

Modellbaserad prediktion av säkerhetspolitiska händelser

JONAS CLAUSEN MORK, CARINA GUNNARSON,
MICHAEL JONSSON OCH FARZAD KAMRANI



**Jonas Clausen Mork, Carina Gunnarson,
Michael Jonsson och Farzad Kamrani**

Modellbaserad prediktion av säkerhetspolitiska händelser

Title	Model-based prediction of security policy-relevant events
Rapportnr/Report no	FOI-R--4954--SE
Månad/Month	Juni/June
Utgivningsår/Year	2020
Antal sidor/Pages	83
ISSN	1650-1942
Kund/Customer	Försvarsdepartementet
Forskningsområde	Säkerhetspolitik
FoT-område	Inget FoT-område
Projektnr/Project no	A11943
Godkänd av/Approved by	Malek Finn Khan
Ansvarig avdelning	Försvarsanalys

Bild/Cover: doctor-a, Deviant Art

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk, vilket bl.a. innebär att citering är tillåten i enlighet med vad som anges i 22 § i nämnd lag. För att använda verket på ett sätt som inte medges direkt av svensk lag krävs särskild överenskommelse.

This work is protected by the Swedish Act on Copyright in Literary and Artistic Works (1960:729). Citation is permitted in accordance with article 22 in said act. Any form of use that goes beyond what is permitted by Swedish copyright law, requires the written permission of FOI.

Sammanfattning

Denna rapport undersöker fyra modellbaserade konfliktprediktionssystem – ICEWS, PVRAM, CRAM och ViEWS – avseende övergripande arkitektur, inriktning och teknikval. Även FOI:s produktion inom afrikanska säkerhetspolitiska studier har studerats. Dessa två delstudier ligger sedan till grund för ett resonemang kring möjligheten, och nyttan med, att FOI skulle bygga upp en egen förmåga att konstruera och använda prediktiva system inom det säkerhetspolitiska området.

Till slutsatserna hör att flera teknologiska framsteg, i synnerhet under det senaste decenniet, påtagligt har förbättrat möjligheterna att konstruera kapabla säkerhetspolitiskt inriktade prediktiva system. Det finns också en växande mängd praktiska erfarenheter av att konstruera dylika system både inom den akademiska forskningen och inom privata företag.

Att döma av inriktningen på FOI:s afrikanska säkerhetspolitiska studier så är prediktiva bedömningar något som återkommande spelar en viktig roll. Bedömningen är därför att den säkerhetspolitiska analysverksamheten har nytta av en förstärkt prediktiv förmåga.

FOI bedöms slutligen ha goda förutsättningar för att såväl konstruera som äga prediktiva system inom det säkerhetspolitiska området, även om utvecklingen av dem sannolikt innebär betydande utmaningar.

Nyckelord: prediktion, säkerhetspolitik, statistik, maskininläring

Summary

This report studies four model-based conflict prediction systems – ICEWS, PVRAM, CRAM and ViEWS – with regard to overall architecture, focus and technology choices. FOI's production in the area of African security policy studies has also been examined. These two studies provide the basis for a discussion concerning the possibility, and utility, of FOI creating a capability of its own regarding the construction and use of predictive systems in the domain of security policy.

The study concludes, among other things, that technological progress in several areas have markedly improved the possibilities to construct capable predictive systems focusing on security policy-related events. There is also a growing body of experience from constructing these kinds of systems within the academic sphere, as well as in private companies.

Judging from the focus and content of FOI's publications concerning African security policy, prediction is a recurrent and important theme. The assessment is therefore that the agency's security policy studies would benefit from a strengthened predictive capability.

Finally, the prospects at FOI for construction and management of predictive systems in the domain of security policy look favorable, even if their development is likely to entail considerable challenges.

Keywords: prediction, security policy, statistics, machine learning

Innehållsförteckning

1	Inledning	10
1.1	Syfte	10
1.2	Frågeställningar	11
1.3	Målgrupp.....	11
1.4	Läsanvisningar	11
2	Bakgrund	12
2.1	Exempel på användning av säkerhetspolitiskt inriktade prediktiva system	12
2.2	Numerisk probabilistisk prediktion.....	13
2.3	Modellbaserad, datadriven eller både och?	13
2.4	Manuell och automatisk insamling av händelsedata.....	14
2.5	Att mäta prediktiv prestation.....	16
3	ICEWS – militär förvarning.....	17
3.1	Systemets fokus	18
3.2	Händelsedata	18
3.3	Modeller.....	19
3.3.1	Agentbaserad simulering	19
3.3.2	Regression.....	20
3.3.3	Sammanvägning av ensembler	22
3.4	Prestation	22
3.4.1	Förvirringsmatrisen och relaterade mått.....	22
3.4.2	Brierpoäng	25
3.4.3	Värden	25
4	CRAM och PVRAM – strategiska underlag till departementen	26
4.1	Systemens fokus	26
4.1.1	Political Violence Risk Assessment Monitor	26
4.1.2	Conflict Risk Assessment Monitor	27
4.2	Händelsedata	28
4.3	Modeller.....	29
4.4	Prestation	29
4.4.1	Värden	29

5	ViEWS – med sikte på humanitära insatser	30
5.1	Systemets fokus	31
5.2	Händelsedata	32
5.3	Modeller	32
5.3.1	Modellering och simulering	33
5.3.2	Dynamisk simulering (dynasim).....	34
5.3.3	Ett-steg-fram-modellering	35
5.4	Imputation: hantering av saknade data.....	35
5.5	Prestation	37
5.5.1	Värden	37
6	Teman inom FOI:s analyser avseende den afrikanska kontinenten	38
6.1	Tematiska trender för Afrikaprojektets produktion 2014-2019.....	38
6.2	Framåtblickande analys och prediktion inom Afrikaprojektets studier 41	
6.2.1	Exempel på studier med prediktion som huvudsyfte.....	43
6.2.1.1	R-rapport (2014) "Mellanöstern och Nordafrika i ett 5- 10-årsperspektiv"	44
6.2.1.2	AG Säkpöl: Underlag (2018) "FOI:s uppdaterade underlag inför nästa försvarspolitiska inriktningsbeslut".	44
6.2.1.3	SLO (Strategisk långsiktig omvärldsanalys): Underlag (2018) "Säkerhetspolitiska trender i Afrika söder om Sahara"	45
6.2.2	Analys	45
6.3	Möjliga områden för framttagande av modeller för prediktion	46
7	Förutsättningar att skapa prediktiva system för FOI:s behov	50
7.1	Utmaningar inom modellering	50
7.2	Utmaningar avseende mjukvara	50
7.3	Utmaningar avseende data.....	51
7.4	Utmaningar avseende datakraft.....	51
7.5	Utmaningar avseende utvecklingsresurser	52
8	Slutsatser.....	53
	Referenser.....	54
	Appendix A.....	60

Strictly proper scoring rules	60
Brierpoäng	60
Logaritmiska poäng	61
Uppdelning i komponenter	61
Kalibreringskurva	62
MCC och F-poäng	64
Areamått	65
Referenser	67
Appendix B	68
Random forest	68
Beslutsträd som beslutsstöd	68
Beslutsträd som prediktiv modell	69
Ensembler av beslutsträd	70
XGBoost och GBM	70
Boosting	70
Gradient boosting	71
Artificiella neuronnät	72
Referenser	75
Appendix C	76
Appendix D	77
Vidare reflektioner kring användning av kvantitativa data inom Afrikaprojektets analyser	77
Användning av kvantitativa data för att förbättra enfallsstudier	77
Användning av kvantitativa data för att förbättra prediktion i regionala studier	79
Användning av kvantitativa data för långsiktiga, prediktiva studier	80
Referenser	81
Appendix E	82
Kodning av publikationer inom Afrikaprojektet 2014-2019	82

Akronymordlista

ACLED	Armed Conflict Location & Event Data (databas)
ANN	Artificiellt neuronnät (eng. <i>Artificial neural network</i>)
AUC-PR	Area under Curve – Precision Recall (prestationsmått)
AUC-RO.	Area under Curve – Receiver Operating Characteristic (prestationsmått)
AUPR	Area under Precision Recall (Curve) (prestationsmått)
AUPRC	Area under Precision Recall Curve (prestationsmått)
AUROC	Area under Receiver Operating Characteristic (Curve) (prestationsmått)
CAMEO	Conflict and Mediation Event Observations (händelseklassificering)
COW	Correlates of War (projekt, databas)
CRAM	Conflict Risk Assessment Monitor (prediktivt system)
DARPA	Defense Advanced Research Projects Agency
EBMA	Ensemble Bayesian Model Averaging (metod)
FN	False negative
FP	False positive
FPR	False positive rate (prestationsmått)
GBM	Gradient Boosted Machine (metod)
GDELT	Global Database of Events, Language, and Tone (databas)
HCSS	Hague Centre of Strategic Studies
ICEWS	Integrated Crisis Early Warning System (prediktivt system)
MCC	Matthews Correlation Coefficient (prestationsmått)
PREVIEW	PREDiction VISualization Early Warning (projekt)
PRIO	Peace Research Institute Oslo
PVRAM	Political Violence Assessment Monitor (prediktivt system)
TERRIER	Temporally Extended, Regular, Reproducible International Event Records (databas)
TN	True negative
TP	True positive

TPR	True positive rate (prestationsmått)
UCDP	Uppsala Conflict Data Program (projekt, databas)
UCDP-GED	UCDP-Georeferenced Event Dataset (databas)
ViEWS	Violence Early Warning System (projekt, prediktivt system)
XGBoost	eXtreme Gradient Boost (metod)

1 Inledning

Vi lever sedan några decennier i en teknologisk revolution driven av en ständigt ökande beräknings- och datalagringskapacitet. Under det senaste årtiondet har denna utveckling börjat få genomslag inom säkerhetspolitisk analys, även om statistisk bearbetning av säkerhetspolitiska frågeställningar har en betydligt längre historia.

I denna rapport kommer vi att belysa några utvalda projekt som exemplifierar denna utveckling samt ge en introduktion till de datalogiska verktyg som utgör kärnan i en ny generation analytiska system. Genom bruk av detaljerad konflikthistorik tillsammans med en rad olika typer av data – i synnerhet politiska och ekonomiska – har kvalificerade modellbaserade konfliktprediktionssystem realiserats i (åtminstone) USA, Nederländerna och Sverige.

För att tydliggöra vilken möjlig påverkan dessa system kan ha på FOI:s analysverksamhet har vi också undersökt bruket och behovet av prediktiv analys inom myndighetens Afrikaprojekt.¹

1.1 Syfte

Ambitionen med denna rapport är att underlätta en diskussion, både inom FOI och mellan FOI och myndighetens uppdragsgivare, kring nyttan med dagens datadrivna analysmöjligheter inom det säkerhetspolitiska området. Genom att öka medvetenheten om fältets nuvarande tillstånd, möjligheter och begränsningar, är förhoppningen att denna rapport ska möjliggöra strategiska ställningstaganden hos lämpliga beslutsfattare kring, i första hand, användning av datadrivna och modellbaserade prediktiva system.

¹ FOI:s forsknings- och studieverksamhet avseende Afrika har pågått i över 10 år med nationella uppdragsgivare såsom Försvarsdepartementet, Utrikesdepartementet och Försvarsmakten, men även med internationella uppdragsgivare såsom FN. För mer information, se www.foi.se/afrika.

