機械学習等を活用した地理的犯罪予測に基づく金属盗パトロール実験

野 貴泰*・関口 雅基・大山 智也*****

Metal Theft patrol experiment based on geographical crime prediction using machine learning

Takahiro No*, Masaki Sekiguchi**, Tomoya Ohyama***

Abstract: In recent years, metal thefts have been increasing due to the soaring price of metal, and the damage caused by metal thefts in 2023 amounted to approximately 13.3 billion yen. The theft of metal cables from solar power plant facilities accounts for a high proportion of metal theft damage. Criminologists argue that crimes do not occur randomly, but are concentrated in small places. Intensive police intervention in places with a high risk of future crime is effective in deterring crime. Therefore, in order to efficiently allocate the police's limited human resources and contribute to effective crime prevention, a geographical crime prediction method using hotspot analysis and machine learning was used to identify places with a high risk of future metal thefts, and intensive patrols were conducted in Ibaraki Prefecture. A difference-in-difference analysis (DID) revealed that metal thefts decreased in treatment stations compared to control stations.

Keywords: 金属盗（metal theft），地理的犯罪予測（geographical crime prediction），機械学習（machine learning），差の差分析（difference-in-differences）

１．はじめに

近年，金属の価格高騰を背景に，金属盗が増加しており，令和 5 年に発生した金属盗の被害金額は約 133 億円であった．金属盗被害の中で，太陽光発電所施設に対する金属ケーブル窃盗の被害が占める割合は高く，ケーブルが盗まれることで長期間にわたる発電停止による経済的損失も発生していることから，太陽光発電所における金属盗の抑止対策の強化が求められる (警察庁，2025) ．よって，多くの太陽光発電所に対して，警察の限られた人的リソース等を最大限に活用する取組みが必要である．

近年，警察の限られたリソースをより効率的に活用するために，特定の場所に焦点を当てた防犯活動の研究が進んでいる．犯罪学者は，犯罪が集中している場所，犯罪リスクが高い場所に警察のリソースを集中させることで効率的に犯罪を減少させる，警察による場所ベースの防犯活動の重要性を主張している (Sherman and Weisburd 1995) ．さらに，環境の特性や，過去の犯罪データを統計的手法を用いて，ある場所での将来の犯罪の発生の可能性を予測する「地理的犯罪予測」を活用して，効率的なリソース配置に寄与する，より積極的な警察活動 (プロアクティブ・ポリシング) の有効性が報告されている (Weisburd and Majmundar 2018) ．この場所ベースの防犯戦術を用いて，暴力，強盗等の罪種を対象に実証研究が多く実施されているが，太陽光発電所における金属盗のような，比較的最近に見られるようになった特異な窃盗犯罪では，適用された類例がない．そこで，本研究は，太陽光発電所における金属盗のリスクエリアを特定し，警察による集中的なパトロールの実施による犯罪抑止効果を明らかにすることを目的とする．

２．方法

地理的犯罪予測に基づき，太陽光発電所における金属盗のリスクエリアを特定する．次に予測された場所での警察による集中的なパトロールを実施し，最後にパトロール実施前後の金属盗認知件数を比較し，犯罪抑止効果を評価する．

２.１　分析対象

本研究は，茨城県内の水戸市，筑西市，桜川市，石岡市，鉾田市，土浦市，行方市，稲城市等を管轄する 8 つの警察署を分析対象とする．対象警察署は，茨城県内で令和 5 年の太陽光発電所における金属盗認知件数が上位の 8 警察署である．

対象犯罪データは，2021 年 10 月から 2025 年 1 月の期間に 8 警察署で認知した太陽光発電所における金属盗認知件数 (2,461 件) である．発生日時，発生場所，手口等が記録され，ジオコーディングされた発生点を対象とする．また，地理的犯罪予測に必要な環境要因データは，R3 土地利用 3 次メッシュデータ (森林，荒地等) ，R2 国勢調査メッシュデータ (一般世帯数総数，65 歳以上人口等) ，R4 気象データ (降水量，平均気温) ，太陽光発電所

