Predictive policing with Risk Terrain Modeling and Harm-based resource allocation

Group 25
Emma Verhagen, Maxim Even, Zeki Pamir Arikan, Artur Jedlinski, Yue Peng, Thorvald Rovers

Abstract

The UK police experience increasing pressure to allocate limited resources efficiently while preserving fairness and accountability. This study investigates how geospatial, socio-economic, and historical crime factors can be combined to support proactive police resource allocation across English LSOAs. By combining Risk Terrain Modeling, public UK crime data, socio-economic indicators, and machine learning, we predict monthly crime demand and incorporate a harm-based allocation strategy. The model achieved an R^2 of 0.49, indicating moderate predictive performance, and identified stable crime hotspots throughout the forecast period. To address ethical concerns, the system was evaluated using the ALGO-CARE principles. Although the project faced some limitations, the results highlight the potential of combining risk terrain modeling, machine learning, and harm-based allocation to support more proactive and evidence-based policing.

1 Introduction

Police departments in the UK encounter rising dynamics in the patterns of crime alongside major reductions in police budget. As law enforcement agencies cannot react to all forms of demand uniformly at all places, neighborhood policing capacities remain finite by nature. This leads to the major dilemma faced by police decision makers regarding optimal organizing and capacity allocation. In view of these constraints, shifting from a reactive strategy to a proactive one becomes necessary to combat crime more efficiently. (Laufs et al., 2020) However, re-organization of the current model of policing entails difficult practical and normative decisions, specifically concerning efficiency versus regional equity. First-tier stakeholders, such as the Home Office and local police chiefs, consider cost-cutting measures and extremely efficient use of scarce capacities as their foremost concerns (Mahmood, S., 2026). Efficiency, in turn, demands an allocation of scarce resources in the areas traditionally experiencing high levels of crime. Second-tier stakeholders like local councils and human rights activists argue that an over-concentration of policing efforts results in an algorithmic feedback loop and excessive policing of disadvantaged regions. To resolve these conflicting perspectives, this project provides research-based advice for policy making. The overarching research question guiding this analysis is: How can geospatial, socio-economic, and historical crime factors be integrated into machine learning models to support efficient, fair, and ethically responsible police resource allocation across English LSOAs under limited operational capacity?

To answer this, the project utilizes a multi-stage analytical pipeline aggregated at the Lower Layer Super Output Area (LSOA) on a monthly basis. The methodology begins with Risk Terrain Modelling (RTM) to identify stable environmental and socio-economic risk indicators. These stable indicators are subsequently fed into machine learning models to forecast localized crime numbers, which are weighted according to their characteristics in order to perform harm-based indexing that reflects the severity and operational demand of the predicted offenses. Finally, an allocation engine distributes policing resources across LSOAs in proportion to their harm scores.

Finally, an interactive dashboard is deployed to visualize harm scores and resource allocation shares. The dashboard also displays detailed statistics per LSOA, and evaluates its own predictive accuracy regionally, hence functioning as a strategic decision-support tool that enables human commissioners to plan proactive resource deployment while maintaining robustness.

2 Related Work / Background

The existing literature highlights the paradigm shift within law enforcement from responding to crime to anticipating future crime demands. Prediction models normally work by examining past police-recorded crime information. However, traditional models are often criticized for confusing real crime volume with "discovered incidents", crimes recorded simply because officers happened to be in those locations. Due to this reliance on self-generating discovery data, these models risk creating runaway feedback loops that repeatedly send police back to historically heavily patrolled areas.

