

# Modeling Spatiotemporal Armed Conflict Dynamics via Physics-Informed Diffusion

Anonymous Author(s)

## Abstract

Accurate early warning of armed violence requires models that capture *how* conflict spreads, not just *where* it has happened. We propose a physics-informed diffusion partial differential equation (*diffusion-PDE*) framework that models the evolution of latent conflict drivers across a learned, spatially heterogeneous potential field. Crucially, our framework accounts not only for smooth diffusion but also for complex dynamics like collective relocation and endogenous conflict generation. We implement this framework in *ConflictNet*, an end-to-end, three-stage solver composed of a spatiotemporal encoder, a diffusion-PDE solving block, and a predictive decoder. Evaluated on the ViEWS gold-standard benchmark, ConflictNet yields a 56% relative accuracy gain over strong baselines and reveals interpretable diffusion patterns that align with social science theory. By linking a sophisticated process model with data-driven inference, our work bridges the gap between machine learning and social science theory, offering a new path to build more mechanistic and accurate forecasting systems.

## Keywords

Conflict Prediction, Spatiotemporal Data Mining, Physics-Informed Learning, Diffusion Modeling

## ACM Reference Format:

Anonymous Author(s). 2025. Modeling Spatiotemporal Armed Conflict Dynamics via Physics-Informed Diffusion. In *Proceedings of 32nd SIGKDD Conference on Knowledge Discovery and Data Mining, 2026 (SIGKDD 2026 Research Track)*. ACM, New York, NY, USA, 16 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 Introduction

Armed conflict often manifests as a recurrent and spatially diffuse phenomenon, driven by the complex interplay of structural, environmental, and strategic factors [84, 86]. Since 2000, political scientists have developed a variety of prediction frameworks to forecast the spread of armed conflict, supporting early warning and the planning of preventive actions to mitigate damage. These frameworks range from classical game-theoretic models to more recent ensemble-based approaches, such as ViEWS [34].

ViEWS has become the de facto “gold standard” in conflict prediction [32, 33]. The project established the current benchmark, popularized the cell-month forecasting paradigm, and demonstrated

the value of open, continuously updated warning systems. In this paper, we adopt the same task setting as ViEWS. Specifically, we use *PRIO-GRID*, a globally uniform  $0.5^\circ \times 0.5^\circ$  latitude–longitude lattice developed by the Peace Research Institute Oslo [78]. It is widely adopted as the standard spatial unit for geo-referenced conflict studies. In this framework, each grid cell represents one spatial unit, and each calendar month corresponds to one temporal step. Thus, our target is the probability that at least one conflict event occurs in a given *cell-month* [32].

Although existing methods perform well at coarse scales, two limitations remain. First, they rely heavily on pre-defined features. Second, they fail to model the underlying spatial-temporal dynamics of conflict propagation explicitly. Current machine learning methods predominantly treat conflict forecasting as a statistical pattern-matching task. This perspective overlooks a critical reality: conflict spread follows physical and strategic constraints rather than random patterns. Existing spatial dependence models rely excessively on simplistic adjacency or distance metrics [57], neglecting the physical dynamics of actor mobility, resource flows, and grievance transmission [17, 24]. This oversimplification limits interpretability and transferability.

We propose a *physics-informed* diffusion partial differential equation (*diffusion-PDE*) framework to model spatiotemporal armed conflict dynamics. *Diffusion-PDE* models the evolution of latent conflict drivers such as displaced civilians, redeploying armed groups, or price shocks. It derives transition probabilities for conflict factors based on a Boltzmann distribution over spatially heterogeneous *potential fields* (e.g., mountain barriers, transport corridors, or border zones), allowing the model to capture both smooth diffusion and barrier effects shaped by terrain, governance, or social fragmentation [28, 52]. The resulting transfer dynamics yield a PDE that governs the propagation of conflict over time and space. To enhance modeling robustness, we incorporate regularization terms that account for collective rigid-body movement (e.g., sudden group redeployments) and endogenous conflict generation.

Building on this framework, we introduce *ConflictNet*, an end-to-end solver for the diffusion-PDE. The framework extracts latent conflict signals, evolves them under the dynamics of the potential field through a diffusion-PDE process, and outputs the probabilities of violence for each cell-month. By modeling conflict spread as a structured physical process rather than a purely data-driven pattern, ConflictNet not only enhances accuracy but also improves interpretability. When evaluated on the PRIO-GRID benchmark maintained by ViEWS, it achieves a 56% relative gain over the “gold standard” and reveals stable diffusion patterns across regions and years.

The development of predictive models is driven by a compelling humanitarian rationale: enabling a shift from reactive to proactive interventions that save lives. Accurate, high-resolution forecasts provide the foresight needed for preventive diplomacy, optimized

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

SIGKDD 2026 Research Track, Jeju, Korea

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-XXXX-X/2018/06

<https://doi.org/XXXXXXX.XXXXXXX>

peacekeeping, and pre-positioned humanitarian aid to mitigate violence before it escalates [71, 88]. Accordingly, our framework is designed as an analytical tool for conflict prevention and management, aligning with the core mission of established early-warning systems [33].

We summarise our contributions as follows:

- Propose a **physics-informed diffusion-PDE** and instantiate it in **ConflictNet**, a PDE solver that runs end-to-end to model armed conflict diffusion dynamics.
- Achieve a **56 % relative accuracy gain** over the ViEWS benchmark on *PRIO-GRID*, outperforming all other strong baselines.
- Reveal **stable diffusion patterns** and interpretable latent flows that align with collective-mobilisation theory.
- Bridge the gap between data-driven forecasting and theory-driven process modeling, offering a new path for integrating machine learning with social science theory.

## 2 Related Works

Forecasting armed conflict sits at the intersection of political-science theory, spatiotemporal machine learning and, increasingly, physics-informed modelling. We situate our contribution against advances in each strand.

**Conflict forecasting.** Forecasting political violence has long been central to peace research, with early approaches grounded in structural explanations and game theory. A shift occurred as machine learning entered mainstream political science through the work of Ward et al. [83], bringing attention to the predictive potential of data-driven approaches. With improved spatial data and computational infrastructure, forecasting at finer spatiotemporal scales—particularly the PRIO-GRID cell-month level—has become feasible [55, 72].

Notable contributions include ViEWS [33], which introduced an ensemble forecasting framework and set the standard for country-month-level prediction. However, its performance at the PRIO-GRID cell-month level remains limited [35]. Complementary efforts include Python et al.’s [63] use of gradient-boosted trees to forecast terrorism globally with high temporal resolution, demonstrating the applicability of traditional ML models at fine spatial granularity. Despite growing accuracy, most models remain largely descriptive, relying on handcrafted predictors and lacking explicit spatial mechanisms.

**Spatiotemporal modeling.** Recent works have adopted deep learning to capture complex spatiotemporal patterns at the PRIO-GRID level [36]. Chadeaux [19] applied time-series clustering to extract recurring conflict shapes. Lindholm et al. [51] used hidden Markov models with latent cell states to model temporal shifts. Radford [65] framed conflict as a sequence prediction problem using ConvLSTM, while Brandt et al. [8] leveraged spatiotemporal graph convolutional networks to capture evolving conflict dependencies.

Broader deep learning advances have included models for time series [2, 41, 54, 89], models for spatial data [44, 80], and hybrid architectures [73, 75, 76, 91–93] to unify both. While these approaches improve predictive fidelity, they often function as black boxes and lack the inductive bias necessary to capture structured spatial dynamics, thereby limiting explanatory power.

**Physics-informed learning.** A growing body of research has sought to bridge this gap by incorporating physical modeling into neural architectures [87]. This line of work embeds domain knowledge—such as conservation laws or dynamical systems—into deep learning models to improve generalization and interpretability. Examples include knowledge-assisted architectures in materials science [64], quantum chemistry [40, 61], and environmental forecasting [95], where PDE constraints are fused into neural networks to guide learning and prediction.

In conflict modeling, however, such integration remains rare due to the abstract nature of sociopolitical dynamics and the lack of clear physical analogues. Nevertheless, earlier work has shown the promise of spatial process modeling through point processes [94]. Our work extends this frontier by demonstrating how physics-inspired formulations can encode both diffusion and relocation dynamics of conflict factors, leading to improved performance and interpretability in conflict forecasting tasks.

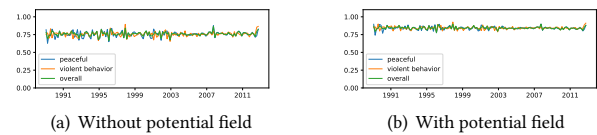
## 3 Preliminary

We frame conflict forecasting as a spatiotemporal prediction task on a regular grid, identifying empirical patterns—local mean-reversion, spatial cohesion, and inter-feature coupling—that justify our physics-informed modeling. Scalars are written in lightface, while vectors are denoted in boldface.

### 3.1 Problem Formulation

Our study domain is the  $0.5^\circ \times 0.5^\circ$  PRIO-GRID [78] lattice that blankets the African continent from 1990 to 2020. Each grid cell carries two feature streams. *Static attributes* describe enduring characteristics—topography, transport networks, night-time lights, land use, and other long-horizon factors. *Dynamic signals* record month-to-month shifts such as protest activity, population displacement, and commodity-price shocks. A complete feature table is provided in Supplementary Material Table 4.

Given all observations up to the present month, the task is to predict whether a cell will register at least one conflict event in the next 1, 3, 6, or 9 months. We score performance at this *cell-month* level, aligning with the ViEWS early-warning benchmark while retaining full spatial detail.

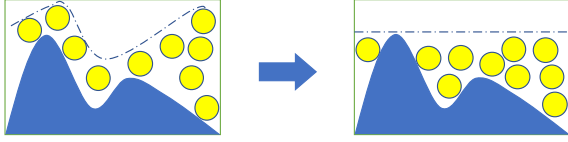


**Figure 1: Month-ahead changes in conflict intensity against the ratio between each cell and the average of its immediate neighbours. It is computed across three versions: peaceful protests, violent protests, and total protest counts. The figure indicates that cells with above-average violence tend to decline in the next month, while below-average cells rise—i.e. activity drifts back toward the local mean.**

### 3.2 Empirical Motivations from Data

Our analysis of conflict-related data [32–34] reveals patterns motivating our physics-informed approach. We summarize three core observations below.

**Local mean-reversion.** Many dynamic features show strong temporal locality. These features tend to revert toward the spatial mean of neighboring cells over time. Natural disasters [72] diffuse smoothly across flat potential fields. More volatile factors, such as forced migration and protest events [68], regress erratically, particularly on rugged terrains with spatial heterogeneity. Historical data confirm this theory (Figure 1). Most grid cells exhibit conflict intensity reverting toward the local mean. The tendency strengthens when accounting for the potential field. Figure 2 visualizes this phenomenon: feature dynamics resemble balls rolling toward equilibrium on a landscape.



**Figure 2: Illustration of feature diffusion. In the presence of a spatial potential field  $\phi(\mathbf{x})$ , diffusion proceeds until the gradient force  $\nabla u(\mathbf{x}, t)$  balances the external force  $-\nabla\phi(\mathbf{x})$ , reaching equilibrium.**

**Local Transition Scaling.** As Figure 1 shows, conflict intensity regresses to the mean within a cell’s immediate neighbourhood. The pattern largely disappears beyond second-order neighbours, indicating that adjustment happens through short, local steps rather than long-range jumps. This locality accords with socio-political theory: troop movement, resource transport, and grievance diffusion are bounded by logistics, terrain, and information frictions, creating steep cost gradients that limit spatial reach [18]. Similar short-range contagion has been documented for disaster shocks [72] and protest waves [68].

**Inter-feature coupling.** Violent conflict rarely stems from a single driver. Multiple factors—opportunity, grievance, and organizational capacity—must converge to trigger conflict [18, 21]. Conflict-relevant features show strong interdependence. Environmental shocks and mobilization-related variables correlate closely, as Figure 5 in Supplementary Material demonstrates. Such relationships demand models capable of capturing multivariate co-movement across time and space.

These empirical observations motivate the structured diffusion framework. The next chapter formalizes this framework through mathematical assumptions, leading to the derivation of an explicit expression for conflict evolution.

## 4 Methodology

Armed conflict propagates through structured latent drivers. These drivers include displacement, arms flow, and mobilization. Logistics, infrastructure, and command delays constrain the propagation [4, 12]. Unlike random diffusion, conflict spreads at finite speeds. Physical space constraints improve modeling of directional shifts and temporal lags compared to classical probabilistic approaches [9, 43].

We derive an explicit diffusion partial differential equation (diffusion-PDE) to model the evolution of conflict, incorporating the spatiotemporal dynamics of latent drivers and their constraints. The diffusion-PDE captures the complex relationships and physical constraints that govern conflict dynamics. To solve them efficiently, we integrate the equations into a deep learning framework. This approach improves both the accuracy and robustness of conflict forecasting by enabling data-driven estimation while preserving the physical realism of the underlying process.