1.2 Frågeställningar

De frågor vi söker svar på är de följande:

1. Hur är systemen ICEWS, CRAM, PVRAM och ViEWS uppbyggda, och vilka statistiska och datalogiska tekniker har man använt i dem?
2. Vad vet vi om dessa systems prediktiva kapacitet?
3. Kan FOI:s Afrikaprojekt redan idag nyttja prediktioner av detta slag i sin analysverksamhet?
4. Vilka händelser, som i dagsläget inte är föremål för motsvarande prediktioner, skulle det vara av störst nytta för Afrikaprojektet att få analyserade av ett prediktivt system?
5. Är FOI en lämplig konstruktör och ägare av denna typ av system?

1.3 Målgrupp

I detta arbete riktar vi oss till tre huvudsakliga grupper:

1. Chefer, analytiker och forskare vid FOI, i synnerhet inom det säkerhetspolitiska området
2. Intresserade läsare vid Regeringskansliet
3. Intresserade läsare vid andra svenska myndigheter

1.4 Läsanvisningar

Efter denna inledning ger vi, i kapitel 2, en snabb överblick över några av rapportens centrala teman, däribland den form av prediktion som studeras och bruket av de studerade systemen.

I kapitel 3-5 beskrivs sedan de fyra utvalda prediktiva systemen, de datalogiska och matematiska tekniker de gör bruk av, samt något om deras kända prestanda.

Kapitel 6 behandlar arbetet inom FOI:s Afrikaprojekt och innehåller också en diskussion kring den möjliga nyttan med prediktiva system i säkerhetspolitisk analys.

I kapitel 7 diskuteras förutsättningarna för FOI att utveckla praktiskt användbara prediktiva system under de närmaste 1-2 åren.

Avslutningsvis presenteras arbetets slutsatser (kapitel 8), referenser och ett femdelat fördjupande appendix.

2 Bakgrund

I detta kapitel ger vi läsaren en introduktion till bruket av automatiserad prediktion inom säkerhetspolitik samt vissa nyckelbegrepp och centrala frågeställningar inför de efterföljande kapitlen.

2.1 Exempel på användning av säkerhetspolitiskt inriktade prediktiva system

De tre system som figurerar i denna rapport kan, förenklat, sägas tillhöra tre olika domäner: den militära, den politiska och den akademiska.

ICEWS-systemet utvecklades av företaget Lockheed Martin med forskningsmedel från den amerikanska statliga försvarsforskningsfinansiären DARPA och har därefter tagits i bruk vid några amerikanska militärkommandon, s.k. *combatant commands* eller COCOMs.^{2,3}

Hague Centre for Strategic Studies (HCSS) är en nederländsk tankesmedja som i första hand använder sina prediktiva system – CRAM och PVRAM – själva i samband med att man författar rapporter åt de nederländska utrikes- och försvarsdepartementen.⁴

ViEWS är ett akademiskt projekt vid Uppsala universitets institution för freds- och konfliktforskning. Till de målgrupper för prediktionerna från ViEWS som har diskuterats hör NGO:er, nationella regeringar och FN.⁵

Mot slutet av projektperioden har vi också uppmärksamats på att det tyska utrikesdepartementet driver ett utvecklingsprojekt vid namn PREVIEW inriktat mot datadriven prediktion och nulägesuppfattning.⁶

Vårt projekt har inte syftat till att skapa någon komplett bild av bruket av denna typ av system. Den begränsade genomgången till trots kan det vara värt att

² O'Brien (2010).

³ Se t.ex. <https://www.conseil-vista.com/>.

⁴ Personlig kommunikation med HCSS.

⁵ Hegre et al. (2019).

⁶ "Krisenfrüherkennung, Konfliktanalyse und Strategische Vorausschau", <https://www.auswaertiges-amt.de/de/aussenpolitik/themen/krisenpraevention/-/2238138>.

konstatera att vi inte sett något material under arbetets gång – akademiskt eller annat – som indikerar att andra aktörer på hög politisk eller militär nivå har utvecklat motsvarande typ av prediktiva system, även om det givetvis är möjligt att så är fallet. Bedömningen är därför att vi fortfarande befinner oss på ett tidigt stadium avseende införandet av avancerad modellbaserad prediktion inom det säkerhetspolitiska området, globalt sett.

2.2 Numerisk probabilistisk prediktion

När vi i fortsättningen talar om prediktion (och prediktiva system) så menar vi en verksamhet där man kopplar sannolikheter (därav *probabilistisk*) uttryckta med siffror (därav *numerisk*) till framtida händelser. För detta används normalt antingen procenttal, t.ex. ”Sannolikheten för regn är 32 %”, eller decimaltal mellan 0 och 1, t.ex. ”Sannolikheten för regn är 0,32”.

Andra sätt att uttrycka en prediktion kan vara med hjälp av *verbala* sannolikheter – så som ”mycket osannolikt”, ”troligt” eller ”oväntat” – eller med någon form av rangordning.

Man kan också frånga sannolikheterna helt och göra helt oreserverade uttalanden om framtiden, t.ex. ”Det kommer att regna imorgon”. I det fallet skulle man kunna tala om *icke-probabilistisk* prediktion.

Vi kommer i denna rapport varken att gå in på sannolikhetsteori eller de olika förtjänster (och problem) som olika sätt att uttrycka prediktioner har, med ett undantag: när numeriska sannolikheter används inom prediktion öppnar det möjligheter för ingående och noggrann utvärdering av prestation hos såväl personer som system. Detta är avgörande om vi vill jämföra olika sätt att ta fram prediktioner med varandra. Som vi ska se kan det dock vara utmanande att jämföra olika prediktiva prestationer även i de fall då sannolikheterna uttrycks numeriskt.

2.3 Modellbaserad, datadriven eller både och?

Titeln på denna rapport innehåller uttrycket ”modellbaserad prediktion”, men det hade nästan lika gärna kunnat heta ”datadriven prediktion”.

Modellbaserad, i detta sammanhang, innebär att prediktionen tas fram genom explicita matematiska beräkningar, och att dessa beräkningar utförs med hjälp av en eller flera modeller som har tagits fram specifikt för det fenomen man studerar. Dessa matematiska modeller kan vara av många olika slag, och processen att ta fram en modell skiljer sig ibland kraftigt mellan olika modelltyper. I genomgångarna av de tre systemen kommer vi att förklara de

olika modelltyperna översiktligt. För den intresserade finns viss fördjupning i appendix.

De system vi har studerat är datadrivna på åtminstone två olika sätt. Det första av dessa sätt är att modellerna typiskt kräver en viss mängd (in)data för att kunna producera ett resultat. Det andra är att uppbyggnaden av de modeller som används kräver en viss (ofta mycket stor) mängd data.

Dessa två aspekter av att vara datadriven följs ofta åt, men det finns undantag. Det går att bedriva datadriven prediktion utan explicita matematiska modeller, t.ex. när en människa tar in stora mängder data och omsätter det till en prediktion. Det går också att bedriva modellbaserad prediktion som inte är datadriven i vanlig mening, t.ex. när man konstruerar en modell helt utifrån teoretiska hänsyn och sedan genomför simuleringar (där man genererar sina egna indata).

2.4 Manuell och automatisk insamling av händelsedata

För att tillgodose behovet av data i *konstruktionen* av modeller krävs åtminstone en uppsättning data avseende det fenomen man vill predicera. För de system vi har studerat har detta utgjorts av en särskild typ av händelsedata: konfliktdata.

Det är alltid vanskligt att försöka återge ett akademiskt områdes historik, men Ward pekar på Richardsons bok *The Statistics of Deadly Quarrels* från 1960 som startskottet för en serie initiativ att dels systematiskt samla in data om krig, dels försöka förstå – och även förutsäga – deras förekomst, upptakt, längd, slut och spridning. Två tongivande projekt inom detta fält är Correlates of War, eller COW, som började 1963 och Uppsala Conflict Data Program, eller UCDP, som började 1979.^{7, 8, 9} Båda är datainsamlingsprogram som sedan har legat till grund för ett stort antal forskningsartiklar kring konflikters dynamik, vissa av dem med en prediktiv ansats.¹⁰

⁷ Ward et al. (2013a).

⁸ Diehl (1992).

⁹ Bussman, Dorussen & Gleditsch (2014).

¹⁰ Ward et al. (ss. 476-477, 2013a).

Både UCDP och COW, liksom det nyare dataprojektet Armed Conflict Location & Event Data, eller ACLED¹¹, är exempel på *manuellt kodade* händelsedatabaser. Det innebär att databaserna utökas genom att människor systematiskt läser texter från utvalda källor och sedan upprättar databasposter utifrån ett noga specificerat kodningssystem. UCDP och COW är i första hand databaser för väpnade konflikter, medan ACLED innehåller en bredare palett av händelser.

I kontrast till detta står *automatisk kodade* (eller *automatkodade*) databaser, där det istället är mjukvara som står för textbehandlingen och extraheringen av de element (typiskt händelsebeskrivningar) som man vill ha ut ur grundmaterialet, ofta nyhetsartiklar.¹² Trots att automatkodning av händelsedata har studerats åtminstone sedan 90-talet så tycks strategin inte vara något som har spritt sig till centrala europeiska projekt som UCDP:s eller Peace Research Institute Oslos (PRIO:s) olika datainsamlingsinitiativ¹³, utan i vår genomgång är det fortsatt (i huvudsak) i USA som vi har sett såväl utveckling som användning.

Frågan om vilken typ av data – automatkodad eller manuellt kodad – man vill använda vid konstruktion av modeller handlar till stor del om en avvägning mellan kvalitet och kvantitet. Automatiska system kan generera mycket stora databaser med stor hastighet, men har länge haft svårt att konkurrera med manuellt kodade databaser avseende kvalitet. Detta kvalitetsgap har minskat med tiden, men det är i dagsläget inte helt klart om något existerande automatkodande system når ända upp till mänsklig förmåga.¹⁴ Med detta sagt så är det inte heller trivialt att på förhand bedöma hur god kvalitet på händelsedata som krävs. I våra fyra studerade system har man i tre av dem valt att använda automatkodade data, medan man i det fjärde använder manuellt kodade data.

Om vi går bortanför konstruktionen av modeller och istället tittar på vilka data som modellerna kräver för att göra sina prediktioner så finns det en tydlig slutsats att dra: ju snabbare uppdateringshastighet som efterfrågas, om modellen kräver händelsedata, desto starkare blir argumentet för automatkodade databaser. Som exempel så uppdateras UCDP:s snabbaste databas – UCDP-Candidate –

¹¹ Raleigh et al. (2010).

¹² En centralfigur när det gäller automatkodade händelsedatabaser är Philip A. Schrodt som har lett eller varit delaktig i nästan alla större projekt inom området fram till några år sedan, däribland KEDS (Schrodt, Davis & Weddle, 1994), TABARI (Gerner et al., 2002), CAMEO (ibid.), ICEWS (O'Brien, 2010) och GDELT (Leetaru & Schrodt, 2013).

¹³ <https://www.prio.org/>.

¹⁴ Se t.ex. <https://dataverse.harvard.edu/file.xhtml?persistentId=doi:10.7910/DVN/28075/GBAGXI&version=22.0>.

månatligen, ACLED uppdateras veckovis och ICEWS-systemets automatkodade databas uppdateras i nära realtid.

