SOCIAL SCIENCES

Predicting non-state terrorism worldwide

Andre Python[1,2]*, Andreas Bender[3], Anita K. Nandi[2], Penelope A. Hancock[2], Rohan Arambepola[2], Jürgen Brandsch[4], Tim C. D. Lucas[5]

Several thousand people die every year worldwide because of terrorist attacks perpetrated by non-state actors. In this context, reliable and accurate short-term predictions of non-state terrorism at the local level are key for policy makers to target preventative measures. Using only publicly available data, we show that predictive models that include structural and procedural predictors can accurately predict the occurrence of non-state terrorism locally and a week ahead in regions affected by a relatively high prevalence of terrorism. In these regions, theoretically informed models systematically outperform models using predictors built on past terrorist events only. We further identify and interpret the local effects of major global and regional terrorism drivers. Our study demonstrates the potential of theoretically informed models to predict and explain complex forms of political violence at policy-relevant scales.

INTRODUCTION

Research on armed conflict and insurgency has led to the development of predictive models informed by theory (*1*–*7*), which includes a recent successful research program that applies machine learning algorithms to predict conflict at fine spatiotemporal scales (*8*–*10*). This literature highlights that theoretically informed modeling strongly improves the predictive performance of machine learning techniques. However, this important insight from armed conflict research (*11*, *12*) has not found its way into terrorism studies, yet. Terrorism research, by and large, has focused on explanatory models, applying statistical approaches to capture and quantify the effects of drivers of terrorist attacks in space and time (*13*–*15*). There is one notable exception; Ding *et al.* (*16*) predict terrorism at a fine spatial scale. However, the authors aggregated the data on a yearly level, which ineluctably ignores the short-term dynamics of terrorism and cannot provide the necessary information for policy makers to implement targeted and rapid counterterrorism measures. Consequentially, there is a need to develop an interpretable modeling framework to predict terrorist events at fine spatial and temporal scales that can help policy makers implement efficient interventions and assess and develop theories at relevant scales.

Here, we develop a framework to predict and explain the occurrence of terrorism perpetrated by non-state actors outside legitimate warfare worldwide a week ahead and at a fine spatial scale. We compare the results of a flexible spatial statistical model [generalized additive model (GAM)] and two machine learning approaches, including an efficient implementation of gradient-boosted trees [extreme gradient boosting (XGB)] (*17*) and a random forest (RF) algorithm (*18*). Our models are informed by an advanced theoretical understanding and explicitly incorporate (i) structural features—time-invariant variables that account for the effect of, e.g., rough terrain or per capita gross domestic product (GDP)—and (ii) procedural features—dynamic variables to account for the temporal dependency of terrorism (*15*). We compare their predictive performance with benchmark models that include structural and/or procedural features. Parsimonious models using procedural features exclusively may show high predictive performance in short-term forecasting of conflict events [e.g., civil war forecast at country-month level (*19*)]. The results are used to provide guidance on the choice of features associated with short-term predictions of terrorist events, which remains under debate in the conflict literature (*19*).

We consider terrorist attacks perpetrated between 2002 and 2016, divided into 13 regions worldwide, which include all subcontinental regions defined in the Global Terrorism Database (GTD) (*20*) and West Africa (Fig. 1, regions A to M). Predictive models are built for each region, which allow us to identify, assess, and compare the role of major terrorism drivers across different regional contexts and reduce the computational requirements. In addition, training machine learning models for separate regions allows the algorithms to select different hyperparameters for different regions (e.g., regions with lower prevalence require greater regularization). The spatiotemporal domain is represented by regular cells at 0.5° spatial resolution from PRIO-GRID (*21*) that cover all inhabited areas (with population density above five persons per square kilometer) where terrorist events can potentially occur and be reported with high spatial accuracy in the GTD. The grid cells are replicated for each week from 2002 to 2016 across all regions worldwide. The total number of week cells considered in this study are 21,108,045 (26,551 grid cells × 795 weeks). Countries that did not encounter terrorist events (or those where no terrorist event was reported in GTD) over the study period are not considered (see the Supplementary Materials for details on the data preparation).

A definition focusing on attacks against civilians is in line with recent research (*22*, *23*) but necessitates modifications of the raw GTD data. Therefore, we first gather data on terrorism from the GTD dataset from 2002 to 2016 for each region and adapt it to our definition (*20*). Second, we discretize each region using PRIO-GRID, a 0.5° spatial grid that provides various geographic and socioeconomic variables worldwide (*21*). Third, we extract in all week cells in each region worldwide: (i) The occurrence of terrorism using GTD data (2002–2016) and (ii) feature values from a set of predictors gathered at various spatiotemporal lags. The choice of the structural and procedural features included in the predictive model builds on our conceptual understanding of terrorism as a tactic, and we identified several variables theoretically associated with its probability of occurrence (see list of features in table S1).

We define terrorism as politically motivated attacks outside legitimate warfare (i.e., targeting noncombatants) perpetrated by non-state

[1]Center for Data Science, Zhejiang University, Hangzhou, P.R. China. [2]Li Ka Shing Centre for Health Information and Discovery, Oxford University, Nuffield Department of Medicine, Big Data Institute, Oxford, UK. [3]Department of Statistics, LMU Munich, Munich, Germany. [4]Bonn International Center for Conversion (BICC), Bonn, Germany. [5]School of Public Health, Imperial College London, London, UK.
*Corresponding author. Email: apython@zju.edu.cn

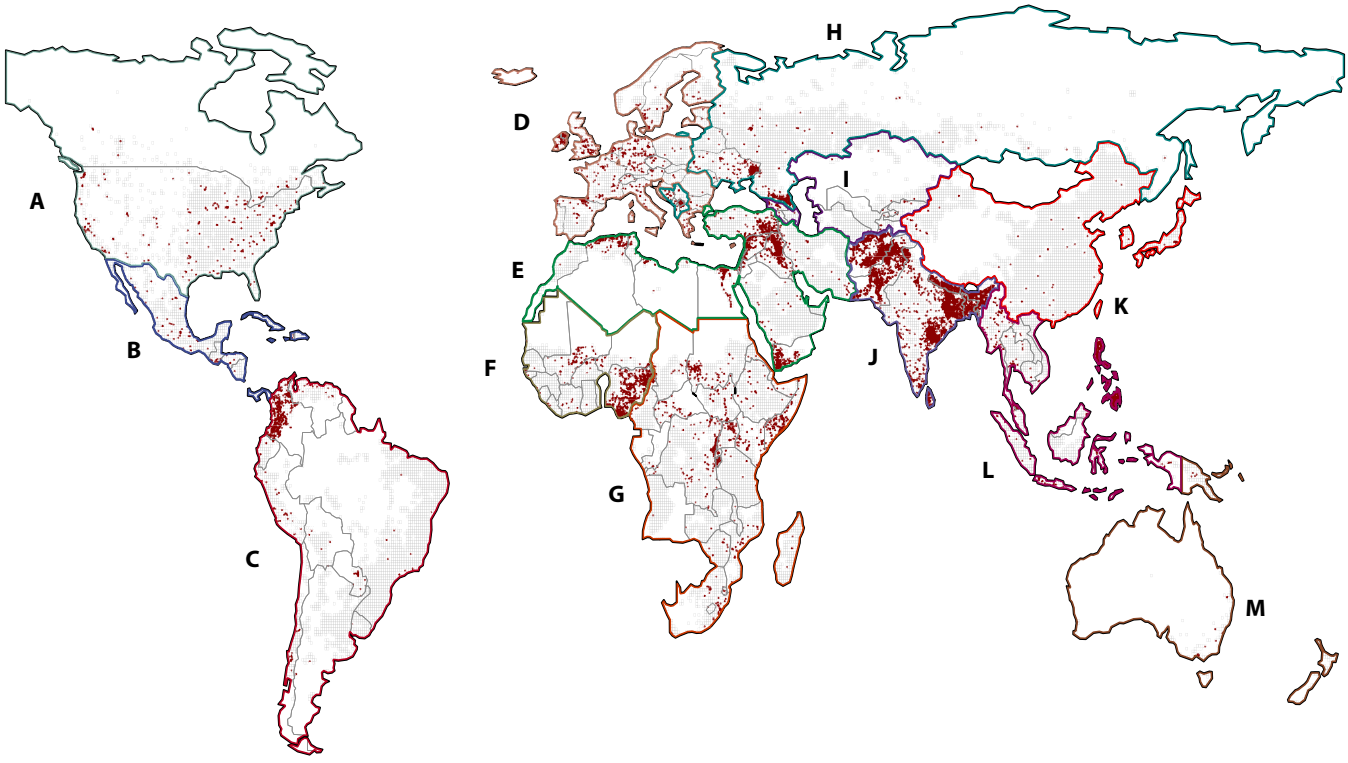

Fig. 1. Location of terrorist events (2002–2016) and spatial domain (grid cells) worldwide. Location of terrorist events (red dots) from the GTD (*50*) that occurred from 2002 to 2016 in North America (**A**), Central America and Caribbean (**B**), South America (**C**), Europe (EU28 and Schengen area) (**D**), Middle East and North Africa (MENA) (**E**), West Africa (**F**), sub-Saharan Africa (**G**), Russia and Eastern Europe (**H**), Central Asia (**I**), South Asia (**J**), East Asia (**K**), Southeast Asia (**L**), and Australasia and Oceania (**M**). The spatial domain (PRIO-GRID cells, gray squares) (*21*) covers all inhabited areas worldwide (population density above five persons per square kilometer).

actors to communicate to a wider audience. Here, terrorism is understood as separate from armed conflict. While armed conflict is a violent process between two or more armed actors, terrorism is a tactical tool used either inside or outside of armed conflicts that targets the unarmed (*24*).

Terrorism, defined as violence against civilians, is a weapon of the weak. Non-state actors who are unable to overpower their adversaries militarily use this tactic to compensate for their weakness. Terrorism thus may have many purposes (*25*), but its two core strategies are either coercing or provoking an opponent or a population (*22*, *26*). In coercion strategies, terrorists use violence to threaten further violence in the future. The aim here is to coerce the target audience into giving in to a demand of the terrorists (*27*, *28*). Provocation, to the contrary, seeks to lure the attacked groups into a counter attack that in the end benefits the terrorists (*22*, *29*). In both cases, terrorism is a communication tool. Violence itself is used to send messages to an observing population. As a consequence, we first identify structural variables that increase the possibility of successfully communicating with an audience for the purpose of either provocation or coercion. Note, however, that the variables included here also have relevance for other theoretical considerations. To capture the communicative value of a location, we include closeness to capital, distance to large cities, road density, and population density as appropriate variables. High population density and facilitated access by roads or closeness to large cities will, in general, increase the ease by which the communicated message travels through the audience. Thus, these locations should constitute more attractive targets (*30*, *31*).

