LEVERAGING STREET LEVEL PREDICTIVE MODELLING WITH GREEN INFRASTRUCTURE FOR URBAN RESILIENCE

Shawn Tew 5925673
Msc Geomatics | TU Delft

CONTENT

- **Introduction**
- **Related work**
- **Research Objectives**
- **Methodology**
- **Results**
- **Next Steps**

TOPIC RELEVANCE

Problem statement

- Evaluating localised green infrastructure impact is complex
- Lack of standardised methods to evaluate green infrastructure effectiveness across different urban contexts

Importance of the topic

- Green infrastructure is a proven tool for enhancing urban climate resilience
- Data driven GI planning is essential for long-term sustainable cities

Thesis contribution

- Aims to develop a modular workflow for green infrastructure analysis in urban systems
- Supports planners in prioritising green interventions where they are most effective

MAIN RESEARCH QUESTION

"How can predictive modelling techniques be used to evaluate the effectiveness of GI in improving urban resilience at street level across different dimensions of the built environment"

SUB RESEARCH QUESTION 1

What aspects of GI can be assessed in a multifunctional way to evaluate their contributions to urban resilience across multiple dimensions?

SUB RESEARCH QUESTION 2

How can environmental, biodiversity, and morphological features be combined to comprehensively assess and enhance urban resilience?

SUB RESEARCH QUESTION 3

How can predictive modelling techniques, such as machine learning, be used to evaluate and optimise GI for improving urban resilience?

SUB RESEARCH QUESTION 4

How can the presence of spatial autocorrelation in the residuals affect the reliability of the model's predictions?

SUB RESEARCH QUESTION 5

How can the proposed framework be applied to assess and prioritise GI interventions in Rotterdam?

ROTTERDAM AS A CASE STUDY

Environmental Challenges:

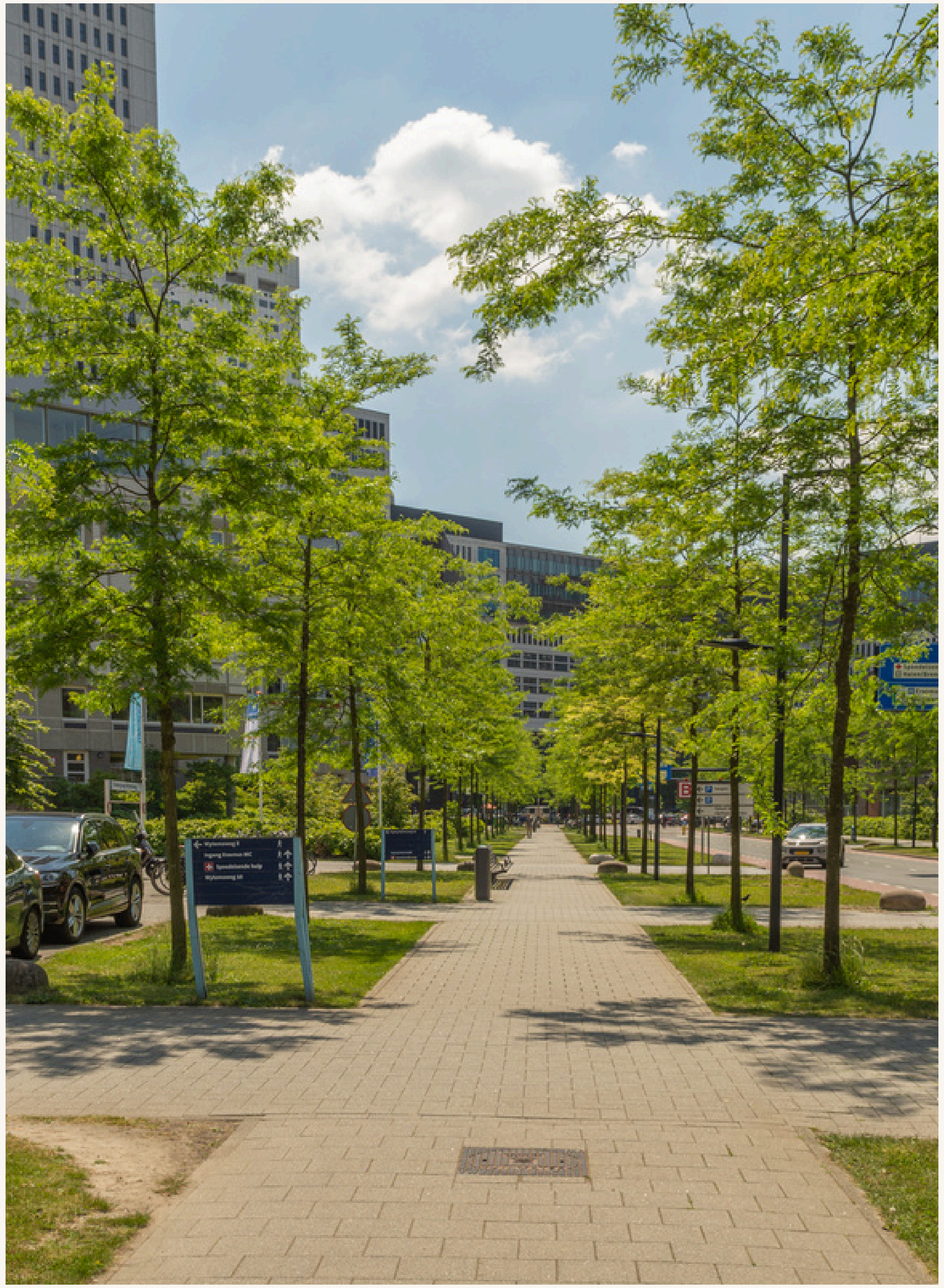
- Increased flood risks
- Increasing heat stress
- Varying Biodiversity

Current Initiatives:

- Rotterdam Climate Adaptation Strategy
- Blue Green infrastructure planning strategy

Research Gap:

- Some, but not many current data driven approaches in solving environmental challenges



GI AND URBAN RESILIENCE

Urban resilience

- A cities ability to absorb, adapt and recover from environmental, social and economic deterrents while retaining functionality

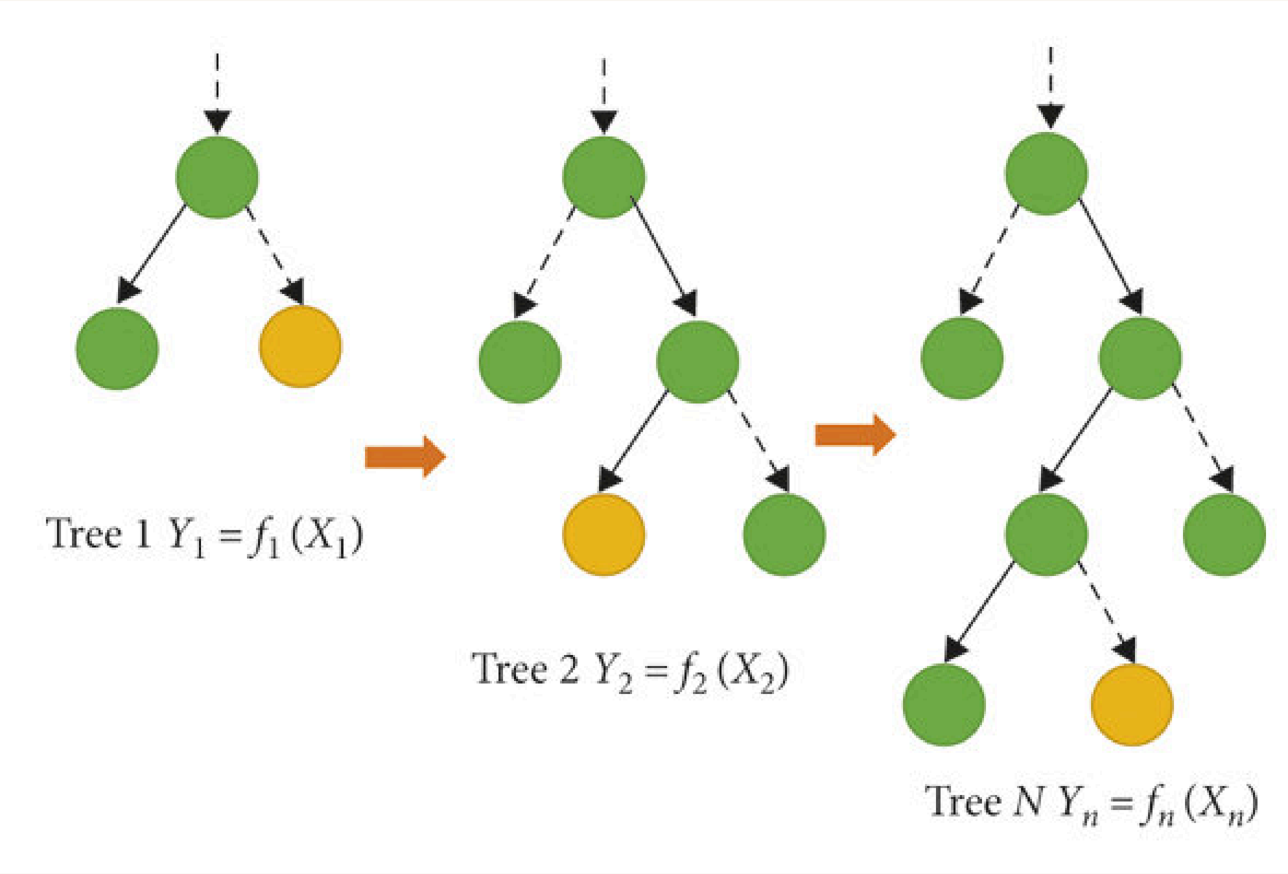
Role of GI:

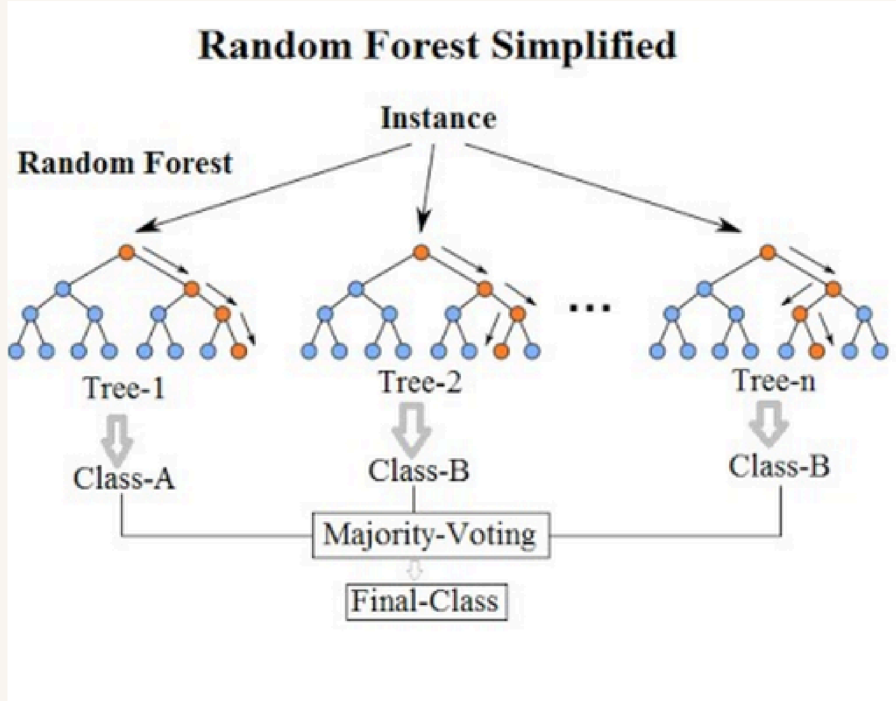
- Enhancing urban resilience
- Supressing built environment related vulnerabilities
- e.g Environmental, Biodiverse, and Morphological challenges.

