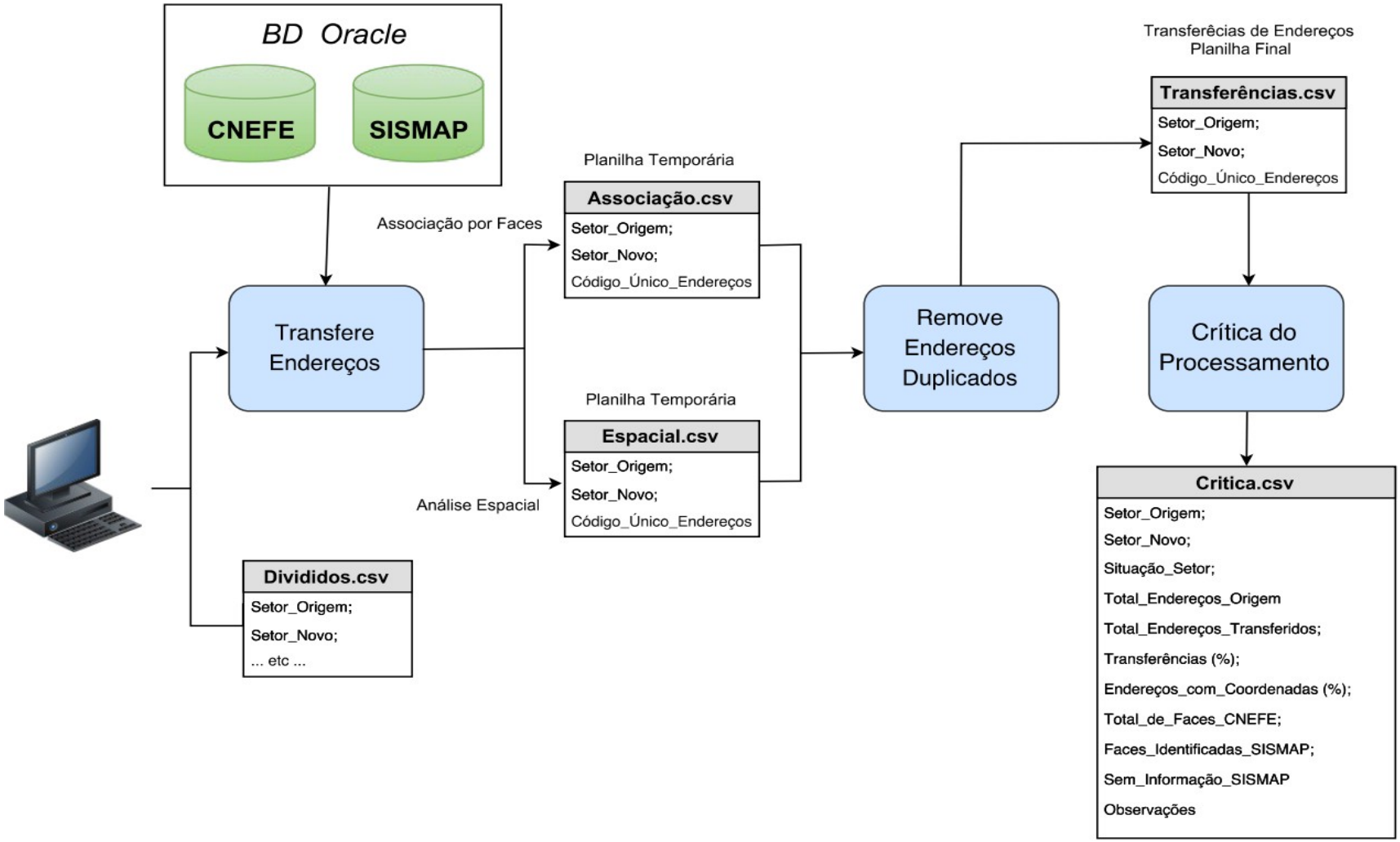
Transferência de endereços entre setores

Atualização de endereços em setores que foram divididos



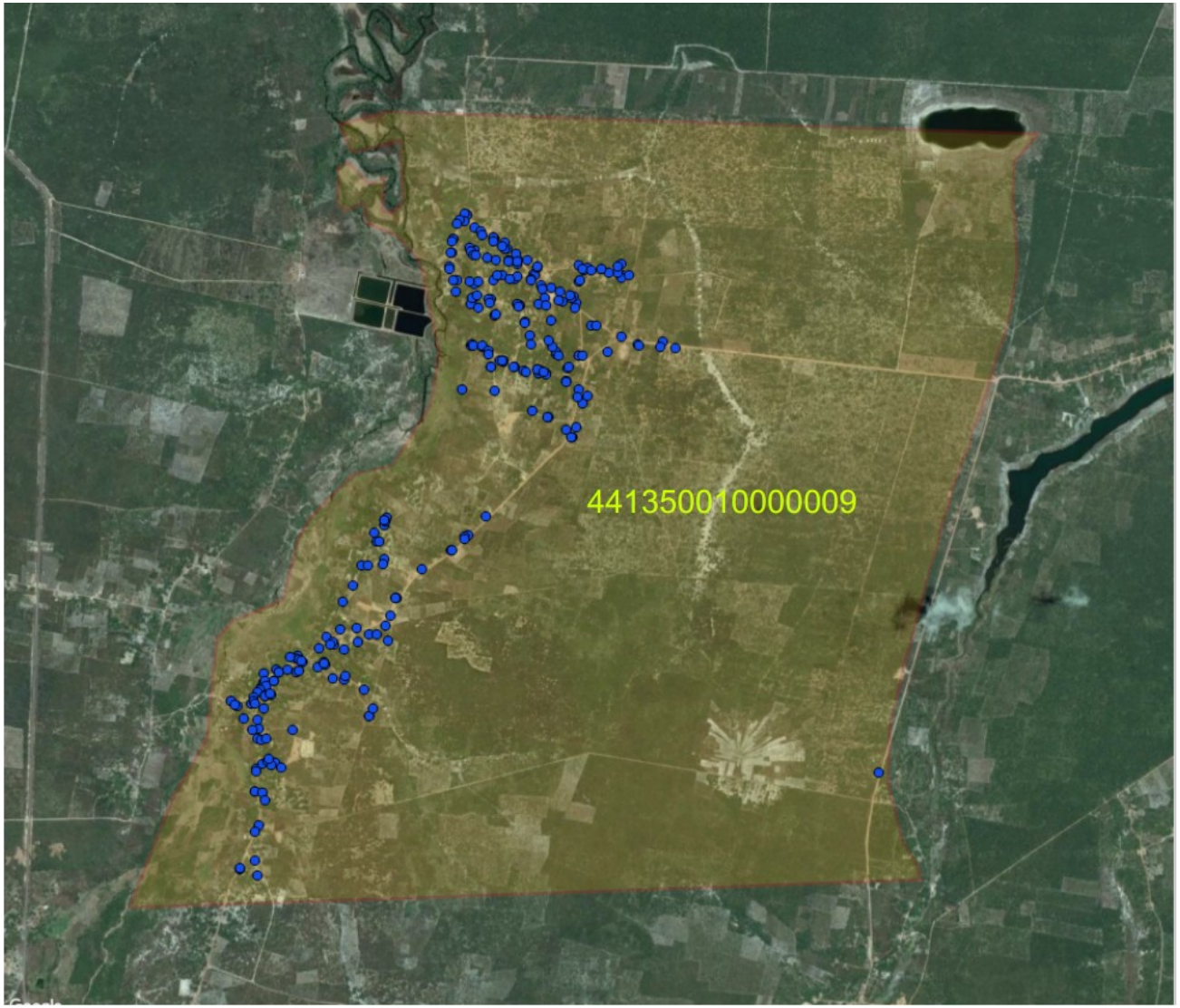
Transferência de endereços entre setores