Second, terrorists often operate in remote and difficult to access regions where state authorities cannot reach them easily. Therefore, we included variables that indicate a geographical advantage for terrorists, such as closeness to international borders, mountainous coverage, and altitude. While these locations do not provide the terrorists with high impact targets, they may be closer to their home base and thus easier to reach for attacks. We also included a variable on the locations of drug cultivation as drugs usually represent potential financial resources for terrorist groups. It is worth stressing here that the variables associated with the communication value of a location can similarly reflect geographical advantages for terrorists as, for instance, low values in road density will also identify remote areas.

Third, coercive attacks have a higher impact in areas where they hurt the target the most. Coercion seeks to maximize the expected costs of future attacks, and high costs in the present can credibly establish a threat of high costs in the future. For this reason, we included proxies of economic activity from satellite night lights and gross cell product data, the infant mortality rate as a proxy for well-being, and data on some key resources that can provide a considerable income to governments. We account particularly for the presence of valuable resources, such as onshore petroleum, diamond, gem, and gold deposits. Again, some of the variables included can have multiple interpretations. Particularly, gross cell product and satellite night lights data also identify remote areas.

Fourth, provocative attacks seek to engender a violent reaction either by the government or the local population. Previous research showed that terrorists particularly seem to target those locations that have a high probability of escalating into more violent conflict because this creates more recruits for the terrorists' organizations (*22*, *32*). In line with this research, we include variables on past conflict events (*32*) and politically excluded ethnic groups (*33*–*35*). We use a yearly lagged sum of conflict events (state violence only) and integrate the number of politically excluded ethnic groups.

Fifth, on the country level, we also consider the political system and GDP as important factors. While there is considerable controversy about the association between political regimes and terrorism, it is generally assumed that terrorist strategies are affected by the level of democracy (*36*). Therefore, we include a proxy for estimating liberal democracy. In addition, the overall economic development of a country may have an influence on terrorism (*37*). Although we already included a local level variable on GDP, the development of a country's economic system may help to predict the location of terrorist attacks.

Last, the repeated occurrence of terrorism is highly dependent on terrorist organizations gaining a local foothold. Once terrorist organizations establish a presence in a given area (either in a country or in a specific location), we expect that attacks will be highly autocorrelated. Therefore, we built procedural variables associated with previous attacks using various temporal and spatial lags. These variables are particularly important in areas of high terrorist activity as they pick up information about the terrorist group's internal calculation. Previous attacks reveal where terrorists are able and willing to strike and can thus indicate where future attacks may occur.

The resulting data that include temporally and spatially lagged features for each week cell from 2003 to 2016 are used to train the investigated models over several years, with weekly predictions carried out a year ahead to assess the out-of-sample predictive performance of the models (workflow in fig. S1 and predictive maps in fig. S4). For each region (Fig. 1, regions A to M), we compare the predictive performance on hold-out data between 2012 and 2016 of models using structural and procedural features (XGB, RF, and GAM) with those from benchmark models using procedural features only [autoregressive models of order 1 AR(1) and of order 2 AR(2)].

RESULTS

To estimate the general predictive performance of the models, we compute receiver operating characteristic (ROC) curves, which plot the true-positive rate (also called sensitivity or recall) against the false-positive rate (1-specificity or probability of false alarm) at various thresholds. Hence, we compute the area under the ROC curves (AUROCs), which summarizes the model's performance over a range of thresholds, with values ranging from 0 (worse prediction) to 1 (perfect prediction), with 0.5 for a random classifier. The average AUROC values of the best models range from 0.81 to 0.97 across all regions worldwide, which represents a good overall predictive performance (*19*).

Given the potential large damages caused by terrorist events, it is important that the precision of the models is high enough so that a large fraction of them can be prevented. The precision of a model refers to its ability to correctly predict the proportion of week cells that encountered terrorism over the total number of predicted positive cases. Obtaining high precision remains challenging, especially in regions that exhibit a low level of terrorism prevalence. Prevalence is defined as the proportion of positive cases over the total number of positive and negative cases. The AUROC metric is not suitable to assess the precision of the models if the data are imbalanced. In all regions, most week cells did not encounter terrorism during the study period, which leads to a considerable excess of negatives over positives. To account for data imbalance, we compute the area under the precision-recall (PR) curves (AUPRCs) and report the complementary predictive performance metrics that account for the precision of the models (tables S2 and S3). In contrast to the ROC curves, the baseline for a PR curve (random classifier) is determined by the prevalence (*38*). For an imbalanced class distribution with the ratio of positives over negatives of, e.g., 1 to 9, the baseline function and the corresponding AUPRC are equal to $1/(9 + 1) = 0.1$. A perfect AUPRC is reached if all week cells that encountered terrorism are predicted as positive events without accidentally predicting any week cell that did not encounter terrorism as positive.

Tables S2 and S3 show that XGB and RF have the best AUPRC values overall, with AUPRC values ranging from 0.2 to 2.0% in regions with very low prevalence (A, B, I, K, and M) and from 17 to 51% in regions with high prevalence (E, J, and L). The important variation in the AUPRC values associated with terrorism prevalence indicates that predicting correctly a high fraction of positive week cells while keeping a low proportion of "false alarm" week cells is more difficult to achieve in areas where terrorism is rare. However, the AUPRC values for XGB and RF remain usually relatively high. For example, in Central Asia (region I) where the prevalence of terrorism is very low (0.022%), the AUPRC of the best model (RF) is about 63 times higher than the AUPRC for the baseline model. To account for the capability of the models to detect positives, we compute AUPRCs and alternative predictive performance metrics (see tables S2 and S3 for detailed results).

We observe that XGB, RF, and GAM show better predictive performance scores than the benchmark models [AR(2), AR(1) extra, and AR(1)] in most regions based on AUROC, AUPRC, and additional performance metrics (see Materials and Methods). The high predictive performance of XGB, RF, and GAM is particularly noticeable in regions with higher level of terrorism prevalence. While XGB, RF, and GAM show relatively high AUROC scores, the machine learning algorithms (XGB and RF) appear to show a better model precision overall, with top AUPRC values obtained in a larger number of regions compared to GAM. Thus, we compute the Brier skill score (BSS) (*39*) to assess both the calibration (statistical consistency between the probabilistic forecasts and the observations) (*40*) and the sharpness (concentration of the predictive distributions) (*41*). We observe that all models perform poorly in regions with very low terrorism prevalence. However, in regions with medium-high prevalence, XGB and GAM show the best BSS scores (tables S2 and S3). These results suggest that, in regions with a relatively high level of terrorism prevalence, models that incorporate both structural and procedural features are likely to produce more reliable and accurate short-term predictions at fine spatial scales than models that would only include procedural features.

While GAM and RF show relatively high performance in some of the investigated predictive scores, XGB systematically ranks among the best predictive models, this independently of the metric used and the region considered. Therefore, we select XGB to assess the role of important variables in each region (A to M) worldwide. For each regional model, we compute the importance (in percent)

of each feature (20 structural and 14 procedural features) using the variable importance metric. This metric represents the relative influence of each variable in the model based on criteria that include its role on a split in a specific tree and the improvement of predictive performance over all trees. It is scaled with a maximum value of 100 (*42*). It characterizes the overall effect of a feature in comparison with other features in the model. In contrast to a regression coefficient, it does not provide the absolute magnitude and the direction of the effect.

To assess the global importance of specific drivers of terrorism, we reported the average importance of each variable across all regions. Figure 2 shows the structural (label in black) and procedural (label in gray) variables ranked by their average importance across all regions (column A:M). Time since previous terrorist event is the most important variable overall using this metric. Because previous attacks capture multiple facets of the terrorists' strategic consideration (e.g., target value versus reachability), it is no surprise that they have the highest predictive power in most regions worldwide. In addition, five variables also rank very high on a global scale. First, three variables that we see associated with the communicative value of a location (population density, travel time to the nearest large city, and road network density) appear to be important in the prediction of terrorism in most regions. Second, two variables associated with the coercive value of a location (satellite night lights and gross cell product) are also strongly related to terrorism across regions. Quite interestingly, variables associated with the provocative value or with easy-to-hide-in locations do not feature in this upper group and neither do country-level variables.

We further explore the regional variation of the marginal effects of the four most important features (Figs. 3 to 6) across all regions (see figs. S9 to S38 for additional features). In our study, we do not distinguish lethal from nonlethal terrorist attacks. However, lethal events are more likely to be reported and are therefore less susceptible

