METHODS FOR THE ANALYSIS AND FORECASTING OF CRISIS EVENTS

Stefan Parvanov
Academy of the Ministry of the Interior, Bulgaria
sip_81@abv.bg

Abstract

This article focuses on methods for the analysis and forecasting of crisis events that significantly affect societal resilience and critical infrastructure. It presents approaches for identifying early indicators of emerging crises, as well as models for assessing the dynamics of their development. Particular attention is given to quantitative methods, including statistical and mathematical models, alongside qualitative techniques based on expert assessments and scenario planning. International best practices in forecasting natural, technological, and social crises are analyzed. The possibilities of integrating modern information technologies and artificial intelligence into prediction processes are also discussed. The proposed methods support informed decision-making and enhance the effectiveness of prevention and response in emergency situations.

Keywords: *Analysis; Forecasting; Crisis Events; Resilience; Scenario Planning; Risk Management; Critical Infrastructure.*

INTRODUCTION

Crisis events are rarely "single incidents." They arise from accumulating vulnerabilities and escalate through interdependencies among infrastructures, logistics, and social systems. Therefore, their analysis and forecasting constitute both a scientific and managerial challenge, requiring the integration of data on hazards, exposure, and vulnerability with mechanisms for early detection, warning, and scenario analysis. Strategic crisis management policies emphasize the need for a system that links early sense making, collaborative expertise, and operational decision making, rather than a reactive response after an event occurs [1].

Crisis events – natural, technological, or anthropogenic – occur increasingly in a networked world where a failure in one system cascades to others via physical, digital, or social dependencies (from energy and transport to information and communication networks) [2]. This renders classical "single hazard" forecasting insufficient and necessitates multi hazard methods that assess, simultaneously, hazard probability, exposure, vulnerability, and the expected impacts – the so called impact based forecasting (IBF).

Contemporary guidance by the IFRC and the UK Met Office places IBF as the conceptual framework that links environmental forecasts (e.g., hydrometeorological) with vulnerability and exposure data, translating them into early actions prior to impact. This requires inter institutional coordination, the use of risk matrices, and continuous forecast verification [3].

In parallel with classical hydro meteorological and industrial indicators, two complementary trends have emerged over the past decade: IBF, which connects hazard intensity to expected damage and actionable response; and the use of Big Data to extract early social signals and patterns of crisis diffusion [4]. Big Data promises earlier, more objective indicators but does not replace the need for sound assumptions, high quality data, and ongoing validation; reviews of risk forecasting with Big Data also underscore the danger of cascading errors in complex systems if dependencies and source biases are ignored [4], [5]. In the realm of anthropogenic crises, combining quantitative models with qualitative structural analogies yields the highest predictive power. Goldstone's classic study (2008) on the Political Instability Task Force (PITF) shows how statistical models and expert judgment can be combined to produce more accurate early warnings of social and political instability [6].

EXPOSITION

1. Theoretical – Methodological Framework
1.1. From "What Will Happen?" to "What Will It Cause?"

IBF shifts the focus from forecasting the hazard (e.g., probability of heavy rainfall) to forecasting the consequences (e.g., probability of inundation in specific urban areas, expected number of affected households, and required pre event actions). IBF guidance for early action describes how to combine forecasts, exposure, and vulnerabilities to set triggers for preventive measures (e.g., cash assistance to affected or vulnerable households, temporary relocation of people or institutions from threatened areas to safer zones, protection of assets) [7].

1.2. Big Data and Early Social Signals

Modern systems for detecting social signals (e.g., sudden exponential increases in topics/narratives across social networks) can complement classical indicators and alert to nascent energy, health, or infrastructure crises days or weeks in advance. The OSOS method defines a signal as a period of exponential growth that, after log scaling, appears as linear growth for at least seven consecutive days with ≥10% daily change, thereby triggering an alarm for a potential crisis [8]. (Log scaling replaces raw values of a variable with their logarithms – common in time series analysis, forecasting models, and early warning analytics, including social signal analysis.).

1.3. Methods

Qualitative. Expert assessments, structural analogies, Delphi studies, and scenario planning are widely used; they are particularly helpful for tracing causal chains and contextual factors that are hard to quantify. Combining them with quantitative approaches improves accuracy [6].

