Predicting Domestic Extremism and Targeted Violence: A Machine Learning Approach

Iris Malone  
*George Washington University*

Anastasia Strouboulis  
*George Washington University*

National Counterterrorism Innovation, Technology, and Education Center

Follow this and additional works at: [https://digitalcommons.unomaha.edu/ncitereportsresearch](https://digitalcommons.unomaha.edu/ncitereportsresearch)  
Please take our feedback survey at: [https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE](https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE)

**Recommended Citation**  
[https://digitalcommons.unomaha.edu/ncitereportsresearch/25](https://digitalcommons.unomaha.edu/ncitereportsresearch/25)

This Report is brought to you for free and open access by the National Counterterrorism Innovation, Technology, and Education (NCITE) at DigitalCommons@UNO. It has been accepted for inclusion in Reports, Projects, and Research by an authorized administrator of DigitalCommons@UNO. For more information, please contact [unodigitalcommons@unomaha.edu](mailto:unodigitalcommons@unomaha.edu).
Predicting Domestic Extremism and Targeted Violence
A Machine Learning Approach

Iris Malone and Anastasia Strouboulis
# Table of Contents

EXECUTIVE SUMMARY ........................................................................................................ 3

INTRODUCTION ...................................................................................................................... 3
  RESEARCH QUESTIONS ........................................................................................................ 4
  METHODOLOGY .................................................................................................................. 4

DEFINITIONS ........................................................................................................................... 5
  DOMESTIC EXTREMISM ....................................................................................................... 5
    Types of Domestic Extremism ............................................................................................ 5
  TARGETED VIOLENCE ......................................................................................................... 6
    Types of Targeted Violence ............................................................................................... 6

OVERVIEW OF DOMESTIC VIOLENT EXTREMISM AND TARGETED VIOLENCE IN THE U.S. ......................................................................................................................... 7
  TYPES OF RMVE ................................................................................................................ 7
  TRENDS IN RMVE .............................................................................................................. 9
  CHANGE IN OFFLINE AND ONLINE ORGANIZATION .................................................... 11
  TYPES OF TERRORISM AND TARGETED VIOLENCE ...................................................... 12

RISK FACTORS ....................................................................................................................... 15
  SOCIAL RISK FACTORS .................................................................................................... 16
  ECONOMIC RISK FACTORS ............................................................................................ 17
  INDIVIDUAL BACKGROUND RISK FACTORS ................................................................ 18
  DEMOGRAPHIC RISK FACTORS ...................................................................................... 19

METHODOLOGY ..................................................................................................................... 20
  DATA SOURCES .................................................................................................................. 20
  RESEARCH DESIGN ............................................................................................................ 21

APPLICATIONS ....................................................................................................................... 22
  PREDICTING DOMESTIC EXTREMISM .............................................................................. 22
  PREDICTING TARGETED VIOLENCE ................................................................................. 26

EXTENSION: IDENTIFYING VULNERABILITIES IN THE MARITIME DOMAIN ...................... 28

CONCLUSION AND POLICY RECOMMENDATIONS ................................................................ 30

APPENDIX ............................................................................................................................... 32
About the authors: Iris Malone is an Assistant Professor of Political Science and International Affairs at the Elliott School of International Affairs, George Washington University (irismalone@gwu.edu). Anastasia Strouboulis is a MA Canditate at the Elliott School of International Affairs, George Washington University (astrouboulis@gwmail.gwu.edu)

About NCITE: The National Counterterrorism Innovation, Technology, and Education (NCITE) Center was established in 2020 as the Department of Homeland Security Center of Excellence for counterterrorism and terrorism prevention research. Sponsored by the DHS Science and Technology Office of University Programs, NCITE is the trusted DHS academic consortium comprised of over 60 researchers across 18 universities and non-government organizations. Headquartered at the University of Nebraska at Omaha, NCITE seeks to be the leading U.S. academic partner for counterterrorism research, technology, and workforce development.

Acknowledgement: This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number 20STTPC00001-01.

Disclaimer: The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security or George Washington University.
Executive Summary

- This report summarizes the results of two machine learning prototype models that forecast the location of (1) domestic extremist groups and (2) active shooting incidents.
  - The domestic extremism model forecasts a group’s area of operations with 96% accuracy and 85% sensitivity rate.
  - The active shooter model forecasts incident locations with 91-92% accuracy and 51-71% sensitivity rate.
- The results suggest community-level risk factors are highly predictive of extremist operations and incidents.
- Prioritizing resources towards high-risk areas and supporting community-based awareness programs may mitigate these vulnerabilities.

The report applies machine learning (ML) techniques to forecast where domestic extremist groups and active shooter incidents are most likely to occur in the United States. Identifying high-risk areas for these emerging threats is important for effective counterterrorism and conflict prevention, but complicated by the fact that policymakers often need to detect these threats at a stage when there might not be overt warning signs of violence. This report addresses this gap and directly supports Strategic Goals 1.1 and 1.2 in the June 2021 National Strategy for Countering Domestic Terrorism by providing “data-driven guidance on how to recognize potential indicators of mobilization to domestic terrorism.”

We develop and test two prototype machine learning models based on existing research about the causes of radicalization, ideologically-motivated violent extremism (IMVE), and targeted violence. First, we input information about these potential risk indicators as well as data about extremist actors and violent incidents to map patterns between 2017-2020. We then use this information to forecast which areas are at highest risk for extremism and active shooter incidents. As an extension, we also identify which areas in the maritime domain are most likely to experience active shooter incidents. The model’s high level of accuracy suggests that these risk indicators are highly predictive of extremist operations and incidents.

Overall, these models provide guidance for practitioners about where extremist actors and violent incidents are most likely to emerge moving forward.

Introduction

In recent years, the threat of domestic radicalization and violent extremism (DVE) has eclipsed that of Al-Qaeda and the Islamic State. While the United States prioritized counterterrorism and countering violent extremism operations following 9/11, new IMVE—ranging from white supremacist ideologies to anti-government militias to so-called “lone wolves”—has steadily become a more prominent national security priority. In its annual threat assessment released in October 2020, the Department of Homeland Security (DHS) warned that “racially and ethnically motivated violent extremists (RMVEs) will remain the most persistent and lethal threat in the homeland.”

As evidence, the Southern Poverty Law Center’s annual assessment on domestic extremist

---

groups reported over 650 “hate groups” operating in 2020 across all 50 states, highlighting the breadth of the problem (Table 1).

Echoing the DHS assessment, in March 2021, a report from the Office for the Director of National Intelligence (ODNI) wrote that “racially or ethnically motivated violent extremists and militia violent extremists (MVE) present the most lethal DVE threats.” This fear is not unfounded. Between January 1, 2020 and August 31, 2020, the Center for Strategic and International Studies (CSIS) found that white supremacist groups were responsible for 67% of terrorist plots and attacks. The potential for ideologically-motivated domestic extremists to turn to violence and conduct either terrorism, hate crime, or targeted violence is a significant and pressing concern.

