

MODELS MATEMÀTICS PER A LA PREDICCIÓ DELS ROBATORIS AMB FORÇA A DOMICILI A CATALUNYA

Pere Boqué Busquet

Per citar o enllaçar aquest document:

Para citar o enlazar este documento:

Use this url to cite or link to this publication:

<http://hdl.handle.net/10803/674301>

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.



Tesi Doctoral

**Models matemàtics
per a la predicció dels robatoris
amb força a domicili a Catalunya**

Pere Boqué Busquet



TESI DOCTORAL

**Models matemàtics
per a la predicció dels robatoris
amb força a domicili a Catalunya**

Pere Boqué Busquet

2021

Programa de doctorat en Dret, Economia i Empresa

Dirigida per:

Dr. Marc Saez Zafra i Dra. Laura Serra Saurina

Memòria presentada per optar al títol de doctor/a per la Universitat de Girona



El Dr. Marc Saez Zafra i la Dra. Laura Serra Saurina, de la Universitat de Girona i membres del Grup de Recerca en Estadística, Econometria i Salut (GRECS),

CERTIFIQUEN:

Que el treball titulat “Models matemàtics per a la predicció dels robatoris amb força a domicili a Catalunya”, que presenta Pere Boqué Busquet per a l’obtenció del títol de doctor, ha estat realitzat sota la seva direcció.

I, perquè així consti i tingui els efectes oportuns, signem aquest document.

Girona, 21 de juliol de 2021

AGRAÏMENT

Gràcies Marc i Laura, els meus directors de tesi, per proposar-me iniciar aquesta recerca acadèmica i per la paciència i dedicació que heu tingut al llarg d'aquest camí. No me n'hagués sortit sense les grans línies mestres del Marc ni sense les meravelloses correccions i apreciacions de la Laura. Formeu un gran tàndem, i us estic eternament agraït.

Gràcies a la Maria Agüareles i a la Marta Pellicer per estendre'm la mà i obrir-me les portes de la Universitat de Girona. Vosaltres veu iniciar tot això i segur que ho continuarem.

Gràcies Marina i Jofre, els meus fills, per entendre que el pare us manlevés una part del temps que us corresponia i que espero retornar-vos. Estic molt orgullós de vosaltres i desitjo que al llarg de la vida també trobeu en l'estudi i el coneixement un plaer imprescindible.

Gràcies a la meva família pel vostre suport incondicional. Als meus pares Onofre i Carme, per ser sempre al peu del canó, i a la meva germana Montse, una font d'energia i inspiració inesgotable.

Gràcies molt especials a la Marta Sendra, el meu puntal vital i directora a l'ombra d'aquesta tesi. Quina sort haver pogut compartir, debatre, discutir i riure amb tu les principals idees i dubtes d'aquesta recerca, i escoltar els teus sempre savis consells.

Gràcies al George Mohler, matemàtic de referència en la predicció delictiva, per compartir el teu coneixement sense traves ni esperar res a canvi.

Gràcies als meus companys de la Unitat d'Anàlisi i del Departament d'Interior, per l'interès i suport que m'han mostrat en aquest treball. Gràcies Francesc Guillén per la teva ajuda, pel teu mestratge, per introduir-me les grans idees del món de la gestió de la seguretat, per comptar amb mi i per permetre'm participar en els projectes europeus.

També gràcies als companys de l'ABP Berguedà, i molt especialment a l'Escamot 1, per l'acolliment que m'heu fet i per la vostra implicació diària i exemplar amb la feina de patrullar.

I, finalment, gràcies als Mossos d'Esquadra. Una institució encara jove, que l'any 2009 finalitzava el desplegament territorial a Catalunya i que té el repte de convertir-se en una policia moderna i puntera. Espero que aquesta tesi hi contribueixi, i per això us la dedico.

Moltíssimes gràcies a tots,

Pere

LLISTAT D'ARTICLES

Aquesta tesi es presenta en format monografia que incorpora els resultats en forma d'articles.

ARTICLE 1 (PUBLICAT)

Títol: “Surfing” burglaries with forced entry in Catalonia. Large-scale testing of the theory of near repeat victimisation

Autors: Boqué P, Serra L, Sáez M

Revista: European Journal of Criminology

Factor d'impacte [2019]: 1.671; (Q1 Social Sciences (Law), position 37 out of 751)

DOI: 10.1177/1477370820968102

ARTICLE 2 (ENVIAT)

Títol: Need to go further: using INLA to discover limits and chances of spatiotemporal prediction in heterogeneous environments.

Autors: Boqué P, Sáez M, Serra L

Enviat a la revista: “European Journal of Criminology”

ARTICLE 3 (PENDENT D'ENVIAR)

Títol: Burglaries constellations: applying Principal Components Analysis (PCA) to discover the hot trail left by the all-in-one spatiotemporal repeat pattern in heterogeneous environments

Autors: Boqué P, Serra L, Sáez M

LLISTAT DE FIGURES I TAULES

Llistat de figures

Figura 1. -Evolució dels robatoris amb força a Catalunya (2009-2015)	1
Figura 2. -Ubicació de Catalunya al mapa d'Europa	4
Figura 3. -Mapa de Catalunya	5
Figura 4. -Mapa de les Regions Policials de Catalunya	5
Figura 5. -Mapa de distribució dels principals tipus de cobertes del sòl a Catalunya	6
Figura 6. -Imatge aèria d'una zona residencial de Long Beach (Los Angeles, EUA)	7
Figura 7. -Imatge aèria d'una zona residencial de Barcelona (Catalunya)	7
Figura 8. -Mapa de la Regió Policial Metropolitana Nord.....	8
Figura 9. -Mapa de Catalunya amb una malla de cel·les de 5kmx5km sobreposada	23
Figura 10. -Exemple de ratxes de fets setmanals per diferents cotes d'activació	25
Figura 11. -Exemples gràfics de simulacions de la distribució de Poisson.....	27
Figura 12. -Representació gràfica del coeficient de variació (CV)	29
Figura 13. - Diagrama d'una loteria (Formulació agregada)	37
Figura 14. - Diagrama d'un procés estocàstic (Formulació desagregada)	38

Llistat de taules

Taula 1. - Tipus de victimitzacions repetides.....	9
Taula 2. - Exemple de ratxes de fets setmanals per diferents cotes d'activació.....	25
Taula 3. - Explicació de les 3 situacions possibles en funció del CV.....	30

RESUM

El *predictive policing* és una orientació estratègica global de la policia que posa èmfasi en l'anticipació dels delictes per prevenir-los. El seu desenvolupament prové de varis camps del coneixement, en què la criminologia ha creat un cos teòric sobre els patrons delictius en l'espai i el temps, la matemàtica i l'estadística han modelat aquests patrons i les TIC n'han fet possible la implementació als cossos policials. La recerca segueix oberta, acompanyada d'incerteses sobre la seva modelització i aplicació. Aquesta tesi fa una revisió del coneixement científic, teòric i pràctic, relacionat amb el *predictive policing* i les possibilitats d'aplicar-lo a Catalunya per predir els robatoris amb força a domicili, que van tenir un increment important entre els anys 2009 a 2015 i, des de llavors, s'han mantingut amb uns índexs delictius elevats.

S'ha analitzat el patró espai temps de la victimització repetida i quasi repetida. S'ha comprovat que els robatoris a Catalunya es distribueixen temporalment en onades no aleatòries quan s'analitzen les sèries en cel·les grans, de 5km de costat. A petita escala, en cel·les de menys d'1km², aquest patró és menys detectat. Això ha suggerit que el patró de victimització quasi repetida a Catalunya es manifesta de forma diferent del que és habitual en altres països, on bàsicament s'ha detectat en una escala micro i en entorns urbans. Aquest diferència es relaciona amb l'heterogeneïtat de la geografia catalana, especialment la urbana i residencial, de manera que s'ha proposat ampliar el concepte de repetició en la proximitat geogràfica, al de repetició en zones diferents, també allunyades, però que tenen semblances des del punt de vista de l'oportunitat delictiva. Els grups criminals escollirien aquestes zones d'antuvi i les victimitzarien durant períodes de varies setmanes, generant les onades de robatoris mitjançant un nou patró espai-temps que es proposa: l' "all-in-one".

S'ha comprovat la viabilitat d'aquest nou patró per explicar, predir i prevenir els robatoris amb força a domicili a Catalunya. Aquest patró genera xarxes de zones disperses, relativament semblants, que tenen tendència a activar-se a la vegada i que s'han anomenat "constel·lacions de robatoris". Es suggereix que en entorns geogràfics semblants al de Catalunya, com pot ser gran part del sud d'Europa, la distribució espaciotemporal dels robatoris també podria seguir aquests mateixos patrons proposats.

RESUMEN

El *Predictive policing* es una orientación estratégica global de la policía que pone énfasis en la anticipación de los delitos para prevenirlos. Su desarrollo proviene de varios campos del conocimiento, en el que la criminología ha creado un cuerpo teórico sobre los patrones delictivos en el espacio y el tiempo, la matemática y la estadística han moldeado estos patrones y las TIC han hecho posible la implementación en los cuerpos policiales. La investigación sigue abierta, acompañada de incertidumbres sobre su modelización y aplicación. Esta tesis hace una revisión del conocimiento científico, teórico y práctico, relacionado con el *Predictive policing* y las posibilidades de aplicarlo en Cataluña para predecir los robos con fuerza en domicilio, que tuvieron un incremento importante entre los años 2009 a 2015 y, desde entonces, se han mantenido con unos índices delictivos elevados.

Se ha analizado el patrón espacio tiempo de la victimización repetida y casi repetida. Se ha comprobado que los robos en Cataluña se distribuyen temporalmente en olas no aleatorias cuando se analizan las series en celdas grandes, de 5km de lado. A pequeña escala, en celdas de menos de 1km², este patrón es menos detectado. Esto ha sugerido que el patrón de victimización casi repetida en Cataluña se manifiesta de forma diferente de lo que es habitual en otros países, donde básicamente se ha detectado en una escala micro y en entornos urbanos. Esta diferencia se relaciona con la heterogeneidad de la geografía catalana, especialmente la urbana y residencial, por lo que se ha propuesto ampliar el concepto de repetición en la proximidad geográfica, al de repetición en zonas diferentes, también alejadas, pero que tienen semejanzas desde el punto de vista de la oportunidad delictiva. Los grupos criminales escogerían estas zonas de antemano y las victimizarían durante periodos de varias semanas, generando las olas de robos mediante un nuevo patrón espacio-tiempo que se propone: el "all-in-one".

Se ha comprobado la viabilidad de este nuevo patrón para explicar, predecir y prevenir los robos con fuerza en domicilio en Cataluña. Este patrón genera redes de zonas dispersas, relativamente similares, que tienen tendencia a activarse a la vez y que se han denominado "constelaciones de robos". Se sugiere que en entornos geográficos similares al de Cataluña, como puede ser gran parte del sur de Europa, la distribución

espacio-temporal de los robos también podría seguir estos mismos patrones propuestos.

ABSTRACT

Predictive policing is a global strategic orientation of the police that emphasizes the anticipation of crimes to prevent them. Its development comes from various fields of knowledge, in which criminology has created a theoretical body on criminal patterns in space and time, mathematics and statistics have modelled these patterns and ICTs have made possible their implementation in police forces. The research remains open, accompanied by uncertainties about its modelling and application. This thesis reviews the scientific, theoretical and practical knowledge related to predictive policing and the possibilities of applying it in Catalonia to predict residential burglaries with forced entry, which had a significant increase between 2009 and 2015 and, since then, they have remained at high crime rates.

The spatiotemporal pattern of repeated and near-repeated victimization has been analysed. It has been verified that the residential burglaries in Catalonia are temporarily distributed in non-random waves when analysing the series in large cells, of 5km on each side. On a small scale, in cells less than 1km², this pattern is less detected. This has suggested that the pattern of near repeated victimization in Catalonia is manifested differently than is usual in other countries, where it has basically been detected on a micro scale and in urban environments. This difference is related to the heterogeneity of Catalan geography, especially with respect to urban and residential, so it has been proposed to expand the concept of repetition in geographical proximity, to repetition in different areas, also far away, but which have similarities from the point of view of criminal opportunity. Criminal groups would choose these areas in advance and victimize them for periods of several weeks, generating waves of burglaries through a proposed new spatiotemporal pattern: the "all-in-one."

The viability of this new pattern to explain, predict and prevent burglaries at home in Catalonia has been verified. It generates networks of relatively similar scattered areas that tend to activate at the same time and have been called "burglaries constellations." It is suggested that in geographical environments similar to Catalonia, such as much of southern Europe, the spatiotemporal distribution of residential burglaries could also follow these same proposed patterns.

TAULA DE CONTINGUTS

AGRAÏMENT	III
LLISTAT D'ARTICLES.....	V
LLISTAT DE FIGURES I TAULES.....	VI
RESUM	VII
RESUMEN.....	VIII
ABSTRACT	X
TAULA DE CONTINGUTS.....	XI
1 INTRODUCCIÓ	1
1.1 Els robatoris amb força a domicili a Catalunya.....	1
1.1.1 <i>El tipus delictiu</i>	1
1.1.2 <i>Dia i l'hora del fet</i>	2
1.1.3 <i>Tipus de domicili i el seu entorn</i>	2
1.1.4 <i>Mesures de seguretat i formes d'accés amb força als domicilis</i>	3
1.1.5 <i>Autors dels robatoris</i>	3
1.1.6 <i>La denúncia i el registre dels fets en les bases de dades policials</i>	3
1.2 El territori.....	4
1.2.1 Catalunya	4
1.2.2 La Regió Policial Metropolitana Nord	7
1.3 Les teories de l'oportunitat delictiva i la conducta criminal.....	8
1.4 La teoria de la victimització quasi repetida	9
1.4.1 <i>Risc estàtic</i>	10
1.4.2 <i>Risc dinàmic</i>	10
1.5 Policing - La prevenció policial de la delinqüència	11
1.5.1 <i>El model actual de la policia</i>	11
1.5.2 <i>Predictive Policing</i>	11
1.5.3 <i>La prevenció operativa: patrulles uniformades</i>	13
1.5.4 <i>La prevenció actual dels robatoris a domicili</i>	14
1 HIPÒTESIS.....	15
2 OBJECTIUS.....	16
3 METODOLOGIA	17
3.1 Background metodològic.....	17
3.1.1 <i>Procés puntual</i>	17
3.1.2 <i>Aleatorietat espacial completa</i>	17

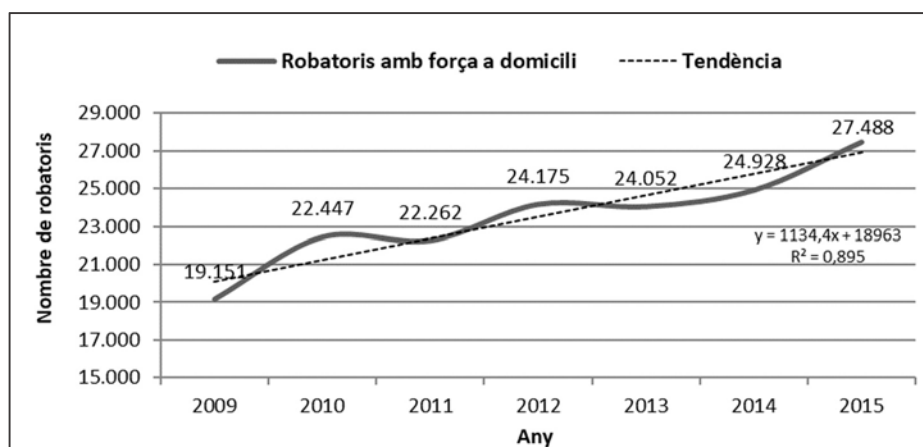
3.1.3	<i>Processos de Poisson no-homogenis</i>	18
3.1.4	<i>Estimació de la intensitat espacial</i>	19
3.1.5	<i>Processos Puntuals Espaciotemporals</i>	19
3.1.6	<i>Separabilitat espaciotemporal</i>	20
3.1.7	<i>Autocorrelació temporal i processos autoregressius (AR)</i>	20
3.1.8	<i>Test de Knox: estudi de la dependència espaciotemporal</i>	20
3.1.9	<i>Model de Short</i>	21
3.1.10	<i>Self-exciting Point Process</i>	21
3.2	Configuració espai-temps	22
3.3	Mapes delinqüencials	24
3.4	Distribució de les sèries de robatoris.....	25
3.4.1	<i>El test de ratxes</i>	25
3.4.2	<i>Ajustament a la distribució de Poisson</i>	26
3.4.3	<i>Criteris d'optimitat</i>	29
3.5	Model Log Gaussià de Cox i el mètode Integrated Nested Laplace Approximation (INLA)	31
3.6	Aplicació del mètode INLA amb el paquet estadístic R	33
3.6.1	<i>Construcció de la malla de punts (mesh)</i>	33
3.6.2	<i>Selecció dels priors</i>	34
3.6.3	<i>Comparativa de models</i>	34
3.6.4	<i>Del camp latent a la predicció pràctica</i>	35
3.7	Anàlisi de components principals	35
3.8	Formulació probabilística del patró “all-in-one”	37
3.9	Altres metodologies.....	38
4	RESULTATS	40
4.1	Article 1.....	41
1	Introduction	43
2	Data and methodology	45
2.1	Data.....	45
2.2	Methods.....	46
3	Results.....	49
4	Discussion	54
5	Conclusions	59
6	References	62
7	Appendix.....	65
4.2	Article 2.....	74

1	Introduction	76
2	Data and methodology	77
2.1	Data.....	77
	The territory: Catalonia.....	77
	Crime data.....	78
2.2	Methods.....	78
2.2.1	INLA specifics when applied	80
3	Results.....	81
3.1	General stability of the data	81
3.2	Predictive models with R-INLA	84
3.2.1	INLA summary results for North Metropolitan Police Region	86
3.2.2	INLA summary results for Girona Police Region	88
3.2.3	INLA summary results for Barcelona Metropolitan Police Region	90
4	Discussion	93
5	Conclusions	98
6	References	100
7	Appendix A: Figures, graphics and tables	105
8	Appendix B: INLA summary results for Eixample district of Barcelona	108
4.3	Article 3.....	110
1	Introduction	112
2	Data and methodology	113
2.1	Data.....	113
2.2	Methods.....	114
3	Results.....	115
4	Discussion	120
5	Conclusions	124
6	References	125
7	Appendix.....	129
5	DISCUSSIÓ	130
6	CONCLUSIÓ	137
7	BIBLIOGRAFIA.....	139

1 INTRODUCCIÓ

1.1 Els robatoris amb força a domicili a Catalunya

Des de l'any 2009 la Policia de la Generalitat – Mossos d'Esquadra (PG-ME) està desplegada a tot el territori de Catalunya on, entre altres, hi té encomanada la funció de garantir la seguretat pública. Això comporta una tasca preventiva que eviti o minori la comissió d'infraccions penals, entre les quals hi ha els robatoris amb força a domicili. L'estadística policial entre els anys 2009 a 2015 mostra una tendència creixent d'aquesta tipologia delictiva. Les accions policials, algunes de les quals específicament dissenyades i dirigides a evitar aquests robatoris, no aconsegueixen neutralitzar l'increment de fets. Aquestes accions solen agrupar-se en dos grans eixos, per una banda les patrulles uniformades fan major presència, identificacions i controls policials en les zones on s'han detectat més casos, guiats, normalment, per mapes delinqüencials. Per altra banda, les unitats d'investigació intenten esbrinar la identitat dels autors, a vegades organitzats en grups criminals, amb la finalitat de detenir-los, desarticular-los i posar-los a disposició dels òrgans judicials.



Font: Elaboració pròpia segons les dades facilitades per la Policia de la Generalitat-Mossos d'Esquadra

Figura 1.-Evolució dels robatoris amb força a Catalunya (2009-2015)

La predicció delictiva pot ser una eina que ajudi a millorar les accions preventives i les investigacions policials. Alguns països han aplicat programes de predicció dels robatoris amb força a domicili amb resultats satisfactoris.

1.1.1 El tipus delictiu

El delictiu de robatori amb força a domicili està tipificat en el codi penal espanyol (Llei Orgànica 10/1995, de 23 de novembre) en el títol XIII, capítol II.

Artículo 237: Son reos del delito de robo los que, con ánimo de lucro, se apoderaren de las cosas muebles ajenas empleando la fuerza en las cosas para acceder o abandonar el lugar donde se encuentran o violencia o intimidación en las personas, sea para cometer el delito, para proteger la huida, o sobre los que acudiesen en auxilio de la víctima o que persiguieren.

Artículo 238: Son reos del delito de robo con fuerza en las cosas los que ejecuten el hecho cuando concurra alguna de las circunstancias siguientes:

1º Escalamiento

2º Rompimiento de pared, techo o suelo, o fractura de Puerta o ventana.

3º Fractura de armarios, arcas u otra clase de muebles u objetos cerrados o sellados, o forzamiento de sus cerraduras o descubrimiento de sus claves para sustraer su contenido, sea en el lugar del robo o fuera del mismo.

4º Uso de llaves falsas

5º Inutilización de sistemas específicos de alarma o guarda

El delictes es compleix quan es donen les circumstàncies descrites en els articles anteriors. Altrament, es podran tipificar de forma diferent. En el cas que es faci força en les coses però els lladres no aconseguixin entrar al domicili o no s'enduguin res, el fet es pot tipificar com a delictes de danys, en alguns casos, o com a temptativa de robatori en altres. En aquest segon cas, es considerarà robatori amb força a domicili a efectes d'estadística policial. També es pot donar el cas que els lladres aconseguixin entrar al domicili sense fer ús de la força, és a dir, sense que es donin cap de les cinc circumstàncies de l'article 238 del codi penal. Seria, per exemple, quan un lladre aprofita que hi ha una porta oberta i accessible del domicili i sostreu alguna cosa del seu interior. En aquest cas es tractaria d'un delictes de furt.

1.1.2 Dia i l'hora del fet

En el moment que es denuncia un robatori es fa constar l'hora en què suposadament ha ocorregut. Al tractar-se de robatoris amb força, els lladres eviten ser detectats tant pels propietaris del domicili com per terceres persones que podrien alertar a la policia, de manera que, sovint, no se sap l'hora exacte del fet i, a vegades, tampoc el dia. La majoria de robatoris són en horari diürn i en primeres residències, llavors el moment del robatori es pot establir en un dia concret i en una franja de poques hores en funció de quan els propietaris van abandonar al domicili al matí i quan hi van tornar. A la denúncia es fa constar un dia i hora d'inici del fet i un dia i hora de finalització del fet. En el cas de robatoris en segones residències, hi pot haver varis dies o setmanes de diferència entre les dues dates.

1.1.3 Tipus de domicili i el seu entorn

Urbanísticament i geogràfica, Catalunya és un país divers i heterogeni. Les ciutats i pobles solen visualitzar-se com un mosaic de zones relativament petites, a vegades coincidents amb el límit dels barris, que mostren tipus de construccions com edificis antics, alguns construïts fa varis segles, zones residencials de pisos construïts els anys d'immigració interna (1950-1980), zones de torres i xalets, zones de cases adossades, zones de pisos i barris moderns construïts els últims 20 anys, amb alguns gratacels, zones d'apartaments turístics, zones rurals, masies i cases de pagès disseminades, etc.

Relacionat amb els tipus de domicilis i entorns també hi ha els diferents nivells de vida dels seus ocupants o propietaris que, a la vegada, es relaciona amb el tipus de botí que els lladres esperen obtenir.

1.1.4 Mesures de seguretat i formes d'accés amb força als domicilis

La varietat d'entorns i construccions va aparellada amb la varietat de mesures de seguretat, que estan condicionades pels tipus d'accessos als domicilis, com disposar o no de porter automàtic, tipus de porta, tipus de pany, tipus de tancaments de les diferents obertures, accessos a la teulada, celobert, jardí o pàrquing, presència de balcons o terrasses, facilitats d'escalament, ascensor, tanca exterior, càmeres de videovigilància, gossos, etc.

La forma més comuna d'accés al domicili és mitjançant el forçament de la porta. Existeixen diferents mètodes de fer-ho en funció del tipus de porta. Alguns mètodes coneguts són el sistema de la radiografia, l'ús de rossinyols o el trencament del bombí amb una eina tipus bec de lloro. Una altra forma habitual d'accés és forçant les finestres, a vegades amb escalament previ.

1.1.5 Autors dels robatoris

Els autors que s'han pogut identificar cometent robatoris amb força a domicili a Catalunya tenen perfils diversos. Alguns d'ells es qualifiquen de delinqüents locals, que sovint actuen sols i d'una forma poc planificada i oportunista. Altres formen part de grups més o menys organitzats, actuen de forma més planificada i escullen els domicilis on robar en funció de la seva experiència prèvia d'èxits i fracassos. Alguns formen part de grups criminals especialitzats, sovint composts amb persones no nacionals, amb alta mobilitat al territori, que a vegades es desplacen amb vehicles d'alta gamma robats o amb vehicle de lloguer i que canvien sovint els seus membres per no ser detectats o considerats grups organitzats.

Tanmateix, l'autoria dels robatoris amb força a Catalunya només es coneix en aproximadament el 10% dels casos, que són els que es consideren policialment resolts.

1.1.6 La denúncia i el registre dels fets en les bases de dades policials

En el moment que una persona detecta que ha estat víctima d'un robatori amb força a domicili, decideix si denuncia els fets a la policia o no. En alguns casos, quan no s'observen danys o aquests són menors, i/o quan no han sostret res de l'interior del domicili o el que han sostret és de poc valor, la víctima no denuncia els fets.

Una altra casuística habitual és que la víctima del robatori avisi a la policia trucant al telèfon d'emergències 112 just en el moment en què s'adona del que ha passat. La policia es persona al lloc i normalment aixeca una acta de comprovació dels fets on hi consten els danys ocasionats, en cas que n'hi hagin. Aquesta actuació policial no és equivalent a la denúncia, sinó que queda registrada com a incident sense denúncia. El fet no quedarà recollit com a robatori amb força a domicili si no s'acaba personant la víctima o denunciant en una comissaria per formalitzar la denúncia.

Tots els fets que es denuncien o que han generat incidents (amb o sense denúncia) queden registrats en les bases de dades policials. El registre d'aquestes dades consisteix en informació estructurada i altra en format de text obert. El sistema informàtic de la PG-ME permet consultar i extreure informes de les dades estructurades, mentre que les dades no estructurades o en format de text obert s'hi ha d'accedir de forma individual i concreta.

Referent als robatoris amb força a domicili, l'estructuració de les dades conté els següents camps:

- Tipus de fet (delicte de robatori amb força)
- Lloc del fet (municipi, adreça, coordenades)
- Tipus de domicili (pis, casa, primera o segona residència)
- Data i hora d'inici del fet
- Data i hora de finalització del fet
- Data i hora de la formalització de la denúncia

Pel que fa a les dades estructurades associades als incidents són:

- Tipus d'incident (robatori amb força o alarma de robatori en habitatge)
- Lloc de l'incident (municipi, adreça, coordenades)
- Tipus de lloc (Domicili particular o Altres habitatges)
- Data i hora d'inici de l'incident
- Data i hora de finalització de l'incident

El sistema informàtic de la PG-ME disposa de dues bases de dades diferents, una per delictes i una d'incidents, no relacionades internament de forma automàtica.

1.2 El territori

1.2.1 Catalunya

Catalunya és un país de l'Europa mediterrània, a la costa oriental de la península Ibèrica, de la que n'ocupa un 5,5% de la superfície. Limita al nord amb la serralada pirinenca i França, a l'est amb la mar mediterrània, i a l'oest i al sud amb l'estat espanyol. La superfície és de 3.895km², i té una població de 7,7 milions d'habitants.



<https://ca.wikipedia.org/wiki/Catalunya#/media/File:E.U-Catalonia.png>

Figura 2.-Ubicació de Catalunya al mapa d'Europa



Font: Institut Cartogràfic i Geològic de Catalunya – ICGC

Figura 3.-Mapa de Catalunya

El territori es divideix en un total de 9 regions policials (Figura 4). Cada regió policial agrupa un conjunt d'àrees bàsiques policials (ABP), que són la unitat bàsica d'implantació territorial, des d'on surten la majoria de les patrulles, i que està definida per criteris geogràfics i policials i solen incloure diversos municipis.

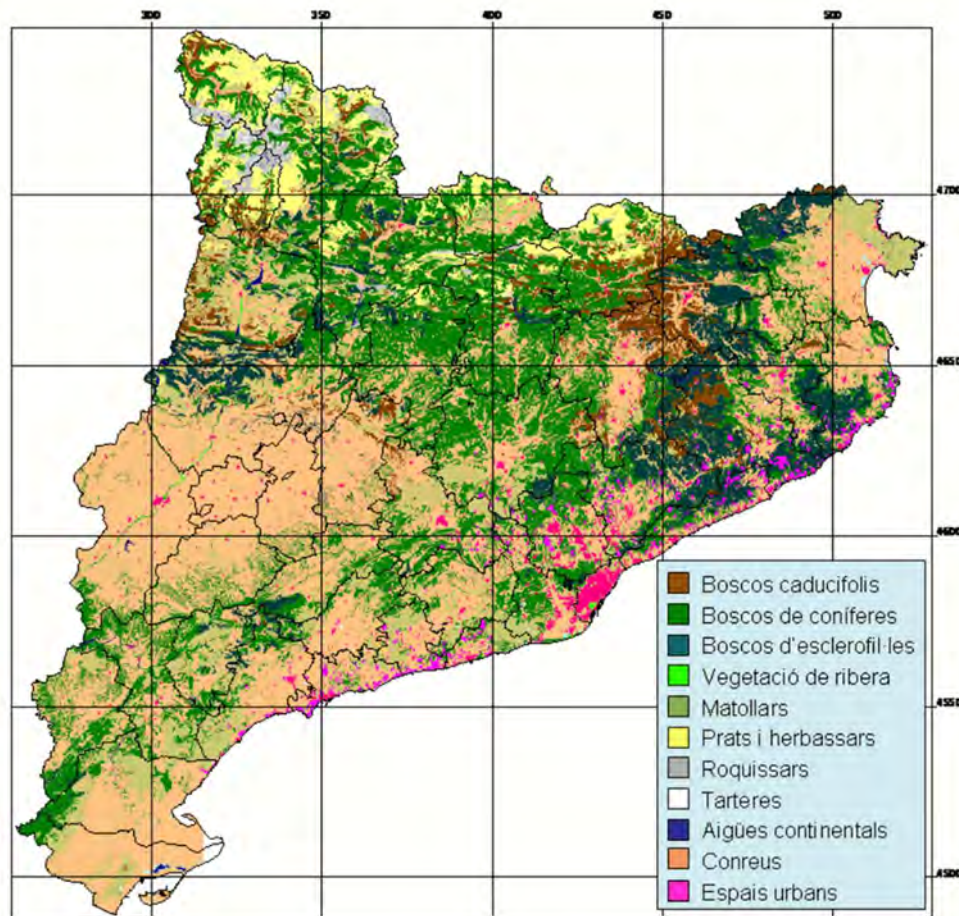


Font: Elaboració pròpia

Figura 4.-Mapa de les Regions Policials de Catalunya

Aquestes regions tenen una geografia diversa: la Regió Policial Metropolitana Barcelona té pràcticament tot el territori urbanitzat, mentre que la Regió Policial Pirineu Occidental és una zona muntanyosa, amb poques poblacions i disperses. En general, tant a la Figura 3

com a la Figura 5, s'observa l'heterogeneïtat geogràfica del territori català. La major concentració urbana es troba a l'entorn metropolità de Barcelona i al litoral on, a la vegada, es localitzen la majoria de robatoris amb força a domicili.



Font: CREAF - Centre de Recerca Ecològica i Aplicacions Forestals

Figura 5.-Mapa de distribució dels principals tipus de cobertes del sòl a Catalunya

Al litoral també és on es concentra l'habitatge turístic i de segona residència. A l'interior de Catalunya es distingeixen entorns diversos, amb abundants zones rurals i conreades però també ocupades per ciutats grans i mitjanes amb entorns industrials i zones residencials tipus urbanitzacions. Tot i que Barcelona és la ciutat catalana on més robatoris a domicili hi ha, aquest estudi ha focalitzat l'atenció en les vuit regions restants, que són les que, aparentment, tenen més tendència a observar onades de robatoris.

L'heterogeneïtat dels entorns residencials catalans, segurament mesurable amb indicadors urbanístics específics, es palesa fàcilment mitjançant l'observació i coneixement directe d'aquest entorn, especialment quan es compara amb el d'altres països. En molts d'aquests, com els anglosaxons, l'exemple paradigmàtic serien grans les extensions residencials de cases distribuïdes de forma regular i homogènia com els mostrats a la figura 6.



Font: Google Maps

Figura 6.-Imatge aèria d'una zona residencial de Long Beach (Los Angeles, EUA)



Font: Google Maps

Figura 7.-Imatge aèria d'una zona residencial de Barcelona (Catalunya)

Per contra, una visualització d'aquests entorns residencials a Catalunya, mostra grans diferències (Figura 7), sent l'heterogeneïtat la seva principal característica.

1.2.2 La Regió Policial Metropolitana Nord

La Regió Policial Metropolitana Nord (Figura 8) s'ha escollit per a fer un estudi més detallat dels models predictius dels robatoris amb força a Catalunya. Hi ha varis motius per fer-ho, un dels quals és que s'hi localitzen un 25% del total d'aquest tipus de robatoris, també perquè és representativa de l'heterogeneïtat d'entorn característica de Catalunya i, finalment, perquè al ser un territori més reduït permet interpretar millor alguns aspectes de la modelització predictiva.



Font: Elaboració pròpia a partir d'un mapa de l'Institut Cartogràfic i Geològic de Catalunya – ICGC

Figura 8.-Mapa de la Regió Policial Metropolitana Nord

Aquesta regió policial té una superfície de 1.864km² i 2,2 milions d'habitants, en un territori divers, amb zones situades a la conurbació de Barcelona, amb ciutats densament poblades com Badalona o Santa Coloma de Gramenet, grans zones industrials, eixos de comunicació importants com l'autopista AP7, zones litorals residencials i turístiques, abundants urbanitzacions, així com ciutats mitjanes com Terrassa, Sabadell, Granollers o Mataró, i zones rurals amb pobles i cases disseminades.

1.3 Les teories de l'oportunitat delictiva i la conducta criminal

En les últimes dècades, el corrent criminològic preponderant és el que justifica la conducta criminal com una actitud racional, fruit del mateix tipus de decisions i mecanismes psicològics que les persones utilitzen en qualsevol àmbit, no exclusivament relacionat amb la delinqüència. Des d'aquest punt de vista, les accions dels delinqüents poden ser previsibles si es coneixen aquests patrons d'actuació "objectius". El marc teòric general de les principals teories criminològiques que modelen la conducta criminal és la teoria de l'oportunitat delictiva (Gottfredson and Hirschi 1990), ben recollit i explicat en l'article "Opportunity Makes the Thief. Practical theory for crime prevention" (Felson and Clarke 1998). Els seus plantejaments s'articulen mitjançant la teoria de les Activitats Rutinàries (Cohen and Felson 1979), la de l'Elecció Racional (R. V. Clarke and Cornish 1986), la teoria de la Foraging Behaviour (Bernasco 2009), la teoria dels Crime Patterns (P. Brantingham and Brantingham 1995; P. J. Brantingham 1984) o de la concentració delictiva/Crime concentration (Farrell 2015).

D'aquestes teories més generals, se'n desprenen altres que investiguen aspectes més concrets de la conducta criminal, com pot ser l'estudi sobre els desplaçaments que fan els delinqüents per anar a delinquir, normalment des del seu domicili fins al lloc on cometen l'il·lícit. Per la seva rellevància amb la present tesi, també s'han consultat aquestes teories del que en anglès solen anomenar "Journey to crime" (Bernasco and Nieuwebeerta 2005; Michael Townsley and Sidebottom 2010; Nee 2015; Van Deiver, Van Daele, and Van Der Beken 2015; Wollinger et al. 2018).

Una altra derivada d'aquestes teories és la de la victimització repetida i quasi repetida, fonamental en l'estudi de la predicció delictiva.

1.4 La teoria de la victimització quasi repetida

Els conceptes de victimització repetida i quasi repetida provenen de la teoria criminològica que porta el mateix nom, coneguda també o inclosa dins el concepte més ampli de victimització reiterada. En el llibre "Criminología ambiental: ecología del delito y de la seguridad" (Vozmediano Sanz and San Juan Guillén 2010) els autors classifiquen els tipus de repeticions en la següent taula:

Tabla 2. Tipología bidimensional de la victimización repetida.			
	Objetivo	Localización	Infractor
Repetición en el espacio.	Misma vivienda, persona, vehículo, negocio, etc. <i>hot spots</i>	Repeticiones concentradas en el espacio: <i>hot spots</i> , instalaciones peligrosas.	Mismo infractor repitiendo el delito en el mismo lugar.
Repetición en el tiempo.	Repetición rápida del delito contra el mismo objetivo, por ejemplo violencia doméstica.	Una repetición rápida no siempre implica a los mismos infractores, por ejemplo, robos a un negocio por distintas personas.	Puede ser el mismo infractor que realiza repeticiones rápidas; ej. violencia doméstica.
Repetición del tipo de delito.	Mismo tipo de delito cometido contra objetivos similares.	Un área no muy extensa y definida; por ejemplo zonas de trapicheo de drogas.	Por ejemplo, hurtos repetidos en las tiendas de ropa por parte de adolescentes.
Repetición táctica.	Repetición virtual en la que se usa la misma táctica (<i>modus operandi</i>).	Misma estrategia puede ser útil en un mismo lugar, por ejemplo, carterismo en un mercado en la calle.	Infractor que emplea repetidamente la misma estrategia (carterista, robo de coches, etc.)

Fuente: Adaptada de Farrell & Pease (2008. p 122).

Font: Llibre "Criminología ambiental: ecología del delito y de la Seguridad", UOC

Taula 1.- Tipus de victimitzacions repetides

Pel que fa als robatoris amb força a domicili, diferents estudis indiquen que la repetició en l'espai i el temps, en domicilis semblants i amb les mateixes tàctiques, ofereix avantatges als lladres, augmentant els beneficis i la probabilitat d'èxit i disminuint els costos i els riscos.

Alguns dels articles de referència que indaguen sobre les causes de la victimització quasi repetida en el cas dels robatoris amb força a domicili són:

- Like taking candy: why does repeat victimization occur? (Farrell G., Pease, and Phillips 1995)

- New insights into the spatial and temporal distribution of repeat victimization. (S. D. Johnson, Bowers, and Hirschfield 1997)
- Infectious burglaries. A test of the near repeat hypothesis. (M. Townsley 2003)
- The stability of space-time clusters of burglary. Br J Criminol (Shane D Johnson and Bowers 2004)
- Mapping crime: understanding hot spots. (J E Eck et al. 2005)
- When does repeat burglary victimization occur? (Sagovsky and Johnson 2007)
- A statistical model of criminal behavior. (Short et al. 2008)

En aquesta tesi, es parteix del plantejament de la teoria de la victimització quasi repetida en el context dels robatoris amb força a domicili i centra l'estudi en els suposats clústers espai-temps que pronostica i que poden permetre predir els robatoris a llarg i, especialment, a curt termini. Aquests clústers s'analitzaran independentment de l'autoria, tot i que l'anterior teoria defensa que normalment seran comesos pels mateixos autors. Si aquests patrons s'observen, no són aleatoris, i tenen certa regularitat, es podran crear models que permetin predir-los i que esdevinguin una eina útil perquè la policia millori la prevenció.

En aquest sentit, s'ha estudiat en diferents països, especialment els últims 15 anys. Un article de referència és "Space-time patterns of risk: a cross national assessment of residential burglary victimization" (Shane D. Johnson et al. 2007), que va utilitzar dades dels robatoris amb força a domicili de localitats d'Austràlia, Països Baixos, Nova Zelanda, Regne Unit i Estats Units d'Amèrica.

En general, aquests estudis sobre la victimització repetida i quasi-repetida estableixen que els autors dels robatoris tornen a robar al domicili que ja han robat o als que estan pròxims, i que en un determinat lapse de temps i espai, que sol ser d'unes dues setmanes i un radi d'uns 250 metres, els mateixos autors hi cometen varis robatoris. Incorporar la variable temps als mapes delinqüencials permet fer-los dinàmics i observar clústers de punts en l'espai-temps. Així, detectada la localització d'un primer robatori, permet fer prevenció en el seu entorn durant el lapse de temps en què hi hagi risc de repetició.

Es distingeixen dues components del risc, l'estàtic i el dinàmic.

1.4.1 Risc estàtic

S'anomena risc estàtic (Tseloni and Pease 2003) el que depèn de factors d'entorn, com les mesures de seguretat de l'edifici, la facilitat d'accés, el tipus de botí que hi pot haver o el nombre de robatoris que hi ha hagut en el passat (varis anys). Aquest risc es pot establir com a fix.

1.4.2 Risc dinàmic

L'altre component del risc és el dinàmic (Kate J Bowers and Johnson 2004; Ornstein and Hammond 2017), que depèn dels robatoris que s'estan fent al domicili o als domicilis propers, els quals, en virtut de la teoria de la victimització repetida, n'augmenten el risc durant cert temps.

Els models de predicció s'articularen a partir de la idea de risc, és a dir, de la mesura de la probabilitat que robin en els domicilis.

1.5 Policing - La prevenció policial de la delinqüència

1.5.1 El model actual de la policia

Sent aquesta tesi d'un Programa de Doctorat en Dret, Economia i Empresa i en la línia de recerca sobre Mètodes Quantitatius en la Recerca per l'Economia i l'Empresa, sembla pertinent contextualitzar-la mínimament en aquest sentit. La Policia de la Generalitat – Mossos d'Esquadra és un servei públic i com a tal és gestionat i administrat pel Govern de Catalunya. La Policia està subjecta a les normes i a les lleis vigents, amb reglaments propis, i dirigida per càrrecs i institucions públiques. Les diferents disposicions legals que afecten la policia n'expressen un determinat model de gestió i de compromís amb la societat que serveix. El model de policia a Catalunya està en consonància, a grans trets, amb els models de policia d'altres països Europeus, i ha sofert una evolució o transformació paral·lela els últims anys (Lasierra 2006). A grans trets, el model del que es podrien anomenar “policies modernes” es caracteritza per una cessió de part de les seves funcions tradicionals a altres serveis públics o privats i als propis ciutadans. Aquesta cessió és en forma de col·laboració, major relació i intercanvi d'informació o posada en comú de tot allò que s'emmarca en el concepte ampli de “seguretat” o “inseguretat” ciutadana. Aquest enfocament més obert està en consonància amb la idea de resoldre els problemes de seguretat, de tractar-los com a objectes d'anàlisi social però també científic. Tant en l'anàlisi com en la solució d'aquests problemes, hi intervindran diferents actors, policials i no policials, de manera que es generarà un debat d'idees sobre el què, i el com resoldre'ls.

Aquest debat d'idees és el que millor representa la situació actual del model de policia. Actualment, els Mossos d'Esquadra participen en projectes europeus que discuteixen sobre les millors pràctiques o les més aconsellables per aconseguir el que anomenen Model de Seguretat Europeu. Un d'aquests, dins del programa de la Comissió Europea HORIZON 2020, és el projecte “Cutting Crime Impact – CCI (“CCI Project - We Aim To Cut Crime Impact” n.d.). Aquest projecte es divideix en quatre grans àmbits:

- *Predictive Policing*
- *Community Policing*
- *Crime Prevention through urban design and planning (CP-UDP)*
- *Measuring and mitigating citizens' feelings of insecurity*

Aquests àmbits il·lustren l'orientació actual dels problemes de seguretat, i mostren la rellevància que es dona a la predicció delictiva.

1.5.2 Predictive Policing

Les policies modernes tenen en la prevenció la seva principal vocació, referint-s'hi a vegades amb el terme *proactive policing* (Weisburd et al. 2019). La icònica imatge de la policia encalçant un lladre que fuig amb el botí, típica del model de policia reactiva, ha deixat pas a la policia proactiva, que analitza i vol solucionar els problemes, preventivament, per mirar d'evitar-ne la repetició. Anticipar-se, entendre els factors i

mecanismes que generen la delinqüència, i com es relacionen en l'espai i el temps, han passat a ser la prioritat. Saber el què, el per què, l'on i el quan. Una anàlisi exhaustiva de la criminalitat, de les seves causes, interaccions, detonants, dinàmiques i distribucions espaciotemporals.

D'aquest coneixement més profund dels fenòmens criminals i del convenciment que n'és la clau per gestionar la seguretat pública de la policia del futur, en va sorgir el concepte *intelligence led policing* (J. H. Ratcliffe 2014; J Ratcliffe 2008), la policia dirigida per la intel·ligència. En la mateixa línia, altres plantejaments estratègics globals com el *Problem-Oriented Policing (POP)* (Policia orientada a la solució dels problemes) ("Center for Problem-Oriented Policing | ASU Center for Problem-Oriented Policing" n.d.; Goldstein 1990; Michael Townsley et al. 2003; Borrion et al. 2020; Sidebottom et al. 2020) i el *Comunity / Proximity Policing* (Policia comunitària o de proximitat) (Fielding 2005; Tilley 2012; John E. Eck and Rosenbaum 2014), a la vegada relacionats amb l' *enviromental criminology* (criminologia ambiental) (Vozmediano Sanz and San Juan Guillén 2010) i la *situational crime prevention* (prevenció situacional del delicte) (Soto Urpina 2016). Més recentment, s'ha afegit a la llista el *Predictive Policing* (Policia predictiva) (Perry et al. 2018), relacionat a la vegada amb el *hot spot policing* (prevenció a les zones de més risc) o amb el concepte més general de *placed-based policing* (prevenció basada en el lloc).

Amb major o menor mesura, aquestes estratègies entenen la criminalitat com una manifestació objectiva d'unes problemàtiques relacionades amb l'entorn geogràfic, humà, econòmic i social, de manera que hi ha la confiança que coneixent i intervenint en aquests àmbits es pugui reduir la delinqüència. I, de fet, aconseguir un objectiu més ampli, que és que la gent se senti segura, millorar el sentiment de seguretat i benestar.

Augmentar la visibilitat de les patrulles en un lloc concret, interaccionar amb els habitants per entendre les problemàtiques locals, analitzar les vulnerabilitats de seguretat dels espais públics, reunir-se amb les entitats d'un barri, etc. ja formen part de les estratègies preventives habituals per millorar la seguretat d'una zona. La necessària col·laboració dels ciutadans, dels comerços, les empreses, les entitats socials o serveis públics diversos, en la gestió de la seguretat pública, converteix la policia en un actor principal, però no únic.

En aquest marc de treball ampli i obert, les policies no poden ser institucions tancades en si mateixes, sinó obertes a les demandes de seguretat i al coneixement en general, especialment a les noves tecnologies i eines d'anàlisi. A la vegada, l'era del Big Data, el Data Science i el Machine Learning fan imprescindible la col·laboració amb les universitats, els centres de recerca i les empreses del sector de les TIC.

Tot i així, aquestes col·laboracions tenen pros i contres. El *màrqueting* empresarial, els diferents interessos de les parts implicades o, simplement, la manca d'un llenguatge comú, pot comportar frustracions i fracassos inesperats. Problemàtiques que també poden estar relacionades amb l'aplicació prematura de mètodes científics en àmbits encara poc preparats per fer-ho. Fins i tot la ciència més propera a la policia, la criminologia, té dificultats per accedir i influenciar el dia a dia dels cossos de seguretat. En aquesta línia, també és una dificultat els canvis d'orientació estratègica de la policia que es poden derivar

del *predictive policing*, que a la vegada impliquen canvis operatius, tàctics i organitzatius en unes estructures ja funcionals en la prevenció de la delinqüència, que es poden sentir qüestionades. Són coneguts alguns antecedents en altres països en què constaten que les policies reben el *predictive policing* més com una imposició tecnològica innecessària que com un ajut a les seves tasques i funcions preventives.

Un altre aspecte controvertit i que pot ser important per valorar l'èxit d'estratègies com el *predictive policing* és la inexistència d'una mesura objectiva de la seguretat, imprescindible per valorar-ne la seva gestió. Aparentment, l'estadística policial hauria de complir aquesta funció, però és sabut que les dades policials no sempre són representatives d'allò que aparentment recullen, no són fruit del disseny d'una mostra estadística, sinó simplement el que ha quedat registrat a les bases de dades, el que s'ha denunciat o que, d'alguna manera, la policia n'ha tingut coneixement. Se sap que en alguns tipus delictius existeix una important xifra negra, les persones no denuncien o avisen a la policia de tot allò que els passa relacionat amb la inseguretat. Per altra banda, existeixen enquestes de victimització, que compleixen requisits estadístics de representativitat i fiabilitat, que poden ser dissenyades ad hoc per fer aflorar la xifra negra o altra informació no continguda en les bases de dades policials. Aquestes enquestes solen tenir per objectiu el coneixement estratègic, de fenomen, i la seva realització sol estar contemplada en els plans estratègics de seguretat per mesurar el grau de seguretat del país en diferents àmbits. Actualment, l'Enquesta de Seguretat Pública de Catalunya té una periodicitat bianual, i es complementa amb altres enquestes específiques com l'Enquesta de Violència Masclista.

1.5.3 La prevenció operativa: patrulles uniformades

L'actual model de desplegament dels Mossos d'Esquadra a Catalunya es basa en la distribució dels recursos humans i materials arreu del territori. Existeixen 59 comissaries que són la seu de les Àrees Bàsiques Policials (ABP), unes divisions administratives, normalment coincidents amb els límits comarcals o municipals, que són el nucli de la gestió policial de la seguretat ciutadana. El territori de cada ABP sol dividir-se en diferents sectors de patrullatge on, habitualment, sempre hi haurà una patrulla uniformada per atendre els requeriments de seguretat. Aquestes patrulles uniformades solen estar formades per dos agents de policia uniformats que circulen amb un vehicle logotipat. A banda dels recursos propis de cada comissaria, esporàdicament, els territoris són reforçats amb agents dels serveis regionals o centrals. La planificació dels serveis sol recaure en el cap de seguretat ciutadana del territori i en el cap de torn, i focalitzen l'atenció en les problemàtiques locals i diàries. En general, es distingeix entre els serveis de reacció i els serveis proactius. Els de reacció fan referència a l'atenció de tots els incidents de seguretat que puguin sobrevenir, mentre que els serveis proactius o preventius són els dirigits a evitar la comissió de nous delictes o altres problemàtiques de seguretat.

Per a una correcta gestió dels recursos i de la prevenció, la policia disposa de serveis d'anàlisi policial. Aquests serveis no solen estar ubicats en les ABP, sinó en les Regions Policials o en els Serveis Centrals. Fan un diagnòstic de la situació de la seguretat i detecten amenaces i riscos. Aquesta informació serà traspasada als responsables de la seguretat de cada territori.

1.5.4 La prevenció actual dels robatoris a domicili

Les accions preventives actuals per combatre els robatoris amb força a domicili solen consistir en la presència de patrulles uniformades en els llocs recentment victimitzats o que històricament ho han estat més. Aquesta presència es pot fer circulant lentament amb el vehicle policial per la zona o fent-hi patrulles a peu. Això es sol combinar amb controls policials a la mateixa zona o a les vies d'accés, i amb la identificació de persones que la policia considera que poden ser autores de robatoris. Per altra banda, des d'una vessant més estratègica, es solen donar consells de seguretat especialment als col·lectius que es poden considerar de més risc. En el cas dels robatoris amb força a domicili, els consells són referents a tot allò que poden fer els propietaris o habitants d'un domicili per minimitzar el risc que els lladres hi entrin. Tant respecte a la millora de la seguretat física de l'edifici (accessos, panys, càmeres de vigilància, etc.) com respecte a la conductes preventives de caire personal (en cas de sortir de vacances, en la relació amb els veïns, etc.).

Hi ha pocs estudis sobre l'efecte que tenen aquestes accions preventives de la policia per prevenir els robatoris, llevat del seguiment de l'estadística policial. Des d'aquesta perspectiva, els últims 10 anys no es poden valorar favorablement. Més enllà de l'estadística policial, el més preocupant són els brots d'inseguretat que poden generar les conegudes "onades" de robatoris. La concentració d'aquests, tant en l'espai com en el temps, augmenten la possibilitat d'alarma social i, per tant, evitar les onades de robatoris hauria de ser un objectiu prioritari de la prevenció.

1 HIPÒTESIS

Hipòtesi principal: Els robatoris amb força a domicili a Catalunya segueixen patrons repetitius en l'espai i el temps.

Hipòtesis secundàries:

1. L'acotació de l'entorn territorial on es produeix la repetició ve determinada per les característiques geogràfiques del lloc i del domicili robat, seguint la norma de victimitzar, preferiblement, llocs semblants als que ja s'han robat (victimització quasi-repetida).
2. L'interval temporal en què es produeixen les repeticions sol ser inferior a les tres setmanes, i està condicionat pel factor territorial, per l'èxit dels robatoris executats i per la pressió o presència policial.
3. La presència policial, més o menys freqüent i més o menys propera als llocs on els lladres tenen intenció de robar, els pot dissuadir de fer-ho.