Atualização de endereços em setores que foram divididos



Transferência de endereços entre setores

- Por coordenadas (análise espacial)

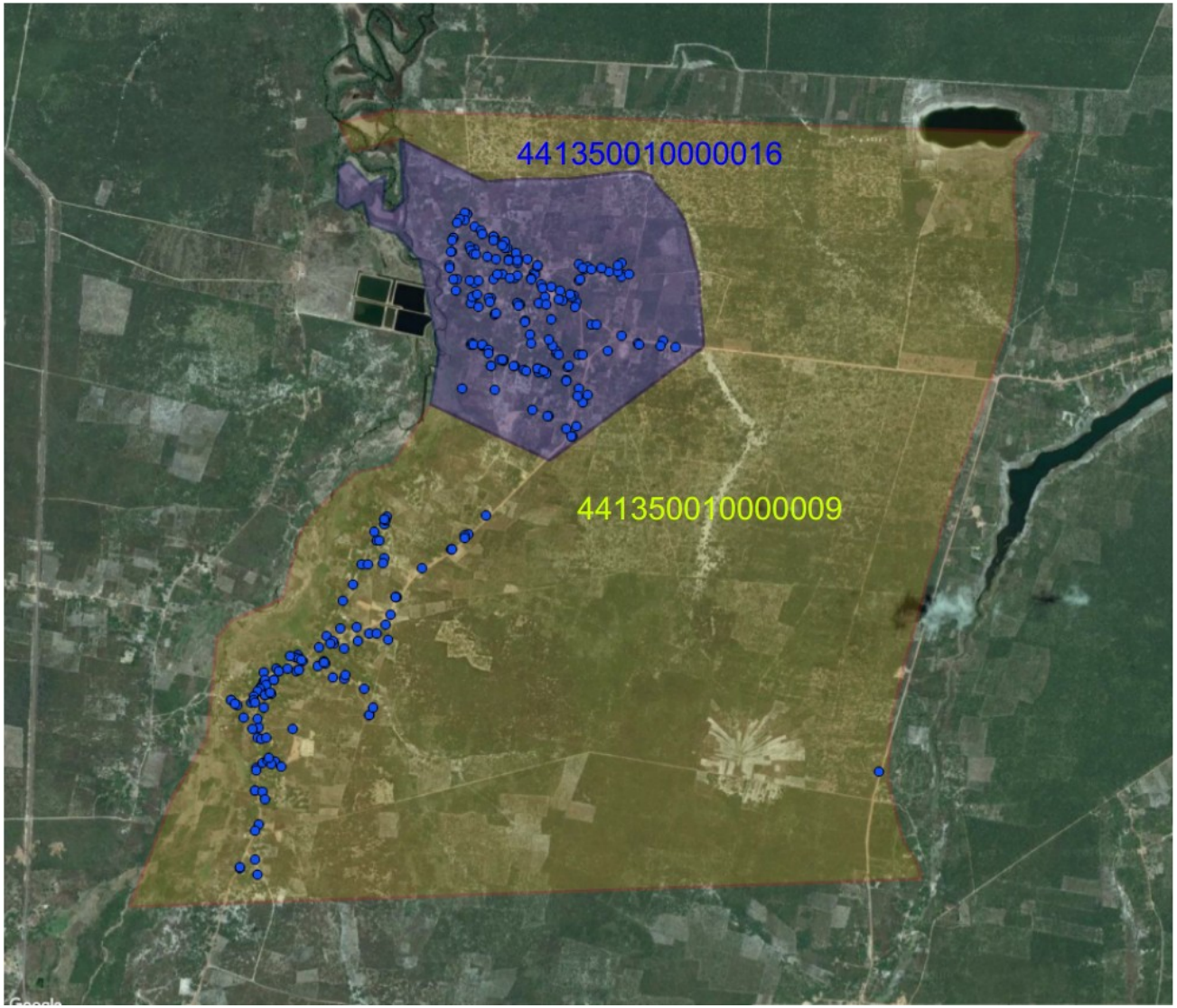


Endereços no CNEFE

Polígono doSetor no SISMAP

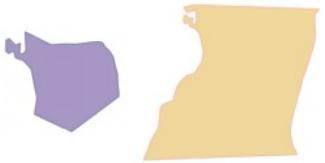
Transferência de endereços entre setores

- Por coordenadas (análise espacial)



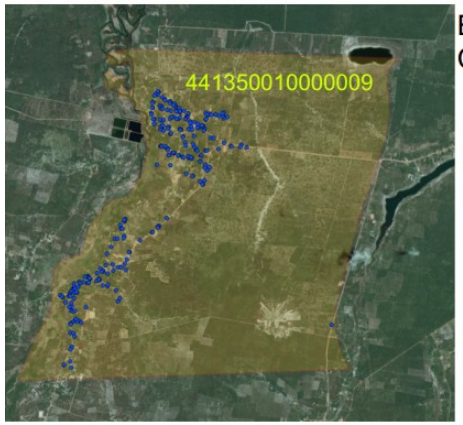
Endereços no CNEFE

Polígono doSetor
no SISMAP



Transferência de endereços entre setores

- Por coordenadas (análise espacial)



SISMAP

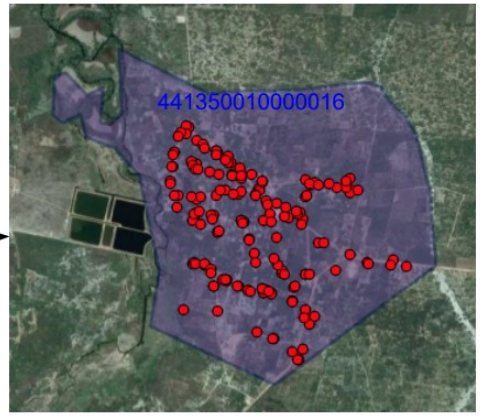
Contorno do Setor Novo

Endereços com Coordenadas. (Pontos)

