MOMENTKA: Použití strukturálních podmínek a QCA k odhalení zemí náchylných k eskalaci konfliktu

TAKING A SNAPSHOT: Discovering Escalation Prone Countries Through Structural Conditions and QCA

Kateřina Fridrichová[a], *Adriana Ilavská*[b]

Abstrakt

Tento článek představuje metodu předvídání eskalace konfliktů na základě dlouhodobých, strukturálních podmínek z otevřených zdrojů za použití kvalitativní komparativní analýzy (QCA). Kombinací těchto dvou prvků vzniká model a metoda predikce, která je v efektu predikce srovnatelná s dalšími, daleko sofistikovanějšími a lépe financovanými systémy a nabízí citlivější odhad eskalace.

Abstract

This paper introduces a method of conflict escalation prediction based on long-term structural conditions, drawn from open sources, and QCA (Qualitative Comparative analysis). The combination of the two components offers a model of prediction, which is comparable to other more sophisticated and resourced systems. Moreover, it offers a more sensitive estimation of an escalation.

Klíčová slova

Předvídání; eskalace konfliktů; strukturální podmínky; včasné varování; csQCA.

Keywords

Prediction; Conflict Escalation; Structural Conditions; Early Warning; csQCA.

---

[a] Department of International Relations and European Studies, Masaryk University. Brno, Czech Republic. fridrichova@fss.muni.cz. Researcher ID: U-3593-2019.
[b] Department of International Relations and European Studies, Masaryk University. Brno, Czech Republic. ilavska@mail.muni.cz. Researcher ID: AAG-6712-2020.

INTRODUCTION

"Happy families are all alike; every unhappy family is unhappy in its own way."

Leo Tolstoy

The so-called Anna Karenina principle points out a critical feature of the social world. There are few ways how to succeed in any given endeavour; the configurations of measures leading to success are few and relatively narrow with minimal deficiencies. However, there are myriads of ways how to fail. There has been a lot of research on both state success and state failure. That is why the search for causes of conflict has been expanding, and it includes dozens of indicators that given data and modelling can predict the conflict outbreak or conflict escalation. Nevertheless, there is no single one that explains the totality of the universe of conflict escalations.

The conflict research is being done in two different but not entirely mutually exclusive modes, because these inform each other. The academic research is interested in methodology, theories, and data collection, evaluation and analysis as a part of effort seeking an explanation of the phenomena. The other mode is applied research – various early warning systems, which do all of the previously mentioned activities to provide policy guidance and help to shape preventive measures. And, of course, there is also the effort of intelligence agencies to inform policymakers about threats to their countries.

The research in this paper falls under the latter, applied category, but draws heavily on the first. The research does not aim at discovering new insight into conflict research but creates a methodology of conflict escalation prediction and identification of threats emanating from such an event for the use of the Czech Ministry of Interior. Moreover, the model needs to be affordable for a country of the size of the Czech Republic and relatively easy to update. A research method able to provide formalised steps to analyse open-source data provided by other countries or international organisations fulfils both requirements mentioned above.

For the purpose of this article, the system is specified for regions currently considered most important for the internal security of the Czech Republic; it is not developed as a global warning system. The fundamental methodological approach is variable and could be adjusted to different regional groupings; specificity is achieved by including countries of the broader neighbourhood of the Czech Republic (Western Europe, post-communist and post-Soviet countries, Middle East and Northern Africa) into the model building. Their list is provided in Appendix 2.

The innovation presented in this paper is a combination of two components previously not used together in the conflict prediction research: long-term structural data and QCA (Qualitative Comparative Analysis). These two components are naturally complementary; the creators of QCA originally devised it as a "macro-qualitative". Its role is to analyse a presence or an absence of features at a "macro" level of large structures like political systems, states or classes.[1] However, there has not been any significant attempt to try to

---

[1] BERG-SCHLOSSER, Dirk - Lasse CRONQVIST. Macro-Quantitative vs. Macro-Qualitative Methods in the Social Sciences - An Example from Empirical Democratic Theory Employing New Software.

use it for conflict escalation prediction. In this respect, our proposed method is exploratory.

The research question for building the model is under what conditions prevailing in the society does conflict escalate? The model takes a form of different paths (configurations of conditions) that had captured a country in a moment of possible conflict escalation in the past. These paths serve as a "snapshot" of society prone to a conflict escalation in the horizon of 1–3 years. As a next step, this snapshot is then applied to a current set of data and generates a list of countries emanating possible threats. We offer a list of countries generated for years 2019 and 2020 in comparison to other systems at the end of the paper (Table 3).