2 OBJECTIUS

L'objectiu principal d'aquesta tesi és estudiar els patrons de repetició dels robatoris amb força a domicili a Catalunya per millorar-ne les estratègies de prevenció policial. Per aconseguir-ho, s'estableixen tres objectius específics:

Objectius secundaris:

1. Detectar les zones amb més probabilitat de repetició en l'espai i el temps i acotar-les territorialment (zona potencial de repetició) i temporalment (interval temporal de la repetició).
2. Classificar aquestes zones segons el nombre i la probabilitat d'aquestes repeticions i segons l'acotació en l'espai i el temps, de manera que es puguin prioritzar els llocs on la prevenció policial pugui ser més efectiva.
3. Construir un model de prevenció policial que optimitzi el coneixement d'aquests patrons repetitius amb la gestió dels recursos policials.

3 METODOLOGIA

Per analitzar les dades dels robatoris amb força a domicili a Catalunya s'ha fet una primera anàlisi descriptiva de la distribució en l'espai i el temps en funció de les divisions administratives (municipis, Àrees Bàsiques Policials i Regions Policials). S'ha estudiat la viabilitat del test de Knox en diferents punts del territori per comprovar el patró de victimització quasi repetida a petita escala. S'ha fet una divisió de l'espai en cel·les de 5km de costat per analitzar les distribucions de les sèries setmanals de robatoris. Concretament, se n'ha estudiat l'ajustament a la distribució de Poisson i l'aleatorietat. Amb els resultats obtinguts, s'ha proposat un algoritme predictiu basat en la victimització quasi repetida a gran escala. Per estudiar les possibilitats predictives a més petita escala, s'ha creat un model log-Cox Gaussià que s'ha aproximat amb la metodologia Integrated Nested Laplace Approximation (INLA). Finalment, s'han analitzat les correlacions de les sèries de robatoris en diferents zones mitjançant l'Anàlisi de Components Principals i s'han proposat nous patrons en l'espai-temps per descriure la distribució dels robatoris i millorar les prediccions.

3.1 Background metodològic

Els robatoris amb força a domicili es distribueixen en l'espai condicionats per diferents factors, el més evident dels quals és la presència d'un domicili. La distribució dels domicilis i de les zones residencials, ja siguin urbanitzacions, disseminats, pobles o ciutats, és de forma espacialment aglomerada, amb diferents graus de concentració. El tipus de domicili, que pot ser una casa de quatre-vents, una casa adossada, o un pis/apartament, també condiciona l'ocupació de l'espai, per l'extensió de terreny a l'entorn de les cases de quatre vents, o del nombre de plantes d'un edifici de pisos.

Tot i les diferents casuístiques, la forma més senzilla d'aproximar la ubicació d'un domicili és mitjançant les seves coordenades geogràfiques, projectades en coordenades cartesianes en el pla $s = (x, y)$. En el cas dels robatoris a domicili, s'ha utilitzat el Sistema de Coordenades Universal de Mercator (*Universal Transverse Mercator*, UTM). D'aquesta manera, s'entén la distribució espacial dels robatoris com un procés puntual.

3.1.1 Procés puntual

Formalment, un Procés Puntual Espacial (Spatial Point Process, SPP) és un procés estocàstic, successió d'observacions d'origen aleatori, que genera localitzacions d'interès dins d'una regió concreta d'estudi (Cabrero Ortega and Garcia Perez 2015).

3.1.2 Aleatorietat espacial completa

El primer objectiu en l'anàlisi de la distribució de les localitzacions en l'espai és comprovar si es distribueixen a l'atzar, o si, contràriament, es distribueixen de forma regular, uniformement espaiats, o es distribueixen de forma agrupada, en clústers.

Quan es distribueixen a l'atzar, es diu que s'observa *Aleatorietat Espacial Completa* (*Complete Spatial Randomness, CSR*), i es formalitza matemàticament com un *Procés de Poisson Homogeni*, que es caracteritza per les següent propietats:

- a) El nombre de localitzacions d'una regió A, d'àrea $|A|$ segueix una distribució de Poisson de mitjana $\lambda|A|$, on λ és la intensitat del procés, és a dir, el nombre esperat de localitzacions per unitat d'àrea.
- b) Donades n localitzacions en una regió A, és a dir, condicionalment a que hi hagi n localitzacions en A, aquestes es distribueixen segons una distribució uniforme sobre A.
- c) En dues regions disjunctes A i B, el nombre de localitzacions en A i el nombre de localitzacions en B són variables aleatòries independents.

Comprovar que un procés puntual espacial segueix, o no, l'Aleatorietat Espacial Completa, es pot fer mitjançant un test χ^2 de bon ajustament a la distribució de Poisson o mitjançant mètodes basats en l'anàlisi de distàncies.

3.1.3 *Processos de Poisson no-homogenis*

Quan el procés puntual no segueix un model de distribució espacial aleatòria, sinó que s'observen agrupaments o clústers de les localitzacions, pot ser que es pugui modelar segons un *Procés de Poisson no homogeni*, en què la intensitat $\lambda(s)$ del procés no és constant sinó que depèn de la localització $s \in A$. Aquest tipus de processos són els que compleixen les següents propietats:

- a) El nombre de localitzacions d'una regió A, $N(A)$, segueix una distribució de Poisson amb mitjana $\int_A \lambda(x)dx$, on $\lambda(x)$ és una funció no-negativa.
- b) Donades n localitzacions $\{x_i\}_{i=1}^n$ en una regió A, formen una mostra aleatòria d'una distribució en A amb funció de distribució de probabilitat proporcional a $\lambda(x)$.
- c) En dues regions disjunctes A i B, les variables aleatòries $N(A)$ i $N(B)$ són independents.

Entre aquest tipus de processos hi ha els *Processos de Cox* i els *Processos de Poisson amb Clústers*.

En els *Processos de Cox*, també anomenats *Processos de Poisson doblement estocàstics*, la intensitat $\lambda(x)$ és aleatòria, generalment amb distribució Gamma. Un cas específic és el *Procés de Cox Log-Gaussià*, en què $\lambda(s) = e^{z(s)}$, amb $\ln(z(n)) \sim N(\mu, \sigma^2)$.

Els *Processos de Poisson Clúster* són models que incorporen explícitament la formació de grups de localitzacions, seguint les següents condicions:

1. Hi ha un procés amb "pares" que segueix un procés de Poisson amb intensitat λ .
2. Cada "pare" produeix un nombre aleatori de "descendents" M d'acord amb una determinada distribució.
3. Les posicions dels descendents respecte els pares són aleatòries i segueixen una distribució bivariada.
4. El procés Poisson-Clúster es correspon amb la posició dels descendents.

En el cas que la distribució bivariada sigui normal, s'anomena *Procés de Thomas*.

3.1.4 Estimació de la intensitat espacial

En el cas d'un procés de Poisson homogeni la intensitat és constant en cada àrea considerada A , de manera que si conté n localitzacions, un estimador serà $\hat{\lambda} = \frac{n}{|A|}$, on $|A|$ representa l'àrea de la regió A .

En el cas de processos de Poisson no homogenis, la intensitat es defineix com,

$$\lambda(s) = \lim_{ds \rightarrow 0} \frac{E[Y(ds)]}{ds}$$

On ds és una regió reduïda a l'entorn de la localització s , i $Y(ds)$ és el nombre de localitzacions en ds . Existeixen diferents possibilitats per estimar-la, distingint entre mètodes paramètrics i no paramètrics. En el cas dels paramètrics, es proposa una funció d'intensitat que depèn d'uns paràmetres, estimables mitjançant la màxima versemblança. Aquesta sistemàtica permet incloure-hi covariables que dependran de la localització.

L'estimació de la intensitat per mètodes no paramètrics es basen en l'Estimador Nucli Suavitzat (*Kernel smoothing*), donat per:

$$\hat{\lambda}(s) = \frac{1}{q(\|s\|)h^2} \sum_{i=1}^n K\left(\frac{\|s - s_i\|}{h}\right)$$

on es suposa que s'han observat n localitzacions s_1, \dots, s_n , on K és la funció nucli considerada (normalment bivariada), $\|s\|$ és la longitud del vector s , $q(s)$ una correcció frontera per compensar els valors que es perden quan s està prop de la frontera de la regió A , i on h és una mesura del nivell de suavitzat (*smoothing*), també anomenada *ample de banda* (*bandwidth*), que es vol considerar.

3.1.5 Processos Puntuals Espaciotemporals

Els Processos Puntuals Espaciotemporals (*Spatiotemporal Point Process, STPP*) es poden definir com a processos puntuals espacials en què se'ls afegeix la variable temps. Formalment (Gabriel, Rowlingson, and Diggle 2013), són una successió aleatòria de punts, $P = \{(s_i, t_i): i = 1, 2, \dots\}$, en què $s_i \in \mathbb{R}^2$ és la localització i $t_i \in T \subset \mathbb{R}^+$ és el moment en què ha ocorregut l'esdeveniment i -èssim. A la pràctica, el conjunt de dades són una mostra de punts $(x_i, t_i): i = 1 \dots n$ fruit d'una realització del procés estocàstic en un domini espaciotemporal restringit, $S \times T$, on S sol ser un polígon i T un interval tancat.

Des de la perspectiva contínua del temps, l'anàlisi dels processos espaciotemporals es pot entendre com una extensió directe del cas espacial. Si es considera $Y(A)$ el nombre d'esdeveniments en una regió arbitrària A , la intensitat del procés seria:

$$\lambda(s, t) = \lim_{|ds| \rightarrow 0, |dt| \rightarrow 0} \frac{E[Y(ds, dt)]}{|ds||dt|}$$

On ds és una regió reduïda a l'entorn de la localització s , dt és un interval petit que conté t , i $Y(ds, dt)$ és el nombre d'esdeveniments en $ds \times dt$.

Des d'una perspectiva discreta del temps, aquest es pot analitzar com una covariable, dins el que s'anomenen Processos Puntuals amb Marques (*Marked Point Process*). També es

poden analitzar com una transició de processos puntual espacials. Des d'aquest últim plantejament (Diggle 2005), si es considera \mathcal{P}_t com el procés puntual espacial en el moment temporal t , l'objectiu d'estudi seria la distribució conjunta de $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_t\}$, que té la forma,

$$[\mathcal{P}] = [\mathcal{P}_1][\mathcal{P}_2|\mathcal{P}_1] \dots [\mathcal{P}_t|\mathcal{P}_{t-1}, \dots, \mathcal{P}_1]$$

Per tant, caldrà descobrir com influeixen les realitzacions passades del procés espacial en les futures, quina és la seva dependència espaciotemporal.

En aquest sentit, es podran plantejar diferents hipòtesis d'estudi, com el compliment de la propietat de Màrkov, en què la probabilitat de transició del procés només depèn de l'estat (o moment) anterior $P(X_{t+1}|X_1, X_2, \dots, X_t) = P(X_{t+1}|X_t)$.

3.1.6 Separabilitat espaciotemporal

Un procés puntual espaciotemporal es diu que és *separable de primer ordre* (Gabriel, Rowlingson, and Diggle 2013) si la seva intensitat $\lambda(s, t)$ es pot factoritzar de la següent manera,

$$\lambda(s, t) = m(s)\mu(t), \quad \forall (s, t) \in S \times T$$

És a dir, les components espacial i temporal són independents.

3.1.7 Autocorrelació temporal i processos autoregressius (AR)

L'autocorrelació temporal és la correlació d'una sèrie temporal discreta amb una versió desplaçada d'ella mateixa.

Donat un procés estocàstic $\{X_t\}_{t \in T}$, s'anomena autocorrelació d'ordre k , ρ_k , a la correlació de qualsevol de les variables aleatòries del procés, distanciades en k instants o retards,

$$\rho_k = \frac{Cov(X_t, X_{t+k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+k})}}, \quad \forall k, t \in \mathbb{Z}$$

A la vegada, un procés autoregressiu d'ordre p , que es denota $AR(p)$, és un procés en el que la variable X_t s'obté mitjançant una regressió sobre els valors passats de la mateixa,

$$X_t = a_t + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p}$$

3.1.8 Test de Knox: estudi de la dependència espaciotemporal

La manera habitual de comprovar el patró de victimització repetida i quasi repetida és mitjançant el test de Knox (E. G. Knox and Bartlett 1964; G. Knox 1963), fàcilment aplicable mitjançant software gratuït ("R: The R Project for Statistical Computing" n.d.; "CRAN - Package Rcrimeanalysis" n.d.; "CRAN - Package Surveillance" n.d.; "CrimeStat: Spatial Statistics Program for the Analysis of Crime Incident Locations | National Institute of Justice" n.d.).

Donat un conjunt d' N robatoris, el test necessita les seves coordenades espai temps (x, y, t) i els intervals espacials (en metres) i temporals (en dies) en els quals es vol comprovar l'existència del patró de repetició (també anomenats valors crítics d'espai i temps) i que representen la distància màxima acceptada entre dos casos per considerar-se pròxims (la

distància espacial pot ser l'euclídea o de Manhattan, i la distància temporal és la diferència en valor absolut). A partir d'aquests valors, es classifiquen els $n = N(N - 1)/2$ parells de robatoris quatre grups: $A = \{\text{parells pròxim en espai i temps}\}$, $B = \{\text{parells pròxims només en l'espai}\}$, $C = \{\text{parells pròxims en el temps}\}$ i $D = \{\text{parells que no són pròxims ni en l'espai ni en el temps}\}$. S'estableix la hipòtesi nul·la que no existeix agregació espai-temps entre els parells de casos, i per contrastar-la es compara amb el nombre d'elements de A respecte el nombre esperat si la hipòtesi nul·la fos certa.

Si el $\text{Card}(A+B)$ i $\text{Card}(A+C)$ són petits respecte n (condició 1), llavors $\text{Card}(A)$ segueix una distribució de Poisson amb mitjana $\mu = \frac{\text{Card}(A+C)\text{Card}(A+B)}{n}$ (David and Barton 1966), i amb aquesta hipòtesi es calcula el p-valor del test: $p\text{-valor} = P\{\text{Poisson}(\mu) \geq \text{Card}(A)\}$. En cas que no es compleixi la condició 1, el p-valor s'obté mitjançant simulacions amb el mètode de Montecarlo.

Aquest mètode ha tingut diverses revisions, sent una de les més recents una adaptació del text de Knox per entorn heterogenis (Briz-Redón, Martínez-Ruiz, and Montes 2020).

3.1.9 *Model de Short*

El pas definitiu cap a la modelització matemàtica es va fer l'any 2008, amb el Model de Short, publicat en l'article "Statistical Model of Criminal Behaviour" (Short et al. 2008).

El Model de Short, formulat en equacions en diferències pel cas discret i en equacions diferencials pel cas continu, és un model de simulació que permet visualitzar la generació de hotspots en l'espai i el temps en funció de certs paràmetres. La seva estructuració i formulació va tenir una gran influència en els models predictius posteriors. Parteix de la configuració a microescala, considerant cases i lladres individuals. Les cases les representa en els nodes d'una malla regular, on els lladres s'hi mouen seguint una random walk. El risc que hi hagi un robatori en una casa depèn de l'atractivitat d'aquesta i de la presència d'un lladre. La dinàmica dels robatoris ve determinada per la component aleatòria del random walk però també pel mecanisme que reinicialitza l'atractivitat de la casa (boosts) quan s'ha robat en alguna del seu voltant. El model s'entén com un procés de Poisson, en què la intensitat de la probabilitat de robar en una casa, dita atractivitat (A), depèn de l'espai (s) i el temps (t), i aquesta dependència és la suma d'una component estàtica A_s , basada en la hipòtesi flag, i una component dinàmica B_s^t , basada en la hipòtesi boost, segons la relació $A_s^t = A_s + B_s^t$. Aquest plantejament, en els seus aspectes generals, s'ha mantingut invariable en tots els models predictius posteriors. Tanmateix, s'ha descartat l'opció de simulació amb equacions diferencials.

3.1.10 *Self-exciting Point Process*

Del mateix grup d'autors, vinculats a la Universitat de Califòrnia – Los Angeles (UCLA), en va sorgir el primer model pròpiament predictiu, formulat pel matemàtic George Mohler. En l'article "Self-exciting point process modeling of crime" (Mohler et al. 2011), s'estableixen els fonaments del programa de *predictive policing* més conegut, el PREDPOL ("PredPol – The Predictive Policing Company" n.d.).

La seva formulació s'inspirava en models matemàtics per predir les rèpliques dels terratrèmols, que també segueixen patrons espai-temps de proximitat. Concretament, es considera un procés puntual amb intensitat $\lambda(t, x, y)$ que és la suma de dues components:

$$\lambda(t, x, y) = \mu(t, x, y) + \sum_{\{k: t_k < t\}} g(t - t_k, x - x_k, y - y_k).$$

On $\mu(t, x, y)$ és la intensitat mitjana, relacionada amb l'històric de dades (background), la component estàtica o risc a llarg termini i amb la hipòtesi flag, i on la funció $g(t, x, y)$ és l'increment de la intensitat (o risc) provocat per robatoris en els punt propers en l'espai i el temps, relacionada amb la dinàmica a curt termini del risc i la hipòtesis boost.

A la pràctica, assumeixen que $\mu(t, x, y) = v(t)\mu(x, y)$, és a dir, que les fluctuacions temporals de la intensitat mitjana són globals, comunes a tot el territori, i a causa d'aspectes com la meteorologia, l'estacionalitat, el dia de la setmana, etc. Pel que fa a la funció $g(t, x, y)$, s'assumeix que decreix concèntricament amb la distància, i es considera isotròpica, és a dir, el seu valor només depèn de la distància h entre els punts (x, y) i (x_k, y_k) .

En el mateix article, expliquen la semblança que té aquest mètode predictiu amb el que fins aleshores s'utilitzava per descobrir els punts de concentració delictiva, els *hotspot maps*. En aquests últims, l'estimació de la intensitat o risc de robatoris es sol fer mitjançant funcions de suavització espacial amb les mateixes propietats que l'anterior funció g aplicades a tots els robatoris durant un cert període temporal, seguint el mètode Kernel Density Estimation (KDE). En ambdós plantejaments, les propietats suposades de la funció de suavització g impliquen una concepció homogènia de l'espai, de manera que la irradiació del risc sempre es fa de forma concèntrica i a la mateixa distància del punt on han robat. La principal diferència és que el model predictiu proposat incorpora una "memòria històrica" del risc, mitjançant la intensitat mitjana (background), i demostren que això li dona un millor encert predictiu.

3.2 Configuració espai-temps

En l'aplicació del test de Knox, clau per a determinar si es compleix el patró de victimització repetida en l'espai-temps, la finestra o lapse temporal i l'ample de banda de l'espai considerat per a cercar repeticions influeix decisivament en el resultat. En una mateixa zona, una configuració espai temps concreta determinarà la presència del patró repetitiu mentre que una altra configuració diferent no.

En general, la configuració dels programes de predicció delictiva basats en la victimització quasi repetida treballen amb cel·les d'entre 250 a 500 metres de costat, i intervals temporals no superiors a una setmana.

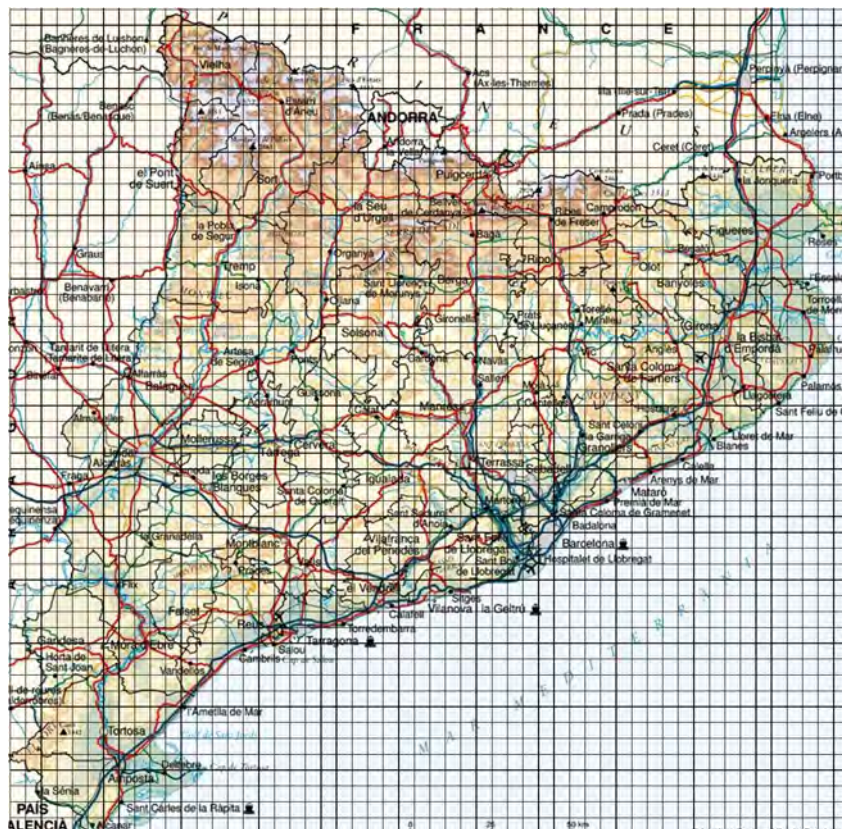
Com a criteris de referència s'han de tenir presents aspectes relacionats amb la conducta criminal, la concentració delictiva i les possibilitats preventives que generaria la detecció d'un possible patró. Criminològicament, s'entén per repeticions o rèpliques els robatoris que provenen d'una mateixa autoria, però, a priori, no es pot descartar un patró estadístic

de concentració en l'espai i el temps que tingués altres explicacions. Des del punt de vista preventiu, interessen les altes concentracions delictives el màxim d'acotades en l'espai i el temps.

Per altra banda, també s'ha de tenir present la fiabilitat i exactitud de les coordenades (x,y) i el temps (t) de cada robatori. En general, el moment del robatori es pot acotar en un interval d'un dia, però depenent de l'ús que es faci d'un domicili i d'altres circumstàncies, aquest interval pot arribar a ser de varis dies o setmanes. Pel que fa a les coordenades, solen tenir una precisió en un interval de 100m, però hi ha casos en què les adreces no estan correctament validades o no tenen tota la informació, com pot ser l'absència del nom del carrer o del número de casa.

Tenint presents les circumstàncies anteriors i les diferents casuístiques, s'han assajat diferents configuracions per determinar-ne una d'òptima per a l'estudi dels robatoris amb força a domicili a Catalunya. Finalment, s'ha escollit una segmentació de l'espai cel·les de 5km de costat i una segmentació temporal consistent en les setmanes naturals (Figura 9).

Es tracta d'una configuració adequada per estudiar els patrons espai temps en un territori extens com és la totalitat de Catalunya. El coneixement policial indica que els autors dels robatoris tenen tendència a desplaçar-se per tot el territori per cercar els entorns que més els afavoreixen per delinquir, de manera que aquesta configuració pot ajudar a detectar patrons de victimització en la proximitat, dins una mateixa cel·la, però també estudiar l'evolució dels principals punts de concentració dels robatoris.



Font: elaboració pròpia dels autors a partir d'un mapa de l'Institut Cartogràfic i Geològic de Catalunya – ICGC

Figura 9.-Mapa de Catalunya amb una malla de cel·les de 5kmx5km sobreposada

Tanmateix, aquesta configuració no és apte per a l'estudi en entorns urbans, densament poblats, com pot ser la conurbació de Barcelona. Per complementar l'estudi, també s'han considerats configuracions espacials més petites, concretament cel·les de 250m, 500m i 1000m de costat.

Pel que fa a l'elecció de les setmanes naturals com a interval temporal, es justifica amb el fet que són els períodes temporals més fàcilment comparables. Si s'escullen intervals diaris, els robatoris hi tenen un comportament diferenciat. És diferent un dia festiu que un de laboral, i diferent un dilluns que un divendres o un dimecres. Per altra banda, és més fiable la dada que informa de la setmana del robatori que la que informa del dia concret.

En general, també s'ha tingut en compte la intensitat dels robatoris en cadascuna de les configuracions possibles. Fora d'entorns estrictament urbans, la intensitat dels robatoris pot ser molt baixa si es redueixen molt els intervals espai-temps, cosa que en complicaria tant la modelització com les possibles estratègies preventives.

3.3 Mapes delinqüencials

Una tècnica d'anàlisi policial habitual són els mapes delinqüencials (John E. Eck et al. 2005; A. S. Chainey 2012), que proporcionen una informació molt valuosa per interpretar els fenòmens criminals. L'explicació del risc estàtic sol tenir relació amb el seu entorn, tant el més immediat (tipus de cases i barri), com el més genèric (lloc turístic, estar al centre o a les afores d'una població, estar pròxim a una via ràpida, etc). Tot aquest coneixement, que es pot arribar a recollir mitjançant capes d'informació, no se sol incorporar als registres policials de forma estructurada i automàtica però sí que es pot visualitzar sobre el mapa.

En aquest estudi, s'han representat les dades delinqüencials sobre mapes de diferent escala com a ajuda a la interpretació del fenomen i dels resultats obtinguts. A escala 1:100 s'hi ha representat el mapa de Catalunya, la graella de cel·les, el mapa de hotcells i altres classificacions de les cel·les. A escala 1:50.000 s'hi ha representat les cel·les individualment. A aquesta escala més petita es pot distingir el tipus de zona on s'ubica l'habitatge victimitzat, si es tracta d'una zona urbana o un disseminat, si són cases o edificis de pisos, si es troba en una urbanització o al centre de la ciutat, si la zona és turística o rural, si està prop de vies de comunicació ràpides, etc. Aquests mapes s'han extret de l'Institut Cartogràfic de Catalunya ("Inici. Institut Cartogràfic i Geològic de Catalunya" n.d.) i s'hi ha sobreposat la graella de cel·les.

També s'han utilitzat les visualitzacions amb els mapes de GoogleMaps des de diferents paquets del programari lliure R ("R: The R Project for Statistical Computing" n.d.), entre els quals els paquets "ggmap" i "ggplot2" (i les seves dependències). De forma més específica, per estudiar la victimització quasi repetida dels robatoris a domicili, s'han utilitzat les representacions gràfiques i dinàmiques, també en forma de mapes, del paquet "Stpp" (Gabriel, Rowlingson, and Diggle 2007).

3.4 Distribució de les sèries de robatoris

Fixades les diferents configuracions espacials, un primer objectiu ha estat estudiar la distribució temporal dels robatoris en cadascuna de les cel·les.

S'ha passat el test de bondat d'ajustament a la distribució de Poisson i s'ha estudiat l'aleatorietat utilitzant el test de ratxes i el concepte de cotes d'activació.

3.4.1 El test de ratxes

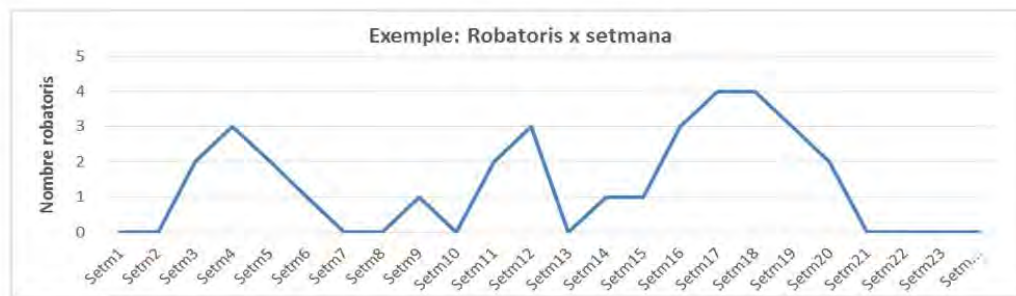
Es vol comprovar que la sèrie de robatoris en una cel·la per determinats lapses de temps no és aleatòria, és a dir, que s'observen ratxes de robatoris en lapses de temps consecutius, seguides de ratxes consecutives sense robatoris o amb molts menys robatoris.

Per fer-ho, s'ha transformat la sèrie temporal en una sèrie de zeros i uns per diferents cotes d'activació: si en un lapse de temps no s'ha assolit la cota, serà 0, altrament 1. Sobre aquestes sèries temporals s'ha comprovat el test de ratxes (Wald and Wolfowitz 1940) que, tot i no ser una prova definitiva d'aleatorietat, n'és un bon indicador.

El test de ratxes dedueix la no aleatorietat quan el nombre total de ratxes de la mostra és pròxim al màxim o al mínim de ratxes possibles segons el total de 0 i 1 disponibles. Fixat el nombre $n = n_1 + n_2$ d'elements, el nombre de zeros (n_1) i el nombre d'uns (n_2), la distribució del nombre R de ratxes té mitjana μ_R i variància σ_R^2 , on:

$$\mu_R = \frac{2n_1n_2}{n_1 + n_2} + 1 \quad \sigma_R^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

L'estadístic $z = \frac{R - \mu_R}{\sigma_R}$ segueix una distribució $N(0;1)$. Si el valor Z és negatiu, indica major agrupament de les ratxes, i si és positiu, indica major dispersió.



Ratxa>0	0	0	1	1	1	1	0	0	1	0	1	1	0	1	1	1	1	1	1	0	0	0	0	
Ratxa>1	0	0	1	1	1	0	0	0	0	0	1	1	0	0	0	1	1	1	1	1	0	0	0	0
Ratxa>2	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0
Ratxa>3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0

Font: elaboració pròpia

Figura 10.-Exemple de ratxes de fets setmanals per diferents cotes d'activació

Test de ratxes

	ratxa>0	ratxa>1	ratxa>2	ratxa>3
Valor de prova: Mitjana	,58	,42	,25	,08
Casos < Valor de prova	10	14	18	22
Casos >= Valor de prova	14	10	6	2
Total casos	24	24	24	24
Nombre de ratxes	9	7	7	3
Z	-1,361	-2,221	-1,413	-1,789
Sig. Assimt. (lateral)	,173	,026	,158	,074

Font: elaboració pròpia

Taula 2.- Exemple de ratxes de fets setmanals per diferents cotes d'activació

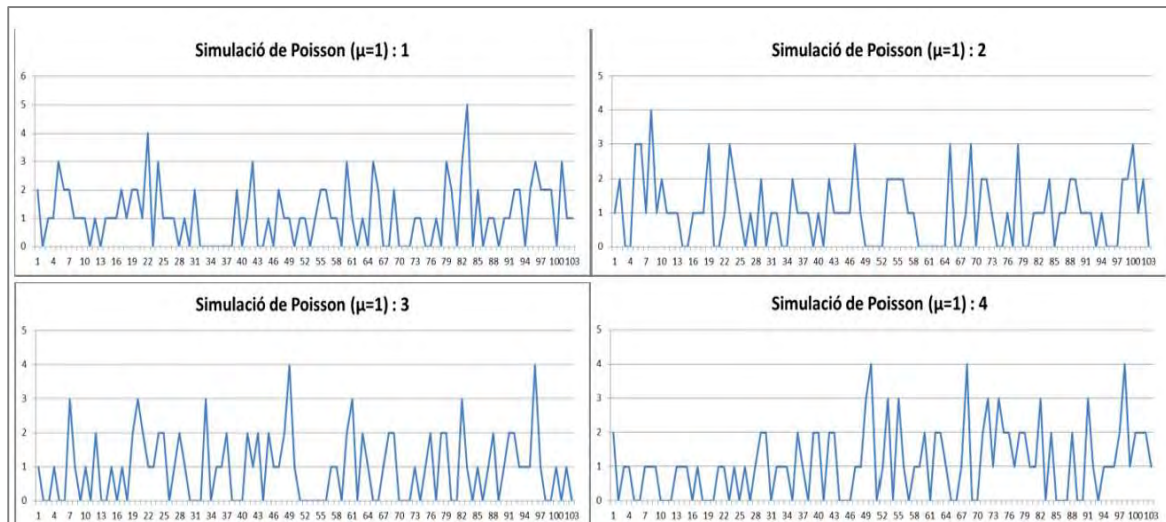
En el cas de l'exemple (Figura 10, Taula 2), la cota que estableix més clarament l'observació de ratxes no aleatòries de robatoris és l'1. En aquest cas, totes les ratxes de robatoris apareixen amb més d'una setmana consecutiva, de la mateixa manera que ho fan les setmanes amb 1 o cap robatori. Clarament en el gràfic s'observen tres "onades" de robatoris en setmanes consecutives. En el cas de la ratxa>2, apareixen dues ratxes amb un sol 1 que acaben decantant cap a la no aleatorietat del test tot i les quatre setmanes consecutives amb més de 2 robatoris. Valoracions semblants es podrien fer pels casos >0 i >3. En tots els casos, el valor negatiu de la Z indica que la no aleatorietat és per un major agrupament de les ratxes respecte el que seria normal en el cas aleatori.

3.4.2 Ajustament a la distribució de Poisson

S'ha fet servir el contrast de bondat d'ajustament a la distribució de Poisson en les sèries de robatoris a domicili. Aquest no té en compte si les dades són o no aleatòries i es limita a fer un recompte de freqüències, independentment de l'ordre en què han aparegut els valors en els diferents lapses de temps. Això no és contradictori amb el fet que una part important de les cel·les segueixin una distribució igual o semblant a la distribució de Poisson.

Segons la teoria de la victimització quasi repetida, hi ha dependència espaciotemporal, és a dir, el fet de robar en un habitatge en el moment t1 condiciona la probabilitat de robar en el mateix habitatge o els habitatges propers en el moment consecutiu t2 . Llavors, la distribució dels robatoris no hauria de ser aleatòria.

De fet, encara que es pogués demostrar que el nombre de robatoris en una cel·la segueix una llei de Poisson, amb esdeveniments independents i per tant aleatoris entre els diferents lapses de temps, es podrien observar onades. Això es pot comprovar amb les simulacions de la distribució de Poisson, que generen onades de forma més o menys freqüent (Figura 11).



Font: elaboració pròpia

Figura 11.-Exemples gràfics de simulacions de la distribució de Poisson

En les onades hi conflueixen factors aleatoris i factors no aleatoris. Amb l'objectiu de distingir-los, mesurar-los i comparar-los, es proposa construir el que s'han anomenat "matrius força no aleatòria" de l'onada, que recullen la probabilitat de transició entre les diferents cotes d'activació habituals en lapses temporals consecutius. Aquestes matrius, permeten visualitzar d'una forma compacte i resumida, per cadascuna de les cel·les i les seves cotes d'activació habituals, quina és la intensitat o "força" no aleatòria, mesurant al mateix temps tant la intensitat de la ratxa com l'ajustament global a la distribució de Poisson. Aquestes matrius són un complement als contrastos anteriors.

La definició que s'ha pres per aquests conceptes és la següent:

Espai o cel·la activada en un període de temps t: es diu que un espai o cel·la està activada en un període de temps t quan, durant aquest període, hi ha hagut algun robatori.

Cotes d'activació dels espais o cel·les: cota mínima del nombre de robatoris a partir del qual s'activarà la cel·la o espai. Activació >n : s'activarà a partir d'n+1 robatoris en el període t. Quan no s'especifiqui la cota d'activació s'entendrà que és >0.

Onada d'un espai o cel·la en lapses de temps consecutius: és la probabilitat d'observar ratxes d'activacions per cotes semblants en una cel·la o espai, és a dir, ratxes de cotes altes d'activació seguides de ratxes de cotes baixes d'activació. També es pot dir que és la probabilitat que els valors alts d'activació en un lapse de temps, provinquin de valors alts d'activació en el lapse de temps anterior. El càlcul dependrà del llindar d'activació i del lapse de temps considerat:

$$\frac{\#(\text{Lapses activa} > n, \text{si activa} > n \text{ en el lapse anterior})}{\#(\text{Lapses activa} > n)} := \text{autocorrelació} > n$$

També es calcularà l'onada d'una cel·la per diferents nivells d'activació:

$$\frac{\#(\text{Lapses activa} > n1, \text{si activa} > n2 \text{ en el lapse anterior})}{\#(\text{Lapses activa} > n1)} := \text{probabilitat} > n1 \text{ provingui de} > n2$$

Matriu onada: conjunt d'onades d'una cel·la per tots els llindars d'activació considerats i ordenats de la següent forma,

$$\begin{bmatrix} > n & \dots & > n \text{ amb } > 0 \\ \vdots & \ddots & \vdots \\ > 0 \text{ amb } > n & \dots & > 0 \end{bmatrix}$$

En aquesta matriu els valors de les files són creixents, ja que els "casos favorables" augmenten mentre que els "casos possibles" es mantenen constants.

Comparativa amb la matriu "onada" de Poisson

Donada una cel·la o espai, es pot contrastar la matriu onada obtinguda de la mostra, amb mitjana μ , amb la que s'obtingria d'una distribució teòrica de Poisson de la mateixa mitjana. En el cas teòric, la probabilitat que un estat $>n_1$ provingui d'un estat $>n_2$, és igual a $P\{>n_2\}$, per la aleatorietat i, per tant, independència entre els esdeveniments entre els diferents lapses de temps. Així, la matriu onada de Poisson seria la corresponent a les probabilitats d'observar els següents estats:

$$Poiss \begin{bmatrix} > n & \dots & > 0 \\ \vdots & \ddots & \vdots \\ > n & \dots & > 0 \end{bmatrix}$$

Aquesta matriu recull el que es pot anomenar part aleatòria de l'onada, i es pot restar de la matriu onada, amb el que es pot anomenar "força no aleatòria de l'onada":

$$Força \text{ no aleatòria de l'onada} := \begin{bmatrix} > n & \dots & > n \text{ amb } > 0 \\ \vdots & \ddots & \vdots \\ > 0 \text{ amb } > n & \dots & > 0 \end{bmatrix} - Poiss \begin{bmatrix} > n & \dots & > 0 \\ \vdots & \ddots & \vdots \\ > n & \dots & > 0 \end{bmatrix}$$

Si s'obtenen valors positius en alguns elements de la matriu, indica que la probabilitat d'observar onades en aquests és més gran que en la distribució teòrica, que s'anomenarà força no aleatòria de les dades en aquesta cel·la o espai

Probabilitat de transició: Donada una cel·la o espai, es consideraran els diferents nivells d'activació com a successos o estats en què pot trobar-se i que poden ser "0", ">0", ">1", ..., ">n". La probabilitat de transició serà la probabilitat de passar d'un estat a un altre. El càlcul dependrà del llindar d'activació i del lapse de temps considerat:

$$\frac{\#(Lapses \text{ activa } > n_1, \text{ si activa } > n_2 \text{ en el lapse anterior})}{\#(Lapses \text{ activa } > n_2)} := \text{prob. transició de } > n_2 \text{ a } > n_1$$

Matriu de transició : aquestes probabilitats ordenades en files i columnes de la mateixa manera i notació que la matriu de ratxes conformen el que es dirà matriu de transició:

$$\begin{bmatrix} > n & \dots & > n \text{ si } > 0 \\ \vdots & \ddots & \vdots \\ > 0 \text{ si } > n & \dots & > 0 \end{bmatrix}$$

En aquesta matriu els valors de les columnes són creixents, ja que els "casos favorables" augmenten mentre que els "casos possibles" es mantenen constants.

Comparativa amb la matriu de transició de Poisson

Per diferenciar els valor predictius motivats per les onades no aleatòries de robatoris respecte els valors que serien obtinguts amb una distribució teòrica de Poisson, es pot fer la mateixa operació que anteriorment. En el cas teòric, la probabilitat de passar d'un estat $>n_2$ a un estat $>n_1$, és igual a $P\{>n_1\}$.

$$\text{Força no aleatòria de transició} = \begin{bmatrix} > n & \dots & > n \text{ si } > 0 \\ \vdots & \ddots & \vdots \\ > 0 \text{ si } > n & \dots & > 0 \end{bmatrix} - \text{Poiss} \begin{bmatrix} > n & \dots & > n \\ \vdots & \ddots & \vdots \\ > 0 & \dots & > 0 \end{bmatrix}$$

Si en aquesta operació de matrius s'obtenen valors positius en alguns elements, indica que la probabilitat de transició entre aquests és superior que en el cas de la distribució teòrica, és a dir, aquella transició d'estats serà més freqüent observar-la que en la distribució teòrica, perquè les dades estaran més agrupades en ratxes.

Comparativa entre la informació donada per la matriu onada i la matriu de transició

Entre els estats que es detectin onades no aleatòries, generalment també la probabilitat de transició entre ells serà superior a la distribució teòrica. Això indica que, de fet, tant la força de l'onada com la força de transició entre aquests estats tenen valor predictiu. Tant una com l'altra indicaran no aleatorietat i major agrupament entre determinats estats.

3.4.3 Criteris d'optimitat

Per valorar quines són les situacions òptimes per predir i prevenir els robatoris amb força a domicili en una determinada cel·la, s'ha considerat el coeficient de variació, definit com a $CV = \sigma/\mu$, respecte el qual altres treballs sobre la victimització quasi repetida han pres com a referent per detectar zones amb patrons de repetició (Saldana et al. 2018). El CV és una mesura de l'altura habitual de les "crestes" de les onades en relació a l'altura de la superfície (mitjana) respecte el fons. Quan aquestes dues altures són semblants, la situació és òptima per predir, sempre i quan les onades siguin no aleatòries per major agrupament.

La representació gràfica del CV (Figura 12) il·lustra les tres situacions bàsiques que es poden donar (Taula 3).



Font: elaboració pròpia

Figura 12.-Representació gràfica del coeficient de variació (CV)

Fixat un espai (e) i uns lapses de temps (t), queden determinades la mitjana μ i la desviació típica σ , i, per tant, el CV, que indicarà que la segmentació espai-temps és l'òptima quan CV s'aproximi a 1. En el cas que la distribució de freqüències sigui compatible amb una Poisson, el CV serà òptim si $\sqrt{\mu} \approx 1$, que es podria acceptar per valors de μ en l'interval [0.6 , 1.5] i, per tant, quan el CV estigui en l'interval [0.77 , 1.22] (En aquest estudi, finalment, s'ha considerat l'interval [0.7,1.3]).

Amb aquestes consideracions, s'estableix el criteri de cel·la o espai òptim per a fer predicció quan $CV \approx 1$ i que s'observi no aleatorietat per major agrupament.

CV	Situació	Predicció/Prevenició
CV<<1	La desviació típica és bastant inferior a la mitjana i, per tant, les onades de robatoris només es visualitzen "a la superfície", registrant en cada lapse de temps un nombre mínim i bastant superior a 0 robatoris. Podria ser el cas representat a la figura 2. Si només es parés atenció policial en aquest espai quan es superés la cota d'activació que produís ratxes no aleatòries, s'hi deixaria d'anar en molts lapses de temps en què també hi hauria "bastants" robatoris.	No és la configuració més desitjable a efectes de predicció i prevenició: Hotspot massa extens
CV~1	La mitjana i la desviació típica són de magnitud semblant. Quan les distribucions tinguin freqüències coherents amb les de Poisson, llavors això passarà quan l'arrel quadrada de la mitjana sigui pròxima a 1. Es podria acceptar per mitjanes entre 0.6 i 1.5. En aquesta situació, prenent com a referència les freqüències de Poisson, és molt probable que hi hagi lapses de temps sense robatoris. Per exemple, si la mitjana és 0.7, en la meitat de lapses de temps no hi haurà robatoris, i si la mitjana és 1, en un 37% dels lapses tampoc n'hi haurà.	Situació més desitjable pel fet que quan es superi el valor de la cota d'activació adequada, hi haurà alta probabilitat d'onada de robatoris, especialment si les dades són no aleatòries per major agrupament, mentre que quan no sigui així, hi haurà pocs o cap robatori. Òptima per predicció i prevenició
CV>>1	La mitjana és molt més petita que la desviació típica i per tant, en general la sèrie de robatoris tindrà molts zeros i alguns pics puntuals, a vegades amb repetició, però que tendiran a ser considerats com a aleatoris.	Situació no adequada perquè dificulta o fa impossible la predicció, de manera que en aquest cas s'hauria d'ampliar l'espai i/o els lapses temporals. Molt baixa intensitat de robatoris

Taula 3.- Explicació de les 3 situacions possibles en funció del CV

3.5 Model Log Gaussià de Cox i el mètode Integrated Nested Laplace Approximation (INLA)

Per contrastar de forma genèrica el pes de les components estàtica i dinàmica i el patró habitual d'irradiació del risc en la proximitat geogràfica, s'ha plantejat un model spatio-temporal Log Cox Gaussian Process (Diggle et al. 2013; Serra et al. 2014). Aquest model s'ajustarà mitjançant la metodologia Integrated Nested Laplace Approximations - INLA (Lindgren and Rue 2015; Blangiardo et al. 2013).

Com s'ha comentat en el punt 3.1.3, aquest processos són un cas particular d'un Procés de Poisson no homogeni en què la intensitat és aleatòria i de la forma $\lambda(s, t) = e^{z(s, t)}$, amb $\ln(z(n)) \sim N(\mu, \sigma^2)$.

El model parteix d'una aproximació contínua del fenomen, de naturalesa discreta, dels robatoris amb força. Aquesta aproximació es construeix amb la suposició de l'existència d'un risc latent de robatori en l'espai (i el temps) de naturalesa gaussiana, és a dir, que els llocs propers en l'espai (i el temps) tenen un risc semblant (seguint la coneguda "Llei de Tobler" (Tobler 1970)). Així, qualsevol punt en l'espai (i el temps) se li pot assignar un risc mitjançant la interpolació dels punts propers.

Matemàticament, es formula com un procés latent espacio-temporal,

$$Z(s_{ij}; t_j) = Y(s_{ij}; t_j) + \epsilon(s_{ij}; t_j)$$

on $t \in \{t_1, \dots, t_T\} \in D_t$ són els temps en què s'han fet les observacions en el domini temporal D_t , del nombre de robatoris $Z(s_{ij}; t_j)$, en un conjunt de punts de l'espai $s \in \{s_1, \dots, s_M\} \in D_s$, on D_s és el domini espacial, $Y(s_{ij}; t_j)$ és el procés latent i $\epsilon(s_{ij}; t_j)$ és un error iid amb mitjana 0 i variància σ_ϵ^2 .

L'objectiu és predir el valor del que s'anomena camp latent $Y(s_0; t_0)$ a la localització espai temps (s_0, t_0) en funció del vector d'observacions conegudes Z .

Aquest camp latent es pot descompondre en dues parts, una part estàtica i una dinàmica, $Y(s; t) = \mu(s; t) + \eta(s; t)$, $\forall (s; t) \in D_s \times D_t$, on $\mu(s; t)$ representa la mitjana del procés, interpretable com la component estàtica o no aleatòria, i $\eta(s; t)$ representa la part aleatòria o dinàmica del procés, amb mitjana 0 i amb dependència espacial i temporal. L'estimació del camp latent en el punt $(s_0; t_0)$, $\hat{Y}(s_0; t_0)$, es farà mitjançant la recerca d'una funció de predicció que minimitzi l'error quadràtic mitjà, $E \left(Y(s_0; t_0) - \hat{Y}(s_0; t_0) \right)^2$.

Una adaptació d'aquest plantejament per als casos en què la variable d'interès és un comptatge, com és el cas dels robatoris, és una formulació jeràrquica amb probabilitats condicionades i amb la introducció d'una funció link:

$$\begin{aligned} Z(s; t) | Y(s; t), \gamma &\sim \pi(Y(s; t), \gamma), & s \in D_s, & t \in D_t \\ g(Y(s; t)) &= x(s; t)' \beta + \eta(s; t), & s \in D_s, & t \in D_t \end{aligned}$$

On π és una distribució de probabilitat, amb un paràmetre d'escala γ i una mitjana $Y(s; t)$. La funció $g(\cdot)$ és el la funció link que transforma la mitjana de la variable resposta en una suma d'un terme que recull els efectes fixos $x(s; t)' \beta$ (que pot incloure covariables), i un terme que recull els efectes aleatoris $\eta(s; t)$ que es poden modelar mitjançant una matriu de covariàncies espacio-temporal.

Aquesta matriu de covariàncies és la que modelarà la dependència dels punts en l'espai i el temps. A la pràctica, sovint es considera separable, de manera que per una banda es modela la covariància en l'espai (dependència espacial) i per l'altra la covariància en el temps (dependència temporal).

La formulació del model pel cas dels robatoris amb força serà considerant que els comptatges segueixen la distribució de Poisson, i que la funció link és el logaritme (Log Gaussian Cox Process):

$$Z(s; t) | Y(s; t), \gamma \sim \text{Pois}(Y(s; t), \gamma)$$

$$\log(Y(s; t)) = x(s; t)' \beta + \eta(s; t)$$

$$s \in D_s, t \in D_t$$

La funció $\eta(s; t) = \Phi(s; t) \alpha(s; t)$, on $\Phi(s; t)$ és una matriu $m_t \times n_\alpha$ de funcions bàsiques espacials, i $\alpha(s; t) \sim \text{Gau}(0, C_\alpha)$ són els coeficients aleatoris associats en què es suposa que la matriu de covariància és separable, producte de la matriu de covariància espacial i la matriu de covariància temporal.

Com a covariància de l'espai es prendrà la funció Matérn, que estableix que si la distància entre dos punts és $d = \|s_i - s_j\|$, la seva covariància és $C_\nu(d) = \frac{2^{1-\nu}}{\Gamma(\nu)} (kd)^\nu K_\nu(kd)$, on Γ és la funció gamma, K_ν és la funció modificada de Bessel, $k > 0$ és un paràmetre d'escala i $\nu > 0$ és un paràmetre de suavització. Aquesta covariància és estacionària i isotròpica, és a dir, només varia en funció de la distància dels punts.

Un cop ajustats els paràmetres, es determinarà el grau de dependència espacial entre dos punts mitjançant l'estimació de la distància fins on s'irradia el risc de nous robatoris causat pels robatoris d'una cel·la concreta. En aquest sentit, el mètode INLA retorna l'estimació de la distribució del paràmetre l anomenat *practical range*, on $l = \sqrt{8\nu}/k$, que és la distància a partir de la qual la correlació espacial serà inferior a 0.1.

Pel que fa a la covariància temporal, es prendrà una sèrie del tipus AR(1) (Geurts, Box, and Jenkins 1977), de manera que un sol paràmetre, ρ , indicarà el grau de correlació temporal entre lapses successius. El fet que s'hagi comprovat que les dades solen estar agrupades en onades de robatoris fa esperable que aquest paràmetre sigui pròxim a 1 si es consideren lapses temporals d'una setmana.

Amb aquesta metodologia s'estudiaran diferents models de predicció dels robatoris en les regions policials, amb una configuració espai temps de cel·les de 500m de costat i lapses setmanals. Es tracta d'una configuració entre la micro i la macro escala, que

permet fer una lectura de les dependències en l'espai i el temps tant en entorns urbans grans com en entorns rurals. El resultat de l'ajustament d'aquests models ha de permetre contrastar la viabilitat dels plantejaments basats en la irradiació del risc en la proximitat geogràfica.

Per aquests contrastos globals, s'ha optat per períodes temporals de 8 setmanes, en què set s'utilitzen per estimar els paràmetres del model i la vuitena setmana es fa la predicció. Les dades de partida són les sèries de robatoris en totes les cel·les de 500m que van patir algun robatori durant l'any 2019.

3.6 Aplicació del mètode INLA amb el paquet estadístic R

El mètode INLA s'implementa en el programa estadístic R des del paquet anomenat "INLA" o indistintament "R-INLA", des de l'enllaç específic <http://www.r-inla.org>. Concretament, les instruccions serien:

```
# Set INLA repository
options(repos = c(getOption("repos"),
  INLA="https://inla.r-inla-download.org/R/stable"))

# Install INLA and dependencies (from CRAN)
install.packages("INLA", dep = TRUE)
```

3.6.1 Construcció de la malla de punts (mesh)

El mètode INLA requereix la definició d'una malla de punts on aproximar el camp latent. Aquesta malla de punts no és irrellevant, ja que pot condicionar els paràmetres estimats pel model, com són les autocorrelacions temporal i espacial. La malla ha de ser prou fina com per modular els efectes espacio-temporals i prou basta com per evitar càlculs redundants que augmentin innecessàriament el temps de computació. També dependrà de la precisió en l'espai i temps de les dades d'entrada i del sentit de l'aproximació contínua del fenomen estudiat.

En el cas dels robatoris amb força a Catalunya i des de la perspectiva de l'estudi global que s'està portant a terme, s'han escollit majoritàriament cel·les de 250 o 500m pel còmput del nombre de robatoris a tot el territori (250m si l'entorn estudiat és predominantment urbà i 500m si inclou altres tipus de zones de forma dispersa en un territori més o menys extens). Aquesta definició del grid-cell d'entrada és el punt de partida de la malla. Malles més fines que el grid-cell d'entrada o gaire més bastes no tenen sentit.

INLA facilita la funció *inla.nonconvex.hull* per construir la malla segons una Constrained Refined Delaunay Tringulation (CRDT) (Krainski et al. 2019), que compleix la propietat de maximitzar els angles dels triangles que formen la triangulació (i minimitzar la longitud dels costats), per evitar errors d'arrodoniment. El punt de partida de la triangulació pot ser una frontera o polígon que contingui els punts estudiats, però s'hi poden incloure els punts del grid-cell d'entrada (localitzacions dels robatoris).

