

CAS final report

Stability and Change 2022–2023

Report by Nils Lid Hjort and Håvard Hegre

Project leaders

November 2023

Abstract

The Stability and Change project aimed at and succeeded in fruitfully bringing together core groups of international top-level scholars from two different fields – those of (a) peace-and-conflict research, with associated themes from political science, and (b) methodological and applied statistics. Themes of clear shared interest for (a) and (b) include (i) modelling and analysis of armed conflict, taking the special statistical features of violence data into account; (ii) unravelling key patterns which lead to higher risk for escalation, building upon relevant time series of violence and influential covariates; (iii) modelling, estimating, and understanding the heavy-tailed power-law distributions prevalent with violence data; (iv) how to combine information from partly very different data sources; (v) identification of trends and changepoints, e.g. regarding democracies and consequences thereof; (vi) how to count the not yet counted (as in casualties in wars); and (vii) predicting future levels of violence, along with adequate assessment of uncertainty.

1 General description: Armed conflicts, wars, statistics, predictions, uncertainty

The *Stability and Change* work group brought together researchers from statistics on the one hand and political science and conflict research on the other to explore a number of research questions related to understanding and possibly predicting armed conflict. Conflict data exhibit a number of challenges that were topics of the collaboration. One such challenge is the distribution of the data,

as most cases of interest have no observed conflict whereas the actual conflict cases exhibit a heavily right-tailed distribution. Both inferential and predictive analysis is difficult with such data, and even assessing aggregate trends and changes in them is far from straightforward. Another challenge is that a considerable fraction of the violence tends to go unnoticed. Seeking out methods to fill in the gaps in the available data or to systematically handling these gaps where they occur has been another topic (cf. the PredUnc, N000, and DemoIndex workshops described below). Relatedly, the group also explored various methodological approaches to predicting with uncertainty.

Yet another challenge relates to the fact that we mainly observe the actual violence in armed conflict, whereas the causes of this violence such as the strategic considerations of the actors involved can only be inferred through modeling the conflict processes theoretically. Professor Braumoeller was leading this work, which was also the content of the ProcMod workshop.

2 Participants and research activities

From work carried out and still ongoing from the group's long-term and crucial participants, on the statistics methodology side, we briefly give these pointers. Jonathan Williams worked on conformal prediction methods (Williams, 2021), with Gudmund Hermansen and others with hidden Markov models for modelling and analysing both aspects of civil wars and ceasefires (Williams et al., 2023) as well as on computational aspects of fitting Markov processes, with a continuous time, time-inhomogeneous transition rate matrix (Kendall et al., 2022), and with Hermansen, Randahl and Vegelius to build an R package to incorporate a variety of different Markov modeling approaches to prediction and uncertainty quantification. Williams also teamed up with Anders Hjort, for new types of applications of Conformal Prediction methods, with A. Hjort now spending the autumn semester 2023 at North Carolina State University.

Also, work has been initiated, with Céline Cunen, Hjort, and Williams, for building model selection criteria for 'counting the not counted' challenges. Idris Eckley, with collaborators Martin Tveten, Per August Moen, Ingrid Glad, have worked with models and algorithms for detection of both changepoints and anomalies, partly with industrial applications in mind, but where methods also have clear potential for being used for time series of conflict. Cunen and Hjort have started a project where changes are seen as having happened over certain time periods, inviting the notion of Windows of Change, as opposed to the traditional changepoint literature view. Finally, Aaron Clauset has pursued parallel research lines with

Kristian Skrede Gleditsch exploring the origins of the power-law nature of violence statistics, including (i) the study of violent attacks by terrorist organisations, and (ii) the escalation dynamics within civil and interstate wars (Clauset et al., 2023).