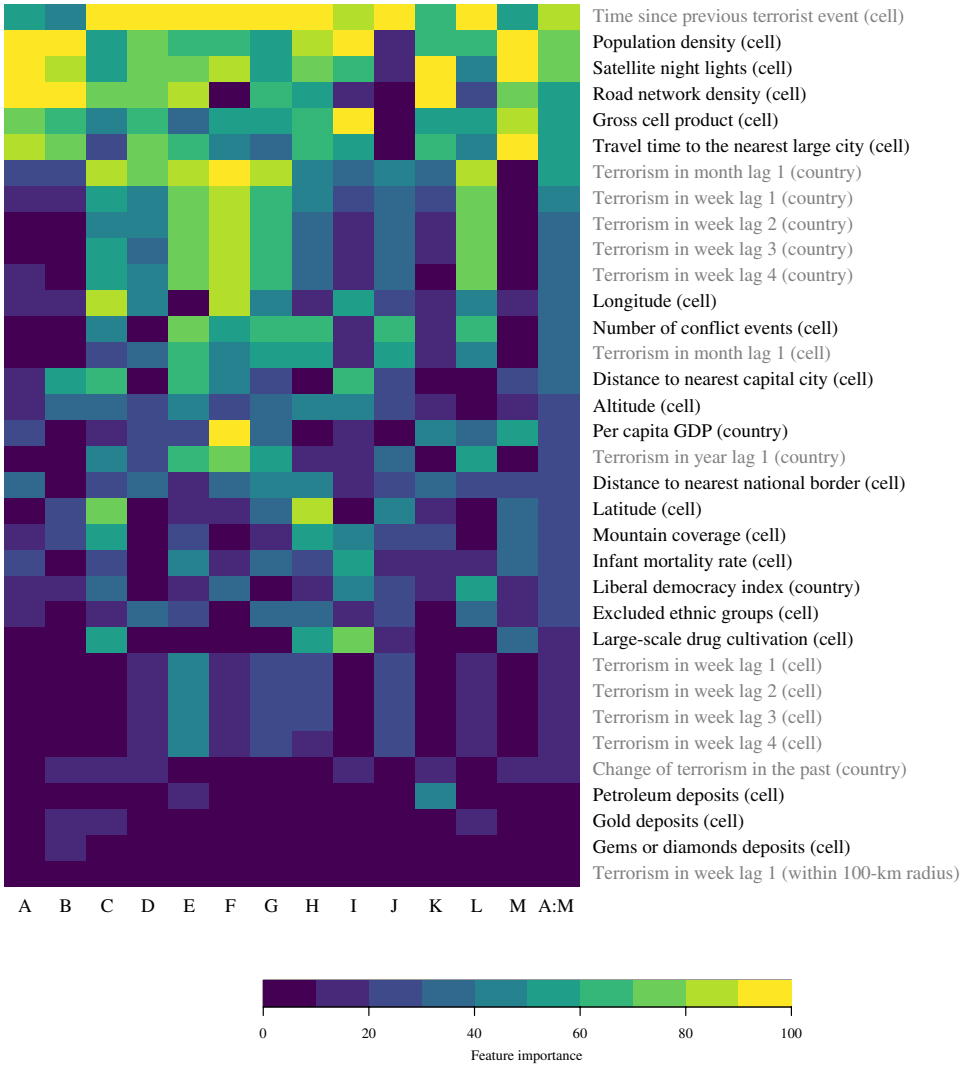

Fig. 2. Assessment of the role of features (variable importance) estimated in the 13 regions. The plot shows the importance of each structural (label in black) and procedural (label in gray) feature (total of 34 features) in each region (columns A:M) and the average across all regions (column A:M): North America (**A**), Central America and Caribbean (**B**), South America (**C**), Europe (EU28 and Schengen area) (**D**), Middle East and North Africa (MENA) (**E**), West Africa (**F**), sub-Saharan Africa (**G**), Russia and Eastern Europe (**H**), Central Asia (**I**), South Asia (**J**), East Asia (**K**), Southeast Asia (**L**), and Australasia and Oceania (**M**). The variable importance is computed from the model with the best overall predictive performance: XGB. Country effects (dummy) are not displayed.

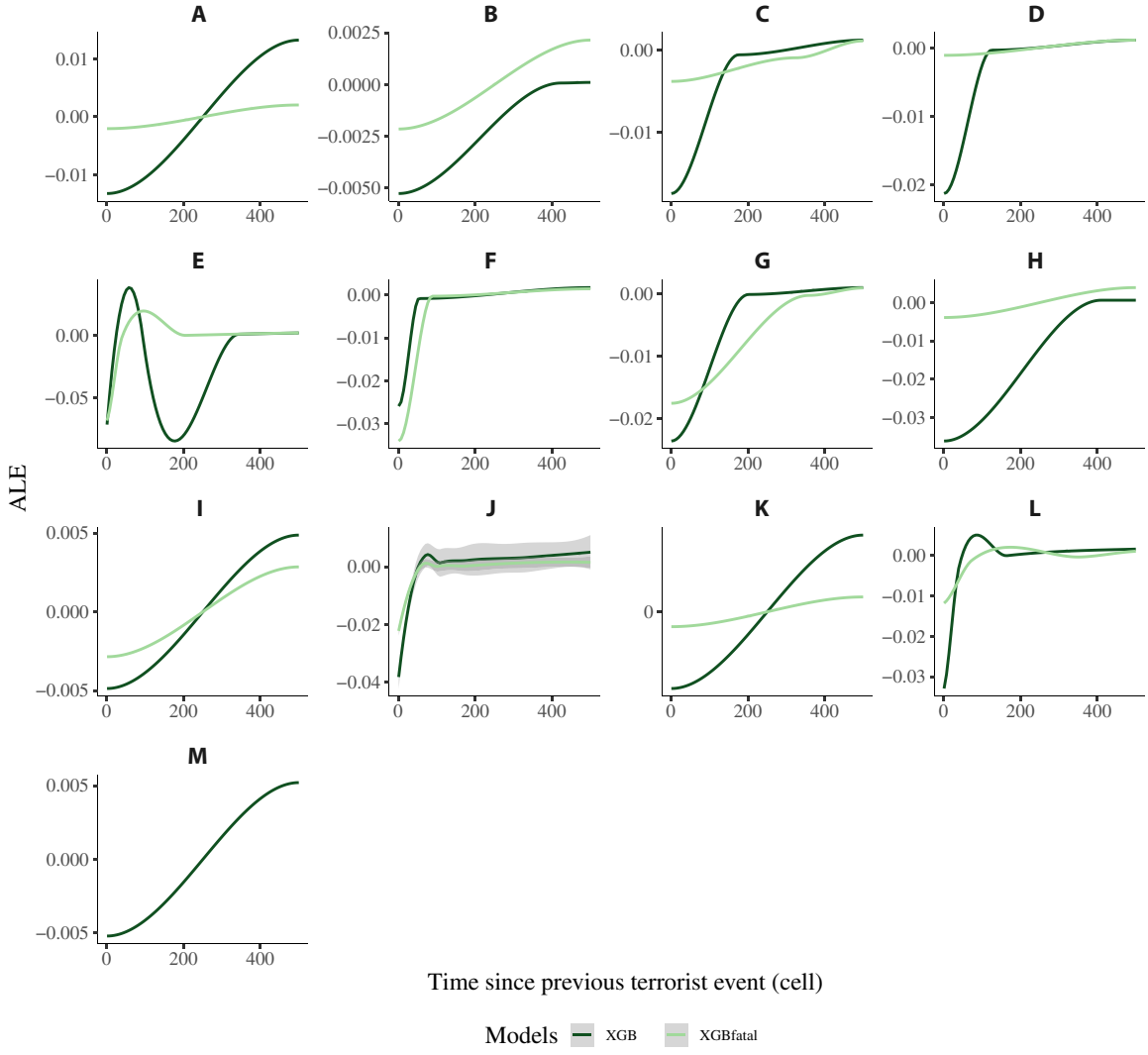

Fig. 3. ALE plot assessing the effect of time since previous terrorist event on the predicted probability of terrorism across all regions. The centered ALE plot [smooth function (loess) with span = 0.8] estimates the ALE on the basis of the results of the XGB model (2012 to 2016) with all events (XGB) and lethal events only (XGBfatal) in 13 regions worldwide: North America (**A**), Central America and Caribbean (**B**), South America (**C**), Europe (EU28 and Schengen area) (**D**), Middle East and North Africa (MENA) (**E**), West Africa (**F**), sub-Saharan Africa (**G**), Russia and Eastern Europe (**H**), Central Asia (**I**), South Asia (**J**), East Asia (**K**), Southeast Asia (**L**), and Australasia and Oceania (**M**). The ALE shows the marginal difference in prediction with an incremental change in the feature. The y axis represents change in the predicted probability of terrorism occurrence. The x axis represents time since previous terrorist event (cell) in weeks. ALE values are not computed for XGBfatal in region (M) (no variation in the response in one or more training datasets). Gray areas are 95% confidence intervals.

than nonlethal events to reporting biases (*43*). To assess the robustness to reporting bias associated with the lethality of the attacks, we compare the main results that use XGB applied to both nonlethal and lethal attacks (XGB) with the results using XGB applied to lethal attacks only (XGBfatal). Details on the sensitivity analysis are provided in the Supplementary Materials.

The four variables that show the highest importance level across all regions (overall average) include a procedural feature (at cell level) [time since previous terrorist event (Fig. 3)] and three structural variables (at cell level) [population density (Fig. 4), satellite night lights (Fig. 5), and road network density (Fig. 6)]. To further analyze their effects on terrorism, we used the accumulated local effect (ALE) plots, which are more robust to the presence of collinearity in the predictors and faster to compute than the partial dependence (PD) plots (*44*). For a given feature, the value of the ALE on the y axis represents the main feature effect at a specific value on the x axis on the predicted occurrence of terrorism compared to the overall prediction. For example, a point (x, y) in an ALE plot with an ALE value $y = -0.25$ and feature satellite night lights with a value $x = 3$ indicates that, for a luminosity value equal to 3, the predicted probability of terrorism decreases by 25% compared to the average prediction of the occurrence of terrorism.

As we indicated above, many variables are associated with more than one theoretical concept. Thus, it is very difficult to interpret the findings with regard to theory, and we will be cautious in our discussion. Nevertheless, the results from the ALE plots provide an interesting array of information we want to at least tap into. Furthermore, the ALE plots are more reliable in regions where the models show a good predictive performance. For the sake of transparency, we show the ALE plots for all regions. However, we do not interpret