METHODOLOGY AND DATA

Structural data

Structural data are long-term conditions represented by indicators that change only slowly and describe the whole populations over longer-time periods (such as GDP, GINI index, annual water stress indicators, and population size). These data were the basis of the first generation of early warning systems and many of the conflict prediction models, as demonstrated, for example, in the work of Paul Collier.[2] The nature of the data makes them useful for some tasks and unsuitable for others. Their long-term, almost invariant nature does not capture the dynamics of the conflict escalation for early warning purposes and the indicators are connected to the outcome only indirectly,[3] since they are proxy measures.[4] These then become better or worse predictors of a conflict; the so-called greed model (represented by indicators such as per capita income, economic growth, and educational attainment) outperforms grievance variables (such as political victimisation, access of power, economic inequality). However, the question remains whether this represents the real dynamics of the conflict escalation or captures a state of the country/society at the moment of escalation. If there are correlations of these variables, their potency changes in time and sometimes fluctuates dramatically,[5] although most of

Historical Social Research / Historische Sozialforschung [online]. 2005, Vol. 30, No. 4 (114) [cit. 30.06.2020]. Available from: https://www.jstor.org/stable/20762081; BERG-SCHLOSSER, Dirk - Sven QUENTER. Macro-Quantitative vs Macro-Qualitative Methods in Political Science - Advantages and Disadvantages of Comparative Procedures using the Welfare-State Theory as an Example. *Historical Social Research / Historische Sozialforschung* [online]. 1996, Vol. 21, No. 1 (77) [cit. 12.09.2016]. Available from: http://www.jstor.org/stable/20756014

[2] COLLIER, P. Greed and Grievance in Civil War. *Oxford Economic Papers*. 2004, Vol. 56, No. 4. DOI: 10.1093/oep/gpf064

[3] O'BANNON, Brett R. 'Monitoring the Frog' in Africa: Conflict Early Warning with Structural Data. *Global Responsibility to Protect*. 2012, Vol. 4, No. 4. DOI: 10.1163/1875984X-00404004

[4] DAVENPORT, Christian - Mark I. LICHBACH - David A. ARMSTRONG. *Conflict Escalation and the Origins of Civil War*. University of Maryland, 2005.

[5] BRAUMOELLER, Bear F. *Only the Dead: The Persistence of War in the Modern Age*. New York, NY: Oxford University Press, 2019, p. 139.

the research presupposes their impact to be stable[6] and linear[7], e.g., joint democracy as a cause of conflict since 1820. In this sense, the configurative nature of the research does not depend on best predicting proxies, rather on a coverage of wider variety of social life, capturing the atmosphere in society which could contribute to conflict escalation. Chosen variables represent the main areas of indicators, which were often suggested by previous research, and they underline the focus on dynamics of societal changes.

Early warning systems of newer generations supplement the long-term conditions by short-term variables sometimes referred to as "triggers", "accelerators", or "catalysts" when trying to monitor and evaluate the situation as it is developing on the ground. Where underlying societal conditions of malcontent, bad economic performance and state repression were structural conditions in the society, the self-immolation of a small businessman after his run-in with the police became a trigger of an uprising in the case of Tunisia. However, triggers are difficult to define; their ontological and epistemological status differs from the long-term ones. They are a contested notion in sociology and also in the study of revolutions.[8] However, it is not just the sociologists, statisticians Blyth and Taleb point out the fact that one can hardly blame the last grain of sand for causing the pile to crumble, instead of the structure of the pile.[9] Analogically, the notion of triggers appears in the conflict research. But one thing is sure: they are impossible to predict[10]. As a consequence, the systems of early warning, which need short-term input of real-time reporting on the ground, media monitoring and other similar techniques to capture the already ongoing events as they escalate, they are not modelling some kind of trigger events. The early warning systems then integrate these observations as an additional input of data.

But even those events remain context-dependent – the material needs to be flammable for a spark to set it alight and countries need to be ripe for a conflict escalation for some triggering event to set it off. A draught is a cyclical phenomenon in the Middle Eastern region, but only under specific government policies and the inability of compensating for its impacts seems to be what sparked the revolution in Syria:

> *"The humanitarian crisis that followed the 2006-10 drought can thus be seen as the culmination of 50 years of sustained mismanagement of water and land*

---

[6] Ibid., pp. 136–137.

[7] MCLOUGHLIN, Stephen. From Reaction to Resilience in Mass Atrocity Prevention: An Analysis of the 2013 UN Report "The Responsibility to Protect: State Responsibility and Prevention". *Global Governance* [online]. 2016, Vol. 22, No. 4 [cit. 08.09.2020]. Available from: https://bit.ly/3mdwJjs

[8] KUNTZ, Philipp - Mark R. THOMPSON. More than Just the Final Straw: Stolen Elections as Revolutionary Triggers. *Comparative Politics* [online]. 2009, Vol. 41, No. 3 [cit. 08.09.2020]. Available from: https://bit.ly/3a1VtsE

[9] TALEB, N. Nicholas - Mark BLYTH. The Black Swan of Cairo: How Suppressing Volatility Makes the World Less Predictable and More Dangerous. *Foreign Affairs* [online]. 2011, Vol. 90, No. 3 [cit. 31.05.2018]. Available from: https://bit.ly/376IYKf

[10] Ibid.

resources, and the dead end of the Syrian government's water and agricultural policies."[11]

Similarly, the theorised trigger of "stolen elections" is discussed primarily in the context of an electoral-authoritarian regime, where population and opposition feel like possibilities are taken away from them and the governing regime starts fracturing. This, in turn, leads to population mobilisation and attempts to change the results.[12]

This context-dependence returns the role of the structural conditions as a possible tool for prediction under a suitable methodology. O'Bannon suggests that structural data could be supplemented by choice of a particular species of conflict (in his case, the herder-farmer conflict in Africa), which would be an analogy to monitoring the so-called "indicator species" in the field of conservation biology.[13]

"As with the frog that biologists monitor so closely looking for signs of environmental disturbance, emergent, low-intensity conflicts indicate the presence of social, economic, cultural and/or political factors normally resistant to simple observation, but which are known to be associated with violent conflict."[14]

When it comes to conflicts that influence the security of the Czech Republic, no single type of conflict has been yet identified that could play this indicator role. Moreover, given the diversity of the regions involved (Western Europe, post-communist countries, Middle Eastern and North African countries), diversity of situations for conflict escalation is expected.