To deal with the problem of feedback loops in crime prediction models, the literature recommends using risk terrain modeling (RTM) and geospatial crime analysis (Marchment and Gill, 2021). This model is different from traditional models based on previous crime levels or even person-based profiling since RTM assesses environmental factors. It leverages spatial regression to detect the presence of stable physical environments like population density, Index of Multiple Deprivation (IMD), or geographical closeness to certain locations such as transport nodes or pubs that serve as environmental generators of crime. By focusing on place vulnerability, RTM ensures an analytical process that is largely immune to any form of surveillance bias. Once the stable environmental factors have been identified, they can be combined with historical data to train machine learning algorithms (e.g., Ordinary Least Squares, random forests, gradient boosting) to forecast localized demands at the LSOA-month level. However, there needs to be a balance in the selection of algorithms based on performance and interpretability. Highly complex algorithms, such as deep neural nets, are not considered appropriate for public-sector applications because they cannot easily explain how they arrive at their predictions. Interpretable models make it easier to move towards harm-based indexing, where crime predictions are ranked according to social impact rather than incident volume.

Such applications raise significant ethical and normative questions that have to be closely governed (Gstrein, 2019). The literature emphasizes using systematic frameworks like the ALGO-CARE model to manage the competing interests of efficiency and equality. This ensures that predictive systems do not override human decision-making, remain transparent and lawful, and account for the gap between police-reported events and actual crimes to prevent the over-policing and victimization of marginalized socio-economic neighborhoods.

3 Stakeholder Analysis

Any attempt to modify the current system of policing will generate several issues from both practical and ethical perspectives due to the fact that different organizations have varying priorities concerning the effectiveness of the system and its geographical balance. The first order decision-makers include the Home Office, NPCC, regional police services, PCCs, and independent supervisory bodies, namely HMICFRS. These organizations have practical power and their main driving force behind any decisions made is efficiency and effective use of capabilities. In the case of the police service, the only thing that matters is the maximum extraction of value from existing resources, which usually implies focusing existing capabilities on high-crime areas. This means that the optimal equation for the police force would be focusing all capabilities on urban LSOAs with high crime rates.

In contrast, the secondary stakeholders such as local councils, communities and individuals affected by the policing strategy, civil rights organizations, and researchers have no decision-making authority but are greatly affected by these policing strategies. These groups emphasize the importance of taking the entire community into consideration when prioritizing, arguing that prioritizing only the areas depending upon their risk score can cause stigma towards some neighborhoods. They advocate for long-term prevention and accountability, regionality, and fairness, often mentioning that volume policing can create an algorithmic cycle and overpolice socioeconomically disadvantaged areas.

Essentially, today's policing strategy has its built-in trade-offs between efficiency and fairness. In order to reconcile the two contradictory views, the use of a harm-based index is proposed as an important policy recommendation. The former can be addressed by concentrating more attention on high-demand areas; on the other hand, the latter would be satisfied by putting restrictions on the maximum deployment frequency in low-risk hotspots.

4 Data

Data sources The analysis is conducted at the Lower Layer Super Output Area (LSOA) level.

LSOAs are small geographic units used for statistical analysis in England and Wales (Office for National Statistics, 2021). In this project, only English LSOAs are included as the scope of this project is limited to England.

The Police.uk Crime Data dataset contains monthly street crime records (Police.uk, 2026). Data is used from April 2023 to March 2026 and it contains information such as crime type, location, month and LSOA code. The dataset serves as the primary source for the target variable. Individual crime records are aggregated into monthly crime counts for each LSOA.

The Index of Multiple Deprivation (IMD) dataset contains socio-economic indicators at the LSOA level (Ministry of Housing, Communities and Local Government, 2019). Previous research has identified environmental and contextual factors as relevant predictors of crime risk (Caplan et al., 2011, 2017). The dataset contains several elements of deprivation, including income, employment, education, health, and housing/living environment. Through the LSOA code, selected IMD indicators are combined with environmental features and used in the construction of the RTM risk score.

OpenStreetMap (OSM) Points of Interest (POIs) provide environmental and contextual features. The dataset consists of locations such as shops, pubs, schools, transport, and leisure venues. POIs are aggregated at the LSOA level and used as environmental features in the construction of the RTM risk score.

LSOA reference datasets are used to facilitate data integration by linking geographic information across sources. The lookup table resolves differences between the 2011 and 2021 LSOA coding systems, allowing data from multiple sources to be merged. Furthermore, LSOA boundary data are used to perform spatial joins and aggregate environmental features such as Points of Interest at the LSOA level.