CNEFE

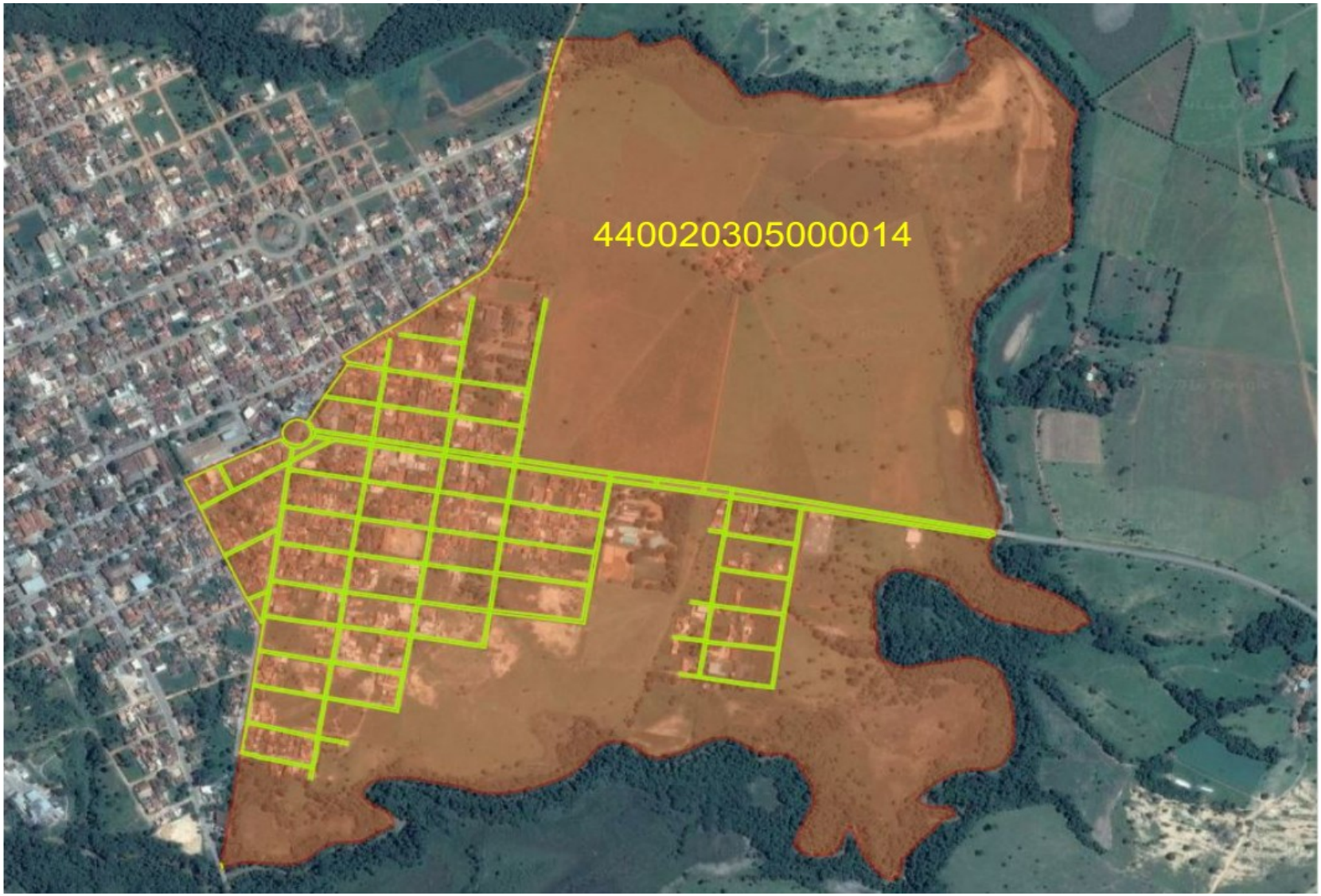
Análise Espacial (Interseção dos Pontos)

	Tabela_de_Transferências
	441350010000009; 441350010000016; Código_Único_Endereço



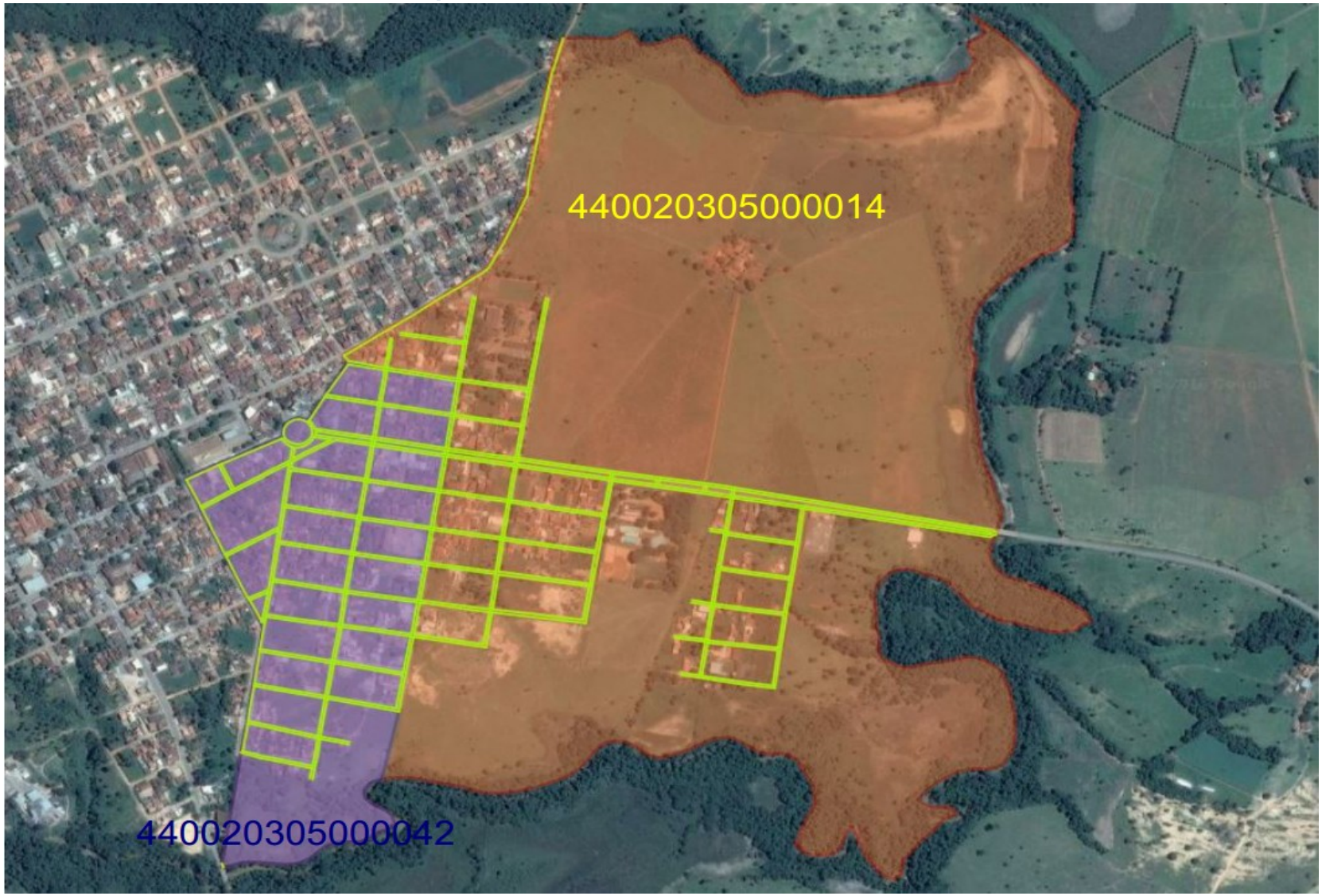
Transferência de endereços entre setores