2.5 Att mäta prediktiv prestation

En av de första sakerna att beakta när man talar om prestation i *probabilistiska* prediktiva sammanhang är att ”en gång är ingen gång”. Att en person eller ett system har satt en hög sannolikhet på något som sedan händer eller en låg sannolikhet på något som inte händer – vilket är bra – eller det omvända – vilket är dåligt – vid ett isolerat tillfälle ger oss ytterst lite att gå på. Det är först när många prediktioner har gjorts som vi kan börja urskilja tur från skicklighet, och otur från brist på skicklighet. Av den anledningen är det delvis missriktat att fråga sig t.ex. hur systemet/personen förutsåg den arabiska våren/Rysslands annektering av Krim/folkmordet i Rwanda. Även om prediktionerna avseende en handfull betydelsefulla händelser skulle vara korrekta (eller inkorrekta), så kan systemet (eller personen) helt enkelt haft tur (eller otur) de gångerna.

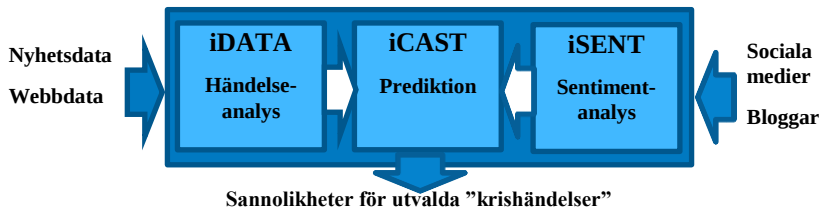
Det finns, som kommer att illustreras i de följande kapitlen, många sätt att mäta prediktiv prestation, eller snarare olika *aspekter* av prediktiv prestation. Vidare är vissa sätt mer eller mindre användbara beroende på vilka händelser vi vill försöka förutsäga. De system vi har studerat i denna rapport är inriktade mot prediktion av konflikter, och just konflikter är ett utmärkt exempel på detta. Problemet ligger i det faktum att konflikter, oavsett hur omvärlden och historien ibland kan upplevas, är ovanliga. Ett lätt sätt att lyckas förhållandevis väl med konfliktprediktion är att *alltid* sätta en väldigt låg sannolikhet, något som *i snitt* kommer att fungera bra. Samtidigt kan vi redan på förhand säga att vi med den strategin *aldrig* kommer att kunna förutsäga en konflikt, i meningen att vi har satt en hög sannolikhet på konflikt och detta sedan realiserar. Det finns lösningar på dessa utmaningar, men det understryker vikten av att välja ”rätt” sätt att utvärdera prestation.

En annan komplikation är att vad som är att betrakta som en *bra* prestation är områdesbundet. Anta t.ex. att vi har ett system som ger sannolikheter för finanskrascher, ett annat som ger sannolikheter för väpnade konflikter och ett tredje som ger sannolikheter för resultat i fotbollsmatcher. I det läget kan vi inte på något enkelt sätt jämföra dessa med varandra i prestationsavseende, utan vi får jämföra ett system för finanskrascher med andra system (eller personer) som förutsäger finanskrascher o.s.v. Det som däremot typiskt går att etablera, även i frånvaro av andra system att jämföra med, är om man slår slumpen (eller något trivalt tillvägagångssätt). *Hur* mycket bättre än slumpen som prediktioner inom ett område kan bli är sedan något som helt enkelt måste utforskas genom att man försöker.

3 ICEWS – militär förvarning

2007 initierade den amerikanska statliga försvarsforskningsfinansiären Defense Advanced Research Projects Agency¹⁵ (DARPA) ett projekt kallat Integrated Crisis Early Warning System¹⁶ (ICEWS) för att konstruera ett system med kapacitet att förutsäga politisk instabilitet.¹⁷ Av de forskargrupper som deltog i projektets första fas utsågs en vinnare som fick fortsätta med utvecklingen av sitt system: Lockheed Martin Advanced Technology Laboratories, eller LM-ATL. Arbetet med ICEWS var en fortsättning på ett projekt som hade gått under namnet PRESAGE och innebar att man kombinerade ett antal nya angreppssätt i ett sammanhållet system.¹⁸

ICEWS är ett system utformat för att beräkna sannolikheter för olika former av politisk instabilitet. Från ett tidigt fokus på Asien har systemet idag global täckning och producerar sannolikheter för ett antal utvalda händelser (se nedan) i 167 av världens 195 länder.¹⁹



Figur 1 Översikt över ICEWS

¹⁵ Enligt Wired spenderade DARPA USD 38M på projektet (<https://www.wired.com/2011/02/pentagon-predict-egypt-unrest/>).

¹⁶ Uttalas "i-queues" eller "IQs".

¹⁷ O'Brien (2010).

¹⁸ Predicting Stability through Analyzing Germane Events (<https://news.lockheedmartin.com/2007-10-24-Lockheed-Martin-to-Develop-Early-Warning-System-That-Predicts-Crisis-Events-and-Stability-of-Countries>).

¹⁹ Ward et al. (2013b).

Den variant av systemet som finns idag hos Lockheed Martin har fyra komponenter, varav tre används (se figur 1) för att skapa prediktioner: iData, iSent och iCast. Den fjärde, iTrace, ger visualiseringar av nuläge och trender.²⁰

3.1 Systemets fokus

Själva prediktionsmotorn finns i modulen **iCast**. Den producerar löpande prediktioner för följande händelser inom de kommande 6 månaderna²¹:

1. Inhemsk politisk kris (*domestic political crisis*): betydande motstånd mot regeringen, men inte på nivån av uppror eller revolt (exempelvis maktstrider mellan två politiska fraktioner som involverar betydande strejker eller våldsamma sammanstötningar mellan anhängare)
2. Uppror (*rebellion*): organiserat motstånd där målet är att söka autonomi eller självständighet
3. Revolt (*insurgency*): organiserat motstånd där målet är att störta centralregeringen
4. Etniskt eller religiöst våld (*ethnic or religious violence*): våld mellan etniska eller religiösa grupper och som inte är specifikt riktat mot regeringen
5. Internationell kris (*international crisis*): konflikt mellan två eller fler stater, eller höjd spänning mellan två eller fler stater vilken kan leda till konflikt

3.2 Händelsedata

Modulen **iData** analyserar i första hand nyhetsdata på engelska, spanska, portugisiska och arabiska, och har genererat en databas med över 25 miljoner händelser från år 1991 och framåt. Varje händelse har klassificerats som tillhörande en av drygt 300 händelsetyper från taxonomin Conflict and Mediation Event Observations (CAMEO).²² För att möjliggöra extraktion av händelser av typen ”vem gjorde vad mot vem?” ur nyhetsflödet har iData en förteckning av över 50 000 entiteter, t.ex. Sveriges regering, Kinas försvarsmakt eller ISIS.

²⁰ <https://www.lockheedmartin.com/en-us/capabilities/research-labs/advanced-technology-labs/icews.html>.

²¹ O'Brien (s. 90, 2010).

²² Gerner et al. (2002).

Databasen som iData har skapat var initialt endast tillgänglig för amerikanska myndigheter, men finns numera öppet publicerad på Harvard Dataverse.²³

Den andra datainsamlingsmodulen är **iSent**, vilken främst inhämtar material från sociala media. Likt iData extraherar iSent händelser, men istället av typen ”vem sa vad om vem?”.²⁴

3.3 Modeller

Prediktionerna produceras av modellensembler som drivs av data dels från iData, dels från en stor uppsättning andra datakällor, och över 80 olika modeller används i systemet. ICEWS ensembler innehåller med största sannolikhet fler modelltyper än de nedanstående, men dessa var de som diskuterades i O'Briens artikel från 2010.

3.3.1 Agentbaserad simulering

Agentbaserad simulering (ABS) är en datorbaserad beräkningsmetod för modellering av komplexa system, där enheter (t.ex. individer, grupper, organisationer eller system) modelleras som autonoma mjukvaruagenter som interagerar med varandra. ABS är en nerifrån-och-upp-metod (*bottom-up method*) där interaktioner mellan agenterna drivs av reaktioner eller autonoma handlingar enligt agenternas inbyggda regler (som inte nödvändigtvis är ekvations- eller sannolikhetsbaserade). ABS används ofta för att studera hur individuellt beteende på mikronivå kan producera strukturer och regelbundenheter på makronivå.²⁵

En av flera modeller som ingår i ICEWS är den agentbaserade simuleringsmodellen FactionSim.²⁶ FactionSim är en modell som gör det möjligt att skapa konfliktscenarier med olika grupperingar som har sina ledare, följare, gruppstillgångar och resurser. Dessa agenter agerar autonomt och följer sina regler och beslut på mikronivå baserade på vad situationen kräver. Makrobeteenden uppstår som ett resultat av interaktioner mellan dessa entiteter. Agenterna är komplexa och har värderingar (mål, standarder, preferenser, kulturella och etiska värden, och personlighet). Agenterna tillhör grupper som

²³ Harvard Dataverse, som drivs av Harvarduniversitetet, är en elektronisk lagringsplats för forskningsdata. Se <https://dataverse.harvard.edu/dataverse/icews>.

²⁴ O'Brien (2013).

²⁵ Holland & Miller (1991).

²⁶ Silverman et al. (2007).

har resurser, hierarkier av ledarskap och följare. Agenterna och de grupperna de tillhör upprätthåller dynamiska relationer med varandra. Relationerna utvecklas eller modifieras baserat på vilka händelser som utspelar sig. Simuleringen innefattar också samhälleliga företeelser som t.ex. institutionell korruption eller diskriminering. FactionSim syftar till att fånga det sociokulturella ”spel” som fokuserar på konkurrens mellan grupper om kontroll av resurser, t.ex. ekonomiska och politiska resurser.²⁷

En annan modell som ingår i ICEWS är PS-I (Political Science-Identity), en agentbaserad simuleringsmodell som ursprungligen utvecklades av Ian Lustick och Vladimir Derhachev för att operationalisera, förfinas och testa konkurrerande versioner av konstruktivistisk identitetsteori.²⁸ Huvudsyftet med denna plattform har varit att tillgängliggöra en användarvänlig beräkningsmiljö för att utforska politiska fenomen för politiska och andra samhällsforskare utan programmeringsbakgrund. Agenter med en repertoar av identiteter interagerar i ett bestämt område och påverkas av värden knutna till särskilda identiteter. Dessa värden förändras över tid och simulerar därmed förhållanden där individer kan uttrycka latenta identiteter eller anskaffa sig nya identiteter.²⁹

3.3.2 Regression

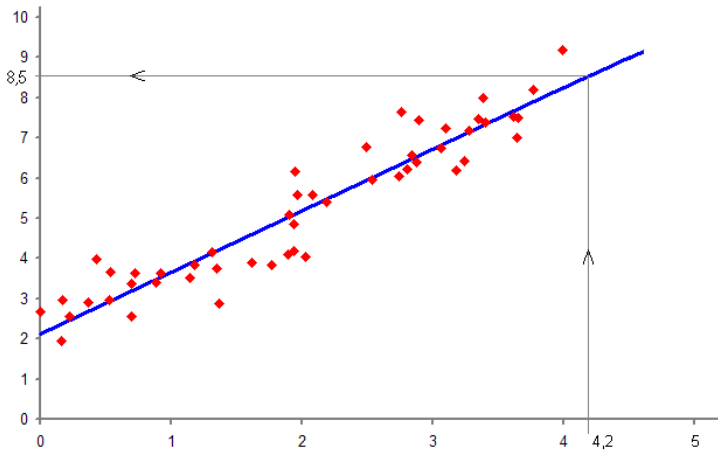
I tre av rapportens fyra studerade system förekommer regressionsmodeller, och då i form av *logistisk regression*. Regressionsanalys (eller regression) är en process där man uppskattar ett samband mellan en beroende variabel, t.ex. regnmängd inom 6 timmar, och en eller flera oberoende variabler, t.ex. lufttryck, luftfuktighet och molngeometri. Detta kan göras för att skapa en förståelse av hur värdet på den beroende variabeln ändras när var och en av de oberoende variablerna ändras. Har man tagit fram en regressionsmodell kan den användas för att predicera värdet på den okända beroende variabeln utifrån de oberoende variablerna.