*　正会員　警察大学校警察情報通信研究センター（Police Info-Communications Research Center）
〒183-8558　東京都府中市朝日町 3-12-1　E-mail：t.no.kt@nparc.npa.go.jp
**　非会員　警察大学校警察情報通信研究センター（Police Info-Communications Research Center）
***　正会員　東京大学 情報基盤センター（Information Technology Center, The University of Tokyo）

の発電出力データである．データソースは，オープンデータの国土交通省の国土数値情報，経済産業省の再生可能エネルギー発電設備情報等を利用した．

２.２　金属盗リスクエリア

Perry et al. (2013) は，地理的犯罪予測手法を，ホットスポット分析，回帰分析，データ マイニング，Near-repeat 分析，時空間分析，Risk Terrain 分析の 6 つに分類した．このうち，ホットスポット分析とは，犯罪数や犯罪発生率の高い地域を特定する手法であり，一般的なマッピング手法はカーネル密度推定法である (Chainey et al. 2008) ．本研究では，金属盗の予測にカーネル密度推定法 (ガウス型カーネル関数) を採用した．一方，Rummens and Hardyns (2020) は，複数の機械学習モデルを組み合わせるアンサンブル手法が，地理的犯罪予測にも一定の有用性があることを示唆している．そこで，2 つの手法 : 過去の事象の集積を示すホットスポット分析と，環境要因からモデルで推定する機械学習モデルを組み合わせる方法を採用した．

機械学習アルゴリズムには，LightGBM を採用した．なお, random forest, XG Boost 等と予測精度を比較した結果，最良であった (詳細は紙幅の都合で省略) ．

カーネル密度推定法で推定した予測エリアと LightGBM で推定した予測エリアの和集合をとった（予測エリアを 1，非予測エリアを 0 としたとき，2 つの手法で 1/1, 1/0, 0/1 となるユニットを全てリスクエリアとした）ものを金属盗リスクエリアと定義する．

まず，ホットスポット分析において，金属盗を予測するために，PAI を用いてパラメータ調整を行った．将来の犯罪を予測する能力を評価する指標 PAI (Prediction Accuracy Index) を式 (1) に示す (Chainey et al. 2008) ．Hit Rate は，研究エリアの全犯罪数に対する，予測される場所における実際の犯罪数の割合，Area Percentage は研究エリア面積に対する予想される場所の面積の割合である．Hit Rate が高く，Area Percentage が小さいほど予測精度は高い．

$$PAI = \frac{HitRate}{AreaPercentage} = \frac{\left(\frac{n}{N}\right) * 100}{\left(\frac{a}{A}\right) * 100} \qquad (1)$$

n は予測場所の犯罪数，N は研究エリア内の犯罪数，a は予測場所の面積，A は研究エリア面積である．

予測期間の 1 か月 (2024 年 4 月) に対する学習データ期間をそれ以前の 24 か月とし，バンド幅は 1.0 km，1.5 km，2.0 km，2.5 km の 4 パターンで変化させた．これらのうち，最も PAI が高くなるバンド幅を最終パラメータとして，カーネル密度推定法を用いて予測エリアを特定した．なお，セルサイズは固定 100 m とし，犯罪密度 (件/m^2) は平均値から標準偏差の 1 倍以上高い値を扱った．

次に，機械学習アルゴリズムの LightGBM を用いて，データセットから金属盗の有無を予測するモデルを構築した. データセットは, 目的変数には金属盗データ (2023 年 10 月から 2024 年 3 月) ，説明変数には目的変数の以前の 3，6，12，24 か月の金属盗データと上記の環境要因データで構成される．

予測対象のグリッドは，介入対象の警察署管轄エリア及びそのエリアの周囲 1 km バッファの領域に存在する太陽光発電所を含むグリッドとした．空間解析単位は，太陽光発電所の発電出力から求めた面積を考慮し，500×500 m のグリッドセルとした．