Data Integration and Preprocessing To use the data, it has to be cleaned properly. First, we started cleaning the UK crime data, which consisted of monthly street crime, crime outcome, and stop-and-search records. Only the street crime and outcome datasets were used in this project, while stop-and-search observations were excluded because they could not be linked through a Crime ID. The data were cleaned by removing duplicate records, validating coordinates, handling missing values, and standardizing categorical variables. Only valid English LSOA records from April 2023 to March 2026 were retained. Street crime records were then linked to crime outcome records using Crime ID.

Second, street crime records were linked to crime outcome records using Crime ID as a matching identifier. The street crime dataset was used as the primary table, and the outcome information was added through a left join. The resulting merged dataset provided both crime and outcome information for the analysis.

The IMD dataset was cleaned by verifying the required columns, standardizing the LSOA codes, removing duplicate LSOA entries, checking for missing values, and converting important indicators to numeric values. This process ensured that the IMD data could be linked to the other datasets through LSOA code. To ensure compatibility between data sources, LSOA identifiers were standardised to the 2021 coding system using a 2011-to-2021 lookup table prior to merging, as the IMD dataset uses 2011 LSOA codes while the crime and boundary datasets use 2021 codes.

To create the risk terrain modeling dataset, we used IMD indicators and POI counts per LSOA, these were combined using the respective LSOA. The resulting dataset contains normalized risk scores for each LSOA. Since IMD and POI features are both static, the risk score does not vary over time.

The processed crime data were stored and queried in a DuckDB database during model development. IMD indicators and RTM risk scores were maintained separately.

Feature engineering After data cleaning and integration, it is important to understand which features were used in the model. These features include the target variable, socio-economic indicators, environmental characteristics, the risk terrain modeling (RTM) risk score, and temporal information.

The target variable is the variable that the model aims to predict. In the project, the target variable is the monthly crime count for each LSOA. These crime counts are used as a measure for police demand.

Furthermore, the Index of Multiple Deprivation (IMD) dataset was used to create the socio-economic features in the model. The IMD dataset contains many deprivation domains, but only the income score and employment score were selected.

These two variables showed the strongest correlation with crime rate during exploratory analysis. The remaining IMD domains were excluded to reduce noise. The income score and employment score were incorporated indirectly into the model through the RTM risk score rather than separate model inputs.

Environmental features were used from the OpenStreetMap (OSM) dataset. The OSM data provided 19 Point Of Interest (POI) categories and were aggregated at the LSOA level. The Point of Interest are used to capture environmental features that may be associated with crime patterns. Feature selection was conducted using Ordinary Least Squares (OLS) regression. POIs with a p-value below 0.05 were selected, meaning these points have a statistically significant relationship with crime rate. As a result, 15 of the 19 POI categories were included in the RTM risk score.

The RTM risk score was constructed using the 15 selected POI categories together with the income score and employment score from the IMD dataset. The score is calculated using OLS regression and normalized using Min-Max scaling. Since these features are static, the RTM score remains constant over time.

In addition to spatial and environmental features, temporal information was integrated through the monthly structure of the crime dataset. Our model used crime observations for each month between April 2023 and March 2026, allowing the machine learning model to learn patterns over time and predict future crime counts.

Descriptive statistics Table 1 provides an overview of the final dataset used in the analysis.

Table 1: Overview of the final dataset.

Characteristic	Value
Study period	April 2023 – March 2026
Number of months	36
Number of LSOAs	34294
Number of crime records	15738685
Number of observations	388467
Selected POI categories	15
Available IMD domains	7
Selected IMD domains	2
Target variable	Monthly crime count
Geographical unit	LSOA
Temporal unit	Month

5 Methodology

5.1 Risk Terrain Modeling

Risk Terrain Modeling is a geospatial crime analysis framework developed by Caplan and Kennedy that identifies environmental features of the landscape likely to generate or attract criminal activity (Caplan et al., 2011). The core principle of RTM is that crime concentrates at specific locations where the physical features of the built environment create conditions conducive to illegal behaviour.