En enkel men ändå kraftfull regressionsmodell som används flitigt i många tillämpningar är linjär regression där man modellerar sambandet mellan den beroende variabeln och de oberoende variablerna med en linjär funktion (se figur 2).

²⁷ Bharathy & Silverman (2012).

²⁸ <http://ps-i.sourceforge.net>.

²⁹ Lustick, I. (2002).



Figur 2 Linjär regression. De 50 röda datapunkterna illustrerar sambandet mellan en oberoende variabel (den horisontella axeln) och en beroende variabel (den vertikala axeln). Den blå linjen är en uppskattad linjär regression som uttrycker en trolig relation mellan variablerna. En prediktion för värdet 4,2 på den oberoende variabeln ger resultatet 8,5 för den beroende variabeln.

Logistisk regression är som alla andra regressionsanalyser en prediktiv analys som används för att beskriva data och för att förklara förhållandet mellan en beroende binärvariabel (som endast kan anta två möjliga värden, typiskt 0 eller 1) och en eller flera oberoende variabler. Medan man i linjär regression modellerar de förväntade värdena för den beroende variabeln med en kombination av värden på de oberoende variablerna modellerar man i logistisk regression *sannolikheten* eller *oddset* för att den beroende variabeln antar ett visst värde baserat på en kombination av värden av de oberoende variablerna.

En av forskarna inom ICEWS, Michael Ward, inriktade sig under projektet särskilt mot att använda spatiala (rumsliga) förhållanden och olika typer av relationsförhållanden i regressionsmodeller.³⁰

³⁰ Se t.ex. Ward & Gleditsch (2002) och Hoff & Ward (2004).

3.3.3 Sammanvägning av ensembler

Under projektet provade man två olika metoder för att aggregera utfallen från sin mångfald av modeller.³¹ Den första av dessa var ett s.k. *bayesianskt nätverk*³², men man valde senare att byta ut detta mot Ensemble Bayesian Model Averaging (EBMA).³³

En förenklad beskrivning av EBMA är att man, utgående från en uppsättning modeller som alla ger sannolikheter för samma fenomen, viktar varje modell efter prestation och unicitet. Det innebär att modeller som gör bra ifrån sig under metodens utvärderingsperiod får mer påverkan på den sammanvägda sannolikhetsvärderingen. Modeller som ofta är ensamma om att göra bra ifrån sig får också högre vikt. Om flera modeller gör bra ifrån sig i samma situationer kan de få en hög vikt som grupp betraktad, men varje enskild välpresterande modell får ett kompensande ”straff” för att resultatet liknar de från andra modeller.

Den effekt som man ofta ser, och alltså hoppas uppnå med sammanvägning av modellensembler, är att ensemblens prediktioner är bättre än var och en av de enskilda modellernas prediktioner.

3.4 Prestation

När det gäller kapaciteten hos existerande system av den här typen så är det inom ViEWS-projektet och HCSS CRAM vi har sett de mest omfattande och uppdaterade utvärderingarna. Lockheed Martin har däremot inte publicerat några ingående prestationssiffror på länge. Innan vi går in på de siffror som ändå *har* publicerats så behöver vi titta på det som på engelska kallas en *confusion matrix*.

3.4.1 Förvirringsmatrisen och relaterade mått

En vanligt använd uppdelning vid utvärdering av bedömningar i allmänhet, och prediktioner i synnerhet, är nedanstående matris där man ställer bedömningar (där P är ett påstående) mot faktiska sakförhållanden:

³¹ Wedgwood, Ruvinsky & Siedlecki (2012).

³² Mahoney et al. (2011).

³³ Montgomery, Hollenbach & Ward (2012).

Tabell 1 Förvirringsmatris (confusion matrix)

		SAKFÖRHÅLLANDE	
		P är sant	P är falskt
BEDÖMNING	P är sant	SANT POSITIV TP (true positive)	FALSKT POSITIV FP (false positive) Typ I-fel
	P är falskt	FALSKT NEGATIV FN (false negative) Typ II-fel	SANT NEGATIV TN (true negative)

Detta brukar alltså kallas en *confusion matrix* (se Tabell 1), och på svenska används ibland översättningen ”förvirringsmatris”. Om man har en uppsättning bedömningar så räknar man helt enkelt hur många som hamnar i varje kategori och skriver in antalen i matrisen. Från dessa värden kan man sedan beräkna en mängd olika förhållanden som berättar om olika aspekter av en prediktiv prestation. De som vi kommer att lyfta fram särskilt är de följande (förkortningarna TP, FP, FN och TN återfinns i tabellen ovan):

Tabell 2 Fyra vanliga andelsbaserade prestationsmått

Andelen faktiska ”positiva” som identifieras	$\text{True positive rate (TPR), Recall (R)} = \frac{TP}{TP + FN}$
Andelen ”positiva” bedömningar som är korrekta	$\text{Precision (Pr)} = \frac{TP}{TP + FP}$
Andelen faktiska ”negativa” som (fel)bedöms som positiva	$\text{False positive rate (FPR)} = \frac{FP}{FP + TN}$
Andelen korrekta bedömningar överlag	$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP}$

I en vanlig förvirringsmatris är det som sagt *antal* korrekta och felaktiga bedömningar som skrivs in i matrisen, men när det handlar om probabilistiska bedömningar så finns det inget okontroversiellt sätt att välja i vilken cell i matrisen en given bedömning ska hamna. En probabilistisk bedömning som är skild från 1 och 0 är aldrig *helt* korrekt eller *helt* felaktig. Vi kan se problemet lättare med ett exempel.

Anta att vi ska klassificera en fisk i någon av följande kategorier: Påfågelciklid, Torsk, Nilabborre eller Annan. Om vi för varje klassificerad fisk anger sannolikheten för varje kategori, t.ex. $\{p(\text{Påfågelciklid}); p(\text{Torsk}); p(\text{Nilabborre}); p(\text{Annan})\} = \{0,36; 0,40; 0,21; 0,03\}$, så har vi inte gjort något entydigt val av kategori; alla klassificeringar är fortsatt möjliga. Genom att använda olika tumregler, t.ex. att den kategori med högst sannolikhet får bli ”den valda”, så är det möjligt att skapa en traditionell matris, men inget sätt är invändningsfritt.

En annan väg att gå är att skapa en probabilistisk förvirringsmatris, något som också kan göras på flera sätt. En sådan är en snittsannolikhetsmatris.

Tabell 3 Probabilistisk förvirringsmatris

		SAKFÖRHÅLLANDE			
		A	B	C	D
SANNOLIKHETS BEDÖMNING (SNITT)	p(A)	0,45	0,20	0,03	0,14
	p(B)	0,12	0,66	0,30	0,02
	p(C)	0,18	0,03	0,39	0,05
	p(D)	0,25	0,11	0,28	0,79

Ju högre sannolikheter längs diagonalen, desto bättre. Med fler kategorier än två blir distinktionen positiv/negativ inte lika rättfram, men man kan konstruera probabilistiska motsvarigheter till t.ex. TPR och FPR.

3.4.2 Brierpoäng

Brierpoäng, ibland kallade *sannolikhetspoäng*, utgörs av det kvadratiska felet hos en sannolikhetsbedömning.³⁴ Om en händelse inträffar så hade en (i efterhand) perfekt bedömning varit sannolikhet 1 för just den händelsen. Allt under det innebär ett visst fel: differensen mellan 1 och den bedömda sannolikheten. Brierpoängen är alltså kvadraten av detta fel, så en bedömd sannolikhet på 0,68 för en inträffad händelse ger ett fel på $1 - 0,68 = 0,32$, och Brierpoängen blir följaktligen $0,32^2 = 0,1024$. Brierpoäng används typiskt i två varianter där 0 är den bästa poängen för båda, och där 1 respektive 2 utgör de sämsta möjliga utfallen. För mer om Brierpoäng, se Appendix A.

3.4.3 Värdet

Lockheed Martin anger på sin hemsida en sammanslagen *accuracy* på över 80 procent. I en presentation från 2015 anges att systemet har en *accuracy* på över 90 procent och under 20 procent falska positiva.³⁵ Det kan dock vara problematiskt att använda *accuracy* som mått när man studerar ovanliga händelser. Anledningen till det är – som har nämnts tidigare – att faktiska negativa kommer att stå för merparten av alla händelser. I det fallet kan även en trivial prediktionsstrategi, som säger att *inga* positiva fall (krig, uppror etc.) kommer att inträffa, få en hög TN (sant negativa). Detta kan i sin tur leda till ett förhållandevis högt accuracymått.

För att hitta akademiskt publicerade mått måste vi gå tillbaka ända till O'Brien (2010) där modellen testades *out-of-sample*³⁶ på data för 2005-2006, med kvartalsvisa prediktioner. Ensemblens Brierpoäng för *uppror* angavs till strax under 0,2. Vidare beräknades *accuracy*, *recall* och *precision* för sex olika händelseklasser: instabilitetsindex³⁷, uppror, revolt, etniskt/religiöst våld, inhemsk politisk kris och internationell politisk kris. *Accuracy* låg över 0,8 för alla klasser, *recall* över 0,8 för alla utom inhemsk politisk kris, och *precision* över 0,8 för alla utom inhemsk och internationell kris.

³⁴ Brier (1950).

³⁵ "Integrated Crisis Early Warning Systems (ICEWS)", <https://www.lockheedmartin.com/en-us/capabilities/research-labs/advanced-technology-labs/icews.html>.

³⁶ Detta innebär att modellen testas på en datamängd som inte har använts i modelleringen, en s.k. *hold-out sample*. På detta sätt kan man få en uppskattning av modellens kapacitet när den används på nya data.

³⁷ Anges i O'Brien (2010) till 1 = ingen konflikt, 2 = kris, 3 = våldsam kris och 4 = krig.

4 CRAM och PVRAM – strategiska underlag till departementen

Det nederländska forskningsinstitutet Hague Centre for Strategic Studies (HCSS) har utvecklat flera s.k. ”monitorer” för visualisering av olika typer av risk på global nivå. Två av dessa är Political Violence Risk Assessment Monitor (PVRAM) och Conflict Risk Assessment Monitor (CRAM), vilka båda är inriktade på förekomsten av systematiskt våld. Båda dessa har utvecklats inom ramen för ett program kallat StratMon som producerar strategiska bedömningar av globala risker på uppdrag av flera nederländska departement.

4.1 Systemens fokus

De två av HCSS system som vi tar upp nedan skiljer sig åt på en del viktiga punkter, även om inriktningen är snarlik.