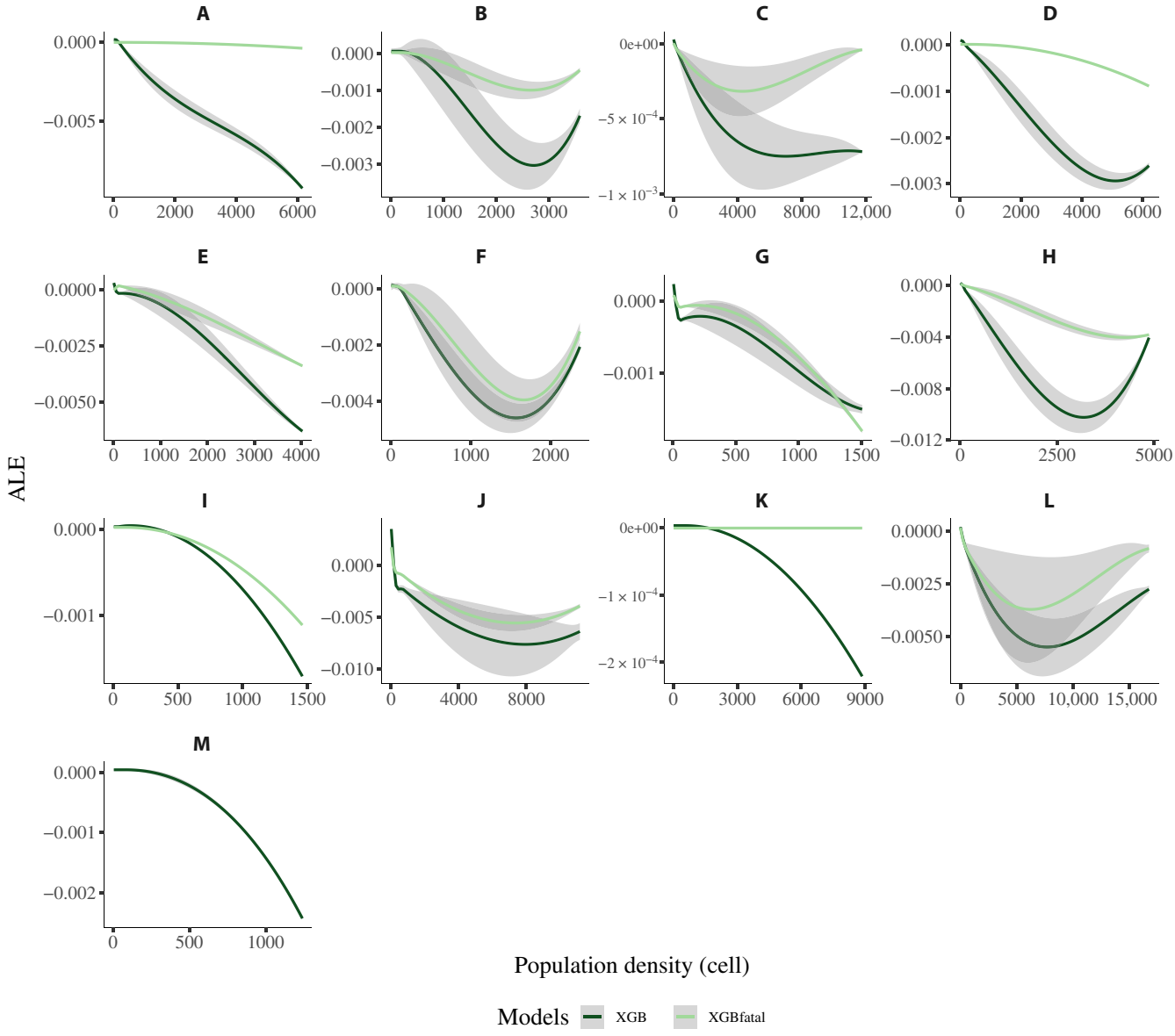

Fig. 4. ALE plot assessing the effect of population density on the predicted probability of terrorism across all regions. The centered ALE plot [smooth function (loess) with span = 0.8] estimates the ALE on the basis of the results of the XGB model (2012 to 2016) with all events (XGB) and lethal events only (XGBfatal) in 13 regions worldwide: North America (**A**), Central America and Caribbean (**B**), South America (**C**), Europe (EU28 and Schengen area) (**D**), Middle East and North Africa (MENA) (**E**), West Africa (**F**), sub-Saharan Africa (**G**), Russia and Eastern Europe (**H**), Central Asia (**I**), South Asia (**J**), East Asia (**K**), Southeast Asia (**L**), and Australasia and Oceania (**M**). The ALE shows the marginal difference in prediction with an incremental change in the feature. The y axis represents change in the predicted probability of terrorism occurrence. The x axis represents population density (cell) in inhabitants per square kilometer. ALE values are not computed for XGBfatal in region (M) (no variation in the response in one or more training datasets). Gray areas are 95% confidence intervals.

the ALE plots obtained in regions (A, B, I, K, and M), which have encountered a very low level of terrorism prevalence. The predictive performance of the models is low in these regions; therefore, the ALE plots are less reliable. Also, some features in these regions may show constant ALE values (see e.g., Fig. 4K, XGBfatal), which should not be interpreted as evidence of the absence of an effect; rather, it is the result of a very low level of prevalence of lethal attacks in the region, which leads to insufficient variability of the feature values for the model to identify a potential signal in the data. Therefore, the interpretation of the ALE plots (below) is exclusively based on the results from regions C to H and J and L.

To begin with, the results associated with time since previous terrorist events (Fig. 3) suggest that, overall, the risk of the occurrence of a terrorist event increases as time between terrorist events increases. In most regions, the ALE plots suggest a sigmoid relationship with caps that vary among regions. In regions strongly affected by terrorism (e.g., Southeast Asia; Fig. 3L), the risk peaks at about 200 weeks while the risk continues to increase until about 400 weeks in regions with a lower terrorism prevalence (e.g., Russia and eastern Europe; Fig. 3H). This finding highlights that terrorists do indeed tend to target the same locations multiple times. The reasons for this can be manifold, but in general, it seems that these locations

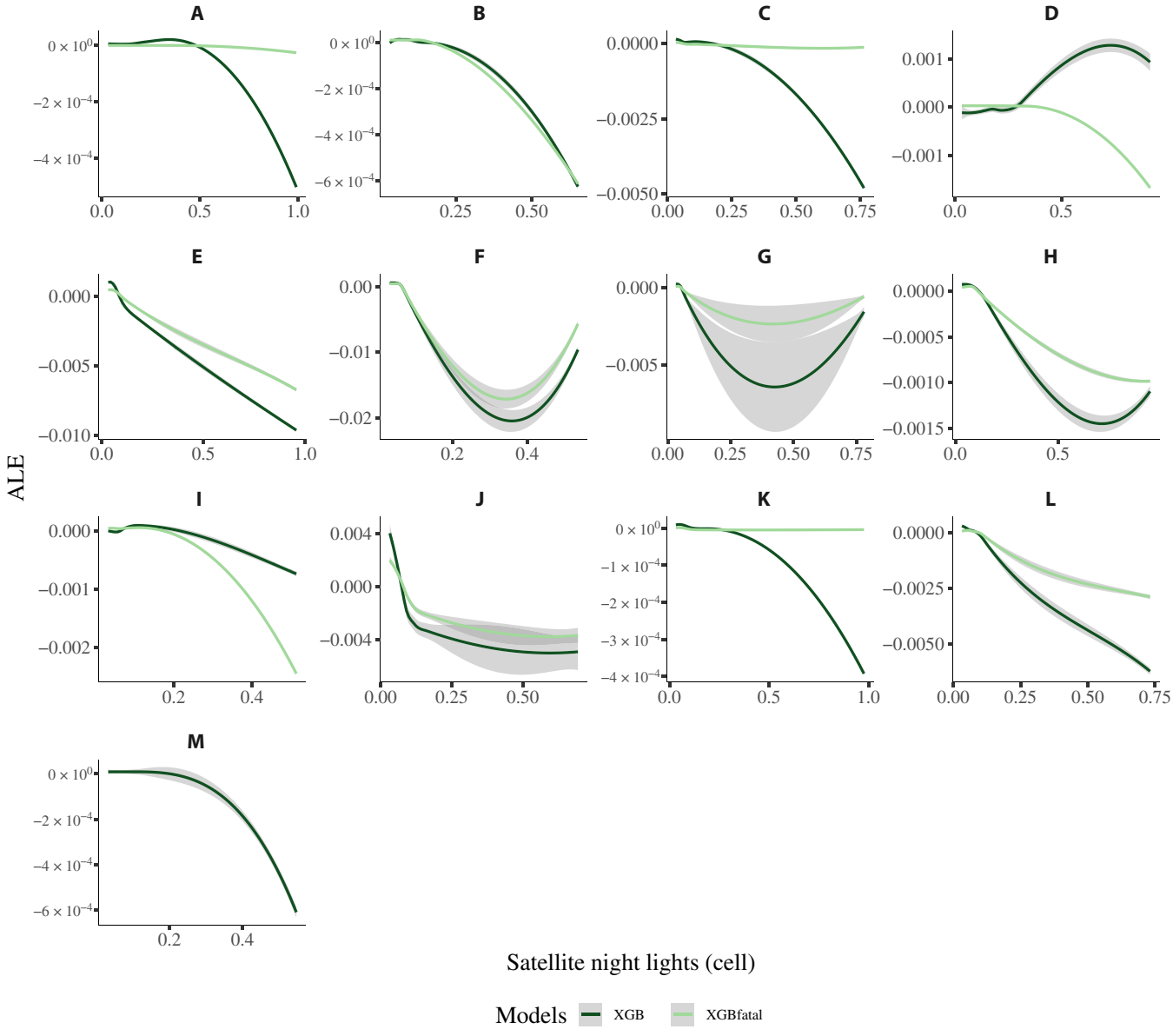

Fig. 5. ALE plot assessing the effect of satellite night lights on the predicted probability of terrorism across all regions. The centered ALE plot [smooth (loess) function with span = 0.8] estimates the ALE on the basis of the results of the XGB model (2012 to 2016) with all events (XGB) and lethal events only (XGBfatal) in 13 regions worldwide: North America (**A**), Central America and Caribbean (**B**), South America (**C**), Europe (EU28 and Schengen area) (**D**), Middle East and North Africa (MENA) (**E**), West Africa (**F**), sub-Saharan Africa (**G**), Russia and Eastern Europe (**H**), Central Asia (**I**), South Asia (**J**), East Asia (**K**), Southeast Asia (**L**), and Australasia and Oceania (**M**). The ALE shows the marginal difference in prediction with an incremental change in the feature. The y axis represents change in the predicted probability of terrorism occurrence. The x axis represents the normalized values of the calibrated satellite night lights (cell) (no unit). Gray areas are 95% confidence intervals. ALE values are not computed for XGBfatal in region (M) (no variation in the response in one or more training datasets).

reflect the strategic equilibrium between reachability (i.e., cost for attacking for the terrorists) and the strategic value (i.e., benefits of an attack) of a location.

In addition, in most regions, the predicted probability of terrorism tends to be negatively associated with population density (Fig. 4), which suggests that the probability of a terrorist attack is higher in less populated areas. Since the urban centers can be seen as more attractive targets, the results may indicate that access to these areas is an important constraint faced by terrorists. This interpretation may be further strengthened by one additional observation. Most regions (C, D, F, H, J, and L) show a U-shaped ALE, with high ALE values associated with very low and very high values of population density. In the alternative model that uses only fatal terrorist attacks (XGBfatal), the predicted probability seems to not fall by as much as with all terrorist attacks included (XGB). However, there is a possibility that this is only due to the fact that lethal attacks are more likely in densely populated areas.

The results associated with satellite night lights (Fig. 5) and road density network (Fig. 6) in most regions seem to be reflecting the same ambivalence as the results of the population density variable.