Each environmental feature is conceptualised as exerting a spatial influence on crime risk, and crucially, it is the co-location of multiple such features, rather than the presence of any single factor alone, that produces areas of heightened vulnerability (Caplan et al., 2017).Environmental features of the built environment can function either as crime generators, which are locations that attract large numbers of people and thereby increase the probability of criminal encounters, or as crime attractors, which are locations specifically known to provide opportunities for offending (Caplan et al., 2011).

By identifying which features are statistically significant predictors of crime risk, RTM moves beyond retrospective hotspot mapping to provide risk assessments substantially more accurate than conventional retrospective mapping techniques (Caplan et al., 2011).

OLS regression was selected as the method for identifying statistically significant environmental predictors, which is similar to the regression-based feature selection approach established in the RTM experiments done by Caplan et al.. OLS was preferred over more complex modelling alternatives due to its interpretability. The results provide a transparent and directly interpretable account of each feature's contribution to crime risk as shown in section 4.

Consistent with the place-based rationale mentioned in Section 2, RTM focuses exclusively on stable environmental characteristics of locations rather than on the demographic attributes of their residents (Caplan et al., 2017). This ensures that the risk scores produced in our study reflect the physical opportunity structure of each LSOA, rather than the characteristics of its population.

5.2 Predictive Modeling

The predictive model is a neural network with learned embeddings for both geographic and crime-type dimensions, trained sequentially across time

using Elastic Weight Consolidation (EWC) (Kirkpatrick et al., 2017).

Architecture. A unified *CrimeEmbeddingModel* is used, which accepts a single (LSOA, crime type, month) triplet and produces a predicted crime count. Two learned embedding layers encode each LSOA and each crime category into dense 16-dimensional vectors. These embeddings are concatenated with three temporal features and the RTM risk score before being passed through two-layer multi-layer perceptron (MLP) with hidden size 128, ReLU activations, and dropout regularization (0.3 and 0.2). The output layer applies a Soft-plus activation, which enforces non-negative predictions consistent with the count nature of the target variable. Training minimizes Poisson negative log-likelihood loss, which is better suited to sparse, over dispersed count data than mean squared error.

Continual Learning with EWC. Because new crime data arrives monthly, the model is trained incrementally rather than retrained from scratch. Retraining on each new month risks *catastrophic forgetting*: the network may overwrite previously learned spatial patterns in order to fit recent observations (Kirkpatrick et al., 2017). EWC addresses this by adding a quadratic regularization penalty to the loss:

$$\mathcal{L}_{total} = \mathcal{L}_{Poisson} + \frac{\lambda}{2} \sum_i F_i \left(\theta_i - \theta_i^*\right)^2 \quad (1)$$

where θ_i^* are the parameter values after the previous training month, F_i is the diagonal Fisher information matrix (estimated from that month's data), and $\lambda = 200$ is the importance hyperparameter. The Fisher diagonal acts as a per-parameter learning rate mask: weights that were important for predicting earlier months are slowed down, preserving accumulated knowledge while still allowing the model to adapt to new data.

The training pipeline proceeds in two phases. A warm-start phase trains for two epochs on each of the twelve months of 2023, seeding the embeddings with full seasonal variation before incremental updating begins. The incremental phase then advances one month at a time through 2024-2025. The model is evaluated on the held-out month, EWC constraints are computed on a 50-batch sample of that month's data, and the model is fine-tuned for five epochs. The 2026 months constitute the out-of-sample forecast horizon and are never used during training.

5.3 Harm-Based Resource Allocation

Allocation of resources was done relative to the predicted harm in each LSOA, which is a measure of the severity of crime in a region. Not all types of crime are equally harmful to society, nor are they equal in the resources they demand. That is why, during the calculation of harm per LSOA, each type of crime that the model predicts was assigned a weight representing these factors. The assignment of weights were done based on their Cambridge Harm Index (CHI) scores (Sherman et al., 2016), scaled by $\sqrt{x}$ to preserve the relative severity of offenses, while also more accurately representing resource demand.