データセットを学習データ，検証データ，テストデータにランダムに分割し, 10 分割交差検証を用いて, 学習・検証データでハイパーパラメータチューニングを，テストデータでモデルの精度評価を行った. 評価指標は F1 スコア，AUC とした．当該モデルを用いて予測エリアを特定した．なお，本研究で使用するデータは金属盗の有無クラス比に偏りがある不均衡データであるため，アンダーサンプリングでバランシングを行った．

最後に，ホットスポット分析と機械学習モデルの組み合わせにおける予測精度を予測期間 1 か月 (2024 年 4 月) として PAI で評価した．

２.３　警察による集中的なパトロールの実施

実験条件は，茨城県内の 8 つの警察署において，実験期間は 8 か月（2024 年 6 月～2025 年 1 月），活動内容はリスクエリア内の太陽光発電所に対するパトカーでの防犯パトロール等 (人員増員なし) である. 指定されたリスクエリアでの防犯パトロールを活動の主とするが，優先すべき事件・事故への対応も可能とした．

金属盗リスクエリアマップは，介入対象の各警察署について，管轄エリア全域と，そのエリア内の交番管轄エリアごとに実験期間の月単位で作成し，提供した．対象月の前々月までの犯罪データでホットスポット分析，機械学習モデルを更新した上で, 各月のマップを作成した．

また, 当該マップには, パトロールの効率化のために，署から各リスクエリアを最短経路で結び，巡回する経路を構築した．各リスクエリアと重なる道路リンクの中央点にポイントを置き，このポイントをパトロールの巡回点とした．本研究では，リスクエリアを最短距離で巡回する経路の構築に，Esri 社の ArcGIS Network Analyst で提供されている VRP (Vehicle Routing Program) を用いた．

２.４　パトロール実験の抑止効果の評価

集中パトロールの抑止効果の評価は，傾向スコアマッチングと差の差分析 (Difference-in-Differences, 以降 DID) を用いた．今回の介入対象となった 8 警察署は，ランダムに選ばれたものではないため，処置群／対照群の単純比較は難しいが，共変量をもとにマッチングした地理的条件の類似した署と，事前／事後の差を比較することで，

介入の効果を可能な限り推定した．また，介入前における処置群と対照群における金属盗認知件数の変化を可視化し，DID の前提条件である平行トレンド仮定を概ね満たしていると仮定した．DID 推定の OLS を式 (2) に示す．

$$y_{ist} = \beta_0 + \beta_1 X_{ist} * T_{ist} + \beta_2 T_{ist} + \varepsilon_{ist} \qquad (2)$$

DID 推定値はβ_1，sは処置・対照の別，y_{ist}は時間tにおける署iの認知件数，X_{ist}は処置・対照の別のダミー変数，T_{ist}は介入前・後のダミー変数，ε_{ist}は誤差項を示す．

処理群と対照群との間のセレクションバイアスを調整するための傾向スコアマッチングでは，介入前の共変量を用いてロジスティック回帰で傾向スコアを求め，1 対 1 の最近傍マッチングを行った．対照群の候補は，本介入実施の影響が少ない，処置群の周辺署を除いた警察署 (茨城県外の署も含む) を選定した．共変量は，R3 土地利用 3 次メッシュデータ (他農用地，荒地) ，R6 署管轄の人口，太陽光発電所の発電出力である．

３．結果と考察

３.１　金属盗リスクエリア

まず，ホットスポット分析において，PAI を用いたパラメータ調整の結果を表 1 に示す．バンド幅が小さいほど，PAI が大きく，本研究の最終パラメータはバンド幅 1 km とした．

次に LightGBM の予測モデルにおいて，テストデータを用いた予測精度の F1 値は 0.40, AUC は 0.67 であった．

また，予測モデルの精度に寄与する特徴量の重要度を count-based feature importance を指標として算出した．特徴量の重要度の 1 位は「太陽光発電所の発電出力」，2 位は「一般世帯数総数」であった．