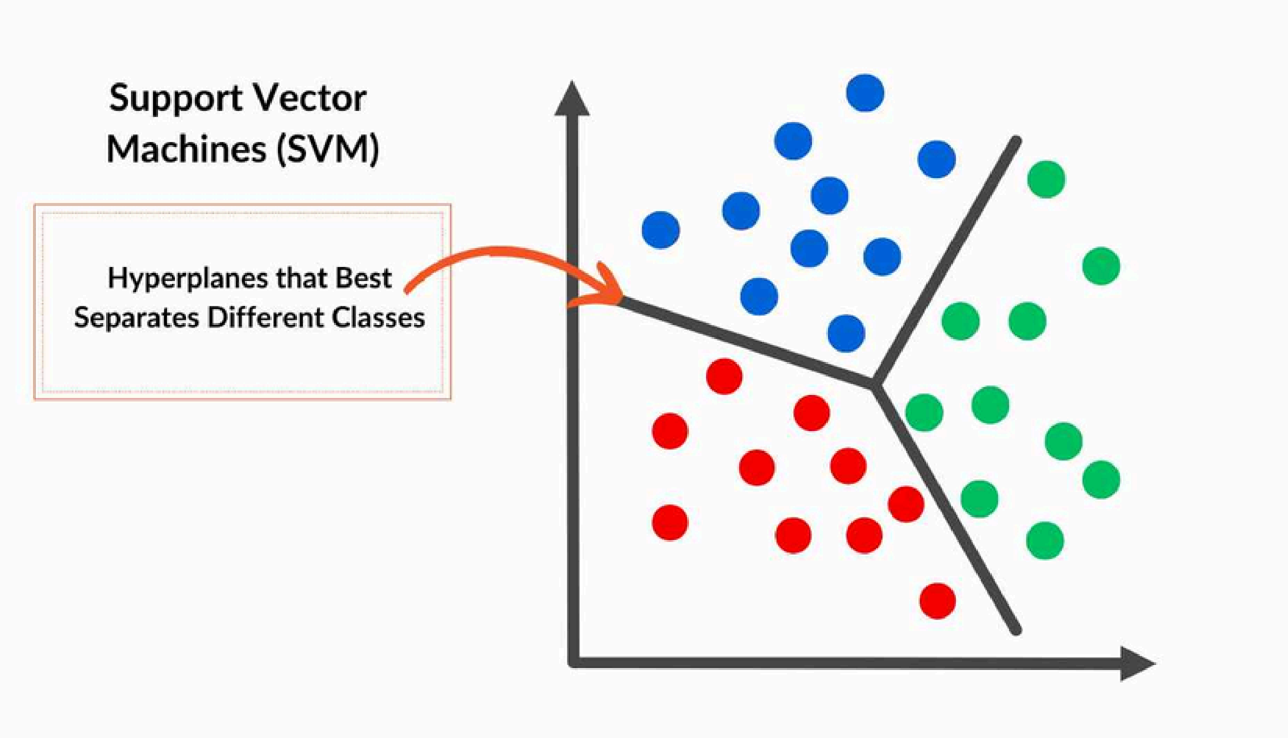
Research Gap:

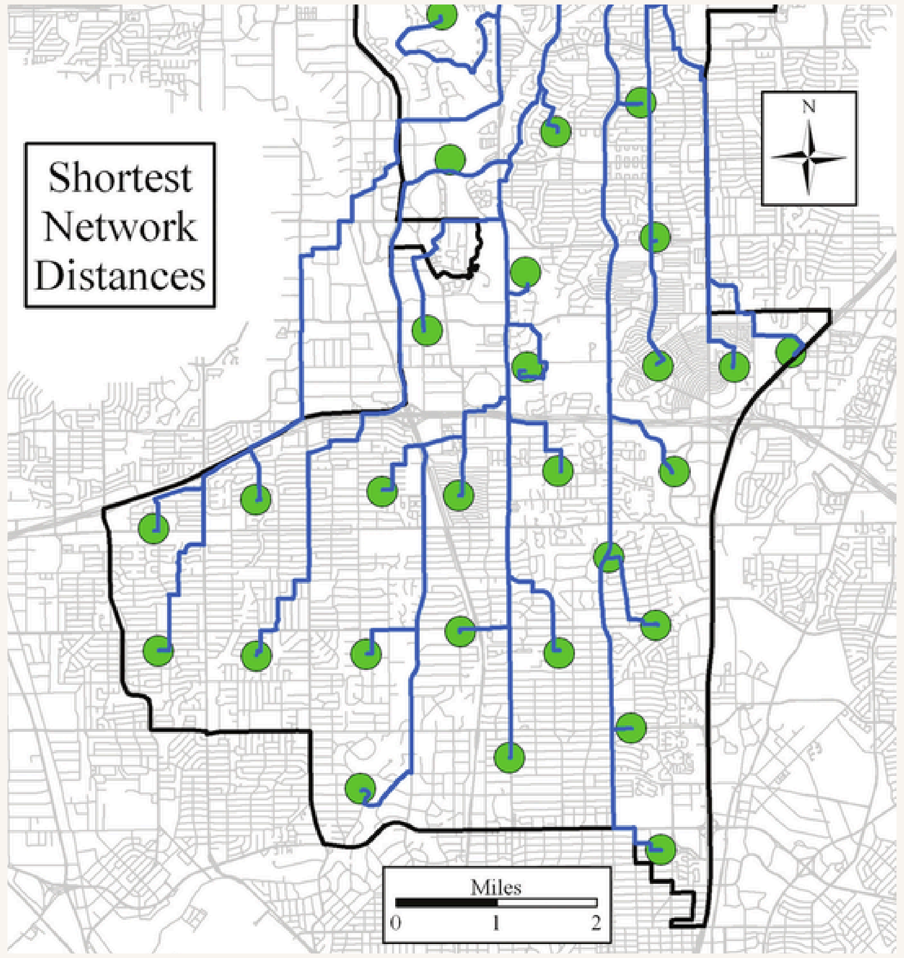
- Limited number of nature based solution that incorporates multiple built environment related dimensions

MODELLING APPROACHES









Machine Learning:

- Random Forest
- Support Vector
- Gradient Boosting

Research Gap:

- Developing an integrated framework that combines spatial analysis and machine learning to evaluate green infrastructure effectiveness.

METHODOLOGY

Data Collection & Cleaning

Feature & Target Extraction

Model Training and Tuning

Case by Case Model Evaluation

DATA COLLECTION

Environmental Data

- Urban Heat island maps
- Perceived temperature maps
- Flood risk maps
- Satellite imagery
- Imperviousness
- NO2 Concentration

Biodiversity Data

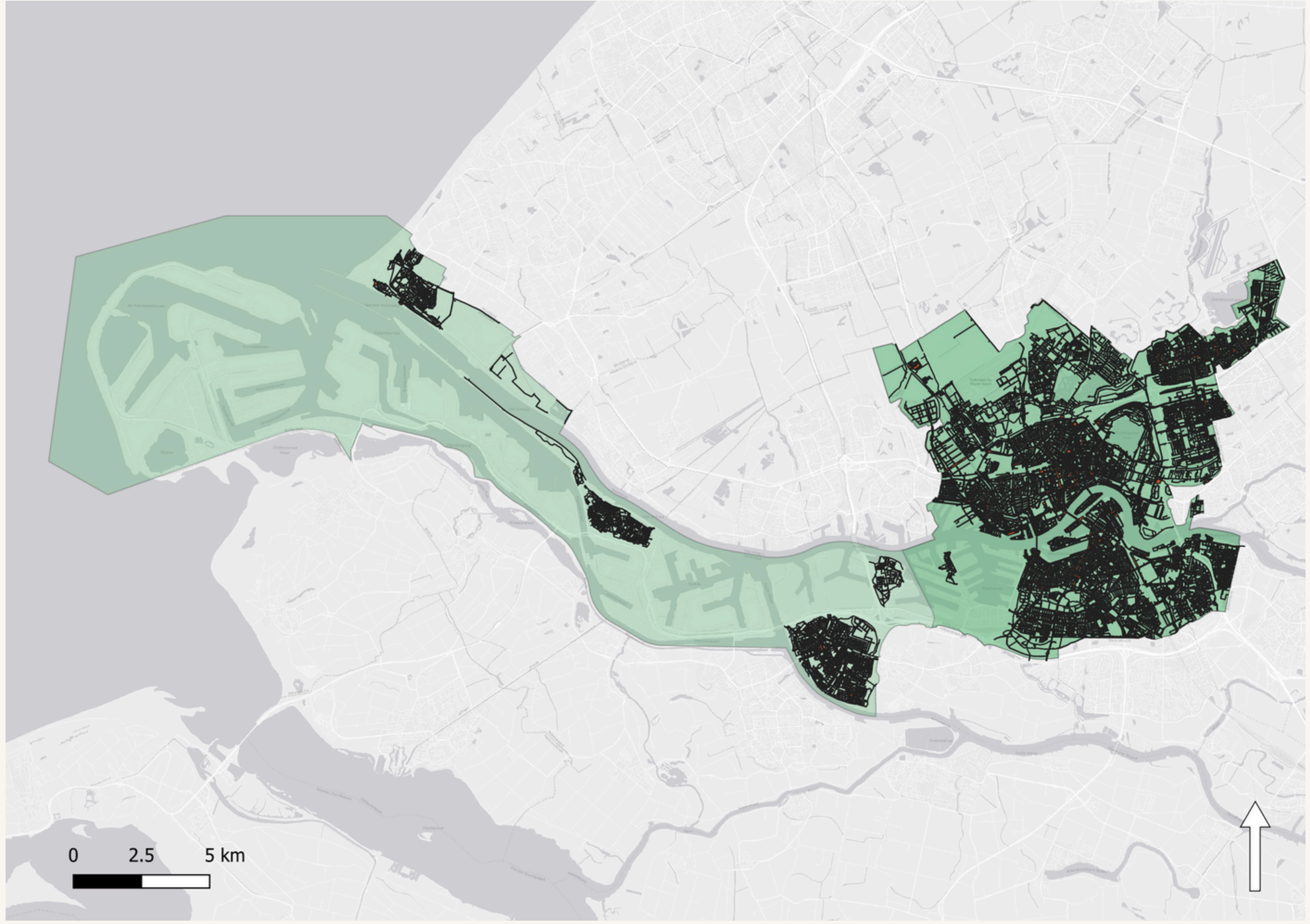
- Tree Data
- Cooling Effect of Green/Blue Spaces

Morphological

- Road Polygons (main spatial unit of analysis)
- Road Centerlines
- Rotterdam Boundary
- Maximum Flood Depth

DATA CLEANING

- Clipping and Masking
- Reprojecting coordinate referencing systems
- Validating Geometry
- Normalising and scaling
- Merging of files



CREATING INDICATORS AND TARGETS

Environmental Data

- Urban Heat Island Effect Values
- Perceived temperature Value
- NDVI Value
- Imperviousness
- **NO2 Concentration**