La funció *inla.convex.hull* té paràmetres que controlen la construcció de la triangulació. Seguint la recomanació d'INLA de fer triangulacions el màxim d'homogènies, s'ha optat per incloure els punts del grid cell d'entrada on hi va haver els robatoris però eliminant els triangles amb costats inferiors als d'una cel·la (normalment de 250 o 500m de costat), i posant un límit a la llargada màxima dels costats dels triangles en les zones sense robatoris. Es tracta dels paràmetres *cutoff* i *max.edge* d'aquesta funció.

3.6.2 Selecció dels priors

Un altre aspecte específic d'INLA prové de la seva vessant bayesiana. Cal especificar informació a priori dels paràmetres objectius del model, concretament de l'autocorrelació espacial i temporal, per a poder fer l'aproximació contínua del model a partir de la solució de l'stochastic partial differential equations (SPDE) .

S'ha optat per fer-ho des de l'òptica dels Penalized Complexity Priors (o PC-Priors) (Daniel Simpson et al. 2017; Fuglstad et al. 2019), els quals permeten aportar la informació a priori com una cua de probabilitat sobre els paràmetres estudiats, del tipus $P(Q(d) > U) = \alpha$ or $P(Q(d) < L) = \alpha$, on U o L és un límit superior o inferior, respectivament, i α és la probabilitat de cua superior o inferior de la distribució anterior.

La funció d'R-INLA per definir el model SPDE en l'opció PC-priors és *inla.spde2.pcmatern*. Els PC-priors provenen dels paràmetres que controlen l'efecte espacial definit per la covariància Matérn: el *prior.range* (r), que fa referència al *practical range* del camp latent, i el *prior.sigma* (σ) que fa referència a la desviació estàndard del logaritme (funció link) dels robatoris en els vèrtexs de la malla.

El prior del *practical range* r es defineix pel conjunt (r_0, p_r) de manera que $P(r < r_0) = p_r$, i el prior de la desviació estàndard σ es defineix pel conjunt (σ_0, p_s) de manera que $P(\sigma > \sigma_0) = p_s$.

Aquests priors, que condicionaran el resultat obtingut, dependran del coneixement previ del fenomen estudiat, i han d'estar en consonància amb la malla construïda. Una malla a petita escala tindrà desviacions típiques més baixes i també la possibilitat d'observar autocorrelació a menys distància.

Pel que fa al prior temporal, també s'especificarà de la forma PC-prior pel conjunt (ρ_0, p_t) on $P(\rho > \rho_0) = p_t$.

Amb aquesta metodologia s'estudien diferents models predictius de robatoris a les regions policials, amb una configuració espai-temps de cel·les de 500m de costat i intervals setmanals. Aquesta configuració es troba entre la micro i la macro escala, cosa que permet fer una lectura de les dependències en l'espai i el temps tant en entorns urbans extensos com en entorns rurals. Únicament per a Barcelona i altres municipis de la seva conurbació s'utilitzen configuracions més petites, de 100m i 250m de costat.

3.6.3 Comparativa de models

Els resultats obtinguts amb INLA es poden valorar des de diferents punts de vista. En ocasions interessa comparar diferents models, i R-INLA ofereix varis estadístics per fer-

ho Marginal likelihood, Contitional Predictive Ordinates (CPO), Predictive Integral Transform (PTI), Deviance Informations Criterion (DIC) o Watanabe-Akaike Information Criterion (WAIC)(Krainski et al. 2019).

Pel cas de la predicció dels robatoris, s'ha prescindit d'aquests estadístics i s'ha optat per per mesurar l'ajustament i la predicció dels models mitjançant el coeficient de determinació o R quadrat (R^2). Les dades obtingudes amb el model aproximat per INLA es projecten en grid-cells de diferents mides. La projecció inicial es fa en el grid-cell de mida més petita (100, 250 o 500m) i es fa el sumatori del nombre de robatoris en aquestes cel·les per obtenir la projecció als grid cells de mida més gran (de 250m a 5 km).

3.6.4 *Del camp latent a la predicció pràctica*

L'interès principal de l'estudi del model log Gaussian de Cox és contrastar-ne la capacitat predictiva en l'espai i el temps. Es necessita una mesura de la precisió d'aquesta predicció comparada amb els valors reals i s'ha optat per calcular-lo segons l'ajustament del nombre de robatoris previstos i reals en cel·les de 5km de costat. La comparació d'aquests dos vectors de dades es fa mitjançant el coeficient de determinació R^2 .

El càlcul del nombre de robatoris en aquestes cel·les de 5km a partir de l'estimació de la variable resposta del camp latent segueix l'estratègia d'integrar o sumar el nombre de robatoris en les 100 subcel·les de 500m de costat que la componen. A la vegada, el nombre de robatoris en cadascuna d'aquestes subcel·les s'obté segons una lectura de la variable resposta (nombre de robatoris) en qualsevol dels seus punts.

Aquesta estratègia és coherent amb les propietats additives de la distribució de Poisson. En cadascuna de les cel·les hi ha un procés de Poisson de mitjana λ_i i en l'espai unió d'algunes d'aquestes cel·les hi haurà un procés de Poisson de mitjana $\sum \lambda_i$.

Al tractar-se d'una aproximació contínua amb una resolució en les dades d'entrada superior de 500m, s'entén que el valor del camp latent de dos punts en distàncies inferiors a 500m és igual o molt semblant, de manera que prenent el valor en qualsevol punt dins un quadrat de 500m s'obté el valor vàlid per aproximar el nombre de robatoris (intensitat mitjana de robatoris per aquella setmana segona la distribució de Poisson).

Nota: és millor el plantejament continu que el lattice (discret) perquè l'estructura de la dependència espacial és més flexible que pel cas d'una matriu d'adjacència.

3.7 Anàlisi de components principals

Amb l'objectiu d'estudiar patrons d'activació setmanal conjunta de les cel·les, s'ha utilitzat l'Anàlisi de les Components Principals (Principal Components Analysis, PCA) (Pearson K. 1901; I.T. Jolliffe 2010), on les variables considerades són les pròpies cel·les. S'ha treballat amb components principals rotades, segons una rotació varimax, per minimitzar el nombre de cel·les amb més pes en cada component, i s'analitza l'ajustament a l'històric de dades mitjançant regressions lineals.

El mètode de les components principals és capaç de detectar la redundància en grans conjunts de variables i permet simplificar-ne l'estructura bo i conservant un alt percentatge de la variància del total d'informació. Iterativament, cerca combinacions lineals de les variables que compleixin el requisit de màxima variància i que les combinacions siguin independents (ortogonals) entre elles. La variabilitat continguda en les variables originals es reparteix en les combinacions lineals, dites components principals o factors, de manera que cadascuna representa un percentatge conegut de la variància global. Les components principals es numeren de forma ascendent, sent la primera component la que té el percentatge més elevat de la variància global i així successivament. Geomètricament, el mètode consisteix en trobar un el·lipsoide els eixos del qual concentrin la major variabilitat.

Sense que hi hagi criteris absoluts, es sol considerar que si la suma dels percentatges de variació de les components principals és superior a un 70% de la global, aquestes representen bé l'explicació de la variància total i, per extensió, el problema o model que es vulgui estudiar. Aquestes combinacions lineals equivalen a índexs abstractes que no sempre tenen una interpretació fàcil. Com a orientació, el coeficient que multiplica les variables en la combinació lineal de cadascuna d'aquestes components mostra el pes que hi té cada variable que, a la vegada, poden ser positius o negatius. Les variables que tenen més pes en cada component principal són les que millor la representen i això pot ajudar a interpretar-les.

En el cas dels robatoris amb força, s'han estudiat les components principals generades per sèries setmanals dels robatoris en les cel·les quadrades de 5km, tant de Catalunya com, més específicament, de la Regió Policial Metropolitana Nord, corresponents a les 51 setmanes completes de l'any 2019. Les variables considerades són les pròpies cel·les, i les components principals atorguen pesos diferents a cadascuna d'elles. Cada component principal es pot representar en un mapa a la manera d'un hotspot map on el color de cada cel·la representa el pes en la combinació lineal de la component principal. Distingint amb colors diferents els pesos positius i negatius, cadascun d'aquest mapes indiquen correlacions del risc entre cel·les.

El mètode s'ha aplicat sobre la matriu de correlacions de les variables, que equival a estandarditzar les variables, de manera que no es té en compte el nombre absolut de robatoris en cada cel·la sinó la seva variació al llarg de les setmanes.

Les components principals obtingudes es poden sotmetre a rotacions que conservin les mateixes característiques d'ortogonalitat i percentatge de variància total explicada però que prioritzin altres aspectes interessants, com pot ser minimitzar el nombre de variables (cel·les) que tinguin un pes alt en cada component. Aquest tipus de rotació, coneguda com a varimax, facilita la interpretació del resultat.

Aquestes components es poden simplificar seguint criteris de reducció de variables, eliminant aquelles cel·les que tenen pesos més baixos i identificant les representatives, per obtenir el que anomenem constel·lacions. Sobre aquestes, s'analitza de nou l'ajustament a les dades mitjançant regressions lineals.

S'ha fet una anàlisi descriptiva del perfil de les cel·les representatives de cada constel·lació, des del punt de vista de l'entorn, la proximitat entre elles i les característiques dels robatoris que s'hi localitzen.

Finalment, s'ha analitzat la capacitat predictiva de les constel·lacions estudiant la no-aleatorietat de les seves onades i l'ajustament de les regressions lineals, sobre les constel·lacions, a les dades reals.

Per tal de contrastar l'estabilitat de les components principals de l'any 2019, aquestes s'han utilitzat, també, per ajustar les dades de l'any 2018.

3.8 Formulació probabilística del patró “all-in-one”

Les constel·lacions de robatoris descrites en el punt anterior es basen en un patró d'actuació dels grups criminals que s'ha anomenat “all-in-one”. Aquest patró estableix que hi ha una elecció prèvia d'un conjunt finit de zones $\{Z_1, Z_2, \dots, Z_n\}$ que victimitzaran de forma contínua durant un període temporal d'unes quantes setmanes.

Per tal de relacionar i valorar aquest patró respecte els resultats obtinguts en l'anàlisi de les dades, se'n fa una proposta de formulació probabilística.

Concretament, es considera que fixades les zones objectiu on desplegar l'onada de robatoris, els lladres escullen les zones on robar, segons una loteria del tipus $\begin{pmatrix} Z_1 & Z_2 & \dots & Z_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$, on Z_i representen les zones de la constel·lació i p_i les probabilitats de ser escollides, amb $\sum_1^n p_i = 1$.

Aquesta formulació és temporalment agregada, per setmanes i zones, amb un model probabilístic que determina les preferències mitjanes durant tota l'onada (Figure 1).

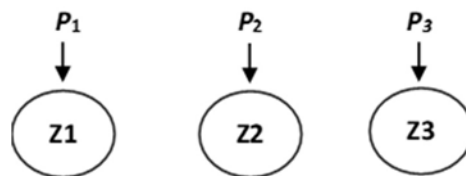


Figura 13.- Diagrama d'una loteria (Formulació agregada)

El model és compatible amb una formulació desagregada, robatori a robatori, on l'elecció de la zona després de cada robatori segueix un procés estocàstic (REF) com el representat a la Figura 2.

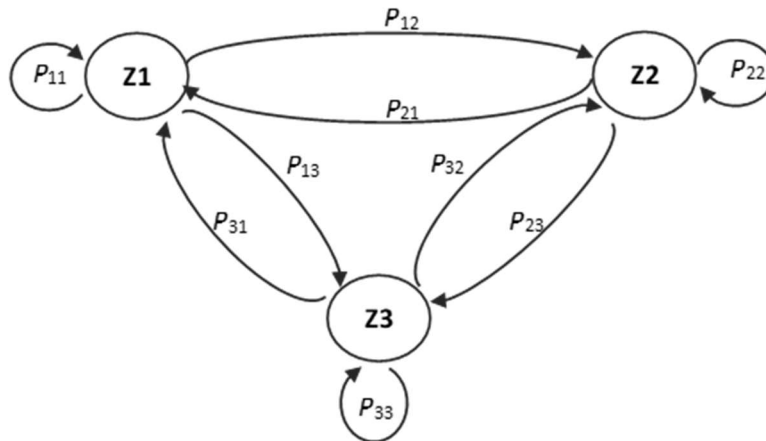


Figura 14.- Diagrama d'un procés estocàstic (Formulació desagregada)

Després d'un primer robatori, l'elecció del següent hauria de determinar si continuar victimitzant la mateixa zona (amb probabilitat P_{ij}) o desplaçar-se a una de les alternatives (amb probabilitats P_{ij}). Si es formula com una cadena de Màrkov, la loteria anterior es podria considerar el límit de les probabilitats de transició, és a dir, les probabilitats "a llarg termini" en què cada zona seria escollida.

Tot i així, és difícil acceptar les condicions d'independència condicional de Màrkov, en què l'estat següent només depèn de l'estat actual. Sembla lògic pensar que si només s'ha fet un robatori en una zona, s'intenti repetir-hi un o dos cops més abans de canviar a una altra. I, contràriament, si ja s'han fet 3 o 4 robatoris a la mateixa zona, sembla més probable canviar. De fet, hi pot haver molts altres factors que influencien aquestes decisions i que en compliquin l'estimació i modelització. Per contra, és relativament senzill estimar el "límit" d'aquestes decisions, en forma de loteria i que representa l'atractivitat mitjana de les zones escollides on robar (p_i). El fet que tant les prediccions com les estratègies preventives habitualment es facin de forma agregada, per setmanes i zones de risc, també afavoreix el plantejament resumit de les loteries.

Anomenant X a la variable aleatòria de Poisson que mesura la intensitat del nombre de robatoris d'un autor o grup d'autors en un determinat lapse temporal, $X \sim Poiss(\lambda)$, i Y_i el nombre de robatoris que aquest autor o grups d'autors farà a la zona Z_i , llavors $Y_i / (X = k)$ segueix una distribució binomial de paràmetres k i p_i , i el nombre de robatoris a la zona Z_i seguirà una distribució de Poisson de paràmetre igual a λp_i . De forma inversa, per la propietat additiva de la distribució de Poisson, la suma de tots els robatoris efectuats pel mateix grup d'autors i en totes les zones prèviament escollides, és la distribució de Poisson X de la intensitat de robatoris del grup durant el període, $\sum Poiss(\lambda p_i) = Poiss(\lambda) \sim X$. En cas que només existís una zona objectiu, Z_1 , la formulació seria equivalent a la clàssica de la victimització quasi repetida.

3.9 Altres metodologies

En el decurs de la recerca, han sorgit diferents metodologies matemàtiques o estadístiques que s'han acabat descartant. Tot i així, algunes d'elles poden ser molt adequades per a seguir investigant els patrons espaciotemporals i, en general, les relacions i agrupacions

entre els robatoris. Entre aquestes tècniques, l'anàlisi de conglomerats (Cluster Analysis) (Everitt et al. 2011) i la teoria de grafs (Biggs, Lloyd, and Wilson 1986; Basart i Muñoz 1993) que permeten estudiar la relació entre robatoris i els clústers de robatoris. Aquestes tècniques, juntament amb l'Anàlisi de Components principals, actualment solen estar incorporades en plantejaments més globals com el *Machine Learning* o el *Deep Learning* (Bishop 2006), tot i que amb nomenclatures diferenciades. Per exemple, els mètodes d'aprenentatge no supervisat serien equivalents a l'estudi de components principals, i l'aprenentatge competitiu o les xarxes de Kohonen serien tècniques d'anàlisi clúster (Pérez Delgado and Martín Martín 2003). També hi ha plantejaments que combinen teoria de grafs amb *Deep Learning* per a l'estudi de relacions espaciotemporals (Y. Zhang and Cheng 2020; B. Wang et al. 2018).

Aquests últims plantejaments més ambiciosos poden tenir l'inconvenient de ser poc transparents en el seu funcionament intern, a vegades qualificant-los com a "caixes negres". En la present recerca, es dona més èmfasi en el fonament i el plantejament dels models, i s'han prioritzat les metodologies simples, específiques i fàcilment interpretables.

4 RESULTATS

Els resultats d'aquesta tesi es presentaran mitjançant tres articles, un publicat i un enviat a la revista i un pendent d'enviar.

Article 1: Boqué P, Serra L, Saez M. "Surfing" burglaries with forced entry in Catalonia. Large-scale testing of the theory of near repeat victimisation, *European Journal of Criminology*. Nov 3, 2020.

Doi: 10.1177/1477370820968102.

[2019 Impact Factor: 1.671; Q1 Social Sciences (Law), position 37 out of 751]

Article publicat:

Article enviat:

Article 2: Boqué P, Saez M., Serra L. Need to go further: the all-in-one repeat pattern for spatiotemporal prediction in heterogeneous environments.

Enviat a la revista: *European Journal of Criminology*.

Article pendent d'enviar:

Article 3: Boqué P, Saez M., Serra L. Burglaries constellations: the hot trail left by the all-in-one spatiotemporal repeat pattern in heterogeneous environments.

Pendent d'enviar a una revista.

4.1 Article 1

“Surfing” burglaries with forced entry in Catalonia. Large-scale testing of the theory of near repeat victimisation

Boqué P, Serra L, Saez M

European Journal of Criminology. 2020 Nov 03. Doi: 10.1177/1477370820968102

Box 1 | Visió general de l'article 1

El que ja se sap sobre la temàtica

- El patró de victimització repetida i quasi repetida s'ha observat en diferents països, especialment en els robatoris a domicili.
- Es poden construir models matemàtics per a la predicció dels delictes.

El que aporta de nou l'estudi

- Es demostra, per primera vegada a Catalunya, que els robatoris amb força a domicili segueixen el patró de victimització quasi repetida.
- La predicció dels robatoris és possible i s'estableixen criteris per determinar quan i on es pot fer.
- Les peculiaritats de l'entorn geogràfic català obliguen a formular els algoritmes predictius des d'una configuració espacial a gran escala.

Boqué P, Serra L, Saez M. “Surfing” burglaries with forced entry in Catalonia. Large-scale testing of the theory of near repeat victimisation, *European Journal of Criminology*. 1-22 Nov 3. 2020

<https://doi.org/10.1177/1477370820968102>

© The Author(s) 2020

Abstract

In recent years, various academic studies have proposed crime forecasting models based on the concept of repeat victimisation. Some of them have been modelled from the area of differential equations and others from the perspective of spatiotemporal statistics, within the framework of point processes. In general, these models have tended towards a certain sophistication in their formulation, which at times impedes understanding of the predictive mechanism and how they adapt to different realities. Predictive models that function well in one environment or society do not appear to do so in others. In this article, the possibility of crime forecasting for burglaries with forced entry in Catalonia is studied from the perspective of near repeat victimisation on a larger territorial scale than is usual. To this effect, the explicative and predictive possibilities of this criminological theory are explored and a predictive system that does not require mathematical or statistical models is proposed. This system is adaptable to different environments and gives police forces the opportunity to improve preventative strategies and to optimise resources using standard tools. Last, the limitations of this approach are debated and new lines of investigation proposed that could increase its predictive capacity without abandoning the concept of repeat victimisation.

Key words: algorithm, heterogeneous territory, near repeat victimisation, police patrol, predictive policing

1 Introduction

Can forced-entry burglaries be predicted? And if so, is this forecast useful for changing police preventative strategies? These are questions that are nowhere near being definitively answered. In recent years, numerous academic studies (Farrell G. and Pease 1993; Farrell G., Pease, and Phillips 1995; Pease 1998; Pease and Farrell 2001; M. Townsley 2003; Michael Townsley, Homel, and Chaseling 2000) have proposed crime forecasting models based on the theory of near repeat victimisation, a variant of repeat victimisation. The latter posits that, after victimising a target, criminals tend to target it again after a short period of time, while in the case of near repeat victimisation the target can also be one with similar characteristics as the first.

This criminological theory was formulated and popularised when it started being applied to different types of crimes almost three decades ago. In the case of burglaries, the theory has helped to understand their distribution and concentration in space and time. Other classic criminological theories such as crime opportunity (Gottfredson and Hirschi 1990), routine activity (Cohen and Felson 1979), rational choice (R. V. Clarke and Cornish 2018) (all of which are mentioned in the article *Opportunity Makes the Thief. Practical theory for Crime Prevention* (Felson and Clarke 1998)) and foraging behaviour (Bernasco 2009) help to substantiate and reinforce it. Two explicative hypotheses are also formulated: the heterogeneity or flag explanation (Tseloni and Pease 2003) and the boost explanation (Kate J Bowers and Johnson 2004; Ornstein and Hammond 2017), also known as sudden increase in risk or dynamic risk. The former explains what is known as ‘static’ risk because it does not change over time or it does so over a long period of time, establishing that there are some houses that are more attractive to burglars than others, be it because they have less security measures, they are easy to access or they promise more valuable booty, and so on. The second hypothesis posits that after a burglar has successfully stolen from a residence, they are likely to burgle either the same residence or a residence close to it again looking to repeat their success by replicating the first burglary.

The repeat and near repeat theory of victimisation has been empirically proven in many countries for residential burglaries (Shane D Johnson et al. 2006; Kikuchi, Amemiya, and Saito 2010; Vijaya Kumar 2011; Z. Wang and Liu 2017). To this effect, the coordinates of place where the burglary took place and the time it occurred (normally during daytime) are needed. Without it always being stated explicitly, it is usually assumed that the environment studied is relatively homogenous. With this information, the Knox test for different spatiotemporal segments is normally applied. This test compares the distribution of the spatiotemporal points (x,y,t) of the burglaries with the distribution of points obtained in the random case, normally using the Poisson distribution as a reference (“Department of Criminal Justice | Near Repeat Calculator” n.d.).

In addition to this statistical test, the phenomenon can also be visualised with crime mapping. A series of weekly and fortnightly hotspot maps indicate that there is a spatiotemporal dynamic appearing and disappearing in different places at different time intervals.

This clustering dynamic for burglaries with forced entry has been compared with contagion and epidemic processes, and also with earthquake aftershocks. These already mathematically modelled phenomena inspired the equations of the near repeat victimisation prediction models. There are models in the area of differential equations that simulate these hotspot generating processes (Short et al. 2008) and there are also models based on the spatiotemporal perspective within the framework of point processes (Mohler et al. 2011). In

general, both approaches have tended towards a certain sophistication in their formulation, which sometimes hampers understanding of their predictive mechanics and their practical adaptation or application to different territorial and social realities.

These patterns have not only been modelled mathematically, but computer software has also been designed to automatically generate forecasts, the most well-known of which is the American PredPol (“Predict Crime | Predictive Policing Software | PredPol” n.d.). In Europe the leading software is the German Precobs (“PRECOBS Ifmpt - Institut Für Musterbasierte Prognosetechnik” n.d.).

These and other programs have been applied to different countries, normally on a local rather than a national scale, and especially in large cities. While the results obtained seem to have been mainly positive, they have also sparked some discussion about the degree to which they contribute to lowering the crime rate. They are generally seen as predictive models that work well in some environments or societies, but do not seem to work, or at least not to work very well, in others. If the models are based on repeat victimisation, clearly part of their success will depend on the intensity of this victimisation. To this effect, a study carried out in Brazil (Spencer Paul Chainey and Figueiredo 2016) has demonstrated the major difficulties involved in observing this phenomenon in cities with a lifestyle and residences dissimilar to those in the Anglo-Saxon world. In Mediterranean and Latin American countries urban zoning and residences and the way of living in them are all different and can affect the phenomenon of near repeat victimisation. In the case of the Brazilian city studied (Bello Horizonte), slums (*favelas*) are interspersed with neighbourhoods comprised of residential skyscrapers, so it is logical to think that in such diverse environments near repeat victimisation manifests itself in dissimilar ways.

Catalonia is another case in point due to its great territorial heterogeneity, both as a whole and in its inhabited area. Catalan villages and cities are usually visualised as a mosaic made up of old towns, modern areas and flats, areas with villas and bungalows, neighbourhoods of mostly terraced houses, areas with tourist apartments and rural zones with scattered residences. Clearly, this heterogeneity affects crime opportunity and is a determinant in shaping crime hotspots (Vozmediano Sanz and San Juan Guillén 2010) and, of course, near repeat victimisation hotspots.

In 2013, given the surge in forced entry burglaries, the Catalan police force Mossos d’Esquadra together with the University of Girona began to study the possibility of adopting some of the existing models and software for predicting burglaries and improving their preventative policing strategies. The study brought to light the difficulty in applying them in such a heterogenous environment and the need to explore other alternatives. Both the dynamics of the waves of burglaries which appeared to be happening in different parts of the territory, and more isolated incidents of residential burglaries, disconcerted police. Their main concern was, and still is, to find a system that would enable them to determine which of the areas affected by burglaries is the most suitable for intensifying preventative actions.

The aim of the present study is just that: to determine the most sensitive areas where waves of burglaries are most likely to occur and to study their temporal dynamic. The aim is to test near repeat victimisation on a larger scale than is habitual, meeting the need to understand the general dynamic of residential burglary hotspots across the territory. The hypothesis is that if near repeat victimisation is observed on a small scale, normally at a distance of just a few hundred meters, then it will also be observed on a large scale (more than 1 km). However, in these more extensive areas more than one focus of near repeat victimisation

may coincide temporally, which could mask the analysis of the phenomenon. The study will demonstrate that criteria for determining when the theory of near repeat victimisation can and cannot be used to make forecasts in a specific area can be established.

While enlarging the scale decreases the precision of a possible forecast, it also increases the chances of observing repetitions. Furthermore, too big a scale could make effective preventative actions impracticable, so the aim is to find the smallest spatiotemporal segmentation where waves of burglaries can be observed which, at the same time, can potentially facilitate optimal preventative policing actions.

The article is structured as follows: presentation of data, method to test near repeat victimisation on a large scale, validation with data, limitations, discussion on the viability of possible predictive models and conclusions.

2 Data and methodology

2.1 Data

The territory: Catalonia

Catalonia is a country in Mediterranean Europe on the east coast of the Iberian Peninsula, occupying 5.5% of this land area. It is bordered by the Pyrenees and France to the north, the Mediterranean Sea to the east and Spain to the west and south.

The territory is divided into a total of 9 police regions. Each region includes a set of Basic Police Areas (ABP), the territory's primary units usually encompassing several municipalities and defined by geographical and policing criteria, from where most of the police patrols are dispatched. These regions are geographically diverse; the territory of the Barcelona Metropolitan Basic Police Area is almost entirely urban, while the Eastern Pyrenees Basic Police Area is mountainous with few, generally small villages. The largest urban centres are the metropolitan area of Barcelona and coastal zones, which are also where most residential burglaries with forced entry occur. Although Barcelona is the Catalan city with the most burglaries, this study focuses on the other eight regions, which are where most waves of burglaries tend to be observed with no or very few reported burglaries for several weeks followed by sudden spates of these crimes over short, consecutive periods of weeks, which end to return to further long periods without any burglaries

Crime data

The data used in this study are the reports of burglaries with forced entry made to the police in Catalonia in 2014 and 2015. The location in Universal Transverse Mercator coordinates and a window of time (generally less than a day) when the burglary is assumed to have occurred are recorded for each burglary. The type of residence burgled – a flat, a house or a farmhouse – and whether it is a first or second residence are also specified. According to this data there were 2,928 burglaries with forced entry in Catalonia in 2014 and 27,488 in 2015, representing an increase of 10.3%.

These data have been provided by the Government of Catalonia Police – Mossos d'Esquadra (PG-ME).

2.2 Methods

The analytical framework followed to study the burglaries was (1) doing a descriptive analysis, (2) segmenting the territory of Catalonia into square cells and the time into fixed time intervals, (3) projecting the spatiotemporal points (x,y,t) of the burglaries onto the square based prisms obtained (4) studying the temporal series associated with each prism obtained from counting the number of burglaries occurring in the total time period considered, (5) studying the non-randomness of the series of data, and last (6) proposing and debating possible predictive algorithms.

The size of the cells for this study were 5km by 5km and the time intervals coincided with the natural weeks. The average weekly burglaries were calculated for each of these cells (Figure 1) and the values obtained were recorded on a histogram. This average served to classify the intensity of the burglaries.

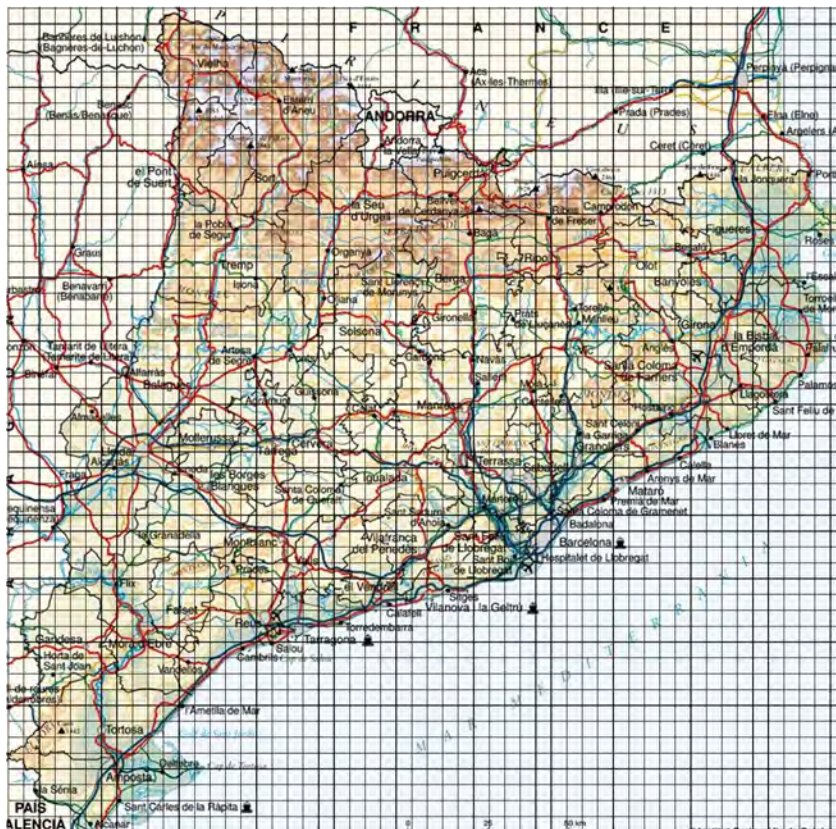


Figure 1: Map of Catalonia with a grid of square cells of 5 km \times 5 km superimposed onto it.

Once these spatiotemporal segments were established, the randomness of the robberies was studied using the runs test and the concept of activation levels (Appendix, Table 2) to verify that waves of burglaries can be observed and predicted. The goodness-of-fit contrast was also used on the Poisson distribution.

The two contrasts provide different, complementary information. According to the theory of near repeat victimisation, there is a spatiotemporal dependency in the sense that a residential burglary occurring in the moment t_1 conditions the probability of there being a burglary in the same or a nearby residence in the consecutive moment t_2 . Hence, the distribution of the burglaries should not be random. This is not contradictory to a large proportion of the cells

remaining distributed the same as or similar to the Poisson distribution. The goodness-of-fit test does not contemplate whether the data are random or not and only counts the frequencies, independently of the order in which the values in the different time intervals have appeared.

In fact, even if it could be proven that the number of burglaries in a cell follows a Poisson law with independent, and therefore, random events between the different time intervals, waves of burglaries could still be observed. This can be tested with the simulations of the Poisson distribution, which generate more or less frequent waves. It could therefore be said that random and non-random factors converge in the waves. With the aim of identifying, measuring and comparing them, we propose constructing what have been called “non-random force matrices” of the wave, which capture the probability of transition between the different usual levels of activation in consecutive time intervals. These matrices enable us to visualise the intensity or the non-random “force” of each of the cells and their usual levels of activation in a compact, summarised way, while at the same time measuring the intensity of the run as a global fit of the Poisson distribution. These matrices complement the previous contrasts.

Last, to evaluate what the optimal situations to forecast and prevent burglaries with forced entry in a particular cell are, the variation coefficient, defined as $CV = \sigma/\mu$, was considered, which has been taken by other research on near repeat victimisation as a benchmark to detect areas with repeat patterns (26).

It could be said that CV is a measure of the usual height of the ‘peaks’ of the waves in relation to the height of the surface (average) with respect to the bottom. When these two heights are similar, the situation is optimal for predicting, provided that the waves are non-random for larger groupings.

The graphic representation of the CV shows the three basic situations that can occur (Figure 2 and Table 1).

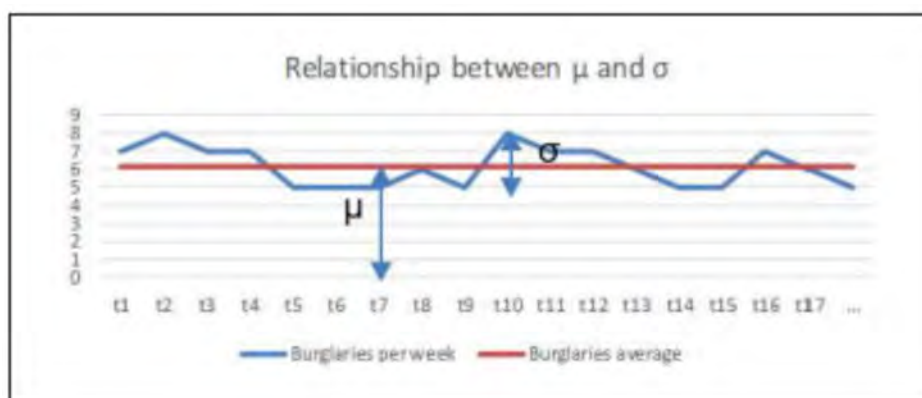


Figure 2: Graphical representation of the coefficient of variation.

Table 1. Interpretation of the coefficient of variation.

CV	Situation	Forecast/prevention
CV<<1	The standard deviation is rather below the average and, therefore, the waves of burglaries are visualized only 'on the surface', registering a minimum number that is significantly higher than 0 burglaries in each time interval. This could be the case represented in Figure 2. If police attention focused only on this space when the activation level that produces non-random runs is exceeded, they would not go to this area during many of the time intervals when there would also be 'quite a lot' of burglaries.	This is not the most desirable configuration for forecasting and prevention: Hotspot too extensive
CV~1	The average and the standard deviation are a similar size. When the distributions have frequencies that are consistent with the Poisson distributions, then this will happen when the square root of the average is near 1. This could be acceptable for averages between 0.6 and 1.5. In this situation, taking the Poisson frequencies as a reference, it is very likely that there are time intervals without any burglaries. For example, if the average is 0.7, in half the time intervals there will be no burglaries, and, if the average is 1, there will be no burglaries in 37 percent of time intervals.	This is the most desirable situation because, when the value of the adjusted activation level is exceeded, there is a high probability of a wave of burglaries, especially if the data are non-random for a larger grouping, whereas when this is not the case there will be few or no burglaries. Optimal for forecasting and prevention
CV>>1	The average is much smaller than the standard deviation and, therefore, in general the series of burglaries will have a lot of zeros and some occasional peaks, sometimes with repetitions, but they will tend to be considered as random.	This is an inadequate situation because it makes forecasting difficult or impossible; in this case the space and/or the time intervals should be extended. Very low intensity of burglaries

Optimal configuration of the spatiotemporal segmentation

Once a space (e) and time intervals (t) are established, the average μ and the standard deviation σ are determined, in addition to the CV, which will indicate that the spatiotemporal segmentation is optimal when the CV is closest to 1. Where the distribution of frequencies is compatible with a Poisson, the CV will be optimal if $\sqrt{\mu} \approx 1$, which could be accepted for values of μ in the interval [0.6, 1.5] and, therefore, when the CV is in the interval [0.77, 1.22] (In the end the interval was considered as [0.7, 1.3] in this study).

Bearing these considerations in mind, the optimal criteria for the cell or space to make forecasts when $CV \approx 1$ is established, and non-randomness is observed for larger groupings.

Analysis of the maps

Crime maps (27-29) are a usual police analysis technique (27-29) that provide very valuable information to interpret criminal phenomena. The explanation of static risk is usually related to both the immediate environment (type of residence and neighbourhood) and the general environment (whether it is a tourist area, if it is in the centre or on the outskirts of a town, whether it is close to a fast road, and so on). All this knowledge, which can be collected through

layers of information, is not usually recorded in police reports in a structured, automatic way but it can be visualised on the map.

In this study, crime data has been represented on different scale maps to facilitate interpreting the phenomenon and the results obtained. The map of Catalonia, the grid of cells, the map of hotspots and other classifications of the cells have been represented on a scale of 1:100. The individual cells have been represented on a scale of 1:50.000, allowing the type of zone of the targeted residence (urban zone or an area with scattered houses, zones with houses or blocks of flats, residential estate, city centre, tourist area, rural area, near fast roads, and so on) to be identified. These maps have been taken from the Cartographic Institute of Catalonia (30) and the grid of cells was superimposed onto them.

3 Results

Function of the distribution of burglaries in fixed time intervals

The daily average of residential burglaries in Catalonia is 70.5 and the histogram shows a relatively symmetrical distribution around this figure, although more frequent values of between 55 and 75 and less frequent but more variable values of between 76 and 125 can be observed (Appendix, Graphic 1)

The same histogram was produced for different police regions (Appendix, Graphic 2 to 6) showing that as the daily average of burglaries in a territory lowers, the frequencies diagram goes from having a shape that is similar to the Normal distribution to having a clearly asymmetrical shape, similar to the Poisson distribution. This property can be observed with any type of territorial division: the lower the number of burglaries there are in a territory, the closer the distribution is to a Poisson; and the more burglaries there are the greater, the tendency towards a Normal distribution.

Taking a segmentation of cells of 5 km x 5 km and time intervals of a week, Figure 3 shows different coloured cells depending on the average weekly burglaries. In the 2 years under study there were burglaries in 807 cells, generally with a very low intensity: in practically 3 out of 4 cells (71.6%) the average weekly burglaries was lower than 0.25, in other words less than an average of 1 burglary per month. Only 3 cells have an average higher than 30 burglaries a week, and they are located in the city of Barcelona. The 17 cells that present an average of between 5 and 30 weekly burglaries were located in Barcelona and some other large cities, some of which have a tourism profile. The rest, a little over one hundred, had averages of between 0.6 and 5 burglaries per week.

Based on the goodness-of-fit test of the Poisson distribution, it was observed that the frequencies of 98.5% of the cells with an average lower than 0.6 were compatible with the Poisson distribution. Of the 64 cells with averages between 0.6 and 1.5, 44 (64.1%) passed the Poisson goodness-of-fit test, and some of the other 20 would have passed it if they had had a less heavy tail: in other words, by eliminating the 2 or 3 weeks when “too high” a number of burglaries were registered. Last, the cells with an average higher than 1.5 had a distribution of frequencies that is more incompatible with the Poisson distribution, such that only 19.4% of the cells passed the test and, among these, there were none with an average higher than 6 burglaries per week.

Hence, it is confirmed that the less burglaries there are in a cell, the more the distribution of their frequencies tends towards the Poisson. Taking this property to the extreme, the conclusion could be reached that the probability of a specific residence being burgled at a specific time follows a Poisson distribution. Then, the added property of this distribution, $Poiss(\lambda)+Poiss(\mu)\sim Poiss(\lambda+\mu)$, would explain that as the number of residences in a territory increases, the probability of the burglaries following a Poisson distribution is equal to the sum of the Poisson distributions for each residence. By means of the Central Limit Theorem, it is deduced that in large territories or in territories with a great density of residences, the Poisson distribution of burglaries will tend towards a Normal distribution.

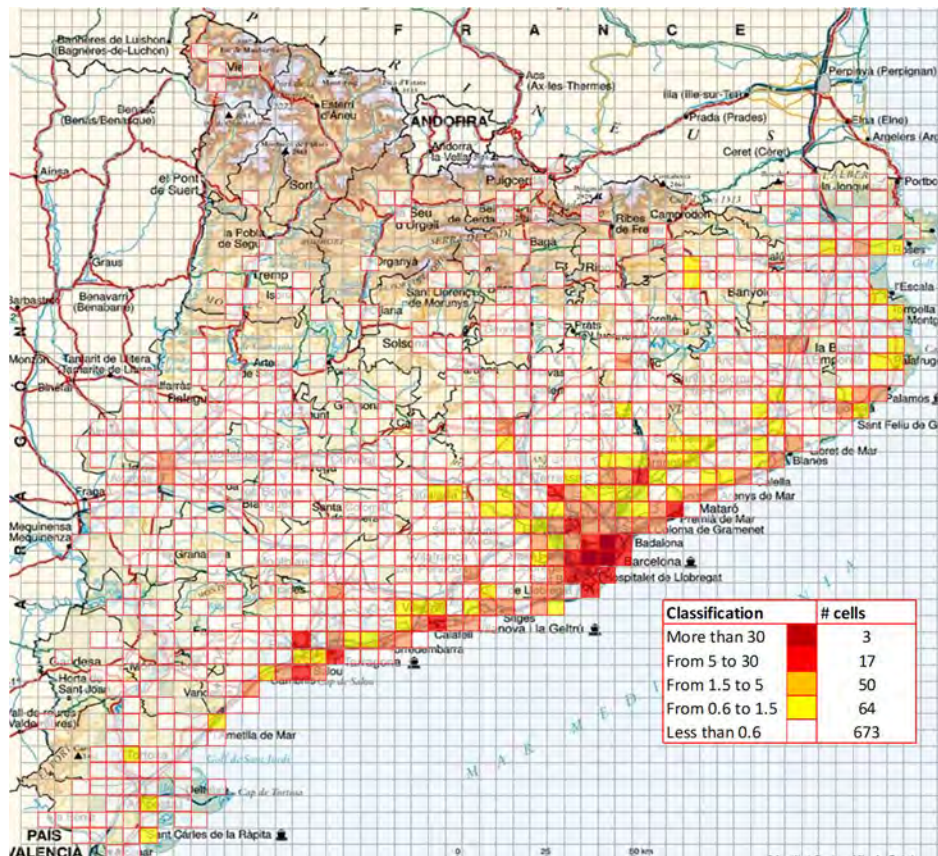
This result is true independently of the time intervals by which the burglaries are counted.

Study of the non-randomness of the burglaries in cells of 5kmx5km and intervals of 1 week

The runs test was applied for time intervals of 1 week and for different levels of activation. The series had 103 values corresponding to the number of whole weeks in the years 2014 and 2015. Some examples of the results are visualised in Appendix 7.6, where the levels of runs are considered up to a threshold of >8 burglaries. With a significance level of less than 0.1, it can be seen how, in some cells, random runs from a certain level to a level of > 8 can be observed, while others only have significance levels of less than 0.1 for some specific, normally consecutive levels. In a few cells, the significance level at which the non-random runs appear is slightly higher, between 0.1 and 0.2, and it is likely that point activations (runs with a single 1) are more frequent in these than in the rest.

The results obtained based on all the 5 km x 5 km cells and time intervals of a week indicate that the cells that are more likely to present non-random runs are those with averages between 1.5 and 8, for levels of >0, >1, >2, >3 and >4. With a significance level lower than 0.2, 82% are non-random; with a significance level of 0.15, 73.1% are non-random; and with a significance level of 0.1, 67.2% are non-random.

Non-randomness is observed in 67.2% of the cells with averages of between 0.6 and 1.5 for runs with levels between >0, >1 and >2, with a significance level of less than 0.2. With a significance level of less than 0.1, this is the case in 45.3% of the cells. Almost half (47.22%) the cells with an average of events between 0.3 and 0.6 have non-random runs for the levels >0 and >1, with a significance level of 0.2. With a significance level of 0.1, this is the case with 38.9% of the cells. Last, the cells that have an average of events below 0.3 usually have too few burglaries to be able to study the non-randomness. Nonetheless, 31.2% have non-random runs for levels >0 and >1, with a significance level of less than 0.2. With a significance level of less than 0.1, this is the case with 22.7% of the cells. In the cases where the average number of burglaries a week is less than 0.1, the runs test is too sensitive to small changes and is therefore not very reliable.



Source: Author's own elaboration on a map from the Cartographic Institute of Catalonia – ICC

Figure 3.- Hotcells of weekly residential burglaries with forced entry in Catalonia (2014/2015)

To these results must be added the fact that in 100% of the cases the value of the statistic of the runs test when there is non-randomness, is negative; in other words, the non-randomness is the result of greater grouping of the data with respect to what would be expected in the random case.

Thus, it can be said that, in general, activation levels that generate non-random runs by grouping the burglaries can be observed. These activation levels are usually near to the average number of burglaries and are normally in the interval $(\mu - \sigma, \mu + \sigma)$.

These results seem to confirm the theory of near repeat victimisation but, as we have already pointed out, the Poisson distributions also generate waves than can appear to be non-random. To test this, random Poisson simulations were carried out, passing the runs test. Specifically, 100 series of 103 values with a theoretical average of 1 were simulated, and runs for the levels >0 , >1 and >2 were sought. The result of the 100 simulated series is that with a significance level below 0.2, 27% of the series are seen to have non-random waves in some of the three cotes (>0 , >1 or >2), whereas this figure is only 12% with a significance level of 0.1. Comparing this percentage with the 67.2% and the 45.3%, respectively, obtained in the cells with averages of between 0.6 and 1.5 burglaries, it is clear that for burglaries with forced entry the possibility of observing waves with a non-random appearance in the phenomena governed by the Poisson law is greater.

This result is even clearer and more differentiated when it is compared with the sign of the statistic of the runs test. In the case of the Poisson simulations, in only 50% is this negative, compared with 100% in the case of the runs of burglaries (in both cases, considering only the cases that are non-random). However, despite this clear result in favour of non-randomness by grouping the runs of burglaries, a certain proportion of ‘false’ non-randomness generated by a random dynamic of the law of Poisson should always be kept in mind. With the aim of isolating these two effects - the false ‘non-random’ and the ‘non-random’ caused by the phenomenon of near repeat victimisation – it is proposed to work with the ‘non-random force’ matrix.

Optimal cells

There are 67 cells that fulfil the optimal cell criteria to forecast and prevent. In the two years in question there were 11,715 burglaries with forced entry in these cells (Table 2) – 25% of the total burglaries in Catalonia – and in all of them the average weekly burglaries was higher than 0.6 (Table 3).

Cells 5 km × 5 km	Cells	Burglaries 2014–15	Percentage
Optimal	67	11,715	25.5
Non-optimal ($\mu > 1.5$)	34	24,303	52.8
Non-optimal ($0.6 < \mu < 1.5$)	33	2,775	6.0
Non-optimal ($\mu < 0.6$)	673	7,225	15.7
TOTAL	807	46,018	100.0

Table 2.- Number of burglaries according to the type of cell

Table 3. Averages of the optimal cells.

Classification of the average	Optimal cells	Min. (μ)	Max. (μ)
Average >1.5	37	1.50	3.38
Average between 0.6 and 1.5	30	0.62	1.43
General total	67		

Table 3.- Averages of the optimal cells

The weekly averages of these cells are between 0.62 and 3.38, and in all of them non-random activation runs are observed for different levels. The cells with an average between 0.6 and 1.5 have non-random activation levels between >0 , >1 and/or >2 , while those with an average higher than 1.5 have non-random activation levels between >0 and >4 .

These 67 cells on the map of Catalonia are represented in Figure 4 together with the non-optimal cells with an average higher than 0.6. The number of burglaries depending on the type of cell are shown in Table 2.

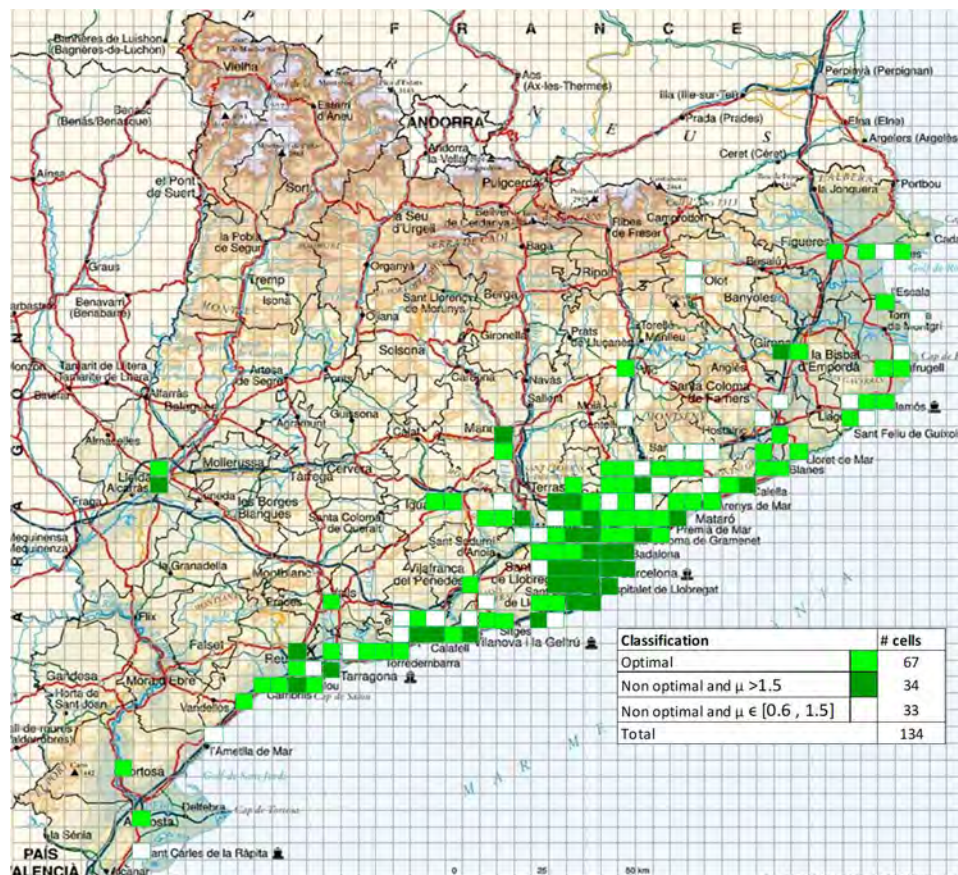
Most of the 34 cells that do not have an optimal configuration and have an average higher than 1.5 (25) have a $CV < 0.7$, despite often showing non-random behaviour. 24,303 burglaries were localised in these 34 cells, representing 52.8% of the burglaries occurring between 2014 and 2015.

Most of the 33 non-optimal cells for forecasting with averages between 0.6 and 1.5 are mostly so because they have a $CV > 1.3$, as the average is quite a lot lower than the standard deviation.

The other 673 cells are non-optimal and have averages of less than 0.6. In these cells in the two years under consideration, there were 7,225 burglaries, representing 15.7% of the total.

The cells with wave dynamics are usually the same ones as have more static risk (hotcells). Hence, those with more non-randomness are the ones with the highest weekly average, between 1.5 and 8.

Up to this point, the system proposed in regular cells has ignored the geographic features of the subjacent territory which, as has been pointed out, is usually heterogenous and can condition both forecast and prevention.



Source: Author's own elaboration on a map of the Cartographic Institute of Catalonia – ICC

Figure 4.- Representation of the optimal and the non-optimal cells

The configuration of time intervals in weeks and areas of 5 km by 5 km could initially be considered excessively large to be able to carry out effective preventative police activity. Nonetheless, prevention can focus on very specific spaces of each of these cells, hugely reducing the area of action.

The appendixes 7.7, 7.8 and 7.9 represent some of the cells according to the classification mentioned above, on a scale of 1:50.000: optimal cells, non-optimal cells with $\mu > 1.5$ and non-optimal cells with $0.6 < \mu < 1.5$. The grey zones on the maps represent urban centres, which in Catalonia are usually comprised of blocks of flats of different heights, while the houses are represented by black square-shaped dots.

The heterogeneity of the territory captured inside the optimal cells can be observed (Appendix). Population clusters can be seen, as well as extensive residential areas (concentrations of houses) and areas with scattered houses, surrounded by varying environments (coast, mountain, plane, woods, industrial areas, and so on), often crossed by relatively important roads (dual carriageway of motorway). Some of the cells have up to three population clusters and others only one. There are also some that contain only neighbourhoods, albeit relatively important ones. Some of these cells have a tourism profile, while others have residential or urban profiles.

However, the most important detail in each of these cells is that a large part of their interior contains no houses, or contains some houses but with a very low density (scattered).

A high density of residences, which subsequently indicates a high probability of burglaries, can generally only be observed in no more than 5 or 6 of the 25 subcells of 1km². Hence, despite a relatively large area in terms of forecasting, preventative action could in fact focus on 25% of the territory of the cell.

The non-optimal cells with $\mu > 1.5$ also show a high heterogeneity (Appendix), but they are usually more populated and tend to contain large towns. In these cases, the territory at highest risk or with the highest concentration of burglaries can cover 50% of the cell or more.

Last, the non-optimal cells with $0.6 < \mu < 1.5$ are likewise heterogenous environments (Appendix), and they are also similar to the optimal cells but, in this case, they are less densely populated with the population clusters covering less than 25% of the cell or consisting of residential estates or scattered houses.

4 Discussion

Through a segmentation into cells of 5 km by 5 km and time intervals of a week, it has been shown that a large part of the series of burglaries with forced entry in residences in Catalonia between 2014 and 2015 follow patterns compatible with near repeat victimisation. This is because the data are not random either in space (due to the heterogeneity of the territory) or in time (due to the formation of waves).