Table 1. Summary Statistics of Domestic Extremism and Active Shooter Incidents

<table>
<thead>
<tr>
<th>Year</th>
<th>Number Extremist Groups</th>
<th>Number Active Shootings</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>660</td>
<td>358</td>
</tr>
<tr>
<td>2018</td>
<td>790</td>
<td>336</td>
</tr>
<tr>
<td>2019</td>
<td>745</td>
<td>417</td>
</tr>
<tr>
<td>2020</td>
<td>650</td>
<td>610</td>
</tr>
<tr>
<td>2021</td>
<td>--</td>
<td>241 (as of May 31)</td>
</tr>
</tbody>
</table>

In 2019, the DHS added “targeted violence” to its prevention mission, stating that “mass attacks are a persistent problem and a grave concern.” The rising number of active shooter incidents reinforces these concerns (Table 1). Though incidents of targeted violence may lack a clear ideological motive, these mass attacks—mainly seen via active shooter incidents—compromise the safety and security of schools, religious institutions, and other public spaces. The events of January 6, 2020 punctuated these concerns when individuals, including members of various extremist groups, breached the United States Capitol to challenge the integrity of the presidential elections. Effectively responding to domestic violent extremism—specifically IMVE—necessitates understanding the risk factors, structures, and processes through which individuals become radicalized and carry out ideologically-motivated violence.

Research Questions

The central research questions this report seeks to answer are:

1. What is the current state of domestic violent extremism in the United States?
2. What are the characteristics of ideologically-motivated extremists and targeted violence, like active shooter incidents?
3. How can machine learning models help predict and prevent the emergence of domestic extremist groups and active shooter incidents?

We answer the first two questions by surveying a growing set of research on emerging trends and key risk indicators. We then use this information to build a machine learning model that forecasts the highest risk areas in the United States for domestic extremism and targeted violence in the near term.

Methodology

This report presents a literature review compiled from various academic sources, think tank reports, government documents, and independent non-profit and non-governmental organizations to answer the research questions. The data in this

---

5 Data on domestic extremist groups comes from the SPLC Hate Groups dataset; data on active shooting incidents comes from the Gun Violence Archive. We explain each data source further in the methodology section.
The report is drawn from multiple sources, including the Armed Conflict Location & Event Data Project (ACLED), the Profiles of Individual Radicalization in the United States (PIRUS), the Empirical Assessment of Domestic Radicalization (EADR), U.S. Census Bureau, U.S. Bureau of Transportation Statistics, the Southern Poverty Law Center (SPLC), the Anti-Defamation League (ADL), and the Gun Violence Archive (GVA). For our machine learning analysis, we develop two prototype models using cross-validation and random forest algorithms.

**Definitions**

**Domestic Extremism**

We define a domestic extremist group as an organization of non-state actors which justify the use of violent and non-violent actions to pursue an ideologically-motivated goal. This is broader than the definition of domestic violent extremists since it includes groups that are not yet violent.

DHS and the Federal Bureau of Investigation (FBI) define domestic violent extremists as individuals based and operating primarily in the United States without direction or inspiration from a foreign terrorist group or other foreign power and who seek to further political or social goals wholly or in part through unlawful acts of force or violence.

This definition aligns with the FBI’s definition of a hate crime which is a “criminal offense against a person or property motivated in whole or in part by an offender’s bias against a race, religion, disability, sexual orientation, ethnicity, gender, or gender identity.”

It also matches growing recognition in the Homeland Security community about the overlap between acts of domestic terrorism and hate crimes. For example, the 2019 DHS Strategic Framework on Countering Terrorism and Targeted Violence broadens understanding about terrorism by stating “hate crimes and non-ideologically motivated large-scale or disproportionately lethal acts of mass violence, including mass attacks, round out the picture of terrorism and targeted violence afflicting the Homeland.”

**Types of Domestic Extremism**

**Domestic Terrorism**

Under 18 U.S. Code § 2331, domestic terrorism is defined as “violent acts or acts dangerous to human life” that occur primarily within U.S. territory. These acts intend to “intimidate or coerce a civilian population,” “influence the policy of a government by intimidation or coercion,” and “affect the conduct of a government by mass destruction, assassination, or kidnapping.” In addition, the FBI defines domestic terrorism as “violent, criminal acts committed by individuals and/or groups to further ideological goals stemming from domestic influences, such as those of a political, religious, social, racial, or environmental nature.”

**Racially motivated violent extremism (RMVE)**

The DHS and FBI define RMVE as the “unlawful use or threat of force or violence in furtherance of ideological agendas derived from bias, often related to race or ethnicity, held by the actor against others...”

---

9 See, for example, “Domestic terrorists...have caused more deaths in the United States in recent years than have terrorists connected to FTOs. Domestic terrorist attacks and hate crimes sometimes overlap, as perpetrators of prominent domestic terrorist attacks have selected their targets based on factors such as race, ethnicity, national origin, religion, sexual orientation, gender, and gender identity.” cited in “Strategic Framework for Countering Terrorism and Targeted Violence.” Department of Homeland Security. September 2019. p. 10.
or a given population group.”

**Hate Groups**
The SPLC defines a hate group as an “an organization or collection of individuals that has beliefs or practices that attack or malign an entire class of people, typically for their immutable characteristics,” including their race, religion, ethnicity, sexual orientation, or gender identity. An organization does not need to engage in criminal, violent, or unlawful conduct to be listed as a hate group by the SPLC.

**Targeted Violence**
According to the DHS, targeted violence is “any incident of violence that implicates homeland security and/or DHS activities in which a known or knowable attacker selects a particular target prior to the violent attack.” Targeted violence is distinct from terrorism in that it includes “attacks that lack a clearly discernible political, ideological, or religious motivation.”

**Types of Targeted Violence**

**Active Shooter Incidents**
Active shooter incidents are one of the most prominent types of targeted violence. The FBI defines an active shooter incident as one where “one or more individuals actively engage in killing or attempting to kill people in a populated area.”

**Mass Shooter Incidents**
A subset of active shooting incidents is mass shootings. These incidents require an active shooter event to result in a minimum number of casualties. The Gun Violence Archive project defines “mass shootings” incidents as “four or more shot and/or killed in a single event, at the same general time and location not including the shooter.” The violence threshold used by this organization maps onto federal statutes and regulations. The Congressional Research Service defines a mass shooting as a “multiple homicide incident in which four or more victims are murdered with firearms, within one event, and in one or more locations in close proximity.” Following the Sandy Hook Elementary School shooting in 2012, Congress lowered the threshold of a “mass killing” to an incident which results in the death of three or more people.

**Hate Crimes**
The DHS strategic framework also includes hate crimes in its definition of targeted violence. The DHS and FBI both define a hate crime as “a criminal offense against a person or property motivated in whole or in part by an offender’s bias against a race, religion, disability, sexual orientation, ethnicity, gender, or gender identity.” Hate crimes are differentiated from regular criminal activities, such as murder, arson, vandalism, or physical assault, by the motivation to commit the crime based on a “bias against people or groups with specific characteristics.”

---

16 Ibid.
Overview of Domestic Violent Extremism and Targeted Violence in the U.S.

Although a range of ideologies motivate DVE, including radical environmentalism, anti-fascism, and anti-government extremism, this report principally focuses on RMVEs because of their growing threat to national security. This section reviews trends in RMVE groups and targeted violence from 2017-2021.