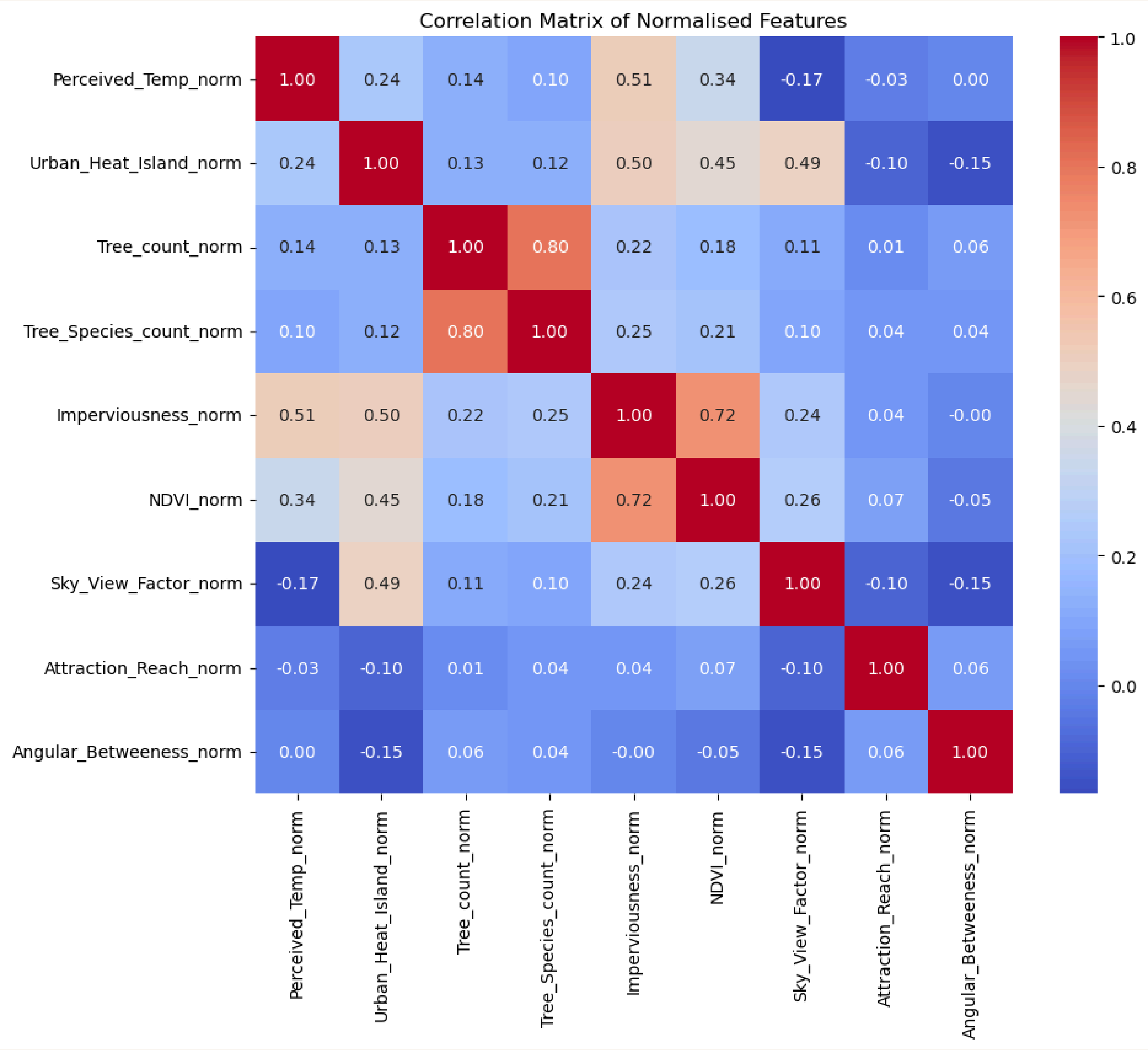
Biodiversity Data

- Tree Population
- Tree Biodiversity
- **Cooling Effect of Green/Blue Spaces**

Morphological

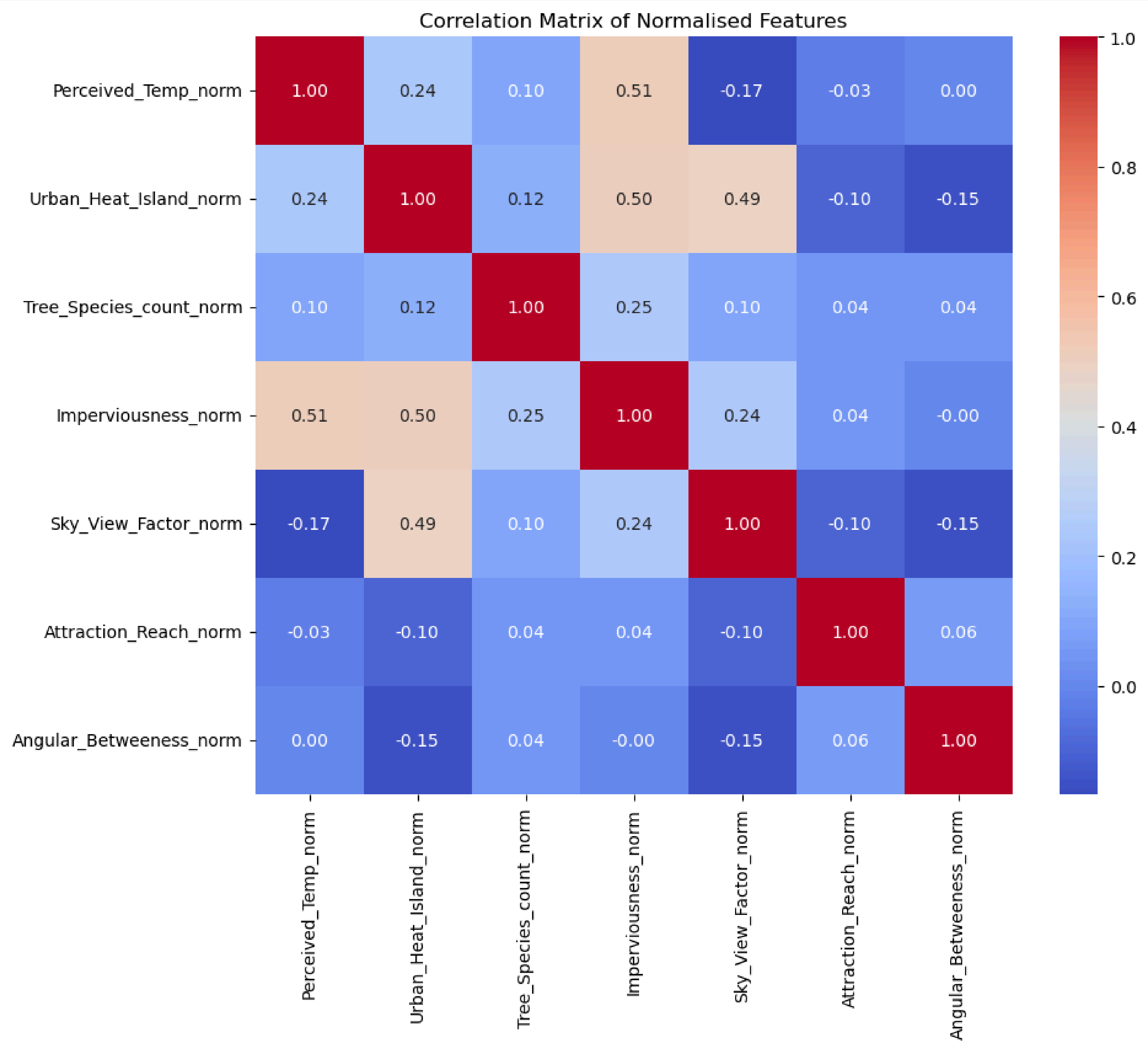
- Attraction Reach
- Angular Betweenness
- Sky View Factor
- **Max Flood Depth**

FEATURE SELECTION



Feature	VIF
NDVI_norm	29.090790
Sky_View_Factor_norm	20.933032
Urban_Heat_Island_norm	12.941318
Perceived_Temp_norm	10.653885
Imperviousness_norm	4.739505
Tree_Species_count_norm	4.313053
Tree_count_norm	3.478288
Attraction_Reach_norm	3.201416
Angular_Betweeness_norm	1.339790

FINAL FEATURE SET



Feature	VIF
Sky_View_Factor_norm	14.274762
Urban_Heat_Island_norm	12.823261
Perceived_Temp_norm	9.121617
Imperviousness_norm	3.799817
Attraction_Reach_norm	3.015511
Tree_Species_count_norm	1.603371
Angular_Betweeness_norm	1.334202

MODELLING APPROACHES

Random Forest Regression

Builds multiple decision trees and aggregates their results. Each tree trained on random subset of data and features

Gradient Boosting Regression

Building trees sequentially, each new tree focuses on correcting error by previous trees

Support Vector Regression

Fits the best hyperplane within a margin of tolerance to predict continuous values, focusing on minimising error outside that margin.

RESULTS CASE 1

(COOLING EFFECTS OF BLUE GREEN INFRASTRUCTURE)

Table 4.9.: Best Hyperparameters and Performance Across Models (Case 1)

Random Forest

Param	Value
n_estimators	200
max_depth	10
max_features	3
min_samples_leaf	4
min_samples_split	5
boostrap	False
-	
R^2	0.44911
MAE	0.04332
RMSE	0.13517

Gradient Boosting

Param	Value
n_estimators	400
max_depth	4
max_features	4
min_samples_leaf	4
min_samples_split	10
learning_rate	0.05
subsample	1
R^2	0.47945
MAE	0.04399
RMSE	0.13139

Support Vector

Param	Value
C	1
epsilon	0.1
kernel	rbf
gamma	scale
-	
-	
-	
R^2	0.39753
MAE	0.06701
RMSE	0.14135

Key takeaways:

- **Gradient Boosting** outperformed all models, best capturing the cooling effects of GI.
- **Random Forest** was a strong alternative, with only slightly lower accuracy.
- **Support Vector** struggled, likely due to limitations in modelling spatial non-linearity.

FEATURE IMPORTANCE CASE 1

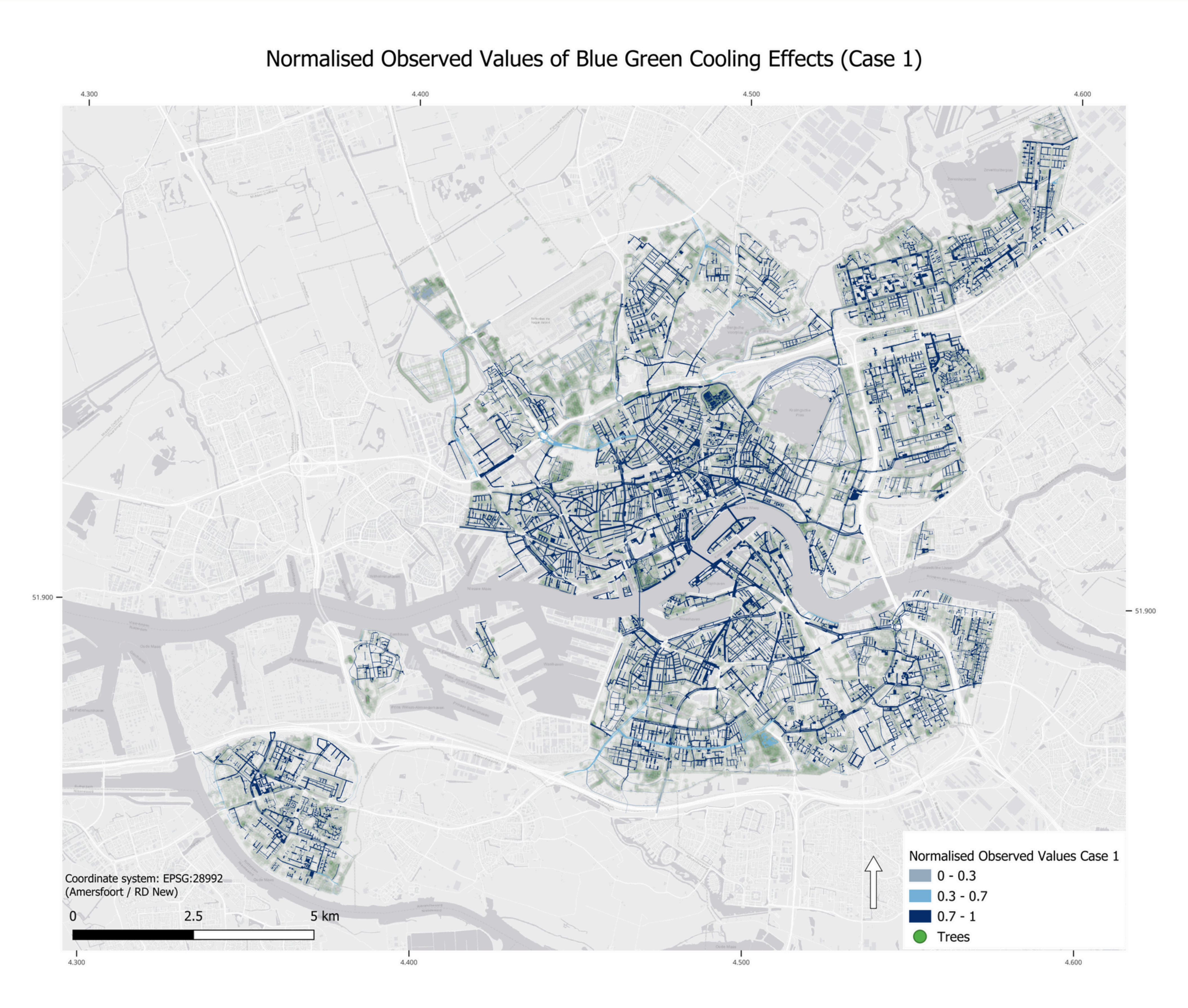
(COOLING EFFCTS OF BLUE GREEN INFRASTRUCTURE)

Table 4.10.: Feature Importance by Model (Case1)

Feature	Random Forest	Gradient Boosting	Support Vector
Urban Heat Island	0.47571	0.48181	0.59311
Imperviousness	0.13748	0.11997	0.20291
Sky View Factor	0.12368	0.12108	0.10446
Angular Betweeness	0.05912	0.06416	0.04773
Attraction Reach	0.10636	0.11158	0.03743
Perceived Temperature	0.06268	0.06032	0.01215
Tree species Count	0.03493	0.04104	0.00221