4.1.1 Political Violence Risk Assessment Monitor

PVRAM (se figur 3) uppdateras dagligen med en prediktion för den kommande månaden. Systemet ger endast sannolikheter för länder som inte redan är i konflikt, så fokus ligger på inledningen av konflikter och oroligheter. En konsekvens av detta är att systemet inte kan användas (i dagsläget) för att få sannolikheter för avslut av konflikter.

Sannolikheterna från systemet avser alltså att minst en händelse tillhörande någon av följande kategorier (med definitioner från Political Instability Task Force, eller PITF³⁸) ska inträffa inom den kommande månaden:

1. Etniska krig: episoder av våldsam konflikt mellan regeringar och nationella, etniska, religiösa eller andra samhälleliga minoriteter (etniska utmanare) i vilken utmanaren söker omfattande förändringar av sin status
2. Revolutionskrig: episoder av våldsam konflikt mellan regeringar och politiskt organiserade grupper som avser att avsätta centralregeringen, att ersätta dess ledare eller att ta makten i en region
3. Folkmord och politiscid: händelser som involverar främjande av, genomförande av och/eller implicit medgivande till en uthållig politik

³⁸ ”INSCR Data Page”, <http://www.systemicpeace.org/inscrdata.html>.

- från styrande eliter eller dessas företrädare vilken resulterar i att betydande delar av en samhällsgrupp eller politiserad grupp dödas
4. Skadliga regimskiften: omfattande, skadliga skiften i styrningsmönster, däri inkluderat omfattande och plötsliga skiften bort från mer öppna, elektorala system till mer slutna, auktoritära system; revolutionära skiften inom politiska eliter och styrelseskick; omstridd upplösning av federerade stater eller utbrytning av ett större område med utomrättsliga medel; samt total eller nära total kollaps av centralmaktens auktoritet och möjlighet att styra

Täckningen är global och den geografiska upplösningen ligger på landsnivå.



Figur 3 Användargränssnitt för PVRAM

För att en konflikt ska inkluderas i PITF:s databas krävs det att de stridande parterna har mobiliserat minst 1000 stridande var, samt att konflikten har resulterat i 1000 dödsfall, med ett minimum på 100 dödsfall under varje år.

4.1.2 Conflict Risk Assessment Monitor

Istället för PITF har HCSS här använt sig av Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP-GED) för sin beroende variabel; det man avser att förutsäga. Systemet anger sannolikheter för att det ska inträffa minst en händelse från minst en av följande kategorier³⁹:

³⁹ Petterson & Eck (2018).

1. Statsbaserad konflikt (state-based conflict): väpnad konflikt där minst en part är en nationell regering, exempelvis konflikt mellan två stater eller mellan en nationell regering och en rebellgrupp
2. Icke-statlig konflikt (non-state conflict): väpnad konflikt mellan organiserade grupper, men där ingen part är en nationell regering
3. Ensidigt våld (one-sided violence): väpnad konflikt mellan en organiserad grupp eller stat på den ena sidan och obehäpnade civila på den andra

Likt PVRAM har CRAM (se figur 4) global täckning, men har en högre geografisk upplösning: administrativ nivå 2, eller regioner.⁴⁰ Det finns också tre olika tidshorisonter: inom 1 månad, inom 6 månader och inom 12 månader. Till skillnad mot PVRAM ger CRAM prediktioner oavsett om ett område upplever en konflikt för närvarande eller ej.



Figur 4 Användargränssnitt för CRAM

4.2 Händelsedata

PVRAM, som byggdes först av de två systemen, använder automatkodade data från systemet Global Database of Events, Language, and Tone (GDELT).⁴¹ Med det nyare CRAM har HCSS breddat mängden källor och tar in automatkodade

⁴⁰ Administrativ nivå 1 motsvarar länder, medan administrativ nivå 2 motsvarar sådant som delstater eller, i Sveriges fall, län.

⁴¹ "The GDELT Project", <https://www.gdelproject.org/>.

händelsedata från ICEWS (se föregående kapitel), Phoenix⁴² och TERRIER⁴³ utöver GDELT och (manuellt kodade data från) UCDP-GED (se föregående avsnitt).⁴⁴

4.3 Modeller

PVRAM använde från början en ensemble av logistiska regressionsmodeller, men har med tiden gått över till att endast använda en uppdaterad variant av sin bästa regressionsmodell. För CRAM har HCSS istället valt att använda tre olika maskininlärningsmodeller vilka beskrivs i Appendix B: *random forest*, XGBoost och Gradient Boosted Machine. Systemet gör inte heller någon sammanvägning av modellerna i en ensemble, utan användaren kan välja mellan de tre modellernas olika prediktioner i användargränssnittet.

4.4 Prestation

Utöver mått kopplade till förvirringsmatrisen (se avsnitt 3.4.1) så har HCSS beräknat ett antal ytterligare: MCC, F-poäng, AUC-RO och AUC-PR. Dessa mått förklaras i Appendix A.

4.4.1 Värden

Vad avser CRAM har vi under projektiden fått ta del av prestationsmått från HCSS, men de finns inte öppet publicerade ännu. Då HCSS inte själva har gjort sina prestationsmått för modellerna öppet tillgängliga ännu har vi valt att inte inkludera systemets värden i denna rapport. Det som kan sägas är att måtten är i paritet med de andra system vi har studerat. För CRAM har HCSS också beräknat en stor bredd av prestationstal: AUC-RO, AUC-PR, F1, F2, F0.5, *precision*, *recall*, *specificity* och MCC.

För PVRAM, utvärderat 2017, var AUC-RO 0,881.⁴⁵ Författarna skriver själva att de bedömer värden över 0,85 som utmärkta, men som vi ska se i nästa kapitel så finns det bland vissa forskare en tvekan kring hur värdefullt AUC-RO-värdet är inom konfliktområdet.

⁴² Phoenix är ett händelsedataprojekt vid University of Dallas, se <http://eventdata.utdallas.edu/>.

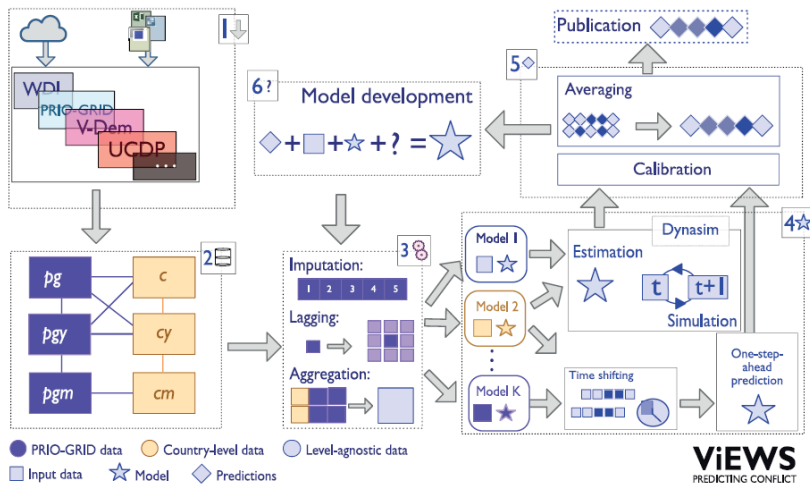
⁴³ Temporally Extended, Regular, Reproducible International Event Records, se www.terrierdata.org.

⁴⁴ Personlig korrespondens med HCSS.

⁴⁵ Rööß et al. (2018).

5 ViEWS – med sikte på humanitära insatser

Violence Early Warning System (ViEWS) utvecklas och drivs av Uppsala universitets institution för freds- och konfliktforskning med hjälp av forskningsmedel från EU.⁴⁶



Figur 5 Schema över ViEWS (Hegre et al., 2019)

Punkterna i schemat visar på de huvudsakliga momenten i projektets arbetsflöde:

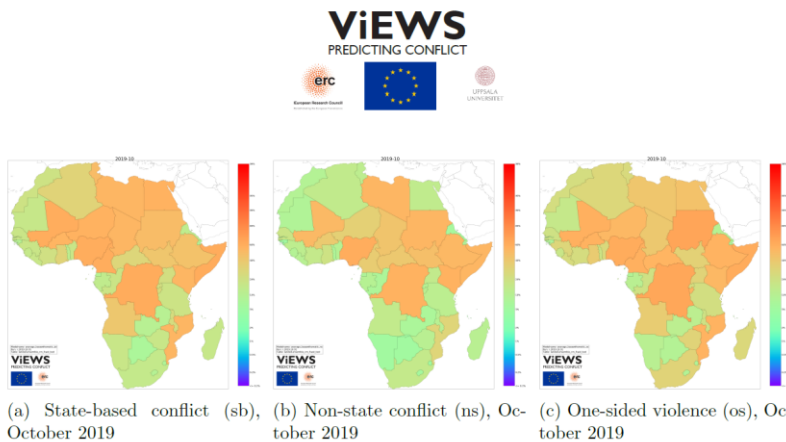
1. Insamling av ingångsdata
2. Bearbetning av data för olika tids- och rumsuppdelningar
3. Bearbetning av data genom komplettering, sammanslagning och förskjutning i tid
4. Modellbaserad prediktion med enskilda modeller
5. Sammanvägning av modellernas resultat
6. Vidareutveckling av prediktionsmodellerna

⁴⁶ Projektet startade i januari 2017 (<https://www.pcr.uu.se/research/views/>). Enligt CORDIS (<https://cordis.europa.eu/project/rcn/205895/factsheet/en>) har ViEWS tilldelats knappt EUR 2,5M för perioden fram till 2021.

5.1 Systemets fokus

Där ICEWS, CRAM och PVRAM alla har global täckning så fokuserar ViEWS istället på den afrikanska kontinenten. ViEWS gör prediktioner på två olika nivåer av geografisk upplösning och ger månatligen uppdaterade prediktioner för var och en av de följande 36 månaderna. De två geografiska nivåer som används är *landsnivån* (se figur 6) och *PRIO-gridnivån* (se figur 7). PRIO-grid är ett geografiskt rutnät som täcker hela jorden och där varje ruta är en halv grad bred och hög, något som vid ekvatorn motsvarar ca 55 x 55 km.⁴⁷

Tuesday 15th October, 2019



Figur 6 Exempel på prediktion för den kommande månaden på landsnivå. Färger närmare rött indikerar högre sannolikheter för våld, medan färger närmare blått indikerar lägre.⁴⁸

De händelser som ViEWS beräknar sannolikheter för är desamma som inom UCDP⁴⁹:

1. Statsbaserad konflikt (*state-based conflict*): väpnad konflikt där minst en part är en nationell regering, exempelvis konflikt mellan två stater eller mellan en nationell regering och en rebellgrupp

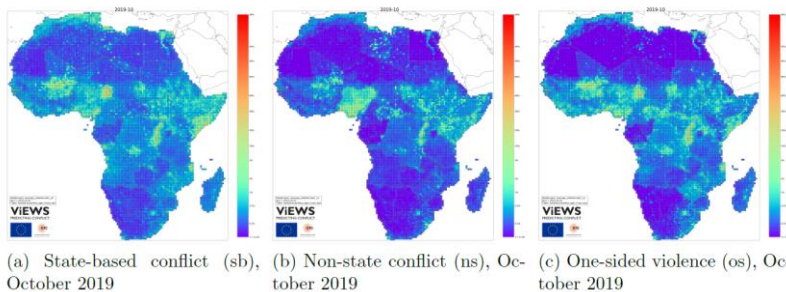
⁴⁷ Tollefsen et al. (2012).