The group also explored several other research topics related to the forecasting of armed conflict beyond 'prediction with uncertainty' (Williams, 2023; Vesco et al., 2023; Croicu, 2023; Hegre, 2023; Randahl, Williams, and Hegre, 2023; Randahl, 2022; Schincariol et al., 2023, and the PredUnc workshop). The group used the year at CAS to plan a prediction challenge aiming to attract solutions to the tricky problem of predicting armed conflict with uncertainty (Hegre et al., 2023). Moreover, the group explored various aspects of evaluation metrics (Colaresi and Hegre, 2023), matching data at different temporal resolutions (Hegre and Colaresi, 2023), the importance of vulnerability factors (D'Angeli and Vesco, 2023), and deep learning models (von der Maase, 2023).

The VIEWS project, directed by Hegre and involving several of the CAS group members (Vesco, Croicu, Dale, Randahl, and Williams), updated its forecasting model during the CAS year.

Finally, the group worked on the links between armed conflict and migration, forced and voluntarily, during the year of activity (Leis and Petrova, 2023). Core group member Bethany Lacina was engaged with a book manuscript, entitled *Strangers and Settlers: Migration and Conflict in a Nativist World* (Lacina, 2023b). The book deals with theories of nativist conflict and also weighs in what kinds of migration data best predict conflict. Maxine Leis and Bethany Lacina also conducted an investigation of the quality of migration data drawn from the Knowledge Centre on Migration and Demography (KCMD) Data Portal hosted by the European Commission. They evaluated the geospatial migration data from KCMD in comparison to national census data in order to inform the VIEWS project's use of population flow data. Finally, at the group's workshop on monitoring and measuring influences of democracy, there was a session on regime datasets tracking migration-related civil rights over time. The most common forms of legal migration restrictions have shifted over the last 150 years. Existing data captures the decline of historical measures but is less informative on contemporary trends (Lacina, 2023a).

3 Colloquia, workshops, seminars, and excursions

3.1 Workshops

We have had as many as five successful workshops during the CAS year, and also a two-day Pre-CAS workshop before the CAS year started. Detailed programmes for these workshops, with titles of talks, various pdfs from presentations, etc., can be had from the project leaders. Additionally, core members gave invited presentations in other fora, including talks at the Department of Mathematics, UiO, and we regularly had smaller in-house special topics discussions.

3.1.1 PreCAS workshop, 11–12 May 2022

The aim of this workshop was to prepare the ground for the CAS year, with a selection of themes presented by the Fellows-to-be. In particular, Hjort and Hegre gave broad introductions; Cunen talked about Statistical Sightings of Better Angels; Clauset on escalation dynamics; Eckley on change-point identification.

Speakers: NL Hjort, C Cunen, J Williams, GH Hermansen, A Clauset, CH Knutsen, D Randahl, KS Gleditsch, H Hegre, H Nygård, EA Stoltenberg, PA Moen, B Lacina, R de Bin, I Eckley, M Croicu.

3.1.2 ProcMod: From Processes to Models. 19–20 October 2022

This workshop focused on ways of understanding data via background processes. As partly opposed to modelling data-as-they-are, one may gain better understanding and insights into seeing data characteristics as consequences of deeper structures. This goes in particular for the interpretation and analysis of the prevalent long-tailed power-law distributions for conflict data. Hjort presented such processes leading to power-laws; Williams talked about hidden Markov models; C. Stoltenberg discussed time-to-event via increasing processes; E. Stoltenberg presented dynamic processes over time for 0-1 data; Braumoeller introduced theoretic models of war; Randahl introduced a new class of Markov models.

Speakers: J Williams, B Braumoeller, SP vd Maase, NL Hjort, M Croicu, D Randahl, C Cunen, GH Hermansen, C Stoltenberg, EA Stoltenberg.

3.1.3 PredUnc: Prediction with Uncertainty. 6–7 December 2022

Themes included general methods for assessing the uncertainty levels of predictions, for processes moving in time and space. Such themes are prevalent and important in many branches of society, not only for conflict prediction. Jullum discussed the famous-in-Norway strømpisene; KA Aastveit explained the methods used by Den Norske Bank; Croicu analysed the consequences for uncertainty of imprecise geographic locations; Vesco a project to elicit uncertainty perceptions from coders in the Uppsala Conflict Data Program. The workshop included an excursion, about 50 metres away from the workshop Røde Kors room, to the White Buses, with Hjort giving a short account of the drama behind this WW2 operation.