Using QCA

As mentioned before, there is more than one way how society can break down, and conflict can escalate, and this needs to be taken into consideration. That is why we propose using QCA (Qualitative Comparative Analysis) to identify conditions which describe a country in a state of conflict escalation.

In the extensive research of conflict escalation, tens of possible causes of conflict escalation had been identified and none of them standing alone can be declared as a sole cause of escalation. Neglecting this would lead to the false premise of causal homogeneity. It is necessary to cover various sets of causes and accept that causes may interact. Quantitative methods such as linear or logistic regression can address this by including interaction effects. However, the number of included interaction effects in one analysis is limited.

---

[11] DE CHÂTEL, Francesca. The Role of Drought and Climate Change in the Syrian Uprising: Untangling the Triggers of the Revolution. *Middle Eastern Studies*. 2014, Vol. 50, No. 4. DOI: 10.1080/00263206.2013.850076
[12] KUNTZ – THOMPSON, ref. 8
[13] O'BANNON, ref. 3
[14] Ibid.

QCA has a unique combination of advantages of both qualitative and quantitative methods. It offers tools for systemised comparisons between a much larger number of cases than comparative case studies. Generalisation is one of the main advantages of quantitative approach but, due to the huge variability across analysed countries, also qualitative insight into individual cases is needed. As QCA is also standing in between qualitative and quantitative approach, it equalises broadness of the former with the deepness of case study knowledge provided by the latter.[15]

Variants of QCA are available, however csQCA (crisp-set Qualitative Comparative Analysis) was chosen as an initial variant for this model. There are several advantages to starting with csQCA, firstly, there is a natural dichotomous character of several conditions (see below: IRE/SEC, TER_CL, CONF_50). Forcing data from these variables into scale instead of keeping them dichotomous would cause serious problems in the results. Moreover, using csQCA helps when dealing with an unequal quality of the data and when adding missing values by a qualitative input. For example, operationalisation of variable "power access inequality" (EPR_UNEQ) to values on scale 0-1 would add even more researchers' bias than the bias we have to face now. In such a case, it is analytically more straightforward to deal with binaries, which emphasise the qualitative aspect of the research. Additional nuancing of the model by employing fuzzy-sets might be a future endeavour.

Using our model is akin to taking a snapshot at a point where society is at and then comparing it to pictures taken in the past. These pictures from the past consist of two piles – those of escalations which resulted in a conflict that produced at minimum 500 BRDs in the timespan of an escalation peak and those where society pacified, did not escalate and no conflict of such scale followed. The result of the comparison is a list of countries that have an escalation-prone setup are chosen for closer analysis.

It is possible to draw on open-source data and the system does not require substantive infrastructure because the databases are maintained by various international organisations and non-governmental and research bodies and enable some compatibility in the data. QCA as a research method has been described[16] and applied[17] to various

---

[15] HUDSON, John - Stefan KÜHNER, 2017. Qualitative comparative analysis and applied public policy analysis: New applications of innovative methods. *Policy and Society*. Vol. 32, No. 4 [cit. 10.11.2020]. ISSN 1449-4035. Available from: doi: 10.1016/j.polsoc.2013.10.001

[16] RIHOUX, Benoît - Gisèle DE MEUR. Crisp-Set Qualitative Comparative Analysis (csQCA). In: RIHOUX, Benoît - Charles C. RAGIN, eds. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. SAGE, 2009.; RAGIN, Charles C. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley; London: University of California Press, 1989.; RAGIN, Charles C. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press, 2008.; SCHNEIDER, Carsten Q. - Claudius WAGEMANN. *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press, 2012.; RIHOUX, Benoît - Gisèle DE MEUR - Sakura YAMASAKI. Addressing the Critiques of QCA. In: RIHOUX, Benoît - Charles C. RAGIN, eds. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. SAGE, 2009.

[17] BINDER, Martin. Paths to intervention: What explains the UN's selective response to humanitarian crises? *Journal of Peace Research*. 2015. DOI: 10.1177/0022343315585847; MELLO, Patrick A. *Democratic participation in armed conflict: military involvement in Kosovo,*

research questions in conflict research numerous times, and therefore only pertinent features will be described in this article.