This result could be surprising. Although the study changes the usual micro level configuration in which the theory of near repeat victimisation is explained, the same spatiotemporal patterns that it describes remain. The phenomenon of repetition, which could be masked by the overlapping waves of burglaries of different groups of authors, finally emerge clearly. In the most attractive cells, waves are observed and are not random, but show stable repeat patterns as expected in the micro case.

It could be deduced that the perpetrators of these waves are the same or that they are mostly the same, despite some occasional overlaps. Otherwise, some coordination between different rival groups would be necessary in order to victimize the same spaces and during the same time periods, which seems quite unlikely.

The fact that burglaries intensity in the optimal cells is relatively low, often with activation levels of 1 or 2 burglaries and executions that generally do not exceed 4 or 5 burglaries per week (less than one burglary per day) reinforces the idea that perpetrators in each of the detected waves is largely the same. Scattered burglaries, from different criminal groups, are

difficult to match stably in space and time to generate non-random repetition patterns. Therefore, it seems clear that the spatiotemporal patterns observed in home burglaries in Catalonia, from a macro perspective, could be explained by the phenomenon of near repeat victimisation.

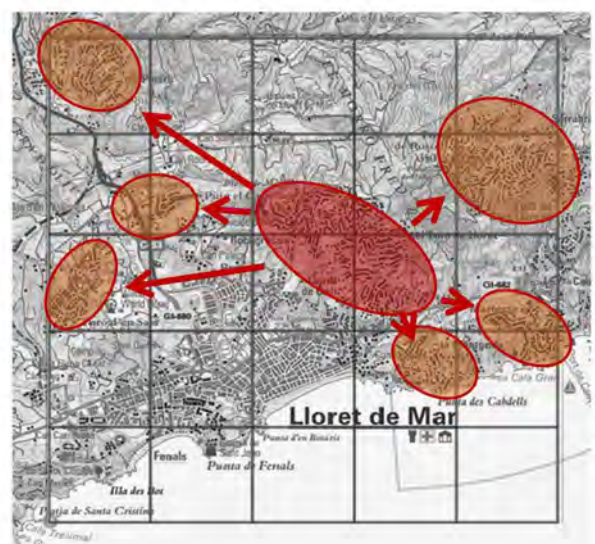
The macro scale option was not the first considered in the study. At the beginning of the research, well-known and available tools were used to check the phenomenon of near repeat victimisation, including “Near repeat calculator” (Jerry Ratcliffe 2009) and other free software. As an example, the space time bandwidths of home burglaries in the city of Barcelona were studied. The result was very different depending on the district considered, motivated, again, by the environment heterogeneity. In an aerial view of the city, it’s possible to observe these constructive changes in the neighbourhoods and districts. This problem, initially only detected in Barcelona, was seen to be a common factor in virtually all inhabited areas of Catalonia.

One of the main limitations of Knox's test-based software is the assumption, usually not explicit, of the space homogeneity. This assumption can be correct when considering cells of up to a few hundred meters, but not more than 1km. At least, not in Catalonia, where within the same 5kmx5km cell there are very different environments of housing and uninhabited areas (Appendix). In these heterogeneous spaces the replicas of burglaries take place, not in a concentric way and at a certain distance from the initial one, but taking various forms according to the territorial disposition of similar inhabited areas. When the initial area is relatively small but, not far away, similar patterns are observed within the same or adjacent cells, the hazard also appears to be able to make jumps to these unconnected but close areas, both in space and similarity (figure 5).

Week number n : Initial burglaries in a residential and touristic area.



Week number $n+1$: Risk irradiation towards connected and unconnected but similar areas



Source: Author's own elaboration on a map of the Cartographic Institute of Catalonia – ICC

Figure 5.- Example of risk irradiation towards unconnected but similar areas in a 5kmx5km cell

These new repeat patterns could be some of the particularities of the phenomenon of near repeat victimization in extensive and heterogeneous spaces.

It was also observed that the average intensity of burglaries in the space-time was relevant, as was the standard deviation, so it was decided to calculate some matrices that would show the viability of each cell or space where a forecast was to be made, the states that generate waves and their non-random 'force'.

With the optimal cell criteria depending on the CV and the non-randomness, the proposal would be to accept the 67 optimal cells to make forecasts with the space-time segmentation considered. With respect to the non-optimal cells with $\mu > 1.5$, a different segmentation that would make them optimal would be required. As it has been shown that these generally behave non-randomly, one option would be to subdivide them into smaller cells or to reduce the time intervals considered.

With respect to the non-optimal cells with averages between 0.6 and 1.5, the opposite is required: increasing the size of the cells and the time intervals. Regarding the first option, some of these cells are surrounded by others with a configuration that is already optimal, or by cells with a much higher average of burglaries, and the proposal is to subdivide these. Therefore, the option of increasing the size should be studied on an individual basis, and the shape obtained may not be square, but could be a juxtaposition of different square cells and subcells. Another option would be to increase the time intervals or to simply discard them for making forecasts.

Of the remaining cells, those with a very low intensity of burglaries that also more often display random behaviour, are not easily mouldable to make a useful forecast because either the territory or the time intervals considered would need to be much larger and this would make the preventative policing task much more complicated. It is therefore suggested that they are discarded for forecasting.

Hence, in a first round, with a strictly defined segmentation and depending on the percentage of events observed in these different sets of cells, a model is achieved that is able to viably forecast and prevent one in four residential burglaries with forced entry. In a second round, acting in the way described in the non-optimal cells with averages between 0.6 and 1.5, a configuration to forecast 2 out of 3 burglaries could be achieved. Hence, overall, 3 out of 4 of the burglaries with forced entry in Catalonia could potentially be forecast.

Thus, with this system of relatively extensive regular cells, a first forecast algorithm can now be designed (Appendix).

This is, therefore, a new option for criminal prediction that is added to the many proposals that have emerged in recent years. But perhaps it is the first to propose working from a macro level perspective, which may enrich the already intense debate on the practical application of predictive policing.

From a micro level perspective, predictions are much more limited in space, which, a priori, would be an advantage when planning preventative actions. However, the number of microcells included in these predictions can be much higher than in the macro case, and with a lower expected number of burglaries in each, given that the repetitions that fall outside the narrow bandwidth considered will not be counted. This means dispersing more preventative actions

and the need for more police resources to try to make these actions relatively intense and effective in each micro risk area.

In the macro level case, 5kmx5km cells include 25 1kmx1km cells, and can be complicated to make intense preventative actions with a deterrent capacity due to the big size of the risk area. But it has been demonstrated that the heterogeneity helps to reduce the space in the cell where there are more likely to be burglaries, which is where the preventative action should be focused. And at the same time, this extensive prevention can have an effect beyond the immediate environment of the first burglaries and be effective even for other future prior burglaries nearby or for those that are not replicas of the previous week.

Hence, in this type of heterogeneous and large-scale environments, the size of the cell by itself is not an indicator of the precision of the forecast or the effectiveness of the prevention.

In fact, the heterogeneity can be studied in greater depth and used to further delimit the area most susceptible to burglaries. According to the theory of near repeat victimisation, criminals repeat their crimes in similar settings, so if they have burgled in a residential area they will probably burgle again in a residential area, and if they have burgled in a block of flat then they will probably burgle again in blocks of flats with similar characteristics, and so on for neighbourhoods of detached houses, etc. This type of information can be very valuable to plan a good preventative strategy as it will facilitate fine-tuning the forecast, thus increasing the accuracy of police action in a relatively extensive risk environment.

Another aspect suggested by the territorial visualisation of cells and related to the reflections made in the previous paragraph, is the limits of the cells. Obviously, if a residential zone has been detected where there have been burglaries and repetitions have been forecast, prevention should not focus strictly on the space in the cell if it has continuity outside its limits. The practical information for the preventative strategy must indicate that the residential estate, neighbourhood, type of neighbourhood in the area, and so on, is highly likely to be the target of burglaries, with a certain flexibility in the interpretation of the limits of the cell.

These examples show that the heterogeneity and the classification of the territory into homogenous areas will determine both forecast and prevention, which is generally assisted by a thorough knowledge of the territory. To this effect, some academic researchers are studying how to incorporate this information into the mathematical and statistical models to improve forecasting (Smith et al. 2010).

The coincidence of both static and dynamic risk in the same area and its interaction explains the crime concentration. This is the Flag-Boost-Interaction or FBI theory (Farrell and Pease 2017) . Police preventative actions must be adapted to the concurrence of both risks in the cells. On the one hand, the environment must be worked from the point of view of the static hotspots to attempt to devise preventative strategies that reduce the risk, looking to find the most influential factors that will enable them to reduce the crime opportunity (type of residences, the buildings' security measures, proximity to fast roads, etc), for example using Risk Terrain Modelling (Caplan and Kennedy 2010). The long-term hotspots or hotcells contribute useful information for structural type prevention, as they detect the most targeted places over time. On the other hand, the study of the dynamics of burglary hotspots such as the one proposed here must help to decide on sporadic preventative actions to stop waves of burglaries being

generated. The two strategies must complement one another, and they must involve different police preventative actions normally in the same sensitive cells or environments.

Thus, the theory of near repeat victimisation provides a valid general pattern that must be adapted to each specific territory, as there is always a random component and a non-random component in terms of both space and time, which can only be understood locally. In general, the non-random tendency of the burglaries is observed in relatively large areas, while in small areas or in a specific residence the occurrence of burglaries is totally random. An explanation for this could be that criminals decide to burgle depending on 2 basic decisions: which area to burgle and when, and which specific residence to burgle within the area chosen on a specific day and time. The non-random component is associated with the first decision, which is normally maintained over a time interval of some 2 or 3 weeks. The chosen area can normally be delimited with homogeneity criteria for the territory. The criminals must study the area before committing the first burglary, and if it is successful they will almost certainly do it again with high expectations of similar success. The random spatial component is due to not knowing which specific residence in the area will be the next target, and the random temporal component is the impossibility of knowing on which day and at which specific moment the next robbery will occur within this wave.

With this approach, concepts like similarity and proximity are important, both in terms of space and time. Some generalisations of this theory could contemplate the hypothesis that criminals can select more than one unconnected area to commit the burglaries 'at the same time' (on the same day or in the same weeks). These different chosen areas are likely to have similar characteristics, so geographical proximity could be extended to proximity via road. For example, if two of the chosen areas are connected by a fast road and the travelling time between them is therefore relatively short, they could be considered to be 'near' each other from the perspective of this theory.

If this hypothesis was proven to be correct, a classification as detailed as possible of the homogenous areas of a territory could be made to calculate the minimum time it takes to go (via a land route) from one of these areas to another, and then to construct a matrix of proximities that summarise this information, combining geographical distance with environmental similarity. Hence, when there is a burglary in a specific area, the theory of 'widespread' near repeat victimisation would indicate that there is an increased likelihood of burglaries in that area and in the 'nearby' areas according to this new distance.

From this perspective, it would be interesting to study the topological aspects of the theory of near repeat victimisation. Some studies refer to this need to review the types of spatiotemporal patterns applied in the analysis of criminology, among them those to do with the relationship between the objects represented (points, lines, trading estates) (Leong and Sung 2015)

Apart from classifying the territory into homogenous areas, more detailed information about each robbery could also be a great help to improve forecasts. This would include the modus operandi, the type of booty stolen, how the residence was broken into, etc. The police currently use this information to investigate the crime, but it could also be useful to make a more accurate forecast and to guide prevention. Nonetheless, if the critical area is homogenous all the residences are likely to continue being a possible target due to similarity (the same home

security systems, similar expected booty, and so on). Hence, the degree to which forecasts would be improved is unclear. Furthermore, it is likely that this information is somehow already included in the information that enables us to classify the area as homogenous.

As already pointed out, the data considered in this study regarding burglaries only takes space and time into consideration. While space is reliable data with precise coordinates, time can be more imprecise. Forced entry burglaries generally occur when the owners are not at home, so it is often difficult to be accurate about the exact time of day it happened beyond the day and a window of time (morning, afternoon or night). And when the burglary occurs in an unoccupied second residence, the time window is even more expansive and can be as much as several weeks. In these cases, the data about the burglary is only approximate. In 2014 and 2015, one in four burglaries happened in second residences, which are usually only found in specific, delimited areas with a tourism profile.

This temporal limitation will condition a possible forecast which, in the best of cases, cannot go beyond the day the burglary may occur. For this study, it was decided to use intervals of a week.

Another limitation of this type of study is the black figure. The police only know the reported events, and there is little information about the black figure. Nonetheless, we can assume that the unreported burglaries are the less serious ones. Some of these may be attempted burglaries or ones where the damage or the value of the goods stolen was low. It can also be assumed that, in general, they have a spatiotemporal distribution that runs parallel to the known facts, with a similar static and dynamic risk.

These limitations were considered during the course of this study. One reported burglary more or less can cause a variation of 0, 1 or 2 more runs, depending on the level, and can modify the results of some statistical contrasts. This is why to confirm the results, the relatively high significance levels of 0.1 and 0.2 and other statistical parameters such as the sign of the value of the statistic of the runs test, were considered.

Regarding the significance level of the runs test, if what is contrasted is that the data are non-random for larger grouping, then the region of acceptance is only the left tail of the $N(0;1)$ where there will be all the cases in which the number of runs is less than expected, which will imply that the test statistic is negative, as happened with 100% of the cases where non-randomness was observed (except for the cases already discarded due to a low intensity of burglaries). Hence, the significance level of the results obtained is, in fact, half of that shown.

Last, despite these limitations, the data were shown to be sufficiently robust as to demonstrate the general dynamic of the burglaries in space-time on a large scale, allowing for some first exploratory studies about the patterns of near repeat victimisation in Catalonia.

5 Conclusions

This study showed that from a macro perspective the burglaries with forced entry in Catalonia follow a spatiotemporal pattern that is compatible with repeat victimisation. The proposed approach enables us to explore useful predictive models for preventing waves of burglaries. The methodology used is generic and can be applied to any territory, at least as a starting point and before adopting more complex predictive models. On applying this to Catalonia, it is shown that

a large number of burglaries – up to a maximum of 3 out of 4 - can potentially be forecast and prevented, from the point of view of the viability of both forecasting and preventative actions. The information about the dynamics of the hotspots justifies the concentration of potential resources in specific places and times.

Nonetheless, to obtain the best results, preventative actions must address both the static and the dynamic risk with different preventative strategies and actions. It has been observed that cells with the highest static risk, with a relatively high and stable average burglaries intensity, usually coincide with the optimal ones to predict. That is, those that generate a dynamic pattern in the form of non-random waves, which is in line with the FBI theory (Farrell and Pease 2017) and explanations of crime concentration.

The study was not a micro-analysis and was not carried out on a small scale, but in this sense, there are limitations to do with the classification of the homogenous areas. Without discarding the possibility of forecasting burglaries in Catalonia much more accurately in both space and time using complex mathematical models, it is likely that better results would be obtained from combining macro statistical information with micro territorial and police information, and that this approach would significantly reduce the number of residential burglaries. Here, we'd like to mention the last sentence of the article "Crime concentrations: Hot Dots, Hot Spots and Hot Flushes" (Ignatans and Pease 2018). They say: "Predictive patrolling is possible, and simple locally grown software and expertise is generally to be preferred to commercial products because of the in house learning and skills enhancement which it brings".

Heterogeneity inside large cells can help to make predictions more precise and accurate, as patterns of near repeat victimisation are likely to be in similar areas to those of initial burglaries. In practice, this limits the risk area to one or a few of the 1km² sub-cells.

The experience derived from this study carried out in Catalonia may be common to that of other territorial law enforcement agencies in similar settings. This may be the case in a large part of the southern regions of Europe. So global prediction and prevention strategies are also possible in there, using the large-scale or macro level option. The heterogeneity of the environment, towns, cities, neighbourhoods, and homes seems to be key to finding the optimal space-time configuration to achieve a useful prediction for the design of preventative strategies.

In any case, police forces are generally recommended to carry out specific, local or regional studies of their territory and its particular problems before adopting any strategic preventative strategy guided by predictive models

The criteria for adopting one model or another should not be so much the precision of predictions, in the exclusive sense of a very narrow space time bandwidth, but the preventative options that offers to police. Models that generate a myriad of micro-cells where there may be some very few repeats from recent days burglaries, can complicate the police effort to prevent all them. Otherwise, this study would propose what we can call the surfer's strategy, this is, to act on the higher wave crests, even if this means ensuring larger spaces. The reward will probably be the great descent down the wave, because these will probably concentrate a big part of the total burglaries of the territory during the week and, as we've said, during the year.

Following this macro focus of the theory of repeat victimisation, some future lines of research are suggested that revise the concept of 'proximity' and formulate the theory in heterogeneous environments, so that replicas in unconnected but similar areas can be considered (figure 5). When homogeneous, small-sized areas are distributed over relatively large spaces, it seems likely that repeats will not fit in a single area and simultaneous waves may be reproduced in several. In determining which areas may be affected simultaneously, the similarity of environment and the geographical distance and displacement between them will play an important role. This could lead to new risk irradiation patterns that need to be studied in depth to improve both predictions and preventative strategies.

6 References

1. Farrell G., Pease K. Once bitten, twice bitten: Repeat victimisation and its implications for crime prevention. *Police Res Group, Crime Prev Unit Pap* 46 [Internet]. 1993;(46):38. Available from: <http://hdl.handle.net/2134/2149%5Cnhttp://core.ac.uk/download/pdf/67259.pdf>
2. Farrell G., Pease K, Phillips C. Like taking candy : why does repeat victimization occur ? *Loughbrgh Univ Institutional Repos.* 1995;35(3).
3. Pease K. Repeat Victimisation: Taking Stock [Internet]. *Crime Detection and Prevention Series.* 1998. 48 p. Available from: http://www.popcenter.org/problems/domestic_violence/PDFs/Pease_1998.pdf
4. Pease K, Farrell G. Repeat Victimization. *Crime Prev Stud* [Internet]. 2001;(February 2015). Available from: <http://link.springer.com/10.1007/978-1-4614-5690-2>
5. Townsley M. Infectious Burglaries. A Test of the Near Repeat Hypothesis. *Br J Criminol* [Internet]. 2003 Jun 1 [cited 2017 May 19];43(3):615–33. Available from: <https://academic.oup.com/bjc/article-lookup/doi/10.1093/bjc/43.3.615>
6. Townsley M, Homel R, Chaseling J. Repeat burglary victimisation: Spatial and temporal patterns. *Aust New Zeal J Criminol.* 2000;33(1):37–63.
7. Gottfredson MR, Hirschi T. *A general theory of crime.* Stanford University Press; 1990. 297 p.
8. Cohen LE, Felson M. Social Change and Crime Rate Trends: A Routine Activity Approach. *Am Sociol Rev.* 1979;44(4):588.
9. Clarke R V., Cornish DB. Modeling Offenders’ Decisions: A Framework for Research and Policy [Internet]. Vol. 6, *Crime and Justice.* The University of Chicago Press; 1986 [cited 2018 Aug 22]. p. 147–85. Available from: <https://www.jstor.org/stable/1147498>
10. Felson M, Clarke R V. Opportunity Makes the Thief Practical theory for crime prevention. *POLICE Res Ser Pap* 98, Home Off [Internet]. 1998 [cited 2018 Mar 17]; Available from: <http://www.popcenter.org/library/reading/pdfs/thief.pdf>
11. Bernasco W. FORAGING STRATEGIES OF HOMO CRIMINALIS: LESSONS FROM BEHAVIORAL ECOLOGY. *CRIME PATTERNS Anal* [Internet]. 2009 [cited 2018 Mar 17];2(5). Available from: <http://wimbernasco.nl/Manuscripts/English/Bernasco2009a.pdf>
12. Tseloni A, Pease K. Repeat personal victimization: “Boosts” or “Flags”? *Br J Criminol.* 2003;43(1):196–212.
13. Bowers KJ, Johnson SD. Who Commits Near Repeats? A Test of the Boost Explanation. *West Crim Rev* [Internet]. 2004 [cited 2018 Mar 17];5(3):12–24. Available from: http://westerncriminology.org/documents/WCR/v05n3/article_pds/bowers.pdf

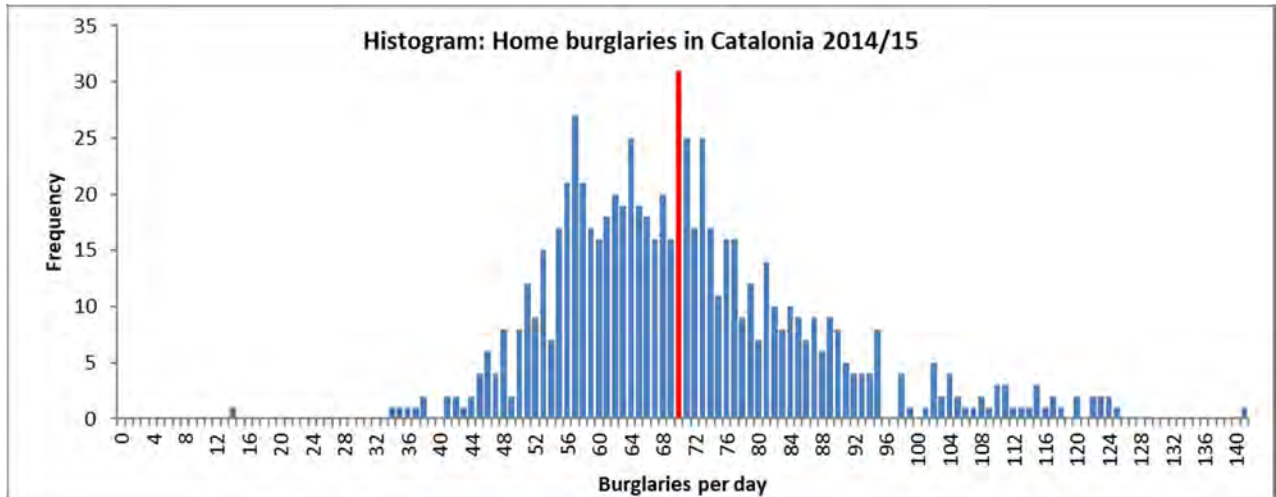
14. Ornstein JT, Hammond RA. The Burglary Boost: A Note on Detecting Contagion Using the Knox Test. *J Quant Criminol* [Internet]. 2017 Mar 14 [cited 2017 May 19];33(1):65–75. Available from: <http://link.springer.com/10.1007/s10940-016-9281-1>
15. Johnson SD, Bernasco W, Bowers KJ, Elffers H. Space-Time Patterns of Risk : A Cross National Assessment of Residential Burglary Victimization Space-Time Patterns of Risk : A Cross National Assessment of Residential Burglary Victimization. 2006;(0):1–47.
16. Kikuchi G, Amemiya M, Saito T. A Spatio-Temporal Analysis of Near Repeat Victimization in Japan. *Eighth Natl Crime ...* [Internet]. 2010; Available from: <http://www.ucl.ac.uk/scs/events/mapping-conf/conf-2010/downloads-2010/4b-kikutchi.pdf>
17. Vijaya Kumar M. Spatial-Temporal Analysis of Residential Burglary Repeat Victimization: Case Study of Chennai City Promoters Apartments, INDIA. *IJCTA* [Internet]. 2011 [cited 2018 Mar 17]; Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.208.7766&rep=rep1&type=pdf>
18. Wang Z, Liu X. Analysis of Burglary Hot Spots and Near-Repeat Victimization in a Large Chinese City. *ISPRS Int J Geo-Information* [Internet]. 2017;6(12):148. Available from: <http://www.mdpi.com/2220-9964/6/5/148>
19. Department of Criminal Justice | Near Repeat Calculator [Internet]. [cited 2018 Mar 20]. Available from: <http://www.cla.temple.edu/cj/center-for-security-and-crime-science/projects/nearrepeatcalculator/>
20. SHORT MB, D'ORSOGNA MR, PASOUR VB, TITA GE, BRANTINGHAM PJ, BERTOZZI AL, et al. A STATISTICAL MODEL OF CRIMINAL BEHAVIOR. *Math Model Methods Appl Sci* [Internet]. 2008 [cited 2017 May 19];18(supp01):1249–67. Available from: http://www.math.ucla.edu/~lchayes/lchayes_website_folder/new_publications_folder/crime_stat.pdf
21. Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE. Self-Exciting Point Process Modeling of Crime. *J Am Stat Assoc* [Internet]. 2011;106(493):100–8. Available from: <http://www.tandfonline.com/doi/abs/10.1198/jasa.2011.ap09546>
22. Predict Crime | Predictive Policing Software | PredPol [Internet]. [cited 2018 Mar 20]. Available from: <http://www.predpol.com/>
23. PRECOBS Ifmpt - Institut für musterbasierte Prognosetechnik [Internet]. [cited 2018 Mar 20]. Available from: <http://www.ifmpt.de/>
24. Chainey SP, Figueiredo B. Examining the extent of repeat and near repeat victimisation of domestic burglaries in Belo Horizonte , Brazil. *Crime Sci*. 2016;1–10.
25. Vozmediano Sanz L, San Juan Guillén C. *Criminología ambiental : ecología del delito y de la seguridad*. Editorial UOC; 2010.
26. Aguarales M, Pellicer M, Ripoll J. An age-structured population approach for the mathematical modeling of urban. 2017;

27. Andrew M, Llb WS, Re M, Barracks D. A Comparison of Hotspot Mapping for Crime Prediction. 2010;(01635 204301):2010.
28. Chainey AS. JDI Briefs : Predictive mapping for policing. Blog post. 2012;(April):1–5.
29. Gonzales AR, Schofield RB, Hart S V. Mapping Crime : Understanding Hot Spots. Natl Inst Justice [Internet]. 2005;79. Available from: <http://discovery.ucl.ac.uk/11291/1/11291.pdf%5Cnhttp://discovery.ucl.ac.uk/11291/>
30. ICGC - Vissir3 [Internet]. [cited 2018 Mar 22]. Available from: <http://www.icc.cat/vissir3/>
31. Smith LM, Keegan MS, Wittman T, Mohler GO, Bertozzi AL. Improving Density Estimation by Incorporating Spatial Information. EURASIP J Adv Signal Process [Internet]. 2010 Dec 27 [cited 2018 Mar 17];2010(1):265631. Available from: [https://asp-
eurasipjournals.springeropen.com/articles/10.1155/2010/265631](https://asp-eurasipjournals.springeropen.com/articles/10.1155/2010/265631)
32. Farrell G, Pease K. Preventing repeat and near repeat crime concentrations. In: Handbook of Crime Prevention and Community Safety. 2017.
33. Caplan JM, Kennedy LW. Risk Terrain Modeling Compendium for Crime Analysis. Rutgers Cent Public Secur. 2011;
34. Leong K, Sung A. A review of spatio-temporal pattern analysis approaches on crime analysis. Int E-Journal Crim Sci. 2015;9:1–33.
35. Ignatans D, Pease K. Crime concentrations: Hot dots, hotspots, and hot flushes. In: The Oxford Handbook of Environmental Criminology. 2018.

7 Appendix

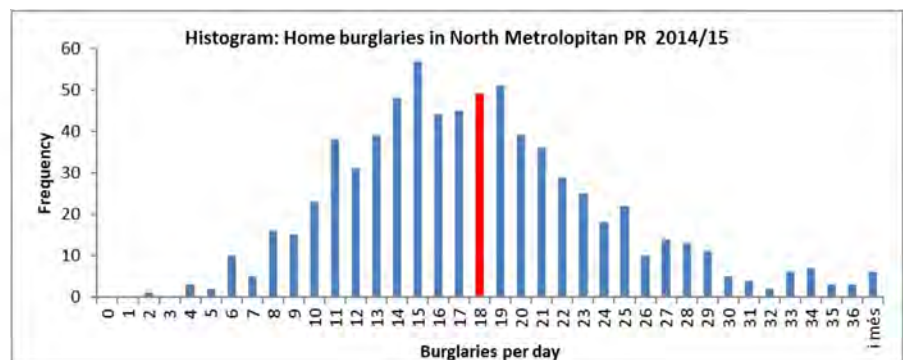
Average distribution (Overall average is marked in red)

Graphic 1: Histogram of daily frequency of burglaries in Catalonia



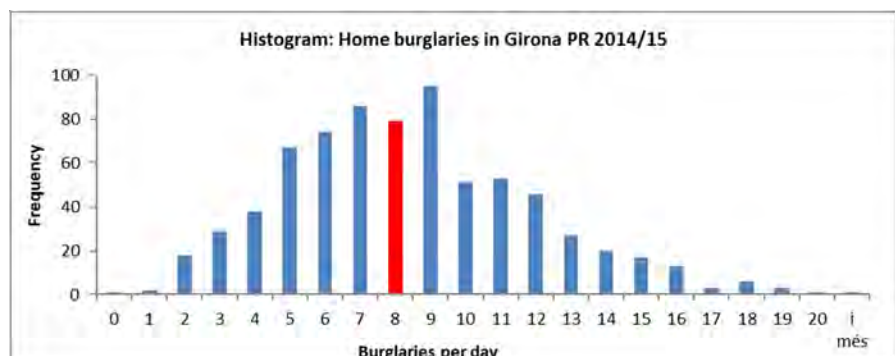
Graphic 2: Histogram of daily frequency of burglaries in North Metropolitan PR

North Metropolitan PR



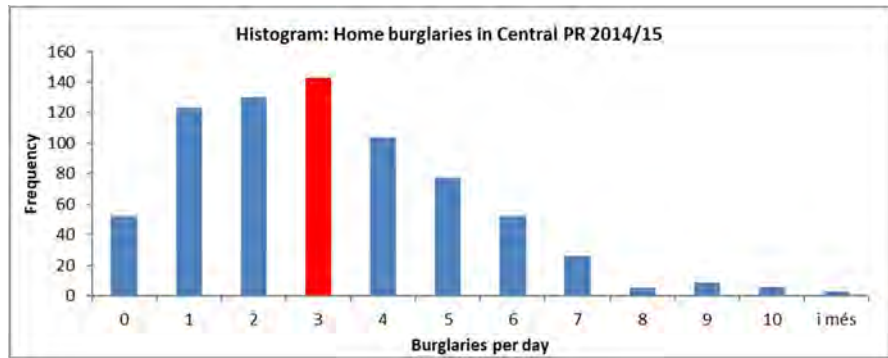
Graphic 3: Histogram of daily frequency of burglaries in Girona Police Region

Girona PR



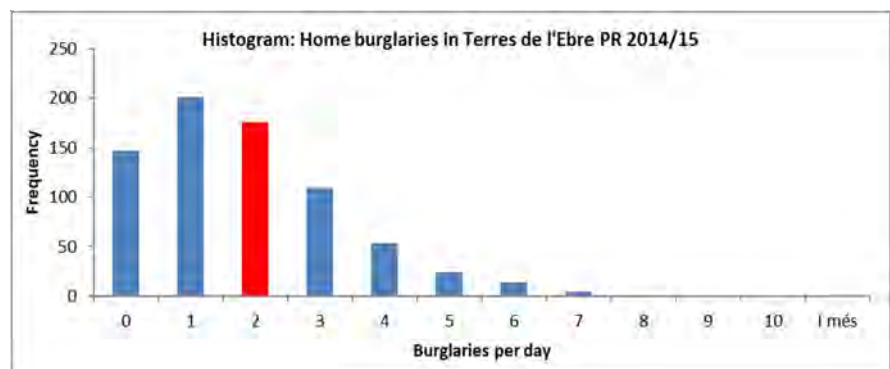
Graphic 4: Histogram of daily frequency of burglaries in Central Police Region

Central PR



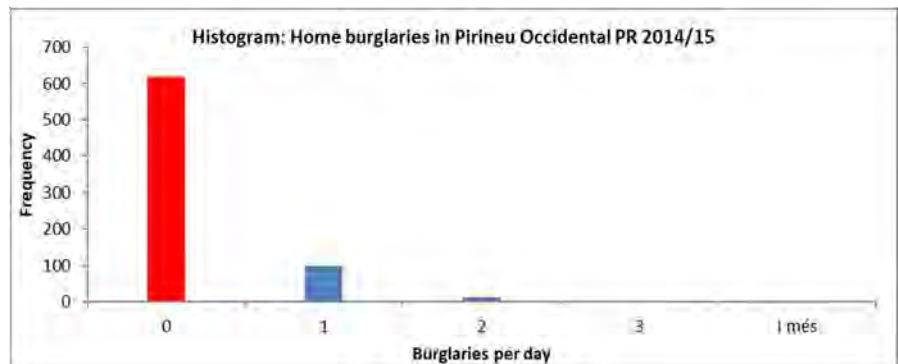
Graphic 5: Histogram of daily frequency of burglaries in Terres de l'Ebre PR

Regió Policial Terres de l'Ebre



Graphic 6: Histogram of daily frequency of burglaries in Pirineu Occidental PR

Regió Policial Pirineu Occidental



Graphic 7: Distribution of the weekly average of burglaries in the 5kmx5km cells

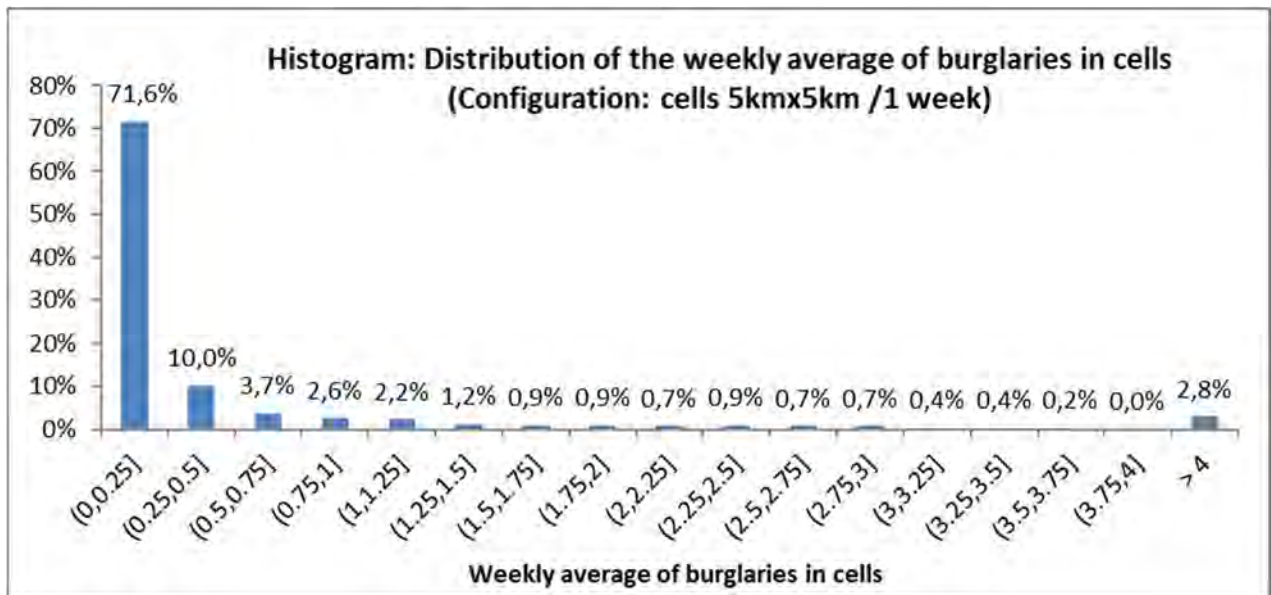


Table 2: Examples of the results obtained with the runs test in some of the study cells.

Runs test	Runs>0	Runs >1	Runs >2	Runs >3	Runs >4	Runs >5	Runs >6	Runs >7	Runs >8
Test value (median)	,50	,50	,50	,50	,50	,50	,50	,50	,50
Total cases	103	103	103	103	103	103	103	103	103
Cell S32	Runs test (μ, σ)=(7.2 , 2.9)								
Number of runs	3	3	7	19	29	29	33	35	32
Z	,141	,141	,335	-,855	-2,417	-3,311	-3,528	-3,289	-3,107
Sig. asimpt (bilateral)	,888	,888	,738	,392	,016	,001	,000	,001	,002
Cell R43	Runs test (μ, σ)=(3.6 , 2.5)								
Number of runs	15	33	35	30	33	27	21	13	7
Z	,771	-1,038	-2,915	-4,343	-2,759	-2,250	-1,217	-,834	,335
Sig. asimpt (bilateral)	,441	,299	,004	,000	,006	,024	,224	,405	,738
Cell U37	Runs test (μ, σ)=(3.3 , 2.2)								
Number of runs	9	24	48	46	37	27	19	11	5
Z	,440	-2,907	-,496	-,465	-,471	-,843	-,022	,548	,233
Sig. asimpt (bilateral)	,660	,004	,620	,642	,638	,399	,982	,584	,816
Cell T35	Runs test (μ, σ)=(2.9 , 2.1)								
Number of runs	15	39	49	39	28	16	12	9	7
Z	,771	1,422	-,593	-,658	-,913	-3,003	-1,636	,440	,335
Sig. asimpt (bilateral)	,441	,155	,553	,510	,361	,003	,102	,660	,738
Cell E50	Runs test (μ, σ)=(2.1 , 2.1)								
Number of runs	29	35	33	21	11	11	5	3	3
Z	-1,320	-3,360	-2,486	-2,239	-4,052	-2,438	-3,458	-5,705	-5,705
Sig. asimpt (bilateral)	,187	,001	,013	,025	,000	,015	,001	,000	,000
Cell L36	Runs test (μ, σ)=(1.5 , 1.3)								
Number of runs	34	48	25	13	9	1	1	1	1
Z	-,446	-,617	-2,590	-2,788	-1,687	No runs	No runs	No runs	No runs
Sig. asimpt (bilateral)	,655	,538	,010	,005	,092				
Cell R36	Runs test (μ, σ)=(1.9 , 1.8)								
Number of runs	42	52	39	27	19	11	3	3	3
Z	,093	,063	-1,371	-1,258	1,004	,548	,141	,141	,141
Sig. asimpt (bilateral)	,926	,950	,170	,209	,316	,584	,888	,888	,888
Cell N49	Runs test (μ, σ)=(2.2 , 2.0)								
Number of runs	37	49	43	32	22	14	8	6	3
Z	,440	-,290	-1,384	-,720	-,764	-1,232	-,962	-1,562	,141
Sig. asimpt (bilateral)	,660	,772	,166	,472	,445	,218	,336	,118	,888

Examples of optimal cells (scale 1: 50,000)

On the map there is a grid of 25 cells of 1 km² that delimits the represented 5km side cell.

Cell in North Metropolitan PR - Rubí



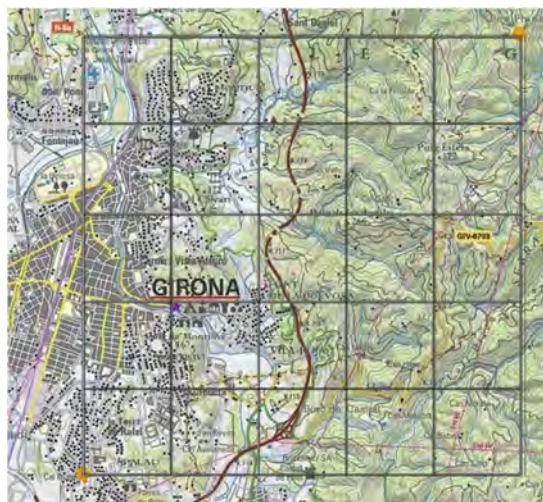
Source map: Cartographic and Geological Institute of Catalonia – ICC

Cell in Girona PR- Lloret de Mar



Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in Girona PR - Girona



Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in Girona PR - Blanes



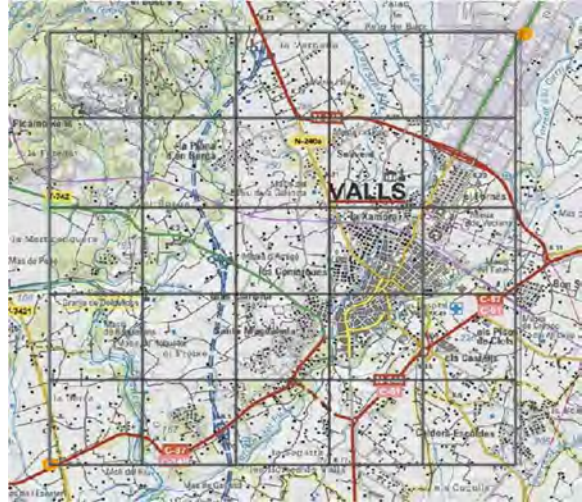
Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in Terres de l'Ebre PR – Amposta



Source map: Cartographic and Geological Institute of Catalonia – ICC

Cell in Camp de Tarragona PR-Valls



Source map: Cartographic and Geological Institute of Catalonia – ICC

Cell in Camp de Tarragona PR- Torredembarra



Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in North Metropolitan PR – Caldes de Montbui



Source map: Cartographic and Geological Institute of Catalonia - ICC

Examples of non-optimal cells with $\mu > 1.5$ (scale 1: 50,000)

Cell in North Metropolitan PR-Terrassa



Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in Camp de Tarragona PR - Tarragona



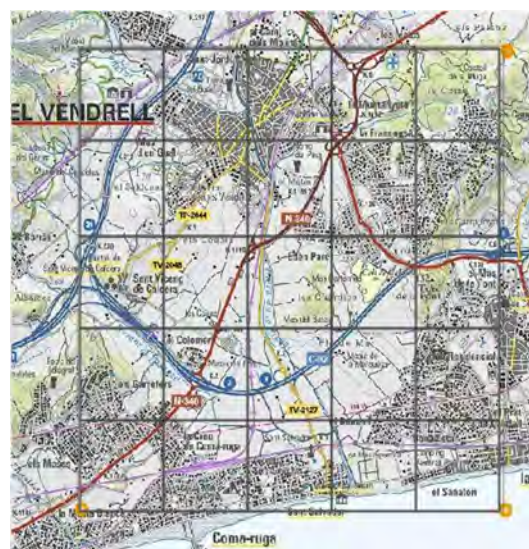
Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in North Metropolitan PR-Granollers



Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in Camp de Tarragona PR – El Vendrell



Source map: Cartographic and Geological Institute of Catalonia - ICC

Examples of non-optimal cells with $0.6 < \mu < 1.5$ (scale 1: 50,000)

Cell in Camp de Tarragona PR-Els Pallaresos



Source map: Cartographic and Geological Institute of Catalonia – ICC

Cell in North Metropolitan PR-Cardedeu



Source map: Cartographic and Geological Institute of Catalonia – ICC

Cell in Girona PR – Nord d’Aro



Source map: Cartographic and Geological Institute of Catalonia - ICC

Cell in South Metropolitan PR-Gavà Mar



Source map: Cartographic and Geological Institute of Catalonia - ICC

Predictive algorithm in extensive regular cells

First part: finding optimal cells

Step 1: Define the territory where you want to predict and study a cell size and temporality that, a priori, seems suitable for preventive purposes in most cases. This will be the starting configuration.

Step 2: Calculation of the CV for each cell and selection of those in which this parameter is in the range [0.7, 1.3]

Step 3: In the selected cells, search for activation levels that generate non-random waves according to the runs test. In general, these levels will be in the range $(\mu - \sigma, \mu + \sigma)$, and are usually between > 0 and > 4 .

Step 4: Calculation of wave and transition matrices and the corresponding force matrices.

Step 5: By observing the force matrices, determine the states that generate non-random runs with greater intensity, from which the cell will be activated in the subsequent time lag. These states will be called predictive states.

Step 6: Group the discarded cells:

- Group 1: Cells that can have an optimal configuration for smaller cell sizes or shorter time lapses.
- Group 2: Cells that can have an optimal configuration for bigger cell sizes or longer time lapses.

Step 7: Begin the study from step 1 for the new configurations that are considered appropriate.

Step 8: The first part of the algorithm will end when no optimal cells are found for any configuration from the point of view of the prediction utility, in the sense that it can guarantee an effective preventive police action.

Second part: choosing target cells or spaces to make dynamic prevention

Step 9: This part depends on each police force, their specific goals, their commitments, their work dynamics or their resources. A priori, the most efficient strategy would be to try to prevent the maximum number of burglaries with the minimum of resources, and this would involve concentrating actions in those cells or optimal spaces where more burglaries are expected. That is, select from top to bottom until there are available resources for that week.

Third part: Design of specific preventative actions and strategies for each of the target spaces or cells

Step 10: This part is also up to each police force. We've mentioned that it's important not only to make dynamic prevention but static or long-term actions also, and that environment heterogeneity can help to better predict inside the cells where to focus.

4.2 Article 2

Need to go further: using INLA to discover limits and chances of spatiotemporal prediction in heterogeneous environments.

Boqué P, Serra L, Saez M

Enviat a la revista: European Journal of Criminology.

Box 2 | Visió general de l'article 2

El que ja se sap sobre la temàtica

- Existeixen models predictius a petita escala, en entorns urbans, però no en entorns extensos, heterogenis i amb perfil residencial divers.
- Els grups criminals que roben en domicilis a Europa són organitzacions complexes i de composició internacional, que fan desplaçaments llargs per delinquir, amb un comportament itinerant i especialitzat.

El que aporta de nou l'estudi

- Es demostra que el patró de victimització quasi repetida restringit a la micro-escala és insuficient per modelar i predir els robatoris amb força a domicili en entorns heterogenis extensos.
- Es proposa el nou patró espaciotemporal de victimització "all-in-one", que amplia el concepte de repetició en la proximitat al de repetició en diverses zones, normalment semblants i que poden ser allunyades.

Need to go further: using INLA to discover limits and chances of spatiotemporal prediction in heterogeneous environments.

Pere Boqué^{1,2}, Marc Saez^{3,4}, Laura Serra^{3,4}

¹ Policia de la Generalitat – Mossos d'Esquadra

² PhD student, University of Girona, Spain

³ CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain.

⁴ Research Group on Statistics, Econometrics and Health (GRECS), University of Girona, Spain

Corresponding author: Laura Serra-Saurina

Research Group on Statistics, Econometrics and Health (GRECS) and CIBER Epidemiology and Public Health (CIBERESP)

University of Girona

Carrer de la Universitat de Girona 10, Campus de Montilivi

Girona, 17004, Spain.

Email: laura.serra@udg.edu

Abstract

Near repeat victimization patterns have enabled the construction of burglaries predictive models. Implemented in different countries, results obtained have not always been in line with initial expectations, to the point where their real effectiveness has been called into question. The possibilities of predicting crime remain under study to improve preventive policing strategies. The present study aims to discover the limitations and chances of burglaries predictive models based on the spatiotemporal pattern of risk spread in geographic proximity. A spatiotemporal log-Gaussian Cox process is contemplated to model the generic near repeat victimization scenario, adjusted using the Integrated Nested Laplace Approximation – INLA methodology. This approach is very suitable for studying and describing this type of phenomenon, obtaining models that perfectly fit the data history. However, prediction obtained with INLA is quite monotonous, of low variability and does not reproduce well the local and short-term dynamics of burglaries for predictive purposes. The conclusion is that predictive models cannot be restricted exclusively to distance decay risk, but they must be designed to detect other types of spatiotemporal patterns which, among other possibilities, leave open the possibility of correlating distant events and clusters. Although other studies have already pointed this problem, it is proposed to go one step further and clearly extend the near-repeat spatial patterns to achieve better prediction results.

Key words: spatiotemporal patterns, log-Gaussian Cox process, INLA, predictive policing, near repeat victimization

1 Introduction

It has been a decade since a group of researchers from Los Angeles surprised criminologists and police by proposing a crime prediction model (Mohler et al. 2011). The analogy between cluster generation systems for crimes and those for earthquake aftershocks led to a mathematical formulation capable of predicting the dynamics of crime clusters in space and time. The key to the analogy and to the predictive model is the criminological theory of repeat victimization (Farrell G. and Pease 1993), which in the case of burglaries establishes that each burglary has the effect of increasing the risk of new burglaries in the same area (M. Townsley 2003). Through the summation of these increments in risk, which are normally radial and adjusted with functions that are inversely proportional to distance, these mathematical models update the risk map for the following days or weeks.

Different police, business, and academic initiatives have pooled ideas to develop crime prediction programs, all with different focuses and methodological peculiarities. However, none have yet managed to gain general acceptance. There is continued controversy over methods, concepts, reliability, effectiveness, the biases of police data, ethical issues, and the problems of implementation (Weathington 2020; Yang 2020; Ratcliffe et al. 2020; Egbert and Krasmann 2020; Cinelli and Manrique 2019; Shapiro 2019; Kaufmann, Egbert, and Leese 2019; Meijer and Wessels 2019; Bennett Moses and Chan 2018; Gerstner 2018; Michael Townsley 2018; Seidensticker, Bode, and Stoffel 2018; EUCPN 2016). The entire predictive process, from its founding ideas and approaches to its methods and results, must continue to be constructed and reviewed. In this line, the present paper has two main objectives.

The first is to contrast the possibility to predict residential burglaries in heterogeneous environments, when considering the classic hypotheses of near repeat victimization, which are distance decay risk in geographic proximity and temporal autocorrelation. Large-scale patterns of burglaries waves observed in Catalonia (Boqué, Serra, and Saez 2020), despite being compatible with the theory of near repeat victimization, lead to questioning of different aspects in terms of burglars' decision model when choosing their targets. The hypothesis is that, in the general framework of optimal foraging theory (Krebs and Davies 1993; Bernasco 2009; Johnson 2014) environments with small, heterogeneous and scattered residential areas, would force criminals to move and burglar in several areas, within the same criminal initiative. The impact on crime patterns would be a lower concentration of near-repeats in space. To test it, a model based on risk decay in geographical proximity will be proposed and analysed with real data.

The second objective is to discuss the role that the two main spatiotemporal risk factors, the static and the dynamic risk, play in these predictive models, which often overlap and confuse. The static risk is related to the crime opportunity of the area and the flag hypothesis (Tseloni and Pease 2003), while the dynamic risk is related to burglars' rational decision model and the boost hypothesis (Bowers and Johnson 2004). More recently, a degree of interaction between the 2 hypotheses called the Flag and Boost Interaction (FBI) has been proposed (Farrell and Pease 2017), in addition to a theory that extends the boost hypothesis to a network approach (Lantz and Ruback 2017).

The question is how to model these factors. The static one is easier to model because of the long-term stability of the data, enabling us to represent and estimate them well using the calculation of means. However, the question is to what point can models based on this component be adjusted and to what extent do they model the dynamic factor of burglaries. To this effect, patterns of risk spread around recent burglaries have been verified (Johnson et

al. 2007), while at the same time the idea of whether we should continue underpinning only the hypothesis of geographical proximity to relate repeat burglaries is being questioned (Wang and Zhang 2020). We'll see that even the very idea of the initial burglary that triggers the whole process of repetitions may be questionable.

Proposals will be made concerning the questions outlined above, allowing us to deepen our understanding of burglaries spatiotemporal patterns. We'll attempt to identify the most suitable model to improve predictions in environments that are not strictly urban, are heterogeneous in nature, and have relatively low or irregular intensity. Results obtained will allow us to discuss some of the problems observed in previous models and to propose explanatory and methodological alternatives.

2 Data and methodology

2.1 Data

The territory: Catalonia

Catalonia is a country in Mediterranean Europe on the east coast of the Iberian Peninsula and occupying 5.5% of this land area. It is bordered by the Pyrenees and France to the north, the Mediterranean Sea to the east, and the rest of Spain to the west and south. Catalan population (7.7 million inhabitants) is concentrated in 30% of the territory, mainly in the coastal plains. Two thirds of the population live in the urban area of Barcelona. Apart from the capital and its conurbation, in general, cities and residential areas are small or medium-sized and with a diverse and heterogeneous urbanism.



Source: Cartographic and Geological Institute of Catalonia – ICGC

Figure 1: Typical heterogeneous urbanism of Catalan populations

The territory is divided into 9 police regions (Figure 1). Each region includes a set of Basic Police Areas (ABP), the territory's primary units, defined by geographical and policing criteria and usually encompassing several municipalities. These regions are geographically diverse; the territory of the Barcelona Metropolitan Police Region is almost entirely urban, while the Eastern Pyrenees Police Region is mountainous with few, generally small villages.



Source: Author's own elaboration

Figure 2.- Location of Police Regions (PR) of Catalonia

Crime data

Data used in this study are the reports of burglaries with forced entry made to the police in Catalonia from 2014 to 2019. Location in UTM coordinates and a window of time (generally less than a day) when the burglary is assumed to have occurred are recorded for each burglary. The type of residence burgled – a flat, a house, or a farmhouse - and whether it is a first or second residence are also specified.

The series of burglaries were segmented in space and time in different sized square cells (5km, 1km, 500m, 250m, 100m) and in weekly intervals. Profile characteristics used were type of building (house or flat), type of residence (first or second), time window of the burglary (morning, afternoon, or night), the day of the week (mid-week or weekend), and the monthly distribution.

These data were provided by the Government of Catalonia Police (Mossos d'Esquadra (PG-ME)), but they were not extracted using the same criteria as applied to the official data made public annually by the same police force ("Mossos d'Esquadra - Portal Dades Obertes," n.d.). In addition, some data that were not correctly geolocated had to be discarded. Therefore, while the data are approximately the same, the total number of burglaries stated in this paper does not exactly match the official data, although the results of the analysis are not affected.