- Por associação de quadra-face



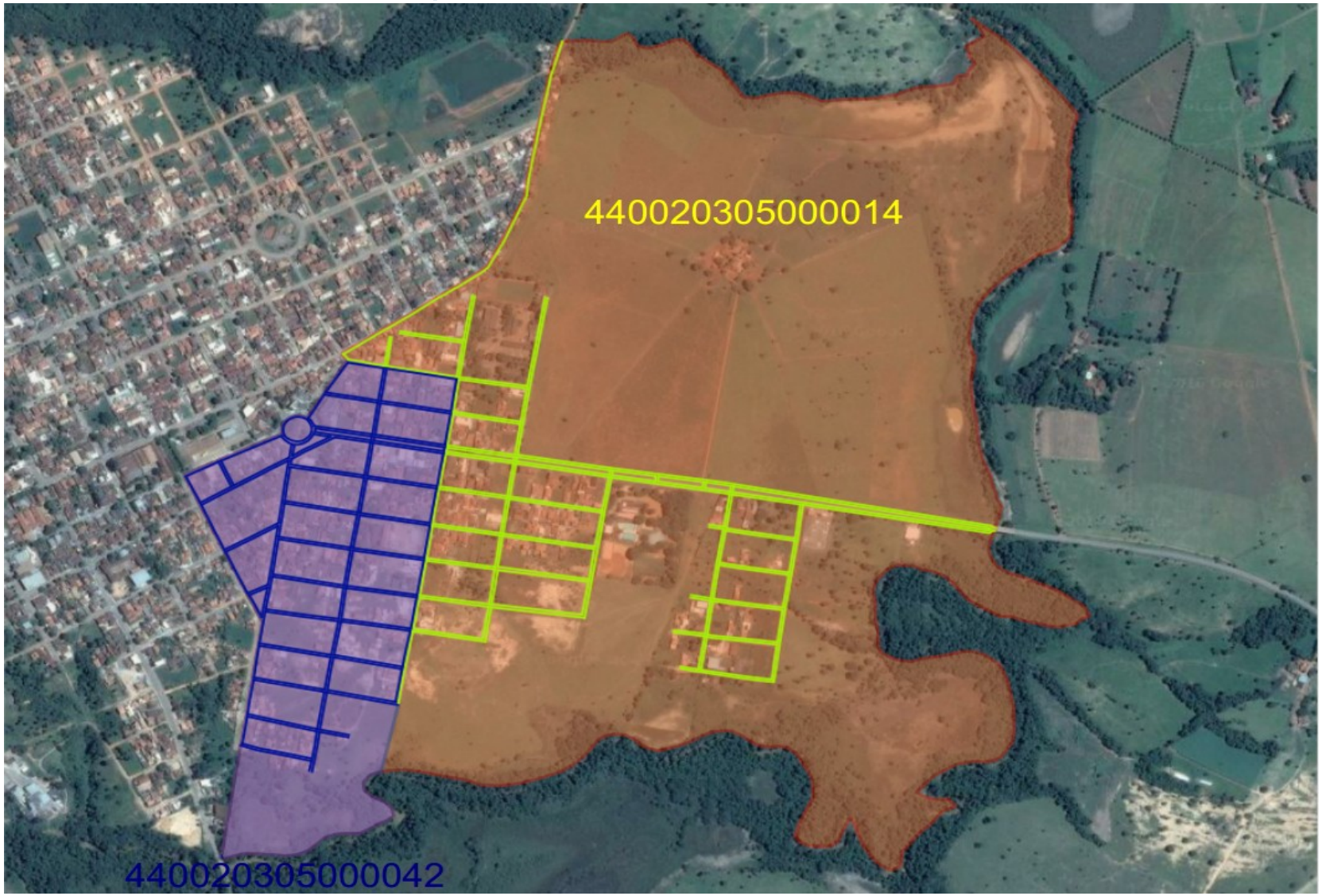
Transferência de endereços entre setores