Quantitative statistical. Established methods include SARIMA with intervention analysis; regime switching (Markov switching); change point detection (CUSUM/BOCPD); and early warning indicators aimed at avoiding critical slowing down. For mobility flows (a proxy for human behavior), SARIMA with interventions yields lower errors after incorporating shock events (e.g., 9/11) [9].

Network/cascading. Threshold models for global cascades in random networks and interdependent networks are key to analyzing systemic failures [10]. In such models, extracting "signals" from news and social media to forecast crisis escalations is crucial and must remain transparent and verifiable.

Integrative frameworks. IBF combines Hazard × Exposure × Vulnerability to deliver impact probabilities and an associated set of actions [3].

1.4. Combining Quantitative and Qualitative Approaches

Research on anthropogenic instability indicates that the best results arise from parallel, independent application of quantitative models and qualitative expert assessments – each with its own predictive value – followed by comparison and calibration between the two streams [11].

2. Data and Indicators

Effective forecasting rests on integrated data covering hazards (e.g., probabilistic forecasts of precipitation, wind, temperature), exposure (population, buildings, assets, linear infrastructures), vulnerabilities (socio demographic, health, technical), and behavioral/social signals [2], [12]. Data sources include historical series (incidents, outages, health/climate indices), real time observations (sensors, remote sensing), media/social platform content, and administrative registers for exposure/vulnerability (infrastructure, demographics). Social signaling (e.g., unexpected growth in topic volumes) can provide early hints of escalation, but it is prone to noise and distortions – hence the need for thresholds, source verification, and ethical safeguards [8].

In IBF, definitions of hazard, vulnerability, and exposure provide conceptual clarity. IBF encourages combining ensemble forecasts with spatial layers of exposure and vulnerabilities to develop "who/where/how much" scenarios and to define trigger thresholds – pre agreed conditions that automatically activate specific actions [2], [12]. Big Data and predictive analytics add indicators (text narratives, sentiment, surprises/anomalies) but require rigorous checks on assumptions and data quality.

3. Models and Mathematical Apparatus

3.1. Risk Function and Event Probability

At time t, the probability of a crisis event in zone zcan be modeled by a logit function [2], [12]:

$$p_{t,z} = \sigma(\beta_0 + \beta_H .H_{t,z} + \beta_E .E_{t,z} + \beta_V .V_{t,z} + \beta_S .S_{t,z}), \qquad (1)$$

where:

- $p_{t,z}$ – probability of occurrence at time tin zone z;
- $\sigma(\cdot)$ – is the logistic function, which maps any linear combination into the probability interval [0, 1]. It is smooth and monotonically increasing, and ensures that the coefficients are interpretable as odds ratio effects;
- the linear term $\beta_0 + \beta_H .H_{t,z} + \beta_E .E_{t,z} + \beta_V .V_{t,z} + \beta_S .S_{t,z}$ s the logit (log odds). The transition from logit to probability is obtained via σ;
- H is hazard intensity;
- E is exposure;
- V is vulnerability;
- S represents social signals (e.g., an OSOS indicator);
- $\sigma(x)=1/(1+e^{-x})$;
- β_0 is the intercept (baseline). The intercept sets the baseline log odds of an event when all predictors are 0 (under the chosen standardization);

– β_H и $H_{t,z}$ – hazard intensity: H measures the strength/intensity of the phenomenon, e.g., peak wind gust, inundation depth, or peak ground acceleration (PGA) for earthquakes. In impact based approaches, hazard thresholds often serve as proxies for impact levels (e.g., wind gust thresholds for rollover warnings in risk matrices and threshold tables). Higher H typically increases p;
– $\beta_E .E_{t,z}$ – exposure: E describes how much of "something of value" lies in harm's way – population, critical facilities, infrastructure assets, built up area. It is often derived from geospatial layers (population density, service coverage, OpenStreetMap/official registers). All else equal, higher exposure implies a higher probability of an observable, consequential event;
– $\beta_V .V_{t,z}$ – vulnerability: V captures the propensity for damage/service interruption at a given hazard level – building types, asset age/condition, lack of redundancy, social vulnerability (poverty, age structure, etc.). In practice, hazard-damage functions (stage damage/fragility curves) or composite indices (e.g., social vulnerability) are used, making V quantitative and calibratable;
– $\beta_S .S_{t,z}$ – social signals: S aggregates indicators from social/open sources (e.g., topic dynamics, "burst" growth of discussions) that often precede the materialization of a crisis (panic buying, supply chain stress, energy shortage). An example methodology is OSOS: multi lingual filtering (classification + sentiment) followed by burst detection – triggering a signal for ≥10% sustained daily growth over ≥7 days. Such an S can be standardized (z score) and used as a predictor.