QCA is a method based on mathematical principles of Boolean algebra and set theory[18] suitable for capturing causal complexity in terms of identifying necessary conditions and combinations of sufficient conditions for escalation.[19] It offers a formalised way to describe each case of conflict escalation and escalation opportunity within an indicated timeframe (1989–2016) by a series of structural conditions. These are operationalised based on previous theoretical research on conflict escalation and collected from reliable and sustainable databases (See Table 2: Explanation [Source]). After that, the collected raw data is calibrated into membership of sets (0 or 1), based on a qualitative anchor drawn from the qualitative research (see Table 2: Calibration and footnotes for these) and the qualitative nature of the observed data.

Table 1: Example of a truth table

Y_UNEMP	POV	UNEQ	POP_GR	…	MILIT	Out-come	N. of cases	PRI
1	0	1	0	…	1	1	4	0.75
1	0	0	0	…	1	C	3	0.33
1	0	0	0	…	0	0	6	0
…								

Source: Authors.

This calibrated data table is then transformed into rows of a *truth table*.[20] Each row is evaluated for its *sufficiency* for the outcome (PRI measure, see Table 1). The sufficiency measure is based on the ratio of cases with the same combination of conditions that lead to the conflict escalation to the total number of cases with that combination of conditions. Sufficient rows (those with the assigned outcome 1) are then used for the minimisation procedure using software which reduces the long description only for those features of the row that are insufficient by themselves but are a necessary part of a compound condition that is unnecessary but sufficient for the outcome.

The main principle is to match the configurations in rows with membership to the set of outcomes (1) or non-membership to the set of outcomes (0). There is a possibility that the same configuration is matched with both 1 and 0 according to cases with that particular configuration. Based on the PRI measure (Table 1), it is not clear whether a significant portion of cases with the configuration led to the outcome or did not lead to

Afghanistan, and Iraq. Basingstoke, Hampshire: Palgrave Macmillan, 2014.; HAESEBROUCK, Tim. NATO Burden Sharing in Libya: A Fuzzy Set Qualitative Comparative Analysis. *Journal of Conflict Resolution*. 2016. DOI: 10.1177/0022002715626248; HAESEBROUCK, Tim. National Behaviour in Multilateral Military Operations. *Political Studies Review*. 2016. DOI: 10.1177/1478929915616288
[18] MAHONEY, James - Gary GOERTZ. A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research. *Political Analysis*. 2006, Vol. 14, No. 3. DOI: 10.1093/pan/mpj017
[19] SCHNEIDER – WAGEMANN, ref. 16, p. 9
[20] Ibid., pp. 119–190

the outcome. The usual threshold is a minimum of 0.8 for the outcome and maximum of 0.2 for non-outcome. However, this estimate is left to the researcher.[21] The configurations with the PRI measure falling within the interval (0.2, 0.8) are labelled as contradictory (C).

In this particular research, the causal analysis is not at the core, but the possibility of escalation. When trying to do causal analysis, the contradictory rows (their status as sufficient row is indeterminate) are then solved by various additional steps towards clarity. In our model, the overall coverage of the model (how many cases it includes) is more important than the consistency of the causal paths. We include rows with the consistency of 0.5 (there was an escalation in half of the cases in the row). After all, there is no standardised threshold that is binding for a researcher, and although higher measures are recommended, QCA is left to the researcher's consideration.[22] A society in this situation is conflict-prone at least as much as it is not and should appear in the list of countries indicated for further analysis.

The type of the solution utilised for this model is the *most parsimonious solution*, which is an algorithm that takes all rows of the truth table sufficient for the outcome (all rows with the consistency of minimum 0.5) and all logical remainders (the possible combinations which did not show in the empirical cases) for *Boolean minimisation*. The effect is a simplification of the combinations of conditions for the model (paths to the outcome).

The outcome of the Boolean minimisation is the recipe or the model for escalation, which is then (in the case of our project) used to sift recent data on the monitored countries. The user collects the same type of data for all states in the dataset for the last year (again country-year) and calibrates it according to the calibration anchors for the QCA model. Then the outcome model and the recent data are compared to each other, and only those countries which show similar combinations are included in the list for further analysis.

The *most parsimonious solution* then generates a higher number of recent cases that might inflate the false positives.

QCA is static – the principle of the analysis does not capture the dynamics of the case, which might be a problem when trying to discover the causality of a conflict escalation. This drawback does not pose a problem for our research because (in most cases) the yearly update of the source data and their long-term stability fit the use of the method for warning within 1-3-year horizon. This estimate is based on one-year lag on some of the variables (see Table 2: Data collected in time) and the span of the update of the open-source databases and the amount of data the model can absorb. In other words, results from creating these "snapshots" enables comparison with a past picture, but the conflict-prone setting of society does not mean the escalation is immediate, it can develop over time.

The conflict may intensify slowly and develop for years. Therefore, the predictions are not relevant only for the year of the snapshot but also for few following years. Conflict-

---

[21] SCHNEIDER – WAGEMANN, ref. 16, pp. 127-128
[22] Ibid.; SCHNEIDER, Carsten Q. - Claudius WAGEMANN. Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets. *Comparative Sociology*. 2010, Vol. 9, No. 3. DOI: 10.1163/156913210X12493538729793

prone moods in special conditions in society does not last forever and the so-called conflict heat, if not transformed into escalation, would eventually fade out.