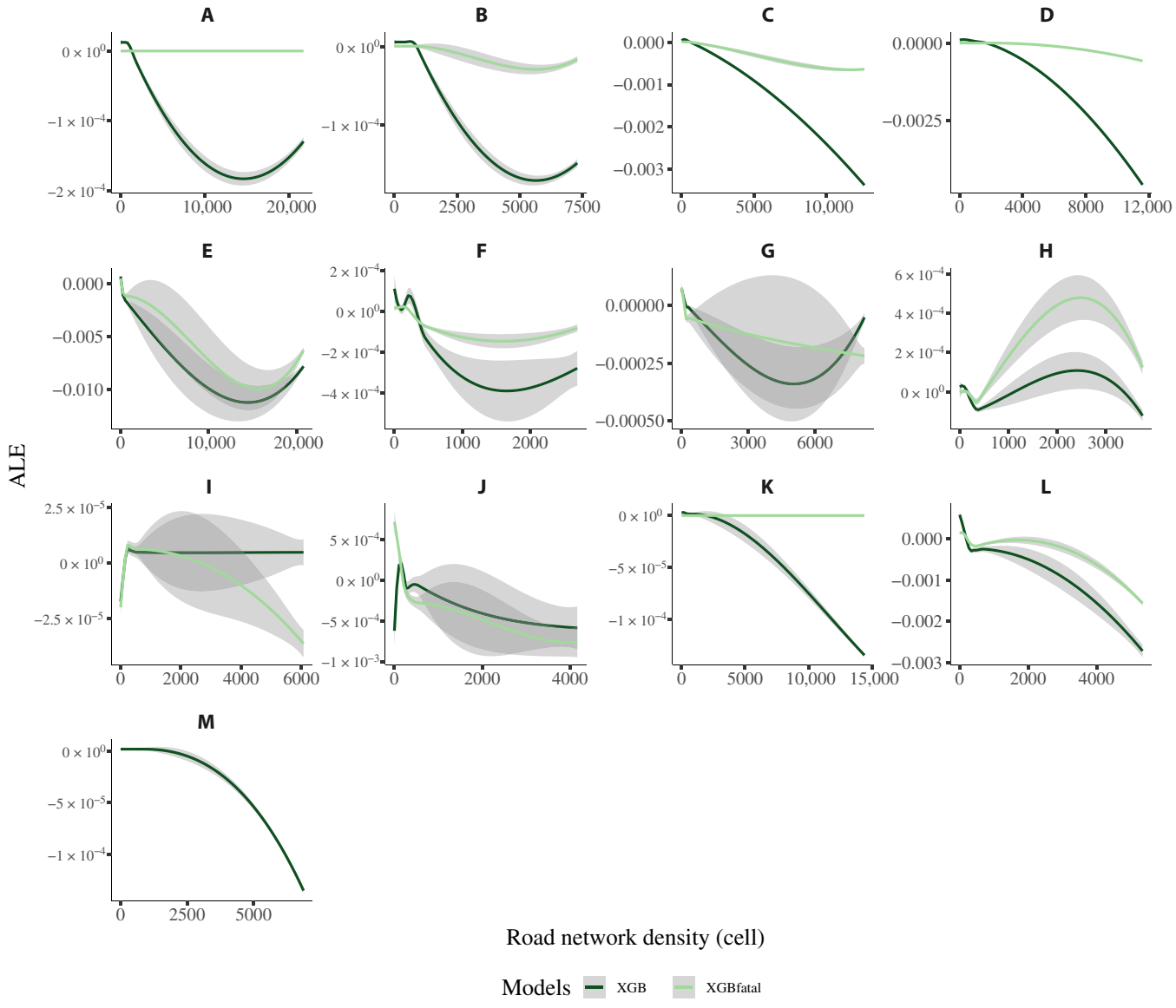

Fig. 6. ALE plot assessing the effect of road network density on the predicted probability of terrorism across all regions. The centered ALE plot [smooth (loess) function with span = 0.8] estimates the ALE on the basis of the results of the XGB model (2012 to 2016) with all events (XGB) and lethal events only (XGBfatal) in 13 regions worldwide: North America (**A**), Central America and Caribbean (**B**), South America (**C**), Europe (EU28 and Schengen area) (**D**), Middle East and North Africa (MENA) (**E**), West Africa (**F**), sub-Saharan Africa (**G**), Russia and Eastern Europe (**H**), Central Asia (**I**), South Asia (**J**), East Asia (**K**), Southeast Asia (**L**), and Australasia and Oceania (**M**). The ALE shows the marginal difference in prediction with an incremental change in the feature. The y axis represents change in the predicted probability of terrorism occurrence. The x axis represents road network density (cell) in meters per square kilometer. Gray areas are 95% confidence intervals. ALE values are not computed for XGBfatal in region (M) (no variation in the response in one or more training datasets).

For instance, particularly for the regions of West Africa and sub-Saharan (regions F and G), the ALE values align with a U-shape function, which suggests that areas with high and low human activities are more at risk than areas with a medium range of human activity. However, two ALE plots seem to stand out. For region D, Europe, satellite night lights are positively associated with the predicted probability of terrorist attacks (see XGB, both lethal and nonlethal events are considered), and for region H, Russia and eastern Europe, road networks appears as an inverted U-shape relationship. While an interpretation of the mechanisms associated with these idiosyncrasies would go beyond the scope of this paper, they highlight the fact that the effects of some predictors may be shaped by the context of each region.

The ALE plots of the remaining features are shown in the Supplementary Materials (figs. S9 to S38). However, it is worthy to mention at least two exemplary aspects to show what kind of information these plots contain. First, we can take a look at the country-level indicator of liberal democracy. This variable does not appear as an important feature across regions (Fig. 2). However, a closer look at the ALE plots (fig. S15) reveals strong regional differences. Both the

direction and the magnitude of the effects of the liberal democracy variable on the predicted probability of terrorism may considerably vary among regions. This is an interesting, if not important, finding to further probe the not yet understood interplay between democracy and terrorism. Second, we want to highlight the variable on conflict events. This variable ranges in the upper third in the cross-region comparison of variable importance (Fig. 2). The ALE plots (fig. S36) show that the predicted probability of terrorism decreases with increasing numbers of conflict events for all but one region. In region J, the ALE plots suggest a U-shape relationship. This is evidence that terrorism (understood as violence against civilians) does not take place in the same areas as armed conflict events. Terrorism may be more likely in areas where conflict is relatively low (*45*).

DISCUSSION

Our work demonstrates that models informed by theory can produce interpretable and accurate predictions of terrorism at fine spatiotemporal scales, particularly in areas already experiencing relatively high levels of terrorism activity. Predictions at fine spatial and temporal resolutions are especially informative for policy makers and can help them in designing and implementing measures to curb terrorism. However, several caveats are warranted. The tradeoff between the sensitivity and the specificity of predictive models has important consequences on policy making. It is reasonable to believe that policy-relevant predictive models should minimize the number of false negatives because the potential damages caused by terrorist events may exceed the costs to prevent the attacks. However, enforcing a low false-negative rate may lead to a considerable increase in the quantity of resources that need to be attributed to prevent terrorist events. In all investigated regions and time periods, the data are strongly imbalanced with about 99.8% of week cells not experiencing terrorism at all (table S4). In this case study, a slight reduction in the proportion of false negatives would ineluctably lead to an important increase in the number of false positives. Therefore, it lies with the analyst to adjust this threshold to reflect the specific policy-making choices, which may vary according to the available resources and counterterrorism strategy used for each country and region.

Furthermore, a large deployment of resources guided by accurate predictions may not suffice to efficiently prevent terrorist attacks. Terrorist groups could relocate their attacks in more vulnerable areas in adaptation to spatially targeted counterterrorism measures. However, if prediction identifies the most cost-effective locations for terrorists to attack, then policing those areas will at least increase the cost for terrorists and might lead to a reduction of terrorist activity overall.

Progress in terrorism research remains constrained by important factors, such as the absence of consensus on the definition of terrorism, data inaccuracy, and reporting bias (reinforced in autocratic regimes), along with incomplete theory (*46*–*49*). Therefore, obtaining accurate long-term predictions of terrorism in various contexts and at fine spatial scales might remain very difficult. While we assess the robustness of our findings to an alternative model that considers lethal events only, the terrorism metric used in this work—a binary variable associated with the occurrence of at least one terrorist attack in a week cell—does not account for the magnitude of the attacks, such as the number of causalities per week cell or the number of attacks per week cell. Alternative specifications of the response might lead to different results (*15*). Furthermore, our work considers a PRIO-GRID week spatiotemporal level of resolution. Thus, research using data aggregated at different spatial and/or temporal levels of aggregation might contrast with our findings.

Terrorism is driven by complex factors unknown to the observer, such as changes in the strategy and tactics, competition within or among terrorist groups, and classified counterterrorism operations. Nevertheless, we brought evidence that short-term predictions in restricted areas can be of sufficient spatial and temporal accuracy to benefit both the research and policy communities as tactics, strategies, and the presence of terrorist organizations can be expected to remain stable in this time frame (*12*). This way, our models allow us to identify local drivers of terrorism at fine spatial and temporal scales, which can serve as a decision-making tool for improving the design and assessment of counterterrorism policies, and ultimately contribute to reducing the number of victims because of terrorism.

We believe that interdisciplinary collaboration in the field of terrorism, data science, statistics, and computer science, along with an increase in quantity and quality of geolocated data on terrorism and its drivers, will improve the predictive performance and interpretability of advanced statistical and algorithmic models and provide practitioners active on the ground with increasingly efficient counterterrorism tools. Overall, we demonstrate that terrorist attacks can be predicted a week ahead within fine spatial grid cells with a relatively high accuracy in areas that are affected by a relatively high level of terrorism. We also show that predictive models that include spatial and temporal dependencies along with structural and procedural features tend to exhibit a better predictive performance of terrorist events at PRIO-GRID week level than autoregressive models that use procedural features only. We hope that our work may open the way to the development of predictive machine learning approaches to better understand terrorism and prevent events that take the lives of thousands of people every year.

MATERIALS AND METHODS

Data

We used the GTD dataset to gather spatial coordinates (longitude and latitude) and temporal information (day, month, and year) of terrorist attacks perpetrated by non-state actors outside legitimate warfare that occurred from 2002 to 2016 in the world (*50*). GTD is currently the most extensive and comprehensible disaggregated database on terrorism worldwide. Our working definition of terrorism considers acts as terrorism if they are perpetrated by a non-state actor, outside of legitimate warfare (i.e., targeting noncombatants), politically motivated, and with the intention to communicate to a larger audience. In GTD, terrorist events are defined as "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation." We include terrorist events on the basis of six attributes (i to vi): (i) "The incident must be intentional," (ii) "the incident must entail some level of violence or immediate threat of violence, and (iii) "the perpetrators of the incidents must be subnational actors" (state terrorism is excluded) (*20*). To be in line with our definition, we enforce criterion three (crit3) to be true so that (iv) only terrorist events that occurred outside the context of legitimate warfare activities are considered. In addition, we consider events that satisfy at least one of the following criteria: (v) "The act must be aimed at attaining a political, economic, religious, or social goal;" and (vi) "there must be evidence of an intention to coerce, intimidate,

or convey some other message to a larger audience (or audiences) than the immediate victims" (*20*).

In line with previous research work, we included 20 structural (time-invariant and annual variables) and 14 procedural (time-dependent variables (table S1) for which there are theoretical grounds to believe that they are associated with previous occurrence of terrorism (*15*, *32*). Note, however, that the variables discussed here often have relevance for more than one theoretical consideration, which limits our ability to interpret the findings, and we are cautious in our assessment.