For each LSOA, harm is calculated by multiplying the predicted number of offenses per crime type with the weight assigned to that type, and then summing these weighted crime scores. Resource allocation was then determined by calculating each LSOAs share of total predicted harm, and distributing resources proportionally.

5.4 Evaluation Metrics

To evaluate both the predictive power and operational utility of the model, we analyze three core dimensions: accuracy, efficiency, and fairness. First, we measure prediction accuracy on unseen test data using *Mean Absolute Error* (MAE) for typical deviations, *Root Mean Squared Error* (RMSE) to penalize large outliers, and the *coefficient of determination* (R^2) to compare across crime types with different base rates. Second, we evaluate operational efficiency by simulating resource allocation, measuring the proportion of crime demand captured across different deployment budgets. Third, we assess regional fairness using the Gini coefficient of deployment frequencies across socio-economic deprivation deciles, ensuring resources are not inequitably concentrated in a few areas. Finally, we calculate these metrics across tight budget levels (1%, 5%, 10%, and 20% of LSOAs) to observe how performance degrades under resource scarcity, helping identify whether volume-based or harm-weighted strategies offer the best balance of efficiency and equity.

5.5 Dashboard

The dashboard was created to visualize the outputs of the model. The main goal is to allow stakeholders to explore prediction results and compare historical crime risk across England, in order to as-

sist in proactive resource allocation. The dashboard was developed using Python Dash.

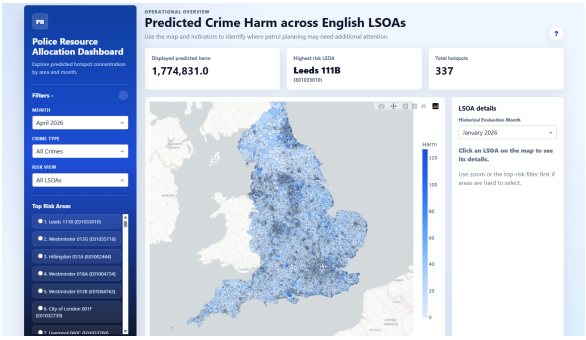

Figure 1: Overview of the dashboard

5.5.1 Layout

The dashboard consists of three main sections: Filters, Map, and Details. A fixed panel on the left allows users to select the forecast month, choose a crime type, and inspect a list of the highest-risk LSOAs. When all crime types are selected, the ranking is based on predicted harm score. When a specific crime type is selected, the ranking is based on the predicted crimes for that category. The centre section contains KPI indicators together with an interactive map of England. On the right side, a details panel displays additional information for the selected LSOA. A general overview of the dashboard is shown in Figure 1.

5.5.2 Map

The Map section displays a choropleth map of English LSOAs, shaded proportionally to predicted harm score. Users can hover over LSOAs to display brief statistics, and can click for detailed statistics. The map allows intuitive comparison between predicted crime demand across England.

5.5.3 Details Panel

The Details section contains statistics for the selected LSOA, such as predicted crime counts, rank, hotspot status, predicted harm score, and allocation share. Historical evaluation results are also displayed for the selected month, which can be selected through the month dropdown at the top of this section. The evaluation section contains predicted and actual crime counts, absolute error, hotspot classification, and a simple reliability indicator based on prediction error all for the chosen month. A bar chart displays the relationship between predicted and actual crime levels for selected month.

5.5.4 Data Flow

The dashboard loads all the data used during initialization, from precomputed datasets. Users interactions are handled through Dash callbacks, and changes made in the filter panel update the map and all displayed data. This allows users to explore different forecast periods and crime types without reloading the application.

6 Results

Model performance The final model achieved an R^2 score of 0.49, indicating that 49% of variation in crime counts can be explained by the predictors included in the model. The model obtained a Mean Absolute Error (MAE) of 0.67 and a Root Mean Squared Error (RMSE) of 2.01. The difference between the RMSE and MAE suggests that prediction errors vary across observations, with some areas being more difficult to predict than others. Overall, these values indicate moderate predictive performance and show that the model captures general crime patterns reasonably well. The machine learning model predicted 5.47 million crimes compared to 5.61 million observed crimes, corresponding to an underestimation of approximately 2.6%. This small underestimation indicates limited systematic bias.

