Fire Safety Journal 157 (2025) 104510

Contents lists available at ScienceDirect

Fire Safety Journal

journal homepage: www.elsevier.com/locate/firesaf

# The use of machine learning for accidental dwelling fire prevention

M. Taylor [a,*], E. Dean [b], J. Fielding [c], R. Lyon [a], D. Reilly [a], H. Francis [a], V. Kwasnica [a]

[a] Department of Computer Science and Mathematics, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK
[b] Greater Manchester Fire and Rescue Service, 146 Bolton Road, Swinton, Manchester, M27 8US, UK
[c] Merseyside Fire and Rescue Service, Bridle Road, Liverpool, L30 4YD, UK

ARTICLE INFO

*Keywords:*
Machine learning fire prevention

ABSTRACT

In this article the use of machine learning for fire prevention support is examined over the period 2010 to 2024 based on a case study in a fire and rescue service in Northwest England. Machine learning was used to develop a multiple linear regression model of accidental dwelling fire risk at the Lower Super Output Area of geography. This was enhanced by using machine learning to develop a k-means cluster analysis model of communities at the finer grained Output Area level. Over the study period the percentage decrease in accidental dwelling fires in the area studied was 44.2 % compared to a decrease of 27.5 % in England as a whole which appeared to indicate that the more precise targeting of fire prevention resulting from statistical models using machine learning had a positive effect on the effectiveness of fire prevention activities.

## 1. Introduction

Fire prevention information systems need to evolve in order to support the fire prevention environment in which they operate. Adetoro [1] advocated the use of machine learning by fire departments to support fire safety. A multiple linear regression statistical model using supervised machine learning was developed to predict the number of accidental dwelling fire incidents (ADFs) per Lower Super Output Area [2] using socio-economic data [3]. The model used a standard ordinary least squares approach to minimise the sum of squared residuals (the differences between the observed values and the values predicted by the regression model) using supervised machine learning available in the SPSS statistical package. A Lower Super Output Area (LSOA) is a geographical unit used in the UK typically containing 400 to 1200 households and 1000 to 3000 residents. A cluster analysis statistical model using unsupervised machine learning was then developed to identify different population segments using finer grained geographical area data at the Output Area [4] level [5]. The K-means cluster analysis model used standard Euclidean distance to measure the similarity (or dissimilarity) between cases (calculating the straight-line distance between points in multidimensional space to assign cases to clusters) using unsupervised machine learning available in the SPSS statistical package. An Output Area is the smallest geographic unit used for statistical purposes in the UK typically containing 40 to 250 households and 100 to 625 residents. The fire and rescue service (FRS) concerned covers a county in the Northwest of England with an area of over 600 km$^2$ containing mainly higher density urban areas, as well as some semi-rural and rural areas. Fire prevention for domestic fires increasingly concerns statistical analysis of the circumstances of fires, and socio-economic causal factors [5,6]. Machine learning for fire prevention research has mainly concentrated on forest fire prediction [7]. Machine learning approaches have also been used for specific aspects of domestic fire prevention including modelling cooker top fires based on sensor data [8], and modelling structure related fire casualties [9]. The originality of the research is the examination of statistical modelling using machine learning for supporting fire prevention activities.

## 2. Research method

The research involved developing statistical models using machine learning to enhance the targeting of fire prevention over the period 2010 to 2024. The first phase involved the development of a multiple linear regression statistical model using supervised machine learning for predicting ADF incidents by LSOA based upon causal factor socio-economic variables. The casual factor variables were variables that had been identified by previous research as being associated with ADF incidence [3,5]. The second phase concerned a k-means cluster analysis statistical model that used unsupervised machine learning for community profiling

* Corresponding author.
*E-mail addresses:* m.j.taylor@ljmu.ac.uk (M. Taylor), deane@manchesterfire.gov.uk (E. Dean), johnfielding@merseyfire.gov.uk (J. Fielding), r.lyon@ljmu.ac.uk (R. Lyon), d.reilly@ljmu.ac.uk (D. Reilly), h.francis@ljmu.ac.uk (H. Francis), v.kwasnica@ljmu.ac.uk (V. Kwasnica).

https://doi.org/10.1016/j.firesaf.2025.104510
Received 17 July 2025; Received in revised form 18 August 2025; Accepted 19 August 2025
Available online 22 August 2025

by Output Area to support the targeting of fire prevention activities to specific vulnerable community groups.