金属盗とこれらの特徴量との空間的関連性を Lee's L 統計によって測定した．Lee's L 統計は，2 つの連続型変数の空間的関連度を-1～1 の範囲で計算し，Lee's L 統計値が 1 に近い場合，変数間に強い正の相関関係，-1 に近い場合，変数間に強い負の相関関係を示す．連続型変数への変換は，Box-Cox 累乗関数を用いた．2 変量の空間的関連性を計算した結果を表 2 に示す. Lee's L 統計値から，「太陽光発電所の発電出力」は弱い正の相関が，「一般世帯数総数」は弱い負の相関があるとわかる．

日常活動理論では，犯罪が発生するためには，有能な監視者が不在の場所と時間で，動機ある犯罪者が適切なターゲットに出会うことで発生すると説明している (Cohen and Felson 1979) ．太陽光発電所の出力が大きく，人口が少ない場所は，窃盗対象である金属ケーブルが多く，監視性が低いことから，犯罪の発生と関係があると考えられる．

表 1　バンド幅の変化による PAI

BandWidth (km)	PAI	Hit Rate
1.0	7.14	0.57
1.5	6.04	0.60
2.0	4.97	0.56
2.5	4.27	0.52

表 2　2 変量の空間的関連性

特徴量	Lee's L	Pearson 相関係数
太陽光発電所発電出力	0.11**	0.12
一般世帯数総数	-0.16**	-0.20

*p<0.05, **p<0.01, ***p<0.001

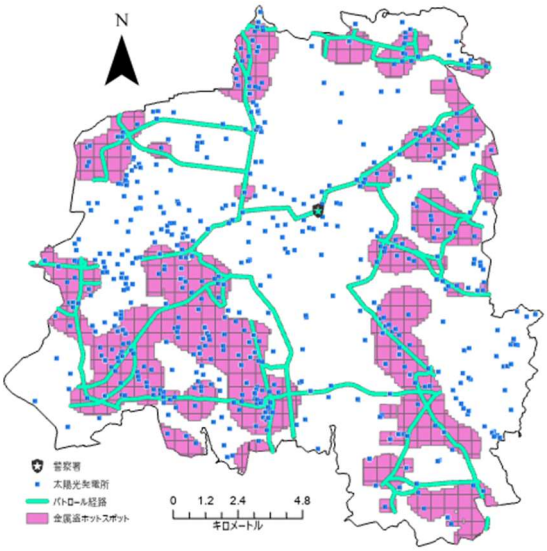


図 1　金属盗リスクエリア (A 警察署)

最後に，介入対象の 8 警察署のうち，A 警察署における金属盗リスクエリアを図 1 に示す．リスクエリア内の点群は，発生が予測された太陽光発電所である．各リスクエリアを結ぶ線は，警察署から最短経路でリスクエリアを巡回する経路である．

太陽光発電所は，A 警察署管轄エリア内の東・西部，中心部に多く分布していることがわかる．予測エリアは，東・西部に離散的に存在しており，結果，対象警察署内のリスクエリアとして予測されたグリッドに含まれる発電所は全体の約 1/3 であった．

また，ホットスポット分析と機械学習モデルを組み合わせて作成した予測エリアにおける予測精度は，PAI 8.08, Hit Rate 0.71 であった．PAI は，ホットスポット分析単独に比べて，機械学習との組み合わせによって 13%上昇した．Rummens and Hardyns (2020) の指摘のとおり，金属盗においても地理的犯罪予測手法のアンサンブルは有効で

表 3　バランステスト結果

共変量	処置群 (N=8) 平均 (SD)	対照群(N=8) 平均 (SD)	t-test
R6署内人口 (千人)	117.39 (89.73)	110.34 (61.09)	-0.18
農地 (km^2)	62.61 (24.43)	61.30 (36.75)	-0.08
荒野 (km^2)	2.64 (1.48)	2.34 (2.61)	-0.26
太陽光発電所発電出力 (mw)	5701.43 (5992.60)	6295.28 (7864.20)	0.16

*p<0.05, **p<0.01, ***p<0.001

表 4　DID 推定結果

Dependent variable	metal theft
Month×Treatment	-3.93(4.71)
Month	-2.59(3.33)
Treatment	11.03(3.33) **
Constant	8.55(2.36) **
R squared	0.34