As shown in Figure 2, most points follow the expected trend. Predicted values generally increase when actual values increase and the model captures low and medium crime counts reasonably well. However, prediction errors become larger for areas with very high crime counts and a few outliers remain visible. In summary, the model successfully learns the general relationship between the predictors and crime counts but extreme crime levels remain difficult to predict.

Model performance varied across crime categories. The highest predictive performance was achieved by Violence and Sexual Offenses (R^2 = 0.52), while Bicycle Theft and Weapon Possession were more difficult to predict. In general, more common crime categories achieved higher predictive performance, whereas less common crime categories were more difficult to model accurately. This suggests that rare crimes have greater variability than more frequently occurring crimes.

Forecasted crime demand Crime forecasts were generated for January to June 2026. As shown in Figure 3, predicted crime demand remains relatively stable throughout the forecast period, with

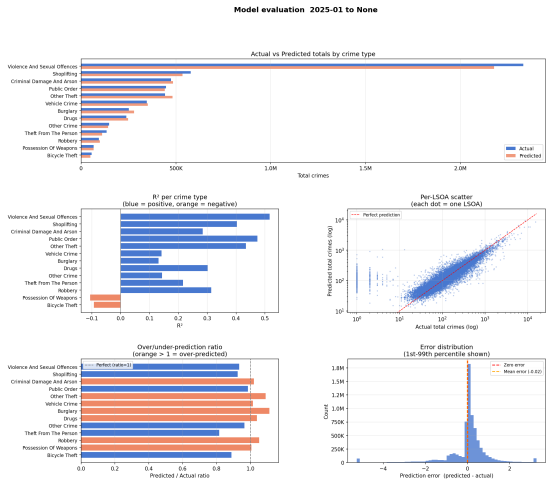

Figure 2: Evaluation plot showing the relationship between actual and predicted crime counts.

no major changes in spatial crime distribution.

Crime demand is concentrated in a relatively small number of LSOAs. Several hotspot regions are clearly visible, but most LSOAs remain low risk. This indicates that crime is spatially clustered rather than evenly distributed.

In addition, several hotspot locations persist across multiple months, while only limited changes occur throughout the forecast period. These high-risk areas remain largely unchanged, suggesting that crime demand remains spatially stable over time.

The identified hotspots represent locations with higher predicted future crime demand and may therefore support future police resource allocation. By identifying areas with persistently high predicted crime levels, the model can be utilized as a decision-support tool for proactive policing and planning.

7 Ethical Considerations

Predictive policing risks automating existing social inequalities. To mitigate this, we embedded the ALGO-CARE framework into our design. Regarding Advisory and Ownership, our tool serves strictly as guidance; it cannot independently dispatch resources, ensuring commanding officers retain ultimate decision-making authority. Regarding Lawful and Granular, all data utilized are publicly available, anonymous crime statistics aggregated at the LSOA-month level.

A primary risk in predictive models is the "feedback loop", where historical police presence is mistaken for actual crime demand. Relying on such data causes algorithms to repeatedly target areas with high past patrol frequency. To address Accuracy and Responsible dimensions, we explicitly mitigate this by excluding stop-and-search records, which prevents the model from reinforcing surveillance biases.

Furthermore, we balance Efficiency versus Geographic Equity. While including socio-economic variables like the Index of Multiple Deprivation (IMD) increases predictive validity, an efficiency-only model would disproportionately funnel resources into disadvantaged areas, risking systemic over-policing and discrimination.

Finally, to meet the Challenge and Explainable criteria, our system avoids "black-box" approaches. By using Risk Terrain Modeling (RTM) to select environmental features, the reasoning behind police placement becomes objective and transparent. This explainability ensures that interventions in a specific LSOA can be reviewed and challenged, preventing the opacity common in complex machine learning models used in public governance.