With regard to the variables selected, we grouped them into six categories. Since terrorism is mainly a communication strategy that either seeks to coerce or to provoke (*22*), we first identified structural variables (time-invariant variables) that heighten the demonstration effects of violence, either coercive or provocative. We considered population density derived from the 2000 Gridded Population of the World (GPW v4.10) (*51*) per percentage of land area (*52*) along with travel time to cities of more than 50,000 inhabitants (*53*), distance to capital data (*52*) [from 2002 to 2014 and extrapolated the missing years (2015–2016)], and density of road at grid-cell level from the Global Roads Inventory Project dataset (*54*) to capture spatial variations in the ease of sending a message to the audience of the attack.

Second, we included variables that indicate a geographical advantage for terrorists. These areas are not necessarily of a particular interest for terrorists but easy to reach without government interference. We included closeness to international borders (*52*) [from 2002 to 2014 and extrapolate the missing years (2015–2016)], mountainous coverage from (*55*), and altitude from the Shuttle Radar Topography Mission (SRTM) digital topographic data (SRTM version 2) (*56*). Terrorist groups may use borders as safe havens to exchange information, money, and weapons and can benefit from anonymity. Along the border between countries that lack resources to control their borders, terrorist groups can take the opportunity to plan and perpetrate attacks (*57*). Some terrorist groups get tactical advantages over governmental forces in mountainous areas. Security forces are often disadvantaged in rough terrain areas because of their lack of knowledge of the local environment (*15*). In addition, we included a variable on the locations of drug cultivation as drugs usually represent potential financial resources for terrorist groups and thus indicate their presence. We considered it as a time-invariant variable (2000) because the data are temporally limited until 2002 (*58*).

Third, we included variables of relevance for coercion strategies. Coercion as a terrorist strategy seeks to impose high costs on the target audience to pose a credible threat of even higher costs in the future. Therefore, terrorists will target those areas where they can impose great costs. We extracted (i) the calibrated National Oceanic and Atmospheric Administration satellite lights at night (version 4 DMSP-OLS, from 2000 to 2012 and extrapolated the data until 2016) (*59*), a widely used proxy for human economic activity (*60*, *61*), and (ii) the gross cell product (USD) using the G-Econ dataset v4.0 (*62*). The data are provided at 5-year interval from 1990 to 2005; therefore, we used data from 2000 as a time-invariant variable. Missing data in the territorial border between Libya and Chad have been spatially interpolated. We extracted the (iii) per capita GDP in current USD at country-year level (indicator NY.GDP.PCAP.CD) from the World Bank (*63*), (iv) data on human well-being (infant mortality rate in 2000) from (*64*, *65*), and (v) data on some key resources that can be a considerable income for governments. To account for the presence of valuable resources, we gathered dummy variables on the presence of onshore petroleum deposits (2000) (*66*), primary and secondary diamond deposits (before 2002) (*67*), gem deposits (2000) (*68*), and gold (placer, surface, and vein) deposits (*69*, *70*). We used the gem and petroleum data from 2000 as time-invariant variables since they are available until 2004 and 2003, respectively. Furthermore, we merged the gem and diamonds datasets into one variable (gem) given important spatial sparsity of positive cases across various regions worldwide. We gathered gold deposit data from 2002 to 2012 and extrapolated values until 2016.

Fourth, we included variables of relevance for provocation strategies. Provocation strategies have the intention to provoke violent reactions from either the government or the local population. This means that terrorists will target particularly those locations that have a tendency to escalate into violence. For instance, locations with previous conflicts and ethnic animosities have been identified in the literature. We thus included a variable on previous armed conflict. We extracted conflict data from The Uppsala Conflict Data Program (UCDP) Georeferenced Event Dataset (version 20.1) (*71*, *72*) and computed a temporally lagged sum of conflict (state violence only) for each week cell. Furthermore, we included the number of politically excluded ethnic groups from the GeoEPR/EPR dataset (*73*, *74*), since violence is more likely to break out in regions where ethnic groups are marginalized (*33*–*35*), we spatially interpolated missing values and extrapolated values from 2014 to 2016 (data available until 2013).

Fifth, we expect some country-level features to influence the occurrence of terrorism. Particularly, the political system of a country is important. While the exact nature of the relationship between democracy and terrorism is still heavily debated, it is clear that the political system has some role to play (*36*, *75*). Therefore, we include a proxy for estimating liberal democracy. We gathered the liberal democracy index variable from the V-DEM dataset (11.1) (*76*, *77*). To account for additional unobserved characteristics of the different countries, we included a national-level dummy variable. Furthermore, to avoid potential information leakage, we lagged (year-1) all structural variables with a yearly temporal resolution.

Last, the repeated occurrence of terrorism is highly dependent on terrorist organizations gaining a local foothold. Once terrorist organizations establish a presence in a given area (either in a country or in a specific location), we expect that attacks will be highly autocorrelated. Therefore, we built procedural variables associated with previous attacks using various temporal and spatial lags. These variables are particularly important in areas of high terrorist activity as they pick up information about the terrorist group's internal calculation. Previous attacks reveal where terrorists are able and willing to strike and can thus indicate where future attacks may occur. Table S1 provides a list of all procedural and structural features.

Model evaluation

We assess the predictive performance of the investigated models based on several relevant predictive performance metrics and using visualization tools. First, we estimate the PR curves (fig. S2) and compute the AUPRCs to assess the accuracy of the prediction of positive cases (tables S2 and S3). The PR curves evaluate the proportion of true positives among positive predictions based on two evaluation measures: recall and precision. Recall is the number of correct positive predictions divided by the total number of positives. Precision is a performance measure of positive predictions defined as the number of correct positive predictions divided by the total number of positive predictions (*38*).

While the baseline (from a random classifier) for a ROC curve is represented by the identity function ($y = x$) (dotted line, fig. S3), the baseline for a PR curve (dotted line, fig. S2) is determined by the prevalence of positive events, i.e., the number of positive cases over the total number of positive and negative cases. The baseline function associated with a PR curve in a perfectly balanced class distribution would be the constant function $y = 0.5$. For an imbalanced class distribution with, e.g., a ratio of positives over negatives of 1 to 9, the baseline function would be determined as $y = 0.1$. While the AUROC of a random classifier is equal to 0.5 (area below $y = x$), the AUPRC of a random classifier is equal to the y-position of the PRC baseline, which is determined by the prevalence of positive events in the data (*38*). Given the large class imbalance in the data observed in all regions, the baseline function is very close to zero and therefore indistinguishable from the x axis in the PR curves plots (fig. S2).

The analysis of the ROC curves and the AUROCs remain the most widely used predictive metrics used in conflict forecast models (*19*). Complementary to the PRC, we show the ROC curves (fig. S3) and report the AUROC values (tables S2 and S3) for each model and region. The ROC curves are constructed by taking the predicted probabilities that there is at least one terrorist attack and calculating for a large number of threshold values: (i) the sensitivity (or recall) and (ii) specificity (number of correct negative predictions divided by the total number of negatives). The AUROC summarizes the model's performance over a range of thresholds.

We further compare the predictive of the models using complementary predictive performance metrics computed on the test set. We compute (i) the F score (F_1) (*78*, *79*), (ii) the Matthews correlation coefficient (MCC) (*80*), and (iii) the BSS. Tables S2 and S3 report the score values (in percent for a better legibility) for each model and region. Higher score values indicate better performance. In contrast to the PR and ROC curves, F_1 and MCC scores are derived from elements of the confusion matrix estimated given a threshold to classify predictive probabilities into a binary output. To account for the important discrepancies in the terrorism prevalence (prev) among regions (from 0.005% positive cell weeks in region K to 1.04% in region J), we compute the F_1 and MCC scores using false-negative thresholds adapted for each region, with the following values: 20% in regions (E, J, and L) with medium-high terrorist activity (prev $\geq$ 0.2%), 50% in regions (C, D, and F to H) with low terrorist activity (0.03% < prev < 0.2%), and 75% in regions (A, B, I, K, and M) with very low terrorist activity (prev $\leq$ 0.03%).

F_1 provides a measure of the accuracy of the models. It is computed as the harmonic mean of precision and recall, defined as two times the product of the precision and recall divided by the sum of the precision and recall. It takes values between 0 (worst value) and 1 (best value). The score reported by F_1 is affected by data imbalance, which is observed in our case study, with a number of week cells without terrorist attacks (true negatives, TNs) much larger than those with attacks (true positives, TPs). Since F_1 does not account for the number of TP, it provides an overoptimistic estimation of the predictive ability of the models on week cells that did not encounter attacks.

Complementary to F_1, we report the MCC, which evaluates the correlation between the actual outcome and the predictions. The metric is computed as follows

$$\mathrm{MCC} = (\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}) / \sqrt{(\mathrm{TP} + \mathrm{FP})(\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})},$$

with false positives (FPs), and false negatives (FNs). MCC takes values between −100% (total disagreement between the predictions and true values) and 100% (perfect prediction). Values close to zero indicate that the model does not perform better than a random guess. An MCC score is high (close to 100%) only if the model shows high performance in predicting both the positive and the negative observations, which provides a suitable alternative metric to reflect the overall predictive performance of models in the context of data imbalance (*81*).

While F_1 and MCC rely on a threshold associated with the confusion matrix, BSS is computed directly from the predicted probabilities generated by the models and computed as follows (*39*, *82*):

$$\mathrm{BSS} = 1 - \mathrm{BS}/\mathrm{BS}_{\mathrm{ref}}, \text{ where } \mathrm{BS}_{\mathrm{ref}} \text{ is defined as } \mathrm{BS}_{\mathrm{ref}} = 1/n \sum_{i=1}^{n} (\bar{o} - o_i)^2,$$

with actual outcome o_i and the average actual $\bar{o}$ outcome, and represents predictions from a baseline model without features (i.e., it predicts the prevalence of terrorism in the training data). Here, perfect predictions have a BSS of 100%, while predictions no better than the reference have a BSS of 0%. If BSS is negative, then this indicates that the investigated model performs worse than the baseline reference model, with $\mathrm{BS}_{\mathrm{ref}} > \mathrm{BS}$. Most models perform poorly in regions with very low terrorist activity (A, B, I, K, and M) and exhibit a relatively low performance in regions with low terrorism activity (C, D, and F to H). However, XGB and GAM show the best BSS scores in regions with higher terrorism prevalence (E, J, and L).

While XGB, RF, and GAM show high AUROC values in most regions, XGB and RF shows a better overall performance associated with the detection of positive cases. We notice that XGB shows a slight advantage over RF by ranking among the best scores in a larger number of regions based on the following metrics (and number of regions with best score in parentheses): (AUPRC, 10 of 13; $F1$, 8 of 13; MCC, 9 of 13; BSS, 12 of 13). In all regions, the F_1 and MCC scores of XGB, RF, and GAM systematically rank among the top. As expected, the predictive performance of XGB, RF, and GAM assessed with metrics that account for the detection of positive cases (AUPRC, F_1, MCC, and BSS) shows better results in regions with the highest levels of terrorism prevalence (regions E, G, and L).