⁴⁸ "ViEWS monthly forecasts, October 2019", <http://files.webb.uu.se/uploader/1576/MonthlyForecastsShort-October-2019.pdf>.

⁴⁹ Petterson & Eck (2018).

2. Icke-statlig konflikt (*non-state conflict*): väpnad konflikt mellan organiserade grupper, men där ingen part är en nationell regering
3. Ensidigt våld (*one-sided violence*): väpnad konflikt mellan en organiserad grupp eller stat på den ena sidan och obeväpnade civila på den andra

I alla tre händelsetyper går gränsen för väpnad konflikt vid minst 25 dödsfall under ett kalenderår.



Figur 7 Exempel på prediktion för den kommande månaden på PRIO-gridnivå⁵⁰

5.2 Händelsedata

Till skillnad från ICEWS har ViEWS i skrivande stund inte någon modul för automatkodning av händelsedata, utan använder sig av en kombination av de manuellt kodade databaserna UCDP-Candidate, UCDP-GED (se avsnitt 4.1.2) och ACLED (se avsnitt 2.4). I likhet med ICEWS används dock sedan en mängd andra datakällor för att komplettera händelsedata.

5.3 Modeller

2018 använde systemet 54 olika modeller vilka sedan sammanvägdes i de olika ensemblerna.⁵¹ Utöver logistisk regression (se avsnitt 3.3.2) och random forest

⁵⁰ "ViEWS monthly forecasts, October 2019", <http://files.webb.uu.se/uploader/1576/MonthlyForecastsShort-October-2019.pdf>.

⁵¹ Hegre et al. (2019).

(se Appendix B) så används två tekniker som beskrivs nedan.⁵² Först presenteras dock modellering och simulering mer allmänt.

5.3.1 Modellering och simulering

En modell är en representation av ett system eller en process i syfte att studera systemet och besvara någon eller några specifika frågeställningar om systemet under kontrollerade förhållanden. För de flesta studier räcker det att enbart ta hänsyn till de aspekter av systemet som är relevanta för det specifika problem som studeras och därför är modeller nästan uteslutande förenklade avbilder av systemet. Modeller kan vara fysiska eller logiska (matematiska). Exempel på fysiska modeller är småskaliga fysiska modeller som kan användas för att utvärdera konstruktioner av flygplan eller bilmodeller i vindtunnlar och fartygsmodeller i bassänger.

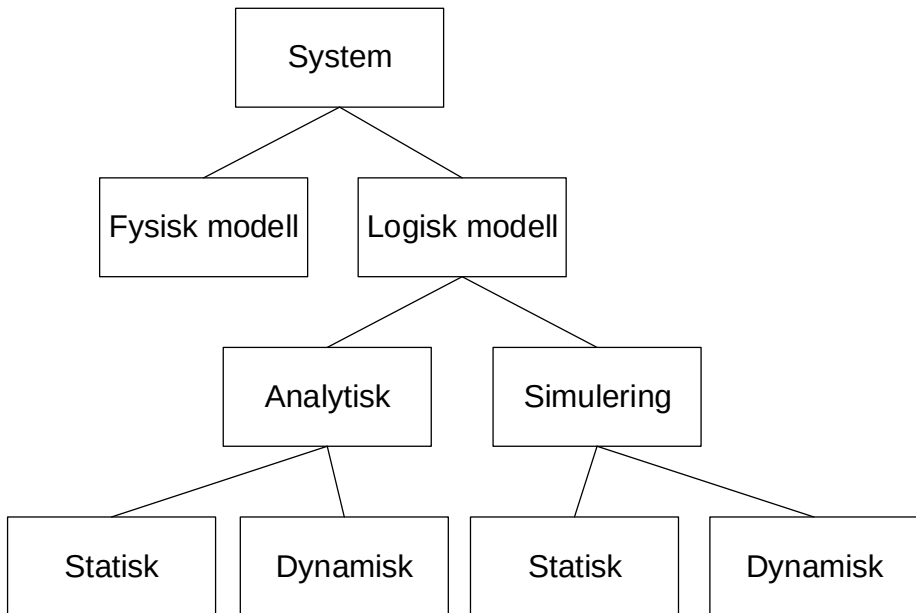
Logiska modeller representerar ett system genom kvantitativa förhållanden som är möjliga att manipulera och ändra för att se hur modellen reagerar. Dessa modeller specificerar viktiga komponenter och gör antaganden om och approximationer av hur systemet fungerar. Modellen behöver vara tillräckligt detaljerad för att tillåta giltiga slutsatser om det verkliga systemet, men behöver alltså inte vara en exakt återskapning av det ursprungliga systemet. Med andra ord är det inte en rikedom av detaljer som avgör kvaliteten av modellen. Framgångsrik modellering är en *för ändamålet giltig* förenkling av det verkliga systemet inom en viss experimentell ram.

Vissa logiska modeller går att studera analytiskt genom t.ex. differentialekvationer, köteori och linjär programmering, men i de flesta fall är modellerna så komplexa att simulering är det enda alternativet.

Simuleringsmodeller är en abstraktion av ett system med hjälp av logiska enheter, t.ex. datastrukturer som efterliknar systemets komponenter, en uppsättning instruktioner och begränsningar som definierar hur enheterna interagerar. Genom att utföra instruktionerna manipuleras enheterna i modellen, vilket öppnar för att studera hur systemet ändras under olika förhållanden.

Logiska simuleringar kan även delas in i statiska och dynamiska simuleringar. En schematisk figur över olika modellklasser kan ses i figur 8.

⁵² Man har också övervägt att använda artificiella neuronnät, se Appendix B.



Figur 8 Olika klasser av simuleringsmodeller

Statiska simuleringar saknar begreppet tid; omvänt är dynamiska modeller de som modellerar tid och tidsberoende egenskaper hos ett system. Inom statiska simuleringar används oftast s.k. *Monte Carlo-metoder*. I dessa förlitar man sig på upprepade slumpmässiga samplingar för att få numeriska resultat. En dynamisk simulering är istället baserad på en modell av tid som framskrider medan olika processer fortskrider och ändrar systemets tillstånd.

Ett exempel på en statisk simuleringsmodell är ett datorprogram som beräknar sannolikheten för att vinna ett kasinospel. Ett program som räknar ut den genomsnittliga kölängden vid en bankkassa givet fördelningar för kundernas ankomst och servicetid (d.v.s. den tid som krävs för att kunden hinner avsluta sitt bankärende) är istället ett exempel på en dynamisk simuleringsmodell.

Hegre et al. (2019) använder sig av två alternativa strategier för att skapa långsiktiga prognoser: dels en form av dynamisk simulering, vilken de kallar *dynasim*, dels en metod som huvudsakligen är baserad på regressionsanalys, vilken de har döpt till *ett-steg-fram* (eng. *one-step-ahead*). I de två följande avsnitten beskriver vi kort dessa två metoder.

5.3.2 Dynamisk simulering (dynasim)

Även om dynasim delar namn med dynamisk simulering som beskrevs ovan skiljer den sig på några punkter. Metoden går i korthet ut på att beräkna

prognoser för tidssteg efter den första månaden – en månad i taget från månad 2 till månad 36 – genom att probabilistiskt ”rulla fram” varje månads prediktion till den efterföljande. Från varje månatlig prediktion dras ett värde från den aktuella sannolikhetsfördelningen och används som ingångsvärde för nästkommande månads prediktion.

Anta att vi är vid tidpunkt t och är intresserade av att predicera N månader framåt. Med vår tränade modellensemble tar vi fram prognosen för den första månaden, $t+1$. För att producera prognoser för nästkommande månad, $t+2$, behöver man indata som ännu inte är tillgängliga och som dessutom själva är funktioner av nuvarande indata (t.ex. tid från senaste konflikt, avstånd till närmaste konfliktszon). För varje otillgängligt variabelvärde genereras sannolikhetsfördelningar för tiden $t+2$. Simulerade värden dras sedan från dessa fördelningar och registreras som ingående variabler för månadens ($t+2$) prediktion. Denna procedur upprepas många gånger för varje enskild månad i en s.k. *Monte Carlo-metod*. Med ett simulerat resultat för $t+2$ går man sedan vidare på samma sätt till $t+3$ och vidare till slutet av prognosfönstret ($t+N$).

5.3.3 Ett-steg-ram-modellering

I ett-steg-ram-modellering prediceras varje steg i framtiden oberoende av andra steg, i motsats till dynasim där de predicerade variablerna propageras sekventiellt. Detta görs genom att skapa en särskild statistisk modell – skild från modellensemblen för det studerade fenomenet – för varje steglängd.

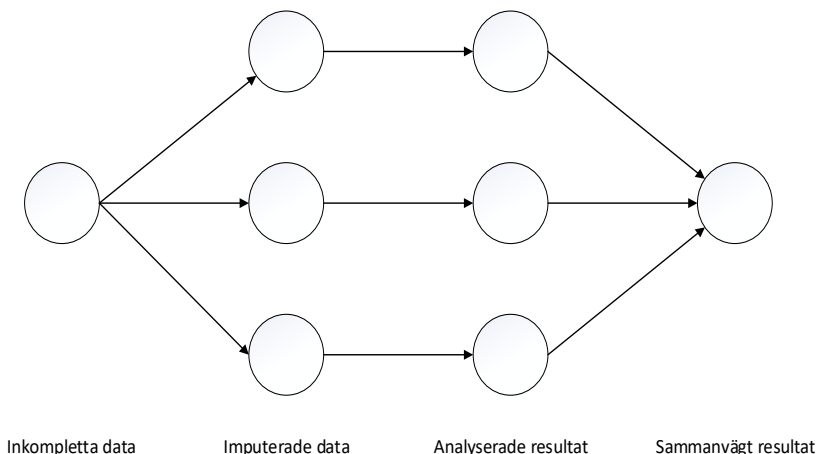
Om man exempelvis vill predicera för var och en av N månader in i prognosfönstret, tränar man N modeller för $t+1$, $t+2$, ..., $t+N$. Modellen för $t+10$ skulle då, t.ex., ta invariablerna för modellensemblen vid tiden t och producera en sannolikhetsfördelning över möjliga värden på invariablerna vid $t+10$. Sedan kan man använda dessa fördelningar för att göra en prediktion för månad $t+11$, med hjälp av sin ”vanliga” modellensemble.

5.4 Imputation: hantering av saknade data

Metoderna som ViEWS-projektet använder kräver att indata för förutsägelserna är kompletta. Om en variabel saknar värden för en period fungerar inte algoritmerna. Dessutom kan saknade data skapa snedvridning (bias) i modellernas prediktioner.⁵³

⁵³ Allison (2009).

För att hantera avsaknaden av data utför ViEWS s.k. *imputationer*, d.v.s. att man på olika sätt genererar värden för att ersätta de saknade. Mer specifikt ersätts saknade värden fem gånger med genererade värden, vilket skapar fem olika kompletta datauppsättningar. För dynasim används alla fem datauppsättningarna samtidigt och resultaten beräknas med Rubins regel.⁵⁴ På grund av tekniska utmaningar i ett-steg-fram-modelleringen används dock bara en av datauppsättningarna för de prognoserna.