Types of RMVE
The ideological landscape of contemporary domestic extremism is complex and multi-faceted. Individuals and groups often adopt multiple ideologies, making it hard to identify individual and group-level risk factors for ideological violence. Extremists illustrate ideological fluidity; an RMVE could be mainly motivated by Islamophobia, while Boogaloo Boys have participated in BLM protests to advance their anti-law enforcement views. The heterogeneity of the RMVE ideology presents a formidable challenge for detecting high-risk personnel and deterring potential attacks. In general, the most common types of ideologically-motivated extremism today are RMVEs, which can include white supremacists, neo-Nazi, neo-Confederate, anti-immigrant, and anti-Muslim ideologies (Table 2).

Table 2. Overview of RVME Categories

<table>
<thead>
<tr>
<th>Types of RMVE</th>
<th>General Description</th>
<th>Example Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Supremacy (WSE)</td>
<td>Actors with ideological agendas derived from bias, often related to race or ethnicity, held by the actor against others, including a given population group. WSEs promote the superiority of the white race.</td>
<td>The Base, Klu Klux Klan, Neo-Nazi</td>
</tr>
<tr>
<td>Neo-Nazis</td>
<td>Actors who promote anti-Semitic and neo-fascist beliefs. They often call to replace democratic institutions with increased authoritarianism and will often harbor white supremacist views.</td>
<td>Atomwaffen Division, Vanguard America, Traditionalist Workers Party</td>
</tr>
<tr>
<td>Neo-Confederate</td>
<td>Actors who promote the preservation of Confederate monuments, memorials, and re-igniting a separatist campaign for the South to secede. They often harbor white supremacist views.</td>
<td>League of the South, Identity Dixie</td>
</tr>
<tr>
<td>Anti-immigrant and anti-Muslim</td>
<td>Actors who promote xenophobic and nativist views. Anti-immigrant hate groups go further by espousing racist propaganda.</td>
<td>ACT for America, Center for Security Policy, Soldiers of Odin</td>
</tr>
</tbody>
</table>
The Intelligence Community’s May 2021 Strategic Intelligence Assessment states that “RMVEs, primarily those advocating for the superiority of the white race, likely would continue to be the most lethal DVE threat” to the United States.23 Although RMVE groups and individuals engage in both non-violent and violent activity, 2019 was the most lethal year for DVE since 1995. That year, RMVE attacks resulted in 19 fatalities.24

According to PIRUS data, the majority (65%) of far-right extremists adhere to some expression of white supremacy.25 For example, the SPLC tracked 25 Ku Klux Klan groups and 128 white nationalist groups active in 2020.26 White supremacist extremists (WSEs) believe that people of European descent are inherently superior to others and should therefore dominate social, political, and cultural institutions. “Accelerationist” beliefs also permeate many RMVE philosophies, such as the Base. According to this view, the current system of Western government is “irreparable and therefore violent action is needed to precipitate societal collapse to start a race war.”27 Accelerationism inspired deadly RMVEs/WSEs attacks, including the October 2018 shooting at Pittsburgh’s Tree of Life synagogue, the April 2019 shooting at a synagogue in Poway, California, and the August 2019 shooting at a Walmart in El Paso.

Some white-supremacist groups adopt neo-Nazi or neo-Confederate positions. Neo-Nazi groups typically promote anti-Semitic and neo-fascist beliefs. There are numerous political organizations such as the Traditionalist Workers Party, National Socialist Movement, and Vanguard America which recruit and raise attention to their cause. In this sense, they mirror many of the far-right ultranationalist political parties in European politics, such as the Nordic Resistance Movement (Sweden), Golden Dawn (Greece), and the Third Path (Germany). Neo-Confederate groups organize around the “Lost Cause,” or mythology surrounding the Confederacy and causes of the U.S. Civil War. Some groups try to preserve Confederate memorials and other symbols of the Confederate movement, which are associated with racist and white supremacist beliefs. Other groups such as Identity Dixie and the League of the South, discuss the need for a renewed separatist campaign to break away from the United States.

A final sub-group within the far-right movement are anti-immigrant and anti-Muslim extremists. While these views are commonly integrated within white supremacy, some adherents are “animated directly in opposition to people who are or are perceived to be immigrants or are of the Islamic faith.”28 PIRUS found that nearly 30% of extremists in the dataset from 2015-2018 were motivated by anti-immigrant and anti-Muslim views, compared to 8% from 2006-2015.29 The FBI reported that of 1,650 religious-bias hate crimes in 2019, 60.3% were anti-Jewish, and 13.3% were anti-Islamic.30 The SPLC reports that in 2020, 19 anti-immigrant groups and 72 anti-Muslim groups were active.31 Anti-Muslim groups believe that Muslims undermine America’s political system

26 “Hate Map.” Southern Poverty Law Center. https://www.splcenter.org/hate-map
29 Ibid, p. 2.
31 “Hate Map.” Southern Poverty Law Center. https://www.splcenter.org/hate-map
and aim to replace it with Sharia law.\textsuperscript{32} Groups such as ACT for America and the Center for Security Studies have sought to deepen their connections with elected officials, particularly during the Syrian refugee crisis and the COVID-19 pandemic.

**Trends in RMVE**

Figure 1 plots the location of all ideologically-motivated domestic extremist groups which formed between 2017-2020 according to the SPLC Hate Group dataset. We note three interesting trends in domestic extremist groups based on the results. First, we find extremist groups present in all 50 states, but the majority in the South Atlantic, Deep South, and Tex-ar-kana region. Second, we find that these groups often organize in relatively more suburban and rural environments. While some groups form in metropolitan areas, like the Rise Above Movement in Los Angeles or Atomwaffen in Tampa, we also find many instances of groups forming in other areas such as Grass Valley, California (“Asatru Folk Assembly”) or Lake City, Florida (“League of the South”).

---

Finally, we find changes in the number of different extremist groups (Figure 2). The number of neo-Nazi affiliated groups has declined since 2018, which may be related to the arrest of several members of Atomwaffen and the Base.33 The number of neo-Confederate groups and anti-immigrant groups has also declined slightly in the last two years. Meanwhile, the number of white nationalist groups remains high.34 This reflects recent ODNI estimates that white supremacy remains a significant threat with the potential to escalate in the future.35 As such, we run a separate model that tries to estimate where WSE groups, specifically, are likely to form.

These findings from SPLC reporting corroborate other RMVE studies. According to the PIRUS database, the number of far-right groups now represents the largest proportion of domestic extremists, including the far-left, Salafi-jihadi adherents, and single-issue extremists.36 Between 2013 and 2017, the number of far-right extremists increased from 30 to 166.37 Other research finds similar results. For example, the CSIS Transnational Threats dataset covers January 1994 to May 2020; during this period, right-wing terrorists perpetrated the majority (57%) of all attacks and plots, followed

---

36 National Consortium for the Study of Terrorism and Responses to Terrorism’s (START) Profiles of Individual Radicalization in the United States (PIRUS). https://www.start.umd.edu/profiles-individual-radicalization-united-states-pirus-keshif
37 Ibid.
by 25% committed by left-wing individuals. Far-right extremists perpetrated two-thirds of the attacks and plots in the United States in 2019 and over 90% between January 1 and May 8, 2020.