### 4.1 Physics-Informed Diffusion-PDE

Earlier machine-learning approaches often modeled armed conflict as pure diffusion in the intensity space [83, 86]. Recent theoretical work in political science argues that macro-level conflict patterns arise not from intensity alone, but from the spatial propagation of latent drivers, such as population movements and weapons flows [26]. Motivated by this, we define **conflict factors** as a low-dimensional latent vector learned by the spatiotemporal encoder from multi-source input features. Rather than being a physical entity, it serves as a compact representation of all factors driving conflict dynamics at location  $\mathbf{x}$  and time  $t$  as  $\mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^l$  (e.g., resource availability, mobilization capacity, grievances). It is the core state variable that is propagated by the diffusion-PDE module.

Prior studies show that spatial heterogeneity strongly influences conflict flow, with movements gravitating toward regions of strategic significance [1, 13, 59]. Similar to particles in a physical field, conflict factors tend to concentrate in regions of lower potential. In statistical physics and increasingly in social systems modeling [37, 58], the Boltzmann distribution describes how agents or particles probabilistically settle in space under the influence of a spatial **potential field**  $\phi$ . Potential field is defined as a learnable spatial embedding generated by a dedicated neural network branch from static features (e.g., terrain, borders, ethnic settlement patterns). Instead of a pre-defined physical field, the potential field represents the inherent susceptibility or resistance of each location  $\mathbf{x}$  to conflict, effectively encoding how persistent factors create barriers or corridors for its propagation. We adopt this principle to characterize equilibrium distributions of conflict factors:

**Assumption 4.1 (Boltzmann Distribution [29, 37, 42, 58, 60]).** At equilibrium in a potential field  $\phi(\mathbf{x})$ , the spatial density distribution  $\pi(\mathbf{x})$  of conflict factors follows the Boltzmann distribution governed by a constant  $\beta$ :

$$\pi(\mathbf{x}) \propto e^{-\beta\phi(\mathbf{x})}. \quad (1)$$

The coefficient  $\beta$  quantifies the sensitivity of conflict factors to the potential field  $\phi(\mathbf{x})$ . Higher values of  $\beta$  imply greater concentration of conflict factors in low-potential regions—interpreted as areas of strategic or symbolic importance. Thus,  $\beta$  can be viewed as a proxy for the latent propensity toward violence, reflecting individuals’ willingness to engage in or support violent behavior—a widely recognized determinant of conflict onset and escalation [30, 53, 70].

Assumption 4.1 is reasonable because it mathematically formalizes a core principle of strategic behavior: actors probabilistically favor locations that maximize their advantage, which the Boltzmann distribution treats as low-potential-energy states. This assumption translates the rational, goal-oriented behavior of conflict actors



into a principled statistical physics framework; further discussion on the theoretical grounding of all assumptions is available in the Supplementary Material D.

Based on Assumption 4.1 and the framework of [74], we derive in Proposition 4.1 the transition probabilities of conflict factors by imposing the balance condition on the underlying Markov process.

**Proposition 4.1.** In a potential field  $\phi(\mathbf{x})$ , a probability density  $P_\phi(\mathbf{x}, \mathbf{r}, \Delta t)$  that a conflict factor moves from location  $\mathbf{x}$  to  $\mathbf{x} + \mathbf{r}$  within time  $\Delta t$  is given by:

$$P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) = \frac{1}{Z} \exp \left[ -\frac{\beta}{2} \{ \phi(\mathbf{x} + \mathbf{r}) - \phi(\mathbf{x}) \} \right] P(\|\mathbf{r}\|, \Delta t), \quad (2)$$

where  $Z$  is a normalization constant, and  $P(\cdot, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a normalized kernel function.

From Proposition 4.1, we observe that when  $\phi(\mathbf{x}) \equiv 0$ , the transition probability reduces to the normalized kernel  $P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) \propto P(\|\mathbf{r}\|, \Delta t)$ , indicating that movement is isotropic and depends solely on the transition distance  $\|\mathbf{r}\|$ , not on the initial location  $\mathbf{x}$ . Conversely, when a potential field  $\phi(\mathbf{x})$  is present, the transition probability  $P_\phi$  is modulated by an exponential term that reflects the local potential energy difference.

Empirical evidence shows that conflict diffusion is constrained by frictional limits—logistical, infrastructural, and organizational factors prevent long-range propagation over short intervals [9, 43]. As the time horizon shortens from  $\Delta t$  to  $\Delta t/2$ , typical transition distances contract accordingly, reflecting actors' limited capacity to mobilize or redeploy [3, 39, 42]. To capture this, we adopt the classical assumption from random process theory:

**Assumption 4.2 (Local Transition Scaling Assumption [47, 90]).**

$$\int_{\mathbb{R}^2} \|\mathbf{r}\|^2 P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} \propto \Delta t, \quad \int_{\mathbb{R}^2} \|\mathbf{r}\|^3 P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = o(\Delta t). \quad (3)$$

The Local Transition Scaling Assumption requires the mean squared displacement to scale linearly with time ( $\propto \Delta t$ ), ensuring normal diffusion, while demanding the third moment to be sublinear ( $o(\Delta t)$ ), guaranteeing local symmetry. These conditions maintain proper diffusivity and eliminate directional bias in infinitesimal displacements, fundamental for modeling standard stochastic processes. Many standard kernels satisfy this condition. For example, the two-dimensional Gaussian heat kernel  $P(\|\mathbf{r}\|, \Delta t) = (4\pi A \Delta t)^{-1} \exp\{-\|\mathbf{r}\|^2 / (4A \Delta t)\}$  and the uniform kernel over a shrinking ball  $P(\|\mathbf{r}\|, \Delta t) = \{\pi R^2(\Delta t)\}^{-1} \mathbf{1}\{\|\mathbf{r}\| \leq R(\Delta t)\}$ , with  $R(\Delta t) = \sqrt{2A \Delta t}$ , both satisfy Assumption 4.2. Detailed derivations are provided in Appendix A.

When  $\phi(\mathbf{x}) \equiv 0$ , symmetry implies that the expected of transition vector  $\mathbf{r}$  equals

$$2 \int_{\mathbb{R}^2} \mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = \int_{\mathbb{R}^2} \mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} + \int_{\mathbb{R}^2} -\mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = 0. \quad (4)$$

Assumption 4.2 is reasonable because it embeds real-world logistical constraints into the model, ensuring that conflict propagates through local, incremental steps rather than implausible long-range jumps. This aligns with extensive empirical evidence on the contagion patterns of insurgencies, making the resulting diffusion process more physically realistic. Furthermore, Proposition 4.2 characterizes the corresponding covariance structure.

**Proposition 4.2.** Under Assumption 4.2 and the condition of Proposition 4.1, there exists a constant  $D > 0$  such that:

$$(2Z\Delta t)^{-1} \int_{\mathbb{R}^2} \mathbf{r} \mathbf{r}^\top P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = D \mathbf{I}_2. \quad (5)$$

Proposition 4.2 implies that, in the absence of a potential field, conflict factor movement is isotropic, uncorrelated, and directionally uniform. The diffusion coefficient  $D$  governs the variability of the displacement vector  $\mathbf{r}$ : larger values of  $D$  correspond to broader dispersion and more frequent transitions. In social terms,  $D$  reflects the degree of organizational cohesion: higher  $D$  implies weaker group control, greater individual autonomy, and increased difficulty in sustaining coordinated insurgent action [52, 81].

Having analyzed the micro-level transition of individual conflict factors, we now examine the evolution of their macroscopic density  $\mathbf{u}(\mathbf{x}, t)$ , focusing on how localized interactions aggregate into large-scale patterns. Beyond individual motion, the dynamics of conflict are shaped by interdependencies among factor types; for instance, personnel movement often induces correlated shifts in the distribution of weapons [6, 16].

We model this interaction through a linear coupling mechanism:

**Assumption 4.3 (Coupling of Transitions [23, 45]).** There exists a constant matrix  $\xi$  where  $\xi_{ii} = 0$  for all  $i$ , such that a transition in  $u_j$  proportionally induces a transition in  $u_i$ , scaled by the coupling coefficient  $\xi_{ji}$ . Then:

$$\mathbf{u}(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} \mathbf{P}_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) \odot \mathbf{u}(\mathbf{x} - \mathbf{r}, t) d\mathbf{r} - \xi^\top \mathbf{J}(\mathbf{x}, t), \quad (6)$$

where  $\mathbf{J}(\mathbf{x}, t)$  represents the flux of  $\mathbf{u}$  at location  $\mathbf{x}$  time  $t$ .

Assumption 4.3 explains that the density at the next moment consists of two components: the spontaneous transition of the conflict factors, given by  $\int_{\mathbb{R}^2} \mathbf{P}_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) \odot \mathbf{u}(\mathbf{x})$ , and the induced linear coupling term  $-\xi^\top \mathbf{J}(\mathbf{x}, t)$  from other types of conflict factors. The assumption is reasonable naturally, because it reflects the interdependent nature of conflict drivers, a central tenet of modern conflict theory. This allows the framework to capture the multivariate reality of conflict escalation instead of treating each driver as an isolated phenomenon.

By combining the above Assumptions and Propositions, we derive the transfer equation governing conflict factors with particle-like behavior:

**Theorem 4.3.** Under condition of Proposition 4.1, 4.2 and Assumption 4.1, 4.2, 4.3, the transfer equation of  $\mathbf{u}(\mathbf{x}, t)$  is given by:

$$\frac{\partial \mathbf{u}}{\partial t} = (\xi^\top + \mathbf{I}_l) [\mathbf{D} \odot (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} \mathbf{u}) + \mathbf{D} \odot \beta \odot \{(\nabla_{\mathbf{x}} \mathbf{u})^\top \nabla_{\mathbf{x}} \phi\}], \quad (7)$$

where  $\nabla_{\mathbf{x}} \mathbf{u} = [\nabla_{\mathbf{x}} u_1, \dots, \nabla_{\mathbf{x}} u_l]$  and  $\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} \mathbf{u} = (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_1, \dots, \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_l)^\top$ .

Here,  $\nabla_{\mathbf{x}}$  denotes the spatial differential operator: applied to a scalar field, it returns the gradient; applied via dot product, it returns the divergence. The first term in Equation (7),  $\mathbf{D} \odot (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} \mathbf{u})$ , represents spontaneous diffusion driven by local density gradients [14, 66]. The second term,  $\mathbf{D} \odot \beta \odot \{(\nabla_{\mathbf{x}} \mathbf{u})^\top \nabla_{\mathbf{x}} \phi\}$ , captures directed relocation, arising from the alignment of the density gradient and the potential gradient  $\nabla_{\mathbf{x}} \phi$ , modulated by sensitivity  $\beta$ . This implies that conflict tends to flow toward lower-potential regions, especially where gradients are steep.

## 4.2 Regularization

Section 4.1 derives the diffusion mechanism of conflict-related factors based on the Boltzmann distribution assumption. Building on historical observations of conflict [4, 12], scholars have widely regarded this particle-like behavior as the dominant mode of transmission in most settings [5, 27, 38, 62, 79].

However, in rare instances of large-scale conflict, exceptional events, such as the coordinated movement of sizable armed groups, may drive a transition from diffuse dynamics to macroscopic, rigid-body-like motion, where the group advances as a cohesive whole [26, 35, 94]. To distinguish this behavior from the main diffusive term  $\mathbf{u}(\mathbf{x}, t)$ , we denote the rigid-body-like component as  $\mathbf{w}(\mathbf{x}, t)$ .

To model this phenomenon, we discretize the spatial domain into cells, each serving as a container for local rigid-body dynamics. Within each cell, conflict factors are assumed to move under the influence of a potential field  $\phi(\mathbf{x})$ , resulting in macroscopic behavior driven by force-based interactions.

Let  $\mathbf{A}(t, \Delta t, \mathbf{x})$  denote the displacement of conflict factors from the initial time  $t$  and initial position  $\mathbf{x}$  over a time interval  $\Delta t$ . The rigid-body-like transfer is then governed by:

$$\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mathbf{w}^{-1} \odot (\xi^\top \mathbf{D} \odot \beta \odot \nabla \phi). \quad (8)$$

Once the displacement  $\mathbf{A}$  is determined, the distribution  $\mathbf{w}(\mathbf{x}, t)$  is updated to reflect the new positions of the conflict factors. The updated distribution at time  $t + \Delta t$  is given by:

$$\mathbf{w}(\mathbf{x}, t + \Delta t) = \int_{\mathbf{A}(t, \Delta t, \mathbf{x}') = \mathbf{x}} \mathbf{w}(\mathbf{x}', t) A(t, \Delta t, \mathbf{x}') d\mathbf{x}'. \quad (9)$$

Here, the integral is taken over all initial positions  $\mathbf{x}'$  whose displacement  $\mathbf{A}(t, \Delta t, \mathbf{x}')$  maps to the target position  $\mathbf{x}$ .