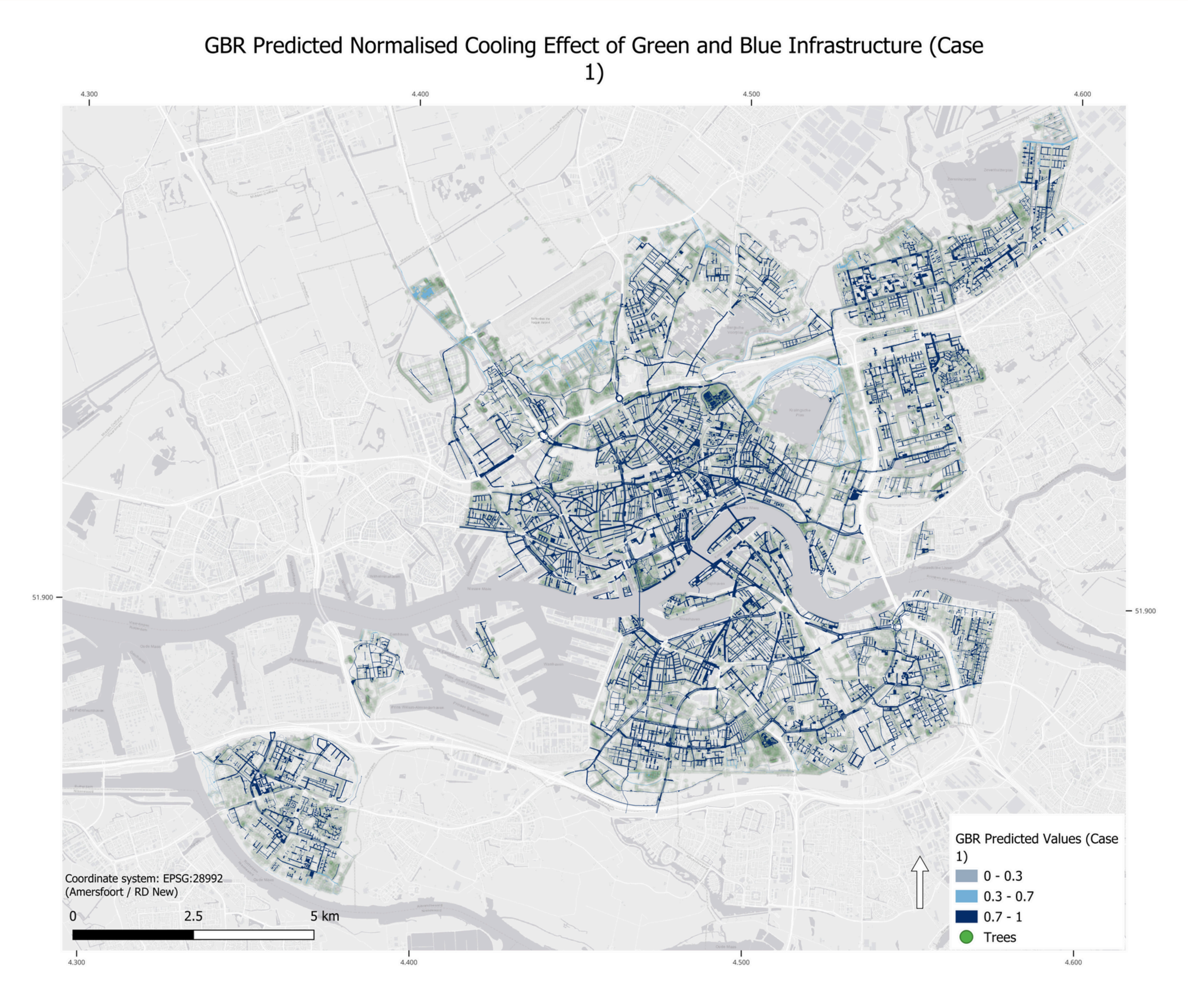
Key takeaways:

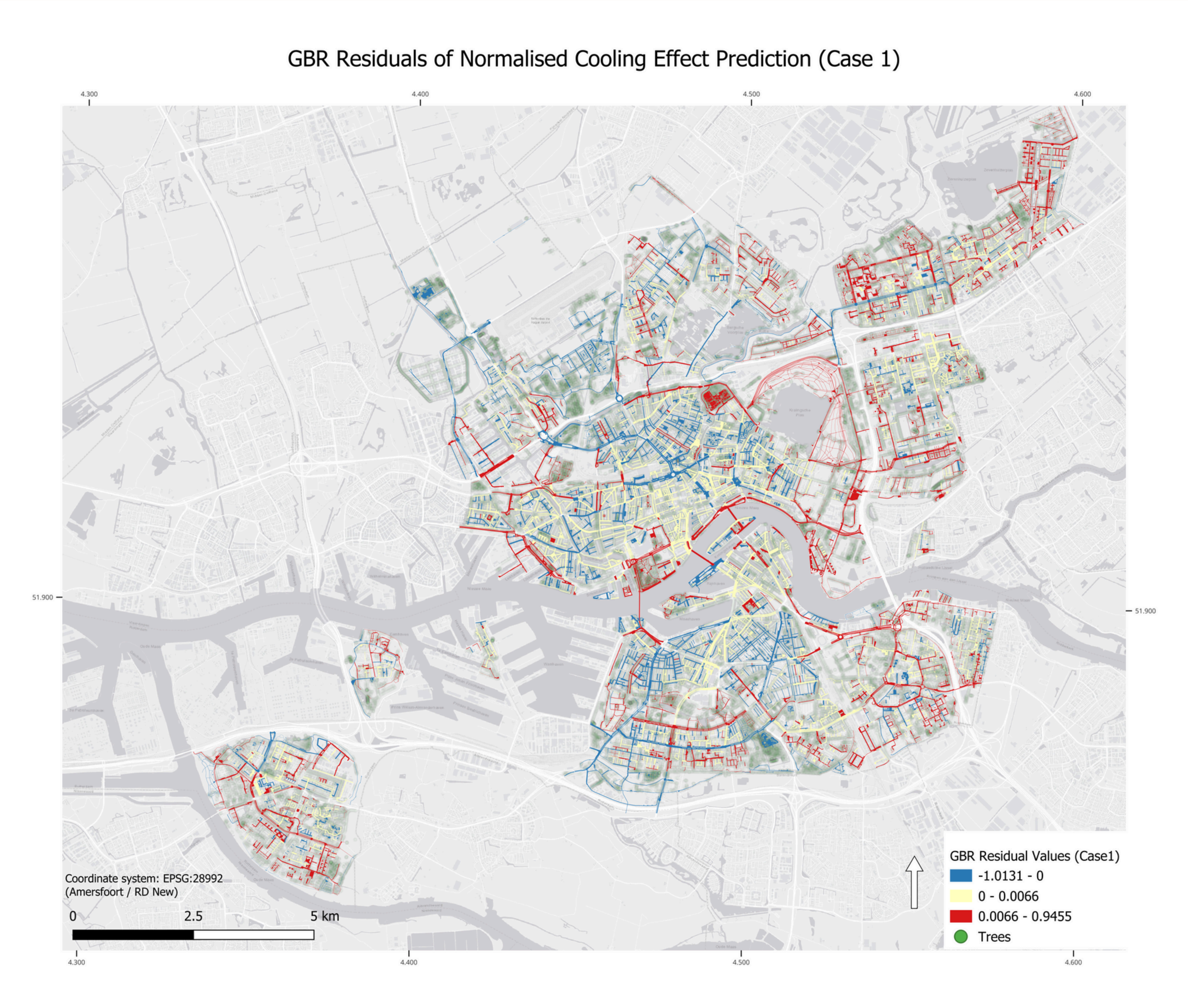
- **Urban Heat Island** is the dominant predictor
- **Imperviousness and Sky View Factor** are consistently important in tree based models.
- **Support Vector** focuses heavily on one feature



Key takeaway:

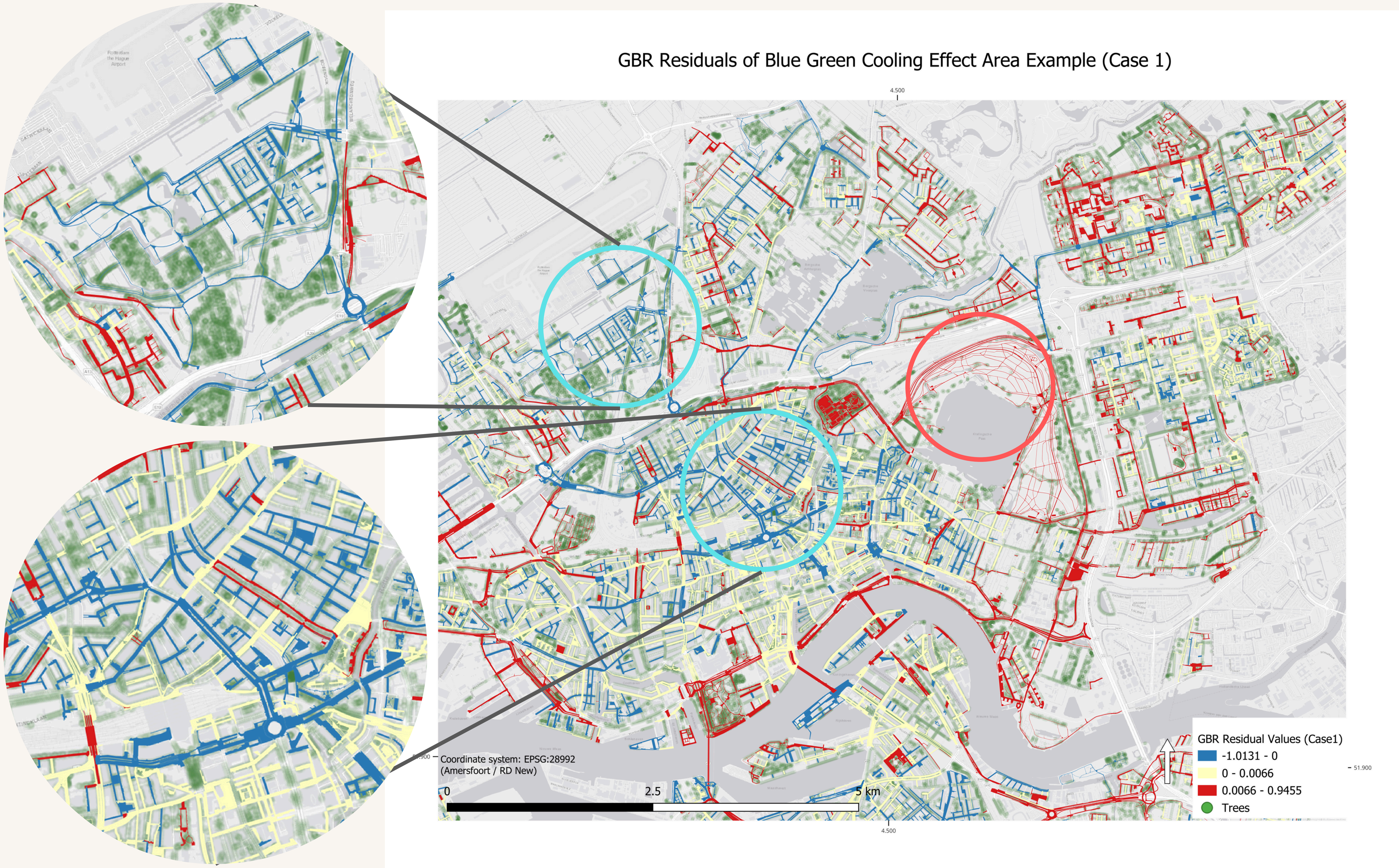
- **Stronger Cooling effects in greener areas**

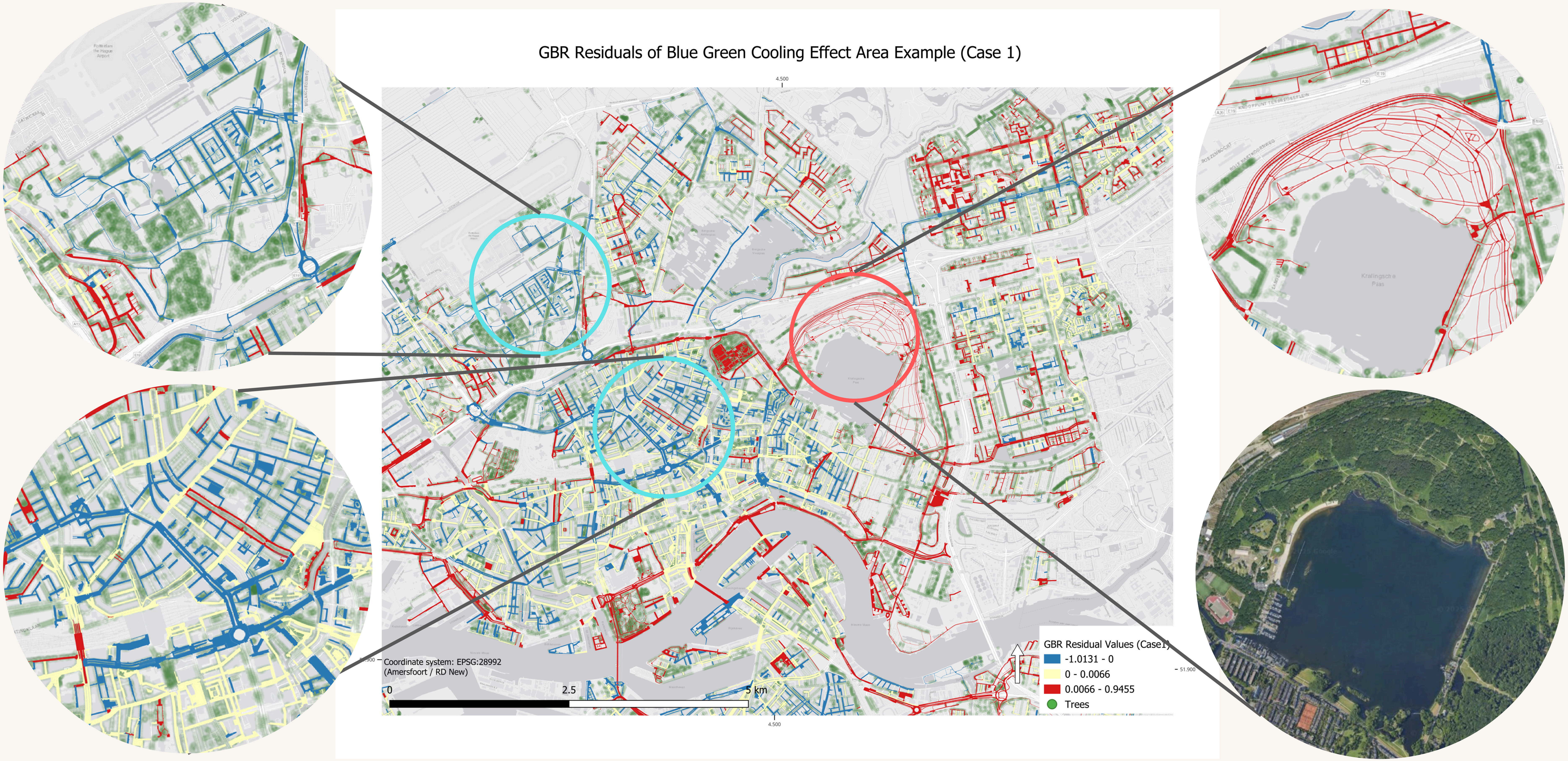

GBR Predicted Normalised Cooling Effect of Green and Blue Infrastructure (Case 1)
4.300
4.400
4.500
4.600
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Predicted Values (Case 1)
0 - 0.3
0.3 - 0.7
0.7 - 1
Trees