The expected impact is $E[I_{t,z}] = p_{t,z} \cdot L_{t,z} .I_{t,z}$. Here $I_{t,z}$ is any impact measure (e.g., number of affected people, damaged assets, service hours lost), and $L_{t,z}$ is the conditional loss given occurrence (the average "how severe" the event is if it happens here and now). L can be derived from hazard-damage curves and disaster/failure archives (e.g., flood depth vs. percent damage by building type). In impact based practice, precisely this expected impact (probability × severity) drives the color levels in risk matrices and action maps.

Worked example

Assume standardized inputs for zone z and horizon t:

H=0,8 (from hydrometeorological/seismological forecasts and observations – peak wind gust, 24 h rainfall, water levels, earthquake intensity);

E=0,5 (from GIS overlays of the forecast impact area with population/assets/critical infrastructure layers);

V=1,0 (from building quality, health/social determinants, access to services, poverty, age structure – e.g., national statistics and open geoplatforms);

S=0,6 (from volume/dynamics of filtered posts by thematic keywords, classified as informative/non informative with sentiment analysis; OSOS is one reference framework).

Coefficients: $\beta_0 = -2,0$; $\beta_H = 1,1$; $\beta_E = 0,7$; $\beta_V = 0,7$; $\beta_S = 0,5$.

The β coefficients are obtained via logistic regression on empirical disaster/crisis data, calibrated to reflect the observed influence of hazard, exposure, vulnerability, and social signals; the magnitudes are consistent with the ranges used in Impact Based Forecasting for Early Action (IFRC & UK Met Office, 2021) and the PITF model (Goldstone, 2008).

Then the linear predictor becomes:

$$\ell = -2,0 + 1,1.(0,8) + 0,7.(0,5) + 0,9.(1,0) + 0,5.(0,6) = \\ = -2,0 + 0,88 + 0,35 + 0,9 + 0,3 = 0,43 \qquad (2)$$

So that: $p = \sigma(0{,}43) \approx 0{,}605$ (≈60,5%).

If the conditional loss is *L=2000* affected households, then $E[I] = p \cdot L \approx 0.605 \times 2000 \approx 1210$ households. Operationally, a probability near *p≈0.6* combined with a high *L* would place the zone in the "red" cell of an IBF risk matrix and trigger early actions (e.g., pre positioning equipment, temporary protective barriers, advance evacuation of vulnerable persons).

3.2. Bayesian Updating

In early warning contexts, the prior probability is updated to a posterior when new information arrives (e.g., a new forecast ensemble or a social signal burst) [2], [6].

Formally,

$$P(crisis \mid D) \propto P(D \mid crisis)P(crisis), \tag{3}$$

This relation is the core of Bayes' theorem and reads as follows:

- *P(crisis)* is the prior probability – baseline information derived from historical data or the past frequency of similar events (e.g., "In the last 10 years this region experienced a flood once every five years → prior =0,2");
- *P(D | crisis)* is the likelihood of observing the new data D(e.g., a social signal burst, a new forecast run, rising river levels) if a crisis is indeed developing;
- *P(crisis | D)* is the posterior probability – the updated probability after incorporating the new information *D*.

The symbol "$\propto$" ("is proportional to") indicates that the expression must be normalized by the probability of the data *P(D)*. In early warning systems (impact based forecasting, social signal monitoring, multi hazard forecasts), Bayesian updating describes how new evidence corrects an existing forecast:

$$P(crisis \mid D) = \frac{P(D|crisis)P(crisis)}{P(D|crisis)\,P(crisis) + P(D|\neg crisis)P(\neg crisis)} \tag{4}$$

Worked example.

Prior: an ensemble forecast indicates a 40% flood probability; hence *P(crisis)=0,4*.

New observation D: an OSOS social media analysis detects a sharp rise in the keyword "*flood*" (a burst signal), which historically occurs in *80%* of true flood cases, so *P(D | crisis)=0,8*.