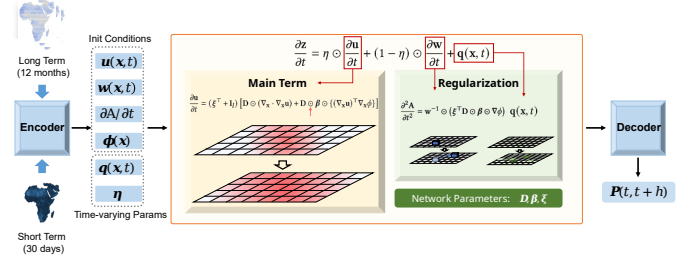
Empirical studies demonstrate that real-world conflicts combine microscopic diffusion of localized events with macroscopic, coordinated movement of armed groups [65, 94]. This motivates modeling conflict factors with two distinct transport modes: diffusive and rigid-body-like movement, controlled by a mixing probability vector  $\eta \in [0, 1]^l$ . Conflict intensity also evolves through endogenous processes and external influences [13, 15, 79, 85], represented by a **source term**  $\mathbf{q}(\mathbf{x}, t)$ . The source term in the diffusion-PDE represents the endogenous generation of new conflict potential at a specific location and time. It captures conflict dynamics that cannot be explained by propagation from neighboring areas, such as outbreaks driven by purely local factors.

Let  $\mathbf{z}(\mathbf{x}, t)$  represent the total conflict factors at each location. The complete transfer dynamics follow:

$$\frac{\partial \mathbf{z}}{\partial t} = \eta \odot \frac{\partial \mathbf{u}}{\partial t} + (1 - \eta) \odot \frac{\partial \mathbf{w}}{\partial t} + \mathbf{q}(\mathbf{x}, t), \quad (10)$$

where  $\mathbf{u}$  and  $\mathbf{w}$  correspond to the diffusive and rigid-body components respectively.

The structure permits several important exceptional cases: when  $\eta \rightarrow 1_l$ , it reduces to pure diffusion, which is suitable for modeling spontaneous unrest; when  $\eta \rightarrow 0_l$ , it describes coordinated troop movements—a rare but notable scenario. In practice, we encourage larger values of  $\eta$  during the solution process to favor diffusion-dominated dynamics, which aligns with the empirical observation that most conflicts evolve in a diffusive manner [5, 27, 38, 62, 79]. The coupling terms allow smooth transitions between these regimes,



**Figure 3: ConflictNet architecture.** The encoder extracts spatiotemporal features and debiases them via sample reweighting. The debiased representations are used to estimate initial conditions and time-varying parameters through MLPs, while intrinsic physical parameters are optimized in the network. These components are integrated into physics-informed transfer equations to forecast conflict dynamics.

adapting flexibly to the observed spatiotemporal patterns of conflict evolution.

## 4.3 ConflictNet

Building on the physical modeling in Sections 4.1 and 4.2, we introduce *ConflictNet*, a deep learning framework designed to solve the transfer equations. It integrates grid-based features to model the spatiotemporal evolution of armed conflict. The overall architecture is shown in Figure 3.

**Static-Dynamic Encoder with De-bias Module.** Our architecture begins by separating input features into static and dynamic components based on temporal variability, with detailed criteria provided in Appendix E. Static features undergo processing through a Graph Convolutional Network to capture spatial dependencies. In contrast, dynamic features are encoded using a hybrid GCN-LSTM module that jointly models spatial topology and temporal evolution.

To address regional and temporal biases inherent in conflict data, we implement an attention-based de-biasing mechanism. This approach learns prototypical representations by grouping instances according to two dimensions: temporal alignment follows historical conflict triggers documented by [32, 33], while spatial clustering follows regional patterns identified by [78]. Comprehensive implementation specifications for the de-biasing components are available in Appendix F.2. The integrated framework achieves dual objectives: maintaining sensitivity to both long-term monthly trends and short-term daily diffusion patterns, while systematically reducing dataset-specific biases.

**Physics-Informed Diffusion-PDE Solving.** The middle section of Figure 3 illustrates the estimation and solution of the physics-informed transfer equations detailed in Sections 4.1 and 4.2. The components of the equations can be categorized into three types.

The diffusion coefficient  $\mathbf{D}$ , the Boltzmann constant  $\beta$ , and the coupling coefficient  $\xi$  are treated as network parameters, which are directly optimized during training. These parameters capture the intrinsic properties of conflict regions and the interactions between conflict factors. They are assumed to be inherently stable over time because they reflect deep-seated, slow-moving structural characteristics rather than volatile, short-term events [1, 17]. For

**Table 1: Performance of various methods in predicting the probability of conflict in the  $h$ -th month, where  $h = 1, 3, 6, 9$ . Bold indicates the best result, and underlined indicates the second-best.**

	PR-AUC %( $\uparrow$ )				ROC-AUC %( $\uparrow$ )				$F_1$ %( $\uparrow$ )			
	1	3	6	9	1	3	6	9	1	3	6	9
ViEWS	36.4	32.8	31.0	29.9	94.8	92.2	90.4	88.6	34.7	31.4	29.6	28.8
XGBoost	37.4	33.1	30.8	30.1	93.9	93.1	92.4	91.6	35.4	31.4	28.7	27.3
RNN	33.4	29.7	29.1	28.6	92.1	91.6	91.4	91.3	35.7	33.6	32.4	31.9
GRU	35.8	31.5	30.2	29.9	92.2	91.3	91.0	90.8	39.5	36.4	35.8	35.4
LSTM	36.2	32.1	30.1	29.4	92.3	92.0	91.5	91.2	39.4	36.6	35.5	34.8
GCN	35.6	31.2	29.7	28.7	91.8	91.1	90.8	90.5	39.8	36.4	35.5	35.0
GAT	38.9	34.4	32.7	31.5	94.3	94.0	93.5	93.2	40.6	36.1	<u>36.4</u>	35.8
STSGCN	42.6	38.5	35.4	<u>32.7</u>	95.1	94.8	<u>94.6</u>	<u>94.4</u>	42.4	37.2	35.2	<u>36.1</u>
STGNN	43.3	39.2	35.4	31.9	95.2	<u>95.0</u>	<u>94.6</u>	<u>93.9</u>	42.5	38.3	36.1	35.9
Gragh WaveNet	37.1	33.7	31.9	30.2	94.1	93.8	93.7	93.5	39.2	36.9	35.8	34.6
STGCN	40.3	35.8	33.2	31.5	94.5	94.3	94.0	93.8	41.1	36.0	36.1	35.8
ST-GCN	40.8	36.1	33.8	31.4	94.7	94.2	94.0	93.7	41.3	37.9	<u>36.4</u>	36.0
D2STGNN	<u>44.8</u>	<u>39.6</u>	<u>35.7</u>	32.6	<u>95.6</u>	<u>95.0</u>	94.5	94.2	<u>42.7</u>	<u>38.9</u>	36.2	35.9
<b>ConflictNet</b>	<b>56.7</b>	<b>51.2</b>	<b>48.8</b>	<b>47.6</b>	<b>97.0</b>	<b>96.6</b>	<b>96.3</b>	<b>96.2</b>	<b>53.7</b>	<b>50.0</b>	<b>48.2</b>	<b>47.5</b>

example, these parameters are tied to enduring features such as organizational cohesion, the strategic importance of terrain [33, 56], and the logistical interplay between different conflict drivers, which do not typically change on a month-to-month basis.

The initial conditions of the conflict dynamics, including the initial values  $\mathbf{u}(\mathbf{x}, t)$ ,  $\mathbf{w}(\mathbf{x}, t)$ , the rigid-body initial velocity  $\frac{\partial \mathbf{A}}{\partial t}(\mathbf{x}, t)$ , and the potential field  $\phi(\mathbf{x})$ , are estimated from the encoder’s debiased feature representations using an MLP. This approach is theoretically sound and effective because the initial state of conflict dynamics is not arbitrary, but is strongly determined by observable structural conditions and recent historical events[8, 35, 36, 63]. The potential field is conceptually a direct function of enduring geographic and political realities; therefore, estimating it from static features like terrain, distance to borders, and ethnic settlement patterns is a well-grounded task, as these factors are known to create barriers and corridors for conflict. Similarly, the initial distribution of latent conflict drivers. Our model’s dynamic features—such as past conflict intensity and protest activity—serve as direct, observable indicators of where this latent potential is currently concentrated, consistent with theories of conflict inertia and contagion. The MLP, therefore, learns a mapping that is not arbitrary but is well-supported by established conflict theory.

The remaining time-varying parameters, specifically the mixing probability  $\eta$  and the source term  $\mathbf{q}(\mathbf{x}, t)$ , are also estimated using an MLP based on the encoder’s debiased feature representations. These parameters have a minor impact on the equation, as discussed in Section 4.2, and we further constrain their influence through the loss function to ensure the dominance of the diffusion term in the dynamics.

To solve the transfer equation iteratively, we employ a two-scale discretization strategy. Fine time steps are used to solve for  $\frac{\partial \mathbf{z}}{\partial t}$  and update the conflict factor distribution  $\mathbf{z}$ , ensuring dynamic consistency. Coarser time steps are then employed for estimating the conflict probability, which is decoded from the conflict factor

distribution  $\mathbf{z}$  using the shared-weight MLP decoder. This approach reduces the complexity of gradient backpropagation during training.

**Loss Function.** The total loss consists of a primary loss term and three auxiliary regularization terms, with weights controlled by the hyperparameter vector  $\lambda = [\lambda_1, \lambda_2, \lambda_3]^\top$ :

$$\mathcal{L} = \mathcal{L}_f + \lambda_1 \mathcal{R}_{\text{supp}} + \lambda_2 \mathcal{R}_{\text{continue}} + \lambda_3 \mathcal{R}_{\text{ap}}, \quad (11)$$

where  $\mathcal{L}_f$  represents the primary objective function, the  $\mathcal{R}_{\text{supp}}$  limits the diffusion, the  $\mathcal{R}_{\text{continue}}$  maintains diffusion continuity, and the  $\mathcal{R}_{\text{ap}}$  serves as a standard prediction enhancement regularizer. We will introduce and define them rigorously as follows.

- (1) The focal loss  $\mathcal{L}_f$  [49] serves as the primary objective, addressing class imbalance by placing greater emphasis on positive conflict samples, which are sparse in the dataset:

$$\mathcal{L}_f = -\alpha y(1 - P)^\gamma \log(P) - (1 - \alpha)(1 - y)P^\gamma \log(1 - P),$$

where  $P = P(t, t + h)$  denotes the predicted probability of a conflict in month  $t + h$  based on features at month  $t$ , and  $y \in \{0, 1\}$  is the binary ground-truth label indicating whether a conflict occurred. The hyperparameters  $\alpha$  and  $\gamma$  control class weighting and focusing strength, respectively.

- (2) The suppression regularization term  $\mathcal{R}_{\text{supp}}$  penalizes the magnitude of the source term  $\mathbf{q}$ . It encourages the mixing probability  $\eta$  to remain close to 1, thereby reinforcing the dominance of the diffusion-based component  $\mathbf{u}(\mathbf{x}, t)$ :

$$\mathcal{R}_{\text{supp}} = \|\mathbf{q}\| + \|1 - \eta\|.$$

- (3) The cyclic regularization term  $\mathcal{R}_{\text{continue}}$  [48] encourages temporal consistency by enforcing agreement between diffusion trajectories originating from different time points but converging to the same future state:

$$\mathcal{R}_{\text{continue}} = \|\mathbf{z}_{t-h \rightarrow t+h} - \mathbf{z}_{t \rightarrow t+h}\|,$$

**Table 2: Out-of-distribution performance of ConflictNet under spatial and temporal train-test splits. Despite substantial differences in conflict patterns across regions and periods, the model maintains high accuracy and recall, demonstrating strong generalization.**

	Train	Test	PR-AUC %( $\uparrow$ )				ROC-AUC %( $\uparrow$ )				$F_1$ %( $\uparrow$ )			
			1	3	6	9	1	3	6	9	1	3	6	9
	*	*	56.7	51.2	48.8	47.6	97.0	96.6	96.3	96.2	53.7	50.0	48.2	47.5
Spatial	E of 20°E	W of 20°E	53.5	49.0	46.8	44.1	96.5	96.2	95.9	95.6	50.1	48.3	46.7	45.2
	W of 20°E	E of 20°E	53.8	49.2	46.7	43.3	96.5	96.1	95.8	95.6	50.3	48.5	46.2	44.7
	S of 12°N	N of 12°N	51.6	46.4	42.8	39.2	96.3	95.9	95.6	95.5	49.2	47.5	45.9	43.9
	N of 12°N	S of 12°N	45.4	39.9	34.5	31.6	95.6	95.1	95.4	95.0	43.3	39.4	35.8	34.3
Temporal	1990.1–2008.9	2008.10–2020.12	55.5	50.0	47.1	43.8	96.8	96.5	96.2	96.0	51.4	49.5	47.8	46.1
	1990.1–2012.12	2013.1–2020.12	53.2	51.2	48.5	42.7	96.6	96.3	96.0	95.8	50.1	48.3	46.6	44.9

where  $\mathbf{z}_{i \rightarrow j}$  denotes the estimated conflict factor distribution propagated from time  $i$  to time  $j$ .

- (4) Last, the smooth average-precision regularization  $\mathcal{R}_{\text{ap}}$ , implemented via the fast-soft-rank operator [7], aims to improve ranking performance by enabling differentiable sorting [10]:

$$\mathcal{R}_{\text{ap}} = 1 - \frac{1}{\|S_P\|} \sum_{P \in S_P} \frac{\text{rank}(P, S_P)}{\text{rank}(P, S_\Omega)}, \quad (12)$$

where  $S_P$  is the set of positive samples,  $S_\Omega$  is the full set, and  $\text{rank}(x, S)$  denotes the (soft) rank of element  $x$  within set  $S$ .