Changes in Offline and Online Organization

- IMVEs increasingly use the Internet and social media platforms to connect and communicate.
- IMVEs have yet to fully operationalize online tools to plan, prepare, and conduct offline activities.

A major change in the domestic extremism landscape is the rapid growth and adoption of new information technologies to organize. IMVE ideologies are diffuse, as their networks increasingly exist on online platforms, including Facebook, Twitter, YouTube, 8kun (8chan), Telegram, and others.

The online space is now “the most critical mechanism for extreme-right connectivity” for groups within the United States and those with transnational relations. Online platforms are used to form groups, spread extremist ideology, recruit members, share propaganda and violent “how-to” manuals, and glorify and encourage acts of violence. In particular, the shift of extremist groups to secure messaging platforms makes it more difficult for law enforcement to surveil users and identify potential attackers while further radicalizing individuals. The internet’s ubiquity has also prompted a decentralization of far-right groups. In line with the “leaderless resistance” concept, attacks are commonly “planned and orchestrated by a single individual or small network.” PIRUS data shows that of the 151 violent attacks or plots carried out by far-right extremists between 2013 and 2018, most individuals, 74 people, were not members of a group, and 51 were members of “an informal group of fellow extremists.” Under these conditions, radicalization occurs both on and offline.

However, online organizing is still in its infancy. Once exposed to an extremist ideology either on or offline, online platforms provide “the original meeting space” where extremists can connect and communicate with one another. Still, online ideologically-related activity “should not be mistaken for a robust transnational operational online space.” As discussed above, operational aspects of violent extremist attacks are typically planned and carried out by individuals or small groups in primarily offline settings. As one report notes, “the transnational violent extreme right-wing movement is not (yet) able to organize an ISIL-style conveyor belt linking the online operations... to offline actions that may be directly controlling, not just inspiring, certain real-life attack plots.”

This means that still being able to detect the physical location of extremist groups may aid in disrupting potential plots.

41 Ibid, p. 145.
43 National Consortium for the Study of Terrorism and Responses to Terrorism’s (START) Profiles of Individual Radicalization in the United States (PIRUS). https://www.start.umd.edu/profiles-individual-radicalization-united-states-pirus-keshif
Types of Terrorism and Targeted Violence

- The three primary types of terrorism and targeted violence within the DVE literature are terrorist incidents, hate crimes, and active shooter events.
- RMVEs/WSEs typically target racial and ethnic minorities.
- The most common attack methods are firearms, explosives and incendiary devices, and, increasingly, vehicles.
- Active shooter events concentrate in suburban and urban areas.

Domestic extremism poses a concern due to its potential to materialize in overt violence. The most prominent types of violent extremist incidents include terrorist attacks, hate crimes, and active shooter incidents (Figure 3).

Figure 3. Overview of Terrorism and Targeted Violence Categories

The first type of DVE threat is terrorist violence. As of June 2020, CSIS researchers compiled a dataset of 893 terrorist attacks and foiled plots in the United States between January 1994 and May 2020. Far-right extremism first peaked in 1995 with 43 incidents, including the Oklahoma City bombing which killed at least 168

---

people and wounded 680.\textsuperscript{47} However, the number of far-right attacks and plots in 2016, 2017, and 2019 matched or exceeded those of 1995, reaching a peak of 53 far-right-perpetrated incidents in 2017.\textsuperscript{48}

A second and related type of violence is hate crimes. The FBI reports that in 2019, there were 7,103 single-bias incidents involving 8,302 offenses.\textsuperscript{49} 55.8% of these incidents were motivated by a race/ethnicity/ancestry bias, and 21.4% were prompted by religious bias.\textsuperscript{50} While individuals who commit hate crimes may have varying degrees of ideological motivation, SPLC states that their hate group list includes groups that are ideologically based but do not necessarily participate in violent activity.

Figure 4. Hot Spots of Active Shooting Incidents in the United States, 2017-2020

A final type of security threat is premeditated targeted violence, commonly seen through active shooter events. The GVA finds that the number of active shooter incidents inside the United States has risen since 2017. Further, while active shooter events are carried out for various reasons, RMVEs/WSEs have perpetrated some of the deadliest attacks.

Though the definition of a “mass shooting” varies, data compiled by the New York Times shows that between 2011 and August 2019, suspects with ties to white extremism have carried out at least 17

\textsuperscript{50} Ibid.
active shooter attacks.\textsuperscript{51} Hate crimes also overlap with active shooter events. For example, on March 16, 2021, Robert Aaron Long shot eight people at three spas in Atlanta, including six women of Asian descent, stoking concerns over a rise in anti-Asian-related hate crimes.\textsuperscript{52} Active shooters typically target public spaces, including schools, religious congregations, and places of business. Mass public shootings often generate extensive media coverage, producing a “contagion” effect where one incident then prompts another, typically within two weeks.\textsuperscript{53}

Figure 4 plots the location of active shooting incidents between 2017-2020 based on an analysis of the GVA data. We note three interesting trends in the geography of these incidents.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{active-shooter-incidents.png}
\caption{Regional Distribution of Active Shooter Incidents, 2017-2020}
\end{figure}

First, unlike the location of domestic extremist groups, the location of active shooter incidents tends to concentrate in relatively more populous areas. Chicago, DC, New York, and Los Angeles have the largest amounts of gun violence. This is likely attributable to higher crime rates and gang-related activities in urban areas.\textsuperscript{54}

Second, not all states experience mass shootings. Rural states like North Dakota, South Dakota, Vermont, and New Hampshire have zero recorded incidents in the data analyzed. Finally, increases in active shooter incidents from 2018 to 2020 have been driven principally by rising incidents in the Great Lakes, Mid-Atlantic, and South Atlantic regions (Figure 5). This suggests that the risk is relatively concentrated in specific locations.

Within these regions, there are several patterns in the targets of terrorism, hate crimes, and targeted violence. In general, there are three types of targets: individuals, institutions, and public spaces. An analysis of 40 federal cases of RMVE attacks between 2014 and 2019 found that the most targeted sites for attack plots were religious institutions. Of the 40 attacks and plots, 10 targeted Jewish synagogues, 6 targeted Muslim mosques, and 4 targeted historically Black churches. The second most targeted sites were large public events. Similarly, a CSIS analysis found that since 2014, RMVEs/WSEs targeted religious institutions and individuals based on their religion, race, or ethnicity. In addition, between January and August 2020, 50% of far-right extremists—including white supremacists and others who opposed the BLM movement—targeted demonstrators. During this period, far-right actors also targeted government, military, and police targets (18% of incidents) and individuals based on race, gender, and other factors (18% of incidents).

Finally, there is evidence that perpetrators carry out terrorism and targeted violence using similar tactics: firearms, incendiary devices, explosives, and vehicles. The analysis of 40 federal cases of RMVE attacks between 2014 and 2019 found that the most popular attack method was shootings. With one exception, all attack plotters in the dataset who successfully carried out a lethal attack used firearms, underscoring the relationship between ease of access to weapons and potential lethality. The Anti-Defamation League’s Center on Extremism reiterates this finding, stating that “guns have been the murder weapon for the majority of extremist-related killings in every year since 2014,” including being used in 88% of DVE-related deaths in 2020. The same analysis of the 40 RMVE federal cases found that the second most popular method of attack was explosives and incendiary devices. This is consistent with a CSIS finding that from January to August 2020, 25% of far-right attackers used this method. However, during this analysis’ period, vehicles were used in 11 violent far-right attacks—27% of all far-right incidents. This represents a significant increase from only one vehicle-related attack between 2015 to 2019, again highlighting the access-impact nexus.