Speakers: H Hegre, J Vestby, M Jullum, P Vesco, J Williams, A Hjort, C Cunen, NL Hjort, A Hubin, D Randahl, GH Hermansen, KA Aastveit.

3.1.4 N000: Counting the Uncounted. 22 February 2023

The important theme of 'counting the uncounted' relates to several tasks in the studies of war and conflict, as when estimating the number of deaths, but also with applications more broadly to 'missing variables'. Here Patrick Ball, leader of the Human Rights Data Analysis Group (and winner of the Rafto Prize) gave an open and broad presentation. Hjort estimated the number of persons killed in Guatemala, 1978–1996; Spagat talked about such issues for the Iran-Iraq war; and Högladh discussed general methods for the UCDP.

Speakers: NL Hjort, P Ball, H Brunborg, S Högladh, M Spagat, A Hubin, P Vesco.

3.1.5 ChangeTrend, Changes and Trends. 28–29 March 2023

Modelling, analysing, interpreting vital time series, concerning violence, democracy, socio-economic parameters, etc., often involves the challenging tasks of identifying trends and changes. This workshop addressed such, from various perspectives. Hjort discussed the Window of Change; Tveten reported from industrial applications; Nyrup presented a new dataset pertaining to the social profile of governments; Eckley gave methods for anomaly detection; Clauset presented a change-point detection algorithm for the structure of changing social networks.

Speakers: NL Hjort, A Clauset, M Tveten, EA Stoltenberg, P LUTscher, J Nyrup, D Christensen, PA Moen, C Cunen, I Eckley, CH Knutsen.

3.1.6 DemoIndex: Measuring and Monitoring Aspects and Influences of Democracy. 26–27 April 2023

Our final workshop concerned *democracy indexes*, how they are constructed and interpreted, with different components, and how they can be studied, in particular to how they can be seen as influential mechanisms for lowering the risk of conflict. Nord discussed the V-Dem reports, Knutsen the different aspects of subjective vs. objective democracy measures; Nyrup explained aspects of female participation in governments; Randahl analysed regime changes with attempts at prediction.

Speakers: M Nord, CH Knutsen, J Nyrup, D Randahl, F Haass, NL Hjort, B Lacina, KS Gleditsch, V Djuve, J Schmid, M Colaresi, M Rivers.

3.1.7 ANTICIPATE workshop

In addition to the workshops arranged through CAS, most project members attended the workshop arranged by the ERC ANTICIPATE project (with Hegre as PI) at PRIO on March 1–3, with the following presentations:

- Paola Vesco, "The impacts of armed conflict on human development: a review of the literature"
- Maxine Leis and Kristina Petrova (UU), "Determining distress-related mobility at the household level – an interpretative machine learning framework"
- David Randahl, "Prediction of electoral violence"
- Håvard Hegre, "Predicting the number of fatalities in armed conflict, with uncertainty — a prediction competition"
- Mihai Croicu, "A fast spatial reassignment procedure for imprecise armed conflict events"
- Chandler Williams, "Nowcasting UCDP-GED based on UCDP-Candidate"
- Paola Vesco, "The underreported death toll of wars: a probabilistic reassessment from a structured expert elicitation"
- Simon Polichinel von der Maase (PRIO), "Adding a variational autoencoder to ConflictNet"
- Michael Colaresi, "Introducing Multiplex pseudo-Earth Mover Divergence (Multi-pEMDiv) to Score Probabilistic Forecasts of Count and Ordinal Targets in Social, Political, and Networked Contexts"

3.2 Open lectures

Our workshops were essentially in-house (and took place either at CAS or at PRIO / Røde Kors), but were also announced more broadly, and had participants outside the project group. In addition, Bethany Lacina gave an internal lunch seminar at CAS as well as a brownbag at PRIO, and Patrick Ball's presentation on 'counting the not counted' also had a broader audience.