## 5 Experiments

### 5.1 Experimental Setup

**Datasets.** We construct a monthly panel from 1990 to 2020 over Africa using PRIO-GRID [78] with  $0.5^\circ \times 0.5^\circ$  spatial resolution, resulting in 4,785,150 cell-months. Conflict labels are from UCDP-GED [77]. Features include structural indicators (PRIO-GRID, WDI), ethnic exclusion [18], protests (ACLED [67]), and institutions (V-Dem [50]). We provide a complete feature table in Supplementary Material Table 4. The dataset is sparse and heterogeneous across time and space.

**Baselines.** To benchmark our model, we compare it against a diverse set of forecasting methods. (1)**Classic models** such as ViEWS [33] and XGBoost [63] combine social science features with traditional machine learning and remain widely used in conflict prediction. (2)**Temporal models** including RNN [54], GRU [2] and LSTM [41] aim to capture sequential trends but struggle with overfitting in sparse data. (3)**Graph-based models** like GCN [44] and GAT [80] focus on spatial correlations but lack temporal modeling capacity. Spatio-temporal GNNs such as STSGCN [75], STGNN [76], and D2STGNN [73] attempt to integrate space-time dynamics via graph convolutions, yet often rely on black-box architectures that limit interpretability and generalization.

**Experiment Settings.** We construct a spatiotemporal graph where each PRIO-GRID cell serves as a node and is connected to its 24 nearest neighbors to define the spatial structure; time is discretized at the monthly level. Given past  $k = 12$  months of data, the task is to predict conflict occurrence in the  $h$ -th month, where  $h = 1, 3, 6, 9$ . The dataset is split chronologically: training (Jan

1990–Dec 2017), validation (Jan 2018–Dec 2018), and testing (Jan 2019–Dec 2020). The test set spans 24 months, and the validation set spans 12 months. Performance is evaluated using ROC-AUC, PR-AUC, and  $F_1$ -Score.

ROC-AUC measures the model’s discriminative ability across thresholds and corresponds to the area under the receiver operating characteristic (ROC) curve, which plots the true positive rate (TPR) against the false positive rate (FPR):

$$\text{TPR} = \frac{TP}{TP + FN}, \quad \text{FPR} = \frac{FP}{FP + TN},$$

where  $TP$ ,  $FP$ ,  $TN$ , and  $FN$  denote the number of true positives, false positives, true negatives, and false negatives, respectively.

PR-AUC evaluates precision-recall performance under class imbalance and is defined as the area under the precision-recall (PR) curve, where precision and recall are given by:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}.$$

The  $F_1$ -Score, which balances precision and recall, is defined as their harmonic mean:

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

### 5.2 Performance Comparison

We evaluate ConflictNet and all baselines on the ViEWS-formatted PRIO-GRID dataset. Table 1 reports PR-AUC, ROC-AUC, and  $F_1$ -Score across different prediction horizons ( $h = 1, 3, 6, 9$ ).

**ConflictNet outperforms all baselines, with a 56% improvement over ViEWS.** ConflictNet achieves consistent and substantial gains over all baselines. Compared to ViEWS—the field’s most established forecasting model—ConflictNet improves predictive accuracy by 56% in PR-AUC and 55% in  $F_1$ -Score at  $h = 1$ , demonstrating the value of incorporating physically grounded transfer dynamics over purely feature-based statistical learning. Even relative to the strongest deep learning baselines, ConflictNet delivers over 10-point gains across all metrics, highlighting its superior ability to capture structured spatiotemporal propagation.

**Temporal models fail to generalize under sparsity and class imbalance.** Temporal architectures such as RNN, GRU, and LSTM underperform, often failing to exceed classic baselines like



**Table 3: Ablation study results for one-month-ahead prediction. "w/o de-bias" removes the causal debiasing module; "w/o transfer" removes the physics-informed transfer module.**

	PR-AUC %( $\uparrow$ )	ROC-AUC %( $\uparrow$ )	$F_1$ %( $\uparrow$ )
ConflictNet	56.7	97.0	53.7
w/o de-bias	53.4	96.7	50.6
w/o transfer	40.3	94.9	39.1

XGBoost. Modeling long-range sequences from sparse, imbalanced conflict data leads to unstable optimization and poor generalization. Graph-based models (GCN, GAT) offer marginal improvements by encoding local spatial structure but lack dynamic representation, limiting their predictive reach.

**Spatiotemporal GNNs help but still lag behind explicit modeling.** Spatiotemporal GNNs (e.g., STGNN, D2STGNN) improve over single-modality baselines by jointly modeling space and time. However, their reliance on deep implicit architectures still limits performance under real-world sparsity and variability. In contrast, ConflictNet’s explicit formulation of diffusion and re-location mechanisms enables not only higher accuracy but also greater robustness across space and time, confirming the benefits of physics-informed modeling in high-resolution conflict forecasting.

### 5.3 Spatial-Temporal Generalization

To evaluate out-of-distribution performance, we construct train-test splits with maximal spatial and temporal divergence (Table 2). Spatially, we divide Africa along the 20°E meridian and 12°N parallel, separating regions with distinct conflict regimes—e.g., organized insurgencies in the Sahel and Horn versus localized communal violence in the south [22, 25, 31]. Despite these differences, ConflictNet retains over 95% of its in-domain ROC-AUC and achieves 43–50%  $F_1$ -Scores under cross-region transfer, indicating strong spatial generalization.

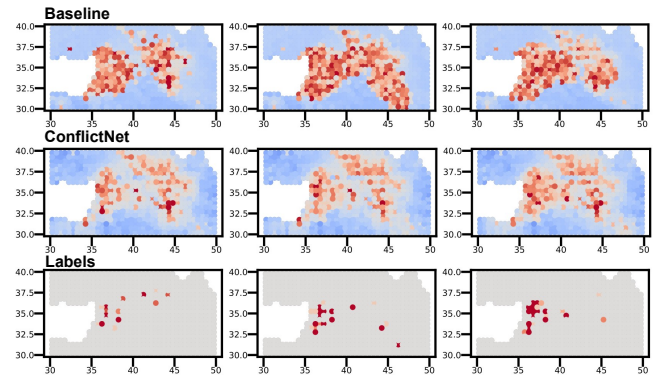
Temporally, we split at September 2008 (global financial crisis) and December 2012 (Arab Spring realignment), which mark major shifts in conflict intensity, actor structure, and external intervention [11, 46, 69, 82]. ConflictNet sustains ROC-AUCs above 96% and PR-AUCs within 2–4 points of in-domain levels. Notably, its PR-AUC under temporal shift remains well above that of all baselines even with access to the full dataset, underscoring the robustness of our approach in detecting rare events under structural change.

### 5.4 Ablation Studies

Table 3 shows ablation results. (1) **Removing the de-bias module** leads to moderate drops across all metrics, confirming its role in mitigating spatiotemporal bias. (2) **Removing the transfer module** causes sharp degradation, reducing performance to baseline levels, which highlights the importance of physics-informed dynamics in capturing conflict propagation and stabilizing training.

### 5.5 Case Study

To assess the plausibility of learned dynamics, we visualize predicted feature diffusion around Damascus from 2012.6 to 2013.6



**Figure 4: Predicted conflict factors (ConflictNet and baseline) and observed events in Damascus, Syria, every six months from June 2012 to June 2013.**

(Figure 4), alongside a baseline and observed conflicts at three-month intervals.

**Temporal dynamics.** ConflictNet exhibits a clear lead-lag structure: predicted feature spread emerges months before observed conflict events. This delay mirrors real-world escalation cycles, where armed groups require time to mobilize after entering new areas. The timing aligns with Syria’s conflict trajectory, expanding outward after mid-2012.

**Spatial continuity.** Compared to the baseline, ConflictNet produces smoother, spatially coherent patterns that reflect realistic regional diffusion rather than artificial pixel-level volatility. This suggests that the model respects physical constraints on actor movement—such as road networks and terrain—while the baseline yields fragmented outputs dominated by local noise.

**Directional accuracy.** ConflictNet correctly predicts conflict propagation from central Damascus toward Rif Dimashq, Homs, and Daraa, matching the documented advance of rebels. In contrast, the baseline fails to reproduce this spatial trajectory and overpredicts activity in unrelated zones.

These results show that physics-informed constraints improve both predictive performance and geographic plausibility, recovering realistic patterns of conflict diffusion.

## 6 Conclusion

Forecasting armed conflict remains a central challenge in peace research and policy. To move beyond statistical pattern-matching, we introduce ConflictNet, a physics-informed deep learning framework. Our framework uses a unified partial differential equation (PDE) to model a rich set of conflict dynamics, shaped by a learnable potential field that reflects geographic and political constraints. By embedding these structured dynamics, we achieve a 56% relative accuracy gain over the ViEWS gold-standard on high-resolution data, reveal interpretable diffusion patterns that align with social science theory, and demonstrate strong generalization across diverse regions and time periods. Ultimately, our work offers a new path for building mechanistic forecasting systems, though capturing fine-grained actor behaviors remains a key avenue for future research.