Risk Factors

A critical dimension of addressing DVE and targeted violence is understanding how individuals adopt extremist ideologies and radicalize to the point of engaging in IMVE. Existing research that answers this question incorporates theories and approaches from several different fields of study, including psychology, criminology, radicalization, and terrorism. Depending on which approach is applied, authors group risk factors for DVE in various ways. We group different individual and community-level risk factors in several categories: social, economic, individual background, and demographic. Figure 6

59 Ibid, p. 4.
63 Ibid, p. 5.
summarizes the findings for significant risk factors for ideologically motivated domestic violent extremism.\textsuperscript{64}

Figure 6. Community and Individual-Level Risk Factors for Domestic Violent Extremism

Social Risk Factors
A significant risk factor for DVE is an individual’s exposure to radical beliefs and behaviors. Social learning theory is a behavioral psychology concept wherein individual behavior is learned through modeling, imitation, and other social interactions.\textsuperscript{65} Patterns of interaction reinforce and influence certain behaviors, thereby impacting an individual’s likelihood of engaging in that behavior. When applied to criminology, social learning theory maintains that radicalization and criminal behavior occurs through learning practical skills, values, and internalizing these belief systems.\textsuperscript{66} Research on the causes of domestic violent extremism builds on this concept. An individual having radical peers and/or clique membership are strongly and positively

\textsuperscript{64}In the Appendix, we include an alternative illustration of these risk factors. This conceptualization captures risk factors across different levels of analysis (individual, community, regional) and how they may interact with each other.
associated with violent extremism. However, some researchers found that having a radical friend who was involved in illegal but non-violent activity and having a radical friend who was involved in a legal activity significantly decreased the odds of committing a violent attack. Still, these results reinforce the notion that radicalization and propensity for violence is a social process.

There is no clear consensus on whether having radical family members is a significant risk factor for extremism. However, this variable was commonly included in analyses, underscoring the potential role that family members can have in the preventing and countering VE process given their proximity to the radicalized individual. Group involvement or membership is another significant risk factor for violent extremism. According to the PIRUS dataset (1948-2018), about 71% of violent far-right extremists were a part of a formal (40.5%) or informal violent extremist organization (30.5%). Compared to PIRUS data from 2013, there has been a decrease in formal group membership and a rise in informal group membership. Other social risk factors that were significant in some analyses include criminal group membership and relationship troubles.

### Economic Risk Factors

In addition to social learning, lack of social control is another prevalent cause of DVE. According to the social control perspective, individuals conform to social norms by developing bonds to prosocial people and institutions. Criminality is, therefore, a result of “weak bonds to family and society, and the absence of positive turning points,” including employment, education, and a sense of personal achievement.

This literature suggests a strong, negative relationship between stable employment history and the propensity for violence. In the analyses where it was included, unemployment was positively associated with the likelihood of violence. Additionally, the literature suggests that lower educational attainment increases the likelihood of violent extremist behavior. Low educational attainment is also more prevalent for far-right extremists than other ideological groups;

---


70 National Consortium for the Study of Terrorism and Responses to Terrorism’s (START) Profiles of Individual Radicalization in the United States (PIRUS), [https://www.start.umd.edu/profiles-individual-radicalization-united-states-pirus-keshif](https://www.start.umd.edu/profiles-individual-radicalization-united-states-pirus-keshif)


the PIRUS dataset shows, for example, that 51.2% of far-right violent extremists had no college experience compared to 21.3% of far-left extremists and 42.9% of Islamist extremists.\(^77\) However, the EADR authors note that low educational attainment is no more common among extremists than the general population, indicating how this factor may interact with others to create a path towards violence.\(^78\) In sum, these social control-related factors suggest that commitment and involvement in prosocial activities restrains radicalized individuals from engaging in violence.

In addition to social control-related risk factors, two other economic indicators for extremist violence are lower socio-economic status and high income inequality. According to the PIRUS dataset, far-right extremists were more likely to come from a lower socio-economic status (25.3%), though most extremists had a middle-class background, including 64% of far-right extremists.\(^79\) In addition, another study found that income inequality produces a significant positive association with mass shootings in the periods of 1990-2000 and 2000-2010.\(^80\) Counties that experienced a one standard deviation increase of inequality observe approximately 0.43 to 0.57 more mass shootings. However, poverty rates did not yield significant results across the data sources. In general, a sense of relative deprivation may generate grievances that, coupled with other factors, could lead to violent outcomes.

### Individual Background Risk Factors

There are also risk factors related to an individual’s background. First, current or past military experience seems to be an increasingly relevant factor.\(^81\) A CSIS analysis found that there has been an increase in the percentage of domestic terrorist plots and attacks perpetrated by active-duty and reserve personnel in recent years. In 2020, 6.4% of all domestic terrorist attacks and plots (7 of 110 total) were committed by one or more active-duty or reserve members, representing a considerable increase from 1.5% in 2019 and zero in 2018.\(^82\) Second, having a criminal history, both non-violent and violent, is a significant risk factor in much of the literature.\(^83\) PIRUS data through 2013 found that 63.1% of far-right individuals had a criminal history, and 25.7% had a violent criminal past—higher than any other ideological group.\(^84\) Researchers also found that individuals motivated by white supremacist views are substantially more likely to engage in criminal behavior before radicalizing than individuals associated with other ideologies.\(^85\) Third, a history of mental or psychological illness has a strong, positive relationship with violent outcomes among radicalized individuals.\(^86\) Though

---


\(^79\) Ibid, 19.


the overall rates of mental illness in the data are low, this risk factor was significant in both qualitative and quantitative analyses. Other risk factors which correlated with violent extremist behavior in some studies were being unmarried and having a history of abuse.

**Demographic Risk Factors**

Finally, there are several demographic indicators of radicalization. First, males are more likely to engage in violence. Among far-right radicalized individuals in the PIRUS dataset, 94.2% are male, higher than all other ideological groups. Similarly, 86% of the individuals arrested for the Capital Hill insurrection were male. Additionally, compared to other ideologies, far-right extremists tend to be older. The average age at which a far-right extremist’s ideology became known is 37.6, compared to 29.7 for extremists on the far-left and 29 for Islamist extremists. For example, 66% of the Capital Hill insurrectionists were older than 34. This is consistent with the finding that while the modal age of radicalization for far-right individuals is similar to that of other groups, far-right individuals have a longer radicalization period, at five or more years.