- Por associação de quadra-face



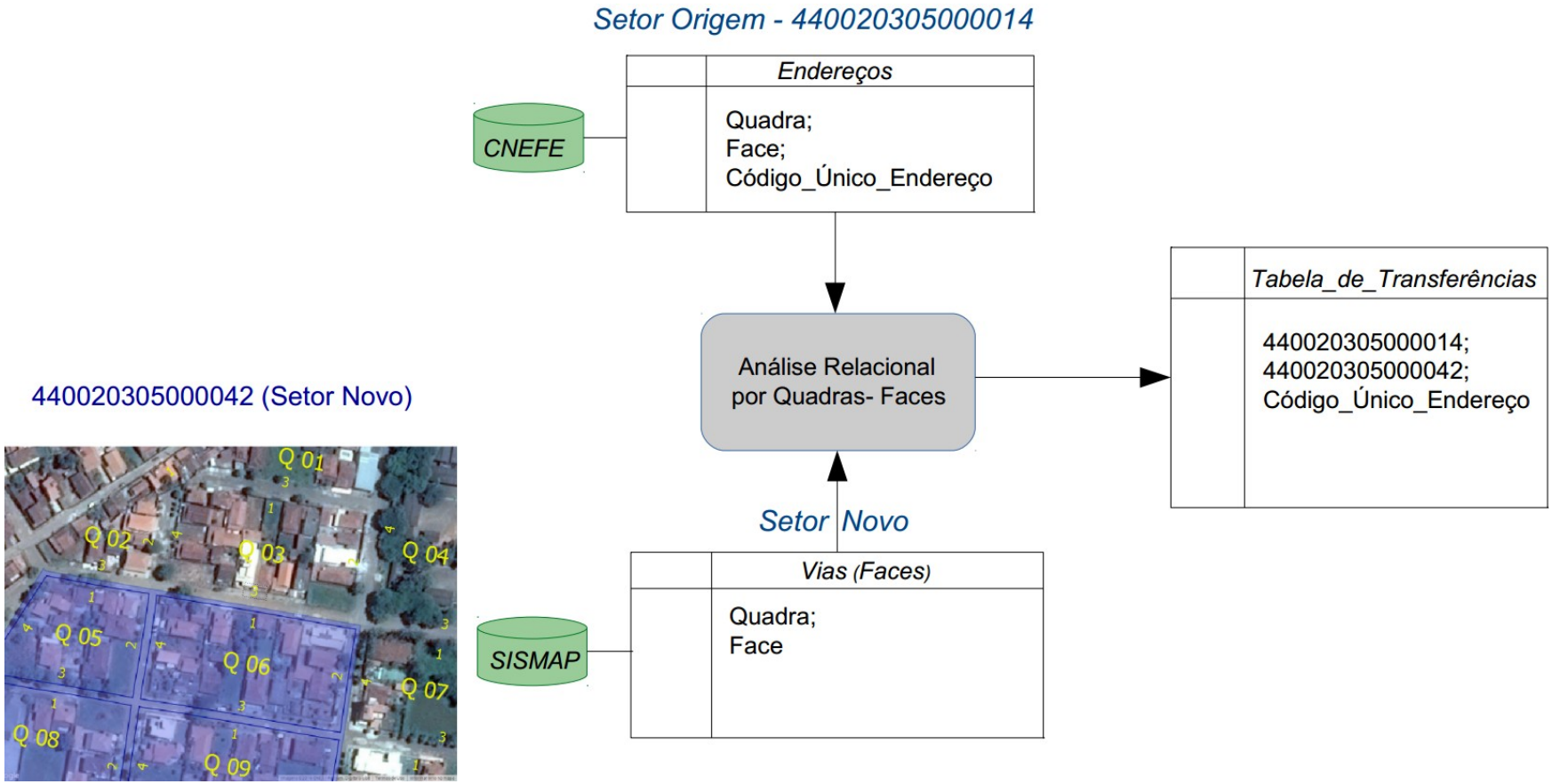
Transferência de endereços entre setores

- Por associação de quadra-face



Transferência de endereços entre setores

- Por associação de quadra-face



Transferência de endereços entre setores

Planilha Resumo do Processamento - Crítica

SETOR_NOVO	SETOR_ORIGEM	Situação	Endereços no ORIGEM	Transferidos	% Transferido	% Coordenadas	Faltas no SISMAP	OBSERVAÇÕES
314070405000038	314070405000010	Urbano	338	0	0	0	Faces não Associadas	
314070405000039	314070405000010	Urbano	338	0	0	0	Faces não Associadas	
314070405000040	314070405000027	Rural	280	1	0	100		17 Já Transferidos no CENEFEUNIV
314070405000041	314070405000035	Urbano	216	15	7	0		
314070405000042	314070405000007	Urbano	467	0	0	0	Faces não Associadas	
314070405000044	314070405000013	Urbano	250	112	45	0		
314070405000045	314070405000018	Urbano	237	23	10	0		
314070405000046	314070405000020	Urbano	290	0	0	0		
314080305000031	314080305000011	Urbano	434	221	51	0		
314080305000032	314080305000012	Urbano	223	0	0	17	Faces não Associadas	
314080305000033	314080305000016	Rural	176	24	14	100		
314080305000034	314080305000025	Urbano	327	10	3	76		107 Já Transferidos no CENEFEUNIV
314080305000035	314080305000007	Urbano	217	48	22	0		
314080305000036	314080305000007	Urbano	217	50	23	0		
314080305000038	314080305000003	Urbano	308	1	0	0		
314080305000039	314080305000004	Urbano	227	0	0	0	Faces não Associadas	
314080305000040	314080305000028	Urbano	235	49	21	0		
314085205000014	314085205000003	Urbano	395	236	60	0		
314085205000015	314085205000007	Rural	358	82	23	97		
314085205000017	314085205000008	Rural	273	0	0	100		
314085205000018	314085205000008	Rural	273	84	31	100		