GBR Residuals of Normalised Cooling Effect Prediction (Case 1)
4.300
4.400
4.500
4.600
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Residual Values (Case1)
-1.0131 - 0
0 - 0.0066
0.0066 - 0.9455
Trees


GBR Residuals of Blue Green Cooling Effect Area Example (Case 1)
4.500
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Residual Values (Case1)
-1.0131 - 0
0 - 0.0066
0.0066 - 0.9455
Trees

GBR Residuals of Blue Green Cooling Effect Area Example (Case 1)




GBR Residuals of Blue Green Cooling Effect Area Example (Case 1)
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Residual Values (Case1)
-1.0131 - 0
0 - 0.0066
0.0066 - 0.9455
Trees

RESULTS CASE 2

(MAX FLOOD DEPTH)

Table 4.12.: Best Hyperparameters and Performance Across Models (Case 2)

Random Forest		Gradient Boosting		Support Vector	
Param	**Value**	**Param**	**Value**	**Param**	**Value**
n_estimators	150	n_estimators	600	C	1
max_depth	12	max_depth	5	epsilon	0.1
max_features	3	max_features	4	kernel	rbf
min_samples_leaf	4	min_samples_leaf	4	gamma	scale
min_samples_split	5	min_samples_split	20	-	
boostrap	False	learning_rate	0.03	-	
-		subsample	1	-	
R^2	0.32554	R^2	0.32106	R^2	0.25918
MAE	0.12573	MAE	0.12698	MAE	0.13562
RMSE	0.15879	RMSE	0.15932	RMSE	0.16642

Key takeaways:

- **Random** forest was the best model predicting max flood depth with moderate acurracy
- **Gradient Boosting** was a strong alternative, with only slightly lower accuracy.
- **Support Vector** struggled again

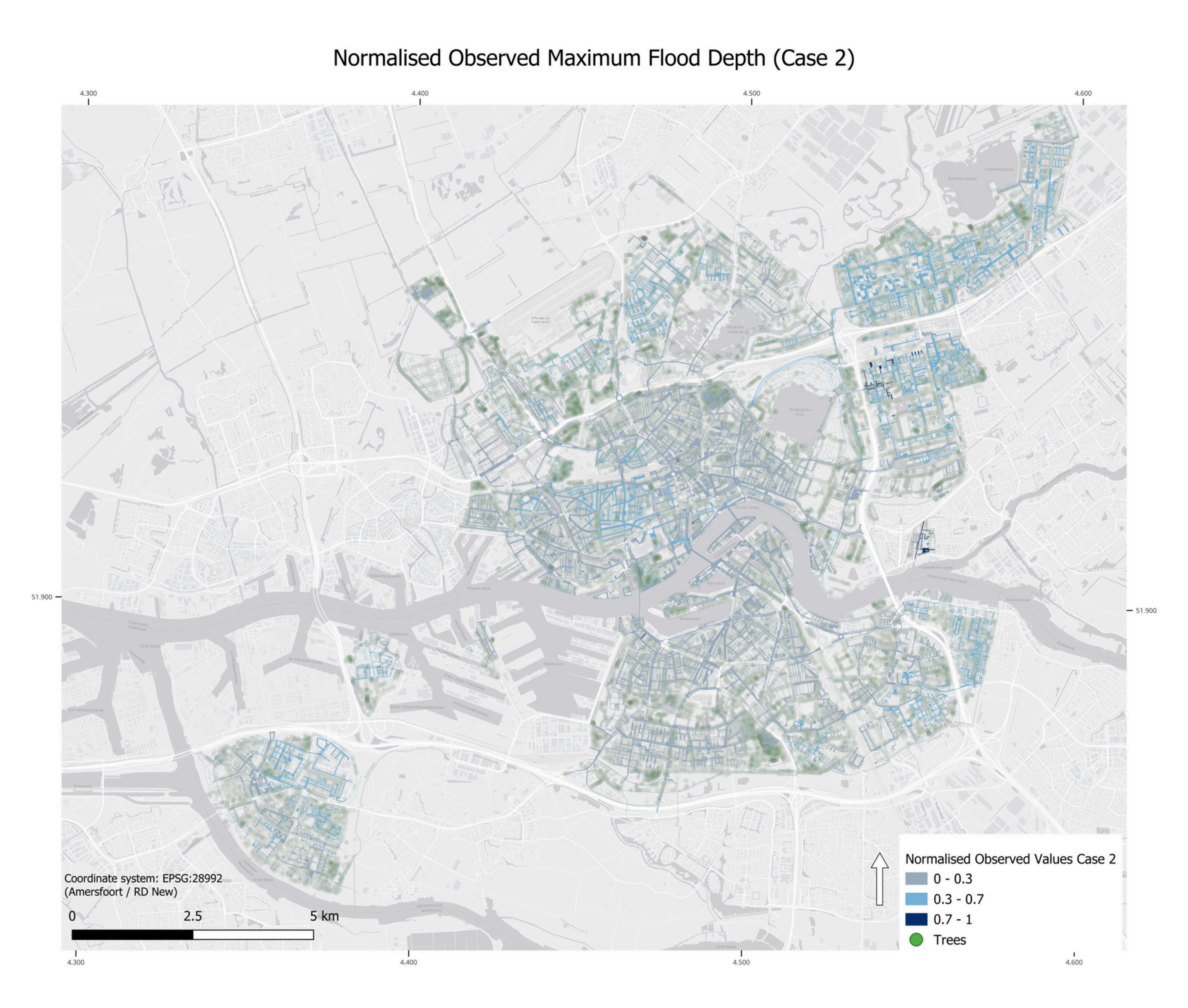
FEATURE IMPORTANCE CASE 2 (MAX FLOOD DEPTH)

Table 4.13.: Feature Importance by Model (Case2)

Feature	Random Forest	Gradient Boosting	Support Vector
Urban Heat Island	0.21551	0.21561	0.17995
Imperviousness	0.09871	0.09140	0.15159
Sky View Factor	0.18333	0.18830	0.21712
Angular Betweeness	0.09180	0.08881	0.06664
Attraction Reach	0.25684	0.25750	0.20884
Perceived Temperature	0.10470	0.11858	0.14590
Tree species Count	0.04907	0.03977	0.02995

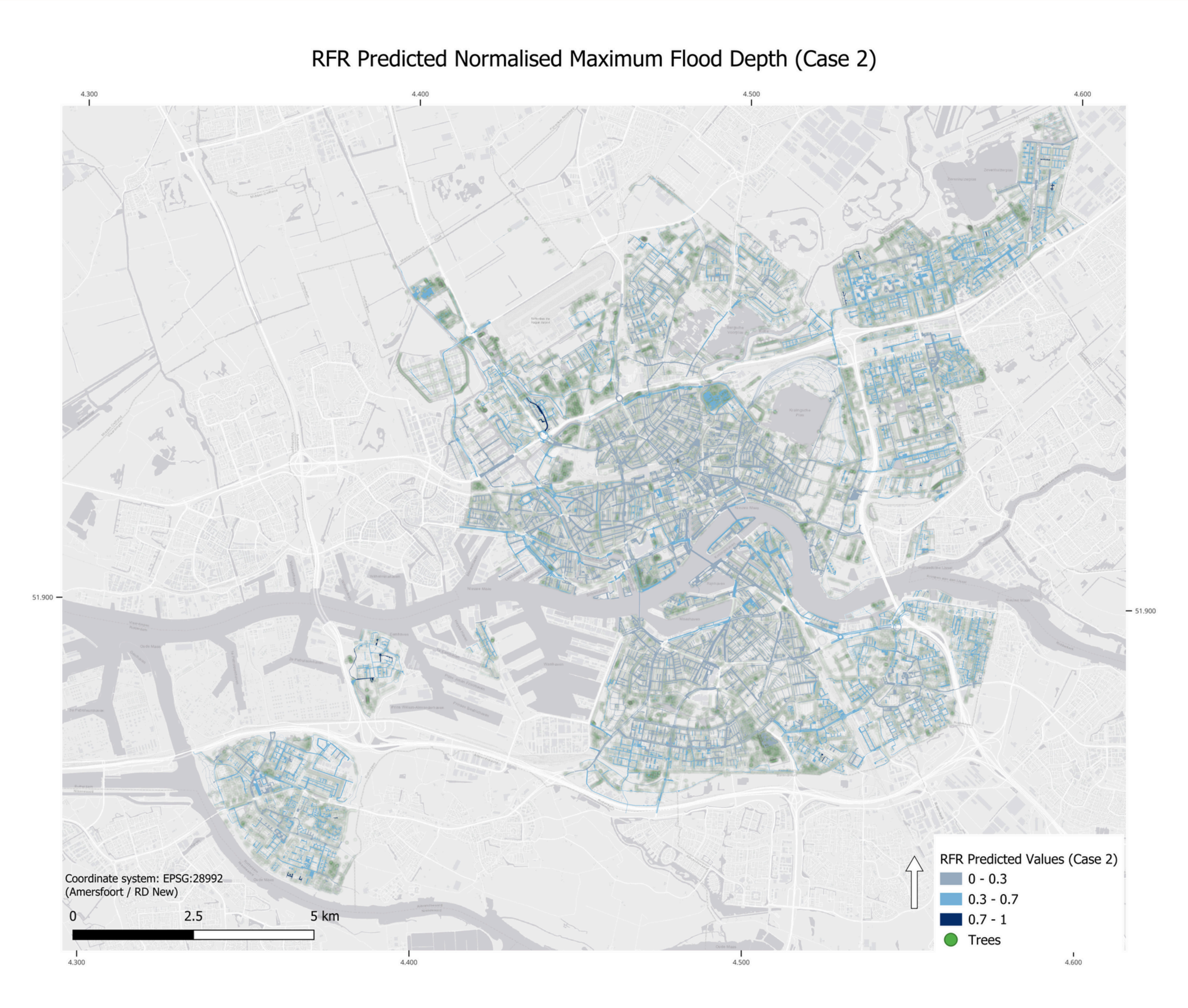
Key takeaways:

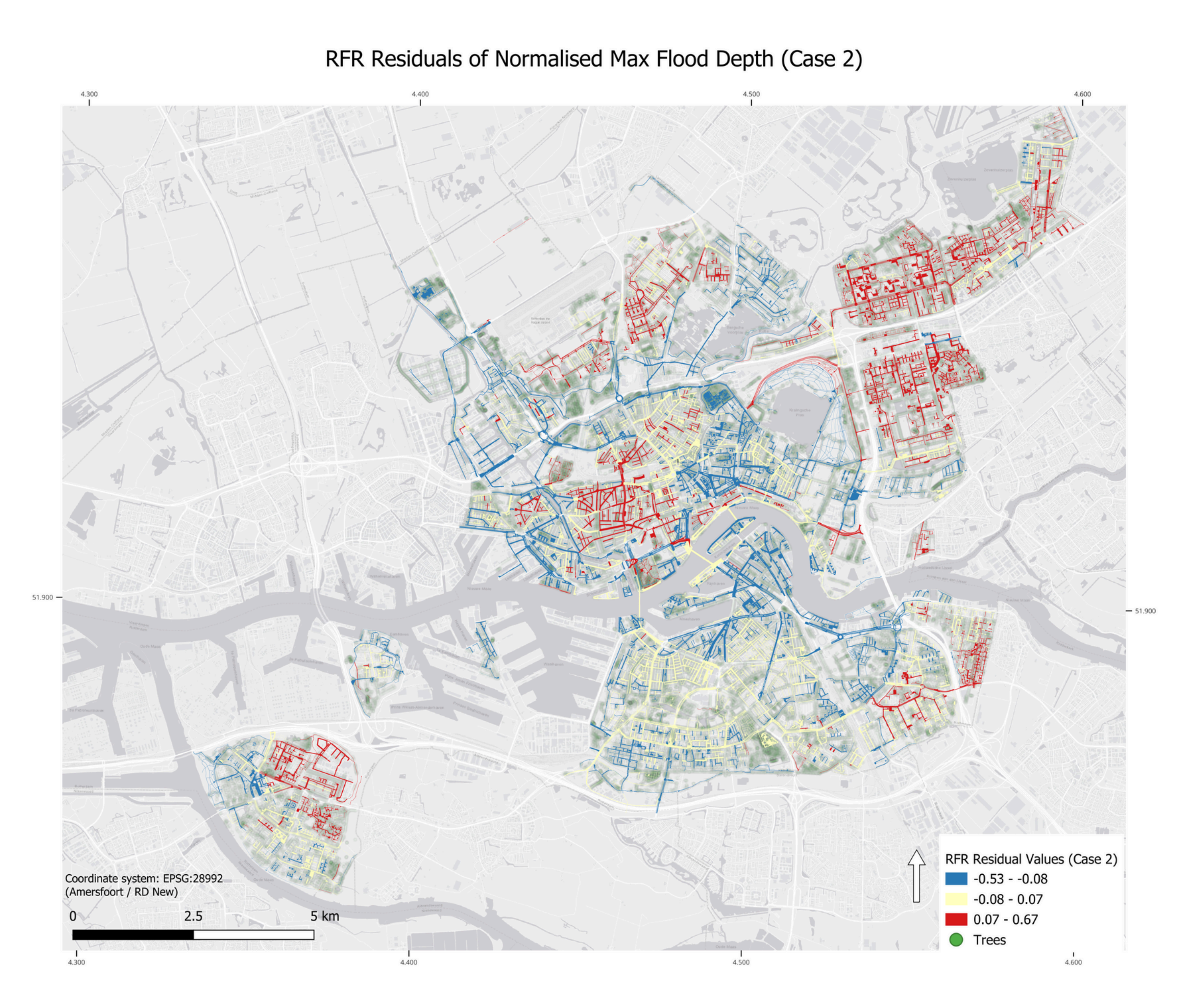
- **Attraction Reach** is the **top** feature across all models.
- **Urban Heat Island and Sky View Factor** are strong predictors.
- **Imperviousness** is important, especially for **SVR**.



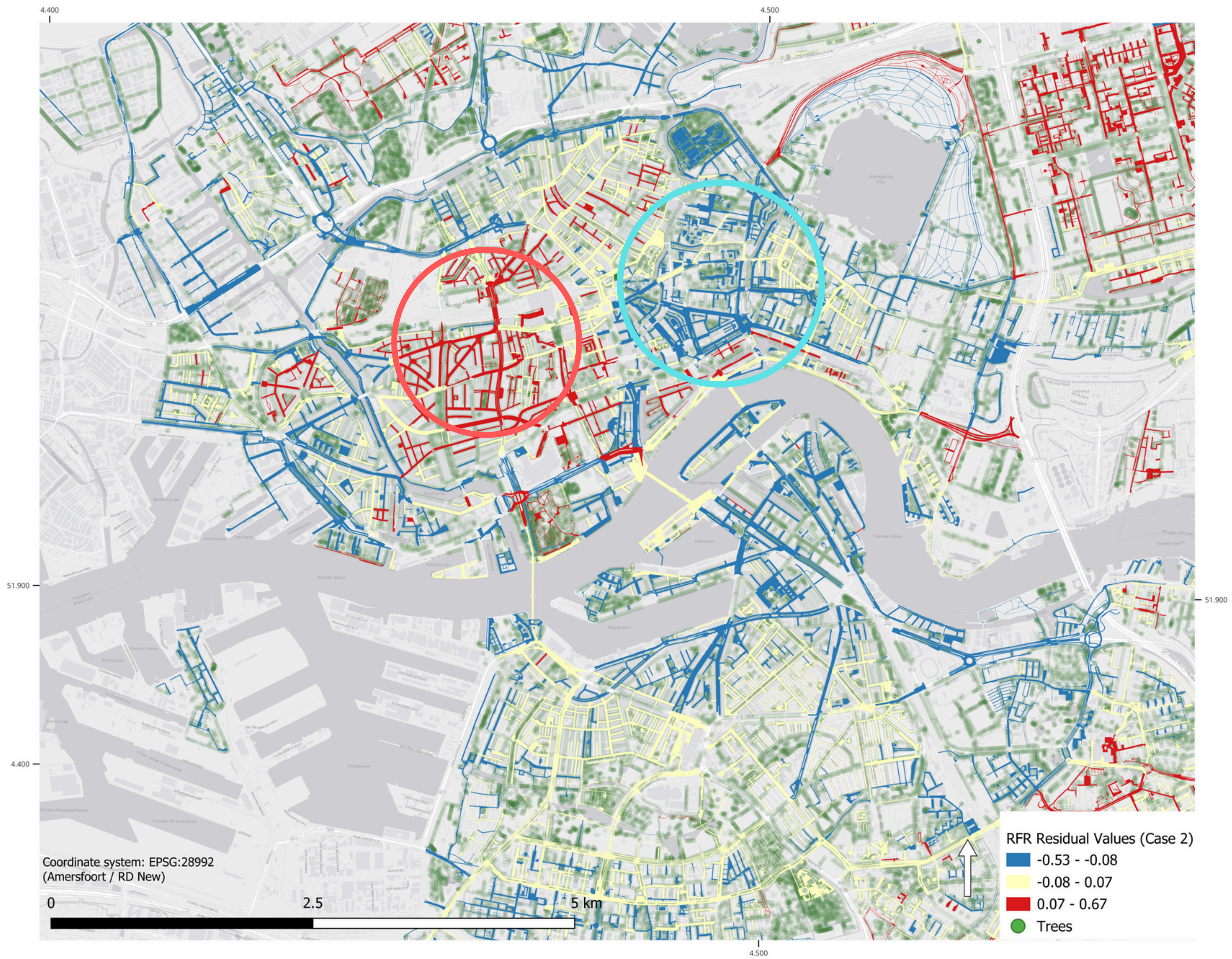
Key takeaway:

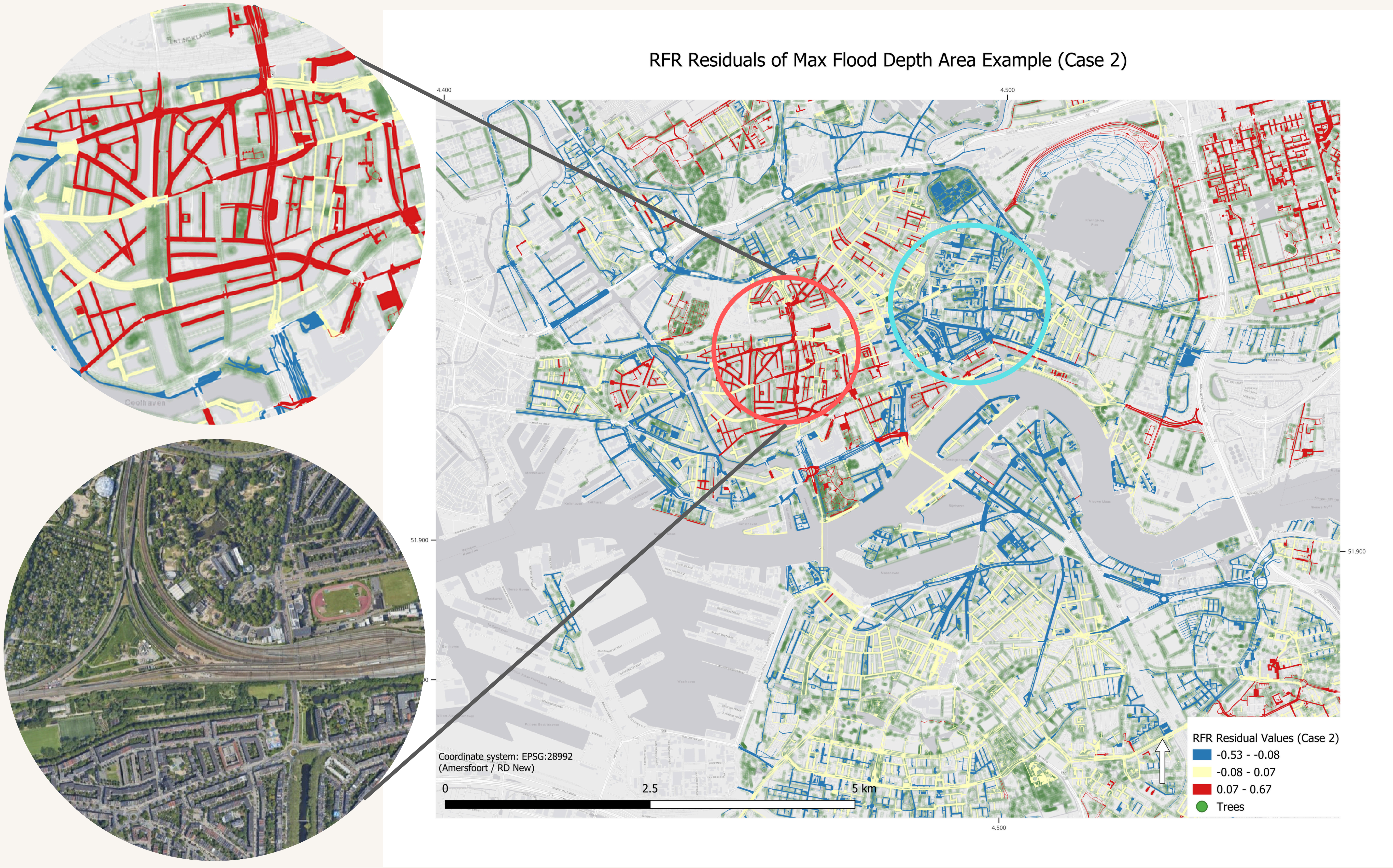
- **Higher max flood depth within city or residential zones**

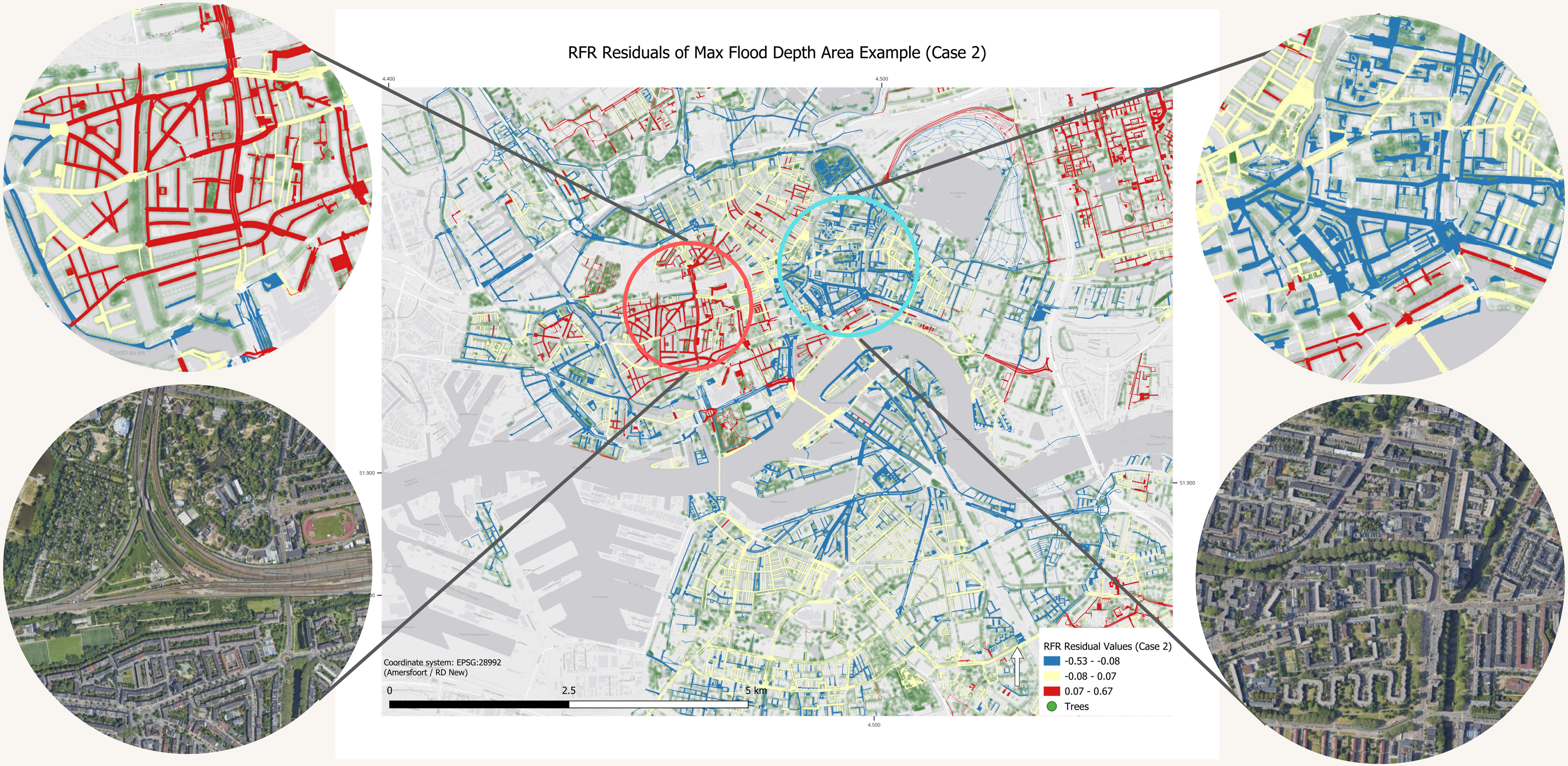

RFR Predicted Normalised Maximum Flood Depth (Case 2)
4.300
4.400
4.500
4.600
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
RFR Predicted Values (Case 2)
0 - 0.3
0.3 - 0.7
0.7 - 1
Trees