2.2 Methods

Methods are divided into two types. First are a set of simple techniques to verify criminal data stability over time, and second is a methodology for modeling spatiotemporal data dynamics.

To test data stability, annual discrepancies are measured using standard deviation (SD) and the coefficient of variation (CV). Using the same criteria, the stability of the spatial distribution of the burglaries inside the 5km cells is determined by means of analyzing the distribution in each of the 25 1km² subcells.

Runs test will also be applied in different cells and spatial configurations to check the validity

and extent of the stability of the non-random wave pattern.

The second and main part of the methodology consists of applying a log-Gaussian Cox process (Diggle et al. 2013; Serra et al. 2014) spatiotemporal model to predict weekly burglaries. This model is adjusted using the Integrated Nested Laplace Approximation – INLA methodology (Lindgren and Rue 2015; Blangiardo et al. 2013), accessible at link <https://www.r-inla.org/home> (“R-INLA Project” n.d.). The validity of this continuous approach for this type of discrete models in which the study variable is a count, it’s been demonstrated (D. Simpson et al. 2015).

The model starts from a discrete-continuous approximation of the phenomenon of home burglaries. This approximation is constructed on the assumption that there is a latent risk of a burglary in space (and time); or in other words, that places close together in space (and time) have a similar risk (following the so-called Tobler’s law (Tobler 1970)). This way, any point in space (and time) can be assigned a risk by means of the next points interpolation.

Mathematically, it is formulated as a latent space-time process,

$$Z(s_{ij}; t_j) = Y(s_{ij}; t_j) + \epsilon(s_{ij}; t_j)$$

where $t \in \{t_1, \dots, t_T\} \in D_t$ are the times when the observations are made in the time domain D_t , of the number of burglaries $Z(s_{ij}; t_j)$, in a set of points in space $s \in \{s_1, \dots, s_M\} \in D_s$, where D_s is the spatial domain, $Y(s_{ij}; t_j)$ is the latent process, and $\epsilon(s_{ij}; t_j)$ is an error iid with a mean of 0 and a variance of σ_ϵ^2 .

The objective is to predict the value of the so-called latent field $Y(s_0; t_0)$ in the space-time localization (s_0, t_0) , depending on the known vector observations Z .

This latent field can be broken down into two parts, a static part and a dynamic part, $Y(s; t) = \mu(s; t) + \eta(s; t)$, $\forall (s; t) \in D_s \times D_t$, where $\mu(s; t)$ represents the average of the process, interpretable as the static and non-random component, and $\eta(s; t)$ represents the random or dynamic part of the process, with an average of 0, and with spatial and temporal dependency. The latent field in point $(s_0; t_0)$, $\hat{Y}(s_0; t_0)$, is estimated by means of searching for a prediction function that minimizes the average quadratic error, $E(Y(s_0; t_0) - \hat{Y}(s_0; t_0))^2$.

An adaptation of this approach for cases where the variable of interest is a count, as is the case with burglaries, is a hierarchic formulation with conditional probabilities and the introduction of a link function:

$$\begin{aligned} Z(s; t) | Y(s; t), \gamma &\sim \pi(Y(s; t), \gamma), \quad s \in D_s, \quad t \in D_t \\ g(Y(s; t)) &= x(s; t)' \beta + \eta(s; t), \quad s \in D_s, \quad t \in D_t \end{aligned}$$

Where π is a probability distribution, with a parameter of scale γ and a mean $Y(s; t)$. The function $g(\cdot)$ is the link function that converts the mean of the response variable into a term that contains the fixed effects $x(s; t)' \beta$ (which can include covariables), and a term that contains the random effects $\eta(s; t)$ that can be modelled using a space-time covariance matrix.

This covariance matrix is the one that will model the dependency of the points in space and time. In practice, it is often considered inseparable, such that both the covariance in space (spatial dependency) and the covariance in time (temporal dependency) is modelled.

The model for the case of forced entry burglaries will be formulated considering that the

counts follow the Poisson distribution, and that the link function is the logarithm (log-Gaussian Cox Process):

$$\begin{aligned} Z(s; t) | Y(s; t), \gamma &\sim \text{Pois}(Y(s; t), \gamma) \\ \log(Y(s; t)) &= x(s; t)' \beta + \eta(s; t) \\ s &\in D_s, t \in D_t \end{aligned}$$

The function $\eta(s; t) = \Phi(s; t)\alpha(s; t)$, where $\Phi(s; t)$ is a matrix $m_t \times n_\alpha$ of basic spatial functions, and $\alpha(s; t) \sim \text{Gau}(0, C_\alpha)$ are the associated random coefficients in which the covariance matrix is assumed to be separable, the product of the spatial and temporal covariance matrices.

The Matérn covariance function is taken as the covariance of the space. This establishes that if the distance between two points is $d = \|s_i - s_j\|$, then its covariance is $C_\nu(d) = \frac{2^{1-\nu}}{\Gamma(\nu)} (kd)^\nu K_\nu(kd)$, where Γ is the gamma function, K_ν is the modified Bessel function, $k > 0$ is a parameter of scale, and $\nu > 0$ is a smoothing parameter. This covariance is static and isotropic, or in other words, it only varies depending on the distance between the points.

Once the parameters are adjusted, the degree of spatial dependency between two points is determined by estimating the distance reached by the radial irradiation of the risk caused by the burglaries in a specific cell. The INLA method gives the estimate of the distribution of the parameter l , called the *practical range*, where $l = \sqrt{8\nu}/k$, which is the distance from where the spatial correlation will be less than 0.1.

Regarding temporal covariance, an AR(1) process is taken (Geurts, Box, and Jenkins 1977), such that a single parameter, ρ , will indicate the degree of temporal correlation between successive intervals. The fact that it has been proven that the data are usually grouped into waves of burglaries means that we expect this parameter to be near to 1 if time intervals of a week are considered.

2.2.1 INLA specifics when applied

INLA method requires a mesh to approximate the latent field (Krainski et al. 2019). This mesh is not irrelevant, as it can condition the parameters estimated by the model, such as temporal and spatial autocorrelations. Mesh should be thin enough to modulate spatiotemporal effects and thick enough to avoid redundant calculations that unnecessarily increase computational time. It will also depend on the accuracy in space and time of the input data and the sense of the continuous approach of the phenomenon studied.

In the case of home burglaries in Catalonia and from the perspective of the global study that is being carried out, cells of 250 or 500m have been chosen to compute them around the territory (250m when environment is predominantly urban, like in Barcelona, and 500m if includes other types of environments, usually in a more extensive area like the other police regions). This input grid-cell definition is the starting point of the mesh. Meshes thinner than the input grid-cell or much coarser don't make sense.

Another specific aspect of INLA comes from its Bayesian orientation. It is necessary to specify a priori information, specifically of the spatial and temporal autocorrelation, in order to be able to make the continuous approximation of the model from the solution of the stochastic partial differential equations (SPDE) (Fuglstad et al. 2019; Daniel Simpson et al. 2017).

These priors, which will condition results, will depend on the previous knowledge of the phenomenon studied, and must be in line with the mesh. A micro-scale mesh will have lower standard deviations and the possibility to observe autocorrelation at a shorter distance, which will not be possible if the mesh configuration is larger.

Different predictive models for burglaries in the police regions are studied using this methodology, with a space-time configuration of cells of 500m per side and weekly intervals. This configuration is somewhere between the micro- and the macro-scales, enabling us to do a reading of the dependencies in space and time in both extensive urban environments and in rural environments. Only for Barcelona and other municipalities in its conurbation, smaller configurations of 100m and 250m per side are used.

Models adjustment must allow to discuss the viability of the approaches based on risk decay in the geographical proximity, with a spatial dependence expected to be up to a few kilometers away.

For these global comparisons, consecutive week time intervals are chosen. Previous weeks are used to estimate the model parameters and make the prediction for the last one. Tests have been made taking 4 to 7 weeks in advance before predict. (It takes at least 2 weeks to estimate the temporal autocorrelation ρ).

The justification for these periods of time is that, usually, near repeat pattern in a victimized place doesn't last more than three or four weeks, and at the same time, allows the model to be adapted to seasonal and trend effects.

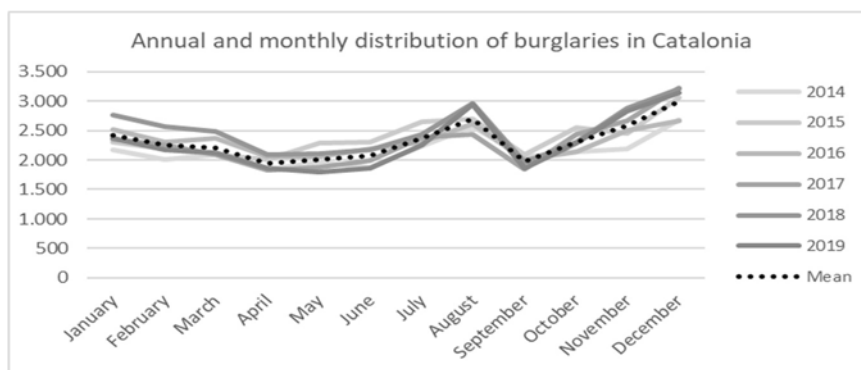
Every week, all model parameters are updated introducing the latest burglaries for next week's prediction, emulating an operative forecasting scenario. The data to be modeled are the series of burglaries in the square cells of 100, 250 or 500m where a burglary took place during 2019.

To compare the models obtained, R-INLA provides some statistics like Marginal Likelihood, Conditional Predictive Ordinates (CPO), Predictive Integral Transform (PTI), Deviance Information Criterion (DIC) or Watanabe-Akaike Information Criterion (WAIC) (Krainski et al. 2019). However, for burglaries prediction, these statistics have been discarded. Then, to measure the adjustment and model's prediction fit the coefficient of determination or R square (R^2) will be used. Data obtained with the INLA approximated model are projected onto grid-cells of different sizes. The initial projection is made in the smallest grid-cell (100, 250 or 500m) and the number of burglaries in these cells is integrated to obtain the projection in the larger grid cells (from 250m to 5km).

3 Results

3.1 General stability of the data

The monthly distribution of the number of burglaries in Catalonia between 2014 and 2019 shows a seasonal monthly pattern (Figure 3, Table 1), with an increased number of burglaries in the summer months and between the end of autumn and the beginning of winter.



Source: Author's own elaboration from data provided by the police

Figure 3.- Graph of the annual and monthly distribution of burglaries in Catalonia

There is little variation in the different years, with a monthly CV of less than 0.1 in all of them. Regarding interannual variation, out of the 6 years considered, 2014 was the year when the least burglaries were recorded, with 25,846, with the mean for the last 5 years stabilizing at around 28,000 annual burglaries. The CVs of the annual total and for the last 5 years are 0.05 and 0.04, respectively.

	January	February	March	April	May	June	July	August	September	October	November	December	Total
Any													
2014	2.178	2.001	2.089	1.842	1.959	1.885	2.235	2.571	2.061	2.147	2.192	2.686	25.846
2015	2.298	2.186	2.095	2.001	2.281	2.298	2.649	2.693	2.093	2.545	2.455	3.060	28.654
2016	2.522	2.312	2.378	2.060	2.057	2.190	2.300	2.613	2.014	2.146	2.503	2.666	27.761
2017	2.354	2.248	2.086	1.823	1.884	1.995	2.387	2.444	1.838	2.437	2.669	3.222	27.387
2018	2.767	2.569	2.477	2.093	2.111	2.179	2.436	2.958	1.966	2.304	2.873	3.204	29.937
2019	2.377	2.181	2.130	1.866	1.790	1.855	2.263	2.939	1.867	2.296	2.827	3.141	27.532
Mean	2.416	2.250	2.209	1.948	2.014	2.067	2.378	2.703	1.973	2.313	2.587	2.997	27.853
SD	205,1	187,9	172,7	118,3	174,7	181,3	152,7	206,6	103,3	158,1	255,6	254,7	1366,8
CV	0,08	0,08	0,08	0,06	0,09	0,09	0,06	0,08	0,05	0,07	0,10	0,09	0,05

Source: Author's own elaboration from data provided by police

Table 1.- Annual and monthly distribution of burglaries in Catalonia

This global profile is common to the 3 metropolitan police regions, but in this case with some slightly higher monthly CVs (Appendix, Figure 1). The rest of the police regions have monthly distributions with seasonal differences, with the summer peak disappearing. The graphics for the police regions with the least burglaries appear to be almost random around a relatively stable monthly average. These patterns can also be seen in the weekly distributions of the burglaries, although in this case the variability is higher than the monthly variability (Appendix, Figure 2).

For smaller territories, weekly distributions of burglaries in the basic police areas are also represented (Appendix, Figure 3). In general, it can be observed how, with a few exceptions, seasonality is diluted or less obvious, while at the same time random appearance increases, with weekly CVs close to or higher than 0.5. Despite this greater variability, data continue to oscillate with respect to a relatively stable weekly average across weeks and years. These characteristics are more accentuated when weekly burglaries in the 5km, 1km, and 250m cells are analyzed (Figure 4).

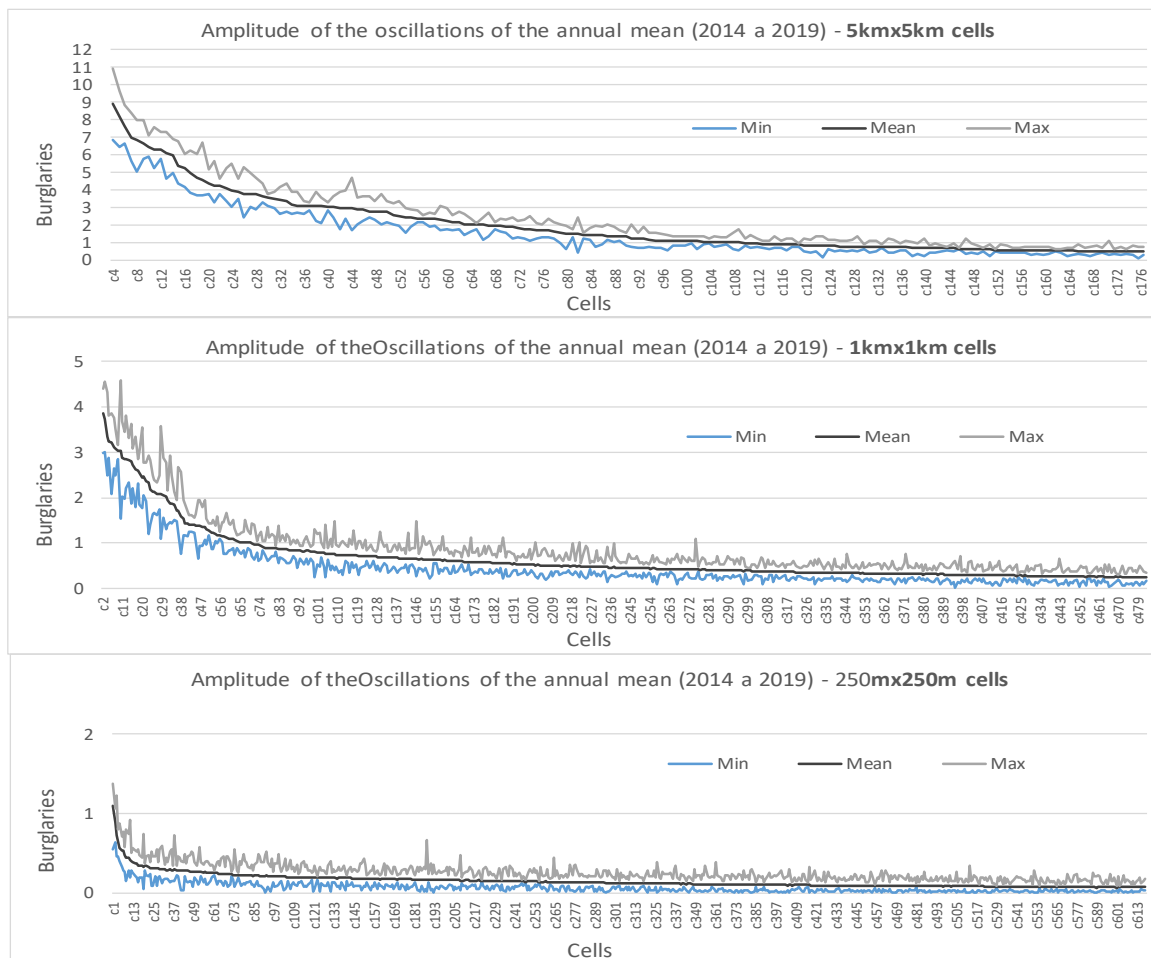


Figure 4.- Amplitude of the oscillations of the annual mean (2014 to 2019) in different size cells

The general pattern, which has been observed previously (Boqué, Serra, and Saez 2020), establishes that burglaries with forced entry in Catalonia follow Poisson distributions with relatively stable means over the years, despite occasional variations or trends.

Regarding the study of the stability of the burglaries profile, the summarized results of the annual discrepancies observed are shown. To avoid statistical distortion of the number of cells with one or very few burglaries, study has been carried out with a sample of 100 square cells of 5km with an annual mean intensity higher than 35 burglaries (weekly mean higher than 0.67) (Table 2).

Characteristics of the burglaries in each cell	Variation in each cell	
	SD<0,1	CV<0,3
Type of building (flat or house)	84% of the cells	93% of the cells
Type of residence (first or second)	74% of the cells	88% of the cells
Burglaries in the morning	87% of the cells	94% of the cells
Burglaries on working days	87% of the cells	95% of the cells

Table 2.- Annual stability of the profile of burglaries in each cell

Also, regarding the stability of the distribution of burglaries inside the 5km cells, the annual percentage in the 25 1km² subcells inside it, are compared. In this case, 425 1km² subcells with an average annual intensity of more than 15 burglaries were chosen. The annual percentage, the mean of this percentage, the standard deviation, and the CV were all

calculated for each of them. The result is that the discrepancy in the standard deviation of these is lower than 0.1 in 94.4% of the cells and, regarding the CV, it is lower than 0.3 in 70% of the cells. The size of these discrepancies can be seen graphically in Figure 5.

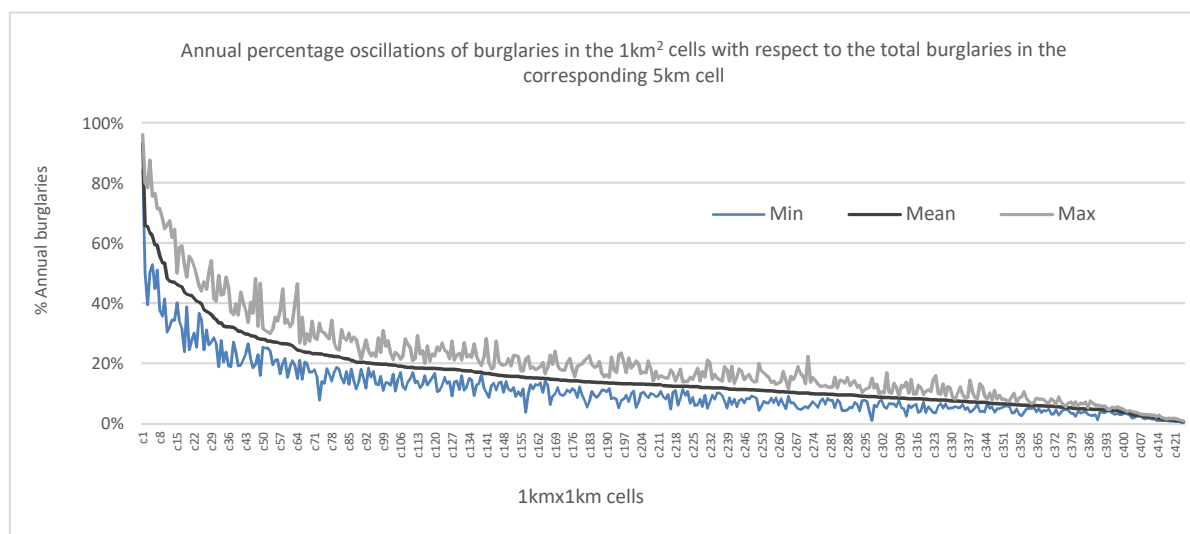


Figure 5.- Annual stability in the spatial distribution of burglaries inside the 5km cells

To check the stability of the non-random waves pattern, the runs test has been applied. The result of the weekly series of burglaries for different size cells can be seen in Table 3.

Cell size	Cell filter for minimum annual instensity	Number of cells	Non-random grouping (waves)	
5km x 5km	>20 annual burglaries	183	91.1%	
1km x 1km	>10 annual burglaries	639	70.3%	
250m x 250m	>4 annual burglaries	1385	44.7%	

Table 3.- Non-random wave patterns in different size cells

The randomness studies for each of the three sizes of cells considered were carried out independently, starting with all the cells each time and not subordinating the analysis of the smaller sizes to the fact that they are included in the larger size non-random cells.

3.2 Predictive models with R-INLA

Predictive models have been built for five police regions that concentrate the majority of home burglaries in Catalonia. These are the three metropolitan regions along with the RPG and the RPCT. The results obtained in all of them are similar and lead to the same conclusions.

INLA output offers, by default, an estimate of *a posteriori* distributions of the main approximate model parameters and its mean. These parameters include temporal autocorrelation (ρ), spatial autocorrelation (l), and the standard deviation of the logarithm of the number of burglaries at the vertices of the mesh (σ). Table 4 shows the minimum and maximum value of these means in the different tests that have been done in non-overlapping weekly periods during 2019.

Police Region	ρ		range (l) (km)		σ	
	min	max	min	max	min	max
Barcelona Metropolitan	0.85	0.95	0.8	1.5	0.90	0.95
North Metropolitan	0.76	0.89	1.4	3.7	1.0	1.3
South Metropolitan	0.58	0.82	1.3	1.9	1.3	1.4
Girona	0.80	0.95	2.0	4.1	1.1	1.2
Camp de Tarragona	0.81	0.97	0.9	1.5	1.3	1.7

Table 4.- Statistical summary of “a posteriori” parameters: minimum and maximum values of the means estimated by INLA in various tests in non-overlapping periods for each of the police regions during 2019

To illustrate it in more detail, a summary of some examples is shown corresponding to the North Metropolitan, Girona and Barcelona Metropolitan police regions.

In the case of the RPMN (Figures 6, 7 and 8), a 7-week predictive scheme was used, where the previous 6 were used to adjust the model parameters, and the prediction was made for the seventh. Data input had an accuracy of 500m. Figure 7 shows the prediction obtained, with an average of the high temporal autocorrelation parameter ($\rho = 0.76$) and a spatial dependence mean with a radius of 1.6 km, distributed between 1 and 2.5 km. In other weekly period, spatial dependence mean of 3.7 km has been obtained. In the case of Girona police region (Figure 10), this value reaches a mean of 4.1km, with an oscillation between 2 and 6km. Due to the territorial profile of these regions and according to the average dimensions of the populations of Catalonia, it can be stated that the risk of burglaries extends to entire populations and could sometimes reach neighboring populations. Similar results have been obtained in the Metropolitan South and Camp de Tarragona police regions.

Figures 8 and 11 show how the fit of the data with this large spatial configuration is good, with high values of the coefficient of determination R^2 . In contrast, in both cases, prediction made for the last week series shows less accuracy. In fact, high temporal autocorrelation implies that prediction ends up being a projection of the previous week’s heat map.

In the case of Barcelona, input data accuracy is 250m, and a predictive scheme of 5 previous weeks has been followed. In this case (Figure 13) a smaller spatial autocorrelation is obtained, with an average of 800m, varying between 400 and 1200m. These dimensions, in the urban environment of Barcelona, can be considered excessively extensive to be able to delimit suitable risk zones to police preventive strategies. In Figure 14, it is observed that the 500m square cells are too small to pick up the spatial dependence effect, with a low R^2 for both the previous weeks’ adjustment and the predictive week. In Figure 15, considering a cell size of 1km², closer to the burglaries practical range of Barcelona, the fit and prediction are better.

3.2.1 INLA summary results for North Metropolitan Police Region

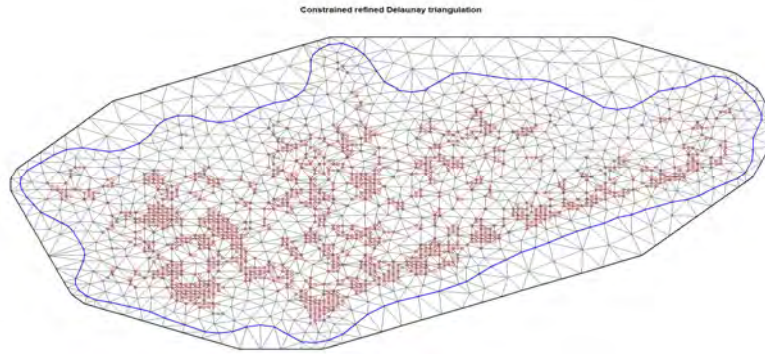


Figure 6.- Mesh for North Metropolitan PR, including burglaries locations (red points)

Prediction: week 8



Mean of estimation parameters (week 2 to 7)

ρ	0.76
Range (l) (km)	1.6
σ	1.3

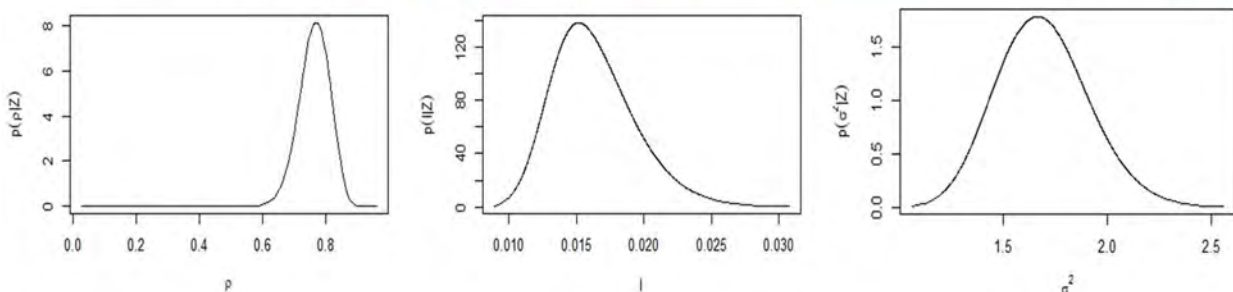
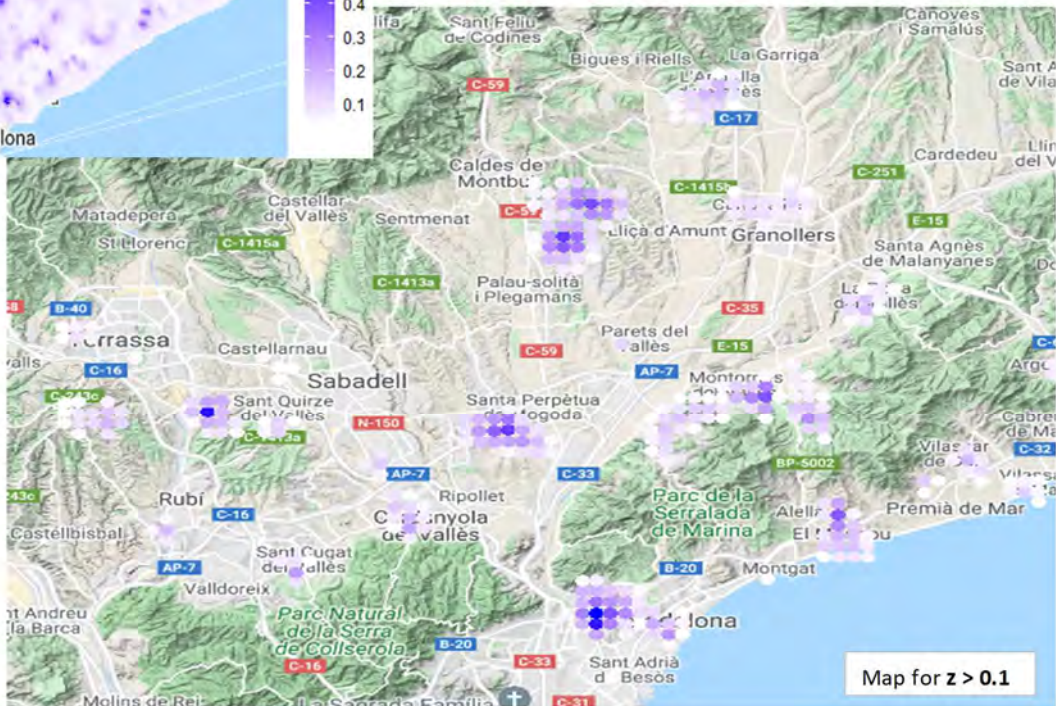


Figure 7.- Example 1.1 (NMPR): INLA predicted burglaries (z) in cells of 500m per side for week 8 with spatiotemporal dependence parameters and variance information (mean and probability distribution). Predictive scheme of the example: 6 previous weeks to adjust parameters before making a prediction for next week.

3.2.2 INLA summary results for Girona Police Region

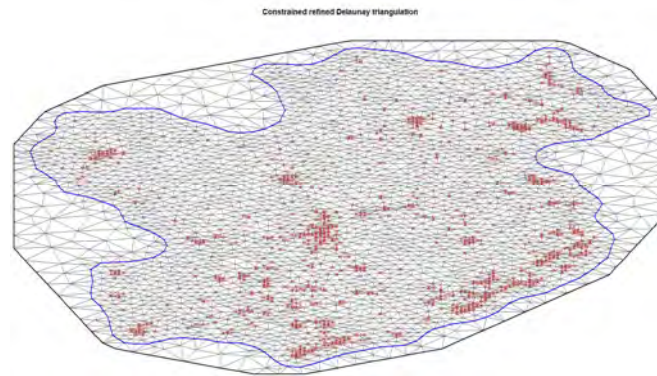


Figure 9.- Mesh for Girona PR, including burglaries locations (red points)

Prediction: week 18

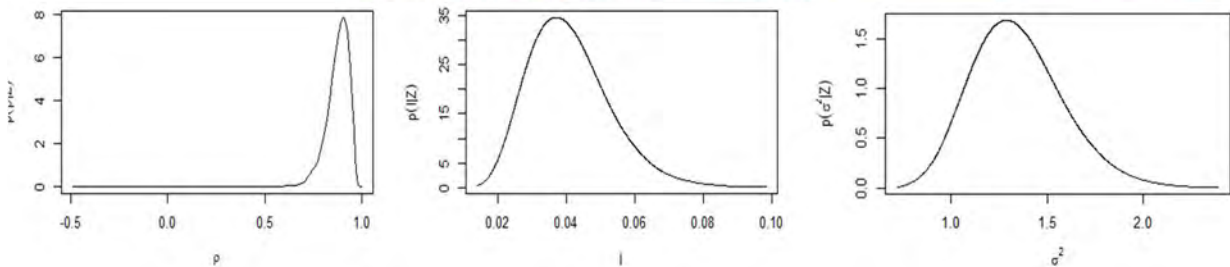
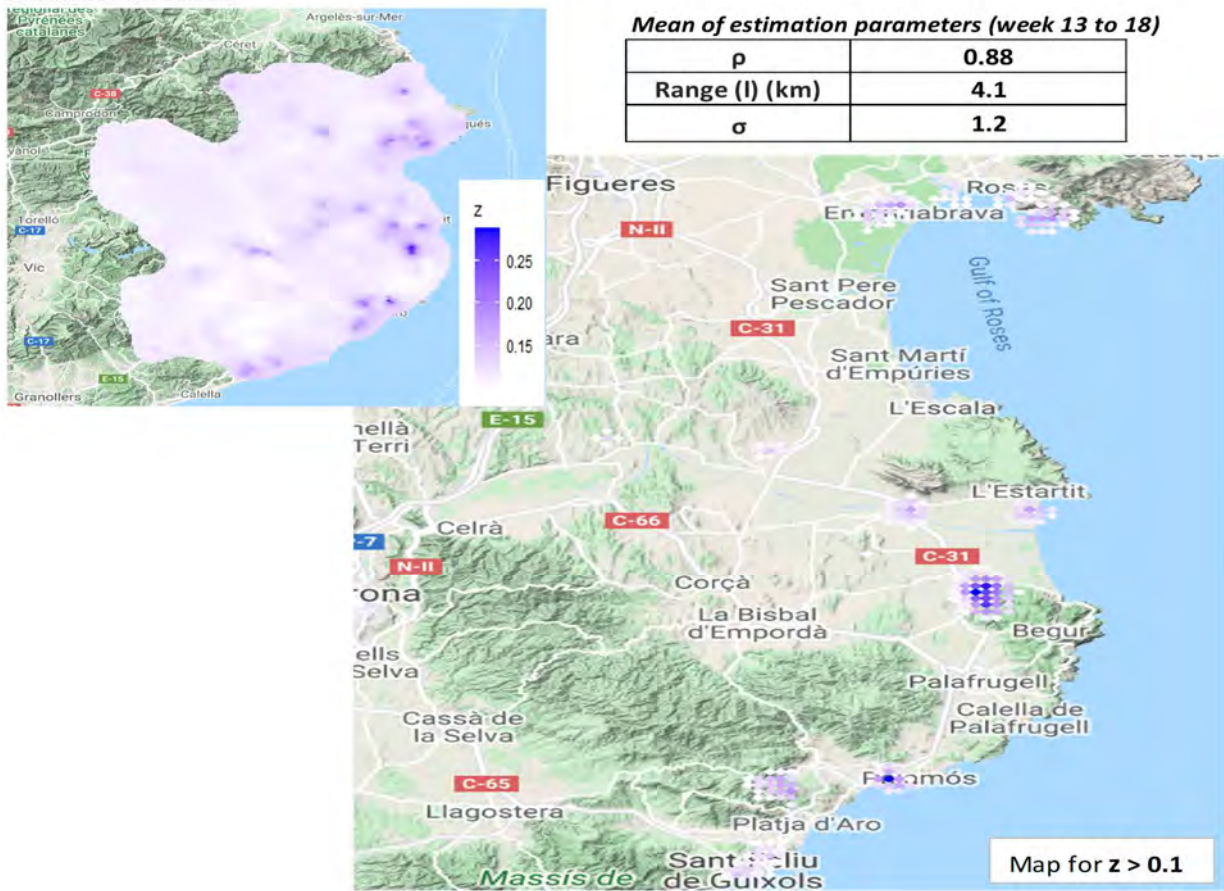


Figure 10.- Example 2.1 (GPR): INLA predicted burglaries (z) in 500m cells for week 18 with spatiotemporal dependence parameters and variance information (mean and probability distribution). Predictive scheme of the example: 5 previous weeks to adjust parameters before making a prediction for next week.

3.2.3 INLA summary results for Barcelona Metropolitan Police Region

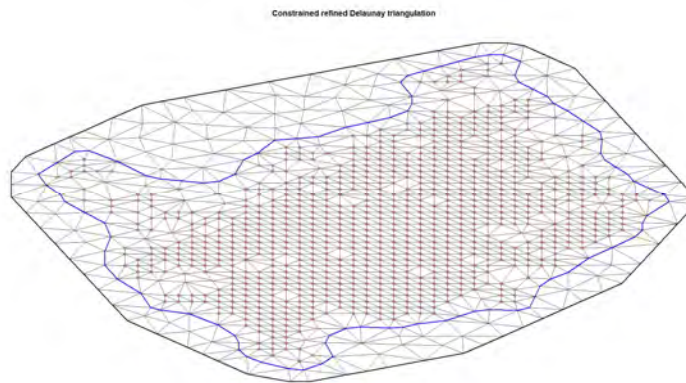
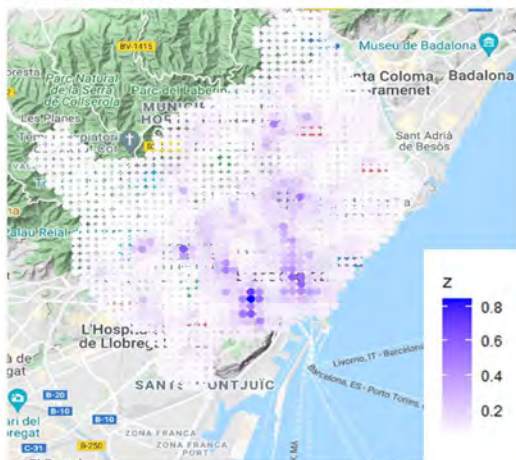


Figure 12.- Mesh for Barcelona MPR, including burglaries locations (red points)

Prediction: week 10



Mean of estimation parameters (week 6 to 10)

ρ	0.95
Range (l) (km)	0.8
σ	0.9

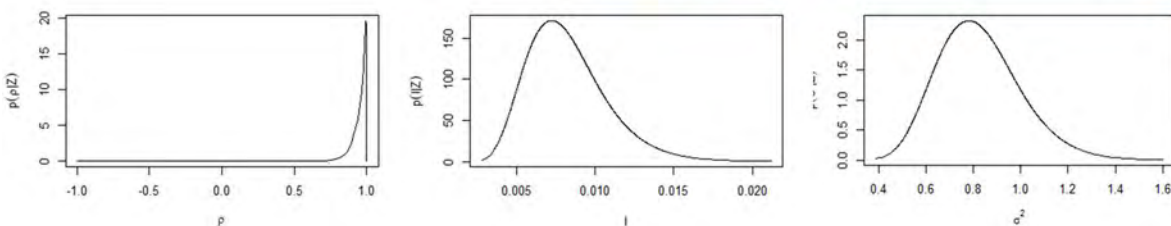
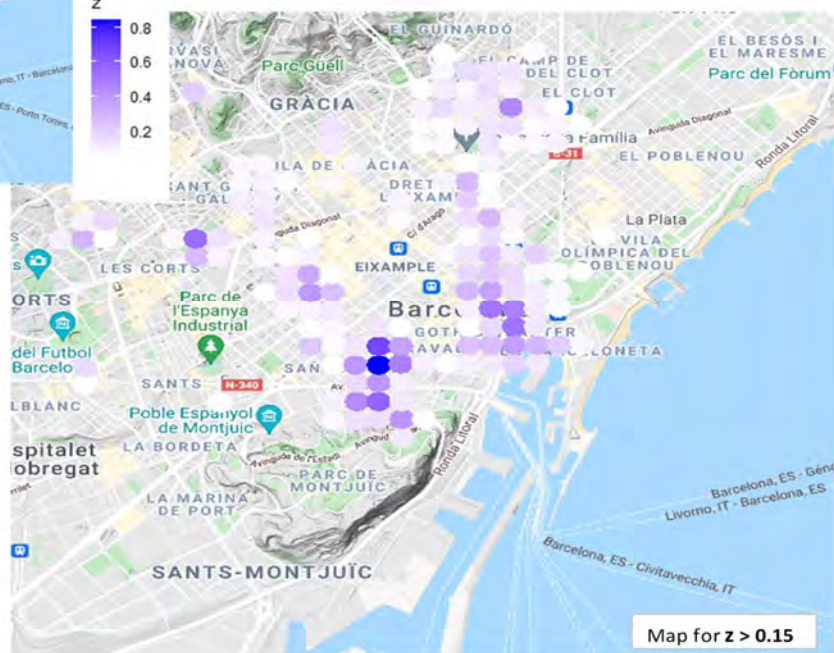


Figure 13.- Example 3.1 (BMPPR): INLA predicted burglaries (z) in 250m per side cells for week 10 with spatiotemporal dependence parameters and variance information (mean and probability distribution). **Predictive scheme of the example:** 4 previous weeks to adjust parameters before making a prediction for next week.

Real data																								Adjusted data																								CC	R ²
week1																								week1																									
Total burglaries 90																								Total burglaries 95																									
A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0,64	0,41
week2																								week2																									
Total burglaries 96																								Total burglaries 97																									
week3																								week3																									
Total burglaries 91																								Total burglaries 97																									
week4																								week4																									
Total burglaries 93																								Total burglaries 97																									
week5																								week5																									
Total burglaries 79																								Total burglaries 95																									

Figure 14.- Example 3.2 (BMPr): INLA output: 500m per side cells individual adjustment per week (6 to 9) and predicted values for 10th week

Real data													Adjusted data														
Total burglaries													Total burglaries													cc	R ²
week1													week1														
A													A												0,79	0,62	
week2													week2														
A													A												0,84	0,70	
week3													week3														
A													A												0,82	0,67	
week4													week4														
A													A												0,82	0,66	
week4													week4														
A													A												0,51	0,26	

Figure 15.- Example 3.2 (BMPR): INLA output: 1km per side cells individual adjustment per week (6 to 9) and predicted values for 10th week

4 Discussion

It has been shown that burglaries in Catalonia are stably distributed in space and time. There are seasonal and trend variations in the average intensity, over the long term, but a stable pattern based on the Poisson distribution is observed. This stability goes hand in hand with a dynamic of weekly waves of burglaries, not random, detected in large and less visible spaces than smaller ones, such as in the 250m per side cells.

These macro waves observed in 5km per side cells, with a heterogeneous environment inside, can have two different explanations. First, they are generated by the sum of independent microwaves located in different areas inside the cell, which can be even several kilometers away from one another. If this were the case, an approach based on the micro-scale near repeat and the Knox test (Briz-Redón, Martínez-Ruiz, and Montes 2020) should correctly model the micro-level burglaries distribution in space and time, and so, also the macro-level. Second, one could consider the overall cell as a near-repeat "unit", so that any subsequent burglary in the cell after those that have originated the macro wave can be considered a replica, without worrying about micro-level patterns inside.

The first explanation, which could be framed within the classic near-repeat victimization theory, has some obvious weaknesses, because it seems contradictory for local micro waves to be independent and synchronized at the same time, ending up generating stable and non-random macro waves.

To test this possible limitation of applying classical near-repeat scenario in Catalonia to predict burglaries, a generic model, based on the concept of the latent field and capable of modelling different situations in which the data presents spatial and/or temporal dependencies, was constructed. Approaches from the perspective of point processes, areal data, grid data, and continuous data are suitably modelled in it (Rue, Martino, and Chopin 2009; Simpson et al. 2015). In this line, the spatiotemporal log-Gaussian Cox model used in this study (Diggle et al. 2013) is the generic approach to model the burglaries near repeat pattern. The Matérn covariance function, considered to represent the spatial correlation between two points, where the main restriction is being inversely proportional to the distance, is also generic. This is likewise the case for the AR(1) autoregressive process, which may be the simplest to model temporal dependency and has been shown to be both sufficient and suitable in this case. Therefore, the starting model is generic and has few assumptions or restrictions, apart from those consistent with the hypothesis of risk spread in the geographical and temporal proximity.

One advantage of modelling based on the continuous latent field is that it allows us to largely avoid the debate about the optimum data aggregation configuration (Malleonid, Steenbeek, and Andresen 2019; Hipp and Kim 2017; Chainey 2013). Cells of 100, 250 or 500m size used in the input give sufficient detail to model burglaries through different residential environments. Based on these, the continuous estimate of the latent field obtained using INLA can be projected to any desired spatial aggregation, to be represented, and analyzed, as we have done in cells of size 250m to 5km.

It's possible to conclude that replicas of initial burglaries, or a part, would occur at a maximum distance from these equivalent to the practical range, so that it would justify a large space configuration in the weekly prediction schemes, such as the case of 5km cells.

Despite the model's flexibility and adaptability, INLA approximation shows an excessively monotonous result, where predicted events in each cell are basically the same in all the time

intervals, with weekly values increasing and decreasing proportionally to the increase and decrease of the total burglaries in the territory.

Hotcell maps of the adjusted values always have the same or a very similar appearance, except for global multiplication by a scale parameter related to the variation in the weekly trend of the total burglaries in the police region. Meanwhile, the hotcell maps of the real values show more variability and a different dynamic. Stability of the data makes the prediction acceptable both overall and according to adjustment parameters like the R^2 , although it clearly does not represent well the local variability or short-term patterns of burglaries distribution in space and time, to the effect that it does not allow us to anticipate where the highest crests of the waves will be.

This limitation in no way contradicts the fact that there is a concentric risk diffusion for each burglary. INLA method detected spatial dependency at distances of between 1 and 6km. The interpretation of this fact is that any burglary has the effect of increasing the risk of new burglaries both in the proximity and some distance away. In other words, the more burglaries there are a week in the police region as a whole, the greater the probability that there will be more burglaries the following week, but anywhere in the region. In heterogeneous environments, it seems that the close homogeneous space in which to make the replicas is too small to fit many, so some will be made in more remote places, which may be similar to the previous areas.

Temporal dependence

With respect to the temporal factor, in all scenarios studied this has been positive and near to 1 (rho parameter $\rho \cong 1$). According to this result obtained with INLA, the best prediction for the following week will be the projection of what has happened in the current week (in a large cell configuration with side sizes of about 2 practical ranges).

Locally, this can be interpreted as a confirmation of the pattern of non-random waves because periods of activation and non-activation, for different thresholds, last for various weeks, or in other words, they correlate among themselves. We could also interpret that the dynamics of activation and deactivation are smooth, always enabling this similarity with the previous period, whether the wave is increasing or decreasing in magnitude.

But in real data the pattern of the waves is not smooth. Sudden increments in the oscillations can be observed, even though they are normally within an already active wave (otherwise, the non-random pattern of the spates would be broken). Seeking to minimize the global error, these models tend towards the mean because they do not capture the local, short-term dynamic well.

The possibility that too large spatial resolution in the input data can mask the real spatiotemporal effect on a smaller scale could be raised. In this sense, tests have been carried out in Barcelona applying 100m grid-cells to model burglaries in the Eixample (Appendix B, figures 1, 2 and 3) and Ciutat Vella districts. The result has been that temporal autocorrelation is lost ($\rho \cong 0$), with mean values close to 0 and which also can be negative (Appendix B, figure 2, bottom), indicating that burglars would avoid returning to the 100m cells victimized the previous week. However, in these micro-scale configurations and in densely populated environments, such as the Eixample district, spatial dependence is detected at distances of less than 250m. Therefore, there is spatial but not temporal clustering, and micro-scale weekly prediction is not possible.

Consequences of low detection of the near repeat pattern in geographical proximity

The consequences of these limitations of the predictive models based on risk decay in geographic proximity are rather a double-edged sword for both *predictive policing* programs (Perry et al. 2018) and the police that enforce them. They tend to recommend going to static hotcells, which is usually where there is, in average, a greater likelihood of burglaries, and they especially recommend this when there have been burglaries during the previous week. While this is not a bad strategy since it focuses police attention on the critical spots, it can almost certainly be improved. In the usual hotcells, there is expected to be special prevention, with a long-term planning horizon, to reduce the crime opportunity of the place (“Center for Problem-Oriented Policing | ASU Center for Problem-Oriented Policing” n.d.); “Risk Terrain Modeling | Official Site - Home” n.d.). It is related to the flag hypothesis and can be detected without the use of predictive programs due to the stability of the data.

Therefore, what appears to be most recommended for designing dynamic preventive strategies is improving the modelling of the dynamic component in a unique manner and without depending on the static component.

To do so, abandoning these micro-scale configurations is recommended. This is also justified by analyzing what happens in the 250m cells because in these, even though the series are non-random and the burglaries appear grouped, they will mostly take place at a rate of 1 per week. According to the Poisson distribution, for cells with averages of 0.2 it is very unusual to observe weeks with a high number of burglaries (above 2). Burglaries considered predictable in a larger configuration will end up isolated in 250m cells which, individually, are not considered to be optimum for making predictions. This restricts predictive possibilities in some high-density urban areas where only a small number of the total burglaries in Catalonia take place.

To improve burglaries prediction, it's necessary to zoom out towards large scales and study spatiotemporal relations that do not rule out distant correlations. It is not sufficient to look for correlations between burglaries located in the same large cell (Boqué, Serra, and Saez 2020); it is about linking distant events and clusters, and looking for correlations between different cells, adjacent or not. The criminological justification for this change of focus in heterogeneous environments can be found in the classical theories that explain crime patterns (Gottfredson and Hirschi 1990; Clarke and Felson 2008; Bernasco 2009), where what is uppermost is maximizing the benefits and minimizing the cost or the risk.

This study is not the first detecting limitations of prediction models based on near-repeat phenomenon. Methods to measure the micro-level predictive capacity of the near-repeat pattern, before deciding whether to adopt it as the reference for making predictions, has been proposed (Groff and Taniguchi 2019). Differences in the intensity of this pattern depending on the environment, has been detected (Spencer Paul Chainey and Figueiredo 2016; Spencer P. Chainey et al. 2018).

In fact, the need to broaden the concept of near-repeat victimization has long been raised (Farrell, Pease, and Fraser 2012). And in the last few years, several researchers have proposed new approaches. One of this is the idea of near repeat chain (Haberman and Jerry H Ratcliffe 2012; Glasner, Johnson, and Leitner 2018), also the approach of conditional spatial Markov chain and spillover effects (Rey, Mack, and Koschinsky 2012; Zhang and Song 2014). More recently, the shift pattern has been proposed (Wang and Zhang 2020; Wang et al. 2019), considered an extension of the near-repeat victimization that links geographically distant events close in time, and in similar environments, due to displacement effects. This last

research (Wang et al. 2019) concludes that adding shift pattern to near repeats increases predictive capacity, and finds that 5km is the average displacement distance (in this research, burglaries are located in a large Chinese city inside a circumference of approximately 10km radius). Displacement is explained (Wang and Zhang 2020) as a response to crime prevention initiatives that motivates offenders to look for alternative places for suitable opportunities. Offenders modify their awareness space (Brantingham and Brantingham 1995; Bernasco and Nieuwebeerta 2005), adding new areas to their “mental map”, in which they can burglar consecutively due to similar geographic characteristics.

The shift pattern approach may be quite appropriate in our case, but there’s a need to go further to adapt it to the context of crime in Europe, especially when trying to model crime beyond urban environments, and in a diverse and heterogeneous territory such as Catalonia.

In Europe, criminal groups that commit burglaries (Wollinger et al. 2018) tends to be of international composition, have complex and diverse forms of organization, are specialized and have a lot of mobility and itinerancy. Studies on journeys to crime (Bernasco and Nieuwebeerta 2005; Van Daele and Vander Beken 2009; Michael Townsley and Sidebottom 2010; Van Deviver, Van Daele, and Van Der Beken 2015) pinpoint different factors that explain why these can be relatively long, and they argue it in a similar way. There is a consensus that specialized groups tend to make longer journeys to access zones that they consider favorable to commit burglaries. They can victimize dispersed areas in one hit on the same day. The decision-making process includes reconnoitering the area, usually some days previously, choosing the areas depending on the environment, and applying their knowledge or previous experiences of burglaries. Once in the place chosen, they decide which specific residence to burgle. They prefer to move around using fast roads (motorways or dual carriageways), and they hope to compensate for the added cost of making the long journey with the lower risk of being detected and with a better expected haul. They evaluate the fact that in remote residential areas there is little social control and they are unknown to the police in the area. The same studies point out that repeat victimization could still be a valid pattern to explain the generation of distant clusters, and that criminal groups act in a serial way, alternating between spates of activity and periods of no activity. They justify the possibility of being inactive for a period with the larger haul expected from this *modus operandi*.

Putting it all together reinforces the initial hypothesis, which now can be reformulated and extended for future research. We name it “*Unus pro omnibus, omnes pro uno*”, in the sense that each burglary will be considered a replica of the whole criminal initiative, and all burglaries will fit in a single spatiotemporal pattern, the “all-in-one”.

The all-in-one spatiotemporal repeat pattern

The “*Unus pro omnibus, omnes pro uno*” hypothesis states that in heterogeneous environments, offenders tend to distribute burglaries in space according to opportunity criteria, with a tendency to concentrate them in several and separate places. Considering the whole, a non-random wave will be observed but it won’t meet the near-repeat pattern in the sense of geographical proximity, but in the sense of co-occurrence and same offenders. Crime concentration areas will be observed, where the classical near-repeat pattern will work. By connecting them via terrestrial, a network will appear, with nodes weighted according to crime concentration and links weighted according to terrestrial communications facilities.

In this extensive and complex network, the spatiotemporal pattern that will link all burglaries will be the so-called **all-in-one spatiotemporal repeat pattern** (Figure 16), which include

near-repeats, shift and co-occurrence patterns. These all-in-one repeats will last a few weeks, as detected in classical near-repeat patterns.