8 Discussion and Conclusion

Main findings Our final model obtained an R^2 of 0.49, Mean Absolute Error of 0.67, and Root Mean Squared Error of 2.01 and identified stable crime hotspots across the forecast period. These findings indicate moderate predictive performance and suggest that crime demand is spatially concentrated and relatively stable over time. The persistence of crime hotspots suggests that targeted strategies may improve operational efficiency. However, concentrating resources in high-risk areas may also raise fairness concerns, highlighting the importance of utilizing the ALGO-CARE framework.

Moreover, the results demonstrate that geospatial, socio-economic, and historical crime factors can be successfully integrated into a predictive policing framework. By combining risk terrain modeling, machine learning, and harm-based allocation, the model supports more strategic / proactive resource allocation while assisting rather than replacing human decision-making. Furthermore, the harm-based allocation technique demonstrates how resource allocation can move beyond crime counts alone by incorporating the societal impact of different crime types.

Limitations Several limitations should be considered when interpreting these findings. First, the model relies on police-recorded crime data, which

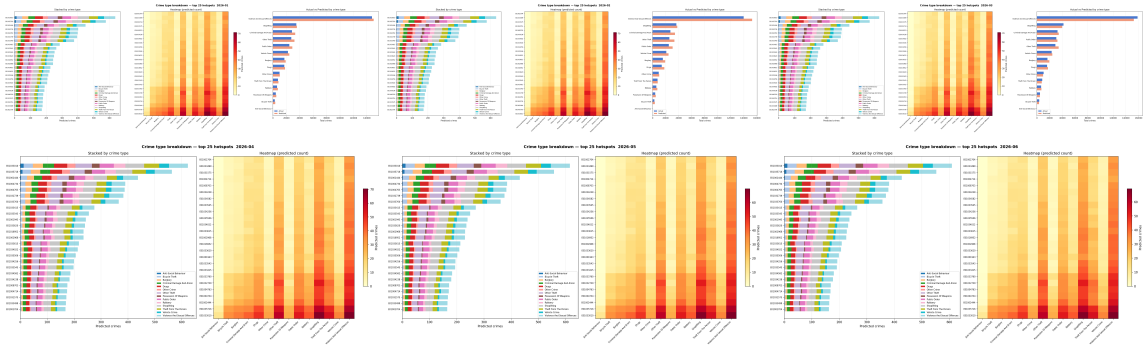

Figure 3: Forecasted crime hotspots from January to June 2026.

may be affected by under-reporting and differences in reporting behaviour across communities. Second, the risk terrain modeling score is based on largely static IMD and OpenStreetMap data. As a result, environmental and socio-economic changes occurring during the study period are not captured. Furthermore, the harm-based allocation strategy depends on weighting choices, where these weights represent a compromise between crime severity and operational police demand and cannot perfectly capture both objectives simultaneously. Furthermore, the dashboard relies on pre-generated prediction files rather than a continuously updated live model, which limits the forecasting period and reduces its usefulness for long-term planning. Finally, although risk terrain modeling provides transparent and interpretable risk factors, the network itself remains partially a black-box model, which limits the explainability of individual predictions.

Future work Since the project was limited in time, several improvements could be made. The model could incorporate dynamic environmental data and automatically update the model when new monthly crime data become available. This would allow predictions to remain current and extend the forecasting period. In addition, the model could be tested in different regions and time periods, while transparency could be improved by more explainable machine learning approaches.

Conclusion In conclusion, geospatial, socio-economic and historical crime factors can be successfully integrated into a machine learning model to support police resource allocation. While limitations and ethical concerns remain, the proposed model demonstrates the potential of combining risk terrain modeling, predictive modeling, and harm-based allocation to support more proactive, efficient, and evidence-based policing.

References

Joel M. Caplan, Leslie W. Kennedy, Jeremy D. Barnum, and Eric L. Piza. 2017. Crime in context: Utilizing risk terrain modeling and conjunctive analysis of case configurations to explore the dynamics of criminogenic behavior settings. *Journal of Contemporary Criminal Justice*, 33(2):133–151.