Additionally, several of us, including group leaders Hegre and Hjort, have given general presentations for broader audiences, including a common presentation on war, peace, statistics, predictions, at the Videnskaps-Akademiet, October 2023, with half an hour of general discussion afterwards.

4 Achievements and evaluation

The original project description was written by Nils Lid Hjort and Håvard Møkleiv Nygård. The main intended focus was on the distributional aspects of armed conflict data and the methodological implications of this for assessing past and future trends as well as causal relations in armed conflict, theoretically and empirically. When Nygård was offered a position outside academia, Hegre stepped in as co-leader of the group. This led to a certain shift in focus from primarily assessing past trends to focusing predicting future ones. This shift came very naturally, however, since the methodological challenges for predictive and inferential modeling are closely related.

Among the most important findings we highlight these.

- Applying models based on statistical theory to link any forecasting model and the prediction errors they generate in a test data set to produce predictions with a valid representation of uncertainty, both based on the conformal prediction approach and on bootstrapping techniques (Randahl, Williams, and Hegre, 2023; Randahl, 2022).
- Assessing the uncertainty of the input data used in armed conflict modeling, including nowcasting models (Williams, 2023), measurement uncertainty (Vesco et al., 2023), and gaussian processes to infer likely geographical location (Croicu, 2023).
- A realization that the recommended best practice of optimizing forecasting model on RLMSE – on the logarithm of the outcome – is counterproductive for armed conflict forecasting and leads to systematic underpredictions (Colaresi and Hegre, 2023).

- A theoretical solution to the problem of matching data measured at different temporal resolutions for forecasting purposes (Hegre and Colaresi, 2023)
- The design and preparation of a new, international prediction competition on the problem of armed conflict forecasting (Hegre et al., 2023).

A primary means of collaboration during the CAS project were the highly intense workshops. Each presentation was subject to detailed discussion. Also important has been several new co-authorships established during the year. Most importantly, being together in the same location for long periods of time offered ample opportunities to consult on topics of joint interest, both methodological and more domain-specific.

These aspects and efforts have also led to work published or in the process of being published in statistical methodological journals. Here we point to Dæhlen, Hjort, and Haff (2023), where advancement of focused model selection criteria are used for fine-tuned analysis of the important battle-deaths time series for interstate wars 1823-2023, and to Christensen, Hjort, and Moen (2023), in which new developments in Bayesian change-point methods are motivated from and illustrated for the same dataset. Furthermore, Williams, Hermansen, Nygård, Clayton, Rustad, and Strand (2023), with direct collaboration between statisticians and PRIO conflict researchers, have developed new methods for Bayesian hidden Markov models, for analysing the effect of ceasefires. Finally, the Hjort and Stoltenberg (2024) book will be published by Cambridge University Press, with several Statistical Stories built around datasets and research themes from Stability and Change.

A series of follow-up projects involving the working group are already in action. Several members of the CAS group will continue to work with the VIEWS group in various projects. Most importantly, a RCN project starting up September 2023 will actively involve Gudmund Hermansen, Céline Cunen, and Jonathan Williams in the project, working on topics initiated at CAS, together with the CAS researchers from PRIO and Uppsala University. Also, among Masters and PhD projects supervised by Hjort at the Department of Mathematics, UiO, are themes related to and building upon research carried out during our CAS year. Hjort is among the PIs of the newly launched Centre of Excellence *INTEGRATE: Norwegian Centre for Knowledge-driven Machine Learning*, 2023–2033, and aims at integrating certain Stability and Change themes in some of the centre's projects.

5 Evaluation of CAS

The CAS concept has been enormously valuable to the group. In particular, it has allowed a great strengthening of the relations between the statisticians interested in armed conflict research and the statistically oriented conflict researchers. Before CAS, these relations were mainly through the interactions between Hjort, Nygård, and Hegre, but the network is now much richer. Oslo remains the hub of this network, but its international edges are now much richer.