Historically fire prevention information systems in the UK utilised spatial analysis of fire incidence [10]. The original information system for fire prevention support within the fire and rescue service studied estimated the fire risk of Lower Super Output Areas by weighting the number of accidental dwelling fires in an area and the indices of multiple deprivation [12]. The indices of multiple deprivation are used in the UK to assess the relative levels of deprivation across small geographical areas using information concerning income, employment, education, health, crime, housing, and living environment.

The data for the machine learning modelling was obtained from the FRS concerned and UK government agencies including the UK Office for National Statistics (ONS), the UK National Health Service (NHS), and the UK Department of Work and Pensions (DWP). An important aspect of using machine learning to develop statistical models supporting fire prevention was checking the assumptions of the statistical techniques used. For example, checking correlations between predictor and predicted variables for the multiple linear regression modelling, and using a z-transformation to avoid variables with larger ranges unduly affecting the cluster analysis. The machine learning models were developed using the Statistical Package for the Social Sciences (SPSS) software.

## 3. Results

Socio-economic data associated with ADFs from external sources including the UK ONS, DWP, and NHS was used. A multiple linear regression statistical model that used supervised machine learning of accidental dwelling fire incidence by LSOA was developed. The correlations between the socio-economic variables and the ADFs per LSOA were examined to assess suitability for inclusion. The predictor variables included: mental health benefit claimants, fitted smoke detectors, severe disability claimants, binge drinkers, lone parents, people living alone, and disability living allowance claimants per LSOA. The machine learning model selection process used forward stepwise and backward stepwise approaches within SPSS. The machine learning algorithm used ordinary least squares regression. The coefficient of multiple linear regression of the model was 0.71 (significant at 0.01 level). The R-squared value (coefficient of determination) for the model was 0.50. The machine learning multiple linear regression model identified LSOAs at higher risk of accidental dwelling fires, and also provided an understanding of why an area might have a higher accidental dwelling fire risk, enabling fire prevention activities to be tailored to the needs of different communities.

An k-means cluster analysis statistical model that used unsupervised machine learning was developed using OA level datasets from the UK ONS, DWP, and HMRC. The cluster modelling enabled the development of detailed community profiles. Twenty different variables were used including: age groups, adult social care, and council tax and housing benefits. Ten community profiles were established including three types of elderly communities, four types of communities with high levels of deprivation, students living in city center locations, young families, and middle-income residents in privately owned properties. The ten community profiles enabled the identification of those individuals with risk factors likely to increase ADF risk including: being over age 65, living alone, disability, and mental health issues.

Over the period studied the number of ADFs in the area studied (FRS) and England is shown in Table 1 as reported in the Fire and rescue incident statistics published by the UK Ministry of Housing, Communities and Local Government [11].

Over the study period the percentage decrease in ADFs in England was 27.5 %, however, the percentage decrease in ADFs in the FRS studied was 44.2 %. This would appear to indicate that the ADF prevention approaches adopted by the FRS supported by machine learning had been effective in terms of reducing the number of ADFs per year. In terms of the practical use of the machine learning based fire prevention, home fire safety checks conducted by the FRS were targeted towards those aged 65+ from 2011 onwards, and from 2017 onwards were further targeted towards those aged 65+ who were living in higher levels of deprivation.