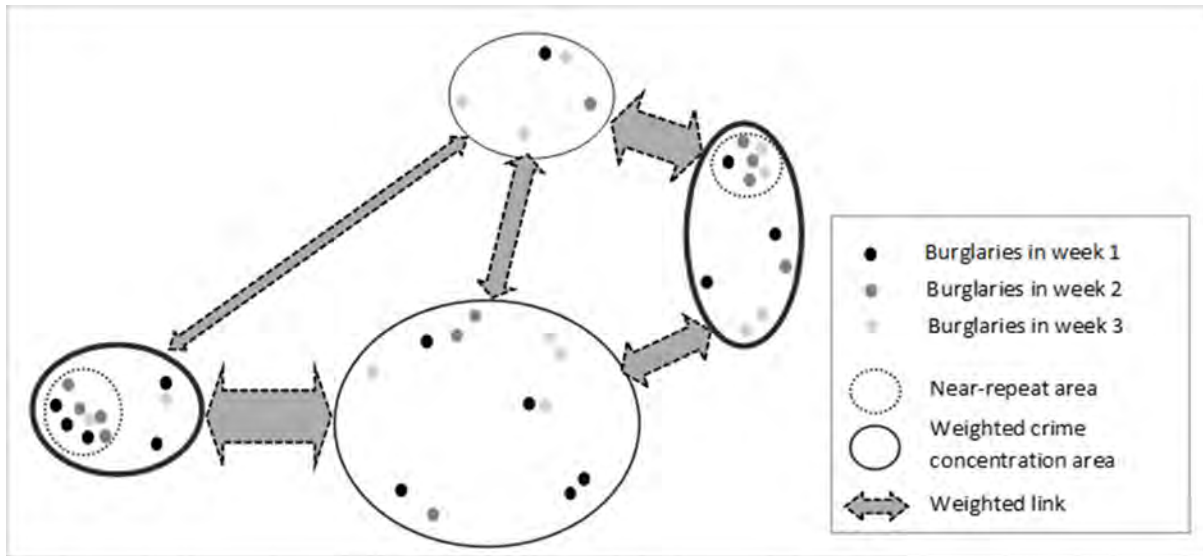


Figure 16.- Schema of the all-in-one spatiotemporal repeat pattern for heterogeneous environments

In the all-in-one repeat pattern, concentration and dispersion are observed equally, and would explain the low detection of the classical near-repeat pattern at a microscale and its high detection on larger scales. In observing Figure 16 it's also highlighted the need to determine which areas are being victimized at the same time, and links between them, in order to predict replicas and avoid displacement effects in case of preventive strategies in any of these areas.

Relation between log-Gaussian Cox Process and Self-Exciting Point Process models

It can be said that crime prediction began with the Self-Exciting Point Process model (Mohler et al. 2011), that PREDPOL assumed and developed. A point process with intensity $\lambda(t, x, y)$ is considered which is the sum of the static and dynamic components:

$$\lambda(t, x, y) = \mu(t, x, y) + \sum_{\{k: t_k < t\}} g(t - t_k, x - x_k, y - y_k).$$

$\mu(t, x, y)$ is the mean intensity, related to the background data, the static component and long-term risk (flag hypothesis), and the function $g(t, x, y)$ is the increase of risk intensity caused by nearby burglaries in space and time, related to the short-term dynamics of the risk (boost hypothesis). In practice, they assume that $\mu(t, x, y) = v(t)\mu(x, y)$, that is, temporal fluctuations of the mean intensity are global, common throughout the territory, and due to aspects such as meteorology or seasonality. As for the function $g(t, x, y)$, it is assumed to decrease concentrically with distance, and is considered isotropic, so that its value depends only on the distance h between the points (x, y) and (x_k, y_k) .

Thus, Self-Exciting Point Process and log-Gaussian Cox process models have much in common, with the main difference being how historical data it's treated. The latter does not functionally distinguish between static and dynamic factors but does play a role in the predictive scheme. The previous weeks that are taken as a reference to approximate parameters model before making a prediction, are affected both by the static factor (mean intensity of the place and time), and by the dynamics of the spatiotemporal autocorrelation.

Therefore, most of the discussion and conclusions that can be reached in this article can be extended to Self-Exciting Point Process approach and equivalents.

Limitations

This study has some limitations. Crime data comes from reported events to police, and so there is an unknown black figure, although in the case of burglaries it's considered to be relatively low compared to other crime typologies. These missing data may be attempted or less serious burglaries, and may be assumed that their spatiotemporal distributions run parallel to reported ones, so don't interfere with main conclusions. With regard to reported burglaries, there is another limitation in relation to the time the burglary occurred. Usually, it's possible to know the day and the time window (morning, afternoon or night), but in some cases, especially when the burglary occurs in an unoccupied second residence, the time window can be as much as several weeks. Percentage of these cases is low, and in this study, we have decided to consider intervals of a week that is a good time window for predicting purposes.

A final note is to justify the absence of covariates in prediction models, which can be considered a limitation. Although it is not always easy to obtain territory layers of information that may be appropriate for burglaries' spatiotemporal modelling, the option of not including covariates has been intentional. The aim of the study was to test the risk decay hypothesis in spatiotemporal proximity, and not the construction of a prediction model that could consider the type of environment, house or burglary to adapt or grade that risk decay. Given now the limitation of this effect, other approaches, such as the proposed "all-in-one pattern", may need covariates for proper contrast and modelling.

5 Conclusions

The phenomenon of residential burglaries in Catalonia follows stable spatiotemporal patterns, related to the static factor, and dynamic waves of burglaries that enable us to predict them in large-scale configurations. When modelling from the micro-scale, limitations appear, and the dynamics of burglaries is not well captured.

To measure to which extent the classical principles of near-repeat victimization theory can be used for predicting purposes, we have considered a log-Gaussian Cox process to estimate and predict the number of weekly burglaries through modeling this phenomenon from the micro-scale and according to risk decay in time and space.

With this approach, INLA methodology has proven to be a very suitable tool for analyzing this kind of spatiotemporal dependencies, offering an output of *a posteriori* parameter distribution that greatly facilitates the interpretation and description of phenomenon. Although in Catalonia results obtained fail in making good weekly predictions, it is highly recommended that any police force interested in studying near repeat victimization use INLA as a first step, to see limits and chances.

Analyzing the results in detail and relying on other research, a hypothesis and a new spatiotemporal pattern have been proposed. We call it the "*Unus pro omnibus, omnes pro uno*" hypothesis, and describes a new framework for explaining space and time dependencies of crime in heterogeneous environments, following the so-called all-in-one repeat pattern. This new framework can be important for understanding serial and specialized burglars' behavior in these environments, which are quite common in Europe. It will also allow a better

understanding of the burglaries' distribution in space and time, and will generate new challenges for crime prediction models and police prevention strategies.

Pattern of near repeat victimization, observed and widely proven around the world (Johnson et al. 2007; Kikuchi, Amemiya, and Saito 2010; Z. Wang and Liu 2017), meets the requirements of being interpretable from the policing perspective and useful for prevention. Even though in this paper it has been demonstrated that its formulation is insufficient to model and predict the general behavior of burglaries in heterogeneous environments, the solidness of the criminological studies that underpin this concept means that it continues to be a basic spatiotemporal pattern for prediction of home burglaries and that maybe, when broadened, as proposed, it will also end up explaining the generation of distant, correlated events and clusters.

6 References

- Bennett Moses, Lyria, and Janet Chan. 2018. "Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability." *Policing and Society* 28 (7): 806–22. <https://doi.org/10.1080/10439463.2016.1253695>.
- Bernasco, Wim. 2009. "FORAGING STRATEGIES OF HOMO CRIMINALIS: LESSONS FROM BEHAVIORAL ECOLOGY." *CRIME PATTERNS AND ANALYSIS* 2 (5). <http://wimbernasco.nl/Manuscripts/English/Bernasco2009a.pdf>.
- Bernasco, Wim, and Paul Nieuwbeerta. 2005. "How Do Residential Burglars Select Target Areas? : A New Approach to the Analysis of Criminal Location Choice." *British Journal of Criminology* 45 (3): 296–315. <https://doi.org/10.1093/bjc/azh070>.
- Blangiardo, Marta, Michela Cameletti, Gianluca Baio, and Håvard Rue. 2013. "Spatial and Spatio-Temporal Models with R-INLA." *Spatial and Spatio-Temporal Epidemiology*. <https://doi.org/10.1016/j.sste.2013.07.003>.
- Boqué, Pere, Laura Serra , and Marc Saez. 2020. "'Surfing' Burglaries with Forced Entry in Catalonia: Large-Scale Testing of near Repeat Victimization Theory." *European Journal of Criminology*, November, 147737082096810. <https://doi.org/10.1177/1477370820968102>.
- Bowers, Kate J, and Shane D Johnson. 2004. "Who Commits Near Repeats? A Test of the Boost Explanation." *Western Criminology Review* 5 (3): 12–24. http://westerncriminology.org/documents/WCR/v05n3/article_pds/bowers.pdf.
- Brantingham, Patricia, and Paul Brantingham. 1995. "Criminality of Place - Crime Generators and Crime Attractors." *European Journal on Criminal Policy and Research* 3 (3): 5–26. <https://doi.org/10.1007/BF02242925>.
- Briz-Redón, Álvaro, Francisco Martínez-Ruiz, and Francisco Montes. 2020. "Adjusting the Knox Test by Accounting for Spatio-Temporal Crime Risk Heterogeneity to Analyse near-Repeats." *European Journal of Criminology*. <https://doi.org/10.1177/1477370820905106>.
- "Center for Problem-Oriented Policing | ASU Center for Problem-Oriented Policing." n.d. Accessed November 10, 2020. <https://popcenter.asu.edu/>.
- Chainey, Spencer. 2013. "Examining the Influence of Cell Size and Bandwidth Size on Kernel Density Estimation Crime Hotspot Maps for Predicting Spatial Patterns of Crime." *BSSLg* 60 (1): 7–19.
- Chainey, Spencer P., Sophie J. Curtis-Ham, R. Mark Evans, and Gordon J. Burns. 2018. "Examining the Extent to Which Repeat and near Repeat Patterns Can Prevent Crime." *Policing* 41 (5): 608–22. <https://doi.org/10.1108/PIJPSM-12-2016-0172>.
- Chainey, Spencer Paul, and Braulio Figueiredo. 2016. "Examining the Extent of Repeat and near Repeat Victimization of Domestic Burglaries in Belo Horizonte , Brazil." *Crime Science*, 1–10. <https://doi.org/10.1186/s40163-016-0049-6>.
- Cinelli, Virginia, and Alberto Manrique. 2019. "El Uso de Programas de Análisis Predictivo En La Inteligencia Policial: Una Comparativa Europea." *Revista de Estudios En Seguridad Internacional* 5 (2): 1–20. <https://doi.org/10.18847/1.10.1>.
- Clarke, R V G, and Marcus Felson. 2008. *Routine Activity and Rational Choice : Advances in Criminological Theory*.

- Daele, Stijn Van, and Tom Vander Beken. 2009. "Out of Step? Mobility of 'Itinerant Crime Groups.'" *Crime, Money and Criminal Mobility in Europe*, 43–70.
- Deviver, Christophe Van, Stijn Van Daele, and Tom Van Der Beken. 2015. "What Makes Long Crime Trips Worth Undertaking Balancing Costs and Benefits in Burglars Journey to Crime." *British Journal of Criminology* 55 (2): 399–420. <https://doi.org/10.1093/bjc/azu078>.
- Diggle, Peter J., Paula Moraga, Barry Rowlingson, and Benjamin M. Taylor. 2013. "Spatial and Spatio-Temporal Log-Gaussian Cox Processes: Extending the Geostatistical Paradigm." *Statistical Science* 28 (4): 542–63. <https://doi.org/10.1214/13-STS441>.
- Egbert, Simon, and Susanne Krasmann. 2020. "Predictive Policing: Not yet, but Soon Preemptive?" *Policing and Society* 30 (8): 905–19. <https://doi.org/10.1080/10439463.2019.1611821>.
- EUCPN. 2016. "Recommendation Paper: Predictive Policing," 13. www.eucpn.org.
- Farrell G., and K Pease. 1993. "Once Bitten, Twice Bitten: Repeat Victimization and Its Implications for Crime Prevention." *Police Research Group, Crime Prevention Unit Paper* 46, no. 46: 38. <http://hdl.handle.net/2134/2149%5Cnhttp://core.ac.uk/download/pdf/67259.pdf>.
- Farrell, Graham, and Ken Pease. 2017. "Preventing Repeat and near Repeat Crime Concentrations." In *Handbook of Crime Prevention and Community Safety*. <https://doi.org/10.4324/9781315724393>.
- Farrell, Graham, Ken Pease, and Simon Fraser. 2012. "Prediction and Crime Clusters," no. 1995: 3862–71.
- Fuglstad, Geir Arne, Daniel Simpson, Finn Lindgren, and Håvard Rue. 2019. "Constructing Priors That Penalize the Complexity of Gaussian Random Fields." *Journal of the American Statistical Association* 114 (525): 445–52. <https://doi.org/10.1080/01621459.2017.1415907>.
- Gerstner, Dominik. 2018. "Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany." *European Journal for Security Research* 3 (2): 115–38. <https://doi.org/10.1007/s41125-018-0033-0>.
- Geurts, Michael, George E. P. Box, and Gwilym M. Jenkins. 1977. "Time Series Analysis: Forecasting and Control." *Journal of Marketing Research* 14 (2): 269. <https://doi.org/10.2307/3150485>.
- Glasner, Philip, Shane D. Johnson, and Michael Leitner. 2018. "A Comparative Analysis to Forecast Apartment Burglaries in Vienna, Austria, Based on Repeat and near Repeat Victimization." *Crime Science* 7 (1). <https://doi.org/10.1186/s40163-018-0083-7>.
- Gottfredson, Michael R., and Travis. Hirschi. 1990. *A General Theory of Crime*. Stanford University Press.
- Groff, Elizabeth, and Travis Taniguchi. 2019. "Quantifying Crime Prevention Potential of Near-Repeat Burglary." *Police Quarterly* 22 (3): 330–59. <https://doi.org/10.1177/1098611119828052>.
- Haberman, Cory P., and Jerry H Ratcliffe. 2012. "The Predictive Policing Challenges of Near Repeat Armed Street Robberies" 6 (2): 151–66. <https://doi.org/10.1093/police/pas012>.
- Hipp, John R., and Young An Kim. 2017. "Measuring Crime Concentration across Cities of

- Varying Sizes: Complications Based on the Spatial and Temporal Scale Employed.” *Journal of Quantitative Criminology* 33 (3): 595–632. <https://doi.org/10.1007/s10940-016-9328-3>.
- Johnson, Shane D. 2014. “How Do Offenders Choose Where to Offend? Perspectives from Animal Foraging.” *Legal and Criminological Psychology* 19 (2): 193–210. <https://doi.org/10.1111/lcrp.12061>.
- Johnson, Shane D., Wim Bernasco, Kate J. Bowers, Henk Elffers, Jerry Ratcliffe, George Rengert, and Michael Townsley. 2007. “Space-Time Patterns of Risk: A Cross National Assessment of Residential Burglary Victimization.” *Journal of Quantitative Criminology* 23 (3): 201–19. <https://doi.org/10.1007/s10940-007-9025-3>.
- Kaufmann, Mareile, Simon Egbert, and Matthias Leese. 2019. “Predictive Policing and the Politics of Patterns.” *British Journal of Criminology* 59 (3): 674–92. <https://doi.org/10.1093/bjc/azy060>.
- Kikuchi, G, M Amemiya, and T Saito. 2010. “A Spatio-Temporal Analysis of Near Repeat Victimization in Japan.” *Eighth National Crime ...*. <http://www.ucl.ac.uk/scs/events/mapping-conf/conf-2010/downloads-2010/4b-kikutchi.pdf>.
- Krainski, Elias T, Finn Lindgren, Daniel Simpson, and Håvard Rue. 2019. “The R-INLA Tutorial on SPDE Models This Work Was Further Developed into the SPDE Book, See [Http://Www.r-inla.org/Spde-Book](http://www.r-inla.org/Spde-Book).” www.r-inla.org/comments-1.
- Krebs, J. R. (John R.), and N. B. (Nicholas B.) Davies. 1993. *An Introduction to Behavioural Ecology*. London [etc.] : Blackwell Scientific Publications. [https://discovery.udg.edu/iii/encore/record/C__Rb1027814__SAAn Introduction to Behavioural Ecology__Orightresult__U__X6?lang=cat](https://discovery.udg.edu/iii/encore/record/C__Rb1027814__SAAn+Introduction+to+Behavioural+Ecology__Orightresult__U__X6?lang=cat).
- Lantz, Brendan, and R. Barry Ruback. 2017. “A Networked Boost: Burglary Co-Offending and Repeat Victimization Using a Network Approach.” *Crime and Delinquency* 63 (9): 1066–90. <https://doi.org/10.1177/0011128715597695>.
- Lindgren, Finn, and Håvard Rue. 2015. “Bayesian Spatial Modelling with R - INLA.” *Journal of Statistical Software* 63 (19). <https://doi.org/10.18637/jss.v063.i19>.
- Mallesonid, Nick, Wouter Steenbeek, and Martin A Andresen. 2019. “Identifying the Appropriate Spatial Resolution for the Analysis of Crime Patterns.” <https://doi.org/10.1371/journal.pone.0218324>.
- Meijer, Albert, and Martijn Wessels. 2019. “Predictive Policing: Review of Benefits and Drawbacks.” *International Journal of Public Administration* 42 (12): 1031–39. <https://doi.org/10.1080/01900692.2019.1575664>.
- Mohler, G. O., M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. 2011. “Self-Exciting Point Process Modeling of Crime.” *Journal of the American Statistical Association* 106 (493): 100–108. <https://doi.org/10.1198/jasa.2011.ap09546>.
- “Mossos d’Esquadra - Portal Dades Obertes.” n.d. https://mossos.gencat.cat/ca/els_mossos_desquadra/indicadors_i_qualitat/dades_obertes/cataleg_dades_obertes/dades-delinqueencials/.
- Perry, Walter, Brian McInnis, Carter Price, Susan Smith, and John Hollywood. 2018. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. <https://doi.org/10.7249/rr233>.

- “R-INLA Project.” n.d. Accessed November 10, 2020. <https://www.r-inla.org/home>.
- Ratcliffe, J., R. Taylor, R. Fisher, J. Koehnlein, A. Askey, K. Thomas, J. Grasso, and K. Bethel. 2020. “The Philadelphia Predictive Policing Experiment.” *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-019-09400-2>.
- Rey, Sergio J., Elizabeth A. Mack, and Julia Koschinsky. 2012. “Exploratory Space-Time Analysis of Burglary Patterns.” *Journal of Quantitative Criminology* 28 (3): 509–31. <https://doi.org/10.1007/s10940-011-9151-9>.
- “Risk Terrain Modeling | Official Site - Home.” n.d. Accessed November 10, 2020. <https://www.riskterrainmodeling.com/>.
- Rue, Håvard, Sara Martino, and Nicolas Chopin. 2009. “Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations.” *J. R. Statist. Soc. B*. Vol. 71.
- Seidensticker, Kai, Felix Bode, and Florian Stoffel. 2018. “Predictive Policing in Germany,” 1–7. <http://nbn-resolving.de/urn:nbn:de:bsz:352-2-14sbvox1ik0z06>.
- Serra, Laura, Marc Saez, Jorge Mateu, Diego Varga, Pablo Juan, Carlos Díaz-Ávalos, and Håvard Rue. 2014. “Spatio-Temporal Log-Gaussian Cox Processes for Modelling Wildfire Occurrence: The Case of Catalonia, 1994–2008.” *Environmental and Ecological Statistics* 21 (3): 531–63. <https://doi.org/10.1007/s10651-013-0267-y>.
- Shapiro, Aaron. 2019. “Predictive Policing for Reform? Indeterminacy and Intervention in Big Data Policing.” *Surveillance and Society*. Vol. 17. <https://doi.org/10.24908/ss.v17i3/4.10410>.
- Simpson, D., J. B. Illian, F. Lindgren, S. H. Sørbye, and H. Rue. 2015. “Going off Grid: Computationally Efficient Inference for Log-Gaussian Cox Processes.” *Biometrika* 103 (1): 49–70. <https://doi.org/10.1093/biomet/asv064>.
- Simpson, Daniel, Håvard Rue, Andrea Riebler, Thiago G. Martins, and Sigrunn H. Sørbye. 2017. “Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors.” *Statistical Science* 32 (1): 1–28. <https://doi.org/10.1214/16-STS576>.
- Tobler, W R. 1970. “A Computer Movie Simulating Urban Growth in the Detroit Region.” *Economic Geography* 46: 234–40. <https://doi.org/10.2307/143141>.
- Townsley, M. 2003. “Infectious Burglaries. A Test of the Near Repeat Hypothesis.” *British Journal of Criminology* 43 (3): 615–33. <https://doi.org/10.1093/bjc/43.3.615>.
- Townsley, Michael. 2018. “Can Predictive Policing Work in Australia?” *Police Science* 2 (1). <https://doi.org/10.4018/978-1-5225-7672-3>.
- Townsley, Michael, and Aiden Sidebottom. 2010. “All Offenders Are Equal, but Some Are More Equal than Others: Variation in Journeys to Crime between Offenders.” *Criminology* 48 (3): 897–917. <https://doi.org/10.1111/j.1745-9125.2010.00205.x>.
- Tseloni, Andromachi, and Ken Pease. 2003. “Repeat Personal Victimization: ‘Boosts’ or ‘Flags’?” *British Journal of Criminology* 43 (1): 196–212. <https://doi.org/10.1093/bjc/43.1.196>.
- Wang, Zengli, Lin Liu, Hanlin Zhou, and Minxuan Lan. 2019. “Crime Geographical Displacement: Testing Its Potential Contribution to Crime Prediction.” *ISPRS International Journal of Geo-Information* 8 (9): 1–12. <https://doi.org/10.3390/ijgi8090383>.
- Wang, Zengli, and Xuejun Liu. 2017. “Analysis of Burglary Hot Spots and Near-Repeat Victimization in a Large Chinese City.” *ISPRS International Journal of Geo-Information* 6

(12): 148. <https://doi.org/10.3390/ijgi6050148>.

Wang, Zengli, and Hong Zhang. 2020. "Construction, Detection, and Interpretation of Crime Patterns over Space and Time." *ISPRS International Journal of Geo-Information* 9 (6). <https://doi.org/10.3390/ijgi9060339>.

Weathington, Katherine. 2020. "A Critical Study of Geospatial Algorithm Use in Crime Analysis and Predictive Policing," no. May.

Wollinger, Gina Rosa, Alicia König, Maximilian Querbach, and Alexandra Röhrig. 2018. "Offender Organization and Criminal Investigations with Regard to Organised Residential Burglary."

Yang, Fei. 2020. "Predictive Policing , Oxford Research Encyclopedia , Criminology and Criminal," no. June: 0–26. <https://doi.org/10.1093/acrefore/9780190264079.013.508>.

Zhang, Haifeng, and Wei Song. 2014. "Addressing Issues of Spatial Spillover Effects and Non-Stationarity in Analysis of Residential Burglary Crime." *GeoJournal* 79 (1): 89–102. <https://doi.org/10.1007/s10708-013-9481-2>.

7 Appendix A: Figures, graphics and tables

Figure 1.- Annual and monthly distributions of burglaries in the Police Regions

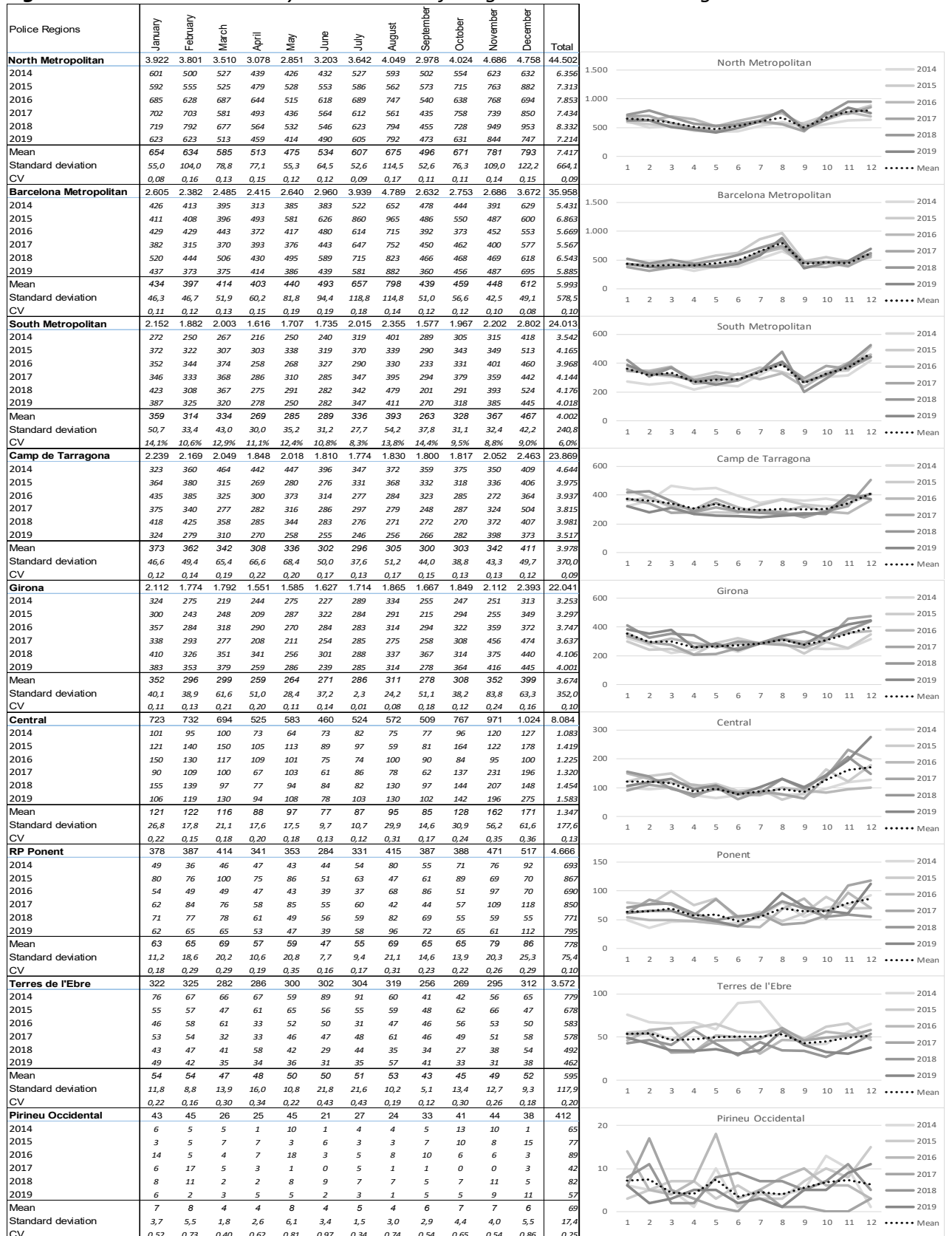
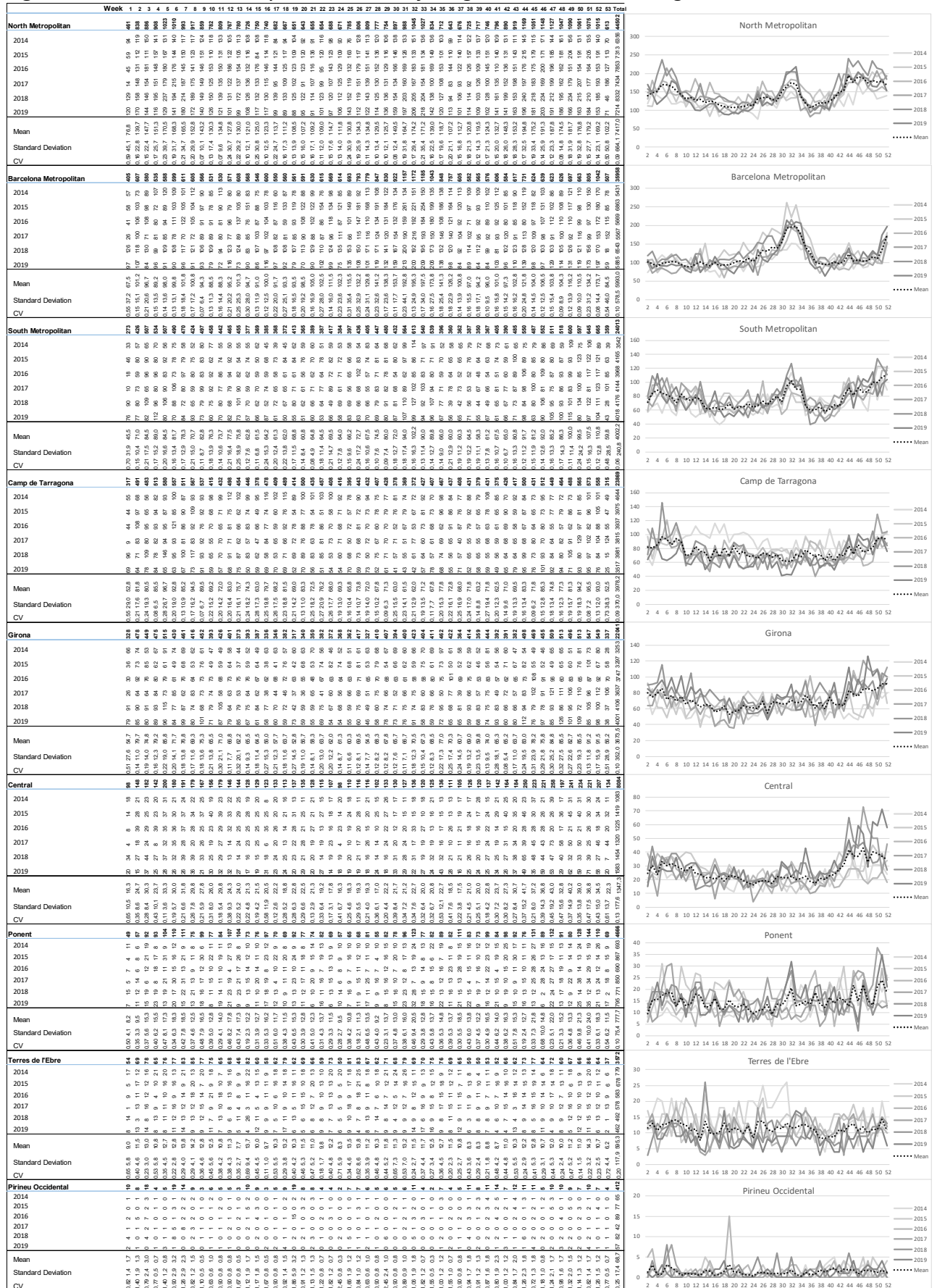
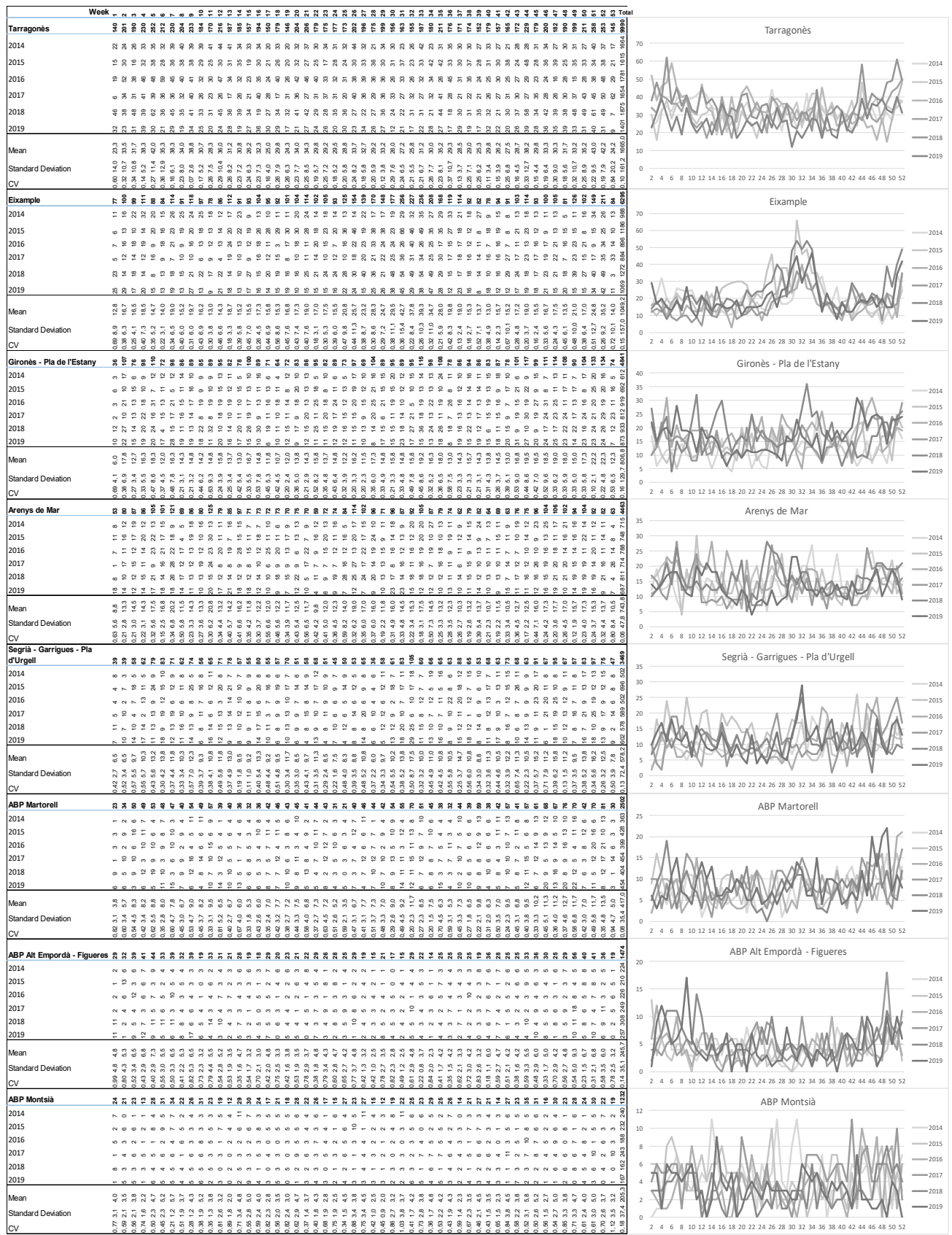


Figure 2.- Annual and weekly distributions of burglaries in the Police Regions



Source: Author's own elaboration from data provided by police (Unofficial data)

Figure 3.- Annual and monthly distributions of burglaries in some of the Basic Police Areas



Source: Author's own elaboration from a data provided by police (Unofficial data)

8 Appendix B: INLA summary results for Eixample district of Barcelona

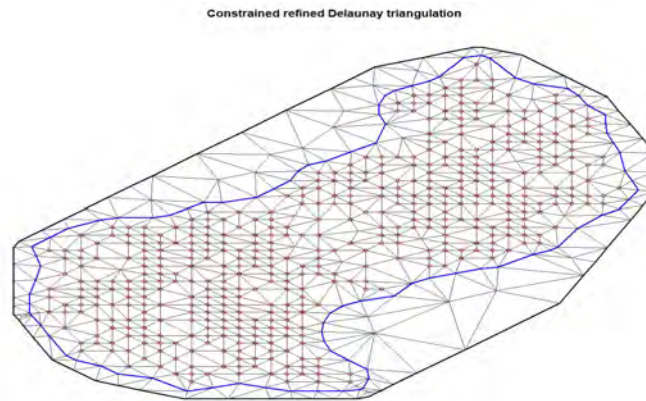
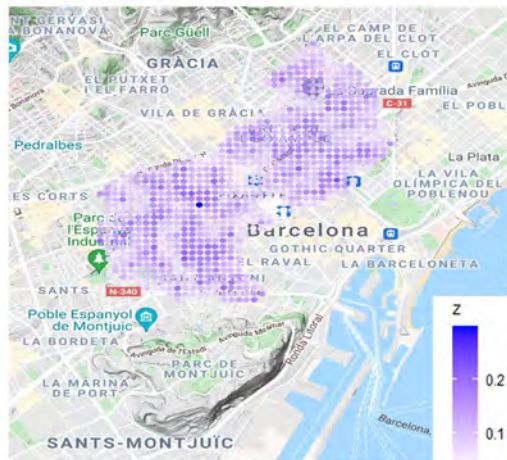


Figure 1.- Mesh for Eixample district of Barcelona (burglaries locations in red points)

Prediction: week 8



Mean of estimation parameters (week 2 to 8)

ρ	0.2
Range (l) (km)	0.11
σ	1.2

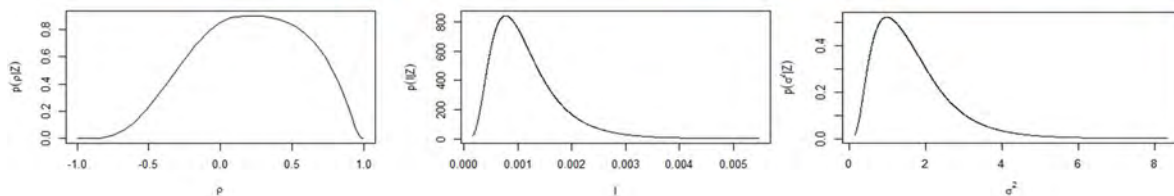


Figure 2.- Example 4.1 (Eixample district): INLA predicted burglaries (z) in 100m cells for week 8 with spatiotemporal dependence parameters and variance information (mean and probability distribution) **Predictive scheme of the example:** 6 previous weeks to adjust parameters before making a prediction for next week.

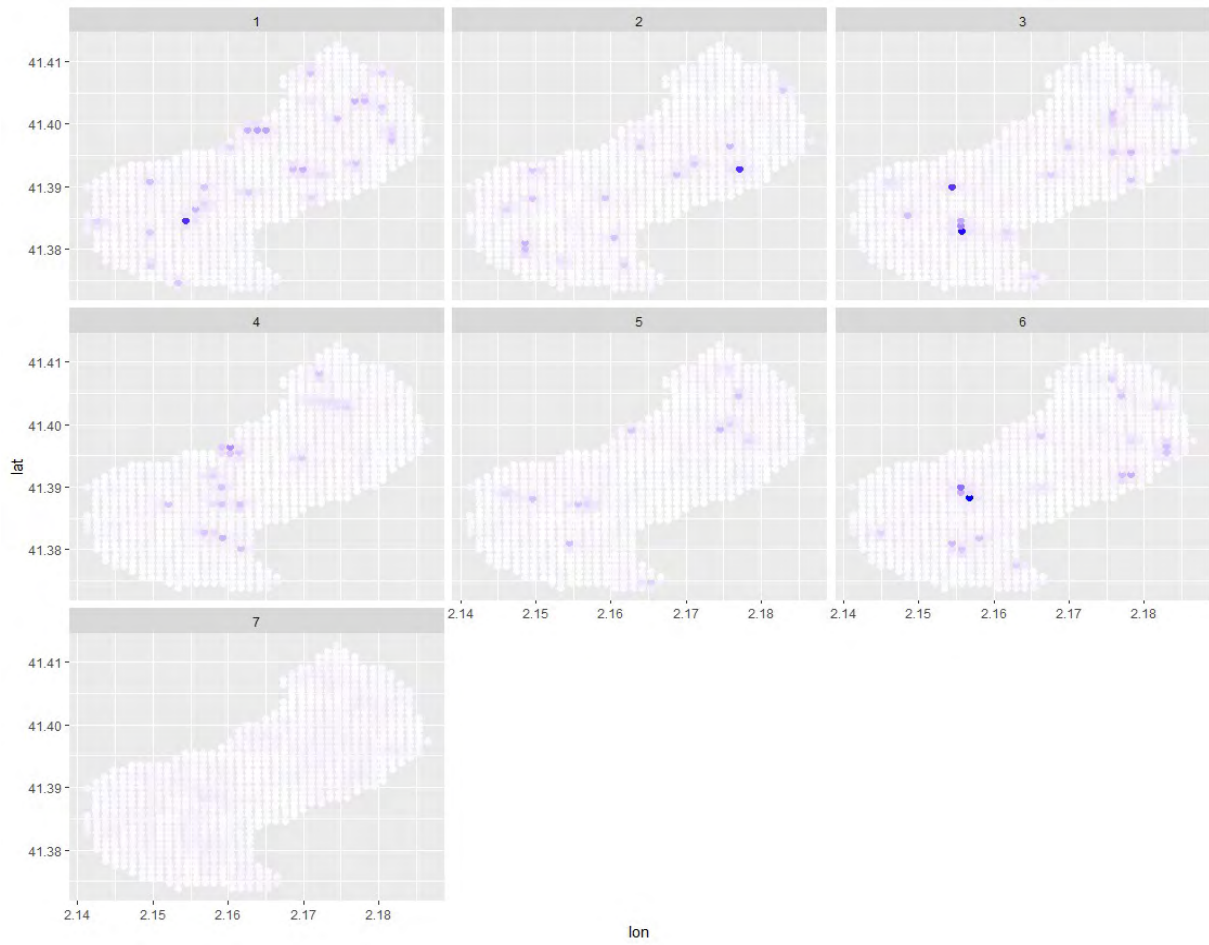


Figure 3.- Example 4.2 (Eixample district): INLA adjusted (1 to 6 graphic) and predicted (graphic 7) burglaries (z) in 100m cells for consecutive weeks.

4.3 Article 3

Burglaries constellations: applying Principal Components Analysis (PCA) to discover the hot trail left by the all-in-one spatiotemporal repeat pattern in heterogenous environments

Boqué P, Serra L, Saez M

Box 3 | Visió general de l'article 3

El que ja se sap sobre la temàtica

- En les configuracions a petita escala, el percentatge de robatoris predits correctament és relativament baix i les policies no aconsegueixen l'efecte preventiu o dissuasiu suficient.
- La conducta criminal és racional, cerca ponderar els possibles guanys amb els costos i riscos, i es guia per l'oportunitat delictiva, cosa que el fa adaptar a l'entorn.

El que aporta de nou l'estudi

- El percentatge de robatoris predits pot augmentar considerablement si es considera el patró "all-in-one" de repetició
- Els grups criminals organitzen ràtzies de robatoris en zones prèviament escollides i predictibles, on hi victimitzen durant unes setmanes.
- Les agrupacions de zones que es solen victimitzar a la vegada, anomenades "constel·lacions de robatoris", formen un conjunt fix, reduït i temporalment estable sobre el qual es poden crear models de predicció.

Burglaries constellations: applying Principal Components Analysis (PCA) to discover the hot trail left by the all-in-one spatiotemporal repeat pattern in heterogenous environments.

Abstract

Spatiotemporal patterns of repeat victimisation in extensive and heterogenous environments are more complex than those observed in large urban environments. When the homogenous space near to the first burglaries is reduced, the near repeat pattern becomes limited and the burglars move on to other areas to continue their criminal activity based on opportunistic criteria and with an attitude consistent with the main criminological theories about criminal behaviour. This strategy, described by the all-in-one spatiotemporal repeat pattern, maintains that selection of victimised areas remains stable throughout the wave of burglaries. The footprint left is a hot trail on heat maps.

Using a varimax rotation and later simplification, the Principal Components Analysis (PCA) method enabled us to approximate these hot trails through weighted groupings of cells called burglaries constellations. Linear regressions on these constellations adjusted adequately to the global and local dynamic behaviour of the weekly series of burglaries. The non-randomness of these series, grouped in waves, allowed us to make predictions.

Similarity between cells and burglaries profiles in each constellation, together with their representation and interpretation on the map, provides valuable additional information to improve both predictions about risk areas and the planning of preventive strategies in heterogenous environments.

Key words: spatiotemporal patterns, predictive policing, near repeat victimisation, all-in-one repeat pattern, heterogeneous environment, burglaries, Principal Component Analysis (PCA)

1 Introduction

Predictive policing (Perry et al. 2018) is a new tool evolving towards integration into other global approaches of the policing business such as intelligence led policing (J. Ratcliffe 2008) and oriented problem solving (Goldstein 1990; “Center for Problem-Oriented Policing | ASU Center for Problem-Oriented Policing” n.d.; Borrion et al. 2020). In this study, the emphasis is on predicting the dynamic of criminal activity in the coming days or weeks. Despite criticisms levelled at this technique in recent years, and the lack of consensus about what the best methods and applications are (EUCPN 2016; Bennett Moses and Chan 2018; Michael Townsley 2018; Seidensticker, Bode, and Stoffel 2018; Gerstner 2018; Cinelli and Manrique 2019; Meijer and Wessels 2019; Shapiro 2019; Kaufmann, Egbert, and Leese 2019; Butt et al. 2020; Egbert and Krasmann 2020; Yang 2020; J. Ratcliffe et al. 2020), efforts are continuing in this area, and there is a constant stream of new proposals.

The focus of this research is the possibilities to predict burglaries in Catalonia, where predictive models based on classic pattern of near repeat victimisation (Farrell G. and Pease 1993; M. Townsley 2003) have been found to be of limited application (REF article 2). Limitations are related to the heterogeneity of residential areas, which are generally small and scattered, limiting and apparently forcing criminals to spread their activity simultaneously over several areas either close by or further afield. This criminal strategy is described by the all-in-one spatiotemporal repeat pattern (REF article 2), which leaves a hot trail on heat maps that can be interpreted as a stable network of areas where burglaries are concentrated, called constellation in the statistical approach based on Principal Components Analysis (Pearson K. 1901; I.T. Jolliffe 2010).

The set of areas that make up each constellation will tend to activate simultaneously, or, in other words, they will be victimised in the same time periods and supposedly by the same burglars, following the “Unus pro omnibus, omnes pro uno” hypothesis (REF article 2). The rationale behind this strategy is consistent with the opportunity theories of crime (Tseloni and Pease 2003), studies of the journey to crime of European criminal gangs (Bernasco and Nieuwbeerta 2005; Michael Townsley and Sidebottom 2010; Van Deviver, Van Daele, and Van Der Beken 2015; Nee 2015; Wollinger et al. 2018), and the optimal foraging theories in general (Krebs and Davies 1993; Bernasco 2009; Johnson 2014).

It seems reasonable to think that heterogeneity of the environment would modify the awareness space (Brantingham and Brantingham 1995; Bernasco and Nieuwbeerta 2005; Curtis-Ham et al. 2020), making it a more complex “mental map” which includes areas scattered across an extensive territory that are attractive for committing burglaries, and also incorporates the access routes to these places and the suitable *modus operandi* to burgle there based on the burglars’ own experience and possibly that of other groups (Farrell and Pease 2017; Lantz and Ruback 2017). Having chosen the area, they will decide which specific residence to burgle and, if they have successfully burgled homes before, they are likely to choose one near to and/or similar to those already victimised (M. Townsley 2003; K. J. Bowers 2005), following patterns of near repeat victimisation.

This hypothetical criminal behaviour can be described following a probabilistic criterion. Having selected a group of areas to burgle over a period of weeks, after each burglary the decision to carry on committing burglaries in the same area or to change to one of the others is made based on a stochastic process. In an aggregated way and by weeks, the attractiveness of each of the areas selected as the limit of the transition probabilities or as the “long-term”

probabilities of this process can be estimated.

The study relates the “Unus pro omnibus, omnes pro uno” hypothesis, with the all-in-one repeat pattern, the hot trails left on heat maps, the probabilities of the different nodes of the constellations being victimised, and the step by step of the stochastic process which models the burglars’ decision.

The objective is to test the validity of this relation of basically equivalent concepts with respect to historical data of burglaries in Catalonia, which have a stable long-term spatial distribution and indicate which places are historically more victimised and any relevant time patterns considerations. Near repeat victimisation criteria will also be considered when estimating these areas where repeat crimes are carried out in residences and environments which are similar to those of previous burglaries. Once the constellations are obtained, the possibilities of constructing predictive models and designing preventive strategies will be analysed.

The paper is structured as follows: presentation of the data, description of the constellation method to study the spatiotemporal patterns of the burglaries, commented results, discussion about new possibilities of predicting and preventing crimes, limitations, proposed new lines of research, and conclusion.

2 Data and methodology

2.1 Data

The territory: Catalonia and the North Metropolitan Police Region

Catalonia is a country in Mediterranean Europe on the east coast of the Iberian Peninsula, occupying 5.5% of this land area. It is bordered by the Pyrenees and France to the north, the Mediterranean Sea to the east, and Spain to the south and west.

The territory is divided into 9 police regions (Figure 1), each of which includes a set of Basic Police Areas (ABP), the territory’s primary units, which are defined using geographical and policing criteria and are usually made up of several municipalities. These regions are geographically diverse with the territory of the Barcelona Metropolitan Police Region almost entirely urban, while the Eastern Pyrenees Police Region is mountainous with few, generally small villages.



Source: Author's own elaboration

Figure 1: Location of Police Regions (PR) of Catalonia

The North Metropolitan Police Region (NMPR) (Figure 2) was chosen for more detailed study. This territory concentrates almost a quarter of the total burglaries in the region and is very representative of the heterogeneity of the environment of Catalonia.



Source: Author's own elaboration using a map from the Cartographic Institute of Catalonia – ICC

Figure 2: Map of the North Metropolitan Police Region (NMPR)

This police region extends over 1,864km² and includes 2.2 million inhabitants. The territory is diverse with areas that include some of Barcelona's suburbs, densely populated towns, large industrial areas, major roads like motorways and dual carriageways, coastal residential and tourist areas, many housing estates, and medium-sized towns and rural areas with scattered villages and residences.

Crime data

The data used in this study are the reports of burglaries with forced entry made to the police in Catalonia in 2018 and 2019, provided by the Government of Catalonia Police (Mossos d'Esquadra (PG-ME)) ("Mossos d'Esquadra - Portal Dades Obertes," n.d.).

The location in UTM coordinates and a time window (generally less than a day) when the burglary is assumed to have occurred are recorded for each crime. The type of residence burgled – a flat, a house, or a farmhouse – and whether it is a first or second residence are also specified.

2.2 Methods

Burglaries coordinates are projected onto a grid of 5km per side cell and analysed by weekly series. In 2019, there were 79 burglaries in these NMPR cells, of which 52 were considered in this study because they had a weekly coefficient variation (CV) greater than 0.6 and a weekly average greater than 0.3, thus preventing an excess of global variability caused by cells with very few burglaries. In the cells considered, there were 6,692 burglaries in 2019, representing 95.8% of the total burglaries in the region (the remaining 4.2% were in the 27 cells not included in the analysis). Among the 52 cells selected for the study, 29 fulfilled the optimality criteria to predict (Boqué, Serra, and Saez 2020) in an individual way.

The Principal Components Analysis method (Pearson K. 1901; I.T. Jolliffe 2010) was used to

study the joint weekly activation patterns of the cells, (Pearson K. 1901; I.T. Jolliffe 2010), where the variables considered were the cells themselves. Rotated principal components were used, following a varimax rotation to minimise the number of cells with the most weight in each component, and the adjustment to historical data using linear regressions was analysed.

These components were simplified following criteria of reduction of variables, eliminating the cells with least weight and identifying the representative ones to obtain what we call constellations, which were then used to again analyse the adjustment to the data using linear regressions.

A descriptive analysis was made of the profile of the most representative cells of each constellation from the perspective of the environment, their proximity to each other, and burglaries.

Last, predictive capacity of constellations was analysed, studying the non-randomness of their waves and the adjustment of linear regressions on constellations to the real data.

Probabilistic approach of the all-in-one spatiotemporal repeat pattern

Assuming the all-in-one repeat pattern scenario (REF art 2), once the criminal group have decided to commit the wave of burglaries on the region, they choose the previously selected areas to burgle following a lottery such as $\begin{pmatrix} Z_1 & Z_2 & \dots & Z_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$, where Z_i represent the areas of the constellation and p_i the probabilities of being chosen, with $\sum_1^n p_i = 1$.

Naming the Poisson random variable that measures the intensity of the number of burglaries committed by a burglar or a group of burglars in a determinate time period as X , $X \sim Poiss(\lambda)$, and Y_i the number of burglaries this burglar or group of burglars will commit in area Z_i , then $Y_i / X = k$ will follow a binomial distribution of parameters k and p_i , and the number of burglaries in area Z_i will follow a Poisson distribution of a parameter equal to λp_i . Oppositely, for the additive property of the Poisson distribution, the sum of all the burglaries committed by the same group of authors and in all the previously chosen areas is the Poisson X distribution of the intensity of the burglaries committed by the group during the period $\sum Poiss(\lambda p_i) = Poiss(\lambda) \sim X$. Where there is only one targeted area Z_1 , the formula will be equivalent to that of the classic near repeat victimisation.

This approach implies that point process generated by burglaries locations will not satisfy one of the non-homogeneous Poisson processes properties, which is the independence of the number of points between two disjoint sets. In this way, correlation between distant and disjoint areas is allowed and studied, overcoming the limitations of classic near repeated victimization theory (REF art 2).

3 Results

Principal components: without rotation and with varimax rotation

The analysis of principal components allowed us to group the original 52 cells into 19 principal components that explained 80.8% of the total variance. The principal components can be represented by a map of hot cells where the value is the coefficient (or weight) of each cell in the component.

Linear regressions on the rotated components

Linear regressions of burglaries for each of the 51 full weeks of 2019 on the principal components with varimax rotation have an average coefficient R^2 of 85.2, and a standard

deviation (SD) of 7.9. In these regressions, the number of significant components each week generally varies between 5 and 11.

Figure 3 shows the graphic example of this adjustment corresponding to the first 10 full weeks of 2019.