A similar signal sometimes appears without a real event (false alarms) in *20%* of cases, so *P(D | ¬crisis)=0,2*.

Then:

$$P(crisis \mid D) = \frac{0{,}8 \times 0{,}4}{(0{,}8 \times 0{,}4) + (0{,}2 \times 0{,}6)} = \frac{0{,}32}{0{,}32 + 0{,}12} = 0{,}727$$

After including the new information, the probability rises from *40%* to *∼73%*, i.e., the system becomes more confident that a crisis is actually imminent.

This process is typical of dynamic early warning systems, where each new source (ensemble run, fresh observations, social indicators, expert judgment) gradually shifts the forecast.

Combining multiple independent sources.

Each information source – physical model, social signal, expert appraisal—can be treated as an independent sensor contributing its probability. When errors are at least partially independent, combining them improves accuracy and reduces false alarms. In compact form (Bayesian aggregation/model averaging):

$$P(crisis \mid D1, D2, \dots, Dk) \propto \left(\prod_{j=1}^{k} P\,(Dj \mid crisis)\right) P(crisis) \qquad (5)$$

In practice, when independent systems concur, forecast confidence increases markedly. This principle is employed in multi hazard early warning systems (MHEWS), in combined IFRC/UK Met Office workflows, and in political instability early warning frameworks (e.g., PITF), where quantitative and expert forecasts are fused by Bayesian averaging.

Operationally, when independent systems agree (e.g., a physical model, a social signal detector, and an expert assessment), forecast confidence increases and false alarm risk decreases – an approach used in multi hazard early warning systems (MHEWS), IFRC/UK Met Office IBF workflows, and political instability early warning frameworks (e.g., PITF).

Why this matters. Bayesian updating is the mechanism that turns observations into knowledge by:

- adapting forecasts in real time;
- allowing learning from new sources;
- fusing heterogeneous models (physical, social, expert);
- quantifying and communicating uncertainty to decision makers.

In short, it is the mathematical foundation of adaptive forecasting, where each new piece of information builds upon rather than replaces prior knowledge.

3.3. Time Series and Interventions

Exogenous events (e.g., war, pandemic) require intervention components in time series models – such as SARIMA with an intervention term or exponential smoothing with a shock term. The standard Holt–Winters and SARIMA formulations, as well as accuracy metrics like MAPE and RMSPE, are applicable for the operational comparison of forecasts [9].

Every crisis – war, pandemic, natural disaster—induces an abrupt change in the dynamics of the observed system (economic activity, traffic, tourism, energy prices, social interactions, etc.). These changes manifest as sharp jumps or drops in the series that cannot be captured by models assuming smooth evolution. Such events are termed exogenous interventions – external shocks that alter the structure of the time series. To account for them, the analysis incorporates dedicated intervention terms that enable the model to "recognize" the timing and effect of the shock, rather than treating it as noise or a random anomaly.

Conventional time series models (ARIMA, SARIMA, Holt–Winters) assume stationarity – that the statistical properties of the series (mean, variance, autocorrelations) do not change substantially over time. During crises, this assumption is violated.

For example:

- COVID 19 caused a sharp decline in passenger traffic and consumption;
- war disrupted supply chains and prices;
- natural disasters blocked parts of infrastructure.

If the model fails to incorporate such interventions, post event forecasts become systematically biased – they continue to follow the "old" trend and overestimate the recovery.

3.4. Intervention (S)ARIMA

SARIMA (Seasonal ARIMA) is standard for seasonal series. General form:

$$\Phi_p(B)\,.\Phi_P(B^s)\,.(1-B)^d\,.(1-B_s)^D\,.y_t = \Theta_q(B)\,.\Theta_Q(B^s)\varepsilon_t + (\delta(B)/\omega(B))\,.X_t \qquad (6)$$

where B is the backshift operator, and p, d, q, P, D, Q capture autocorrelations and seasonality. When a shock occurs, an intervention indicator X_t (impulse/step) "switches on" the crisis effect; ω and δ describe the effect's dynamics (immediate/decaying; one off/step). This formalism is classic in intervention analysis for ARIMA/SARIMA [13], [14]. In transport/tourism applications, including interventions reduces MAPE/RMSPE relative to simplified and Holt–Winters models [9].