Model interpretation

The interpretation of machine learning models is an important step in building confidence in the results of the model. In principle, we can interpret machine learning models on three different scales: (i) global level to assess the model performance, (ii) feature level to identify nonlinear relationships (and potential interactions) between the features, and (iii) data point to understand the role of specific features in the prediction of one or a few observations (*83*). Here, we focus on the feature level to assess the role of important predictors of terrorism.

We compute the ALE plots using the R package iml (*84*). While PD plots (*85*) have been widely used to illustrate the effect of features, PD plots can be misleading if some predictors are highly correlated. In contrast, the ALE method is robust to collinearity in the predictors and substantially less computationally expensive than PD plots (*44*). We select the ALE approach given the relatively large number of features (34 features) considered in each regional model, among which similarity between pairs of features (e.g., terrorism week lag 1 and terrorism week lag 2 or altitude and fraction of mountainous area) inevitably leads to some degree of correlation among features.

The objective of the ALE method is to compute for a given feature X_j (with feature values x_j) the differences in predictions where the feature values are replaced by grid values $z_{k,j}$ for a large number

of intervals k_j. Quantiles are used to define the intervals so that they contain an equal number of observations. The effect of a feature on an observation i in feature interval k_j is computed as the differences in predictions in k_j. The first step to estimate the uncentered local effects $\hat{\tilde{f}}_{j,\mathrm{ALE}}(x)$ (Eq. 1) consists of dividing the range of feature values into k intervals and compute for each interval the differences in the predictions $f(z_{k,j}, x_{\setminus j}^{(i)}) - f(z_{k-1,j}, x_{\setminus j}^{(i)})$, with the i^{th} observation of the subsets of features X_j and $X_{\setminus j}$ denoted as $x_j^{(i)}$ and $x_{\setminus j}^{(i)}$, respectively. The uncentered ALE can be expressed as follows (*44*, *86*)

$$\hat{\tilde{f}}_{j,\mathrm{ALE}}(x) = \sum_{k=1}^{k_j(x)} \frac{1}{n_j(k)} \sum_{i:x_j^{(i)} \in N_j(k)} \left[f\left(z_{k,j}, x_{\setminus j}^{(i)}\right) - f\left(z_{k-1,j}, x_{\setminus j}^{(i)}\right) \right] \quad (1)$$

The accumulated sum of the effects is computed for all observations of a feature in an interval, denoted as neighborhood $N_j(k)$. To obtain the average difference in the predictions (local effects), the sum of the differences in the predictions is divided by the number of observations $n_j(k)$ in the interval. For example, the accumulated effects of a feature X_j in the fourth interval is computed as the sum of the effects in the first, second, third, and fourth intervals. To ease interpretation, we compute the centered ALE (Eq. 2) as follows (*44*, *86*)

$$\hat{f}_{j,\mathrm{ALE}}(x) = \hat{\tilde{f}}_{j,\mathrm{ALE}}(x) - \frac{1}{n} \sum_{i=1}^{n} \hat{\tilde{f}}_{j,\mathrm{ALE}}\left(x_j^{(i)}\right) \quad (2)$$

Here, the estimated effects are centered so that the mean effect is zero. For a given feature, the centered ALE can be interpreted as the feature effect at a given value compared to the average prediction of the terrorism occurrence. For example, an ALE estimate of −0.25 associated with satellite night lights of value 3 indicates that the predicted probability of occurrence of terrorism is reduced by 25% compared with the average prediction.

As a general remark, we note that a detailed interpretation of the ALE plots would require expertise on terrorism in each investigated region and further technical work. Caution is warranted in interpreting the ALE plots. In regions where the models did not show a good predictive performance, the resulting ALE plots are less reliable. While we show all results for the sake of transparency, we discourage readers to interpret ALE plots highlighted in regions (A, B, I, K, and M), which have encountered a very low terrorism prevalence and are associated with a low predictive performance of the models.

Future research work could explore possible feature interactions through two-dimensional ALE plots and/or establish a causal framework to bring further confidence in the interpretation of the results. Future work that aims to improve the interpretability of machine learning algorithms should contribute to make these powerful models more accessible to the research community and policy makers.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/31/eabg4778/DC1

REFERENCES AND NOTES

1. M. Lim, R. Metzler, Y. Bar-Yam, Global pattern formation and ethnic/cultural violence. *Science* **317**, 1540–1544 (2007).
2. A. Zammit-Mangion, M. Dewar, V. Kadirkamanathan, G. Sanguinetti, Point process modelling of the Afghan War Diary. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 12414–12419 (2012).
3. P. Turchin, T. E. Currie, E. A. L. Turner, S. Gavrilets, War, space, and the evolution of Old World complex societies. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16384–16389 (2013).
4. N. B. Weidmann, M. D. Ward, Predicting conflict in space and time. *J. Confl. Resolut.* **54**, 883–901 (2010).
5. H. Hegre, J. Karlsen, H. M. Nygård, H. Strand, H. Urdal, Predicting armed conflict, 2010–2050. *Int. Stud. Q.* **57**, 250–270 (2013).
6. F. D. Witmer, A. M. Linke, J. O'Loughlin, A. Gettelman, A. Laing, Subnational violent conflict forecasts for sub-Saharan Africa, 2015–65, using climate-sensitive models. *J. Peace Res.* **54**, 175–192 (2017).
7. R. A. Blair, C. Blattman, A. Hartman, Predicting local violence: Evidence from a panel survey in Liberia. *J. Peace Res.* **54**, 298–312 (2017).
8. H. Hegre, M. Allansson, M. Basedau, M. Colaresi, M. Croicu, H. Fjelde, F. Hoyles, L. Hultman, S. Högbladh, R. Jansen, N. Mouhleb, S. A. Muhammad, D. Nilsson, H. M. Nygård, G. Olafsdottir, K. Petrova, D. Randahl, E. G. Rød, G. Schneider, N. von Uexkull, J. Vestby, ViEWS: A political violence early-warning system. *J. Peace Res.* **56**, 155–174 (2019).
9. H. Hegre, C. Bell, M. Colaresi, M. Croicu, F. Hoyles, R. Jansen, M. R. Leis, A. Lindqvist-McGowan, D. Randahl, E. G. Rød, P. Vesco, ViEWS$_{2020}$: Revising and evaluating the ViEWS political violence early-warning system. *J. Peace Res.* **58**, 599–611 (2021).
10. H. Hegre, H. M. Nygård, P. Landsverk, Can we predict armed conflict? How the first 9 years of published forecasts stand up to reality. *Int. Stud. Q.* (2021).
11. H. Hegre, N. W. Metternich, H. M. Nygård, J. Wucherpfennig, Introduction: Forecasting in peace research. *J. Peace Res.* **54**, 113–124 (2017).
12. L.-E. Cederman, N. B. Weidmann, Predicting armed conflict: Time to adjust our expectations? *Science* **355**, 474–476 (2017).
13. S. C. Nemeth, J. A. Mauslein, C. Stapley, The primacy of the local: Identifying terrorist hot spots using geographic information systems. *J. Polit.* **76**, 304–317 (2014).
14. M. D. Porter, G. White, Self-exciting hurdle models for terrorist activity. *Ann. Appl. Stat.* **6**, 106–124 (2012).
15. A. Python, J. Illian, C. Jones-Todd, M. Blangiardo, A Bayesian approach to modelling subnational spatial dynamics of worldwide non-state terrorism, 2010-2016. *J. R. Stat. Soc. Ser. A Stat. Soc.* **182**, 323–344 (2019).
16. F. Ding, Q. Ge, D. Jiang, J. Fu, M. Hao, Understanding the dynamics of terrorism events with multiple-discipline datasets and machine learning approach. *PLOS ONE* **12**, e0179057 (2017).
17. T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13 August 2016, pp. 785–794.
18. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
19. R. A. Blair, N. Sambanis, Forecasting civil wars: Theory and structure in an age of "Big Data" and machine learning. *J. Confl. Resolut.* **64**, 1885–1915 (2020).
20. START, Global Terrorism Database (GTD) codebook: Inclusion criteria and variables (2019); www.start.umd.edu/gtd/downloads/Codebook.pdf [accessed 1 March 2021].
21. A. F. Tollefsen, H. Strand, H. Buhaug, PRIO-GRID: A unified spatial data structure. *J Peace Res.* **49**, 363–374 (2012).
22. J. Brandsch, A. Python, Provoking ordinary people: The effects of terrorism on civilian violence. *J. Conflict Resolut.* **65**, 135–165 (2021).
23. V. P. Fortna, Do terrorists win? Rebels' use of terrorism and civil war outcomes. *Int. Organ.* **69**, 519–556 (2015).
24. A. H. Kydd, B. F. Walter, The strategies of terrorism. *Int. Secur.* **31**, 49–80 (2006).
25. A. H. Kydd, B. F. Walter, Sabotaging the peace: The politics of extremist violence. *Int. Organ.* **56**, 263–296 (2002).
26. J. Brandsch, *Killing Civilians in Civil Wars: The Rationale of Indiscriminate Violence* (Lynne Rienner, 2020).
27. R. A. Pape, *Dying to Win: The Strategic Logic of Suicide Terrorism* (Random House, 2005).
28. J. Goodwin, A theory of categorical terrorism. *Soc. Forces* **84**, 2027–2046 (2006).
29. D. A. Lake, Rational extremism: Understanding terrorism in the twenty-first century. *Dialogue IO* **1**, 15–28 (2002).
30. E. Heyman, E. Mickolus, Observations on "why violence spreads". *Int. Stud. Q.* **24**, 299–305 (1980).
31. M. Crenshaw, *The Causes of Terrorism* (St. Martin's Press Inc., 1990), pp. 113–126.
32. A. Python, J. Brandsch, A. Tskhay, Provoking local ethnic violence: A global study on ethnic polarization and terrorist targeting. *Polit. Geogr.* **58**, 77–89 (2017).
33. A. Wimmer, L.-E. Cederman, B. Min, Ethnic politics and armed conflict: A configurational analysis of a new global data set. *Am. Sociol. Rev.* **74**, 316–337 (2009).
34. L.-E. Cederman, A. Wimmer, B. Min, Why do ethnic groups rebel? New data and analysis. *World Polit.* **62**, 87–119 (2010).
35. L.-E. Cederman, K. S. Gleditsch, H. Buhaug, *Inequality, Grievances, and Civil War* (Cambridge Univ. Press, 2013).
36. E. Chenoweth, Terrorism and democracy. *Annu. Rev. Polit. Sci.* **16**, 355–378 (2013).
37. W. Enders, G. A. Hoover, T. Sandler, The changing nonlinear relationship between income and terrorism. *J. Confl. Resolut.* **60**, 195–225 (2016).
38. T. Saito, M. Rehmsmeier, The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE* **10**, e0118432 (2015).