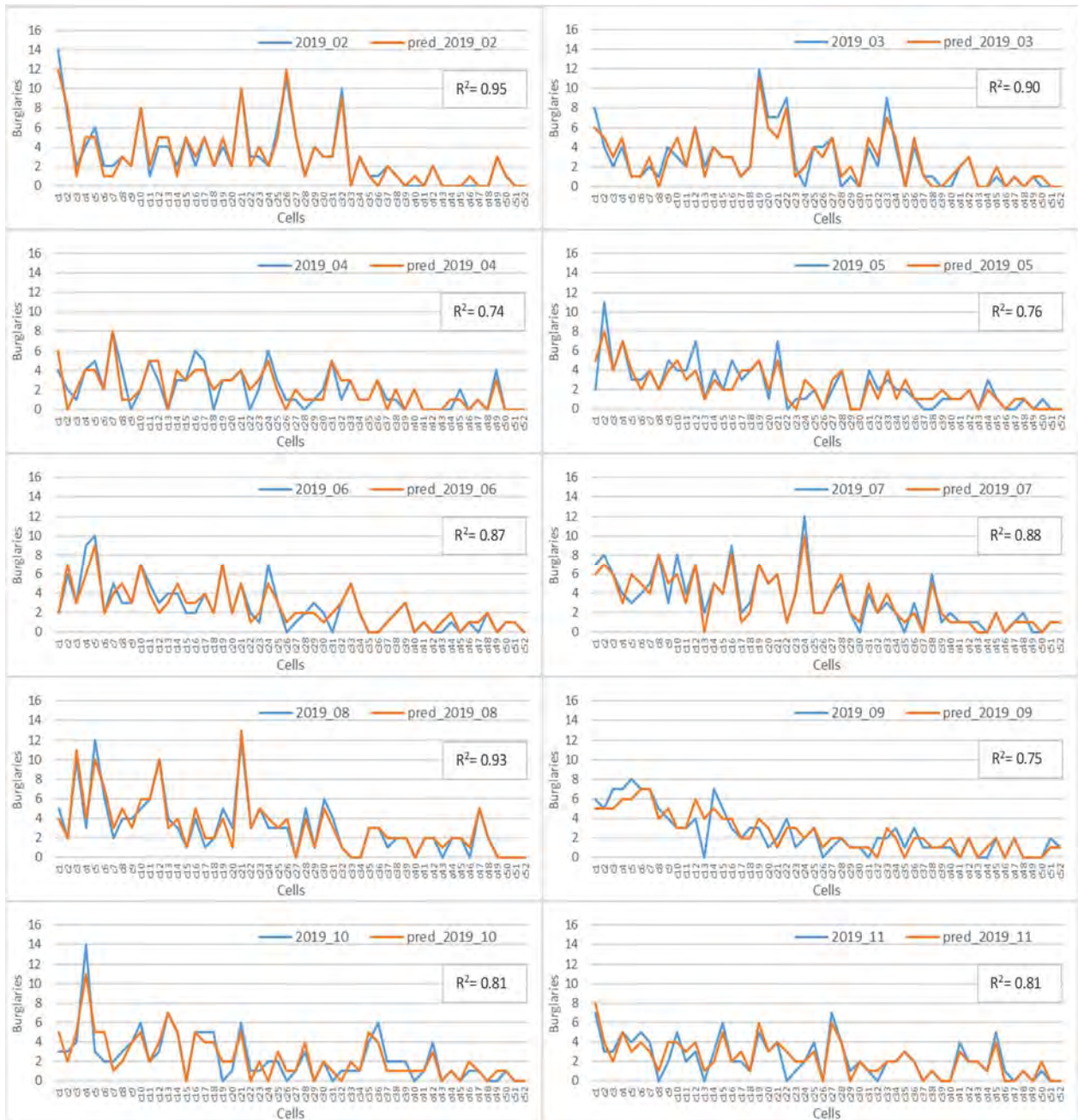


Figure 3.- Graphics of weekly regressions adjustment for the first 10 weeks of 2019

Simplification of the rotated components: burglaries constellations

Principal components obtained can be simplified, eliminating the cells with a relatively low weight. To act with uniform criteria, coefficients of each component are standardised to select the cells with coefficients greater than 0.1. With this simplification, each PC-Varimax is then characterised by a small set of between 3 and 10 cells, which concentrate the greater weights (Table 1).

PC Varimax	Number of cells with standardized weight > 0.1	Number of cells with standardized weight > 0.2	Number of cells with standardized weight > 0.3
PC 1	9	2	1
PC 2	7	3	1
PC 3	6	2	2
PC 4	4	2	1
PC 5	6	2	2
PC 6	8	2	1
PC 7	4	3	2
PC 8	7	2	1
PC 9	6	2	1
PC 10	8	4	1
PC 11	10	2	1
PC 12	5	2	2
PC 13	5	2	2
PC 14	5	2	1
PC 15	4	3	1
PC 16	5	2	1
PC 17	3	2	2
PC 18	7	2	1
PC 19	7	2	1

Table 1.- Summary of threshold weights of PC-Varimax standardized coefficients

We named as “constellations” of burglaries these principal components with varimax rotation, simplified depending on the weights. As said, they can be represented in the form of heat maps (Figure 4).

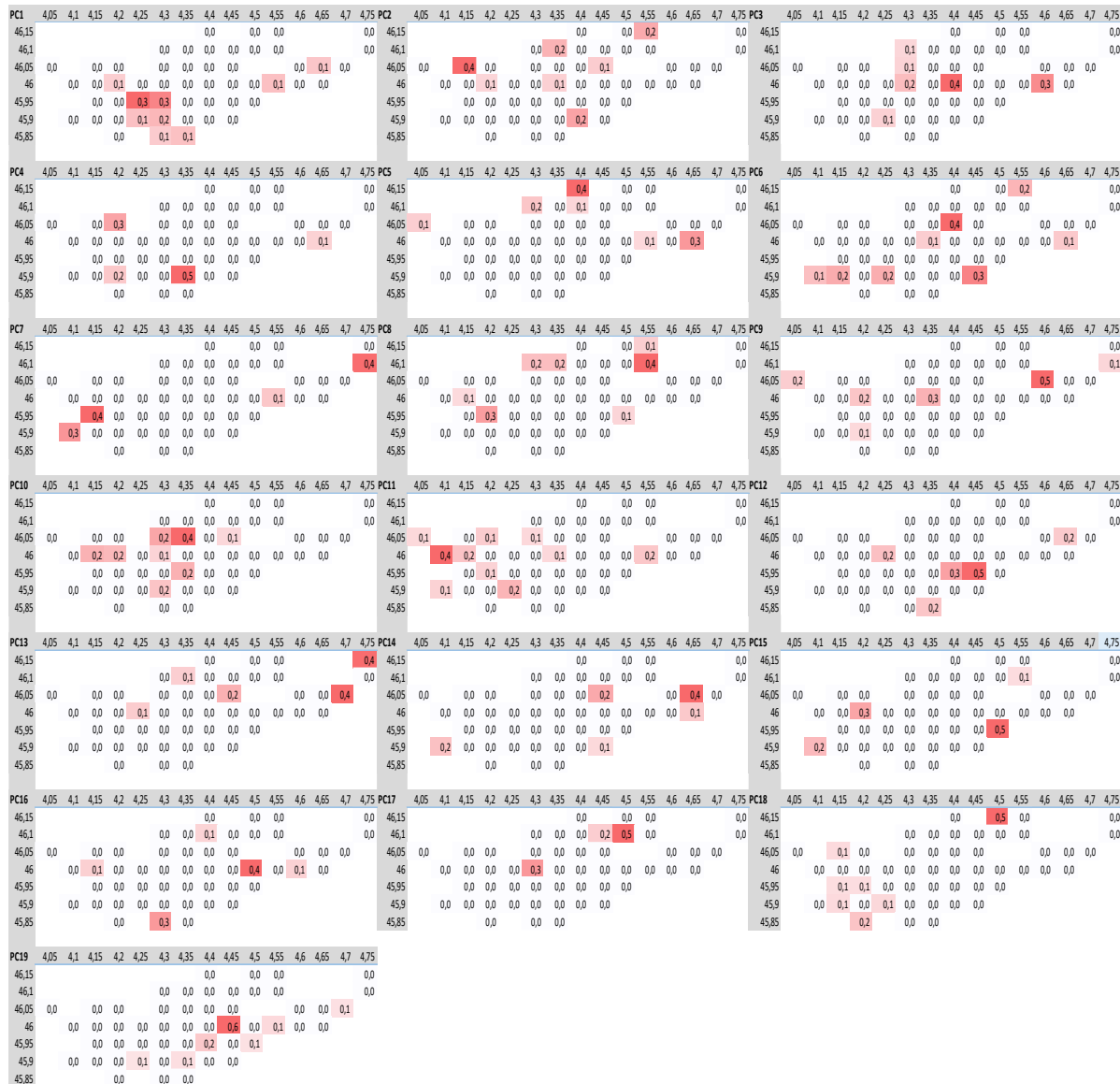


Figure 4.- Burglaries constellations representation (PC with standardized weight>0.1)

Constellations are a statistical object derived from the rotated principal components of the series of weekly burglaries in 5km per side cells. They can be interpreted as groups of cells that tend to activate jointly such that when a constellation is victimised the cells in that constellation are victimised with a probability that is proportional to the weight assigned to them by the constellation.

Each constellation is characterised by a specific set of weighted cells. A total of 34 of the 52 cells are located in more than one constellation. (Table 2). But the two or three cells with most weight in each constellation are generally the least shared ones.

Number of constellations per cell	Number of cells	%
1	18	34,6%
2	17	32,7%
3	9	17,3%
4	4	7,7%
5	3	5,8%
6	1	1,9%
Total	52	100,0%

Table 2.- Summary of overlapping cells in the constellations

Adjustment results of weekly regressions for 2019 on the 19 constellations were an average coefficient R^2 of 71.3 and a SD of 10.6. The number of significant constellations for the weekly model was generally between 1 and 4.

Temporal stability of the constellations

The 19 constellations for 2019 were used to adjust the series of weekly burglaries for 2018. The results showed that the adjustment indicator R^2 went down to an average value of 59.8, with a standard deviation of 8.7.

Series of weekly burglaries in the constellations

Assigning cells to each component meant we could calculate the sum of the series of burglaries in each of them, which was understood as the series of burglaries in the constellation. A statistical summary for 2019 is given in Table 3.

Constellations	Statistical summary of weekly burglaries in the constellations (2019)											
	#Cells	Mean	SD	Min	Max	Range	CV					
	(N)	(μ)	(σ)					(μ/N)	($\mu - \sigma/2$)	($\mu + \sigma/2$)	(max/ μ)	(max/($\mu + \sigma/2$))
Constellation 1	9	36,0	14,5	18	90	72	0,40	4,0	29	43	2,5	2,1
Constellation 2	7	16,9	8,5	4	51	47	0,50	2,4	13	21	3,0	2,4
Constellation 3	6	12,9	6,3	2	33	31	0,48	2,2	10	16	2,6	2,1
Constellation 4	4	11,7	6,5	4	32	28	0,56	2,9	8	15	2,7	2,1
Constellation 5	6	7,7	4,4	1	25	24	0,57	1,3	5	10	3,3	2,5
Constellation 6	8	19,6	7,6	5	45	40	0,39	2,5	16	23	2,3	1,9
Constellation 7	4	6,4	3,2	1	17	16	0,50	1,6	5	8	2,7	2,1
Constellation 8	7	19,6	6,4	8	45	37	0,33	2,8	16	23	2,3	2,0
Constellation 9	6	15,9	6,7	7	36	29	0,42	2,7	13	19	2,3	1,9
Constellation 10	8	23,2	7,9	14	47	33	0,34	2,9	19	27	2,0	1,7
Constellation 11	10	22,6	7,4	10	43	33	0,33	2,3	19	26	1,9	1,6
Constellation 12	5	15,0	6,2	5	36	31	0,41	3,0	12	18	2,4	2,0
Constellation 13	5	9,4	4,8	1	26	25	0,51	1,9	7	12	2,8	2,2
Constellation 14	5	4,0	2,6	0	14	14	0,64	0,8	3	5	3,5	2,6
Constellation 15	4	9,2	4,4	2	19	17	0,47	2,3	7	11	2,1	1,7
Constellation 16	5	16,3	4,9	7	33	26	0,30	3,3	14	19	2,0	1,8
Constellation 17	3	5,5	4,6	0	19	19	0,85	1,8	3	8	3,5	2,4
Constellation 18	7	17,6	6,4	8	38	30	0,36	2,5	14	21	2,2	1,8
Constellation 19	7	19,0	6,4	8	35	27	0,34	2,7	16	22	1,8	1,6

Table 3.- Statistical summary of weekly burglaries in the constellations (2019).

The last column of Table 3 shows the threshold values that can be achieved by the series of

weekly burglaries can be double the value of $\mu + \sigma/2$. The second to last column of the same table shows how the threshold value of the series in various constellations triples the average value (μ). The same table offers the value μ/N , or in other words the average number of burglaries distributed between the number of cells in each constellation.

Randomness study and constellation activation criteria

The randomness of each of the series of burglaries in the 19 constellations was studied. For a better interpretation of the results, first the series of burglaries in each of the constellations were standardised and the runs test subsequently applied. All of them were found to follow non-random grouping patterns, taking the average weekly intensity μ (Appendix, Table 1) as threshold. Activation thresholds of the constellations were generally found in the interval $(\mu - \sigma, \mu + \sigma)$ for the respective averages and standard deviations. Specifically, patterns of waves for thresholds greater than $\mu - \sigma$ were observed in 14 constellations and for thresholds greater than $\mu - \sigma/2$ in 18 of them. For thresholds greater than $\mu + \frac{\sigma}{2}$, the pattern was observed in all the constellations, and for the threshold $\mu + \sigma$ the pattern was observed in all of them except one.

Some general results regarding the intensity of this repeat pattern in terms of the weekly burglaries in 2019 are shown in Table 4.

Burglaries in week t_n	Burglaries in week t_{n+1}	Frequency (with respect to the 52 weeks of 2019)
$> \mu - \sigma/2$	$> \mu - \sigma/2$	81,6%
$> \mu$	$> \mu$	81,2%
$> \mu + \sigma/2$	$> \mu$	90,6%
$> \mu + \sigma/2$	$> \mu - \sigma/2$	95,5%

Table 4.- Some results with respect to the weekly repeat patterns of burglaries

It's seen (Table 4) that if a constellation was activated when the threshold $\mu + \sigma/2$ was surpassed, then a wave was generated that meant that an above average (μ) number of burglaries could be expected the following week in 90.6% of the cases, and in 95.5% of the cases when the threshold $\mu - \sigma/2$ was surpassed. Furthermore, if the constellation was activated when the average (μ) was surpassed, it could be expected that the average would also be surpassed the following weeks in 80% of the cases

Profile of the cells of each constellation

Regarding the study of the profile of the cells with the greatest positive weight in each constellation, it was observed that despite a certain variability, most of them had shared characteristics. Those in first constellation, for example, shared the fact that burglaries tended to take place in flats (71.9%) and terraced houses (17.1%) which were main residences and located in the outlying areas of towns. In the second constellations, the common feature was that burglaries tended to take place in detached houses (45.6%) which were main residences and located on housing estates or similar normally considered to be luxury or upper-class residential areas. The third constellation was like the second, but the locations of the victimised areas changed to be more in tourist environments, and they tended to be second residences. In the fourth, burglaries in flats once again took on the greatest importance (44.4%), but this time they were located mainly in upper class residential areas and housing estates or similar. The profile of burglaries in the fifth constellation was detached houses (55.3%) in areas further away from Barcelona and the major cities on the

outer edge of the police region. Urban burglaries predominated in the sixth constellation, taking place mostly in first residences (60%) and often in the urban centres of medium-sized towns.

4 Discussion

The fact that the weekly burglaries in constellations followed patterns of non-random waves enabled us to make predictions using the system by which they are activated when a certain threshold is surpassed. These fixed thresholds for any of the above-mentioned criteria can be adjusted throughout the year in line with seasonal variations or global trends.

Another predictive option is to look for the correlations between the 19 constellations using the spatial distribution of burglaries in the current week, which is equivalent of observing significant coefficients of the regression model with respect to constellations. The 2 or 3 which, on average, are significant each week will be those that best mould to the distribution of burglaries in the cells. In other words, the most victimised cells will activate the constellations they represent, which will be those that are assigned most weight. To this end, and following the non-random waves pattern, prediction is that the following week these constellations will still be active. The advantage of this system is that it requires neither the global intensity of burglaries in the police regions nor the possible effect or variation of the activation thresholds depending on annual seasonality or other circumstances.

Regardless of which criteria is chosen to make prediction, criticism can be levelled that they all have the failing of not activating constellations despite their having a relatively high number of burglaries, which could even be close to the average. These averages were high in some constellations, with values of 10, 15, 20, or even 36 burglaries in one of them. This would at first appear to not be an optimum way to proceed but, depending on how it is viewed, it may be just that. But the average number of burglaries distributed between the number of cells in each constellation, μ/N (Table 3), can be seen to vary between 0.8 and 4 but is generally between 2 and 3 burglaries for each 5km cell. Therefore, when constellations are not activated, the expected number of burglaries in their cells will normally be less than this average of 2 or 3, which could represent an acceptable cost if we bear in mind that the objective pursued is the deactivation of the crests of the major waves of burglaries.

The clarity of the non-random waves pattern in all the constellations supports the idea that the same burglars are operating in each of them, or at least in a large part of them. It does not seem feasible that criminal groups agree among them to commit burglaries in similar places (despite some differences), and in a systematic way, leading to the generation of waves of such well profiled as those obtained. It is more reasonable to think that the 2 or 3 active constellations per week are being victimised separately by 2 or 3 different criminal groups, one for each, implying that a single criminal group will move around the constellation for some weeks, repeatedly victimising the cells in it and following near repeat, shift, and co-occurrence patterns which, taken together, correspond to the all-in-one repeat pattern.

Similarity in the profile of residence and victimised environment in each constellation was proven, reinforcing the assumption that same burglars were at work, committing repeat burglaries in the general sense, and making relatively long journeys to look for similar targets. This behaviour is also consistent with studies on the specialised, international criminals groups detected in Europe (Bernasco and Nieuwbeerta 2005; Michael Townsley and Sidebottom 2010; Van Deviver, Van Daele, and Van Der Beken 2015; Nee 2015; Wollinger et al. 2018), who have been shown to make relatively long journeys to commit crimes following

the rationale of the theory of opportunism, weighing up the risks and the benefits and tending to go back to the places they have previously victimised.

Prediction based on constellations makes available additional helpful information to design useful preventive strategies. It is known that in each of the 2 or 3 active constellations per week there are 2 or 3 characteristic cells with most weight where the burglaries are expected to be concentrated. Therefore, the risk map of the weekly prediction will show between 4 and 9 cells with a maximum risk integrated into 2 or 3 constellations. If the same author hypothesis holds, each of these will be victimised by a different criminal gang, and the set of cells (and specific areas inside them) that are being victimised, or can potentially be victimised, will be known, as will the profile of residence and environment they are looking for. Where they are going to burgle next can also be predicted when an intense prevention operation is to be carried out in some of the higher risk cells (Guerette and Bowers 2009; J. H. Ratcliffe and Breen 2011; K. Bowers, Summers, and Poynton 2011; Wang et al. 2019). Even the roads the burglars will use can be predicted, as can the optimum points to control and detect their whereabouts. This set of information is much more comprehensive than that gained from a classic near repeat prediction, which focuses attention on a single place of limited radius in a homogenous environment.

Structuring the territory into constellations allows us to extend the areas of the police region that will be incorporated into the predictive model or algorithm. The proposed approach considers 52 of the 5km per side cells (grouped into 19 constellations), a figure which is way higher than the 29 optimum cells that would be considered in the individual predictive model (Boqué, Serra, and Saez 2020). A huge 95.8% of the regions' burglaries took place in these 52 cells in 2019, so if good preventive effectiveness can be achieved there the impact on the total number of burglaries in the region could be important.

Constellations, calculated according to weekly series of burglaries for 2019, also have predictive value for the series of burglaries in 2018, albeit less so. The average of coefficient R^2 of burglaries' weekly adjustment reduces by 11.5 points with respect to 2019. There are various explanations for this. Constellations stability could be related to similarity of the environments and/or the profile of burglaries in the cells that have most weight. When a criminal group victimises a type of cell and environment, they are likely to look for other similar environments to carry on committing burglaries, and these will then form part of the same constellation as the initial cell. According to data analysed at the NMPR, if a criminal group is looking to burgle first residences on upper-class housing estates, they will tend to encroach on the second constellation, while if they are looking to victimise flats that are main residences in urban areas of medium-sized towns, they will tend to encroach on the sixth constellation, and so on. Heterogeneity of the environments clearly has an enormous weight in the constellation's configuration.

Furthermore, criminal groups' knowledge of these environments and the speciality and type of target sought in each can evolve over time (Farrell and Pease 2017; Lantz and Ruback 2017). Some more frequently victimised areas can be abandoned for new ones, with the constellations changing accordingly. These changes can be related to long-term preventive actions, such as increased security measures in a specific area, the installation of security cameras, or pertinent alterations made to the urban landscape or the access routes, etc.

What is clear is that the relatively good adjustment obtained for 2018 using 2019-year' constellations indicates that they have a certain stability and validity over time, although it also points to the need to update them over the months or years.

Bearing in mind the previous discussion points and the fact that the constellations were able to be calculated explicitly, further support is provided for the validity of the probabilistic formulation proposed in the methods section. Criminal groups victimise constellations, so they choose a profile of environment or residence that leads them to select certain familiar areas located in certain cells with the optimum hallmarks for committing burglaries. From here, a stochastic process is set in motion whereby each cell becomes a “state” and movement between them is determined by the probabilities of transition which, in turn, are conditioned by the weight of each of the cells in the constellation. Following an all-in-one pattern, the long-term degree of victimization of the constellation's cells will correspond to its historical appeal.

The basis of the proposal is an aggregated formulation by weeks and areas. The disaggregated formulation, burglary by burglary, could be a Markov chain. Nonetheless, it is difficult to accept Markov's conditions of conditional independence, where the following state will depend solely on the current state. It seems reasonable to think that if just one burglary has been committed in an area, one or two attempts will be made to commit a repeat burglary before changing to another area. And contrarily, if 3 or 4 burglaries have already been made in an area, it seems likely that burglars will then change area. In fact, there can be many other factors influencing these decisions and complicating the estimation and modelling. On the other hand, estimating the “limit” of these decisions, which represents the average attractiveness of the areas chosen to committing burglaries in, is relatively simple using the weight assigned to each cell by the constellations. The fact of habitually carrying out predictions and preventive strategies in an aggregated way, by weeks and risk areas, facilitates this approach.

The consistency between the probabilistic approach and the results obtained make it reasonable to form a hypothesis about the behaviour model of burglars and to make a proposal to extend the concept of near repeat victimisation. It would be interesting to continue with this line of research using data from other territories and environments. It would also be useful to carry out studies based on the simulation of criminal behaviour (Birks, Townsley, and Stewart 2012; Groff, Johnson, and Thornton 2019).

Other studies have analysed spatial temporal patterns of crime from a wide perspective, including near repeat victimisation and patterns related to the journey to crime and co-occurrence (Wang et al. 2019; He et al. 2020; Wang and Zhang 2020). Most of these studies have been based on urban environments of large cities, with methodologies usually related to graph theories, data mining, or machine learning. Based on relationships between pairs of events depending on their spatial and/or temporal proximity, they obtain networks in grid cells that draw out the interactions between cells. Given that they are urban environments, the spatial configuration is usually more detailed, with cells no larger than 500 square meters, while the temporal window is always around 7 days with no need for this period to coincide with natural weeks. These studies also start from the hypothesis that criminal groups generate these new victimisation patterns to adapt to crime opportunities and circumstances which force them to make journeys of sometimes many kilometres to avoid police actions and be able to continue burgling.

It is assumed that results obtained with previous methodologies and the one proposed in this study will obtain similar results. Potential value of constellations approach lies in the simplicity of their calculation using standard and free statistical tools, and the fact that it is related to a broader theoretical approach and supported with new hypotheses about spatial temporal patterns and criminal behaviour.

At this point in the discussion, the only remaining issue is putting the proposal into practice to test its real preventive effectiveness, even though this is easier said than done. Any police officer with the right tools and own resources can apply the constellations method, adapting it to their particular territorial and organisational circumstances. This is an open approach in which predictive model is only a relatively small part of the prediction, which needs to be completed by police. The aim of the predictive output of constellations is not exactitude but a useful global diagnosis to guide prevention dynamic on a weekly basis. This predictive information needs to be integrated into the information derived from intelligence led policing to design the best possible preventive strategy.

All going to plan, we would expect this approach to have an important dissuasive effect. Prevention in terms of constellations implies that criminal groups could not only not act in the cells with the most activity, but neither could they act in similar cells of the same constellation. The possibility of carrying out near repeat burglaries, also in the wider sense, would be eliminated, and not only in the geographical proximity. In the knowledge that near repeat is the keystone to achieving efficiency in the burglary business, victimisations would be expected to fall dramatically. The risk of being detected and the cost in logistics and knowledge about each burglary would increase, making all becoming “initial” or “priors” and located outside the awareness space (in the general sense) of criminal groups. There would be the very optimistic possibility of burglars abandoning their activity, although this may be considered an unrealistic or unachievable goal. It is, nevertheless, the target of some “strike hard” initiatives (Chen and Kurland 2020).

There is enormous unpredictability in the reactions of criminals to ambitious preventive strategies like the one proposed. What would burglars do if their crime pattern were made impossible to implement, and they had to stop carrying out near repeats either in nearby areas or in similar areas further afield? Clearly, burglaries would no longer be predictable, at least for the time being, which should be a point of reflection for all research on predictive policing. The purpose of modelling historical data is to discover the patterns to be broken, which, if successful, to some extent invalidates the predictive model, forcing us to have to reconsider it.

Limitations

This study has some limitations. Crime data comes from events reported to police and so there is an unknown black figure, even though in the case of burglaries it is relatively low compared to other crime typologies. Some of the unreported burglaries may be attempted or less serious ones and since their spatiotemporal distributions can be assumed to run parallel to the reported ones, this limitation does not interfere with the main conclusions. There is also a limitation related to the reported burglaries and concerns the time when they occurred. The day and the time window (morning, afternoon, or night) is usually known, but sometimes it can be as much as several weeks. The percentage of the latter cases is low and the impact on the analysis is reduced by considering weekly time intervals.

Another possible limitation of this study is not having incorporated layers of geographical information in the configuration of the constellations, for which only the place and time data of burglaries are considered. The study of similarities in the profiles of the residences was carried out afterwards, and the geographical environment was interpreted based on how it was represented on maps and local knowledge of the territory. The decision to do so was made consciously. Valid geographic layers configuration information in terms of policing and crime should be done ad hoc, from a perspective of crime concentration and opportunity,

and the best way to approach this is via historical data (Bogucki, Milczek, and Miziula 2020).

5 Conclusions

Committing burglaries in a systematic and planned way in different areas at the same time seems to be the best explanation for the spatiotemporal distribution of the burglaries that take place in Catalonia and, by extension, in similar heterogeneous spaces like those found in a large part of Southern Europe. In these environments, the all-in-one repeat has been proposed as the reference pattern, which extends the concept of repeat burglaries in the proximity to repeat in distant and similar areas previously chosen by criminal groups.

The groupings of areas that tend to be victimised at the same time were named as constellations and they represent burglars' usual preferences in terms of where they carry out their crimes. Clearly non-random, easily predictable waves of burglaries are distributed in the constellations.

The existence of constellations may be surprising since it implies that burglars repeatedly choose the same crime "programme", the same places, the same type of target, and the same routes on their journeys to crime. Furthermore, these crime "programmes" are the same or very similar for all the criminal groups operating in the territory. Therefore, external factors, basically the environment and the type of residence, and factors of rationality of criminal behaviour, justify the existence of constellations in heterogeneous spaces.

It was proposed to model burglars' decisions in this type of environment by means of a stochastic process (Figure 2), following the all-in-one repeat pattern. Observed in an aggregated way by weeks and areas, it tends to be a simple probabilistic model, in the form of lotteries, that correspond to the constellations. This approach allows us to consider the proposed model as a natural extension of the near repeat victimisation model, also from a probabilistic point of view.

The predictive output, enriched by the interpretation of constellations represented on the map, provides useful information for designing preventive strategies for the possible routes the burglars will use and the profile of residences they want to burgle. By adding police information, specific areas that require protection can be demarcated.

Results obtained prompt us to seriously consider the possibility of breaking the criminal groups' operating patterns in these types of environments. However, there are few references and a lot of uncertainties to this effect, and its successful achievement would also have the ironic and adverse effect of breaking the predictive models. However, it seems logical that both the models and the methodologies used to make predictions must be flexible and they need to be a set of analysis tools rather than inflexible, closed approaches. The police essentially need to participate in this analytical process to ensure that common sense prevails and the results are accurately interpreted to guarantee the usefulness of the predictions for planning preventive strategies.

6 References

- Bennett Moses, Lyria, and Janet Chan. 2018. "Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability." *Policing and Society* 28 (7): 806–22. <https://doi.org/10.1080/10439463.2016.1253695>.
- Bernasco, Wim. 2009. "FORAGING STRATEGIES OF HOMO CRIMINALIS: LESSONS FROM BEHAVIORAL ECOLOGY." *CRIME PATTERNS AND ANALYSIS* 2 (5). <http://wimbernasco.nl/Manuscripts/English/Bernasco2009a.pdf>.
- Bernasco, Wim, and Paul Nieuwebeerta. 2005. "How Do Residential Burglars Select Target Areas? : A New Approach to the Analysis of Criminal Location Choice." *British Journal of Criminology* 45 (3): 296–315. <https://doi.org/10.1093/bjc/azh070>.
- Birks, Daniel, Michael Townsley, and Anna Stewart. 2012. "Generative Explanations of Crime: Using Simulation to Test Criminological Theory." *Criminology* 50 (1): 221–54. <https://doi.org/10.1111/j.1745-9125.2011.00258.x>.
- Bogucki, Robert, Jan Kanty Milczek, and Patryk Miziula. 2020. "A Simple Crime Hotspot Forecasting Algorithm." *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020* 21: 23–26. <https://doi.org/10.15439/2020F5>.
- Boqué, Pere, Laura Serra, and Marc Saez. 2020. "'Surfing' Burglaries with Forced Entry in Catalonia: Large-Scale Testing of near Repeat Victimization Theory." *European Journal of Criminology*, November, 147737082096810. <https://doi.org/10.1177/1477370820968102>.
- Borrion, Hervé, Paul Ekblom, Dalal Alrajeh, Aiduan Li Borrion, Aidan Keane, Daniel Koch, Timothy Mitchener-Nissen, and Sonia Toubaline. 2020. "The Problem with Crime Problem-Solving: Towards a Second Generation Pop?" *British Journal of Criminology* 60 (1): 219–40. <https://doi.org/10.1093/bjc/azz029>.
- Bowers, K. J. 2005. "Domestic Burglary Repeats and Space-Time Clusters: The Dimensions of Risk." *European Journal of Criminology* 2 (1): 67–92. <https://doi.org/10.1177/1477370805048631>.
- Bowers, Kate, Lucía Summers, and Suzanne Poynton. 2011. "Spatial Displacement and Diffusion of Benefits among Geographically Focused Policing Initiatives." <https://doi.org/10.4073/csr.20011.3>.
- Brantingham, Patricia, and Paul Brantingham. 1995. "Criminality of Place - Crime Generators and Crime Attractors." *European Journal on Criminal Policy and Research* 3 (3): 5–26. <https://doi.org/10.1007/BF02242925>.
- Butt, Umair Muneer, Sukumar Letchmunan, Fadratul Hafinaz Hassan, Mubashir Ali, Anees Baqir, and Hafiz Husnain Raza Sherazi. 2020. "Spatio-Temporal Crime HotSpot Detection and Prediction: A Systematic Literature Review." *IEEE Access* 8: 166553–74. <https://doi.org/10.1109/access.2020.3022808>.
- "Center for Problem-Oriented Policing | ASU Center for Problem-Oriented Policing." n.d. Accessed November 10, 2020. <https://popcenter.asu.edu/>.
- Chen, Peng, and Justin Kurland. 2020. "The Impact of 'Strike Hard' on Repeat and near-Repeat Residential Burglary in Beijing." *ISPRS International Journal of Geo-Information* 9 (3). <https://doi.org/10.3390/ijgi9030150>.
- Cinelli, Virginia, and Alberto Manrique. 2019. "El Uso de Programas de Análisis Predictivo En La Inteligencia Policial: Una Comparativa Europea." *Revista de Estudios En Seguridad Internacional* 5 (2): 1–20. <https://doi.org/10.18847/1.10.1>.

- Curtis-Ham, Sophie, Wim Bernasco, Oleg N. Medvedev, and Devon Polaschek. 2020. "A Framework for Estimating Crime Location Choice Based on Awareness Space." *Crime Science* 9 (1): 1–14. <https://doi.org/10.1186/s40163-020-00132-7>.
- Deviver, Christophe Van, Stijn Van Daele, and Tom Van Der Beken. 2015. "What Makes Long Crime Trips Worth Undertaking Balancing Costs and Benefits in Burglars Journey to Crime." *British Journal of Criminology* 55 (2): 399–420. <https://doi.org/10.1093/bjc/azu078>.
- Egbert, Simon, and Susanne Krasmann. 2020. "Predictive Policing: Not yet, but Soon Preemptive?" *Policing and Society* 30 (8): 905–19. <https://doi.org/10.1080/10439463.2019.1611821>.
- EUCPN. 2016. "Recommendation Paper: Predictive Policing," 13. www.eucpn.org.
- Farrell G., and K Pease. 1993. "Once Bitten, Twice Bitten: Repeat Victimisation and Its Implications for Crime Prevention." *Police Research Group, Crime Prevention Unit Paper 46*, no. 46: 38. <http://hdl.handle.net/2134/2149%5Cnhttp://core.ac.uk/download/pdf/67259.pdf>.
- Farrell, Graham, and Ken Pease. 2017. "Preventing Repeat and near Repeat Crime Concentrations." In *Handbook of Crime Prevention and Community Safety*. <https://doi.org/10.4324/9781315724393>.
- Gerstner, Dominik. 2018. "Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany." *European Journal for Security Research* 3 (2): 115–38. <https://doi.org/10.1007/s41125-018-0033-0>.
- Goldstein, H. 1990. *Problem-Oriented Policing*. McGraw-Hill. <https://books.google.es/books?id=WCqkQgAACAAJ>.
- Groff, Elizabeth R., Shane D. Johnson, and Amy Thornton. 2019. "State of the Art in Agent-Based Modeling of Urban Crime: An Overview." *Journal of Quantitative Criminology* 35 (1): 155–93. <https://doi.org/10.1007/s10940-018-9376-y>.
- Guerette, Rob T., and Kate J. Bowers. 2009. "Assessing the Extent of Crime Displacement and Diffusion of Benefits: A Review of Situational Crime Prevention Evaluations." *Criminology* 47 (4): 1331–68. <https://doi.org/10.1111/j.1745-9125.2009.00177.x>.
- He, Zhanjun, Liufeng Tao, Zhong Xie, and Chong Xu. 2020. "Discovering Spatial Interaction Patterns of near Repeat Crime by Spatial Association Rules Mining." *Scientific Reports* 10 (1): 1–12. <https://doi.org/10.1038/s41598-020-74248-w>.
- I.T. Jolliffe. 2010. "Principal Components Analysis." *Springer Series in Statistics*, 374–77. <https://doi.org/10.1016/B978-0-08-044894-7.01358-0>.
- Johnson, Shane D. 2014. "How Do Offenders Choose Where to Offend? Perspectives from Animal Foraging." *Legal and Criminological Psychology* 19 (2): 193–210. <https://doi.org/10.1111/lcrp.12061>.
- Kaufmann, Mareile, Simon Egbert, and Matthias Leese. 2019. "Predictive Policing and the Politics of Patterns." *British Journal of Criminology* 59 (3): 674–92. <https://doi.org/10.1093/bjc/azy060>.
- Krebs, J. R. (John R.), and N. B. (Nicholas B.) Davies. 1993. *An Introduction to Behavioural Ecology*. London [etc.] : Blackwell Scientific Publications. https://discovery.udg.edu/iii/encore/record/C__Rb1027814__SAAn Introduction to Behavioural Ecology__Orightrresult__U__X6?lang=cat.
- Lantz, Brendan, and R. Barry Ruback. 2017. "A Networked Boost: Burglary Co-Offending and Repeat Victimization Using a Network Approach." *Crime and Delinquency* 63 (9):

- 1066–90. <https://doi.org/10.1177/0011128715597695>.
- Meijer, Albert, and Martijn Wessels. 2019. "Predictive Policing: Review of Benefits and Drawbacks." *International Journal of Public Administration* 42 (12): 1031–39. <https://doi.org/10.1080/01900692.2019.1575664>.
- "Mossos d'Esquadra - Portal Dades Obertes." n.d. https://mossos.gencat.cat/ca/els_mossos_desquadra/indicadors_i_qualitat/dades_obertes/catalog_dades_obertes/dades-delinqueencials/.
- Nee, Claire. 2015. "Understanding Expertise in Burglars: From Pre-Conscious Scanning to Action and Beyond." *Aggression and Violent Behavior* 20: 53–61. <https://doi.org/https://doi.org/10.1016/j.avb.2014.12.006>.
- Pearson K. 1901. "Pearson, K. 1901. On Lines and Planes of Closest Fit to Systems of Points in Space." *Philosophical Magazine* 2: 559–72.
- Perry, Walter, Brian McInnis, Carter Price, Susan Smith, and John Hollywood. 2018. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. <https://doi.org/10.7249/rr233>.
- Ratcliffe, J., R. Taylor, R. Fisher, J. Koehnlein, A. Askey, K. Thomas, J. Grasso, and K. Bethel. 2020. "The Philadelphia Predictive Policing Experiment." *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-019-09400-2>.
- Ratcliffe, J. 2008. *Intelligence-Led Policing*. Willan. <https://books.google.es/books?id=iPxLvGAAAJ>.
- Ratcliffe, Jerry H., and Clairissa Breen. 2011. "Crime Diffusion and Displacement: Measuring the Side Effects of Police Operations." *Professional Geographer* 63 (2): 230–43. <https://doi.org/10.1080/00330124.2010.547154>.
- Seidensticker, Kai, Felix Bode, and Florian Stoffel. 2018. "Predictive Policing in Germany," 1–7. <http://nbn-resolving.de/urn:nbn:de:bsz:352-2-14sbvox1ik0z06>.
- Shapiro, Aaron. 2019. "Predictive Policing for Reform? Indeterminacy and Intervention in Big Data Policing." *Surveillance and Society*. Vol. 17. <https://doi.org/10.24908/ss.v17i3/4.10410>.
- Townsley, M. 2003. "Infectious Burglaries. A Test of the Near Repeat Hypothesis." *British Journal of Criminology* 43 (3): 615–33. <https://doi.org/10.1093/bjc/43.3.615>.
- Townsley, Michael. 2018. "Can Predictive Policing Work in Australia?" *Police Science* 2 (1). <https://doi.org/10.4018/978-1-5225-7672-3>.
- Townsley, Michael, and Aiden Sidebottom. 2010. "All Offenders Are Equal, but Some Are More Equal than Others: Variation in Journeys to Crime between Offenders." *Criminology* 48 (3): 897–917. <https://doi.org/10.1111/j.1745-9125.2010.00205.x>.
- Tseloni, Andromachi, and Ken Pease. 2003. "Repeat Personal Victimization: 'Boosts' or 'Flags'?" *British Journal of Criminology* 43 (1): 196–212. <https://doi.org/10.1093/bjc/43.1.196>.
- Wang, Zengli, Lin Liu, Hanlin Zhou, and Minxuan Lan. 2019. "Crime Geographical Displacement: Testing Its Potential Contribution to Crime Prediction." *ISPRS International Journal of Geo-Information* 8 (9): 1–12. <https://doi.org/10.3390/ijgi8090383>.
- Wang, Zengli, and Hong Zhang. 2020. "Construction, Detection, and Interpretation of Crime Patterns over Space and Time." *ISPRS International Journal of Geo-Information* 9 (6). <https://doi.org/10.3390/ijgi9060339>.
- Wollinger, Gina Rosa, Alicia König, Maximilian Querbach, and Alexandra Röhrig. 2018.

“Offender Organization and Criminal Investigations with Regard to Organised Residential Burglary.”

Yang, Fei. 2020. “Predictive Policing , Oxford Research Encyclopedia , Criminology and Criminal,” no. June: 0–26. <https://doi.org/10.1093/acrefore/9780190264079.013.508>.

7 Appendix

Table 1: Runs test for the standardized series of weekly burglaries in the constellations

Runs test	Const. 1	Const. 2	Const. 3	Const. 4	Const. 5	Const. 6	Const. 7	Const. 8	Const. 9	Const. 10	Const. 11	Const. 12	Const. 13	Const. 14	Const. 15	Const. 16	Const. 17	Const. 18	Const. 19	
Total cases	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260	260
Test Value ($\mu - 1.5\sigma$)	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50	-1,50
Number of runs	15	18	7	13	7	23	13	9	15	8	20	15	13	1	23	17	1	9	25	
Z	-4,50	-1,92	-4,10	0,40	0,21	0,73	-1,99	-3,13	0,46	-1,92	-2,80	-1,62	-1,99		-5,61	0,53		-3,13	0,80	
Sig. asympt. (bilateral)	,000	,055	,000	,691	,837	,466	,046	,002	,644	,055	,005	,106	,046		,000	,597		,002	,426	
Test Value ($\mu - \sigma$)	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00
Number of runs	41	48	54	55	54	48	52	61	58	50	52	69	51	77	67	65	47	47	51	
Z	-4,35	-3,32	-1,30	-3,01	-3,52	0,07	-3,74	-2,66	-4,21	-3,10	-4,69	-1,40	-2,15	-1,81	-2,08	-0,56	-0,80	-1,74	-2,83	
Sig. asympt. (bilateral)	,000	,001	,193	,003	,000	,945	,000	,008	,000	,002	,000	,162	,032	,070	,038	,574	,421	,081	,005	
Test Value ($\mu - 0.5\sigma$)	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50	-0,50
Number of runs	74	78	86	71	93	98	113	90	72	91	86	91	90	92	107	98	102	75	79	
Z	-5,44	-5,22	-4,30	-6,39	-3,83	-2,54	-1,22	-3,67	-6,69	-4,30	-4,49	-4,20	-2,23	-4,17	-1,68	-2,11	-2,74	-5,71	-5,33	
Sig. asympt. (bilateral)	,000	,000	,000	,000	,000	,011	,224	,000	,000	,000	,000	,000	,026	,000	,093	,035	,006	,000	,000	
Test Value (μ)	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Number of runs	84	88	89	76	93	105	117	86	81	83	87	99	103	108	106	100	92	91	87	
Z	-5,68	-5,16	-4,87	-6,77	-4,50	-3,19	-1,67	-5,50	-6,06	-5,96	-5,24	-3,83	-2,86	-2,84	-2,95	-3,55	-3,67	-4,61	-5,47	
Sig. asympt. (bilateral)	,000	,000	,000	,000	,000	,001	,095	,000	,000	,000	,000	,000	,004	,005	,003	,000	,000	,000	,000	
Test Value ($\mu + 0.5\sigma$)	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	
Number of runs	75	69	59	68	81	77	81	67	69	81	71	83	83	79	90	86	68	77	78	
Z	-3,50	-5,52	-5,89	-5,28	-3,29	-3,30	-4,05	-5,64	-6,56	-3,75	-4,79	-3,09	-2,97	-1,82	-1,84	-3,67	-3,68	-3,82	-3,53	
Sig. asympt. (bilateral)	,000	,000	,000	,000	,001	,001	,000	,000	,000	,000	,000	,002	,003	,069	,066	,000	,000	,000	,000	
Test Value ($\mu + \sigma$)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	
Number of runs	49	41	53	46	49	57	53	53	43	49	49	47	51	59	70	58	54	45	62	
Z	-3,37	-6,43	-4,01	-4,72	-4,72	-3,32	-3,50	-2,62	-6,35	-3,95	-3,67	-2,93	-1,77	-1,05	-2,57	-3,09	-2,97	-4,44	-2,92	
Sig. asympt. (bilateral)	,001	,000	,000	,000	,000	,001	,000	,009	,000	,000	,000	,003	,076	,292	,010	,002	,003	,000	,003	
Test Value ($\mu + 1.5\sigma$)	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	1,50	
Number of runs	29	25	35	22	23	35	31	33	39	39	33	25	33	33	42	29	29	29	24	
Z	-3,35	-4,64	-2,54	-7,80	-5,03	-3,08	-2,42	-1,49	-2,08	-2,08	-1,49	-3,29	-1,49	-2,79	-0,96	-1,11	-4,97	-4,97	-2,16	
Sig. asympt. (bilateral)	,001	,000	,011	,000	,000	,002	,015	,135	,037	,037	,135	,001	,135	,005	,337	,268	,000	,000	,031	

5 DISCUSSIÓ

La discussió d'aquesta tesi seguirà l'ordre cronològic dels tres articles presentats. Després de fer unes valoracions generals respecte als objectius i hipòtesis de la recerca, es discutirà els coneixements previs a l'inici de la investigació i es seguirà amb la discussió global dels resultats que s'han obtingut amb cadascuna de les recerques parcials posteriors.

La primera valoració general és que s'ha demostrat que predir els robatoris amb força a domicili a Catalunya és possible. Es poden anticipar les zones i els moments on hi haurà un risc més elevat de robatoris. Però aquesta predicció està lluny de tenir una configuració espai-temps fina o a micro-escala. En general, els patrons repetitius emergeixen en zones extenses, com les cel·les de 5km de costat, i també en la unió de vàries cel·les o zones disjunctes. En aquests espais, es pot predir el risc de robatoris per a la setmana següent amb una probabilitat alta. Per tant, s'assoleix el primer objectiu específic de la recerca:

Objectiu específic 1: *Detectar les zones amb més probabilitat de repetició en l'espai i el temps i acotar-les territorialment (zona potencial de repetició) i temporalment (interval temporal de la repetició).*

Aquestes prediccions, ponderen les zones de més o menys risc i informen del nombre esperat de robatoris o del nombre mínim de robatoris que hi haurà i amb quina probabilitat. Per tant, també s'assoleix el segon objectiu específic de la recerca:

Objectiu específic 2: *Classificar aquestes zones segons el nombre i la probabilitat d'aquestes repeticions i segons l'acotació en l'espai i el temps, de manera que es puguin prioritzar els llocs on la prevenció policial pugui ser més efectiva.*

Per contra, no s'ha arribat a construir explícitament un model de prevenció policial que optimitzi la gestió dels recursos humans segons aquestes prediccions, que era el tercer objectiu específic. En el Pla de Recerca s'havia previst la possibilitat de fer una prova pilot basada en el *predictive policing* en funció dels patrons espai-temps observats. Tot i l'interès per part del cos de Mossos d'Esquadra, finalment no s'ha donat l'oportunitat de fer-la. Però els resultats obtinguts permetran portar-la a terme així que hi hagi els recursos i l'ocasió. En tot moment, l'orientació de la recerca ha estat pràctica, cercant solucions aplicables per la policia i en consonància amb les seves estratègies preventives habituals. També s'han donat indicacions concretes per coordinar el coneixement policial amb el que es derivi de les sistemàtiques predictives proposades. Els patrons observats, més enllà de la predicció espaciotemporal aporten informació addicional sobre el perfil dels robatoris i de les zones de risc, i també del moviment dels grups criminals per desplaçar-se entre aquestes zones. Això permetrà millorar les estratègies de prevenció policial. De manera que es pot dir que s'ha assolit gran part de l'objectiu principal de la recerca:

Objectiu principal: Estudiar els patrons de repetició dels robatoris amb força a domicili a Catalunya per millorar-ne les estratègies de prevenció policial.

De forma més general, el títol de la tesi, “Fonaments dels models matemàtics i estadístics per a la predicció dels robatoris amb força a domicili a Catalunya”, expressa el moll de l’os o l’esperit de la recerca, estudiar els possibles mecanismes, relacions o estructures de les dades relatives als robatoris amb força a domicili que fan possible construir models i metodologies matemàtiques i estadístiques per a la predicció delictiva.

A l’inici de la recerca, aquests fonaments es basaven en la victimització quasi repetida clàssica, l’opció més clara i utilitzada arreu per predir. Però calia comprovar que aquest patró s’observava també a Catalunya. Tanmateix, el Test de Knox (G. Knox 1963; E. G. Knox and Bartlett 1964; Ornstein and Hammond 2017; Briz-Redón, Martínez-Ruiz, and Montes 2020) aportava resultats incerts, i sorgien dubtes sobre la seva idoneïtat d’aplicar-lo en l’entorn geogràfic català. Prenent qualsevol part del territori s’observava una heterogeneïtat que no quadrava amb el disseny implícitament homogeni del test. Les bandes d’ample espai-temps per a fer els càlculs de l’índex de Knox contenien espais residencials diversos i combinats amb espais no residencials. Per posar un exemple, en una zona relativament petita es podien observar barris de pisos que limitaven amb carrers de cases adossades seguits d’una urbanització i un bosc. I aquest tipus de diversitat era el patró comú. Tornant al punt de vista teòric, les repeticions s’haurien de produir en domicilis semblants, però a Catalunya l’entorn pròxim i semblant quedava molt reduït. I els amplex de banda del test de Knox s’havien d’anar canviant en funció de la zona, sense poder establir paràmetres generals de referència pels models predictius.

Aquest primer contacte amb els patrons de repetició espaciotemporals van fer allunyar la recerca del primer model de referència, el model de Short (Short et al. 2008). Per contra, s’intuïa la necessitat de fer estudis estadístics previs que permetessin entendre millor la mecànica dels patrons de repetició, si és que existien, i l’estimació de diversos paràmetres referents a la dependència espaciotemporal.

En aquesta orientació més estadística, el nou model de referència va ser el Self-Exciting Point Process (Mohler et al. 2011; Reinhart 2017), que va permetre construir el primer model predictiu conegut, el programa PREDPOL (“PredPol – The Predictive Policing Company” n.d.).

En aquest cas, els responsables del programa es van oferir a fer una demostració amb dades delictives de Catalunya. Prenent de referència els delictes patrimonials a Barcelona durant un any, el software facilitava prediccions diàries en cel·les de 250 metres de costat. El programa identificava els hotspots habituals i recomanava a la policia desplegar-hi tasques preventives, especificant els tipus de delictes que s’hi podrien localitzar i la quantitat. Tanmateix, aquest output no semblava aportar valor afegit al coneixement policial ja conegut i que justificqués un canvi d’estratègia tàctica o operativa. A ulls dels propis policies encarregats de la gestió policial, el resultat era poc convincent i potenciava la sensació que s’havia d’assumir com a bo simplement perquè provenia d’unes equacions matemàtiques sofisticades i automatitzades que el feien presumiblement més potent i creïble que les eines d’anàlisi habituals de la policia.

Cercant una explicació als resultats poc convincents del programa PREDPOL a Catalunya, es va arribar a la conclusió que les característiques de l'entorn residencial hi podien tenir un paper molt rellevant. Les imatges de les figures 6 i 7 (Introducció) en són un exemple paradigmàtic i generalitzable.

Aquestes recerques prèvies van reforçar la idea d'estudiar els patrons espai-temps dels robatoris a Catalunya i les seves possibilitats predictives partint des de zero.

Les referències de partida eren els dos models anteriors, el model de Short i el model Self Exciting Point Process. Ambdós compartien alguns aspectes fonamentals. El més rellevant, el patró de victimització repetida i quasi repetida, de manera que el risc sempre s'estenia en la proximitat geogràfica i temporal. També la hipòtesis flag (Tseloni and Pease 2003), que posa de rellevància l'entorn, la teoria de l'oportunitat delictiva, l'estabilitat d'una part important del risc, també anomenat risc estàtic, molt relacionat amb l'històric de dades i les concentracions delictives habituals. Per últim, la hipòtesis boost (Kate J Bowers and Johnson 2004), el risc dinàmic, la noció que cada robatori augmenta el risc de més robatoris al seu voltant, que els robatoris generen rèpliques en els llocs que s'assemblen on havien fet el primer. Amb aquestes dues hipòtesis, els dos models tenien un input de dades senzill, i que no eren altres que les coordenades i el moment en què s'havien fet els robatoris. També compartien el plantejament en una malla de cel·les, tot i que amb configuracions diferents.

Aquests aspectes compartits, d'estructura i hipòtesis del problema i no de les metodologies concretes utilitzades, són les que es van agafar de punt de partida per l'estudi dels robatoris amb força a domicili a Catalunya. Sense tenir clara la configuració espai temps òptima, les dades es van estructurar en sèries temporals de robatoris en malles regulars de diferents mides. Es va observar que la majoria d'aquestes sèries tenien distribucions compatibles amb la distribució de Poisson, però les dades apareixien agrupades, en onades, i per tant, no eren independents o aleatòries. El patró mostrava que a valors baixos li seguien valors baixos, i a valors alts li seguien valors alts. Això va suggerir la idea de transformar les sèries en ratxes. Fixat un valor de tall, normalment pròxim a la mitjana, els valors per sobre i per sota d'aquest valor generaven ratxes no aleatòries. Aquesta circumstància generava un patró predictiu, perquè indica que quan en una de les cel·les considerades es sobrepassa un cert nombre de robatoris, es dispara el risc que el pròxim valor de la sèrie també el sobrepassi. Es tracta d'una regla predictiva que es pot precisar acompanyant-la de la probabilitat concreta que passi allò previst. La millor manera de fer-ho és amb l'històric de dades, perquè es va observar que les sèries de Poisson de les diferents cel·les eren estables. La mitjana que les caracteritza pràcticament no canvia al llarg dels anys, de manera que on hi ha pocs robatoris en l'actualitat, de mitjana hi seguirà havent pocs robatoris en un futur, i al revés. Per tant, les sèries històriques de dades són una bona referència per preveure el comportament futur dels patrons dels robatoris.