## References

- [1] Laia Balcells. 2010. Rivalry and revenge: Violence against civilians in conventional civil wars. *International Studies Quarterly* 54, 2 (2010), 291–313. doi:10.1111/j.1468-2478.2010.00590.x
- [2] Amulya Arun Ballakur and Arti Arya. 2020. Empirical Evaluation of Gated Recurrent Neural Network Architectures in Aviation Delay Prediction. In *2020 5th International Conference on Computing, Communication and Security (ICCCS)*. 1–7. doi:10.1109/ICCCS49678.2020.9276855
- [3] Peter Baudains, Hannah M Fry, Toby P Davies, and Alan G Wilson. 2016. A dynamic spatial model of conflict escalation. *European Journal of Applied Mathematics* 27, 3 (2016), 388–418. doi:10.1017/S095679251500056X
- [4] Mark R. Beissinger. 2007. Structure and Example in Modular Political Phenomena: The Diffusion of Bulldozer, Rose, Orange, and Tulip Revolutions. *Perspectives on Politics* 5, 2 (2007), 259–276. doi:10.1017/S1537592707070774
- [5] Andrea L Bertozzi, Martin B Short, P Jeffrey Brantingham, Georgios Sgourakis, and George Tita. 2010. A PDE model for criminal behavior. *SIAM J. Appl. Math.* 70, 6 (2010), 1984–2006. doi:10.1137/090748295
- [6] Stephen Biddle and Robert Zirkle. 1996. Technology, civil–military relations, and warfare in the developing world. *Journal of Strategic Studies* 19, 2 (1996), 171–212. doi:10.1080/01402399608437665
- [7] Mathieu Blondel, Olivier Teboul, Quentin Berthet, and Josip Djolonga. 2020. Fast differentiable sorting and ranking. In *International Conference on Machine Learning*. PMLR, 950–959.
- [8] Patrick T. Brandt, Vito D’Orazio, Latifur Khan, Yi-Fan Li, Javier Osorio, and Marcus Sianan. 2022. Conflict forecasting with event data and spatio-temporal graph convolutional networks. *International Interactions* 48, 4 (July 2022), 800–822. doi:10.1080/03050629.2022.2036987
- [9] Dirk Brockmann and Dirk Helbing. 2013. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science* 342, 6164 (2013), 1337–1342. doi:10.1126/science.1245200
- [10] Andrew Brown, Weidi Xie, Vicky Kalogeiton, and Andrew Zisserman. 2020. Smooth-AP: Smoothing the Path Towards Large-Scale Image Retrieval. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX* (Glasgow, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 677–694. doi:10.1007/978-3-030-58545-7\_39
- [11] Jason Brownlee, Tarek Masoud, and Andrew Reynolds. 2015. *The Arab Spring: Pathways of Repression and Reform*. Oxford University Press. <https://books.google.com/books?id=KbxBgAAQBAJ>
- [12] Halvard Buhaug and Kristian Skrede Gleditsch. 2008. Contagion or Confusion? Why Conflicts Cluster in Space. *International Studies Quarterly* 52, 2 (2008), 215–233. doi:10.1111/j.1468-2478.2008.00501.x
- [13] Halvard Buhaug and Jan Ketil Rød. 2006. Local determinants of African civil wars, 1970–2001. *Political Geography* 25, 3 (2006), 315–335. doi:10.1016/j.polgeo.2006.02.005
- [14] Lars-Erik Cederman. 2003. Modeling the size of wars: From billiard balls to sandpiles. *American Political Science Review* 97, 1 (2003), 135–150.
- [15] Lars-Erik Cederman and Kristian Skrede Gleditsch. 2009. Introduction to special issue on “Disaggregating civil war”. *Journal of Conflict Resolution* 53, 4 (2009), 487–495. doi:10.1177/0022002709337490
- [16] Lars-Erik Cederman, Kristian Skrede Gleditsch, and Halvard Buhaug. 2013. *Inequality, Grievances, and Civil War*. Cambridge University Press. doi:10.1017/CBO9781139084161
- [17] Lars-Erik Cederman and Nils B. Weidmann. 2017. Predicting armed conflict: Time to adjust our expectations? *Science* 355, 6324 (Feb. 2017), 474–476. doi:10.1126/science.aal4483
- [18] Lars-Erik Cederman, Nils B. Weidmann, and Kristian Skrede Gleditsch. 2011. Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison. *American Political Science Review* 105, 3 (Aug. 2011), 478–495. doi:10.1017/S0003055411000207
- [19] Thomas Chadeaux. 2022. A shape-based approach to conflict forecasting. *International Interactions* 48, 4 (July 2022), 633–648. doi:10.1080/03050629.2022.2009821
- [20] Yingjie Chen, Diqi Chen, Tao Wang, Yizhou Wang, and Yun Liang. 2022. Causal Intervention for Subject-Deconfounded Facial Action Unit Recognition. *ArXiv abs/2204.07935* (2022). <https://api.semanticscholar.org/CorpusID:248227393>
- [21] Paul Collier and Anke Hoefler. 1998. On economic causes of civil war. *Oxford Economic Papers* 50, 4 (Oct. 1998), 563–573. doi:10.1093/oepp/50.4.563 \_eprint: <https://academic.oup.com/oepp/article-pdf/50/4/563/6968570/50-4-563.pdf>
- [22] Jennifer G. Cooke and Thomas M. Sanderson. 2016. *Militancy and the Arc of Instability: Violent Extremism in the Sahel*. Center for Strategic and International Studies. [http://cis-website-prod.s3.amazonaws.com/s3fs-public/publication/160922\\_Sanderson\\_MilitancyArcInstabilitySahel\\_Web.pdf](http://cis-website-prod.s3.amazonaws.com/s3fs-public/publication/160922_Sanderson_MilitancyArcInstabilitySahel_Web.pdf)
- [23] Iain D Couzin, Jens Krause, Nigel R Franks, and Simon A Levin. 2005. Effective leadership and decision-making in animal groups on the move. *Nature* 433, 7025 (2005), 513–516. doi:10.1038/nature03236
- [24] Skyler J. Cranmer and Bruce A. Desmarais. 2017. What Can We Learn from Predictive Modeling? *Political Analysis* 25, 2 (April 2017), 145–166. doi:10.1017/pan.2017.3
- [25] Eric Denéce and Alain Rodier. 2012. *Global Security Risks and West Africa: Development Challenges*. OECD Publishing. [https://read.oecd-ilibrary.org/development/global-security-risks-and-west-africa\\_9789264171848-en#page=37](https://read.oecd-ilibrary.org/development/global-security-risks-and-west-africa_9789264171848-en#page=37)
- [26] Cassy Dorff, Max Gallop, and Shahryar Minhas. 2022. [W]hat Lies Beneath: Using Latent Networks to Improve Spatial Predictions. *International Studies Quarterly* 66, 1 (Feb. 2022), sqab086. doi:10.1093/isq/sqab086
- [27] Joshua M Epstein. 2002. Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences* 99, suppl 3 (2002), 7243–7250. doi:10.1073/pnas.092080199
- [28] James D. Fearon and David D. Laitin. 2003. Ethnicity, Insurgency, and Civil War. *American Political Science Review* 97, 01 (Feb. 2003), 75–90. doi:10.1017/S0003055403000534
- [29] Serge Galam. 1997. Rational group decision making: A random field Ising model at T=0. *Physica A: Statistical Mechanics and its Applications* 238, 1-4 (1997), 66–80. doi:10.1016/S0378-4371(96)00454-2
- [30] Ted Robert Gurr and Will H. Moore. 1997. Ethnopolitical Rebellion: A Cross-Sectional Analysis of the 1980s with Risk Assessments for the 1990s. *American Journal of Political Science* 41, 4 (1997), 1079–1103. doi:10.2307/2960482 Publisher: [Midwest Political Science Association, Wiley].
- [31] Stig Jarle Hansen. 2019. *Horn, Sahel, and Rift: Fault-lines of the African Jihad*. Hurst & Company. <https://books.google.com/books?id=S2-fDwAAQBAJ>
- [32] Håvard Hegre, Marie Allansson, Matthias Basedau, Michael Colaresi, Mihai Croicu, Hanne Fjelde, Frederick Hoyles, Lisa Hultman, Stina Högbladh, Remco Jansen, Naima Mouhle, Sayeed Awn Muhammad, Desirée Nilsson, Håvard Mokleiv Nygård, Gudlaug Olafsdottir, Kristina Petrova, David Randahl, Espen Geelmuyden Rød, Gerald Schneider, Nina Von Uexkull, and Jonas Vestby. 2019. ViEWS: A political violence early-warning system. *Journal of Peace Research* 56, 2 (March 2019), 155–174. doi:10.1177/0022343319823860
- [33] Håvard Hegre, Curtis Bell, Michael Colaresi, Mihai Croicu, Frederick Hoyles, Remco Jansen, Maxine Ria Leis, Angelica Lindqvist-McGowan, David Randahl, Espen Geelmuyden Rød, and Paola Vesco. 2021. ViEWS<sub>2020</sub>: Revising and evaluating the ViEWS political violence Early-Warning System. *Journal of Peace Research* 58, 3 (May 2021), 599–611. doi:10.1177/0022343320962157
- [34] Håvard Hegre, Nils W Metternich, Håvard Mokleiv Nygård, and Julian Wucherpfennig. 2017. Introduction: Forecasting in peace research. *Journal of Peace Research* 54, 2 (March 2017), 113–124. doi:10.1177/0022343317691330
- [35] Håvard Hegre, Håvard Mokleiv Nygård, and Peder Landsverk. 2021. Can We Predict Armed Conflict? How the First 9 Years of Published Forecasts Stand Up to Reality. *International Studies Quarterly* 65, 3 (Sept. 2021), 660–668. doi:10.1093/isq/sqaa094
- [36] Håvard Hegre, Paola Vesco, and Michael Colaresi. 2022. Lessons from an escalation prediction competition. *International Interactions* 48, 4 (July 2022), 521–554. doi:10.1080/03050629.2022.2070745
- [37] Dirk Helbing. 1995. *Quantitative Sociodynamics: Stochastic Methods and Models of Social Interaction Processes*. Springer. doi:10.1007/978-3-642-85295-2
- [38] Dirk Helbing. 2001. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics* 73, 4 (2001), 1067–1141. doi:10.1103/RevModPhys.73.1067
- [39] Dirk Helbing, Illés J Farkas, and Tamás Vicsek. 2000. Simulating dynamical features of escape panic. *Nature* 407, 6803 (2000), 487–490. doi:10.1038/35035023
- [40] Jan Hermann, Zeno Schätzle, and Frank Noé. 2020. Deep-neural-network solution of the electronic Schrödinger equation. *Nature Chemistry* 12, 10 (Oct. 2020), 891–897. doi:10.1038/s41557-020-0544-y
- [41] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* 9, 8 (1997), 1735–1780. doi:10.1162/neco.1997.9.8.1735
- [42] Petter Holme and Fredrik Liljeros. 2015. Mechanistic Models in Computational Social Science. *Frontiers in Physics* 3 (2015), 78. doi:10.3389/fphy.2015.00078
- [43] Elizabeth E. Holmes. 1993. Are Diffusion Models Too Simple? A Comparison with Telegraph Models of Invasion. *The American Naturalist* 142, 5 (1993), 779–795. doi:10.1086/285572
- [44] Bo Jiang, Ziyang Zhang, Doudou Lin, Jin Tang, and Bin Luo. 2019. Semi-Supervised Learning With Graph Learning-Convolutional Networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 11305–11312. doi:10.1109/CVPR.2019.01157
- [45] Stathis N. Kalyvas. 2006. *The Logic of Violence in Civil War*. Cambridge University Press. doi:10.1017/CBO9780511818462
- [46] Andrey Korotayev, Leonid Issaev, and Sergey Malkov. 2022. The Arab Spring: A Quantitative Analysis. In *Handbook of Revolutions in the 21st Century*. Springer, 777–807. [https://www.academia.edu/download/96681268/2022\\_Goldstone\\_et\\_al\\_Revolutions.pdf#page=778](https://www.academia.edu/download/96681268/2022_Goldstone_et_al_Revolutions.pdf#page=778)
- [47] P. L. Krapivsky, S. Redner, and E. Ben-Naim. 2010. *A Kinetic View of Statistical Physics*. Cambridge University Press, Cambridge, UK.
- [48] Zhengqi Li, Qianqian Wang, Forrester Cole, Richard Tucker, and Noah Snavely. 2023. DynIBaR: Neural Dynamic Image-Based Rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 4273–4284.
- [49] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal Loss for Dense Object Detection. In *2017 IEEE International Conference on Computer Vision (ICCV)*. 2999–3007. doi:10.1109/ICCV.2017.324