Joel M. Caplan, Leslie W. Kennedy, and Joel Miller. 2011. Risk terrain modeling: Brokering criminological theory and gis methods for crime forecasting. *Justice Quarterly*, 28(2):360–381.

Bunnik A. Zwitter A. Gstrein, O. J. 2019. Ethical, legal and social challenges of predictive policing. *Católica Law Review*, 3(3):77–98.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. 2017. Overcoming catastrophic forgetting in neural networks.

J. Laufs, K. Bowers, D. Birks, and S.D. Johnson. 2020. Understanding the concept of 'demand' in policing: a scoping review and resulting implications for demand management. *Policing & Society*, pages 895–918.

Mahmood, S. 2026. From local to national: a new model for policing. *GOV.UK*.

Z. Marchment and P. Gill. 2021. Systematic review and meta-analysis of risk terrain modelling (RTM) as a spatial forecasting method. *Crime Science*, 10(12).

Ministry of Housing, Communities and Local Government. 2019. The english indices of deprivation 2019.

Office for National Statistics. 2021. Census 2021 geographies.

Police.uk. 2026. About police.uk data.

Lawrence Sherman, Peter William Neyroud, and Eleanor Neyroud. 2016. The cambridge crime harm index: Measuring total harm from crime based on sentencing guidelines. *Policing*, 10(3):171–183.

A Appendix

Reproducibility Statement

A.1 Github link

https://github.com/SpelGekko/4CBLW020-Group25

A.2 Analysis pipeline

A.2.1 DuckDB

To create a DuckDB database containing historical crime data, first run the `src/processing/dataprocessing-cleaning.py` script on the raw crime datasets obtained from Police.uk. The cleaned files are exported to the processed/cleaned directory. Next, run the `src/processing/dataprocessing-aggregate.py` script to combine the cleaned street crime data with the corresponding crime outcome datasets. The resulting files are stored in the processed directory. Finally, run the `src/processing/dataprocessing-database.py` script, which creates a DuckDB database named `crime.duckdb` from the processed crime files.

A.2.2 RTM

First, OpenStreetMap data for England is downloaded under the name `england.gpkg`. The `extract_pois.py` script extracts 19 candidate POI categories from the OSM POI and transport layers and the extracted POIs are saved in `pois_extracted.gpkg`.

Second, the `count_pois_per_lsoa.py` script performs a spatial join between the extracted POIs and the 2021 LSOA boundary file `LSOA_boundaries.geojson`, producing a count of each POI type per LSOA saved under the name `poi_counts_per_lsoa.csv`.

Third, the `IMD_poi_match.py` script merges the POI counts with the IMD dataset `cleaned_imd.csv`. Since the IMD dataset uses 2011 LSOA codes, a lookup table is used to match the 2011 coding to the 2021 coding system. And the merged dataset is saved under the name `imd_poi_combined_cleaned.csv`.

Finally, the `RTM_regression.py` script fits an OLS regression using the combined dataset `imd_poi_combined_cleaned.csv` and generates a normalised risk score for each LSOA under the name `risk_score_normalized.csv`. It's used as a static input feature in the downstream forecasting model.

A.3 Forecasting Model

Running the training for the model is straightforward. By running `src/machinelearning/train.py` the model automatically starts training in the correct sequence.

To create the predictions run `src/machinelearning/predict.py`.
To create the hotspot files run `src/machinelearning/hotspot.py`. And to create the evaluation metrics, run `src/machinelearning/evaluate.py`.

The trained model will be located in `outputs/model` as a .pt. The evaluation files, metrics file, hotspot files and predictions are all .csv files with the respective names.

A.4 Dashboard

Before launching the dashboard, prediction outputs generated by the model must be manually copied into the dashboard data directory. This includes the hotspot prediction files, evaluation files and type prediction files. Then run the resource allocation module from `src/app/Main.java` to generate `allocation_results.csv`.

Once these files are in place, run `src/visualizations/dashboard/app.py`.
The dashboard automatically loads the prediction outputs and LSOA boundary data from the `data` directory.