RFR Residuals of Normalised Max Flood Depth (Case 2)
4.300
4.400
4.500
4.600
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
RFR Residual Values (Case 2)
-0.53 - -0.08
-0.08 - 0.07
0.07 - 0.67
Trees

RFR Residuals of Max Flood Depth Area Example (Case 2)




RFR Residuals of Max Flood Depth Area Example (Case 2)
4.400
4.500
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
RFR Residual Values (Case 2)
-0.53 - -0.08
-0.08 - 0.07
0.07 - 0.67
Trees


RFR Residuals of Max Flood Depth Area Example (Case 2)
4.400
4.500
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
RFR Residual Values (Case 2)
-0.53 - -0.08
-0.08 - 0.07
0.07 - 0.67
Trees

RESULTS CASE 3

(NO2 CONCENTRATION)

Table 4.15.: Best Hyperparameters and Performance Across Models (Case 3)

Random Forest		Gradient Boosting		Support Vector	
Param	**Value**	**Param**	**Value**	**Param**	**Value**
n_estimators	200	n_estimators	400	C	1
max_depth	12	max_depth	5	epsilon	0.01
max_features	4	max_features	4	kernel	rbf
min_samples_leaf	6	min_samples_leaf	4	gamma	scale
min_samples_split	15	min_samples_split	10	-	
boostrap	True	learning_rate	0.03	-	
-		subsample	0.7	-	
R^2	0.46011	R^2	0.46845	R^2	0.42757
MAE	0.07618	MAE	0.07571	MAE	0.07742
RMSE	0.10082	RMSE	0.10004	RMSE	0.10382

Key takeaways:

- Same as case 1 **Gradient Boosting** at the top
- **Random Forest** again strong alternative, with only slightly lower accuracy.
- **Support Vector** not as bad as in first two cases

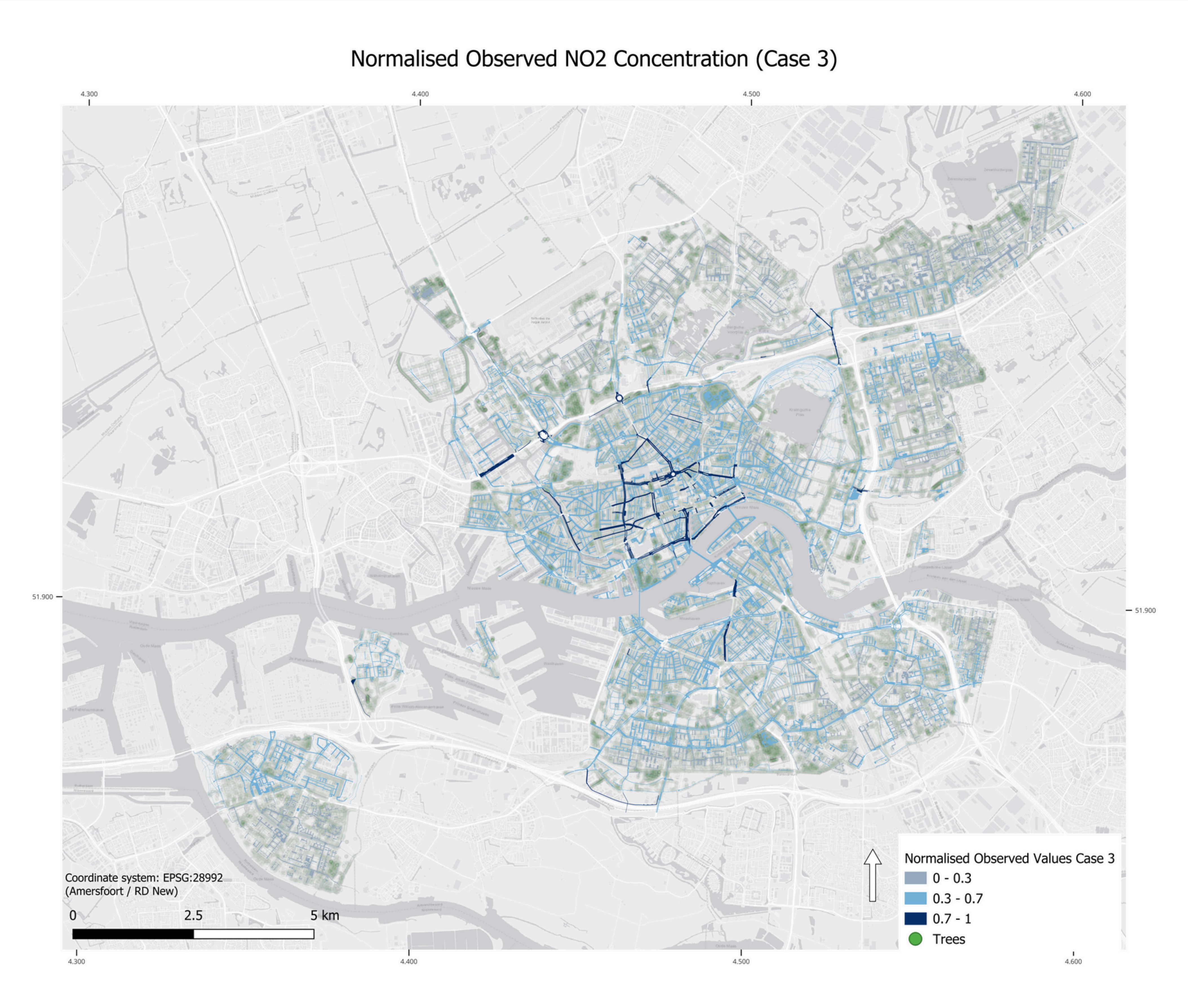
FEATURE IMPORTANCE CASE 3 (NO2 CONCENTRATION)

Table 4.16.: Feature Importance by Model (Case3)

Feature	Random Forest	Gradient Boosting	Support Vector
Urban Heat Island	0.41596	0.37409	0.55818
Imperviousness	0.07900	0.08712	0.11349
Sky View Factor	0.15151	0.15271	0.11337
Angular Betweeness	0.15750	0.15748	0.09257
Attraction Reach	0.09981	0.11084	0.05733
Perceived Temperature	0.06138	0.07782	0.04895
Tree species Count	0.03479	0.03990	0.01609

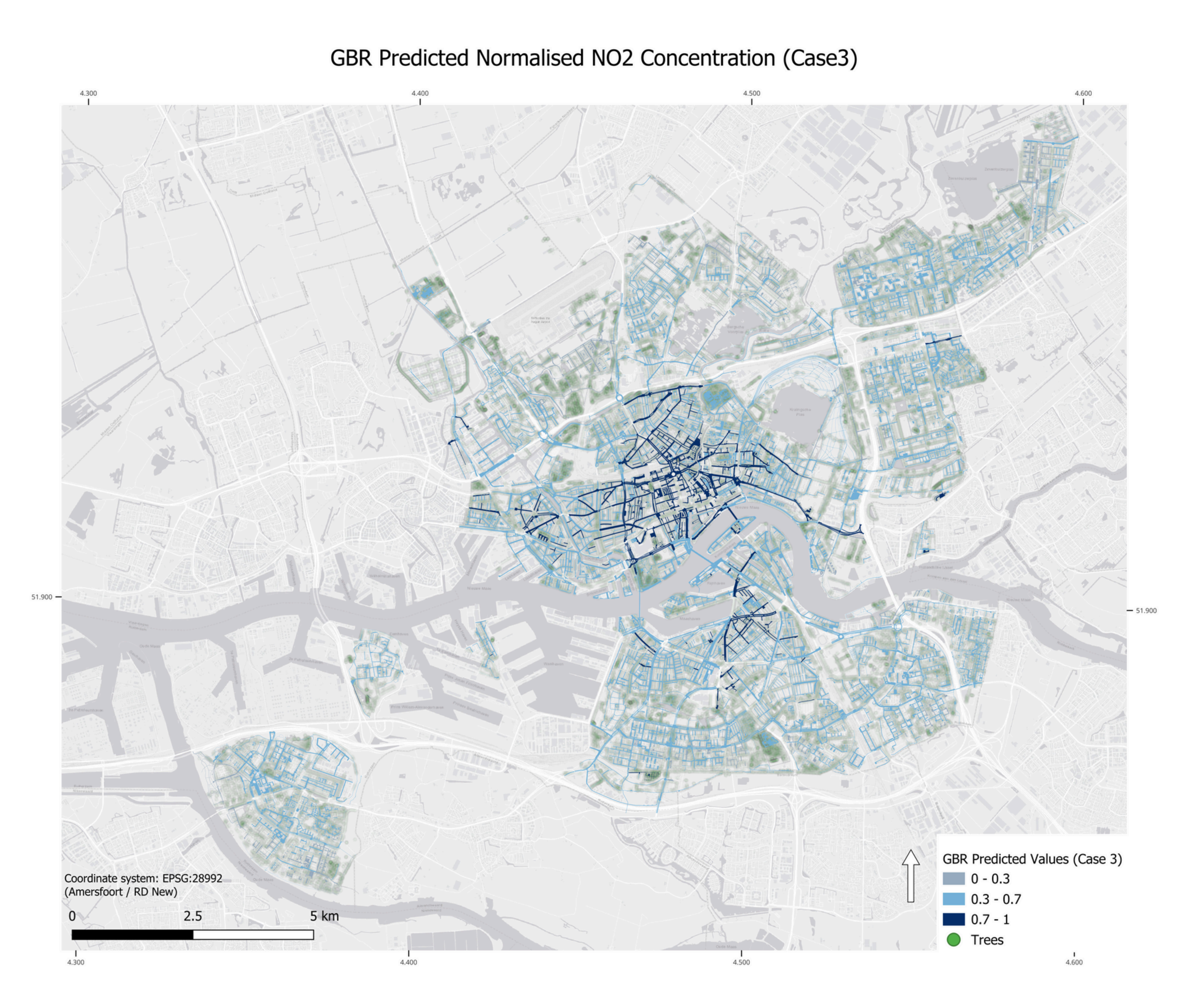
Key takeaways:

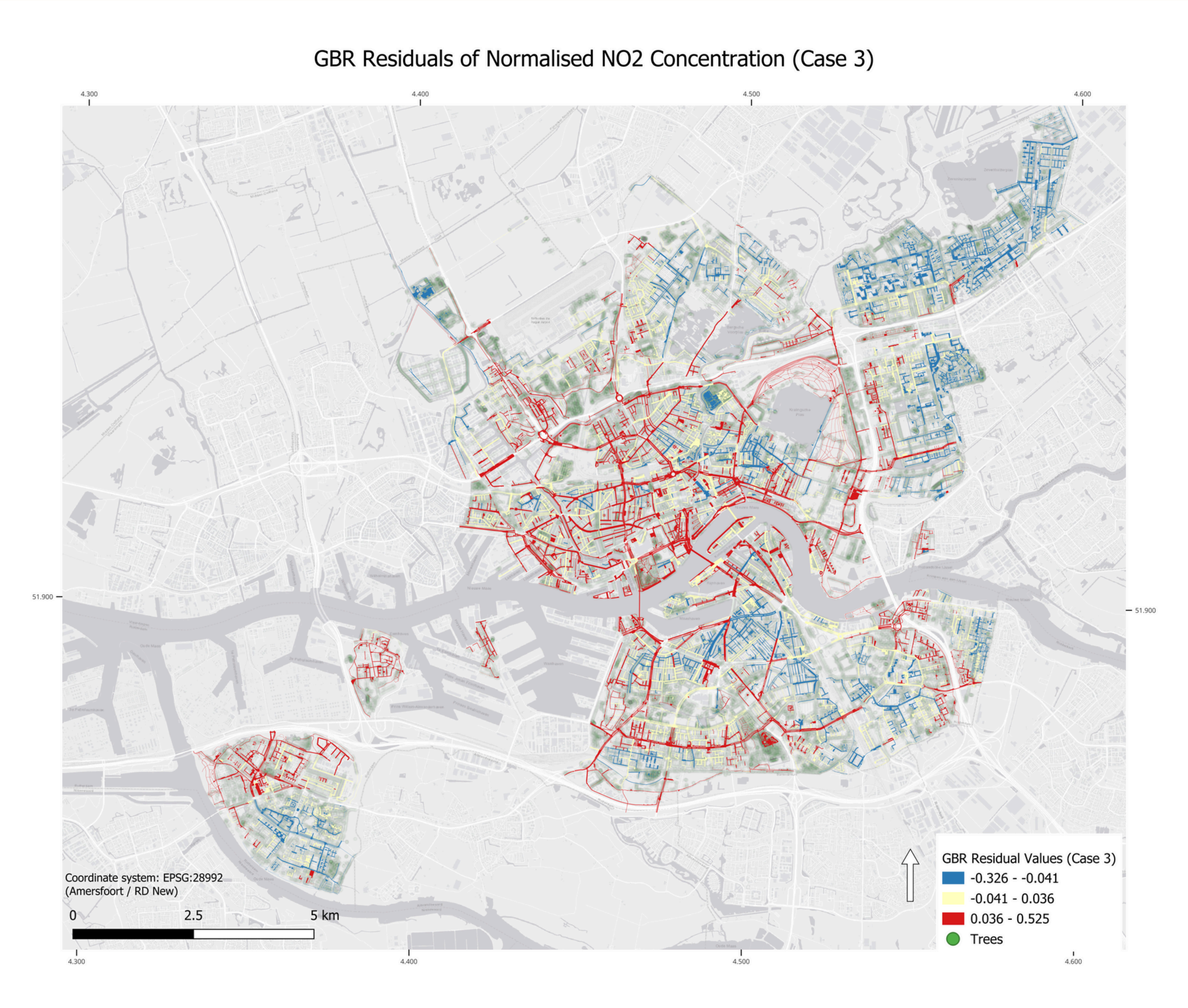
- **Urban Heat Island** is the top predictor
- **Sky View Factor** and **Angular Betweenness** are moderately important.
- **SVR** relies heavily on Urban Heat Island alone.

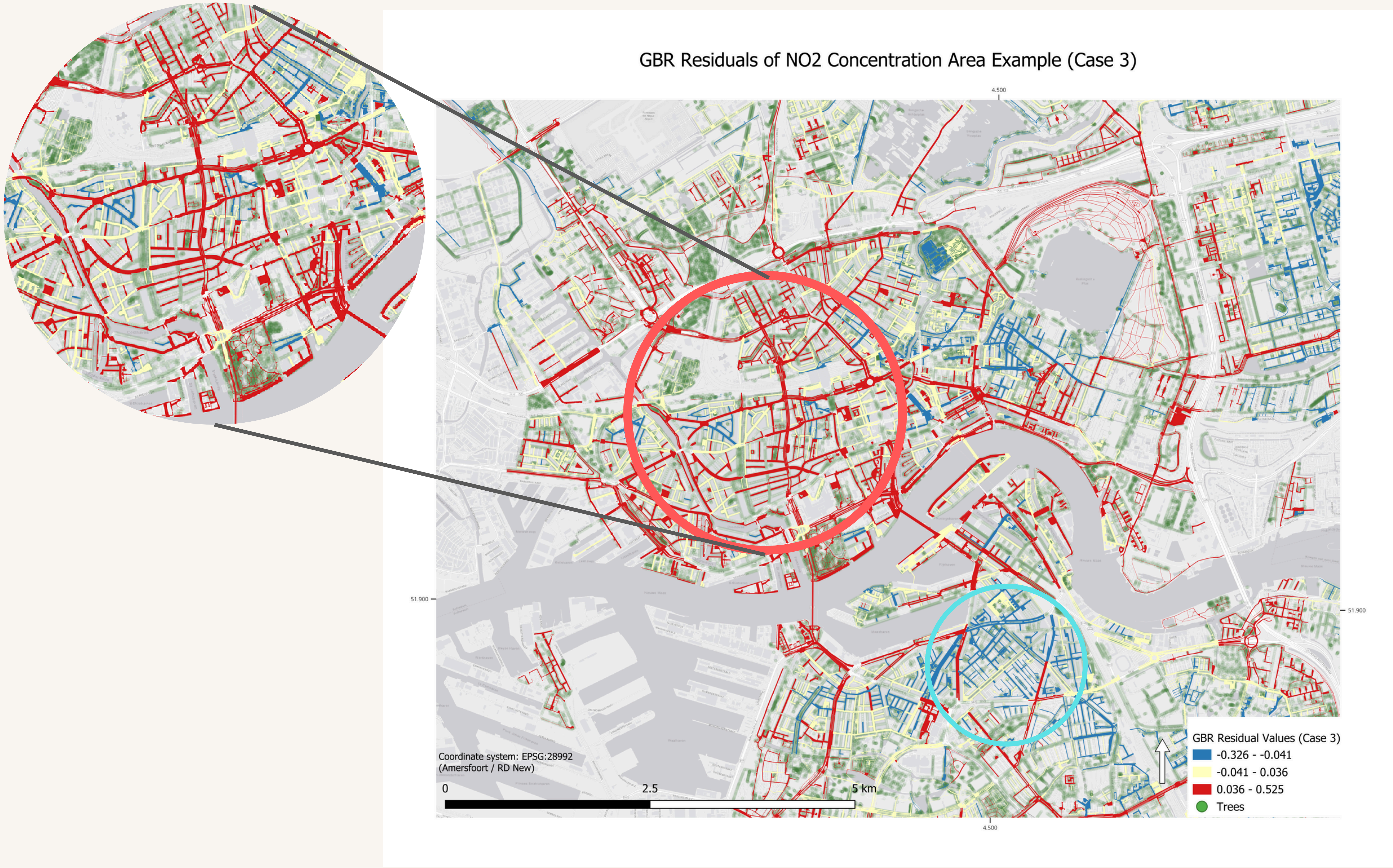


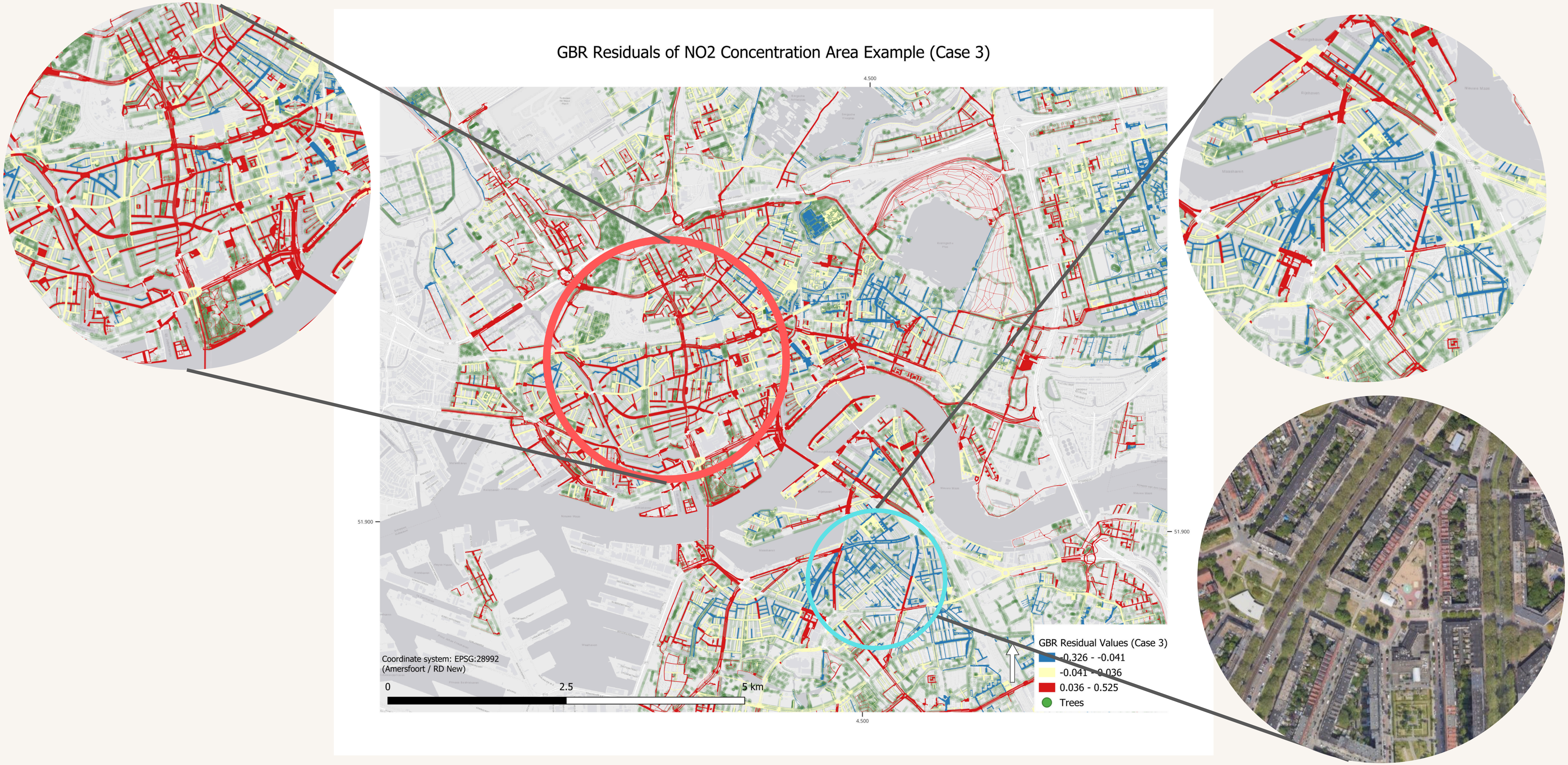
Key takeaway:

- **Higher NO2 concentration near main road networks**


GBR Predicted Normalised NO2 Concentration (Case3)
4.300
4.400
4.500
4.600
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Predicted Values (Case 3)
0 - 0.3
0.3 - 0.7
0.7 - 1
Trees


GBR Residuals of Normalised NO2 Concentration (Case 3)
4.300
4.400
4.500
4.600
51.900
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Residual Values (Case 3)
-0.326 - -0.041
-0.041 - 0.036
0.036 - 0.525
Trees


GBR Residuals of NO2 Concentration Area Example (Case 3)
4.500
51.900
51.900
4.500
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Residual Values (Case 3)
-0.326 - -0.041
-0.041 - 0.036
0.036 - 0.525
Trees


GBR Residuals of NO2 Concentration Area Example (Case 3)
4.500
51.900
51.900
4.500
Coordinate system: EPSG:28992
(Amersfoort / RD New)
0
2.5
5 km
GBR Residual Values (Case 3)
-0.326 - -0.041
-0.041 - 0.036
0.036 - 0.525
Trees

Overall Results

	Case 1 (Blue Green Cooling Efffects)	**Case 2** (Max Flood Depth)	**Case 3** (No2 Concentration)
Best Model	Gradient Boosting	Random Forest	Gradient Boosting
Most Important Feature	Urban Heat island	Attraction Reach	Urban Heat Island
Common Limitation	Limited Green Infrastructure data	Lacked drainage potential features	Lack of wind/traffic flow features
Common residual Pattern	Overprediction in centre/ Underprediction in areas lacking public green data	Over Prediction near water and main road networks	Under Prediction in main road networks
Feature Diversity	Reliant on UHI	Spread through top 3	Reliant on UHI

Key takeaway:

- Model Performance Varies by Target
- Urban Heat Island (UHI) Dominates as a Predictor
- Key Limitation: Missing Thematic Data
- Residuals Reveal Spatial Bias

CONTRIBUTIONS

- Introduced a multi-indicator feature set (environmental, biodiversity, morphological)
- Used multiple models to predict different urban resilience related targets
- Demonstrated benefits of feature extraction and importance of spatial context
- Provides a practical framework to support data driven urban resilience planning

LIMITATIONS & POSSIBLE FUTURE WORK

- Limited Street-Level Data
- Limited Private Green Infrastructure Data
- No Temporal Variables
- No Socio-Economic Features
- Feature Quality and Resolution Mismatch
- Model Generalisation Limits

Thank You for Listening