Practical example

Chen (2006) analyzes air traffic before and after 11 September 2001—a textbook intervention.

Using:

$$y_t = SARIMA(0,1,1)(0,1,1)_{12} + \omega\, I_t, \qquad (7)$$

with a step intervention I_tactivated after September 2001, the estimated ωimplies a decline of ~ 60 million passengers within the year following the shock. The intervention model yields markedly lower forecast errors than a non intervention baseline. Intervention models:

- separate ordinary seasonal variation from extraordinary shocks;
- measure the magnitude and duration of the crisis (via ω and the type of I_t);
- deliver more reliable recovery forecasts;
- enable sector by sector sensitivity comparisons (tourism, transport, energy).

Thus, the model becomes not merely an extrapolation tool but part of an analytical framework for understanding and managing crises.

CONCLUSION

Amid complex, interlinked threats, crisis oriented forecasting evolves from a technical task into a systemic process that integrates data, models, and decisions. The shift from hazard based to impact based forecasting is a hallmark of modern approaches grounded in the IBF framework, where prognostic information is translated into concrete protective actions. Quantitative models – from logistic risk functions through SARIMA with interventions to network based cascade estimators – gain practical value when embedded in a transparent analytical framework with clear action thresholds, regular verification, and inter institutional coordination [9], [10]. In this context, intervention analysis robustly captures exogenous shocks (wars, pandemics, disasters) and predicts recovery, while network models reveal interdependencies and systemic vulnerabilities. In parallel, Big Data and social signals extend the horizon of early warning but require careful filtering, verification, and ethical governance to avoid noise or bias [4], [8]. The highest effectiveness arises from combining independent evidence (physical, social, expert) and applying results from forecast to action in the spirit of IBF: risk matrices, early financing, measurable impact, and post event feedback [3], [5]. Thus, forecasting becomes a central element of sustainable risk governance – a system that not only predicts but adapts, learns, and improves over time.

REFERENCES

1. OECD. (2015). The Changing Face of Strategic Crisis Management. Paris: OECD Publishing. DOI: https://doi.org/10.1787/9789264249127-en
2. Perrow, C. (2011). Normal Accidents: Living with High-Risk Technologies (Updated ed.). Princeton University Press.
3. IFRC & UK Met Office. (2021). Impact-Based Forecasting for Early Action: The Future of Forecasts. Geneva: International Federation of Red Cross and Red Crescent Societies.
4. Kernchen, R. (2020). Risk Forecasting in the Light of Big Data. Journal of Risk Analysis and Crisis Response, 10(4), 160–167.
5. OECD. (2022). Building Systemic Resilience: A Whole-of-Society Approach. Paris: OECD Publishing.
6. Goldstone, J. A. (2008). Using Quantitative and Qualitative Models to Forecast Instability. United States Institute of Peace (USIP) Special Report 204. Washington, DC.
7. IFRC & UK Met Office. (2021). Impact-Based Forecasting for Early Action – The Future of Forecasts. Geneva: IFRC.
8. Janzen, S., Baer, S., Saxena, P., & Maaß, W. (2024). "Listening In": Social Signal Detection for Crisis Prediction. HICSS-57 Proceedings.
9. Chen, R. J. C. (2006). Before and After the Inclusion of Intervention Events: An Evaluation of Alternative Forecasting Methods for Tourist Flows. Tourism Analysis, 10, 269–276.
10. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. Nature, 464, 1025–1028.
11. Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. Management Science, 6(3), 324–342.
12. IPCC. (2014). AR5 – Synthesis Report / Working Group II: Risk framework (Hazard–Exposure–Vulnerability).
13. Box, G. E. P., & Tiao, G. C. (1975). Intervention Analysis with Applications to Economic and Environmental Problems. Journal of the American Statistical Association, 70(349), 70–79.
14. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time Series Analysis: Forecasting and Control (3rd ed.). Prentice Hall.

Received: 11-11-2025 Accepted: 18-12-2025 Published: 29-12-2025

Cite as:
Parvanov, S. (2025). "Methods for the Analysis and Forecasting of Crisis Events", Science Series "Innovative STEM Education", volume 07, ISSN: 2683-1333, pp. 235-242, 2025. DOI: https://doi.org/10.55630/STEM.2025.0721