- [50] Staffan I Lindberg, Michael Coppedge, John Gerring, and Jan Teorell. 2014. V-Dem: A new way to measure democracy. *Journal of Democracy* 25, 3 (2014), 159–169. Publisher: Johns Hopkins University Press.
- [51] Andreas Lindholm, Johannes Hendriks, Adrian Wills, and Thomas B. Schön. 2022. Predicting political violence using a state-space model. *International Interactions* 48, 4 (July 2022), 759–777. doi:10.1080/03050629.2022.2094921
- [52] Christopher Linebarger. 2015. Civil War Diffusion and the Emergence of Militant Groups, 1960–2001. *International Interactions* 41, 3 (May 2015), 583–600. doi:10.1080/03050629.2015.984809
- [53] Aila M. Matanock. 2017. Bullets for Ballots: Electoral Participation Provisions and Enduring Peace after Civil Conflict. *International Security* 41, 4 (April 2017), 93–132. doi:10.1162/ISEC\_a\_00275
- [54] Larry R Medsker and LC Jain. 2001. Recurrent neural networks. *Design and Applications* 5, 64–67 (2001), 2.
- [55] Sidney Michelini, Barbora Šedová, Jacob Schewe, and Katja Frieler. 2023. Extreme weather impacts do not improve conflict predictions in Africa. *Humanities and Social Sciences Communications* 10, 1 (Aug. 2023), 522. doi:10.1057/s41599-023-01996-1
- [56] David Muchlinski, David Siroky, Jingrui He, and Matthew Kocher. 2016. Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data. *Political Analysis* 24, 1 (2016), 87–103. doi:10.1093/pan/mpv024
- [57] Hannes Mueller and Christopher Rauh. 2018. Reading Between the Lines: Prediction of Political Violence Using Newspaper Text. *American Political Science Review* 112, 2 (May 2018), 358–375. doi:10.1017/S0003055417000570
- [58] Alexandre Nicolas and Fahmy H. Hassan. 2023. Social groups in pedestrian crowds: review of their influence on the dynamics and their modelling. *Transportmetrica A: Transport Science* (2023). doi:10.1080/23249935.2021.1970651
- [59] John O’Loughlin, Frank D. W. Witmer, Andrew M. Linke, and Arna Thorwardson. 2010. The locational determinants of civil war violence: Evidence from African conflicts. *Political Geography* 29, 8 (2010), 638–648. doi:10.1016/j.polgeo.2010.07.005
- [60] Ekaterina V. Orlova. 2024. A Novel Brillouin and Langevin Functions Dynamic Model for Two Conflicting Social Groups: Study of R&D Processes. *Mathematics* 12, 17 (2024), 2788. doi:10.3390/math12172788
- [61] David Pfau, James S. Spencer, Alexander G. D. G. Matthews, and W. M. C. Foulkes. 2020. *Ab initio* solution of the many-electron Schrödinger equation with deep neural networks. *Physical Review Research* 2, 3 (Sept. 2020), 033429. doi:10.1103/PhysRevResearch.2.033429
- [62] Mason A Porter and James P Gleeson. 2016. *Dynamical Systems on Networks: A Tutorial*. Frontiers in Applied Dynamical Systems: Reviews and Tutorials, Vol. 4. Springer. doi:10.1007/978-3-319-26629-8
- [63] Andre Python, Andreas Bender, Anita K. Nandi, Penelope A. Hancock, Rohan Arambepola, Jürgen Brandsch, and Tim C. D. Lucas. 2021. Predicting non-state terrorism worldwide. *Science Advances* 7, 31 (July 2021), eabg4778. doi:10.1126/sciadv.abg4778
- [64] Paul Raccuglia, Katherine C. Elbert, Philip D. F. Adler, Casey Falk, Malia B. Wenny, Aurelio Mollo, Matthias Zeller, Sorelle A. Friedler, Joshua Schrier, and Alexander J. Norquist. 2016. Machine-learning-assisted materials discovery using failed experiments. *Nature* 533, 7601 (May 2016), 73–76. doi:10.1038/nature17439
- [65] Benjamin J. Radford. 2022. High resolution conflict forecasting with spatial convolutions and long short-term memory. *International Interactions* 48, 4 (July 2022), 739–758. doi:10.1080/03050629.2022.2031182
- [66] Clionadh Raleigh, Andrew Linke, Håvard Hegre, and Joakim Karlsen. 2010. Introducing ACLED: An armed conflict location and event dataset. *Journal of Peace Research* 47, 5 (2010), 651–660. doi:10.1177/0022343310378914
- [67] Clionadh Raleigh, Andrew Linke, Håvard Hegre, and Joakim Karlsen. 2010. Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research* 47, 5 (Sept. 2010), 651–660. doi:10.1177/0022343310378914
- [68] Espen Geelmuyden Rød, Håvard Hegre, and Maxine Leis. 2023. Predicting armed conflict using protest data. *Journal of Peace Research* (Sept. 2023), 00223433231186452. doi:10.1177/00223433231186452
- [69] Stephen M. Saideman. 2012. When Conflict Spreads: Arab Spring and the Limits of Diffusion. *International Interactions* 38, 5 (2012), 713–722. https://www.researchgate.net/publication/263615147
- [70] Katherine Sawyer and Talbot M. Andrews. 2020. Rebel recruitment and retention in civil conflict. *International Interactions* 46, 6 (Nov. 2020), 872–892. doi:10.1080/03050629.2020.1814765
- [71] Philip A. Schrodt. 2014. Seven deadly sins of contemporary quantitative political analysis. *Journal of Peace Research* 51, 2 (2014), 287–300. doi:10.1177/0022343313517991
- [72] Sebastian Schutte, Jonas Vestby, Jørgen Carling, and Halvard Buhaug. 2021. Climatic conditions are weak predictors of asylum migration. *Nature Communications* 12, 1 (April 2021), 2067. doi:10.1038/s41467-021-22255-4
- [73] Zezhi Shao, Zhao Zhang, Wei Wei, Fei Wang, Yongjun Xu, Xin Cao, and Christian S. Jensen. 2022. Decoupled dynamic spatial-temporal graph neural network for traffic forecasting. *Proc. VLDB Endow.* 15, 11 (July 2022), 2733–2746. doi:10.14778/3551793.3551827
- [74] Kim Sharp and Franz Matschinsky. 2015. Translation of Ludwig Boltzmann’s Paper “On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium”. *Entropy* 17, 4 (2015), 1971–2009. doi:10.3390/e17041971
- [75] Chao Song, Youfang Lin, S. Guo, and Huaiyu Wan. 2020. Spatial-Temporal Synchronous Graph Convolutional Networks: A New Framework for Spatial-Temporal Network Data Forecasting. In *AAAI Conference on Artificial Intelligence*. https://api.semanticscholar.org/CorpusID:214274349
- [76] Bei Su and Wen Zheng. 2020. Traffic Flow Prediction via Spatial Temporal Neural Network “ResLS-C”. In *2020 Eighth International Conference on Advanced Cloud and Big Data (CBD)*. 119–124. doi:10.1109/CBD51900.2020.00030
- [77] Ralph Sundberg and Erik Melander. 2013. Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research* 50, 4 (2013), 523–532.
- [78] Andreas Forø Tollefsen, Håvard Strand, and Halvard Buhaug. 2012. PRIO-GRID: A unified spatial data structure. *Journal of Peace Research* 49, 2 (March 2012), 363–374. doi:10.1177/0022343311431287
- [79] Peter Turchin. 2003. *Historical Dynamics: Why States Rise and Fall*. Princeton University Press, Princeton, NJ.
- [80] Petar Velicković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph Attention Networks. arXiv:1710.10903 [stat.ML]
- [81] Paola Vesco, Håvard Hegre, Michael Colaresi, Remco Bastiaan Jansen, Adeline Lo, Gregor Reisch, and Nils B. Weidmann. 2022. United they stand: Findings from an escalation prediction competition. *International Interactions* 48, 4 (July 2022), 860–896. doi:10.1080/03050629.2022.2029856
- [82] Frederic Volpi and Johannes Gerschewski. 2020. Crises and Critical Junctures in Authoritarian Regimes: Addressing Uprisings’ Temporalities and Discontinuities. *Third World Quarterly* 41, 6 (2020), 1031–1051. doi:10.1080/01436597.2020.1729728
- [83] Michael D Ward, Brian D Greenhill, and Kristin M Bakke. 2010. The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research* 47, 4 (July 2010), 363–375. doi:10.1177/0022343309356491
- [84] Michael D. Ward, Nils W. Metternich, Cassy L. Dorff, Max Gallop, Florian M. Hollenbach, Anna Schultz, and Simon Weschle. 2013. Learning from the Past and Stepping into the Future: Toward a New Generation of Conflict Prediction. *International Studies Review* 15, 4 (Dec. 2013), 473–490. doi:10.1111/misr.12072
- [85] Nils B Weidmann. 2009. Geography as motivation and opportunity: Group concentration and ethnic conflict. *Journal of Conflict Resolution* 53, 4 (2009), 526–543. doi:10.1177/0022002709336450
- [86] Nils B. Weidmann and Michael D. Ward. 2010. Predicting Conflict in Space and Time. *Journal of Conflict Resolution* 54, 6 (Dec. 2010), 883–901. doi:10.1177/0022002710371669
- [87] Jared Willard, Xiaowei Jia, Shaoming Xu, Michael Steinbach, and Vipin Kumar. 2022. Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. http://arxiv.org/abs/2003.04919 arXiv:2003.04919 [physics, stat].
- [88] World Bank. 2018. *Pathways for Peace: Inclusive Approaches to Preventing Violent Conflict*. World Bank, Washington, DC. https://openknowledge.worldbank.org/handle/10986/28337
- [89] Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. 2023. TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis. arXiv:2210.02186 [cs.LG] https://arxiv.org/abs/2210.02186
- [90] Qingbiao Wu and Yueqing Zhao. 2006. The convergence theorem for a family deformed Chebyshev method in Banach space. *Appl. Math. Comput.* 182, 2 (2006), 1369–1376. doi:10.1016/j.amc.2006.05.022
- [91] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. 2019. Graph WaveNet for Deep Spatial-Temporal Graph Modeling. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 1907–1913. doi:10.24963/ijcai.2019/264
- [92] Sijie Yan, Yuanjun Xiong, and Dahua Lin. 2018. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence* (New Orleans, Louisiana, USA) (AAAI’18/IAAI’18/EAAI’18). AAAI Press, Article 912, 9 pages.
- [93] Bing Yu, Haoteng Yin, and Zhanxing Zhu. 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*. International Joint Conferences on Artificial Intelligence Organization, 3634–3640. doi:10.24963/ijcai.2018/505
- [94] Andrew Zammit-Mangion, Michael Dewar, Visakan Kadirkamanathan, and Guido Sanguinetti. 2012. Point process modelling of the Afghan War Diary. *Proceedings of the National Academy of Sciences* 109, 31 (July 2012), 12414–12419. doi:10.1073/pnas.1203177109
- [95] Yuchen Zhang, Mingsheng Long, Kaiyuan Chen, Lanxiang Xing, Ronghua Jin, Michael I. Jordan, and Jianmin Wang. 2023. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature* 619, 7970 (July 2023), 526–532. doi:10.1038/s41586-023-06184-4

# Supplementary Material

## A More Theoretical Results

### A.1 Diffusion of a Single Type of Conflict Factors

We can first consider diffusion of a single type of conflict factors. At this time, diffusion does not involve the coupling and mutual influence multiple types of conflict factors, which is a simpler situation than Theorem 4.3.

**Theorem A.1.** Under Assumption 4.1, 4.2 and condition of Proposition 4.1, 4.2, diffusion of a single type of conflict factors is given by:

$$\frac{\partial u}{\partial t} = D \nabla \cdot \nabla u + \beta D \nabla u \cdot \nabla \phi. \quad (13)$$

### A.2 More Discussions on Assumption 4.2

Before introducing the additional theoretical result, we need a technical lemma.

**Lemma A.2** (expectation of a quadratic form). Under the condition of Proposition 4.1, for any matrix  $\mathbf{B}$  independent of  $\mathbf{r}$  and  $\Delta t$ , we have

$$\int_{\mathbb{R}^2} \mathbf{r}^\top \mathbf{B} \mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = \text{tr}(\mathbf{B} \int_{\mathbb{R}^2} \mathbf{r} \mathbf{r}^\top P(\|\mathbf{r}\|, \Delta t) d\mathbf{r}).$$

**Theorem A.3.** Under Assumption 4.2, we have:

Many classic kernels satisfy Assumption 4.2. Here we can give two classic examples:

**Example A.1** (two-dimensional Gaussian heat kernel).

$$P(\|\mathbf{r}\|, \Delta t) = \frac{1}{4\pi A \Delta t} \exp\left(-\frac{\|\mathbf{r}\|^2}{4A \Delta t}\right), \quad (14)$$

where  $A$  is a constant.

**Example A.2** (the uniform kernel over a shrinking ball).

$$P(\|\mathbf{r}\|, \Delta t) = \frac{1}{\pi R^2(\Delta t)} \begin{cases} 1, & \text{if } \|\mathbf{r}\| \leq R(\Delta t) \\ 0, & \text{otherwise} \end{cases}, \quad (15)$$

where  $R(\Delta t) = \sqrt{2A \Delta t}$ .

**Proposition A.4.** The  $P(\|\mathbf{r}\|, \Delta t)$  defined in Example A.1 and A.2 satisfy Assumption 4.2

## B Proof of the Results in the Supplementary Materials

### B.1 Proof of Lemma A.2

Recall the classic theorem: for random variable  $X$  and matrix  $B$  we have

$$E[X^\top B X] = E[X]^\top B E[X] + \text{tr}(B \cdot \text{cov}[X, X]). \quad (16)$$

Choose  $X = \mathbf{r}$  and we can obtain Lemma A.2.

### B.2 Proof of Theorem A.1

We begin with the integral form of the density evolution:

$$u(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) u(\mathbf{x} - \mathbf{r}, t) d\mathbf{r}. \quad (17)$$

For convenience, we denote  $u(\mathbf{x}, t)$  and  $\phi(\mathbf{x})$  as  $u$  and  $\phi$ , correspondingly.

**B.2.1 Taylor expansion of (17).** First, we perform a second-order Taylor expansion of  $u(\mathbf{x} - \mathbf{r}, t)$  around  $\mathbf{x}$ :

$$u(\mathbf{x} - \mathbf{r}, t) = u - \mathbf{r}^\top \nabla u + \frac{1}{2} \mathbf{r}^\top \nabla^2 u \mathbf{r} + O(\|\mathbf{r}\|^3). \quad (18)$$

Substitute Equation (18) into Equation (17), we have:

$$u(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) \left[ u - \mathbf{r}^\top \nabla u + \frac{1}{2} \mathbf{r}^\top \nabla^2 u \mathbf{r} + O(\|\mathbf{r}\|^3) \right] d\mathbf{r}. \quad (19)$$

Since  $P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t)$  is the probability density, we have

$$\int_{\mathbb{R}^2} u P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) d\mathbf{r} = u \int_{\mathbb{R}^2} P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) d\mathbf{r} = u.$$

Therefore, Equation (19) is simplified to:

$$u(\mathbf{x}, t + \Delta t) = u + \int_{\mathbb{R}^2} P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) \left[ -\mathbf{r}^\top \nabla u + \frac{1}{2} \mathbf{r}^\top \nabla^2 u \mathbf{r} + O(\|\mathbf{r}\|^3) \right] d\mathbf{r}. \quad (20)$$

According to Proposition 4.1, the transition kernel is given by:

$$P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) = \frac{1}{Z} \exp\left(-\frac{\beta}{2} [\phi(\mathbf{x}) - \phi(\mathbf{x} - \mathbf{r})]\right) P(\|\mathbf{r}\|, \Delta t). \quad (21)$$

We first expand the potential difference in the exponent in (21):

$$\phi(\mathbf{x}) - \phi(\mathbf{x} - \mathbf{r}) = \mathbf{r}^\top \nabla \phi + O(\|\mathbf{r}\|^2).$$

Substituting  $y = [\phi(\mathbf{x}) - \phi(\mathbf{x} - \mathbf{r})]$  and expanding the exponential function  $e^y$  at  $y = 0$ , we have:

$$P_\phi(\mathbf{x} - \mathbf{r}, \mathbf{r}, \Delta t) = \frac{1}{Z} \left[ 1 - \frac{\beta}{2} \mathbf{r}^\top \nabla \phi + O(\|\mathbf{r}\|^2) \right] P(\|\mathbf{r}\|, \Delta t). \quad (22)$$

Further, by substituting Equation (22) into the right-hand side of Equation (20), we can obtain

$$\begin{aligned} u(\mathbf{x}, t + \Delta t) - u &= \int_{\mathbb{R}^2} \frac{1}{Z} \left[ 1 - \frac{\beta}{2} \mathbf{r}^\top \nabla \phi + O(\|\mathbf{r}\|^2) \right] \left[ -\mathbf{r}^\top \nabla u + \frac{1}{2} \mathbf{r}^\top \nabla^2 u \mathbf{r} + O(\|\mathbf{r}\|^3) \right] P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} \\ &= \int_{\mathbb{R}^2} \frac{1}{Z} \left[ -\mathbf{r}^\top \nabla u + \frac{\beta}{2} \mathbf{r}^\top \nabla \phi \mathbf{r}^\top \nabla u + \frac{1}{2} \mathbf{r}^\top \nabla^2 u \mathbf{r} + O(\|\mathbf{r}\|^3) \right] P(\|\mathbf{r}\|, \Delta t) d\mathbf{r}. \end{aligned} \quad (23)$$

By Equation (4), we have:

$$\int_{\mathbb{R}^2} -\mathbf{r}^\top \nabla u P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = 0.$$