Figur 9 En schematisk bild över huvudstegen i multipla imputationer där man skapar tre stycken kompletta datauppsättningar om indata saknas. Dessa tre uppsättningar analyserar man sedan, var och en för sig, för att sedan väga samman resultaten. I praktiken väljer man oftast ett större tal än tre för att få mer robusta uppskattningar; i ViEWS skapas fem datauppsättningar.

Figur 9 visar en schematisk bild över huvudstegen vid multipla imputationer. I första steget skapar man flera kompletta datauppsättningar genom att ersätta de saknade värdena med rimliga värden. På liknande sätt som i en del av metoderna ovan dras dessa värden från en sannolikhetsfördelning som är lämplig för den aktuella datauppsättningen. Således är de tre datauppsättningarna identiska avseende observerade data, men skiljer sig åt i de värdena som har lagts till. Storleken på dessa skillnader återspeglar vår osäkerhet om vilket värde som ska tillräknas.

Det andra steget är att utföra prediktionerna för varje given datauppsättning. Detta görs vanligtvis genom att använda den analytiska metoden som vi skulle ha

⁵⁴ Se Rubin (1987).

använt om uppgifterna var fullständiga. Resultaten kommer då att variera eftersom deras indata skiljer sig åt.

Det sista steget är att väga samman de olika uppskattningarna till en kombinerad uppskattning, och dessutom uppskatta variansen. Under lämpliga förhållanden kommer de kombinerade uppskattningarna att ha bättre egenskaper i flera avseenden, jämfört med vad situationen hade varit utan användning av imputation.

5.5 Prestation

Från ViEWS ger man tre nyckeltal: AUROC, AUPR (se Appendix A) och Brierpoäng (se avsnitt 3.4.2).⁵⁵

5.5.1 Värden

Inom ViEWS-projektet mäter man separat för var och en av sina tre händelseklasser: statsbaserat våld, icke-statligt våld och ensidigt våld. För att ha en prestationsmässig referenspunkt har ViEWS-projektet också mätt hur deras prediktioner står sig mot något som skulle kunna tolkas som en basnivåstrategi, nämligen att ansätta en konstant sannolikhet för var och en av händelseklasserna, på den relativa snittfrekvensen under systemets träningsperiod. Referensnivån, kallad *baseline* av ViEWS, är alltså en form av trendframskrivning där snittfrekvensen för den föregående perioden används som sannolikhet för den kommande. Slutligen anger ViEWS också prestation på sina två geografiska aggregeringsnivåer: land och PRIO-gridcell. Se Appendix C för en komplett tabell.

Något som enkelt går att utläsa är att ensemblerna ger förbättringar jämfört med *baseline* för alla kategorier. Man kan också se att AUROC-värdena ligger över 0,9 för alla utom en ensemble, och att AUPR-värdena för PRIO-gridmodellerna är klart lägre än motsvarande för landsnivån. Från ViEWS påpekar man att AUPR är betydligt svårare att höja än AUROC när det gäller konflikthändelser och rekommenderar att särskild vikt ska läggas vid detta mått.

⁵⁵ "How well do we do?", <https://www.pcr.uu.se/research/views/current-forecasts/how-well-do-we-do/>.

6 Teman inom FOI:s analyser avseende den afrikanska kontinenten

I detta kapitel diskuteras arbetet inom Afrikaprojektet och den eventuella efterfrågan på – och nyttan med – prediktiv analys i projektets framtida arbete. Som nämndes i rapportens inledning så har säkerhetspolitisk analys på den afrikanska kontinenten fungerat som rapportens referensområde för att resonera kring betydelsen av datadriven prediktion för FOI.

I de följande avsnitten sammanställs huvuddelen av Afrikaprojektets skriftliga produktion 2014-2019, för att utröna dels beställarens huvudsakliga intresseområden, dels hur studierna vanligtvis utformas vad gäller metod, data och prediktion. Mot bakgrund av detta analyseras därefter huruvida det vore möjligt och användbart att utforma prediktiva system inom Afrikaprojektet, samt inom vilket tematiskt område en pilotstudie lämpligast skulle genomföras.

6.1 Tematiska trender för Afrikaprojektets produktion 2014-2019

Som underlag för analysen har författarna gått igenom 63 skriftliga publikationer, vilket utgör merparten av Afrikaprojektets skriftliga leveranser under åren 2014-2019.⁵⁶ Av dessa utgörs 39 skrifter av kortare memon, samt 24 av längre så kallade R-rapporter. Dessa har kodats för att möjliggöra en sammanställning av tematik, metod, data, teoriansknytning och framförallt användningen av prediktion och existerande kvantitativa databaser.⁵⁷

Med hänsyn taget till att granskningen av varje enskild publikation av tidsskäl varit något översiktlig, kan vissa slutsatser likväl dras med relativ säkerhet.

För det första kan det konstateras att en överväldigande majoritet av studierna har en kvalitativ inriktning – endast en publicerad studie har en renodlat kvantitativ

⁵⁶ Sammanställningen av de granskade publikationerna har gjorts på grundval av Afrikaprojektets hemsida, där även publikationer inom insatsprojektet ingår. I den mån författarna känt till ytterligare studier av relevans har dessa inkluderats i analysen. Översikten gör inte anspråk på att täcka 100 procent av FOI:s Afrikarelaterade studier; däremot torde urvalet täcka en överväldigande majoritet av den skriftliga produktionen inom Afrikaprojektet under perioden 2014-2019.

⁵⁷ Den kompletta sammanställningen kan fås av författarna på begäran. Kodboken återfinns i Appendix E.

ansats, medan två studier använder blandad metodologi. Ett fåtal studier har enskilda inslag av kvantitativ analys. Ett exempel på en kvantitativ studie är Strategisk långsiktig omvärldsanalys Afrika (SLO) som genomförts 2019 på uppdrag av Försvarsmakten.⁵⁸ Den kvalitativa inriktningen på studierna kan i hög grad förklaras av de forskningsfrågor som ställts av beställaren i dialog med Afrikaprojektet.

Givet relativt god tillgång på både kompetens och data så finns det samtidigt en utvecklingspotential när det gäller att utforma mer kvantitativt inriktade studier eller att i större utsträckning använda kvantitativa databaser som komplement i studier som baseras på kvalitativ datainsamling. Fördelarna med blandad metodologi är generella, men extra påtagliga när det gäller prediktion. Kvantitativa data skulle kunna bidra med ökad precision i analysen – om den kvalitativa analysen exempelvis visar på att en ökad konfliktintensitet är sannolik, från vilken nivå utgår detta? Vidare, i den mån lärdomar kan dras inte bara bakåt i tiden, utan även mellan fall, skulle kvantitativa data möjliggöra och fördjupa denna typ av jämförelser. Förenklat uttryckt kan kvantitativa data förbättra både reliabilitet och validitet i prediktioner som till huvudsak baseras på kvalitativa data.

För det andra kan det konstateras att studierna primärt är beskrivande (drygt två tredjedelar).⁵⁹ Metodmässigt ligger tyngdpunkten på dokument- och litteraturstudier (två tredjedelar) snarare än intervjubaserade undersökningar.⁶⁰ Teorianknytningen är förhållandevis svag: drygt två tredjedelar av studierna saknar referenser till teorier eller begrepp.⁶¹

Det bör samtidigt betonas att ambitionsnivån i publikationernas analytiska ansats i hög grad påverkas av studiens omfång och de tidsramar som står till författarnas

⁵⁸ De studier som har kvantitativ karaktär är Lins de Albuquerque (2017); respektive Hull Wiklund & Ingerstad (2015) och Torbjörnsson (2016). I kategorin studier med enskilda inslag av kvantitativ analys återfinns exempelvis Torbjörnsson & Jonsson (2017). Den sistnämnda studien heter "Säkerhetspolitiska trender i Afrika söder om Sahara".

⁵⁹ 41 av publikationerna klassificeras som primärt beskrivande, 14 som både beskrivande och förklarande och 8 primärt förklarande.

⁶⁰ 40 studier klassificeras som primärt skrivbordsstudier, medan 21 studier även redovisar datainsamling genom intervjuer och 9 studier är utvärderande. N=70 förklaras av att utvärderingsstudierna oftast även involverade intervjuer. Dock redovisar memon sällan sina källor och utgör ofta sammanfattningar av R-rapporter. Andelen studier som delvis baseras på intervjuer är därmed sannolikt underrepresenterad.

⁶¹ 42 av texterna (av totalt 63) klassificeras som att de i huvudsak saknar tydlig teorianknytning, 14 har det i viss mån (referenser görs till teori men strukturerar ej analysen). Endast 7 studier har stark eller mycket stark teorianknytning, dvs. där teorin strukturerar analysen.

förfogande. Av de 24 R-rapporterna inkluderar till exempel två tredjedelar datainsamling genom intervjuer.⁶² Av de sju publikationer som hade stark eller mycket stark teoranknytning var sex R-rapporter. Detta betyder att fokus på ett mindre antal, men mer omfångsrika och tidskrävande studier skulle kunna möjliggöra en djupare analytisk ansats. Sammanställningen antyder även att det finns potential för vidareutveckling vad gäller studiernas teoranknytning, vilket skulle kunna bidra till tematiskt mer fokuserade och avgränsade studier.⁶³ Med detta sagt bör högre ambitionsnivå gälla för rapporter än för memon såväl metodologiskt, teoretiskt som empiriskt givet stora skillnader vad gäller studiernas resurser, omfattning och tidsramar. Såväl memon som rapporter måste självklart anpassas efter tillgänglig finansiering, tidsramar och beställarens uttalade intressen.

En tredje slutsats är att studierna överlag består av fallstudier av ett enskilt land, med få jämförande studier och relativt få studier med en regional ansats.⁶⁴ Geografiskt domineras studierna av fyra länder (Somalia, Mali, Libyen och Nigeria) som behandlas i över hälften av projektets publikationer,⁶⁵ eller mer än två tredjedelar av de landspecifika studierna och nästan två tredjedelar av R-rapporterna.⁶⁶ Tematiskt fokuserar studierna på säkerhetspolitiska frågor i bred mening, med en särskild betoning på inbördeskrig, terrorism, fredsfrämjande insatser samt regionala och nationella säkerhetsstrukturer. Denna inriktning speglar projektets syfte, personalens kompetens och beställarens intressen. Det finns likväl en bred uppsättning säkerhetspolitiska frågor av stor relevans som inte behandlas, såsom vapenexport och tillgång till små vapen, protester och elektoralt våld, barnsoldater, legosoldater, samt privata militära företag.⁶⁷ Fram till 2019 var förekomsten av analyser av externa aktörers närvaro och intressen i

⁶² Av de 21 studier som involverade datainsamling var 18 R-rapporter.

⁶³ Över 40 % (10 av 24) R-rapporter saknade exempelvis teoranknytning.

⁶⁴ 41 studier behandlar enskilda länder, insatser eller väpnade grupper i ett enskilt land, 7 har ett jämförande perspektiv och 15 har ett regionalt fokus (inklusive studier av regionala organisationer).

⁶⁵ Somalia behandlas i 13 av de landspecifika publikationerna, Mali i 8, Libyen 8 och Nigeria 5.

⁶⁶ Dessa länder behandlas i 34 av de totalt 48 studier som har ett landspecifikt fokus, inklusive i 15 av de 24 R-rapporterna.