**Table 1**
ADFs in the area studied (FRS) and England 2010/11 to 2023/24.

| Year | FRS | England |
|---|---|---|
| 2010/11 | 1200 | 31718 |
| 2011/12 | 1201 | 30802 |
| 2012/13 | 1133 | 29674 |
| 2013/14 | 1153 | 28615 |
| 2014/15 | 1053 | 28324 |
| 2015/16 | 1090 | 28360 |
| 2016/17 | 998 | 27248 |
| 2017/18 | 927 | 27595 |
| 2018/19 | 900 | 26562 |
| 2019/20 | 869 | 25533 |
| 2020/21 | 800 | 24288 |
| 2021/22 | 839 | 24479 |
| 2022/23 | 777 | 24093 |
| 2023/24 | 670 | 23002 |

Over the period studied the number of ADF injuries in the area studied (FRS) and England is shown in Table 2.

Over the period the percentage decrease in ADF injuries in England was 36.9 %, with a similar percentage decrease in ADF injuries in the FRS of 36.7 %.

Over the time period studied the number of ADF fatalities in the area studied (FRS) and England is shown in Table 3.

Over the period there appeared to be no distinct pattern or trend in the numbers of ADF fatalities in England as a whole or the FRS studied. There was a reasonably strong correlation between the number of ADFs per year and the number of ADF injuries per year of 0.89 with a significance level $< 0.001$ in the fire and rescue service studied. However, there was very little correlation between the number of ADF injuries per year and the number of fatalities ($r = 0.15$, $p = 0.60$), or between the number of ADFs per year and the number of fatalities per year ($r = 0.31$, $p = 0.29$).

Fig. 1 shows the changes in ADFs, ADF injuries and ADF fatalities in the FRS studied over the study period.

## 4. Conclusions

Machine learning can support the development of fire prevention information systems by examining the combinations of different subsets of relevant variables to aid understanding of socio-economic factors associated with ADFs. In terms of developing statistical models using machine learning to support ADF preve ntion, limiting factors included the limited sample size in terms of the number of ADFs per year, and the availability of data at the relevant level of geographic granularity. A

**Table 2**
ADF injuries in the area studied (FRS) and England 2010/11 to 2023/24.

| Year | FRS | England |
|---|---|---|
| 2010/11 | 248 | 6525 |
| 2011/12 | 221 | 6372 |
| 2012/13 | 243 | 5934 |
| 2013/14 | 248 | 5433 |
| 2014/15 | 180 | 5282 |
| 2015/16 | 200 | 5184 |
| 2016/17 | 175 | 4767 |
| 2017/18 | 154 | 4818 |
| 2018/19 | 159 | 4628 |
| 2019/20 | 149 | 4556 |
| 2020/21 | 137 | 4292 |
| 2021/22 | 162 | 4237 |
| 2022/23 | 140 | 4042 |
| 2023/24 | 157 | 4118 |

**Table 3**
ADF fatalities in the area studied (FRS) and England 2010/11 to 2023/24.

| Year | FRS | England |
|---|---|---|
| 2010/11 | 6 | 211 |
| 2011/12 | 4 | 187 |
| 2012/13 | 6 | 173 |
| 2013/14 | 9 | 178 |
| 2014/15 | 10 | 166 |
| 2015/16 | 14 | 189 |
| 2016/17 | 7 | 186 |
| 2017/18 | 4 | 242 |
| 2018/19 | 4 | 164 |
| 2019/20 | 5 | 167 |
| 2020/21 | 7 | 148 |
| 2021/22 | 4 | 186 |
| 2022/23 | 11 | 190 |
| 2023/24 | 1 | 157 |