TOTAIS:	21		6.079	956	15,73%

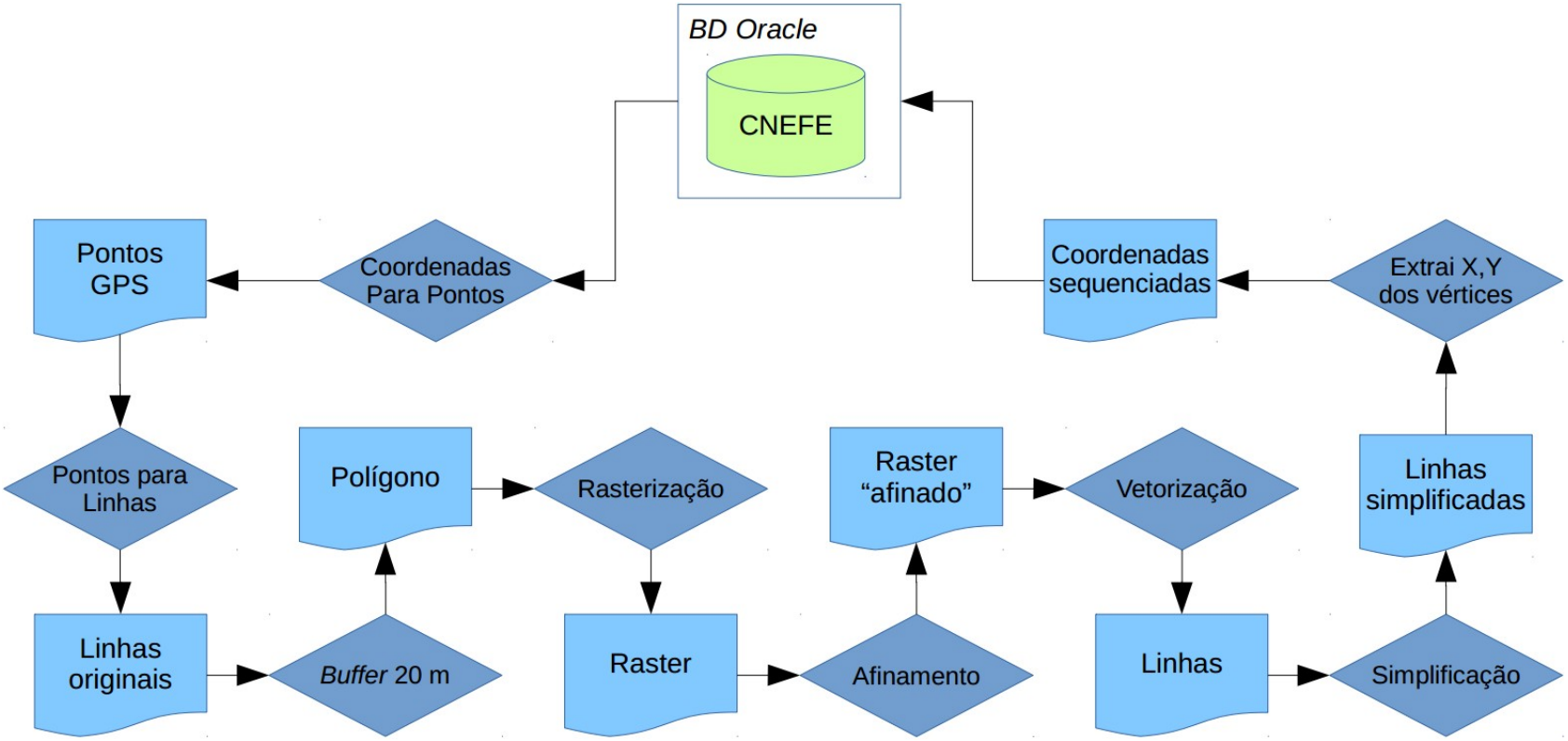
Estudos de caso

Registro do trajeto do atualizador em campo

- Dados de GPS coletados em campo de forma contínua durante a atualização do cadastro de endereços em um setor rural;
- São capturados para fins de supervisão do trabalho em campo e posterior atualização dos insumos para a próxima atualização a ser realizada naquele mesmo setor;
- O atualizador pode passar em um mesmo logradouro (estrada, rio etc.) mais de uma vez durante a atividade, o que gera sobreposição de dados coletados;
- Devido ao erro do GPS (entre 5 m e 10 m), linhas sobrepostas e/ou paralelas são geradas a partir de percursos repetidos em campo;
- Uma estrada deve ser representada somente por uma única linha: necessidade de uma “limpeza” das feições originais.

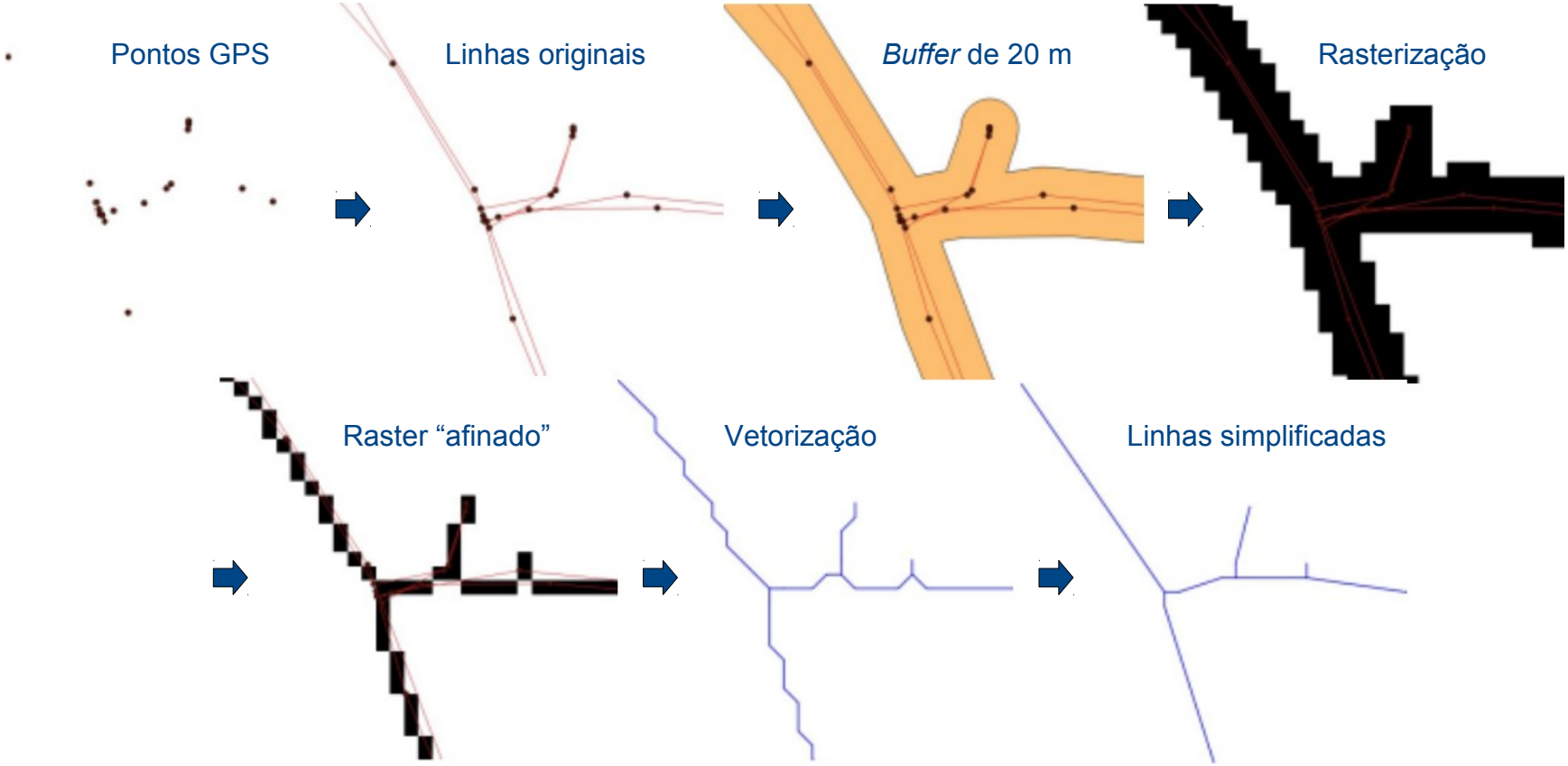
Registro do trajeto do atualizador em campo