We feel compelled to mention that our friend and colleague Professor Bear Braumoeller very unexpectedly died, in May 2023, in the middle of active collaboration with the group. Bear was a first-rate international scholar, always generously interested and supportive, with the Fellows and Visitors, and is greatly missed. In the weeks following this dramatic event CAS was very supportive and helpful, for the project group and for Bear's family.

The facilities at CAS are fabulous (although the temperature kept in the offices was too low for some). The year at CAS could have been even better for a group like ours (which probably was a bit of an outlier). First, given that the group consisted of a relatively small core and a sizeable number of associates it would have been better to have more desk spaces available at any time, in particular for guests coming for a day only. Managing office space was very demanding. The group had Chandler Williams available to help the group leaders with administrative tasks, but the perennial shortage of office space meant that he could not be present at CAS more than a fraction of time. In general, for our group where the group leaders had extensive project management commitments on top of the CAS working group, administrative support was in short supply, suggesting some room for improvement. The costs for this support should come from the project budget and we did charge some of it for this purpose, but we had no good solution to the office constraint problem.

(As a minor footnote to our report we mention that Hjort turned seventy, in the middle of Stability and Change, with the consequence that he worked as group leader for NOK zero for the last five months of the project. We had hoped that a modest agreement, with pensjonslønn NOK 233 per hour for e.g. two days a week, could have been entered into the budget, but that was not granted. Nevertheless, Hjort will carry on with CAS project related work and supervision at the Department of Mathematics, UiO, apparently still for NOK zero.)

References

- Christensen, Dennis, Nils Lid Hjort, and Per August Moen (2023). *Bayesian changepoint analysis for the interstate battle deaths time series, 1823-2023*. Typescript.
- Clauset, Aaron, Barbara F. Walter, Lars-Erik Cederman, and Kristian Skrede Gleditsch (2023). *Escalation dynamics and the severity of wars*. Draft manuscript.
- Colaresi, Michael and Håvard Hegre (2023). *RMSLE versus RMSE: Evalutaion of log-transformed outcomes will not reward the models you think they will when distributions are zero-inflated*. Typescript, Center for Advanced Studies, Oslo.
- Croicu, Mihai (2023). *A Fast Multiple Stochastic Spatial Reassignment Procedure for Imprecise Armed Conflict Events*. Typescript, Uppsala University.
- D'Angeli, Mariagrazia and Paola Vesco (Mar. 2023). *Predicting the impact of armed conflict on vulnerability: a Machine Learning approach*. Presented to 2023 Annual Convention of International Studies Association. Montreal, Canada. URL: https://viewsforecasting.org/wp-content/uploads/Dangeli_Vesco_2022_Vulnerability.pdf.
- Dæhlen, Ingrid, Nils Lid Hjort, and Ingrid Hobæk Haff (2023). "Accurate bias estimation with applications to focused model selection". In: *Scandinavian Journal of Statistics*.
- Hegre, Håvard (2023). *Accounting for uncertainty when forecasting armed conflict*. Lecture at CAS-LMU, Munich, 10 May 2023. URL: https://www.en.cas.uni-muenchen.de/events/lectures_ss_23/ev_hegre/index.html.
- Hegre, Håvard and Michael Colaresi (2023). *Forecasting annual outcomes in VIEWS*. Typescript, PRIO and Uppsala University.
- Hegre, Håvard, Paola Vesco, Michael Colaresi, and Jonas Vestby (2023). *The 2023/24 VIEWS Prediction competition: Predicting the number of fatalities in armed conflict, with uncertainty*. Typescript, PRIO and Uppsala University. URL: https://viewsforecasting.org/wp-content/uploads/VIEWS_2023.24_Prediction_Competition_Invitation.pdf.
- Hjort, Nils Lid and Emil Aas Stoltenberg (2024). *Statistical Inference: 555 Exercises, 66 Stories, and Solutions to All*. Cambridge University Press.
- Kendall, E. B., Jonathan P. Williams, Gudmund Horn Hermansen, F. Bois, and V. H. Thanh (2022). *Beyond time-homogeneity for continuous-time multistate Markov models*. Under review.
- Lacina, Bethany (2023a). *Migration laws in democracy indices: Trends captured and trends missed*.