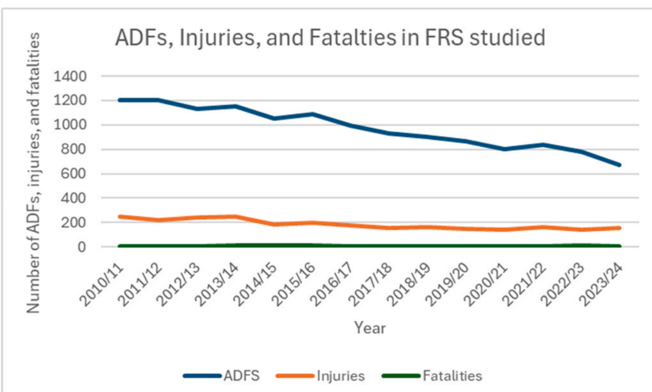


**Fig. 1.** ADFs, ADF injuries and ADF fatalities in the FRS 2010/11 to 2023/24.

limitation of multiple linear regression modelling can be that the variance of the outcome may increase with its mean which could violate the ordinary least squares assumption of homoscedasticity (constant variance of errors). Areas with more households would tend to have more fires and more variability in fire counts. Regressing raw counts might therefore underestimate uncertainty in low-count areas, and over-estimate precision in high-count areas. The percentage decrease in ADFs in the area over the time period studied was 44.2 % compared to a decrease of 27.5 % in England as a whole. This appeared to indicate that the more precise targeting of fire prevention resulting from the use of statistical models using machine learning had a positive effect on fire prevention activities. This research extends and enhances previous research into machine learning use by fire and rescue services which mainly concerned forest fire prevention.

## CRediT authorship contribution statement

**M. Taylor:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **E. Dean:** Writing – review & editing, Investigation, Data curation. **J. Fielding:** Writing – review & editing, Data curation. **R. Lyon:** Writing – review & editing, Methodology, Formal analysis. **D. Reilly:** Writing – review & editing, Methodology. **H. Francis:** Writing – review & editing, Investigation. **V. Kwasnica:** Writing – review & editing, Formal analysis.

## Declaration of competing interest

I can confirm that there are no Conflicts of interest with regard to the submission of the paper entitled "The use of machine learning for accidental dwelling fire prevention" to your journal.

## Data availability

Data will be made available on request.

## References

[1] A. Adetoro, Data analytics and machine learning: revolutionizing fire safety and compliance for U.S. fire departments, Int. J. Stat. Med. Res. 4 (4) (2021) 151–167.
[2] LSOA, Lower Super Output Areas, UK Census 2021 Geographies, UK Office for National Statistics, 2021. https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies.
[3] M. Taylor, E. Higgins, P. Lisboa, V. Kwasnica, An exploration of causal factors in unintentional dwelling fires, Risk Manag. 14 (2012) 109–125.
[4] OA, Output Areas, UK Census 2021 Geographies, UK Office for National Statistics, 2021. https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies.
[5] E. Higgins, M. Taylor, M. Jones, P. Lisboa, Understanding community fire risk—A spatial model for targeting fire prevention activities, Fire Saf. J. 62 (2013) 20–29.
[6] M. Taylor, D. Appleton, G. Keen, J. Fielding, Assessing the effectiveness of fire prevention strategies, Publ. Money Manag. 39 (6) (2019) 418–427.
[7] C. Shi, F. Zhang, A forest fire susceptibility modeling approach based on integration machine learning algorithm, Forests 14 (7) (2023) 1506.
[8] W. Tam, E. Fu, A. Mensch, A. Hamins, C. You, G. Ngai, H. Leong, Prevention of cooktop ignition using detection and multi-step machine learning algorithms, Fire Saf. J. 120 (2021) 103043.
[9] A. Schmidt, E. Gemmil, R. Hoskins, Machine learning based risk analysis and predictive modeling of structure fire related casualties, Mach. Learn. Appl. 20 (2025) 100645.
[10] FSEC, FSEC Toolkit Review Manual, UK Ministry of Housing, Communities and Local Government, 2006. https://www.gov.uk/government/publications/fsec-toolkit-review-manual.
[11] MHLG, Fire and Rescue Incident Statistics, UK Ministry of Housing, Communities and Local Government, 2025. https://www.gov.uk/government/collections/fire-statistics-monitor.
[12] IMD, English indices of deprivation. Ministry of Housing, Communities and Local Government, UK Government, 2019. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019.