Demographic factors related to geography may also have a role in identifying high-risk areas. One study found that, overall, areas with less diversity, more poverty, population change, and education correlate with more extremism. However, other studies which analyze the Capital Hill insurrectionists find other results. The insurrectionist-producing counties that Biden won were more urban, more racially diverse, and had higher employment. Additionally, the rate of insurrection is four times higher in counties where the percentage of non-Hispanic whites declined the most. These findings indicate the importance of local voting patterns relative to demographic change. Overall, they found that 39% of insurrectionists came from “purple” counties, and 18% came from more rural counties. These results are consistent with ACLED’s predictions ahead of the November 2020 election. ACLED reported that violence and unrest among MVE would be more prevalent in swing states, particularly in their state capitals and “periphery” towns where rural and suburban populations could gather. High-risk states were Georgia, Michigan, Pennsylvania, Wisconsin, and Oregon, while moderate-risk states included North Carolina, Texas, Virginia, California, and New Mexico.

Overall, these results highlight a broad range of factors which increase the risk of domestic extremism and targeted violence. Moreover, these
factors can interact in complex, non-linear ways to create many paths towards radicalization and potentially violent extremism. In order to explore how these potential risk factors can interact and amplify the risk of extremism and violence, we develop two supervised machine learning models.

Methodology

One purpose of this report is to forecast the highest risk areas in the United States for domestic extremism and targeted violence in 2021. We accomplish this by developing two supervised machine learning (ML) models to predict where extremism and targeted violence are most likely to occur.

ML is a form of artificial intelligence increasingly used in computer science and social science to solve complex prediction problems. It involves a set of computer algorithms that attempt to “learn” patterns in existing/historical data and extrapolate predictions based on this information. It provides a parsimonious and generalizable model to assist in forecasting risk in future cases. Further, machine learning is strongly aligned with policymaker needs because “predictive heuristics provide a useful, possibly necessary, strategy that may help scholars and policymakers guard against erroneous recommendations.”

Data Sources

We use county-level data from the U.S. Census Bureau American Community Survey dataset to capture many of the risk factors for extremism. This data provides fine-grained information about changing population, socio-economic, demographic, and other factors from year to year. For each model, we examine trends in historical data from 2018-2020 to determine where domestic extremist groups and targeted violence were most likely to occur. We then use these patterns to forecast predicted risk estimates across counties in 2021.

Data on domestic extremist groups comes from the SPLC Hate Groups Dataset. This project records groups which have not necessarily used violence yet, but may potentially do so. We choose to include information about groups with the potential to use violence in order to avoid a selection bias which could come from only studying the most violent groups.

To understand where extremism is most likely to arise, we use SPLC records on the ideology, geographic location, and operational years of approximately 3,752 groups between 2017-2020. The dataset records an observation for every year that the group is operating and is increasingly used to study domestic extremism in the US. A benefit to SPLC data is that it provides detailed data on where these groups form and applies consistent coding rules across time, which mitigates concerns that it could be recording varying numbers of groups as internet and communication technologies make it easier to trace these groups. A limit to this data is that it may include groups with little political aim or opposition against the state. It also includes group operations rather than group formation.

---

Data on active shooter incidents comes from the Gun Violence Archive (GVA) Project. This project records near-real-time data about gun violence using open-source data from approximately 7,500 sources, including local and state police, media, and government sources. The project manually goes through candidate entries and includes snapshots of articles in their listings to validate their entry.

A limit to the GVA is that it does not differentiate between gun violence incidents motivated by criminal, hate, or political ideologies. This means there is a risk that it could overestimate the location of mass shootings in areas highly susceptible to high rates of criminal activity. It could also make it harder to parse out key identifying variables which differentiate between different types of violence. However, surveys of different gun violence datasets consistently find it has the best (1) real-time data accumulation and (2) broadest set of cases. Future research could attempt to go through the narrative records for each incident to clean up and exclude crime-or gang-related activity in order to make a more precise model. As a robustness check, we change the inclusion criteria to filter out smaller incidents which do not match the federal definition of “mass killing.”

For each incident or group entry, we use geospatial software to map the location of these different groups and link their records to different county-level records. This allows us to measure how varying community-level inputs can increase the risk of extremist group operations or mass shooting incidents in a given county. We exclude a minority of entries that cannot be tied to a second-order administrative district (county). In practice, this includes groups with “statewide” activity such as the Patriot Front, a white nationalist group with a large online presence and mixed physical presence across college campuses in Texas. This dataset provides a much stricter criteria for inclusion and thus omits many potential cases of criminal or gang-related activity from the analysis.

Finally, we add open-source data on ports within the US. We collect port data from the U.S. Department of Transportation/Bureau of Transportation Statistics’ (BTS) National Transportation Atlas Database (NTAD) using the ArcGIS online dataset. This captures all navigable rivers, waterways, and ocean entries. We overlay this geospatial information to identify where domestic extremism and targeted violence are most likely to occur within the maritime domain.

Research Design
We build a 10-fold repeated cross-validation (CV) random forest model for analysis. This method partitions the data into a training and test set. We use an established method known as cross-validation to optimize the model’s performance on an out-of-sample test set of cases. Cross-validation estimates a model on the training set then calculates that model’s prediction power when applied to a test set. This process works by training the model on a subset of data and then seeing how well it performs on a reserved validation set of observations. This procedure is repeatedly run using different groups of observations in the training result to ensure individual data points do not unduly influence the results. Following convention, we choose a cross-validation rate of 10 folds to minimize the classification error of the data.

Our ML model also uses a random forest algorithm. This is a non-parametric or “black box” approach, meaning it does not provide insight into what variables are being used—and in what combination—to reach its predicted results. This approach is preferable to traditional regression methods because it can capture potential non-linearities or interactions in the data. If the underlying functional form relating risk factors to domestic extremism and targeted violence is not truly linear, then imposing a more rigid variable selection technique—such as regularization or subset selection—could lead to worse model performance. Scholars increasingly use tree-based methods for ML analysis because this approach requires very few assumptions about the input data and provides model flexibility, resulting in better forecasting abilities.106

Random forests bootstraps (sample) different observations B times from the training set and construct a single decision tree for each sample. Each time the tree considers a new decision rule, it does so by considering a split among a group of randomly sampled variables chosen from our input data. This helps decorrelate the predictions between trees. The algorithm then averages the predictions across trees to make estimates on an out-of-sample training set.

The input variables are county-level risk indicators for extremism and targeted violence (Figure 6). The output variable is whether an extremist group or mass shooting incident occurred in that county in a given year. Since these two phenomena are relatively rare, we correct for class imbalance in the dataset. Specifically, we oversample the outcome of interest in our training set to roughly 50% of the observations. Class imbalance can lead to a highly accurate, but non-informative model. That is, by simply predicting the absence of extremist violence in all counties, the model can appear to achieve an overall high accuracy rate but have a very low sensitivity. To correct for class imbalance and improve the sensitivity of the model, we follow conventional practice and downsample the non-extremist cases to reach relative parity between classes.107

Applications

The final research question of this report asks: how can machine learning models help mitigate vulnerabilities to domestic extremist groups and active shooter incidents? We describe two applications to show how machine learning can leverage existing research on RMVE and targeted violence to identify high-risk areas.

Predicting Domestic Extremism

Can machine learning help practitioners detect where domestic extremists are most likely to operate? In short, yes. The results of a ML prototype model accurately forecast the county-level location of domestic extremist groups with 96% accuracy and an 85% sensitivity rate. This is significant given that extremists are often clandestine when they form and may be hard to detect at their initial mobilization and before there are overt warning signs of violence.