Therefore, Equation (23) is simplified to:

$$\begin{aligned} u(\mathbf{x}, t + \Delta t) - u &= \int_{\mathbb{R}^2} \left[ \frac{\beta}{2Z} \mathbf{r}^\top \nabla \phi \mathbf{r}^\top \nabla u + \frac{1}{2Z} \mathbf{r}^\top \nabla^2 u \mathbf{r} + O(\|\mathbf{r}\|^3) \right] P(\|\mathbf{r}\|, \Delta t) d\mathbf{r}. \end{aligned} \quad (24)$$



**B.2.2 Second-Order Terms in (24).** First we examine the first term in Equation (24):

$$\begin{aligned} \int_{\mathbb{R}^2} \frac{\beta}{2Z} \mathbf{r}^\top \nabla \phi \mathbf{r}^\top \nabla u P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} &= \int_{\mathbb{R}^2} \frac{\beta}{2Z} \nabla \phi^\top \mathbf{r} \mathbf{r}^\top \nabla u P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} \\ &= \frac{\beta}{2Z} \nabla \phi^\top \int_{\mathbb{R}^2} \mathbf{r} \mathbf{r}^\top P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} \nabla u. \end{aligned}$$

By Proposition 4.2, the first term in Equation (24) can be simplified as:

$$\int_{\mathbb{R}^2} \frac{\beta}{2Z} \mathbf{r}^\top \nabla \phi \mathbf{r}^\top \nabla u P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = \beta D \nabla u \cdot \nabla \phi \Delta t. \quad (25)$$

Then we examine the second term in Equation (24):

$$\int_{\mathbb{R}^2} \frac{1}{2Z} \mathbf{r}^\top \nabla^2 u \mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r}. \quad (26)$$

Since matrix  $\nabla^2 u$  is independent of  $\mathbf{r}$  and  $\Delta t$ , according to Lemma A.2, we can rewrite Equation (26) as:

$$\int_{\mathbb{R}^2} \frac{1}{2Z} \mathbf{r}^\top \nabla^2 u \mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = \text{tr}(\nabla^2 u \frac{1}{2Z} \int_{\mathbb{R}^2} \mathbf{r} \mathbf{r}^\top P(\|\mathbf{r}\|, \Delta t) d\mathbf{r}). \quad (27)$$

By Proposition 4.2, we simplify Equation (27) as:

$$\int_{\mathbb{R}^2} \frac{1}{2Z} \mathbf{r}^\top \nabla^2 u \mathbf{r} P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = \text{tr}(\nabla^2 u D \Delta t \mathbf{I}_2) = D(\nabla \cdot \nabla u) \Delta t. \quad (28)$$

**B.2.3 Combine Results.** Substituting Equation (25) and (28) into Equation 24, we have:

$$\begin{aligned} u(\mathbf{x}, t + \Delta t) - u &= \beta D \nabla u \cdot \nabla \phi \Delta t + D(\nabla \cdot \nabla u) \Delta t \\ &\quad + \int_{\mathbb{R}^2} O(\|\mathbf{r}\|^3) P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} \end{aligned} \quad (29)$$

By (3), we have

$$\int_{\mathbb{R}^2} O(\|\mathbf{r}\|^3) P(\|\mathbf{r}\|, \Delta t) d\mathbf{r} = o(\Delta t).$$

It implies

$$u(\mathbf{x}, t + \Delta t) - u = \beta D \nabla u \cdot \nabla \phi \Delta t + D(\nabla \cdot \nabla u) \Delta t + o(\Delta t).$$

Divide both sides by  $\Delta t$  and let  $\Delta t \rightarrow 0$ , we obtain

$$\frac{\partial u}{\partial t} = D \nabla \cdot \nabla u + \beta D \nabla u \cdot \nabla \phi.$$

### B.3 Proof of Proposition A.4

**B.3.1 Gaussian kernel.** In polar coordinates:  $d^2 \mathbf{r} = r dr d\theta$ .

$$\begin{aligned} E[\|\mathbf{r}\|^2] &= \int_0^{2\pi} \int_0^\infty r^2 \frac{1}{4\pi A \Delta t} e^{-r^2/(4A \Delta t)} r dr d\theta \\ &= \frac{1}{2A \Delta t} \int_0^\infty r^3 e^{-r^2/(4A \Delta t)} dr \end{aligned}$$

Let  $u = r^2/(4A \Delta t)$ , then we have  $r^3 dr = 8(A \Delta t)^2 u du$ :

$$E[\|\mathbf{r}\|^2] = \frac{1}{2A \Delta t} \cdot 8(A \Delta t)^2 \int_0^\infty u e^{-u} du = 4A \Delta t$$

since  $\int_0^\infty u e^{-u} du = 1! = 1$ .

The third-order absolute moment is defined as

$$\begin{aligned} M_3(\Delta t) &= E[\|\mathbf{r}\|^3] = \int_{\mathbb{R}^2} \|\mathbf{r}\|^3 P(\|\mathbf{r}\|, \Delta t) d^2 \mathbf{r} \\ &= \int_0^{2\pi} \int_0^\infty r^3 \frac{1}{4\pi A \Delta t} e^{-r^2/(4A \Delta t)} r dr d\theta \\ &= \frac{1}{2A \Delta t} \int_0^\infty r^4 e^{-r^2/(4A \Delta t)} dr. \end{aligned}$$

By making the substitution  $u = r^2/(4A \Delta t)$ , we transform the integral with  $r = 2\sqrt{A \Delta t} \sqrt{u}$  and  $dr = \sqrt{A \Delta t}/\sqrt{u} du$ . The differential element becomes  $r^4 dr = 16(A \Delta t)^{5/2} u^{3/2} du$ . Substituting this back yields  $M_3(\Delta t) = \frac{1}{2A \Delta t} \cdot 16(A \Delta t)^{5/2} \int_0^\infty u^{3/2} e^{-u} du = 8(A \Delta t)^{3/2} \Gamma(5/2)$ , where  $\Gamma(5/2) = 3\sqrt{\pi}/4$ .

The final result is  $M_3(\Delta t) = 6\sqrt{\pi} A^{3/2} (\Delta t)^{3/2}$ , which exhibits the scaling behavior  $O((\Delta t)^{3/2}) = o(\Delta t)$  as  $\Delta t \rightarrow 0$ . This confirms that the Gaussian kernel satisfies  $\int_{\mathbb{R}^2} \|\mathbf{r}\|^3 P d^2 \mathbf{r} = o(\Delta t)$ .

**B.3.2 uniform kernel.** The second-order moment is defined as

$$\mathbf{M}_2(\Delta t) = \int_{\mathbb{R}^2} \mathbf{r} \mathbf{r}^\top P(\|\mathbf{r}\|, \Delta t) d^2 \mathbf{r}.$$

First, we compute the radial second-order moment  $E[\|\mathbf{r}\|^2]$ . In polar coordinates,  $d^2 \mathbf{r} = r dr d\theta$ , yielding

$$\begin{aligned} E[\|\mathbf{r}\|^2] &= \int_0^{2\pi} \int_0^R r^2 \frac{1}{\pi R^2} r dr d\theta \\ &= \frac{2\pi}{\pi R^2} \int_0^R r^3 dr \\ &= \frac{2}{R^2} \left[ \frac{r^4}{4} \right]_0^R = \frac{R^2}{2}. \end{aligned}$$

Substituting  $R(\Delta t) = \sqrt{2A \Delta t}$ , we obtain

$$E[\|\mathbf{r}\|^2] = \frac{2A \Delta t}{2} = A \Delta t.$$

Therefore, this kernel satisfies

$$\int_{\mathbb{R}^2} \|\mathbf{r}\|^2 P d\mathbf{r} \asymp \Delta t.$$

The third-order absolute moment is defined as

$$M_3(\Delta t) = E[\|\mathbf{r}\|^3] = \int_{\mathbb{R}^2} \|\mathbf{r}\|^3 P(\|\mathbf{r}\|, \Delta t) d^2 \mathbf{r}.$$

In polar coordinates,  $d^2 \mathbf{r} = r dr d\theta$ , leading to

$$\begin{aligned} M_3(\Delta t) &= \int_0^{2\pi} \int_0^R r^3 \frac{1}{\pi R^2} r dr d\theta \\ &= \frac{2\pi}{\pi R^2} \int_0^R r^4 dr \\ &= \frac{2}{R^2} \left[ \frac{r^5}{5} \right]_0^R = \frac{2}{5} R^3. \end{aligned}$$

Substituting  $R(\Delta t) = \sqrt{2A \Delta t}$ , we obtain

$$M_3(\Delta t) = \frac{2}{5} (\sqrt{2A \Delta t})^3 = \frac{4\sqrt{2}}{5} A^{3/2} (\Delta t)^{3/2}.$$

Since

$$M_3(\Delta t) = O((\Delta t)^{3/2}) = o(\Delta t) \quad (\Delta t \rightarrow 0),$$

the uniform disk kernel indeed satisfies

$$\int_{\mathbb{R}^2} \|\mathbf{r}\|^3 P d\mathbf{r} = o(\Delta t),$$

which is consistent with the given requirement.

## C Proof of the Results in the Main Paper

### C.1 Proof of Proposition 4.1 (From [74])

We start from Assumption 4.1, the Boltzmann distribution at equilibrium in a conservative potential field  $\phi(\mathbf{x})$ :

$$\pi(\mathbf{x}) \propto e^{-\beta\phi(\mathbf{x})}$$

To preserve this equilibrium, we consider the transition probability with the detailed balance condition:

$$\pi(\mathbf{x})P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) = \pi(\mathbf{x} + \mathbf{r})P_\phi(\mathbf{x} + \mathbf{r}, -\mathbf{r}, \Delta t),$$

Substituting the Boltzmann distribution into the detailed balance condition gives

$$e^{-\beta\phi(\mathbf{x})}P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) = e^{-\beta\phi(\mathbf{x}+\mathbf{r})}P_\phi(\mathbf{x} + \mathbf{r}, -\mathbf{r}, \Delta t).$$

Rearranging, we obtain

$$P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) = e^{-\beta[\phi(\mathbf{x}+\mathbf{r})-\phi(\mathbf{x})]}P_\phi(\mathbf{x} + \mathbf{r}, -\mathbf{r}, \Delta t). \quad (30)$$

Here, we use isotropic core function  $P(\mathbf{r}, \Delta t) = P(\|\mathbf{r}\|, \Delta t)$  to construct  $P_\phi(\mathbf{x}, \mathbf{r}, \Delta t)$  that satisfies Equation (30). We define

$$P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) = \frac{1}{Z}e^{-\frac{\beta}{2}[\phi(\mathbf{x}+\mathbf{r})-\phi(\mathbf{x})]}P(\|\mathbf{r}\|, \Delta t),$$

where  $Z$  is the normalization constant that ensures  $\int_{\mathbb{R}^2} P_\phi(\mathbf{x}, \mathbf{r}, \Delta t) d\mathbf{r} = 1$ . It implies Proposition 4.1.

### C.2 Proof of Proposition 4.2

Denote the left-hand side of (5) as  $\tilde{D}$ . Since  $P(\|\mathbf{r}\|)$  is isotropic, we have

$$\tilde{D}_{ij} = \frac{1}{2Z\Delta t} \int_{\mathbb{R}^2} P(\|\mathbf{r}\|) r_i r_j d\mathbf{r}.$$

For  $i \neq j$ , consider the transformation  $r_i \mapsto -r_i$ . The product  $r_i r_j$  changes sign while the kernel remains unchanged, implying

$$\tilde{D}_{ij} = 0, \quad \text{for } i \neq j.$$

For diagonal elements  $\tilde{D}_{ii}$ , all components contribute equally due to directional symmetry, so  $\tilde{D}_{11} = \tilde{D}_{22}$ . Hence

$$\tilde{D} = D \cdot \mathbf{I}_2, \quad \text{and} \quad \mu = \beta D.$$

### C.3 Proof of Theorem 4.3

By Assumption 4.3, we have:

$$u_i(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} P_{\phi,i}(\mathbf{x} - \mathbf{r}, \mathbf{r}) u_i(\mathbf{x} - \mathbf{r}, t) d\mathbf{r} - \sum_j \xi_{ji} J_j(\mathbf{x}, t) \quad (31)$$

Obviously,  $J_j(\mathbf{x})$  can be written as:

$$\begin{aligned} -J_j(\mathbf{x}, t) &= \int_{\mathbb{R}^2} P_{\phi,j}(\mathbf{x} - \mathbf{r}, \mathbf{r}) u_j(\mathbf{x} - \mathbf{r}, t) - P_{\phi,j}(\mathbf{x}, \mathbf{r}) u_j(\mathbf{x}, t) d\mathbf{r} \\ &= \int_{\mathbb{R}^2} P_{\phi,j}(\mathbf{x} - \mathbf{r}, \mathbf{r}) u_j(\mathbf{x} - \mathbf{r}, t) d\mathbf{r} - u_j(\mathbf{x}, t) \end{aligned}$$