Fluxograma do processo de simplificação dos trajetos



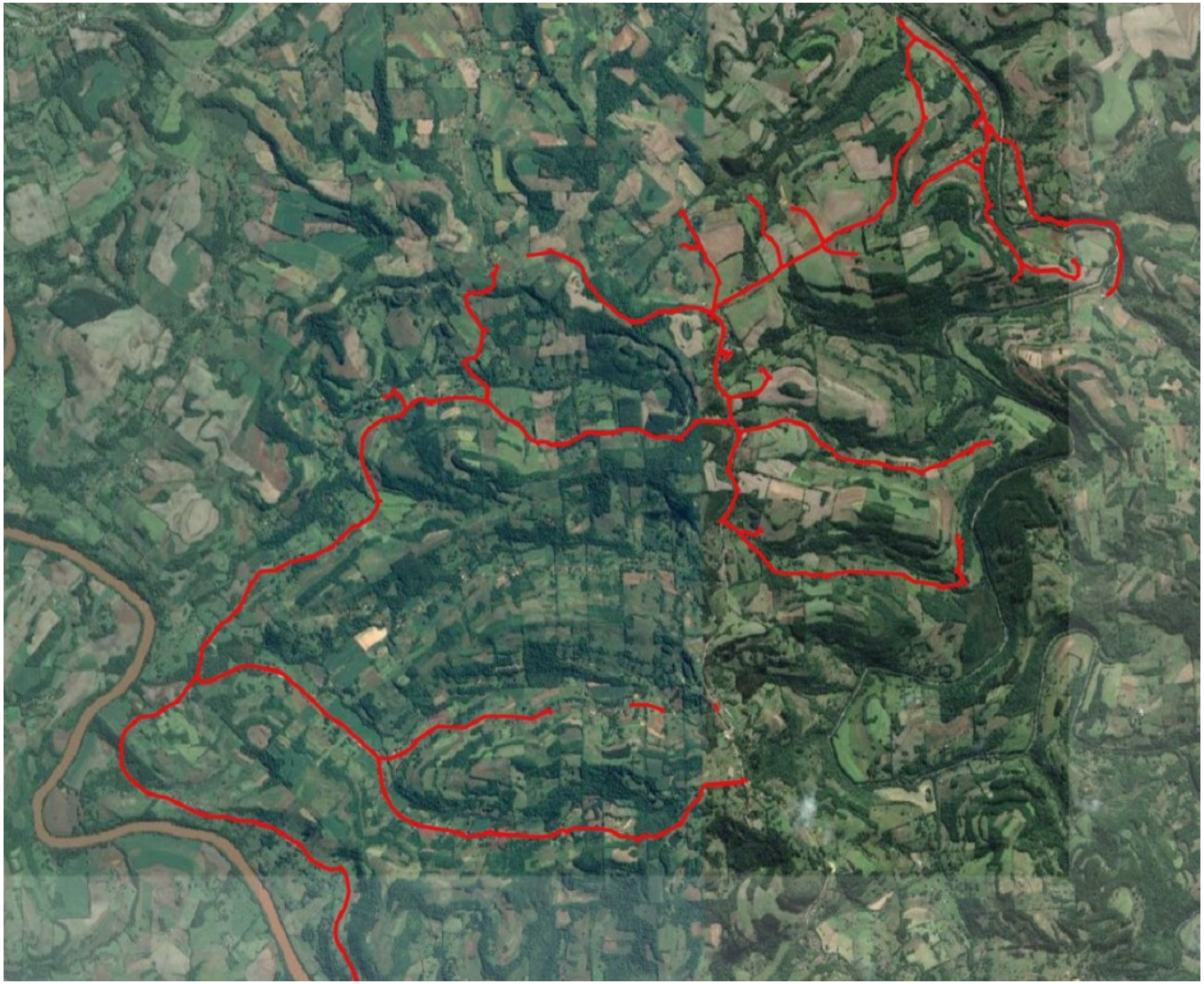
Registro do trajeto do atualizador em campo

Estágios do procedimento de “limpeza” de trajetos



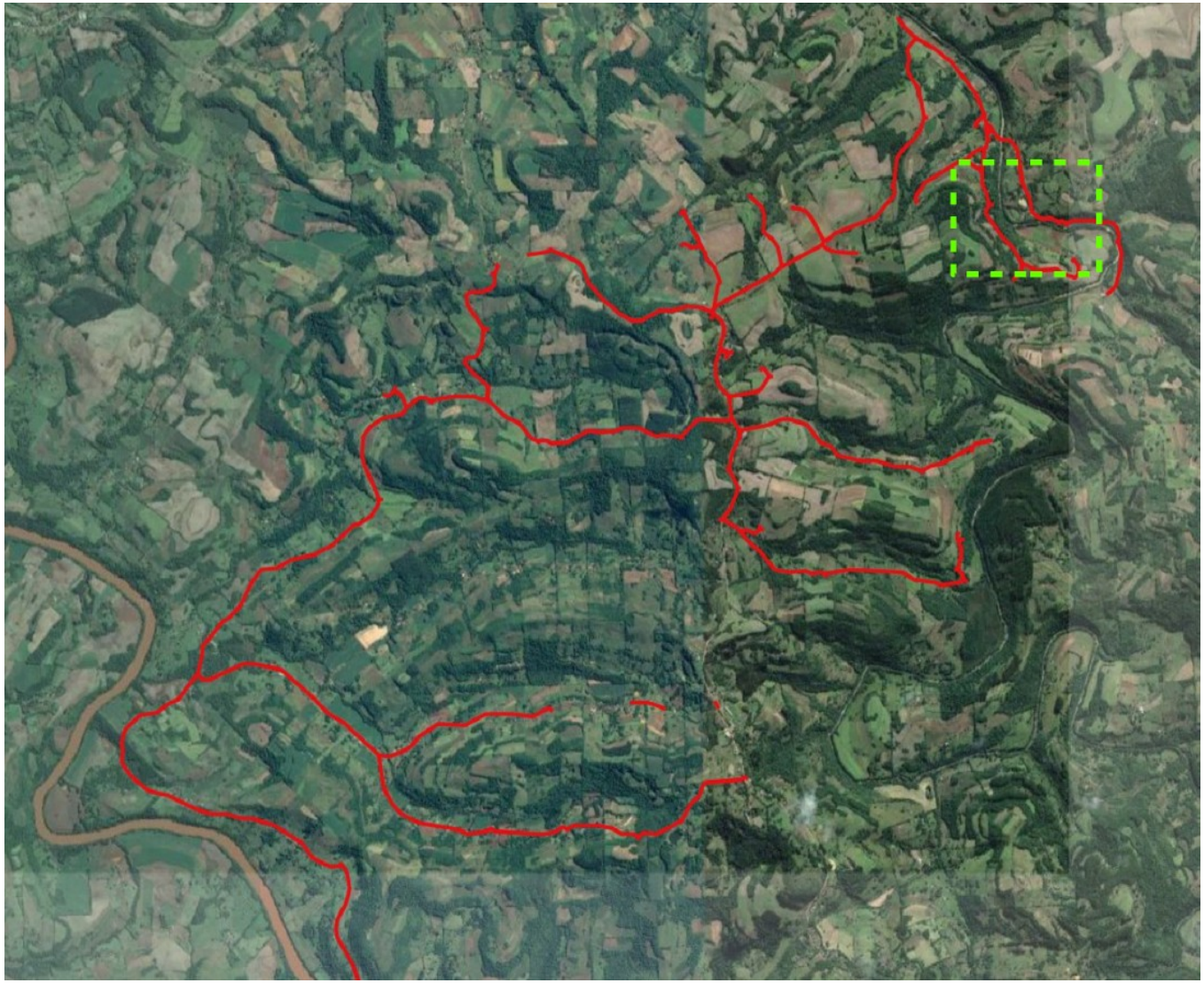
Registro do trajeto do atualizador em campo

Exemplo: setor 431050405000013 (Iraí-RS)



Registro do trajeto do atualizador em campo

Exemplo: setor 431050405000013 (Iraí-RS)