注:Treatment は処置・対照の別のダミー変数, Month は介入前・後 8 ヵ月のダミー変数，Month×Treatment は交互作用変数，括弧は標準誤差，*p<0.05, **p<0.01, ***p<0.001

あると考えられる．

３.２　抑止効果の評価の結果

傾向スコアマッチングによる共変量のバランスチェックを行った結果を表 3 に示す．マッチング後の処置群と対照群の各共変量に有意な差は認めなかった．処置群と対照群に共通する共変量はバランスがとれていることがわかる．

次に，金属盗リスクエリア内での集中パトロールの犯罪抑止効果について，DID 推定結果を表 4 に示す．統計的な有意な結果ではないが，処置群は対照群と比べて 3.93 件/月の減少であった．本研究における集中パトロールの介入実験は一定の抑止効果があると考える．しかしながら，本抑止効果の評価には，以下の限界がある．

第一に，対照群の選定において，茨城県周辺県の警察署を含んでおり，これらの警察署の金属盗抑止対策を制御できていない．よって，対照群の選定等のバイアスによる，パトロール以外の抑止対策等の交絡因子の影響を排除しきれない．第二に，実験期間内において，介入対象の警察署における，金属盗リスクエリア内でのパトロールの活動量を正確に測定できておらず，介入による抑止効果は限定的であると考えられる．

４．おわりに

本研究は，地理的犯罪予測に基づき，太陽光発電所における金属盗のリスクエリアを特定した上で，次に予測された場所での警察による集中的なパトロールを実施し，最後にパトロール実施前後の金属盗認知件数を比較し，犯罪抑止効果を評価することを目的とした．

結果，第一に金属盗リスクエリアの特定において，介入対象の警察署管内エリアでの予測精度は，Hit Rate 0.71 であり，精度良く金属盗を予測できた．また，ホットスポット分析と機械学習 (LightGBM) の組み合わせは，ホットスポット分析単独に比べ, PAI が高くなり，アンサンブルの有効性が示された．さらに，LightGBM における特徴量の重要度及び空間的関連性を測る Lee's L 統計値から，太陽光発電所の発電出力や一般世帯数は，金属盗の発生と関係があることが明らかになった．

第二に，金属盗リスクエリア内での警察による集中的なパトロールを実施し，金属盗の抑止効果を傾向スコアマッチング及び DID を用いて推定した結果，統計的に有意な結果ではないが，一定の条件下で処置群は対照群と比べて減少した．

本研究で明らかになった，機械学習等を用いた金属盗の地理的な予測可能性，また，警察によるプロアクティブな場所ベース防犯戦術による抑止効果は，我が国の金属盗防犯対策の実践に寄与した．

今後，本研究の限界に対応しつつ，地理的犯罪予測の精度向上に向けて，新たなアンサンブル手法の検討や空間的な規模が異なる地域別，罪種別のパトロール抑止効果の差異を明らかにする実証的な検討が必要である．

謝辞

研究にご協力を賜った茨城県警察の関係者の方々に厚く御礼申し上げる．

参考文献

警察庁 (2025) 「金属盗対策に関する検討会報告書」

Chainey, S., Tompson, L., & Uhlig, S. (2008) The utility of hotspot mapping for predicting spatial patterns of crime. Security journal, 21, 4-28.

Cohen, L. E. & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. American sociological review, 588-608.

Majmundar, M. K., & Weisburd, D. (Eds.). (2018) Proactive policing: Effects on crime and communities. National Academies Press.

Perry, W., Mclnnis, B., Price, C., Smith, S., & Hollywood, J. (2013) Predictive policing: The role of crime forecasting in law enforcement operation. Rand Corporation.

Rummens, A., & Hardyns, W. (2020) Comparison of near-repeat, machine learning and risk terrain modeling for making spatiotemporal predictions of crime. Applied spatial analysis and policy, 13(4), 1035-1053.

Sherman, L.W. & Weisburd, D. (1995) General Deterrent Effects of Police Patrol in Crime Hot Spots. A Randomized, Controlled Trial, Justice Quarterly, 12, 635-48.