To yield these results, we build a model which trains on information about the location of domestic extremist groups from 2017-2019 and then test its predictive accuracy based on available information about groups operating in 2020. We find a 96% accuracy rate on the out-of-sample set of cases (Model 1). This means historical patterns in domestic extremist activity can predict which counties are most likely to see future extremism. In practical terms, the model has established that it can accurately predict extremism in 3,017 of 3,143 counties around the United States.


Importantly, the sensitivity of the model is very high. This is significant because accuracy, by itself, might not always be a good measure of model performance. If extremist groups only form in 5% of counties, then the model could appear to have a very high accuracy rate of 95% by simply predicting no group forms in any county. However, this model would provide little information; it would not tell us where groups are most likely to form. Here, the model's high sensitivity rate means it accurately identifies where domestic extremists operate 85% of the time.

Table 3. 10-Fold CV Random Forest Model Results of Domestic Extremism

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremist Types</td>
<td>All</td>
<td>WSE</td>
<td>Neo-Nazi</td>
<td>Neo-Confederate</td>
<td>Anti-Immigrant/Muslim</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Sensitivity Rate</td>
<td>0.85</td>
<td>0.71</td>
<td>0.91</td>
<td>0.62</td>
<td>0.89</td>
</tr>
<tr>
<td>Specificity Rate</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Kappa Score</td>
<td>0.76</td>
<td>0.52</td>
<td>0.60</td>
<td>0.67</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Figure 7. Predicted Risk of Ideologically-Motivated Extremist Groups, 2021
We use Model 1 to estimate the predicted risk of future domestic extremism across counties. Figure 7 maps the predicted risk across counties. We find that counties with the highest risk of domestic extremism have a history of previous group activity. For example, Tampa, Florida first recorded a Proud Boys faction forming in 2018. This faction has been continuously active since. In 2021, a member of this faction was arrested and charged by police for his involvement in the Capitol insurrection. Since the model is based on structural factors like measures of unemployment, income inequality, and educational levels—and these factors tend to change very slowly over time—there is a high baseline risk of continued extremism across counties.

In addition to these counties, the machine learning model also flags several high-risk counties without a history of DVE in 2020 (Figure 8). For example, the model predicts Nevada County, California has a 75% probability of seeing a new domestic extremism group form in the next year. The county previously had one known hate group in 2018, a white ethnonationalist neo-Volkisch group known as Asratu Folk Assembly, but it disappeared in subsequent years. However, environmental risk conditions have not changed much in Nevada County, and there are active neo-Volkisch chapters in neighboring counties.

---


110 For example, the Asratu Folk Assembly has a chapter in Brownsville in neighboring Yuba County.
We conduct several model extensions which specifically aim to predict the risk of white supremacy, neo-Nazi, neo-Confederate, and anti-immigrant/anti-Muslim groups (Models 2-5). The results in Model 2 aim to predict the location of WSE groups like the Base, which had chapters in Georgia and Michigan. The WSE model has a 98% accuracy and 71% sensitivity.

The white nationalist model also finds a history of extremism to be one of the most important indicators of future activity. Areas such as Wayne County, Michigan have high risks of continued WSE because of the presence of many organizations like White Rabbit Radio, the National Socialist Movement, and Northern Hammerskin.

Models 3-5 find similarly high rates of accuracy when examining the propensity of neo-Nazi, neo-Confederate, or anti-immigrant hate groups to organize. In the Appendix, we include the predicted risk estimates for the top 10 counties for each region.

There are a few limits to this analysis. First, this analysis is based on the physical location of known groups. This means it is unlikely to detect lone actors or actors who primarily organize in online spaces. Second, this analysis focuses on DVE and not anti-state extremists. This means it does not pick up activities by militia groups like the Oathkeepers, III Percenters, or other militia movements. Finally, people can move across counties to evade detection. This means due attention should be paid not only to high-risk counties, but the surrounding areas as well.

Real-World Example - The Base:
Two members of the Base were arrested in October 2020 following an incident in Dexter, Michigan (Washtenaw County). Justen Watkins, 25, of Bad Axe (Huron) and Alfred Gorman, 35, of Taylor (Wayne) were both taken into custody and are lodged at the Washtenaw County Jail.
Predicting Targeted Violence

Based on conversations with DHS stakeholders over the last year, a central concern related to domestic violent extremism, terrorism, and targeted violence is active shooter incidents. Understanding where these incidents are most likely to occur can better inform security measures, counter-extremism programs, and other ways to mitigate the risks of violence.

We build a model which trains on information about active shooter incidents from 2017-2019 and then test its predictive accuracy based on available information about incidents in 2020. Model performance metrics are in Table 4. We find a 92% accuracy rate on the out-of-sample set of cases, meaning historical patterns of active shooter incidents predict the location of 2020 incidents nine times out of ten (Model 1). A limit to this model is that it has a lower sensitivity rate. This means that the model was very good at predicting the absence of an active shooter incident in some areas (e.g. North Dakota), but struggled to predict incidents in other areas. This lower sensitivity result might be because we assess the model’s accuracy based on incidents in 2020, which was seen as an outlier in targeted violence due to the pandemic.111

We also run several robustness checks to see how the results change when we only focus on “mass shootings” (Models 2 and 3). We also try to validate the results on data outside of 2020. One critique is that the pandemic reduced the number of conventional active shooter incidents, so one would expect the model to perform poorly in predicting relative numbers that year. We examine how well the model predicts active shooter incidents from January 1 to May 31, 2021, and find comparable accuracy results. The sensitivity of the model—the ability to accurately predict “true positives” for the location of incidents when they happen—is 71%.

Table 4. 10-Fold CV Random Forest Model Results of Active Shooter Incidents

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>Alternate Validation Set (Jan 1.-May 31, 2021)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incident Types</td>
<td>All</td>
<td>Min. 1 Killed</td>
<td>Min. 3. Killed</td>
<td>Alternate Validation Set (Jan 1.-May 31, 2021)</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.92</td>
<td>0.94</td>
<td>0.97</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Sensitivity Rate</td>
<td>0.51</td>
<td>0.44</td>
<td>0.35</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Specificity Rate</td>
<td>0.96</td>
<td>0.96</td>
<td>0.98</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Kappa Score</td>
<td>0.47</td>
<td>0.39</td>
<td>0.32</td>
<td>0.37</td>
<td></td>
</tr>
</tbody>
</table>

As before, we use this model to map the predicted risk of future incidents in 2021 (Figure 9). The results vary slightly from the domestic extremism model. In other words, counties at risk for domestic extremism sometimes, but not always, are at risk for active shooter incidents as well. One of the main deviations is relative population. The active shooter model tends to predict higher risk in more highly-populated areas such as Los Angeles County, King County, or Miami-Dade County.