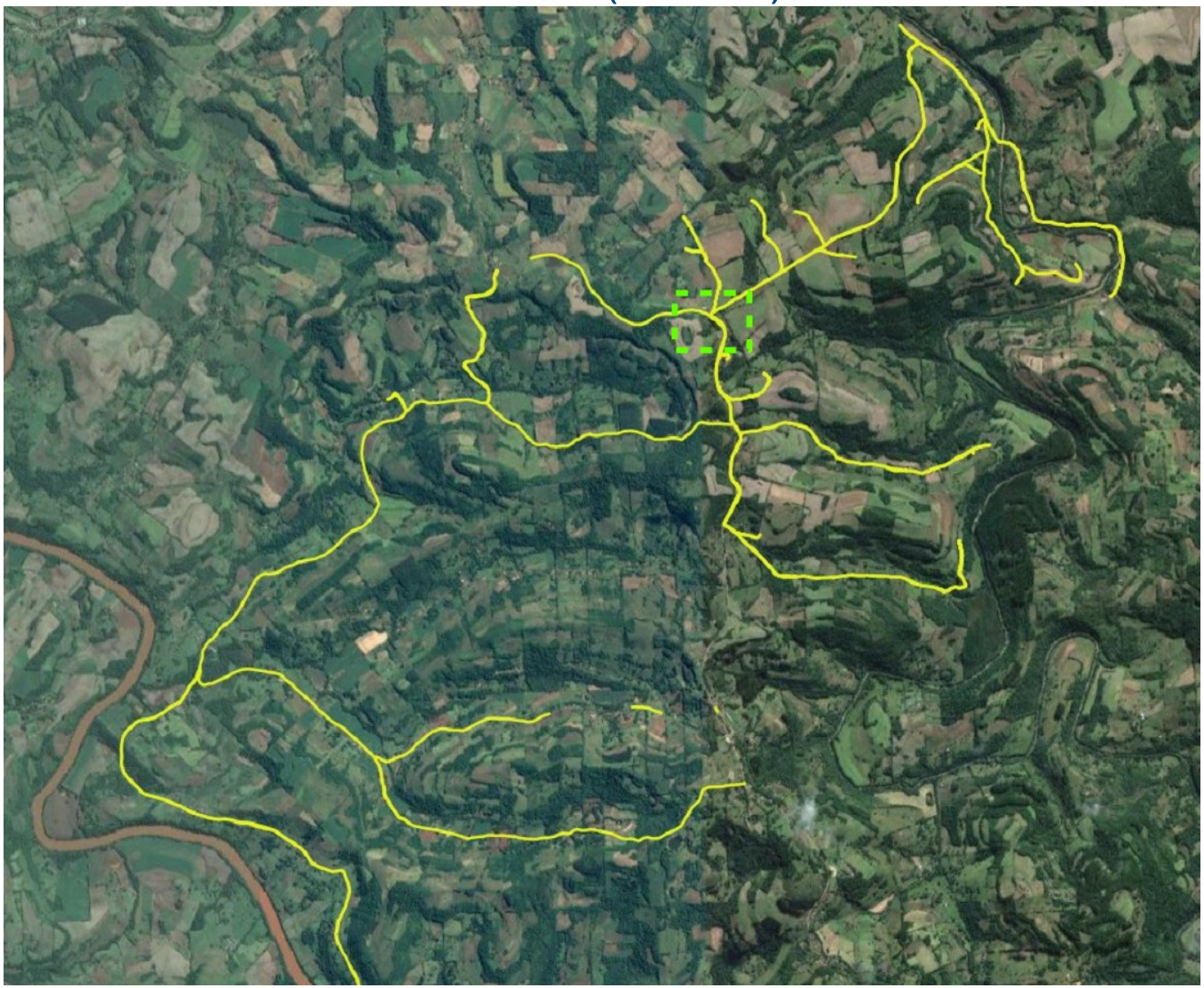


Registro do trajeto do atualizador em campo

Exemplo: setor 431050405000013 (Iraí-RS)

Registro do trajeto do atualizador em campo

Exemplo: setor 431050405000013 (Iraí-RS)



Registro do trajeto do atualizador em campo

Exemplo: setor 431050405000013 (Iraí-RS)

Registro do trajeto do atualizador em campo

Exemplo: setor 431050405000013 (Iraí-RS)

Considerações

- A experiência recente da adoção desta metodologia na automação de consultas e análises espaciais por parte da equipe do CNEFE tem se mostrado bem-sucedida;
- O aumento da qualidade e confiabilidade dos resultados foram significativos, assim como os ganhos no tempo de processamento;
- Confere celeridade às demandas oriundas das atividades de campo, uma vez que são os principais consumidores dos resultados de tais procedimentos;
- A linguagem Python como plataforma integradora para automação de análises espaciais tem se demonstrado bastante robusta e, ao mesmo tempo, de fácil aprendizado através de uma sintaxe simples e intuitiva.

Desdobramentos

- Ampliação do número de procedimentos automatizados no CNEFE através da continuidade do processo de exploração e descoberta de novas ferramentas;
- Teste e potencial adoção de outras bibliotecas Python:
  - para acesso nativo e otimizado às bases Oracle (*cx_oracle*);
  - para estruturação, manipulação e análise de dados em alta performance (*pandas*);
- Disseminação de metodologias e troca de experiências com uso do PyQGIS e outras arquiteturas de 'geoprogramação' entre coordenações ou servidores que estejam desenvolvendo soluções semelhantes no IBGE;
- Possibilidade de padronização de procedimentos entre coordenações que possuam atividades afins ou intimamente correlatas;
- Ampliação da cultura de utilização de geotecnologias abertas no IBGE.

Referências

Goodchild, M. (1992) Geographical Information Science. *International Journal of Geographical Information Systems*, 6:1, 31-45.

IBGE (2016) *Metodologia do Censo Demográfico 2010*. 2ª ed. Rio de Janeiro, 720 p.

Lawhead, J. (2015) *QGIS Python Programming Cookbook*. Packt Publ. Birmingham, Inglaterra.

Sherman, G. (2014) *PyQGIS Programmer's Guide*. Locate Press. Chugiak, EUA.

Contatos

Eduardo Luís Teixeira Baptista
Analista de Geoprocessamento, PR/COC/CNEFE, e-mail: *eduardo.baptista@ibge.gov.br* Tel.: 2142-3520

Fernando Ramalho Gameleira Soares
Analista de Geoprocessamento, PR/COC/CNEFE, e-mail: *fernando.soares@ibge.gov.br* Tel.: 2142-3506

Nilson Antônio de Castro Regal
Analista de Geoprocessamento, PR/COC/CNEFE, e-mail: *nilson.regal@ibge.gov.br* Tel.: 2142-3522

Wolney Cogoy de Menezes
Coordenador Nacional do CNEFE, PR/COC/CNEFE, e-mail: *wolney.menezes@ibge.gov.br* Tel.: 2142-4742