BUILDING THE MODEL

All analytical steps related to QCA and further steps for model building and predictions were performed in R, free software for computing and graphics, with packages QCA v.3.10.[23] and SetMethods v.2.6.[24]

Measuring Escalation and Escalation Opportunities

The conditions (or variables) are calibrated into 1 and 0 according to membership in a particular set and so are analogically the outcomes of the cases.

Escalation (coded 1): UCDP country-year at the start of a period called *escalation peak*. For this UCDP country-year escalation data are collected.

Escalation peak is defined as a period of successive years when conflict reaches 25 BRDs yearly and 500 BRDs in total per the given period. Escalation peak is ended when the conflict falls below 25 BRDs per year.

Setting the threshold of overall BRDs generated during a particular conflict enables the model to differentiate between configurations of conditions according to the level of intensity of a subsequent conflict. A low-level conflict that exceeds 25 BRDs every year for a decade but with a relatively lower number of victims does not generate a relative threat to the Czech Republic compared to a highly intense 2- or 3-year conflict with thousands BRDs.

Escalation opportunities were coded 0. These are defined as cases where the outcome (escalation) could have happened, but it did not.[25] They were operationalised as societal upheaval during a country-year which was qualitatively distinct by its severity and escalation potential (big protests in repressive regimes, political repression, extraordinary moves of populations). Originally, these were coded qualitatively; maximum five such opportunities were chosen for each country. Additionally, country-years with 25 BRDs which did not lead to a prolonged conflict reaching 500 BRDs threshold were also coded as escalation opportunities.

[23] DUSA, Adrian, 2019. QCA with R. A Comprehensive Resource. In: *Springer International Publishing, Cham, Switzerland*. [online]. [cit. 2020-11-14]. Available from: https://bit.ly/378mo44

[24] OANA, Ioana-Elena - Carsten,Q. SCHNEIDER. SetMethods: an Add-on R Package for Advanced QCA. The R Journal [online]. 2018, Vol. 10, No. 1 [cit. 2020-11-14]. ISSN 2073-4859. Available from: doi: 10.32614/RJ-2018-031

[25] MAHONEY, James - Gary GOERTZ. The Possibility Principle: Choosing Negative Cases in Comparative Research. *The American Political Science Review* [online]. 2004, Vol. 98, No. 4 [cit. 17.08.2015]. Available from: http://www.jstor.org/stable/4145330

Measuring the conditions

The structural conditions were chosen for the model to cover several areas of conflict escalation theories and their proxy indicators – economic (poverty, inequality, youth unemployment), societal (power access, inequality amongst ethnic groups, population growth rate, irredentist/secessionist movements), neighbourhood (neighbouring conflict, territorial claims/demands), political (level of democracy, political terror), military (militarisation of the state).

These conditions are chosen as long-term variables correlated with conflict escalation and they are usually chosen according to the theories[26] they represent. But the goal of this model is not to analyse or even weight theories of escalation but to use conditions across societal phenomena to capture the state of society for building a comprehensive picture, instead of fitting it into any of the conflict outbreak or escalation theories, especially under the assumption that there are multiple ways of how countries end up in a conflict escalation. Variety of conditions is then more profitable for this use rather than one particular theory of conflict escalation; as has been pointed out, variables are proxies that might often fit into more than one escalation theory.[27] In the choice of the conditions their availability and sustainability of the open-source databases also played a role.

Every case in the model consists of a country and year; data for variables in those cases were collected mostly in the same year as defined in the case specification, but there are exceptions in 3 variables, as they were chosen to capture the state of society and are all structural and long-term. However, there is a slight difference in the reaction time of each variable, meaning that there is a difference in time in which the impact of changes in variables becomes evident in reality. As a simplified example - low GDP in a particular year has an effect on the society in the very same year, however a level of democracy in a particular year is a driving force for legislative changes the effect of which on society may become eminent only months later. Therefore, some of the variables were lagged when included into the process of model building (see Table 2: Data collected in time). The same logic applies to testing the current state of society – the "snapshot" against the previous escalations. When creating a snapshot of society in 2020, the data from 2020 would be used for 10 variables and the data from 2019 would be used for 3 variables.[28]

---

[26] DIEHL, Paul - Gary GOERTZ. *Territorial Changes and International Conflict (Studies in International Conflict)*. Routledge, 1992.; VASQUEZ, John A. *The Probability of War*. 2004.; TOFT, Monica Duffy. Issue Indivisibility and Time Horizons as Rationalist Explanations for War. *Security Studies*. 2006, Vol. 15, No. 1.; CEDERMAN, Lars-Erik - Andreas WIMMER - Brian MIN. Why Do Ethnic Groups Rebel? New Data and Analysis. *World Politics*. 2010, Vol. 62, No. 1.; DENNY, Elaine K. - Barbara F. WALTER. Ethnicity and Civil War. *Source: Journal of Peace Research*. 2014, Vol. 51, No. 2.; WATTS, Stephen et al. *Understanding Conflict Trends: A Review of the Social Science Literature on the Causes of Conflict*. RAND Corporation, 2017.
[27] DAVENPORT - LICHBACH – ARMSTRONG, ref. 4
[28] In the variable POP_GROWTH, the value from year t-30 is also necessary to calculate population growth over the years as of year t using the compound interest formula.