The model also flagged several areas which had not yet experienced an active shooting in 2021, but were at high risk. This included, for example, Palm Beach County. A few days after the completion of data analysis, an incident occurred at a grocery store in that county.

https://www.nytimes.com/2021/03/19/us/mass-shootings.html
**Real-World Example - Palm Beach County:** A gunman entered a grocery store in Royal Palm Beach (Palm Beach County) on June 10, 2021 and shot two people before killing himself. According to The New York Times, “the authorities said there was no known relationship between the gunman and the two victims. A motive for the shooting still had not been determined.”
Extension: Identifying Vulnerabilities in the Maritime Domain

As an extension, we show how information about domestic extremism and targeted violence can be used to identify high-risk areas in the maritime domain. A 2018 DHS report on soft targets found that “sports venues, schools, and transportation systems” are vulnerable to active shooter incidents due to the high concentration of people within relatively confined spaces.\(^{112}\) Given the Maritime Transportation System spans 25,000 miles of navigable waterways, 3,500 terminals, and other areas, prioritizing sectors of the MTS for these incidents can improve counterterrorism preparedness.

Conversations with key stakeholders in the maritime domain often situated their concerns about homegrown violent extremism in the context of insider threats, infiltrations, and active shooter incidents. For example, there are concerns that an active shooter incident could unfold at a port facility. This would not only endanger the large swath of individuals who work at ports, but also potentially disrupt supply chain and economic commerce flows in and out of the port. Similarly, there are concerns that an active shooter incident may unfold within the maritime transportation system such as a cruise ship or ferry.

Although there is little precedent for these incidents inside the United States, such events have occurred elsewhere. In 1985, members of the Palestine Liberation Front hijacked the Italian cruise ship *Achille Lauro* off the coast of Egypt for three days.\(^{113}\) In 2004, two Palestinian suicide bombers managed to enter Israel’s second busiest port, the Port of Ashdod, detonate their explosives, and kill 10 individuals.\(^{114}\) Finally, that same year, members of the Salafist extremist Abu Sayyaf Group in the Philippines bombed a ferry in Manila, killing 116.\(^{115}\) These events demonstrate the vulnerability of maritime infrastructure to violent extremism. Understanding which maritime vulnerabilities are most susceptible to these incidents is important to better prioritize resources to counter these risks.

We focus on the 805 counties which have at least one piece of maritime infrastructure, according to the Department of Transportation. We then compare the predicted risk of active shooter incidents from Model 1 within those counties. Figure 10 shows how the predicted risk of an incident maps onto the location of different maritime infrastructure.

Similar to the main active shooter model, the model often prioritizes areas with relatively larger populations (e.g. Jefferson County, Kentucky or Cuyahoga County, Ohio). However, the model does provide some surprising predictions. For example, it predicts that the highest at-risk county in California is the 15\(^{th}\) largest county in California: San Joaquin County, which is home to the Port of Stockton. San Joaquin County is one of the more racially diverse counties in California; 40.5% of residents identify as Hispanic or Latino. It also has some of state’s lowest high school graduation rates and below-average median income. While it is hard to know if and how these risk factors contribute to the model’s expected risk, it could explain the model’s forecasts.


As a final assessment, we show these county-level risk assessments correspond to the location of different maritime sites (Figure 11). For example, in Florida the model predicts relatively higher levels of active shooter incidents in the Miami-Dade metropolitan area than the Tampa-St. Petersburg or the Fort Walton Beach-Pensacola panhandle. Since the Port of Miami is a central hub for the cruise industry, this suggests it could be a target for a cruise ship attack like the Achille Lauro incident.

In contrast, in Michigan the model predicts relatively low levels of active shooter incidents along Lake Michigan. This may mitigate concerns of active shooter incidents involving targets such as the cross-lake SS Badger ferry system operating out of Mason County.

Overall, we find that large populous areas remain highly vulnerable to future incidents. However, there are other areas with smaller populations—such as Lafayette Parish, Louisiana (Pinhook Bridge, Bankers Ferry, Private Industry Docks), Liberty County, Texas (Port of Liberty near the Trinity River), and Plymouth County, Massachusetts (USCG Station, Wharf, Marina)—which also have high risk levels. These results demonstrate the potential of machine learning for specific stakeholders in the homeland security enterprise.
Conclusion and Policy Recommendations

Two random forest models to forecast domestic extremism and targeted violence find:

- RMVE extremist groups and active shooter incidents tend to emerge in the same areas from year to year.
- Community-level risk factors forecast high-risk areas with large predictive accuracy.
- ML can prioritize risk across different counties in order to guide counterterrorism efforts.

Overall, this report provides data-driven guidance for practitioners about where extremist actors and violent incidents are most likely to emerge moving forward. We find that community-level risk indicators of radicalization strongly predict which counties are at greatest risk of group operations and active shooter incidents. While machine learning provides results with 92-96% accuracy, the black box nature of these algorithms means it is difficult to know exactly how these factors interact with each other to raise the risk of extremism.

One natural extension of this prototype model could attempt to detect the location of militia violent extremists (MVE). Understanding whether MVE form and operate in the same area as RMVE can help isolate whether there are common causes of radicalization or whether certain attributes make some types of domestic extremism more likely than others.

If community-level risk factors predict risk across different counties, then community-driven solutions can mitigate these risks. We identify three community-driven responses which can address the risks identified here.

Building on the initiatives undertaken by the DHS Center for Prevention Programs and Partnerships (CP3), the results suggest that increasing community awareness and building community resilience can address these risks.
Regional Prevention Coordinators are working through local school boards to connect potentially at-risk individuals and their families to the relevant intervention network.\textsuperscript{116}

Further, ongoing programs in Australia, New Zealand, and Canada may provide a template for building community resilience. For example, increased community engagement in Australia aims to create tailored training and accessible platforms through which parents, families, and other gatekeepers can recognize and react to signs of extremist engagement (see Community Awareness Training (CAT) Manual). The program provides information about the processes by which individuals become radicalized and how to respond. CAT also has “train-the-trainer” sessions for community members and service providers who can share this information with others.\textsuperscript{117}

Finally, it is important to ensure that at-risk individuals have the necessary access to interdisciplinary teams, particularly at a community level. Canada’s established “Community Safety” Hubs are one model for providing integrated, coordinated responses to imminent threat situations within 24 to 48 hours.

While domestic extremism and targeted violence result from a complex interaction of factors, this report demonstrates that there are identifiable patterns in the literature and reliable data-driven models that can inform the creation of effective community-level interventions.

## Appendix

### Table A1. Community- and Individual-Level Risk Factors for Domestic Violent Extremism

<table>
<thead>
<tr>
<th>Community-Level</th>
<th>Individual-Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social</td>
<td>Economic</td>
</tr>
<tr>
<td>Radical peers and/or clique membership</td>
<td>Unstable employment history</td>
</tr>
<tr>
<td>Extremist group membership and/or involvement</td>
<td>Unemployment</td>
</tr>
<tr>
<td>Criminal group membership</td>
<td>Lower educational attainment</td>
</tr>
<tr>
<td>Relationship troubles</td>
<td>High income inequality</td>
</tr>
<tr>
<td></td>
<td>Lower socio-economic status</td>
</tr>
</tbody>
</table>
Figure A1. Top 10 At-Risk Counties for Domestic Extremism by Region, 2021
Figure A2. Top 10 At-Risk Counties for White Nationalist Groups by Region, 2021
Figure A3. Top 10 At-Risk Counties for Anti-Immigrant/Anti-Muslim Groups by Region
Figure A4. Alternative Grouping of Risk Factors for Domestic Violent Extremism