- Lacina, Bethany (2023b). *Strangers and Settlers: Migration and Conflict in a Nativist World*. Precis presented at the European Political Science Association conference, June 2023. Unpublished manuscript.
- Leis, Maxine and Kristina Petrova (2023). *Determining distress-related mobility at the household level - an interpretative machine learning framework*. Uppsala University.
- Randahl, David (2022). *Parametric and non-parametric measures of uncertainty of predictions*. Paper presented at the CAS Stability and Change workshop on prediction with uncertainty, 6-7 December 2022, Oslo, Norway.
- Randahl, David, Jonathan Williams, and Håvard Hegre (2023). *Conformal predictions to estimate uncertainty in conflict forecasts*. Accepted for presentation at APSA 2023 Annual Meeting Exhibition Los Angeles, CA, 31.08.23 – 03.09.23.
- Schincariol, Thomas, Paola Vesco, Thomas Chadeaux, and Bear Braumoeller (Aug. 2023). "Identifying Motifs and Patterns in War Fatalities: A Matrix Profile Application". In: *APSA Annual Convention*. Los Angeles, USA.
- Vesco, Paola, Håvard Hegre, David Randahl, and Stina Höglbladh (2023). *Uncertainty in the coding of fatalities in armed conflicts: an expert elicitation of the UCDP coding team*. Typescript, Uppsala University.
- von der Maase, Simon (2023). *ConflictNet 1.0 - A probabilistic recurrent U-net for globe conflict forecasting*. Presented the ANTICIPATE workshop, PRIO 03.03.2023.
- Williams, Chandler (2023). *Nowcasting Fatalities: nowcasting using UCDP Candidate data*. Presentation at the ANTICIPATE workshop, 1-3 March 2023, Oslo, Norway.
- Williams, Jonathan P. (2021). *Model-free generalized fiducial inference*. Under review.
- Williams, Jonathan P., Gudmund Horn Hermansen, Håvard M. Nygård, G. Clayton, Siri Aas Rustad, and Håvard Strand (2023). *Bayesian hidden Markov models for latent variable labeling assignments in conflict research: application to the role ceasefires play in conflict dynamics*. Typescript.

APPENDIX

Participants and research activities

Fellows (long-term, crucial participants)

Bear Braumoeller	Professor, Ohio State University
Aaron Clauset	Professor, University of Colorado, Boulder
Michael Colaresi	Professor, University of Pittsburgh
Mihai Croicu	PhD Candidate, Uppsala University
Céline Cunen	Senior Researcher, Norwegian Computing Center
James E. Dale	Data scientist, Uppsala University
Idris Eckley	Professor, Lancaster University
Kristian Skrede Gleditsch	Professor, University of Essex
Bethany Lacina	Associate Professor, University of Rochester
David Randahl	Postdoctoral Fellow, Uppsala University
Paola Vesco	Researcher, Uppsala University
Jonathan Williams	Assistant Professor, North Carolina State University

Affiliated researchers

In addition to these crucial core participants, the following have been affiliated researchers:

Patrick Ball	Leader of Human Rights Data Analysis Group
Dennis Christensen	PhD Candidate, University of Oslo (UiO)
Gudmund Horn Hermansen	Associate Professor, University of Oslo (UiO)
Anders Hjørt	PhD Candidate, University of Oslo
Aliaksandr Hubin	Associate Professor, University of Oslo (UiO)
Per August Moen	Postdoctoral Research Fellow, University of Oslo (UiO)
Emil Aas Stoltenberg	Associate Professor Norwegian Business School (BI)
Håvard Strand	Associate Professor, University of Oslo (UiO)
Simon Polichinel von der Maase	Senior Researcher, Peace Research Institute Oslo (PRIO)
Chandler Williams	Research Assistant, Peace Research Institute Oslo (PRIO)
Maxine Leis	PhD Candidate, Uppsala University