Table 2: Conditions and Calibration

Condition	Explanation [Source]	Data collected in time	Calibration[29]
Y_UNEMP	Modelled youth unemployment 15 – 24 years [ILO]	t	<15 % = 0 >15 %=1
POV	Poverty measured through GDP per capita per purchasing power in constant dollars (2011) [WB]	t	<6.800 USD = 1 >6.800 USD = 0
UNEQ	Inequality (GINI post-tax, post-transfer), [SWIID]	t	<33 = 0 >33 = 1
POP_GR	Population growth [WB]	t, t-30	<2 % = 0 >2 % = 1
EPR_UNEQ	Power access inequality. Ratio of marginalised ethnic groups to ethnically relevant population. [ETH Zurich]	t	Inequality exists (= 1) if MEG (marginalized ethnic group) > 0.1 or EGIP (ethnic group in power) = <0.8 Otherwise = 0
IRE/SEC	Existence of irredentist or secessionist movement [Secessionist movement, version 3/qualitative]	t	Movement exists = 1; does not exist = 0
TER_CL	Territorial claim [Paul Hansel TCOW database/qualitative)]	t	Territorial claims from other state or territorial claims to other state = 1 Otherwise = 0
NEIGH_CONF	Neighbouring conflict during the same year or year preceding [UCDP]	t and t-1	Conflict = 1 No conflict = 0[30]
CONF_50	Conflict during last 50 years in the territory of the same state [UCDP]	t	Conflict = 1 No conflict = 0
PT	Political terror scale [PTS]	t-1 (lagged)	3-5 = 1 0-2 = 0

Condition	Explanation [Source]	Data collected in time	Calibration
DEM	Democracy level [POLITY™ IV PROJECT]	t-1 (lagged)	0-5 = non-democratic (0) 6-10 = democratic (1)
W_STR	Water stress [FAO UN]	t	Stress = 1 No stress= 0
MILIT	Militarisation of society [BICC – Bonn International Center for Conversion]	t-1 (lagged)	<600 = 0 >600 = 1

Source: Authors.

THE 500+ MODEL

The cases of 25 BRDs a year but not reaching the 500 BRDs threshold of an escalation peak were then recalibrated and put into the set of escalation opportunities in order to make the model differentiate the potential of the 500 BRDs minimum conflict in the countries of interest from a lower-level conflict, which does not have severe spill-over effects and remains contained on the territory of one country. The severity of the conflict that might threaten the security of the Czech Republic is much lower due to European integration and interconnectedness. However, creating two QCA models to reflect this fact would be methodologically unfeasible at the moment. Splitting countries into two groups would cause a reduction of cases in both models and analysing 12 conditions on fewer cases would lead to unsound results and be against the good practice of QCA. The threshold of 500 BRDs could be changed to achieve higher sensitivity of the model for European countries. After checking the model's results with reality and other models' predictions (Table 3), the current threshold of 500 BRDs appears to be sufficient, and the analysis proceeds with it.[31]

There are 11 paths generated representing about 90% of the cases of conflict escalation. Although higher coverage would be preferable, this is a satisfactory outcome. These paths capture multiple cases clustering the countries and cases into related types of conflict. Since there was no single explanatory theory proposed, the story told by each of the paths is different but awaits theoretical considerations and interpretations.

---

[29] Sources used for choosing and setting calibration thresholds can be found in Appendix 1.
[30] PETTERSSON, Therése - Magnus ÖBERG. Organized violence, 1989–2019. *Journal of Peace Research*. 2020, Vol. 57, No. 4, 597-613. ISSN 0022-3433. Available from:
doi: 10.1177/0022343320934986
[31] The regionality is not erased from the data beyond the model and is taken into consideration in the subsequent analysis, which is beyond the scope of the article.

Table 3: Model of escalation

PATHS	CONSISTENCY
~Y_UNEMP*UNEQ*IRE/SEC*~TER_CL	0.67
~POV*POP_GR*~EPR_UNEQ*NEIGH_CONF	1
POV*~IRE/SEC*CONF_50*PT	1
POP_GR*IRE/SEC*CONF_50*~MILIT	0.63
Y_UNEMP*~UNEQ*CONF_50*PT*MILIT	1
POV*~POP_GR*EPR_UNEQ*~IRE/SEC*NEIGH_CONF	1
Y_UNEMP*~POV*IRE/SEC*PT*DEM*MILIT	0.75
~POV*UNEQ*IRE/SEC*~NEIGH_CONF*~DEM*MILIT	1
Y_UNEMP*~POV*~POP_GR*CONF_50*NEIGH_CONF*PT*~W_STR	0.75
POV*UNEQ*POP_GR*IRE/SEC*TER_CL*NEIGH_CONF*PT*MILIT	0.56
TOTAL CONSISTENCY	0.72
TOTAL COVERAGE	0.9