Aquest escenari predictiu implicava fer el seguiment individual de multitud de sèries, tantes com el nombre de cel·les regulars en què s'havia dividit el territori. Però no en totes aquestes cel·les hi havia robatoris, o n'hi havia molt pocs. I en les que n'hi havia, no en totes s'hi observava el patró no aleatori per agrupació de dades segons el test de ratxes i,

per tant, no en totes es podia predir. En les que sí, hi podia haver varies cotes a partir de els quals es generés l'onada de robatoris. Llavors, si l'onada prevista era de molt pocs robatoris, calia anar-hi a fer prevenció? O era preferible esperar-hi un increment superior? O potser s'havia d'anar a llocs on, encara que no es predís una onada, històricament hi hagués una mitjana alta de robatoris?

Aquestes qüestions van suggerir afegir un nou paràmetre per escollir els llocs òptims per predir, el coeficient de variació $\frac{\sigma}{\mu}$. Quan en una cel·la aquest valor és pròxim a 1, el patró d'onades alterna períodes de molt pocs o cap robatori amb períodes amb bastants robatoris. Per tant, si s'hi fes prevenció quan es detecti l'onada, a la llarga el nombre de robatoris hauria de disminuir considerablement. Aquesta semblaria l'estratègia preventiva adequada. Així, s'establia el criteri de cel·la òptima per predir quan s'hi observa el patró no aleatori d'onades de robatoris, i el coeficient de variació és pròxim a 1.

A les conclusions del primer article també s'apuntaven dubtes raonables sobre la possibilitat que el patró de repetició en escales petites expliqués bé la dinàmica general dels robatoris. Es suggeria considerar com a repeticions els robatoris en espais pròxims dins la mateixa cel·la de 5km de costat, encara que no fossin urbanísticament connexos. Això trencava el plantejament del patró clàssic de victimització quasi repetida, però en podia ser una ampliació natural en entorns heterogenis.

Les dues qüestions plantejades en el paràgraf anterior van motivar el segon article. Calia comprovar fins a quin punt era insuficient el patró de victimització quasi repetida restringit a la proximitat geogràfica i, en cas de confirmar-ne la limitació, indagar alternatives útils per a predir a Catalunya. Per fer-ho, es va construir un model que complís les hipòtesis del patró clàssic per modelar la dinàmica dels robatoris i es va arribar a la conclusió que el resultat el que modelava era, de fet, la part estàtica.

Per tal que el resultat fos vàlid per aquest propòsit, que permetés extreure conclusions generals, el model escollit havia de ser òptim i genèric per modelar fenòmens que consisteixen en comptatges d'alguna variable en l'espai i el temps. Es va optar per plantejar un model Log-Gaussià de Cox espaciotemporal (Serra et al. 2014; Diggle et al. 2013), que es basa amb la idea de l'existència d'una camp latent, continu, del risc de robatori. Un cop aproximat aquest camp latent, es pot calcular el risc en qualsevol punt de l'espai i el temps. Aquest tipus de model és útil per modelar tot tipus de dades espacials, ja siguin processos puntuals, dades en àrees, dades en malles o dades contínues.

Pel cas dels robatoris a domicili, es prepara el model per tal que ajusti la hipòtesi clàssica de la victimització quasi repetida. En aquest sentit, s'imposa una dependència espacial inversament proporcional a la distància, i una dependència temporal respecte el valor anterior de la sèrie (model AR(1)). Amb aquesta informació bàsica i les dades dels robatoris corresponents a l'any 2019, en el segon article es va comprovar que el mètode INLA (Lindgren and Rue 2015; Blangiardo et al. 2013; Bakka et al. 2018) era capaç d'aproximar el model i predir. La sistemàtica predictiva considerada consistia en introduir les dades dels robatoris de les últimes set setmanes, amb les que el mètode INLA actualitzava els paràmetres del model, i en feia una predicció per a la setmana següent.

El resultat del segon article confirma alguns aspectes del plantejament relacionat amb la teoria de la victimització quasi repetida, pel que fa a l'estabilitat del risc estàtic i a la dependència en l'espai i el temps, també a petita i mitjana escala. Però s'obté un tipus de predicció que es podria resumir en "el nombre total de robatoris la setmana següent serà aproximadament igual als de la setmana actual i, probablement, en els llocs on habitualment ja n'hi sol haver", que queda molt lluny de les expectatives predictives inicials.

Aquest segon article va suposar un punt d'inflexió en la recerca. Si el model, la metodologia per aproximar-lo i la sistemàtica per actualitzar-lo i predir, semblava adequada, quin aspecte es podia canviar per millorar les prediccions?

L'aspecte que semblava limitar més les capacitats predictives del model era la dependència espacial, que s'ha exigint que fos en forma de campana, inversament proporcional a la distància. Aquest va passar a ser l'objectiu a millorar. La conclusió d'aquest segon article reforçava i amplificava la suggerida en el primer. Les repeticions pot ser que es facin no només en espais no connexos a l'interior de cel·les de 5km de costat, sinó també en diferents cel·les, fins i tot allunyades, dins d'una mateixa iniciativa delictiva o onada de robatoris.

Per sustentar aquesta hipòtesis, es va consultar la literatura criminològica respecte l'estudi dels desplaçaments que fan els lladres i els grups criminals per delinquir, una camp de recerca conegut en anglès sota el terme "Journey to Crime" (Bernasco and Nieuwbeerta 2005; Michael Townsley and Sidebottom 2010; Nee 2015; Van Deviver, Van Daele, and Van Der Beken 2015; Wollinger et al. 2018). Com en molts camps de la criminologia, els resultats solen ser parcials, pendents de revisar i necessitats de més comprovacions empíriques. Per acotar-ho, en aquesta tesi s'han prioritzat els estudis fets en l'àmbit territorial Europeu, que aporten uns resultats coherents amb l'experiència i coneixement policial a Catalunya. El perfil dels grups criminals que cometen robatoris a domicili a Europa és que són de configuració internacional, amb membres d'una o varies nacionalitats, a vegades amb estructures organitzatives complexes, itinerants, amb alta mobilitat territorial, que es concerten per robar en llocs o perfils de llocs concrets, prèviament coneguts per algun dels membres del grup, que victimitzaran sistemàticament, en sèrie, amb tendència a tornar a les zones prèviament victimitzades. L' "espai de confort" d'aquests grups, el que es coneix en anglès com a "awareness area" (Cohen and Felson 1979; P. Brantingham and Brantingham 1995), passa ser un mapa mental complex de zones oportunes on robar, de com accedir-hi i actuar-hi. Les recerques criminològiques argumenten que aquest patró d'actuació, del desplaçament llarg, els aporta beneficis segons les teories habituals del comportament criminal (oportunitat delictiva, elecció racional o "foragging theory"). Entre altres aspectes, destaquen el fet que desplaçar-se a zones allunyades en disminueix la possibilitat de detecció.

En aquesta manera d'actuar, els delinqüents combinen concentració i dispersió a la vegada. Es tracta d'un patró complex, on es pot intuir la dificultat de determinar els desplaçaments a petita escala dins la xarxa de zones victimitzades. Però sí que podria ser possible detectar totes les zones que s'estan victimitzant en una mateixa iniciativa delictiva. Aquest últim és

l'objectiu plantejat en el tercer article. Partir de la hipòtesis “unus pro omnibus, omnes pro uno” i del patró de repetició “all-in-one” i cercar conjunts de zones amb possible victimització simultània.

Entre les diferents metodologies possibles per fer-ho, s'ha optat per l'Anàlisi de Components Principals (Hotelling 1933; I.T. Jolliffe 2010), considerant les cel·les de 5km de costat com les variables d'entrada per analitzar l'evolució de les sèries de robatoris setmanals. D'aquesta manera, la redundància d'informació que detecta el mètode, agrupant variables semblants, es pot interpretar com correlació o semblança en quan a les activacions temporals setmanals.

Per a una millor interpretació del resultat, s'opta per analitzar una de les regions policials més poblades, amb més percentatge de robatoris respecte el total i representativa de l'entorn heterogeni del territori català, la Regió Policial Metropolitana Nord (RPMN). Amb aquest plantejament, el resultat obtingut és que les 52 cel·les de 5km de costat on hi va haver robatoris a la regió l'any 2019, es poden agrupar en 19 components principals que concentren més del 80% de la variabilitat d'aquestes sèries de robatoris. Cada component és una llista ponderada de la totalitat de cel·les i, considerant una rotació Varimax de les components i una posterior reducció de variables, es poden eliminar cel·les de les components fins a quedar-ne entre 3 i 10 per cadascuna. A aquestes 19 components simplificades resultants se les anomena “constel·lacions de robatoris”.

Aquestes constel·lacions es poden entendre com uns mapes delinqüencials fixes, que es van activant i desactivant seguint ratxes temporals. Quan se n'activa un, els robatoris s'hi repartiran segons els pesos que els assigna la constel·lació, de manera que es pot fer una equivalència entre pesos i la probabilitats de robatoris a cada cel·la.

També es pot interpretar cada constel·lació com un “programa delictiu” concret, que tendeix a ser repetidament escollit pels delinqüents. El motiu es relaciona amb el perfil semblant que tenen els robatoris en els cel·les de major pes de cada constel·lació. Escollir una constel·lació per victimitzar les pròximes setmanes implica escollir un tipus de domicili, un tipus d'entorn i un tipus de botí. A la vegada, implica uns modus operandi concrets, unes rutes de desplaçament determinades, i un perfil d'autors. De fet, es defensa la idea que els autors de cadascuna d'aquestes onades de robatoris en les constel·lacions són els mateixos. Per tant, amb aquesta sistemàtica predictiva s'ofereix un conjunt d'informació complementària a la referent a l'espai i el moment dels robatoris predits que, des de la interpretació policial, pot augmentar considerablement el coneixement per a dissenyar accions preventives adequades per mirar d'evitar-los.

Els bons resultats predictius obtinguts en el tercer article animen a considerar estratègies preventives ambicioses, i a plantejar-se el trencament dels patrons habituals d'actuació dels grups criminals. Això comportaria un impacte gran en la incidència dels robatoris a domicili a Catalunya, tot i que amb un efecte incert referent a la seva possible reacció. Seran necessàries proves pilot per acabar de determinar el seu possible èxit i viabilitat.

Des del punt de vista dels models i les metodologies escollides per a fer predicció dels robatoris a domicili, aquests resultats tenen conseqüències importants. El principal escull

de la majoria d'aquests és la tendència a restringir la dependència espacial a una certa proximitat. Un dels plantejaments genèrics de referència de molts models són els processos puntuals no homogenis de Poisson. Aquests suposen, entre altres propietats, que si A_1 i A_2 són dos conjunts disjunts, llavors les variables aleatòries que mesuren el nombre de robatoris en cadascun, $N(A_1)$ i $N(A_2)$, seran independents. Aquesta propietat contradiria la presència de les constel·lacions i, per tant, els models que semblarien més adequats per a fer predicció dels robatoris amb força a domicili, en entorns heterogenis, s'haurien d'allunyar dels processos estrictament puntuals per evolucionar cap a models en què les estructures espacials de referència siguin xarxes, grafs o similars.

També obre la possibilitat d'utilitzar altra tipus de metodologies, com el Machine Learning, i plantejaments Big Data, en què la informació sembla ser més rellevant que la modelització. Tot i així, l'opció d'aquesta tesi ha estat la modelització, i la vinculació entre les teories criminològiques i els models matemàtics o estadístics. Això ha permès una lectura i interpretació a dues bandes que ha enriquit el debat i la millora dels resultats. Finalment, seguint el principi de parsimònia, s'ha obtingut un model adequat per a fer predicció dels robatoris a domicili a Catalunya utilitzant eines senzilles com l'Anàlisi de Components Principals. Aquest model té una clara interpretació criminològica que aporta informació útil per al desplegament d'estratègies preventives policials i que poden arribar a ser tant útils i necessàries com el propi model predictiu.

6 CONCLUSIÓ

La predicció dels robatoris amb força a domicili a Catalunya és possible. S'ha constatat l'existència de patrons de repetició tant en el curt com en el llarg termini. Hi ha tendència a victimitzar els mateixos llocs al llarg dels anys, de manera que es pot classificar el territori segons el risc permanent o estàtic de patir robatoris a domicilis. En aquestes zones de risc més elevat, s'hi sol robar en forma de ratxes temporalment agrupades, seguint la teoria criminològica de la victimització quasi repetida. Això implica que, després d'un primer robatori en un domicili, és més probable que hi hagi altres robatoris en la proximitat geogràfica.

Prenent una configuració de cel·les de 5km de costat i intervals temporals coincidents amb les setmanes naturals, el nombre de cel·les òptimes per predir a Catalunya seria suficientment gran com per garantir una estratègia predictiva-preventiva que cobriria el 75% dels robatoris. Aquesta sistemàtica a gran escala té a favor el fet que no necessita una modelització estadística sofisticada, i té en contra el fet de no ser a petita escala, com la resta de models predictius coneguts. Això en complica les hipotètiques accions preventives que se'n podrien derivar, que haurien de ser intenses, dissuasives, però distribuïdes en territoris extensos i durant una setmana. Tanmateix, l'entorn heterogeni a l'interior de les cel·les afavoreix acotar les zones de més risc i els intervals temporals en funció d'anàlisis descriptives dels robatoris recents i històrics a cada cel·la concreta que s'hagi de protegir.

Confirmada la dependència espaciotemporal dels robatoris a gran escala i prenent les premisses clàssiques de la victimització quasi repetida, el plantejament d'un model log-Gaussià de Cox revela que les prediccions a mitjana escala (cel·les de 500m, 1km o 5km de costat, en funció de l'entorn) poden ser acceptables segons els valors del coeficient de determinació R^2 , fins i tot analitzant cel·la a cel·la. Però l'aproximació tendeix a la mitjana, mostra menys variabilitat que les dades reals i no reproduïx bé la dinàmica dels robatoris en el curt termini. Es detecta dependència espacial en forma de campana en la immediata geogràfica però també s'observa que un robatori, en qualsevol lloc, té l'efecte d'incrementar el risc de robatoris arreu. A la vegada, es confirma la dependència temporal dels robatoris de la setmana actual amb l'anterior, que remarca el comportament no-aleatori en forma d'onades també a més petita escala.

S'arriba a la conclusió que tant la teoria de la victimització quasi repetida, com el model log-Gaussià de Cox, són insuficients per modelar la dinàmica general del risc dels robatoris. De fet, s'ha comprovat que les localitzacions dels robatoris tampoc segueixen un Procés de Poisson no homogeni, tot i que aquest s'hi pugui aproximar.

En entorns geogràfics heterogenis com el de Catalunya, en què les zones residencials són de tipologies diverses, normalment de dimensions reduïdes i àmpliament distribuïdes pel territori, les rèpliques d'un primer robatori s'estenen més enllà de la proximitat geogràfica. S'han trobat zones allunyades entre si que tendeixen a activar-se conjuntament, en les mateixes onades de robatoris, i que mostren entorns i perfils de domicilis victimitzats

semblants. Això suggereix que es tracta d'un patró delictiu escollit pels grups criminals de forma conscient, racional i planificada, que es pot justificar mitjançant les principals teories criminològiques de la conducta criminal.

Ajuntant els resultats del model estadístic del segon article amb la recerca criminològica del "desplaçament per delinquir", s'estructura una proposta explicativa mitjançant una hipòtesis i un nou patró espai-temps. La hipòtesis l'anomenem "Unus pro omnibus, omnes pro uno", i afirma que en entorns heterogenis, els lladres distribueixen els robatoris en l'espai seguint els criteris habituals d'oportunitat delictiva i, per adaptar-se a la dispersió i variabilitat de les zones residencials, generen petites concentracions de robatoris en diverses zones, distants entre sí, però que potser s'assemblen en quan a l'entorn més immediat i al tipus de domicili victimitzat. Considerant totes aquestes zones victimitzades com una de sola, s'hi observaria una onada clarament no-aleatòria de robatoris. Considerant-ne una de sola, s'hi podria observar, en alguns casos, patrons clàssics de victimització quasi repetida, en la proximitat geogràfica.

El nom d'aquesta hipòtesis es justifica pel fet que cada robatori serà considerat una rèplica dins l'onada global generada pel grup criminal i, a la vegada, tots aquests robatoris es consideraran inclosos en un sol patró en l'espai-temps. A aquest nou patró l'anomenarem "all-in-one". En funció de la configuració temporal utilitzada per analitzar aquests robatoris produïts en una xarxa de zones com la descrita, s'hi observaran patrons de victimització en la proximitat geogràfica (near-repeat pattern), patrons de victimització simultània (co-occurrence pattern), i patrons de salt o desplaçament (shift pattern). El patró de repetició espai-temps "all-in-one" és el que inclou tots aquests tipus de patrons i qualsevol combinació o alternatives. Tot robatori en les zones prèviament escollides pels grups criminals es considerarà una rèplica seguint el patró "all-in-one".

La descoberta d'aquest patró té implicacions directes en l'elecció dels models òptims per predir la delinqüència. Aquests models no han de restringir la possibilitat de correlació espacial entre zones distants, ans el contrari, han d'estar preparats per detectar estructures de xarxes de zones o grafs que tindran tendència a activar-se de forma unificada.

En aquesta tesi, s'ha comprovat que es poden construir models predictius utilitzant plantejaments i metodologies estàndards i senzilles, i amb un input d'informació consistent amb l'històric de dades dels robatoris. Tanmateix, és possible que una major informació recopilada amb sentit policial pugui millorar-los.

La gran dificultat dels models de predicció delictiva ha estat, i és, la seva implementació real en els cossos policials. En aquest sentit, hi segueixen havent incerteses, tant organitzatives com d'impacte en la delinqüència, que només es podran resoldre mitjançant l'aplicació de proves pilot. És d'esperar que els bons resultats obtinguts des del punt de vista acadèmic i metodològic animin a les policies a fer passos decisius en aquesta direcció.

7 BIBLIOGRAFIA

- Bakka, Haakon, Håvard Rue, Geir Arne Fuglstad, Andrea Riebler, David Bolin, Janine Illian, Elias Krainski, Daniel Simpson, and Finn Lindgren. 2018. "Spatial Modeling with R-INLA: A Review." *Wiley Interdisciplinary Reviews: Computational Statistics*. Wiley-Blackwell. <https://doi.org/10.1002/wics.1443>.
- Basart i Muñoz, Josep M. 1993. *Grafs : Fonaments i Algorismes / Josep M. Basart i Muñoz*. Bellaterra : Publicacions de la Universitat Autònoma de Barcelona,. https://discovery.udg.edu/iii/encore/record/C__Rb1032075?lang=cat.
- Bennett Moses, Lyria, and Janet Chan. 2018. "Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability." *Policing and Society* 28 (7): 806–22. <https://doi.org/10.1080/10439463.2016.1253695>.
- Bernasco, Wim. 2009. "FORAGING STRATEGIES OF HOMO CRIMINALIS: LESSONS FROM BEHAVIORAL ECOLOGY." *CRIME PATTERNS AND ANALYSIS* 2 (5). <http://wimbernasco.nl/Manuscripts/English/Bernasco2009a.pdf>.
- Bernasco, Wim, and Paul Nieuwebeerta. 2005. "How Do Residential Burglars Select Target Areas? : A New Approach to the Analysis of Criminal Location Choice." *British Journal of Criminology* 45 (3): 296–315. <https://doi.org/10.1093/bjc/azh070>.
- Biggs, N, E K Lloyd, and R J Wilson. 1986. *Graph Theory, 1736-1936*. Clarendon Press. <https://books.google.es/books?id=XqYTk0sXmpoC>.
- Birks, Daniel, Michael Townsley, and Anna Stewart. 2012. "Generative Explanations of Crime: Using Simulation to Test Criminological Theory." *Criminology* 50 (1): 221–54. <https://doi.org/10.1111/j.1745-9125.2011.00258.x>.
- Bishop, Christopher M. 2006. *Pattern Recognition and Machine Learning*. New York : Springer. https://discovery.udg.edu/iii/encore/record/C__Rb1228769__SBishop, Christopher M.__Orightresult__X4?lang=cat.
- Blangiardo, Marta, Michela Cameletti, Gianluca Baio, and Håvard Rue. 2013. "Spatial and Spatio-Temporal Models with R-INLA." *Spatial and Spatio-Temporal Epidemiology*. <https://doi.org/10.1016/j.sste.2013.07.003>.
- Bogucki, Robert, Jan Kanty Milczek, and Patryk Miziula. 2020. "A Simple Crime Hotspot Forecasting Algorithm." *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020* 21: 23–26. <https://doi.org/10.15439/2020F5>.
- Boqué, Pere, Laura Serra , and Marc Saez. 2020. "'Surfing' Burglaries with Forced Entry in Catalonia: Large-Scale Testing of near Repeat Victimization Theory." *European Journal of Criminology*, November, 147737082096810. <https://doi.org/10.1177/1477370820968102>.
- Borrion, Hervé, Paul Ekblom, Dalal Alrajeh, Aiduan Li Borrion, Aidan Keane, Daniel Koch, Timothy Mitchener-Nissen, and Sonia Toubaline. 2020. "The Problem with Crime Problem-Solving: Towards a Second Generation Pop?" *British Journal of Criminology* 60 (1): 219–40. <https://doi.org/10.1093/bjc/azz029>.
- Bowers, K. J. 2005. "Domestic Burglary Repeats and Space-Time Clusters: The Dimensions of Risk." *European Journal of Criminology* 2 (1): 67–92. <https://doi.org/10.1177/1477370805048631>.
- Bowers, Kate J, and Shane D Johnson. 2004. "Who Commits Near Repeats? A Test of the Boost

- Explanation." *Western Criminology Review* 5 (3): 12–24.
http://westerncriminology.org/documents/WCR/v05n3/article_pds/bowers.pdf.
- Bowers, Kate, Lucía Summers, and Suzanne Poynton. 2011. "Spatial Displacement and Diffusion of Benefits among Geographically Focused Policing Initiatives."
<https://doi.org/10.4073/csr.20011.3>.
- Brantingham, Patricia, and Paul Brantingham. 1995. "Criminality of Place - Crime Generators and Crime Attractors." *European Journal on Criminal Policy and Research* 3 (3): 5–26.
<https://doi.org/10.1007/BF02242925>.
- Brantingham, Paul J. 1984. *Patterns in Crime*. Edited by Patricia L Brantingham. London: Collier Macmillan.
- Briz-Redón, Álvaro, Francisco Martínez-Ruiz, and Francisco Montes. 2020. "Adjusting the Knox Test by Accounting for Spatio-Temporal Crime Risk Heterogeneity to Analyse near-Repeats." *European Journal of Criminology*. <https://doi.org/10.1177/1477370820905106>.
- Butt, Umair Muneer, Sukumar Letchmunan, Fadratul Hafinaz Hassan, Mubashir Ali, Anees Baqir, and Hafiz Husnain Raza Sherazi. 2020. "Spatio-Temporal Crime HotSpot Detection and Prediction: A Systematic Literature Review." *IEEE Access* 8: 166553–74.
<https://doi.org/10.1109/access.2020.3022808>.
- Cabrero Ortega, Yolanda, and Alfonso Garcia Perez. 2015. "Análisis Estadístico de Datos Espaciales Con Qgis y R," 191.
http://portal.uned.es/portal/page?_pageid=93,23376016&_dad=portal&_schema=PORTAL&_piref93_23376061_93_23376016_23376016.next_page=/htdocs/ficha.jsp?IdArticulo=010501OCT01A01.
- Caplan, Joel, and Leslie Kennedy. 2010. "Risk Terrain Modeling Manual: Theoretical Framework and Technical Steps of Spatial Risk Assessment." *Rutgers Center on Public Security*, 1–122.
- "CCI Project - We Aim To Cut Crime Impact." n.d. Accessed February 25, 2021.
<https://www.cuttingcrimeimpact.eu/>.
- "Center for Problem-Oriented Policing | ASU Center for Problem-Oriented Policing." n.d. Accessed November 10, 2020. <https://popcenter.asu.edu/>.
- ChaNEY, Author Spencer. 2012. "JDI Briefs : Predictive Mapping for Policing." *Blog Post*, no. April: 1–5.
- ChaNEY, Spencer. 2013. "Examining the Influence of Cell Size and Bandwidth Size on Kernel Density Estimation Crime Hotspot Maps for Predicting Spatial Patterns of Crime." *BSSLg* 60 (1): 7–19.
- ChaNEY, Spencer P., Sophie J. Curtis-Ham, R. Mark Evans, and Gordon J. Burns. 2018. "Examining the Extent to Which Repeat and near Repeat Patterns Can Prevent Crime." *Policing* 41 (5): 608–22. <https://doi.org/10.1108/PIJPSM-12-2016-0172>.
- ChaNEY, Spencer Paul, and Braulio Figueiredo. 2016. "Examining the Extent of Repeat and near Repeat Victimization of Domestic Burglaries in Belo Horizonte , Brazil." *Crime Science*, 1–10.
<https://doi.org/10.1186/s40163-016-0049-6>.
- Chen, Peng, and Justin Kurland. 2020. "The Impact of 'Strike Hard' on Repeat and near-Repeat Residential Burglary in Beijing." *ISPRS International Journal of Geo-Information* 9 (3).
<https://doi.org/10.3390/ijgi9030150>.
- Cinelli, Virginia, and Alberto Manrique. 2019. "El Uso de Programas de Análisis Predictivo En La Inteligencia Policial: Una Comparativa Europea." *Revista de Estudios En Seguridad*

- Internacional* 5 (2): 1–20. <https://doi.org/10.18847/1.10.1>.
- Clarke, R V G, and Marcus Felson. 2008. *Routine Activity and Rational Choice : Advances in Criminological Theory*.
- Clarke, Ronald V., and Derek B. Cornish. 1986. “Modeling Offenders’ Decisions: A Framework for Research and Policy.” In *Crime Opportunity Theories*, 147–85. <https://doi.org/10.4324/9781315095301-8>.
- . 2018. “Modeling Offenders’ Decisions: A Framework for Research and Policy.” In *Crime Opportunity Theories*, 6:157–95. The University of Chicago Press. <https://doi.org/10.4324/9781315095301-8>.
- Cohen, Lawrence E., and Marcus Felson. 1979. “Social Change and Crime Rate Trends: A Routine Activity Approach.” *American Sociological Review* 44 (4): 588. <https://doi.org/10.2307/2094589>.
- “CRAN - Package Rcrimeanalysis.” n.d. Accessed February 7, 2021. <https://cran.r-project.org/web/packages/rcrimeanalysis/index.html>.
- “CRAN - Package Surveillance.” n.d. Accessed February 7, 2021. <https://cran.r-project.org/web/packages/surveillance/index.html>.
- “CrimeStat: Spatial Statistics Program for the Analysis of Crime Incident Locations | National Institute of Justice.” n.d. Accessed February 7, 2021. <https://nij.ojp.gov/topics/articles/crimestat-spatial-statistics-program-analysis-crime-incident-locations#citation--0>.
- Curtis-Ham, Sophie, Wim Bernasco, Oleg N. Medvedev, and Devon Polaschek. 2020. “A Framework for Estimating Crime Location Choice Based on Awareness Space.” *Crime Science* 9 (1): 1–14. <https://doi.org/10.1186/s40163-020-00132-7>.
- Daele, Stijn Van, and Tom Vander Beken. 2009. “Out of Step? Mobility of ‘Itinerant Crime Groups.’” *Crime, Money and Criminal Mobility in Europe*, 43–70.
- David, F N, and D E Barton. 1966. “TWO SPACE-TIME INTERACTION TESTS FOR EPIDEMICITY*.” *Brit. J. Prev. Soc. Med* 20: 48. <https://doi.org/10.1136/jech.20.1.44>.
- “Department of Criminal Justice | Near Repeat Calculator.” n.d. Accessed March 20, 2018. <http://www.cla.temple.edu/cj/center-for-security-and-crime-science/projects/nearrepeatcalculator/>.
- Deviver, Christophe Van, Stijn Van Daele, and Tom Van Der Beken. 2015. “What Makes Long Crime Trips Worth Undertaking Balancing Costs and Benefits in Burglars Journey to Crime.” *British Journal of Criminology* 55 (2): 399–420. <https://doi.org/10.1093/bjc/azu078>.
- Diggle, Peter J., Paula Moraga, Barry Rowlingson, and Benjamin M. Taylor. 2013. “Spatial and Spatio-Temporal Log-Gaussian Cox Processes: Extending the Geostatistical Paradigm.” *Statistical Science* 28 (4): 542–63. <https://doi.org/10.1214/13-STS441>.
- Diggle, Peter J. 2005. “Spatio-Temporal Point Processes: Methods and Applications.” *Statistical Methods for Spatio-Temporal Systems*, no. June 2005: 19–64. <https://doi.org/10.1201/9781420011050-5>.
- Eck, J E, S Chainey, J G Cameron, and M Leitner. 2005. “UCL Discovery - Mapping Crime: Understanding Hot Spots,” no. May 2014. <http://eprints.ucl.ac.uk/11291/%5Cnpapers3://publication/uuid/8D145E1D-FFF8-48FB-9E57-3FC90AC3C875>.

- Eck, John E., and Dennis P. Rosenbaum. 2014. "The New Police Order: Effectiveness, Equity, and Efficiency in Community Policing." In *The Challenge of Community Policing: Testing the Promises*, 3–24. 2455 Teller Road, Thousand Oaks California 91320 United States : SAGE Publications, Inc. <https://doi.org/10.4135/9781483327006.n1>.
- Egbert, Simon, and Susanne Krasmann. 2020. "Predictive Policing: Not yet, but Soon Preemptive?" *Policing and Society* 30 (8): 905–19. <https://doi.org/10.1080/10439463.2019.1611821>.
- EUCPN. 2016. "Recommendation Paper: Predictive Policing," 13. www.eucpn.org.
- Everitt, Brian S, Sabine Landau, Morven Leese, and Daniel Stahl. 2011. *Cluster Analysis 5th Edition*. Wiley.
- Farrell G., and K Pease. 1993. "Once Bitten, Twice Bitten: Repeat Victimization and Its Implications for Crime Prevention." *Police Research Group, Crime Prevention Unit Paper 46*, no. 46: 38. <http://hdl.handle.net/2134/2149%5Cnhttp://core.ac.uk/download/pdf/67259.pdf>.
- Farrell G., K Pease, and C Phillips. 1995. "Like Taking Candy : Why Does Repeat Victimization Occur ?" *Loughborough University Institutional Repository* 35 (3).
- Farrell, Graham. 2015. "Crime Concentration Theory." *Crime Prevention and Community Safety* 17 (4): 233–48. <https://doi.org/10.1057/cpcs.2015.17>.
- Farrell, Graham, and Ken Pease. 2017. "Preventing Repeat and near Repeat Crime Concentrations." In *Handbook of Crime Prevention and Community Safety*. <https://doi.org/10.4324/9781315724393>.
- Farrell, Graham, Ken Pease, and Simon Fraser. 2012. "Prediction and Crime Clusters," no. 1995: 3862–71.
- Felson, Marcus, and Ronald V Clarke. 1998. "Opportunity Makes the Thief Practical Theory for Crime Prevention." *POLICE RESEARCH SERIES, Paper 98, Home Office*. <http://www.popcenter.org/library/reading/pdfs/thief.pdf>.
- Fielding, Nigel G. 2005. "Concepts and Theory in Community Policing." *The Howard Journal of Criminal Justice* 44 (5): 460–72. <https://doi.org/10.1111/j.1468-2311.2005.00391.x>.
- Fuglstad, Geir Arne, Daniel Simpson, Finn Lindgren, and Håvard Rue. 2019. "Constructing Priors That Penalize the Complexity of Gaussian Random Fields." *Journal of the American Statistical Association* 114 (525): 445–52. <https://doi.org/10.1080/01621459.2017.1415907>.
- Gabriel, Edith, Barry Rowlingson, and Peter J Diggle. 2007. "STPP." In *Hawley's Condensed Chemical Dictionary*. <https://doi.org/10.1002/9780470114735.hawley15198>.
- . 2013. "Journal of Statistical Software Stpp: An R Package for Plotting, Simulating and Analyzing Spatio-Temporal Point Patterns." <http://www.jstatsoft.org/>.
- Gerstner, Dominik. 2018. "Predictive Policing in the Context of Residential Burglary: An Empirical Illustration on the Basis of a Pilot Project in Baden-Württemberg, Germany." *European Journal for Security Research* 3 (2): 115–38. <https://doi.org/10.1007/s41125-018-0033-0>.
- Geurts, Michael, George E. P. Box, and Gwilym M. Jenkins. 1977. "Time Series Analysis: Forecasting and Control." *Journal of Marketing Research* 14 (2): 269. <https://doi.org/10.2307/3150485>.
- Glasner, Philip, Shane D. Johnson, and Michael Leitner. 2018. "A Comparative Analysis to Forecast Apartment Burglaries in Vienna, Austria, Based on Repeat and near Repeat Victimization." *Crime Science* 7 (1). <https://doi.org/10.1186/s40163-018-0083-7>.
- Goldstein, H. 1990. *Problem-Oriented Policing*. McGraw-Hill.

- <https://books.google.es/books?id=WCqkQgAACAAJ>.
- Gottfredson, Michael R., and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford University Press.
- Groff, Elizabeth R., Shane D. Johnson, and Amy Thornton. 2019. "State of the Art in Agent-Based Modeling of Urban Crime: An Overview." *Journal of Quantitative Criminology* 35 (1): 155–93. <https://doi.org/10.1007/s10940-018-9376-y>.
- Groff, Elizabeth, and Travis Taniguchi. 2019. "Quantifying Crime Prevention Potential of Near-Repeat Burglary." *Police Quarterly* 22 (3): 330–59. <https://doi.org/10.1177/1098611119828052>.
- Guerette, Rob T., and Kate J. Bowers. 2009. "Assessing the Extent of Crime Displacement and Diffusion of Benefits: A Review of Situational Crime Prevention Evaluations." *Criminology* 47 (4): 1331–68. <https://doi.org/10.1111/j.1745-9125.2009.00177.x>.
- Haberman, Cory P., and Jerry H Ratcliffe. 2012. "The Predictive Policing Challenges of Near Repeat Armed Street Robberies" 6 (2): 151–66. <https://doi.org/10.1093/police/pas012>.
- He, Zhanjun, Liufeng Tao, Zhong Xie, and Chong Xu. 2020. "Discovering Spatial Interaction Patterns of near Repeat Crime by Spatial Association Rules Mining." *Scientific Reports* 10 (1): 1–12. <https://doi.org/10.1038/s41598-020-74248-w>.
- Hipp, John R., and Young An Kim. 2017. "Measuring Crime Concentration across Cities of Varying Sizes: Complications Based on the Spatial and Temporal Scale Employed." *Journal of Quantitative Criminology* 33 (3): 595–632. <https://doi.org/10.1007/s10940-016-9328-3>.
- Hotelling, H. 1933. "Analysis of a Complex of Statistical Variables into Principal Components." *Journal of Educational Psychology* 24 (7): 498–520. <https://doi.org/10.1037/h0070888>.
- I.T. Jolliffe. 2010. "Principal Components Analysis." *Springer Series in Statistics*, 374–77. <https://doi.org/10.1016/B978-0-08-044894-7.01358-0>.
- Ignatans, Dainis, and Ken Pease. 2018. "Crime Concentrations: Hot Dots, Hotspots, and Hot Flushes." In *The Oxford Handbook of Environmental Criminology*.
- "Inici. Institut Cartogràfic i Geològic de Catalunya." n.d. Accessed February 21, 2021. <https://www.icgc.cat/>.
- John E. Eck, Spencer Chainey, James G. Cameron, Michael Leitner, and Ronald E. Wilson. 2005. "Mapping Crime: Understanding Hot Spots | National Institute of Justice." Vol. August 200. <https://nij.ojp.gov/library/publications/mapping-crime-understanding-hot-spots>.
- Johnson, S. D., Kate Bowers, and A Hirschfield. 1997. "NEW INSIGHTS INTO THE SPATIAL AND TEMPORAL DISTRIBUTION OF REPEAT VICTIMIZATION." *The British Journal of Criminology* 37 (2): 224–41. <https://doi.org/10.1093/oxfordjournals.bjc.a014156>.
- Johnson, Shane D. 2014. "How Do Offenders Choose Where to Offend? Perspectives from Animal Foraging." *Legal and Criminological Psychology* 19 (2): 193–210. <https://doi.org/10.1111/lcrp.12061>.
- Johnson, Shane D., Wim Bernasco, Kate J. Bowers, Henk Elffers, Jerry Ratcliffe, George Rengert, and Michael Townsley. 2007. "Space-Time Patterns of Risk: A Cross National Assessment of Residential Burglary Victimization." *Journal of Quantitative Criminology* 23 (3): 201–19. <https://doi.org/10.1007/s10940-007-9025-3>.
- Johnson, Shane D, Wim Bernasco, Kate J Bowers, and Henk Elffers. 2006. "Space-Time Patterns of

- Risk : A Cross National Assessment of Residential Burglary Victimization Space-Time Patterns of Risk : A Cross National Assessment of Residential Burglary Victimization,” no. 0: 1–47.
- Johnson, Shane D, and Kate J Bowers. 2004. “The Stability of Space-Time Clusters of Burglary.” *The British Journal of Criminology* 44 (1): 55–65. <https://doi.org/10.1093/bjc/44.1.55>.
- Kaufmann, Mareile, Simon Egbert, and Matthias Leese. 2019. “Predictive Policing and the Politics of Patterns.” *British Journal of Criminology* 59 (3): 674–92. <https://doi.org/10.1093/bjc/azy060>.
- Kikuchi, G, M Amemiya, and T Saito. 2010. “A Spatio-Temporal Analysis of Near Repeat Victimization in Japan.” *Eighth National Crime ...* <http://www.ucl.ac.uk/scs/events/mapping-conf/conf-2010/downloads-2010/4b-kikutchi.pdf>.
- Knox, E G, and M S Bartlett. 1964. “The Detection of Space-Time Interactions.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 13 (1): 25–30. <https://doi.org/10.2307/2985220>.
- Knox, George. 1963. “DETECTION OF LOW INTENSITY EPIDEMICITY APPLICATION TO CLEFT LIP AND PALATE.” *Brit. J. Prev. Soc. Med* 17: 121–27. <https://doi.org/10.1136/jech.17.3.121>.
- Krebs, J. R. (John R.), and N. B. (Nicholas B.) Davies. 1993. *An Introduction to Behavioural Ecology*. London [etc.] : Blackwell Scientific Publications.
https://discovery.udg.edu/iii/encore/record/C__Rb1027814__SAn Introduction to Behavioural Ecology__Orightresult__U__X6?lang=cat.
- Lantz, Brendan, and R. Barry Ruback. 2017. “A Networked Boost: Burglary Co-Offending and Repeat Victimization Using a Network Approach.” *Crime and Delinquency* 63 (9): 1066–90. <https://doi.org/10.1177/0011128715597695>.
- Lasierra, Francesc Guillén. 2006. “Les Polítiques Locals de Prevenció i Seguretat a Europa.” *Revista Catalana de Seguretat Pública*, no. 16 SE-Dossier (January). <https://www.raco.cat/index.php/RCSP/article/view/121649>.
- Leong, Kelvin, and Anna Sung. 2015. “A Review of Spatio-Temporal Pattern Analysis Approaches on Crime Analysis.” *International E-Journal of Criminal Sciences* 9: 1–33.
- Lindgren, Finn, and Håvard Rue. 2015. “Bayesian Spatial Modelling with R - INLA.” *Journal of Statistical Software* 63 (19). <https://doi.org/10.18637/jss.v063.i19>.
- Mallesonid, Nick, Wouter Steenbeek, and Martin A Andresen. 2019. “Identifying the Appropriate Spatial Resolution for the Analysis of Crime Patterns.” <https://doi.org/10.1371/journal.pone.0218324>.
- Meijer, Albert, and Martijn Wessels. 2019. “Predictive Policing: Review of Benefits and Drawbacks.” *International Journal of Public Administration* 42 (12): 1031–39. <https://doi.org/10.1080/01900692.2019.1575664>.
- Mohler, G. O., M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. 2011. “Self-Exciting Point Process Modeling of Crime.” *Journal of the American Statistical Association* 106 (493): 100–108. <https://doi.org/10.1198/jasa.2011.ap09546>.
- “Mossos d’Esquadra - Portal Dades Obertes.” n.d. https://mossos.gencat.cat/ca/els_mossos_desquadra/indicadors_i_qualitat/dades_obertes/cataleg_dades_obertes/dades-delinquencials/.
- Nee, Claire. 2015. “Understanding Expertise in Burglars: From Pre-Conscious Scanning to Action and Beyond.” *Aggression and Violent Behavior* 20: 53–61. <https://doi.org/https://doi.org/10.1016/j.avb.2014.12.006>.

- Ornstein, Joseph T., and Ross A. Hammond. 2017. "The Burglary Boost: A Note on Detecting Contagion Using the Knox Test." *Journal of Quantitative Criminology* 33 (1): 65–75. <https://doi.org/10.1007/s10940-016-9281-1>.
- Pearson K. 1901. "Pearson, K. 1901. On Lines and Planes of Closest Fit to Systems of Points in Space." *Philosophical Magazine* 2: 559–72.
- Pease, Ken. 1998. *Repeat Victimisation: Taking Stock. Crime Detection and Prevention Series*. http://www.popcenter.org/problems/domestic_violence/PDFs/Pease_1998.pdf.
- Pease, Ken, and Graham Farrell. 2001. "Repeat Victimization." *Crime Prevention Studies*, no. February 2015. <https://doi.org/10.1007/978-1-4614-5690-2>.
- Pérez Delgado, Ma Luisa., and Quintín Martín Martín. 2003. *Aplicación de Las Redes Neuronales Artificiales a La Estadística*. Salamanca : Editorial La Muralla. https://discovery.udg.edu/iii/encore/record/C__Rb1203787__SAplicaciones de las redes neuronales artificiales a la estadística__Orighresult__U__X3?lang=cat.
- Perry, Walter, Brian McInnis, Carter Price, Susan Smith, and John Hollywood. 2018. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. <https://doi.org/10.7249/rr233>.
- "PRECOSBS Ifmpt - Institut Für Musterbasierte Prognosetechnik." n.d. Accessed March 20, 2018. <http://www.ifmpt.de/>.
- "Predict Crime | Predictive Policing Software | PredPol." n.d. Accessed March 20, 2018. <http://www.predpol.com/>.
- "PredPol – The Predictive Policing Company." n.d. Accessed February 5, 2021. <https://www.predpol.com/>.
- "R: The R Project for Statistical Computing." n.d. Accessed February 7, 2021. <https://www.r-project.org/>.
- Ratcliffe, J., R. Taylor, R. Fisher, J. Koehnlein, A. Askey, K. Thomas, J. Grasso, and K. Bethel. 2020. "The Philadelphia Predictive Policing Experiment." *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-019-09400-2>.
- Ratcliffe, J. 2008. *Intelligence-Led Policing*. Willan. <https://books.google.es/books?id=iPxLvgAACA AJ>.
- Ratcliffe, Jerry. 2009. "Near Repeat Calculator." <http://www.temple.edu/cj/misc/nr/download.htm>.
- Ratcliffe, Jerry H. 2014. "Intelligence-Led Policing." *Encyclopedia of Criminology and Criminal Justice*, 2573–81. https://doi.org/10.1007/978-1-4614-5690-2_270.
- Ratcliffe, Jerry H., and Clairissa Breen. 2011. "Crime Diffusion and Displacement: Measuring the Side Effects of Police Operations." *Professional Geographer* 63 (2): 230–43. <https://doi.org/10.1080/00330124.2010.547154>.
- Reinhart, Alex. 2017. "A Review of Self-Exciting Spatio-Temporal Point Processes and Their Applications," August. <https://doi.org/10.1214/17-STS629>.
- Rey, Sergio J., Elizabeth A. Mack, and Julia Koschinsky. 2012. "Exploratory Space-Time Analysis of Burglary Patterns." *Journal of Quantitative Criminology* 28 (3): 509–31. <https://doi.org/10.1007/s10940-011-9151-9>.
- "Risk Terrain Modeling | Official Site - Home." n.d. Accessed November 10, 2020. <https://www.riskterrainmodeling.com/>.

- Rue, Håvard, Sara Martino, and Nicolas Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations." *J. R. Statist. Soc. B.* Vol. 71.
- Sagovsky, A, and S Johnson. 2007. "When Does Repeat Burglary Victimisation Occur?" *Australian & New Zealand Journal of Criminology* 40: 1–26.
- Saldana, Joan, Maria Agualeles, Albert Avinyó, Marta Pellicer, and Jordi Ripoll. 2018. "An Age-Structured Population Approach for the Mathematical Modeling of Urban Burglaries." *SIAM Journal on Applied Dynamical Systems* 17 (4): 2733–60. <https://doi.org/10.1137/17M1142090>.
- Seidensticker, Kai, Felix Bode, and Florian Stoffel. 2018. "Predictive Policing in Germany," 1–7. <http://nbn-resolving.de/urn:nbn:de:bsz:352-2-14sbvox1ik0z06>.
- Serra, Laura, Marc Saez, Jorge Mateu, Diego Varga, Pablo Juan, Carlos Díaz-Ávalos, and Håvard Rue. 2014. "Spatio-Temporal Log-Gaussian Cox Processes for Modelling Wildfire Occurrence: The Case of Catalonia, 1994–2008." *Environmental and Ecological Statistics* 21 (3): 531–63. <https://doi.org/10.1007/s10651-013-0267-y>.
- Shapiro, Aaron. 2019. "Predictive Policing for Reform? Indeterminacy and Intervention in Big Data Policing." *Surveillance and Society*. Vol. 17. <https://doi.org/10.24908/ss.v17i3/4.10410>.
- Short, M. B., M. R. D'Orsogna, V. B. PASOUR, G. E. Tita, P. J. Brantingham, A. L. BERTOZZI, and L. B. CHAYES. 2008. "A STATISTICAL MODEL OF CRIMINAL BEHAVIOR." *Mathematical Models and Methods in Applied Sciences* 18 (supp01): 1249–67. <https://doi.org/10.1142/S0218202508003029>.
- Sidebottom, Aiden, Karen Bullock, Rachel Armitage, Matt Ashby, Caitlin Clemmow, Stuart Kirby, Gloria Laycock, and Nick Tilley. 2020. "Problem-Oriented Policing in England and Wales 2019."
- Simpson, D., J. B. Illian, F. Lindgren, S. H. Sørbye, and H. Rue. 2015. "Going off Grid: Computationally Efficient Inference for Log-Gaussian Cox Processes." *Biometrika* 103 (1): 49–70. <https://doi.org/10.1093/biomet/asv064>.
- Simpson, Daniel, Håvard Rue, Andrea Riebler, Thiago G. Martins, and Sigrunn H. Sørbye. 2017. "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors." *Statistical Science* 32 (1): 1–28. <https://doi.org/10.1214/16-STS576>.
- Smith, Laura M., Matthew S. Keegan, Todd Wittman, George O. Mohler, and Andrea L. Bertozzi. 2010. "Improving Density Estimation by Incorporating Spatial Information." *EURASIP Journal on Advances in Signal Processing* 2010 (1): 265631. <https://doi.org/10.1155/2010/265631>.
- Soto Urpina, Carles. 2016. "LA PREVENCIÓN SITUACIONAL: BASES TEÓRICAS DE FUNDAMENTO CRIMINOLÓGICO." *Iter Criminis* 15 (Oct-Dic): 85–111.
- Tilley, Nick. 2012. "Modern Approaches to Policing: Community, Problem-Oriented and Intelligence-Led," August, 401–31. <https://doi.org/10.4324/9780203118238-25>.
- Tobler, W R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46: 234–40. <https://doi.org/10.2307/143141>.
- Townsley, M. 2003. "Infectious Burglaries. A Test of the Near Repeat Hypothesis." *British Journal of Criminology* 43 (3): 615–33. <https://doi.org/10.1093/bjc/43.3.615>.
- Townsley, Michael. 2018. "Can Predictive Policing Work in Australia?" *Police Science* 2 (1). <https://doi.org/10.4018/978-1-5225-7672-3>.
- Townsley, Michael, Ross Homel, and Janet Chaseling. 2000. "Repeat Burglary Victimisation: Spatial

- and Temporal Patterns." *Australian and New Zealand Journal of Criminology* 33 (1): 37–63. <https://doi.org/10.1177/000486580003300104>.
- Townsley, Michael, Shane D Johnson, Ken Pease, and Jill Dando. 2003. "PROBLEM ORIENTATION, PROBLEM SOLVING AND ORGANIZATIONAL CHANGE." *Crime Prevention Studies*. Vol. 15.
- Townsley, Michael, and Aiden Sidebottom. 2010. "All Offenders Are Equal, but Some Are More Equal than Others: Variation in Journeys to Crime between Offenders." *Criminology* 48 (3): 897–917. <https://doi.org/10.1111/j.1745-9125.2010.00205.x>.
- Tseloni, Andromachi, and Ken Pease. 2003. "Repeat Personal Victimization: 'Boosts' or 'Flags'?" *British Journal of Criminology* 43 (1): 196–212. <https://doi.org/10.1093/bjc/43.1.196>.
- Vijaya Kumar, M. 2011. "Spatial-Temporal Analysis of Residential Burglary Repeat Victimization: Case Study of Chennai City Promoters Apartments, INDIA." *IJCTA*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.208.7766&rep=rep1&type=pdf>.
- Vozmediano Sanz, Laura., and César. San Juan Guillén. 2010. *Criminología Ambiental : Ecología Del Delito y de La Seguridad*. Editorial UOC.
- Wald, A., and J. Wolfowitz. 1940. "On a Test Whether Two Samples Are from the Same Population." *The Annals of Mathematical Statistics* 11 (2): 147–62. <https://doi.org/10.1214/aoms/1177731909>.
- Wang, Bao, Xiyang Luo, Fangbo Zhang, Baichuan Yuan, Andrea L. Bertozzi, and P. Jeffrey Brantingham. 2018. "Graph-Based Deep Modeling and Real Time Forecasting of Sparse Spatio-Temporal Data." <https://doi.org/10.475/123>.
- Wang, Zengli, Lin Liu, Hanlin Zhou, and Minxuan Lan. 2019. "Crime Geographical Displacement: Testing Its Potential Contribution to Crime Prediction." *ISPRS International Journal of Geo-Information* 8 (9): 1–12. <https://doi.org/10.3390/ijgi8090383>.
- Wang, Zengli, and Xuejun Liu. 2017. "Analysis of Burglary Hot Spots and Near-Repeat Victimization in a Large Chinese City." *ISPRS International Journal of Geo-Information* 6 (12): 148. <https://doi.org/10.3390/ijgi6050148>.
- Wang, Zengli, and Hong Zhang. 2020. "Construction, Detection, and Interpretation of Crime Patterns over Space and Time." *ISPRS International Journal of Geo-Information* 9 (6). <https://doi.org/10.3390/ijgi9060339>.
- Weathington, Katherine. 2020. "A Critical Study of Geospatial Algorithm Use in Crime Analysis and Predictive Policing," no. May.
- Weisburd, David, Malay K. Majmundar, Hassan Aden, Anthony Braga, Jim Bueermann, Philip J. Cook, Phillip Atiba Goff, et al. 2019. "Proactive Policing: A Summary of the Report of the National Academies of Sciences, Engineering, and Medicine." *Asian Journal of Criminology* 14 (2): 145–77. <https://doi.org/10.1007/s11417-019-09284-1>.
- Wollinger, Gina Rosa, Alicia König, Maximilian Querbach, and Alexandra Röhrig. 2018. "Offender Organization and Criminal Investigations with Regard to Organised Residential Burglary."
- Yang, Fei. 2020. "Predictive Policing , Oxford Research Encyclopedia , Criminology and Criminal," no. June: 0–26. <https://doi.org/10.1093/acrefore/9780190264079.013.508>.
- Zhang, Haifeng, and Wei Song. 2014. "Addressing Issues of Spatial Spillover Effects and Non-Stationarity in Analysis of Residential Burglary Crime." *GeoJournal* 79 (1): 89–102. <https://doi.org/10.1007/s10708-013-9481-2>.

Zhang, Yang, and Tao Cheng. 2020. "Graph Deep Learning Model for Network-Based Predictive Hotspot Mapping of Sparse Spatio-Temporal Events." *Computers, Environment and Urban Systems* 79 (April 2019): 101403. <https://doi.org/10.1016/j.compenvurbsys.2019.101403>.