39. G. W. Brier, Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **78**, 1–3 (1950).
40. D. B. Rubin, Bayesianly justifiable and relevant frequency calculations for the applies statistician. *Ann. Stat.* **12**, 1151–1172 (1984).
41. T. Gneiting, F. Balabdaoui, A. E. Raftery, Probabilistic forecasts, calibration and sharpness. *J. R. Stat. Soc. Series B Stat. Methodology* **69**, 243–268 (2007).
42. M. Kuhn, Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
43. N. B. Weidmann, A closer look at reporting bias in conflict event data. *Am. J. Polit. Sci.* **60**, 206–218 (2016).
44. D. W. Apley, J. Zhu, Visualizing the effects of predictor variables in black box supervised learning models. *J. R. Stat. Soc. Series B Stat. Methodology* **82**, 1059–1086 (2020).
45. S. N. Kalyvas, *The Logic of Violence in Civil War* (Cambridge Univ. Press, 2006).
46. S. Atran, R. Axelrod, R. Davis, B. Fischhoff, Challenges in researching terrorism from the field. *Science* **355**, 352–354 (2017).
47. R. English, *Terrorism: How to Respond* (Oxford Univ. Press, 2010).
48. K. Drakos, The size of under-reporting bias in recorded transnational terrorist activity. *J. R. Stat. Soc. A. Stat. Soc.* **170**, 909–921 (2007).
49. K. Drakos, A. Gofas, The devil you know but are afraid to face: Underreporting bias and its distorting effects on the study on terrorism. *J. Confl. Resolut.* **50**, 714–735 (2006).
50. START, Global terrorism database [data file]. *National Consortium for the Study of Terrorism and Responses to Terrorism* (2017).
51. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades NY, USA, Gridded Population of the World, Version 4.10 data sets (GPWv4), https://doi.org/10.7927/H4B56GPT (2017), [accessed 12 August 2018].
52. N. B. Weidmann, D. Kuse, K. S. Gleditsch, The geography of the international system: The cshapes dataset. *Int. Interact.* **36**, 86–106 (2010).
53. D. Weiss, A. Nelson, H. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, E. Poyart, S. Belchior, N. Fullman, B. Mappin, U. Dalrymple, J. Rozier, T. C. D. Lucas, R. E. Howes, L. S. Tusting, S. Y. Kang, E. Cameron, D. Bisanzio, K. E. Battle, S. Bhatt, P. W. Gething, A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
54. J. R. Meijer, M. A. Huijbregts, K. C. Schotten, A. M. Schipper, Global patterns of current and future road infrastructure. *Environ. Res. Lett.* **13**, 064006 (2018).
55. S. Blyth, I. Lysenko, L. Miles, A. Newton, *Mountain Watch: Environmental change & sustainable developmental in mountains*, no. 12 (UNEP-WCMC Biodiversity Series, 2002).
56. T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, D. Alsdorf, The shuttle radar topography mission. *Rev. Geophys.* **45**, 10.1029/2005RG000183 (2007).
57. C. C. B. Kittner, The role of safe havens in islamist terrorism. *Terror. Polit. Violenc.* **19**, 307–329 (2007).
58. H. Buhaug, P. Lujala, Accounting for scale: Measuring geography in quantitative studies of civil war. *Polit. Geogr.* **24**, 399–418 (2005).
59. National Centers for Environmental Information, Version 4 DMSP-OLS nighttime lights time series, http://ngdc.noaa.gov/eog/dmsp/downloadV4composites. html (2014). [Accessed 1 August 2018].
60. P. C. Sutton, C. D. Elvidge, T. Ghosh, Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *Int. J. Ecol. Econ. Stat.* **8**, 5–21 (2007).
61. C. D. Elvidge, J. Safran, B. Tuttle, P. Sutton, P. Cinzano, D. Pettit, J. Arvesen, C. Small, Potential for global mapping of development via a nightsat mission. *GeoJournal* **69**, 45–53 (2007).
62. W. D. Nordhaus, Geography and macroeconomics: New data and new findings. *Proc. Sci. Natl. Acad. Sci. U.S.A.* **103**, 3510 (2006).
63. World Bank, GDP per capita (current USD). DataBank. World Development Indicators (2021). [Accessed 15 April 2021].
64. A. Storeygard, D. Balk, M. Levy, G. Deane, The global distribution of infant mortality: A subnational spatial view. *Popul. Space Place* **14**, 209–229 (2008).
65. Center for International Earth Science Information Network (CIESIN) Columbia Univesity, Poverty mapping project: Global subnational infant mortality rates. *NASA Socioeconomic Data and Applications Center (SEDAC)* (2018). [Accessed 1 April 2021].
66. P. Lujala, J. Ketil Rod, N. Thieme, Fighting over oil: Introducing a new dataset. *Confl. Manag. Peace Sci.* **24**, 239–256 (2007).
67. E. Gilmore, N. P. Gleditsch, P. Lujala, J. Ketil Rod, Conflict diamonds: A new dataset. *Confl. Manag. Peace Sci.* **22**, 257–272 (2005).
68. P. Lujala, Deadly combat over natural resources. *J. Confl. Resolut.* **53**, 50–71 (2009).
69. S. Balestri, Gold and civil conflict intensity: Evidence from a spatially disaggregated analysis. *PEPS* **18**, 1–17 (2012).
70. S. Balestri, GOLDATA: The Gold Deposits Dataset Codebook, Version 1.2, in *UCSC-Cognitive Science and Communication Research Centre, WP 02/15* (2015), pp. 1–30.
71. R. Sundberg, E. Melander, Introducing the UCDP georeferenced event dataset. *J. Peace Res.* **50**, 523–532 (2013).
72. T. Pettersson, M. Öberg, Organized violence, 1989–2019. *J. Peace Res.* **57**, 597–613 (2020).
73. M. Vogt, N.-C. Bormann, S. Rüegger, L.-E. Cederman, P. Hunziker, L. Girardin, Integrating data on ethnicity, geography, and conflict: The ethnic power relations data set family. *J. Confl. Resolut.* **59**, 1327–1342 (2015).
74. J. Wucherpfennig, N. B. Weidmann, L. Girardin, L.-E. Cederman, A. Wimmer, Politically relevant ethnic groups across space and time: Introducing the GeoEPR dataset. *Confl. Manag. Peace Sci.* **28**, 423–437 (2011).
75. K. Gaibulloev, J. A. Piazza, T. Sandler, Regime types and terrorism. *Int. Organ.* **71**, 491–522 (2017).
76. M. Coppedge, J. Gerring, C. H. Knutsen, S. I. Lindberg, J. Teorell, D. Altman, M. Bernhard, M. S. Fish, A. Glynn, A. Hicken, G. Hindle, N. Ilchenko, J. Krusell, A. Luhrmann, S. F. Maerz, K. L. Marquardt, K. M. Mann, V. Mechkova, J. Medzihorsky, P. Paxton, D. Pemstein, Josefine Pernes, Johannes von Römer, B. Seim, R. Sigman, S.-E. Skaaning, J. Staton, A. Sundström, Ei-tan Tzelgov, Yi-ting Wang, T. Wig, S. Wilson, D. Ziblatt, V-Dem [Country-Year-date] Dataset V11.1 (Varieties of Democarcy Project, 2021); https://doi.org/10.23696/vdemds21.
77. D. Pemstein, K. L. Marquardt, E. Tzelgov, Y.-t. Wang, J. Krusell, F. Miri, The V-Dem measurement model: Latent variable analysis for cross-national and cross-temporal expert-coded data. *V-Dem Working Paper* **21**, 10.2139/ssrn.3167764, (2018).
78. L. R. Dice, Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302 (1945).
79. T. J. Sørensen, *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish* (Munksgaard Copenhagen, 1948), vol. 5.
80. B. W. Matthews, Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* **405**, 442–451 (1975).
81. D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **21**, 10.1186/s12864-019-6413-7, (2020).
82. A. A. Bradley, S. S. Schwartz, T. Hashino, Sampling uncertainty and confidence intervals for the Brier score and Brier skill score. *Weather Forecast.* **23**, 992–1006 (2008).
83. T. C. D. Lucas, A translucent box: Interpretable machine learning in ecology. *Ecol. Monogr.* **90**, e01422 (2020).
84. C. Molnar, B. Bischl, G. Casalicchio, iml: An R package for interpretable machine learning. *JOSS* **3**, 786 (2018).
85. J. H. Friedman, Greedy function approximation: A gradient boosting machine. *Ann. Statist.* **26**, 1189–1232 (2001).
86. C. Molnar, *Interpretable Machine Learning* (Leanpub, 2019).
87. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for statistical computing, Vienna, Austria (2017).
88. M. Kuhn, J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, Z. Mayer, Kenkel, the R Core Team, M. Benesty, R. Lescarbeau, A. Ziem, L. Scrucca, Y. Tang, Candan, T. Hunt, *caret: Classification and Regression Training* (2017). R package version 6.0–76.
89. R. L. Winkler, J. Munoz, J. L. Cervera, J. M. Bernardo, G. Blattenberger, J. B. Kadane, D. V. Lindley, A. H. Murphy, R. M. Oliver, D. Ríos-Insua, Scoring rules and the evaluation of probabilities. *Test* **5**, 1–60 (1996).

Acknowledgments: We thank B. Liu, S. Tan, and D. Svanidze for the extensive assistance on illustrations and data preparation. **Funding:** A.P. has been funded by Zhejiang University under grant numbers 2021QN81029 (fundamental research funds for the central universities) and 188170-11103 (Zhejiang University global partnership fund). A.B. has been funded by the German Federal Ministry of Education and Research (BMBF) under grant number 01IS18036A.
Author contributions: A.P., T.C.D.L., A.K.N., P.A.H., R.A., and A.B. designed the research. T.C.D.L., A.B, J.B., and A.P. performed the research and analyzed the data. A.P., T.C.D.L., J.B., A.K.N., P.A.H., R.A., and A.B. wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 7 January 2021
Accepted 15 June 2021
Published 30 July 2021
10.1126/sciadv.abg4778

Citation: A. Python, A. Bender, A. K. Nandi, P. A. Hancock, R. Arambepola, J. Brandsch, T. C. D. Lucas, Predicting non-state terrorism worldwide. *Sci. Adv.* **7**, eabg4778 (2021).