Source: *Authors*

LIST OF COUNTRIES AND COMPARISON TO OTHER SYSTEMS

The abovementioned 500+ model is based on data from 1989 to 2016. The model was built using configurations of structural conditions leading to conflict escalations in a set of countries in Western Europe, post-communist and post-Soviet countries, Middle East and Northern Africa countries (provided in Appendix 2). The most recent data used for the model are from 2016 because more up-to-date data from 2017 and 2018 serve as testing data for an assessment of prediction abilities.[32] The proposed model (paths in Table 2) is used as a sieve for the data from 2017 and 2018 and it generates a list of conflict-prone countries, in which escalation is probable. The state of society in those years is being analysed and compared to the pictures of societies in the model built from the 1989 – 2016 data. If a "snapshot" of the current society in any state matches the picture from the past with escalation as a result, this country is added to the list of potentially dangerous countries (see Table 3: OUR MODEL). These results tell us that in the listed countries, there is a configuration of factors representing escalation of conflict in the past.

---

[32] Application of the proposed model would require extending the dataset used for model building by the years 2017 and 2018 to create a model based on the period 1898–2018. Then the list of conflict-prone countries would be generated based on the 2019 data.

However, potential risk identified by using the 2017 data is not relevant only for the year 2017. The proposed model identifies the rising heat and tensions in societies, which can develop slowly over months or years until they grow into the conflict. Therefore, by creating snapshots in 2017, the following years 2018, 2019 and 2020 are also addressed. The same case of creating snapshots in 2018 alerts on potentially dangerous countries, being relevant also for the years 2019, 2020 and 2021.

The proposed method for prediction is exploratory and evaluation of its performance requires a non-standard method as well. Estimating values similar to goodness-of-fit measures for the model is not possible because of the configurational approach. The model contains configurations which led to a conflict in the past. It would find all the previous cases of escalation because all of them explicitly define the model.

Table 4: Countries Generated by our Model Compared to Two Global Systems – ICRG and GCRI

OUR MODEL	ICRG	GCRI	BRDs from 2017 to 2020
AFG	-	AFG	92466
EGY	-	-	1980
GEO	-	-	0
IRQ	IRQ	-	21203
LIT	-	-	0
MDA	-	-	0
PAK	PAK	-	2095
RUS	-	-	177
SDN	SDN	-	1712
SSD	-	SSD	1856
SYR	SYR	SYR	79147
TUR	-	-	5278
UKR	-	-	1138
TJK	-	-	0
-	-	NGA	6825
-	-	MNG	0
-	SOM	SOM	7996
-	YEM	-	11830
-	ZAR	-	0

Source: Authors

The International Country Risk Guide (ICRG), provided by PRS Group, is one of the principal political risk methodologies. Every year, the set of indicators of risk in different areas (Voice and Accountability, Political Stability and Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law, Control of Corruption) is computed and combined into the overall risk score. Final results are not available, PRS Group provides

free access only to partial scores.[33] ICRG covers 140 countries all over the globe; however, a few countries with a non-negligible impact on the Czech Republic are not included: e.g., Afghanistan, South Sudan, Bosnia and Hercegovina and a few post-communist countries.[34] The advantage of the proposed method over ICRG is its broader coverage of countries and accessibility. Using open-source data it enables the replication of the whole risk computation process with minimal expenses. The crucial question is the precision and reliability of both systems. A partial score for Political Stability and Absence of Violence was used for the comparison with the results of our model. Countries with a score below 0.5 in years 2016-2018 were chosen as potentially risky.

The second system in the comparison is the Global Conflict Risk Index (GCRI). Relying on the results of the regression model,[35] GCRI provides estimates of the probability and intensity of conflict in foreign countries. Quantitative conflict risk model by the European Commission and the proposed method both use structural variables and are based on open-source data. Therefore, using GCRI can solve the problem with accessibility, but still has the issue with coverage as the system is focused solely outside of the EU borders. Potentially risky countries were calculated with the high violent model probability equation for the national dimension with data provided by the European Commission Joint Research Centre. All countries with probability higher than 2% were included.

Table 3 presents the results of the comparison of all three models. The substantial difference seems to be in the sensitivity in some cases. In addition to the list of countries generated by ICRG and GCRI, our model is also alerting to Egypt, Mali, Russia, Turkey, and Ukraine. Conflict situations in all five countries have an impact on the threats for the Czech internal security; therefore, the five particular cases should be reflected when trying to prevent those threats. We achieve higher sensitivity by applying QCA; it enabled us to use different combinations of conditions valid in different parts of the world and apply them outside their regional context.

By including every single case of previous escalation opportunity in the analysed area, the model can also detect minor variables combinations leading to a conflict. The influence of these variables could fade away in the GCRI regression model. The identification of the risk of escalation in countries is systematic, and the procedure is transparent. Transparency is the advantage over ICRG because anyone can perform analysis and even modify and improve it.

Emphasis on the sensitivity also has a drawback in the form of generating false positives – Georgia, Lithuania and Moldova (at least false so far). There are certainly ways of improving the prediction model, such as including an additional condition that would be of shorter duration. Still, the inclusion of triggers themselves is epistemologically impossible since the problem of defining what a trigger is hinders the way forward. After all, disturbance of stable systems, the so-called "tipping points" are impossible to

---

[33] POLITICAL RISK SERVICES. The International Country Risk Guide (ICRG) [cit. 25.09.2020]. Available from: https://bit.ly/342TqAx

[34] Ibid.