Similar to Section B.2, we have:

$$\begin{aligned} \int_{\mathbb{R}^2} P_{\phi,j}(\mathbf{x} - \mathbf{r}, \mathbf{r}) u_j(\mathbf{x} - \mathbf{r}, t) d\mathbf{r} &= u_j(\mathbf{x}, t + \Delta t) \\ &= u_j + \beta_j D_j \nabla_{\mathbf{x}} u_j \cdot \nabla_{\mathbf{x}} \phi \Delta t \\ &\quad + D_j (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_j) \Delta t + o(\Delta t) \end{aligned} \quad (32)$$

and

$$\begin{aligned} \int_{\mathbb{R}^2} P_{\phi,i}(\mathbf{x} - \mathbf{r}, \mathbf{r}) u_i(\mathbf{x} - \mathbf{r}, t) d\mathbf{r} &= u_i(\mathbf{x}, t + \Delta t) \\ &= u_i + \beta_i D_i \nabla_{\mathbf{x}} u_i \cdot \nabla_{\mathbf{x}} \phi \Delta t \\ &\quad + D_i (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_i) \Delta t + o(\Delta t) \end{aligned} \quad (33)$$

Therefore, we have:

$$-J_j(\mathbf{x}, t) = \beta_j D_j \nabla_{\mathbf{x}} u_j \cdot \nabla_{\mathbf{x}} \phi \Delta t + D_j (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_j) \Delta t + o(\Delta t) \quad (34)$$

Substitute Equation 33 and Equation 34 into Equation 31, we have:

$$\begin{aligned} u_i(\mathbf{x}, t + \Delta t) - u_i &= \beta_i D_i \nabla_{\mathbf{x}} u_i \cdot \nabla_{\mathbf{x}} \phi \Delta t + D_i (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_i) \Delta t \\ &\quad + \sum_j \xi_{ji} \beta_j D_j \nabla_{\mathbf{x}} u_j \cdot \nabla_{\mathbf{x}} \phi \Delta t \\ &\quad + \sum_j \xi_{ji} D_j (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_j) \Delta t \\ &\quad + o(\Delta t) \end{aligned} \quad (35)$$

Divide both sides by  $\Delta t$  and let  $\Delta t \rightarrow 0$ , we have:

$$\begin{aligned} \frac{\partial u_i}{\partial t} &= \beta_i D_i \nabla_{\mathbf{x}} u_i \cdot \nabla_{\mathbf{x}} \phi + D_i (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_i) \\ &\quad + \sum_j \xi_{ji} [\beta_j D_j \nabla_{\mathbf{x}} u_j \cdot \nabla_{\mathbf{x}} \phi + D_j (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_j)] \end{aligned} \quad (36)$$

Written in vector form, we have:

$$\frac{\partial \mathbf{u}}{\partial t} = (\xi^\top + \mathbf{I}_l) [\mathbf{D} \odot (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} \mathbf{u}) + \mathbf{D} \odot \beta \odot \{(\nabla_{\mathbf{x}} \mathbf{u})^\top \nabla_{\mathbf{x}} \phi\}], \quad (37)$$

where  $\nabla_{\mathbf{x}} \mathbf{u} = [\nabla_{\mathbf{x}} u_1, \dots, \nabla_{\mathbf{x}} u_l]$  and  $\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} \mathbf{u} = (\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_1, \dots, \nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{x}} u_l)^\top$ .

## D Justification for Model Assumptions

### Justification for Assumption 4.1 (Boltzmann Distribution).

Assumption 4.1 is reasonable because it mathematically formalizes a core principle of strategic behavior in conflict: actors are not random but instead probabilistically favor locations that maximize their advantage. In social science terms, a location’s strategic value—derived from factors like defensible terrain, resource availability, or the presence of sympathetic populations—creates an uneven “landscape” of opportunity. The Boltzmann distribution provides a robust way to model this, treating advantageous locations as low-potential-energy states that actors are more likely to occupy. This assumption directly translates the rational, goal-oriented behavior of conflict actors into a principled and well-established statistical physics framework.

### Justification for Assumption 4.2 (Local Transition Scaling).

Assumption 4.2 is reasonable because it embeds real-world physical and logistical constraints directly into the model. Conflict propagation is not instantaneous. The movement of personnel, resources, and influence is inherently limited by factors such as terrain, infrastructure, and communication speed. This assumption enforces these frictions by ensuring that influence spreads through a series of local, incremental steps over short time intervals, rather than through implausible long-range jumps. This aligns with extensive empirical evidence on the contagion patterns of insurgencies and protest movements, making the resulting diffusion process more physically realistic.

### Justification for Assumption 4.3 (Coupling of Transitions).

Assumption 4.3 is reasonable because it reflects the interdependent nature of conflict drivers, a central tenet of modern conflict theory. Violent conflict is rarely caused by a single factor but rather emerges from the convergence of multiple, correlated dynamics. For example, a flow of displaced people is often linked to shifts in resource control and the movement of armed groups. This assumption explicitly models this co-movement, allowing the framework to capture the multivariate reality of conflict escalation instead of treating each driver as an isolated phenomenon.

## E Data and Features

All inputs are harmonised to the *PRIO-GRID* lattice and monthly time steps for Africa, 1990–2020. Table 4 groups the resulting variables into twelve thematic blocks. Dynamic layers—conflict and protest events from UCDP GED, ACLED, and SCAD—are converted to cell-month counts, lags, moving sums, and distance-weighted kernels to retain short-run memory. Semi-dynamic socio-economic indicators (night-time lights, infant mortality, gross cell product) are updated annually or as new satellite rasters become available. Truly static layers (terrain, distance to borders, mineral deposits, ethnic power relations) enter unchanged. Climate stressors are captured with multi-scale SPEI drought indices and agricultural-season dummies, while land-cover fractions, population density, and travel-time maps provide background capacity and exposure.

In our experiments, the features belonging to the first two categories in Tabel 4 (i.e., Conflict events and Protest events) are regarded as dynamic and jointly encoded via an LSTM-GCN framework. The remaining features are treated as static and encoded independently using a GCN.

## F Details of ConflictNet

### F.1 Settings

To construct the graph  $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{T})$ , we treat each of the  $N$  grid cells as a node within  $\mathcal{V}$ . Nodes are interconnected based on geographic proximity, with each node linked to its nearest 24 neighbors, forming the edge set  $\mathcal{E}$ . This design captures the spatial interrelations essential for conflict prediction. The time dimension is represented by  $\mathcal{T}$ , encompassing  $T$  discrete time points across our study period.

We use data from the past  $p = 12$  months  $[t - 12, t]$  to predict the probability of conflict after  $h$  months,  $h = 1, 3, 6, 9$ . For the practical application of Partial Differential Equations (PDEs) in modeling conflict spread, we discretize temporal data by dividing each month into five equal segments. This approach enhances our model’s temporal resolution, enabling a more detailed analysis of conflict dynamics over time. All the parameters are initialized with Glorot initialization, and we use Adam optimizer. We performed 3-fold cross-validation with the test set starting in Jan 2013, Jan 2016, and Jan 2019, and 24 months in length, the validation set is the 12 months before the test set starting point, and the rest as the training set. To rigorously evaluate our model’s predictive performance, we employ ROC-AUC and PR-AUC metrics for assessing overall accuracy and sensitivity to positive cases, respectively. Additionally, the  $F_1$ -Score is utilized to balance precision and recall, providing a holistic view of model efficacy in predicting conflict occurrences. All programs are implemented using Python 3.8 and PyTorch 1.11.0 with CUDA 11.8 in the ubuntu20.04 system with RTX4090 NVIDIA GPU devices.

### F.2 De-bias Module

We apply spatio-temporal debiasing to the features extracted by the encoder, following the causal representation learning framework proposed in [20]. The goal is to reduce confounding introduced by temporally and spatially specific factors, so that the learned features reflect causal drivers of conflict rather than spurious associations.

Concretely, given a biased feature representation  $f_b \in \mathbb{R}^{d_m}$  output by the encoder, we compute a debiased feature  $f_{db} \in \mathbb{R}^{d_m}$  using a linear combination of the original feature and a context-aware adjustment term:

$$f_{db} \approx W_x f_b + W_s r_b$$

where  $W_x$  and  $W_s \in \mathbb{R}^{d_m \times d_m}$  are learnable projection matrices, and  $r_b$  is a debiasing factor constructed from category-level reference features.

To obtain  $r_b$ , we first define a set of reference features  $\{s_i\}$ , where each  $s_i \in \mathbb{R}^{d_m}$  is the average encoder output over all samples in category  $i$ . Let  $P(i)$  denote the empirical frequency of category  $i$ , and  $\alpha_i$  be the attention weight assigned to category  $i$ . Then the debiasing factor is computed as a weighted sum:

$$r_b = \sum_{i=1} \alpha_i s_i P(i)$$

The weights  $\alpha_i$  are computed using a scaled dot-product attention mechanism:

$$\alpha_i = \text{softmax} \left( \frac{(W_q f_b)^\top (W_k s_i)}{\sqrt{d_m}} \right)$$



**Table 4: Feature categories used in the model. Each row summarises a group of variables.**

Category	Typical features / transforms	Main data sources
Conflict events	UCDP fatality counts, event lags, moving sums, spatial / spatio-temporal kernels, exponential decay weights.	UCDP GED
Protest events	Cell-month counts and fatalities for riots / demonstrations, identical lag and kernel transforms as conflict events.	ACLEd; SCAD
Population density	Log of total population per cell.	GPW v4
Night-time lights	Mean VIIRS / DMSP intensity (calibrated).	NOAA VIIRS
Socio-economic status	Infant mortality, gross cell product, urban share, travel-time to city.	SEDAC; WB; Uchida-Nelson (2009)
Land-cover composition	Forest, cropland, pasture, shrub, urban fractions.	MODIS; PRIO-GRID v2
Climate / drought	SPEI indices, drought frequency, anomalies, 1–48 month lags.	SPEI Monitor
Agricultural seasonality	Growing-season dummy, harvest volume, irrigated vs. rain-fed share.	MIRCA2000; MapSPAM
Physical geography	Mountain share, distance to coast / capital / border, travel-time to nearest city.	UNEP-WCMC; CShapes
Natural resources	Distance to diamond, petroleum, and other mineral deposits.	PRIO-GRID resource layers
Ethnic power relations	Count of excluded or powerless groups within cell.	EPR dataset
Derived spatial kernels	Pure spatial lags, $k$ -nearest event distances, time-since-last-event decays.	Derived from GED / ACLED

where  $W_q, W_k \in \mathbb{R}^{d_m \times d_m}$  are learnable query and key matrices, and  $d_m$  is the feature dimension. This mechanism encourages the model to selectively align each instance with prototypical representations from relevant categories, thus reducing bias introduced by category imbalance or context-specific artifacts.

The categories used for debiasing are constructed along both temporal and spatial dimensions. Temporally, we divide the data into three historical periods, each corresponding to a distinct global conflict environment, as suggested in [32, 33]. This division reflects major shifts in the global security landscape. Specifically, we define

the three periods as: before September 2008, from September 2008 to December 2010, and after December 2010. Within each period, conflict triggers are more homogeneous, allowing the model to better isolate time-invariant factors and reduce temporal confounding in feature representations. Spatially, we group grid cells at the country level, following the assumption that within-country conflict patterns tend to share common underlying drivers (e.g., regime type, socio-political structure), as supported by the PRIO-Grid framework [78]. This yields a coarse-grained spatial categorization that avoids overfitting to fine-grained local idiosyncrasies.

G More Experimental Results

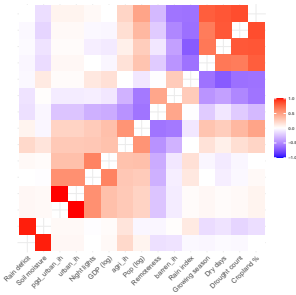


Figure 5: Correlation matrix among selected features, highlighting cross-category coupling between natural and social conditions.

Table 5: Ablation of hyperparameters of the loss function.

$\gamma_1$	$\gamma_2$	$\gamma_3$	$\gamma_4$	PR-AUC	ROC-AUC	F1
100	0.2	100	0.5	56.7	97	53.7
100	0.2	10	0.05	56.6	97	53.9
100	0.2	1000	5	56.7	97	53.4

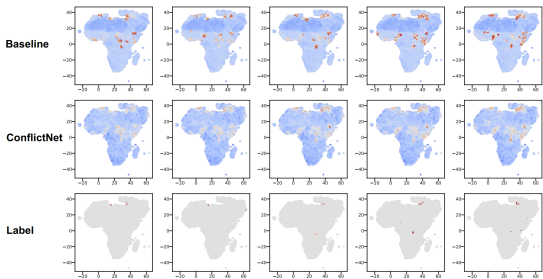


Figure 6: Thermodynamic diagrams depicting the baseline method or ConflictNet conflict factors, along with corresponding labels, in Africa every three months from June 2012 to June 2013.

H Broader Societal Impacts

ConflictNet has the potential to support positive societal outcomes by enhancing early warning systems and informing targeted peace-keeping or humanitarian responses. By modeling the spatial and temporal dynamics of armed conflict, the framework may assist policymakers and international organizations in anticipating escalation, pre-positioning aid, or identifying high-risk zones for intervention. The model’s interpretability also supports academic efforts to better understand the mechanisms of conflict diffusion.

We do not foresee direct negative societal impacts from this work. The model operates at an aggregate, subnational level and does not rely on sensitive individual-level data, reducing risks of misuse or harm. Nonetheless, responsible deployment of conflict forecasting tools—particularly in politically sensitive settings—requires transparency, local engagement, and caution in interpreting probabilistic outputs.