[35] HALKIA, Stamatia et al. *The Global Conflict Risk Index (GCRI) regression model: data ingestion, processing, and output methods.* [online]. LU: Publications Office, 2017 [cit. 25.09.2020]. Available from: https://data.europa.eu/doi/10.2760/303651

predict.[36] However, the possibility of measurable proxies that are less stable in time but possible to capture on a month-to-month basis are still an opportunity for further research. Another possibility would be splitting the countries into more regions, but this would require recalibration of the conditions or even collection of new ones with the defined regionality, which would complicate it beyond its current parsimonious and relatively simple functionality.

Quality of the Model

There were overall 5759 data inputs and we encountered the problem of missing data for some years or some of the countries accounting for 1.72% of data. Variables were affected to a different extent, major influence of missing data was on IRE/SEC and W_STR. Regarding countries, data were missing especially for Yemen and Libya, however, the missing data for these countries did not exceed 40%. Therefore, the first step of filling the missing cells was a qualitatively informed input of data - missing data for some variables were collected from other databases or qualitative sources and their consistency with the available data was checked.

Step two for data that could not be filled by the qualitative method was interpolation using linear regression; step three was the use of random data. Random data were used to fill missing values in raw data, random number had been drawn from uniform distribution, minimum and maximum were based on the min and max values from the values available for the particular variable. Random values account only for 0.45% of the final data. For evaluating the stability of the outcomes, the random input was iterated fifteen times, and the comparison shows a high level of stability of the model – the model is not case sensitive.

Picture 1: Stability Analysis for Random Values
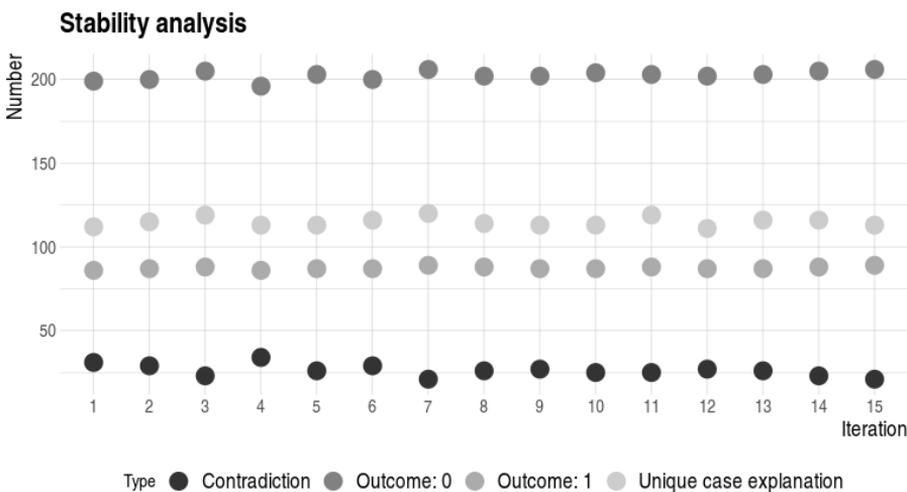
Source: Authors.

---

[36] TALEB – BLYTH, ref. 9

In addition to the stability analysis, another way of ascertaining that the model describes what it should be describing is to run it for escalations of 25 BRDs a year. This step showed that the model certainly differentiates between countries which are stable long-term and where the expectation of conflict escalation is very low (the overwhelming majority of West European countries, Central European countries of the post-communist block) and those that are unstable. In terms of methodology, this gets demonstrated by the fact that the cases in a contradictory row represent the same country, only in different timeframes, but they are not contradictory to cases representing stable countries. The case clustering, according to social reality, works reliably in this application of QCA thanks to the multiple conditions included in the model.

CONCLUSION

The presented method of conflict escalation prediction combines the assessment of structural conditions and QCA results in the identification of countries in which escalation is possible and probable. It leverages the fact that QCA enables to investigate conjunctural combinations of conditions that describe a society going into conflict escalation and the fact that long-term structural conditions are collected by open-source databases and easily available to potential users. Similar kind of risk alert is not new, but the proposed model can offer a few advantages over the established systems when reflecting the specific needs of the Czech Republic.

Our system has the advantages of being a localised and relatively cheap method of conflict escalation prediction, which functions as a source for further probability and intensity estimation and scenario building tied to the threats emanating from the countries of conflict escalation. A specific scenario building method is designed to work with the abovementioned model of prediction, but its description is beyond the scope of this paper. Institutional users, such as intelligence agencies or policy-making bodies, would certainly need to complement this system with regional intelligence gathering on the ground or targeted analyses of specific threats, however, this system offers a systematised way of specifying targets of focus and additional sources.

The model performs well in comparison with much larger and sophisticated systems of conflict prediction that also use long-term structural data. There are still possible ways to improve it, possibly by introducing shorter-term conditions in order to improve the time sensitivity of the predictive model. Additionally, the model might be nuanced by the introduction of fsQCA, but this needs to be tested and improved in further iterations. Still, as demonstrated above, the model shows a lot of promise.