⁶⁷ Flera faktorer som nämnts här kan bidra till att nya konflikter blossar upp, väcks till liv, får regional spridning, blir mer komplexa, svårösta och därmed mer varaktiga över tid, något som kan få konsekvenser för säkerhetssituationen i Europas närhet eller i Europa. Flera europeiska stater är starkt engagerade i Afrika, inte minst genom EU, vilket kan aktualisera nya insatser där Sverige måste ta ställning till ett eventuellt deltagande eller bidrag.

Afrika mycket begränsade.⁶⁸ Det är även relativt få studier som har ett genderperspektiv. Gendertematik berörs även sällan i studier som har ett annat huvudfokus, exempelvis terrorism, men där genderfrågor likväl utgör en relevant aspekt.⁶⁹

6.2 Framåtblickande analys och prediktion inom Afrikaprojektets studier

Vad gäller prediktion ingår detta ofta som ett delmoment i de publikationer som granskats, och då oftast i form av en kort framåtblickande analys i studiens slutsats. Det är mer sällan som prediktion är huvudfokus i studierna, till exempel genom att det uttrycks i studiens syfte eller lyfts fram under en egen rubrik eller i ett delkapitel. Nästan en tredjedel av publikationerna innehåller dock tydliga prediktioner, medan ytterligare två femtedelar gör det i viss mån.⁷⁰ Förekomsten av prediktioner är dock aningen mer framträdande i R-rapporter än i memon.⁷¹ Inriktningen på prediktionerna styrs också av publikationernas syfte. I så kallade insatsutvärderingar inriktas analysen primärt på interna faktorer, såsom sannolikheten att bli operativ inom en viss tidsrymd eller åtgärder som bör vidtas för att åtgärda missförhållanden som annars riskerar att underminera insatsens syfte.

Ett generellt intryck är att det med några undantag saknas en utvecklad metodik för att göra prediktioner; i alla fall är den sällan explicit uttryckt i texten. Även om det förekommer, är det få studier som diskuterar prediktioner och framtiden med ett specifikt *tidsperspektiv* i åtanke, exempelvis kort, medellång eller lång sikt eller annan tidsperiod som definieras.

⁶⁸ I slutet av 2019 publicerades dock R-rapporten Gasinska, K. (red.) et al. (2019), som innebär en ny inriktning i Afrikaprojektet. Studien kommer att följas upp med fördjupad analys av ett urval externa aktörer i Afrika.

⁶⁹ Som exempel kan nämnas förekomsten av kvinnliga självmordsbombare i terrorgrupper eller systematiskt våld som riktas mot kvinnor i konflikter. Kvinnors utbildningsnivå är en ytterligare faktor som kan öka risken för radikaliserings.

⁷⁰ 19 (30 %) av de 64 publikationerna har kodats som innehållandes tydliga prediktioner, medan ytterligare 27 (43 %) gör det i viss mån.

⁷¹ Av de 24 R-rapporterna, bedöms 8 (33 %) presentera tydliga prediktioner, medan ytterligare 13 (54 %) gör det i viss mån.

Studiernas forskningsdesign bygger oftast på *enfallsstudier*, vilket är värdefullt om syftet med studien är att fördjupa vår kunskap om ett specifikt land.⁷² Däremot är det en design som lämpar sig sämre om ambitionen är att göra prediktioner. Med en enfallsdesign finns en risk att studierna blir teoretiskt och empiriskt överlastade, eftersom fokus riktas mot att förstå det enskilda fallet i all dess komplexitet, snarare än att identifiera ett urval kritiska faktorer som har relevans också i andra geografiska kontexter.

En möjlighet, som används i flera studier, är att göra en *historisk analys*, identifiera trender och med denna analys som grund extrapolera utvecklingen i framtiden.⁷³ En generell begränsning med denna typ av analys är att den är tidskrävande och kan präglas av determinism eftersom det inte öppnar upp för alternativa utvecklingsvägar under formativa politiska moment, exempelvis då ett land byter ledare eller då plötsliga revolutioner bryter ut. Det kan också vara svårt att bedöma *vilka* historiska skeenden som påverkar framtida utveckling. Hur långt tillbaka i tiden ska man gå? Den historiska analysen förefaller därmed mer lämpad för att förstå långsiktiga strukturella trender eller förståelsen av ett lands politik, snarare än att förklara dynamiska politiska förlopp. En möjlighet är att förbättra möjligheterna till prediktion genom att kombinera historisk analys med ett medvetet urval indikatorer som både fångar in långsamma skeenden och mer kortsiktig förändring. Om man t.ex. vill fånga in klimatförändringars effekt på konflikt eller migrationsströmmar, är medeltemperatur ett exempel på ett trögörsligt mått som kan kombineras med mer dynamiska mått som perioder utan regn, vattentillgång, skördar eller priser på stapelvaror som ris eller vete.

En annan metod som används är *jämförande analys*. Fall som liknar varandra men som skiljer sig i något väsentligt avseende jämförs i syfte att kunna säga något om framtiden. Som exempel kan nämnas Adriana Lins de Albuquerque (2014), där författaren riktar in analysen på en systematisk jämförelse av regeringarnas implementering av tidigare fredsavtal och kapacitet att integrera Tuaregbefolkningen i Mali och Niger, och där skillnaden i regeringens politik förklarar variationen mellan de två länderna.⁷⁴ Genom att identifiera en kritisk skillnad mellan länderna, har författaren identifierat en kausal mekanism som förklarar ett visst händelseförlopp. Även om denna metod normalt inte används för att göra prediktioner, skulle den kunna användas för att göra en bedömning om risken för rebelluppror i framtiden i länder som med beaktande av övriga

⁷² De metoder som diskuteras i detta avsnitt är välbelagda metoder inom statsvetenskaplig forskning. Se t. ex. Dickovick & Eastwood (2016); Yin (2014).

⁷³ Eriksson (2015).

⁷⁴ Lins de Albuquerque (2014).

relevanta variabler liknar Niger och Mali.⁷⁵ En begränsning är att det av tidsskäl sällan är möjligt att genomföra jämförande studier med fler än två välaggränsade fall. Stor kunskap om ämnesområdet krävs även för att kunna ”matcha” fall där förutsättningarna är relativt lika förutom på den variabel man är intresserad av att undersöka (most-similar design).

Även om prediktion inte alltid är möjlig eller önskvärd, är det tydligt att framåtblickande prognoser ofta utgör ett önskat mervärde av Afrikaprojektets skriftliga leveranser. Det kan även konstateras att möjligheten att sammanställa empiriskt baserade, detaljerade prediktioner sannolikt skulle öka om mer kvantitativa data användes i analyserna, när så är möjligt och lämpligt.

Som nämnts tidigare utgör utvärderingar av internationella insatser ett speciellt område som skiljer sig något från övriga studier inom Afrikaprojektet. Utvärderingarna har en prediktiv ansats som handlar om hur insatsernas målpuppfyllnad kan förbättras. En återkommande utmaning är att ”utvärderaren inte ges tillräckligt med tid för att bekanta sig med den tematiska och geografiska kontext som studien gäller”.⁷⁶ Även här skulle kvantitativa data kunna användas, dels som indirekt utfallsvariabel (d.v.s. som ett mått på om insatsens bredare syfte uppnåts), dels ge utvärderarna enkla metoder för att bekanta sig med konfliktens intensitet och regionala fokus.

6.2.1 Exempel på studier med prediktion som huvudsyfte

I översikten är det några studier som förtjänar att särskilt lyftas fram och där studierna har en explicit och uttalad ambition att uttala sig om framtiden i ett specifikt tidsperspektiv. Av dessa har en studie genomförts på uppdrag av regeringen.⁷⁷ Två studier har genomförts på uppdrag av Försvarsmakten, varav en har publicerats som en R-rapport.⁷⁸

⁷⁵ Andra exempel på jämförande metodik är Eriksson (2014), där fokus är inriktat på regeringarnas hantering av demonstrations- och oppositionsytringar, och Torbjörnsson (2017). Torbjörnsson & Jonsson (2017) är en inomfallsstudie, som dock rymmer jämförande utblickar för att resonera kring Boko Harams utveckling i Nigeria. Ett ytterligare – icke-Afrikarelaterat – exempel är Johan Engvalls studie om två insatser i Georgien och Ukraina, med ett fokus på konventionell rustningskontroll och militära förtroendeskapande åtgärder. Se Engvall (2018).

⁷⁶ Engvall & Berglund (2017).

⁷⁷ Bergenwall & Gunnarson (red.) (2018).

⁷⁸ Bergenwall & Eriksson (2014).

6.2.1.1 R-rapport (2014) "Mellanöstern och Nordafrika i ett 5-10-årsperspektiv"

I denna studie görs en analys av den säkerhetspolitiska utvecklingen i Mellanöstern och Nordafrika med ett 5-10-årsperspektiv. Studiens grepp är brett och omfattar en analys av förändringar i maktbalans mellan staterna i regionen, militära strukturer, statens institutioner, demokrati, multilaterala organisationer, det koloniala arvets betydelse, ekonomiskt beroende samt inrikespolitiska faktorer. Skribenternas ambition är att täcka in växelverkan mellan nationella, regionala och globala trender i syfte att identifiera möjliga utvecklingsvägar. Analysen bygger huvudsakligen på kvalitativ metod och dokumentstudier. Referenser finns till Uppsala universitets konfliktdataprojekt (UCDP). Metoden är sparsamt beskriven med tanke på uppgiftens omfattning och komplexitet. Poängteras bör att ambitionsnivå och svårighetsgrad i denna uppgift är mycket hög och svår att genomföra på ett metodologiskt rigoröst sätt.

6.2.1.2 AG Säkpöl: Underlag (2018) "FOI:s uppdaterade underlag inför nästa försvarspolitiska inriktningsbeslut"

Underlaget har skrivits på uppdrag av regeringen (Försvarsdepartementet). Uppdraget var att inför arbetet med nästa försvarspolitiska inriktningsbeslut ta fram ett relevant underlag om internationella och försvars- och säkerhetspolitiska förhållanden. Underlaget baseras på FOI:s expertis och forskning inom myndighetens samtliga områden med särskild uppmärksamhet åt utvecklingen i världens regioner, däribland Afrika. Underlaget om Afrika omfattar cirka 5 sidor (av totalt 127 sidor text). Tidsperspektivet är att bedöma utvecklingen under de närmaste 5-10 åren. Metoden är kvalitativ och omfattar en analys av "övergripande trender" på kontinenten som helhet, för att därefter identifiera trender i tre subregioner: Nordafrika, Sahel och Afrikas Horn, samt utvecklingen av den afrikanska freds- och säkerhetsarkitekturen. De temata som berörs är demokrati, protester och upplöpp, militärkupper, terrorism, rebeller, radikalisering, väpnade konflikter, ekonomisk utveckling, inkomstklyftor, organiserad brottslighet, rebeller, befolkningstillväxt, klimatförändringar. De länder och regioner som särskilt lyfts fram är utvecklingen i Libyen, Egypten, Somalia, Sahelländerna, Kenya, Sydsudan och Etiopien. Sammanställningen är kvalitativt baserad och bygger på andrahandskällor. Såväl strukturella trender som mer aktörsorienterade faktorer berörs. Referenser till kvantitativa data saknas och det framgår inte av den färdiga produkten vilket källmaterial som har använts.

6.2.1.3 SLO (Strategisk långsiktig omvärldsanalys): Underlag (2018) "Säkerhetspolitiska trender i Afrika söder om Sahara"

Studien har ett långt tidsperspektiv där ambitionen är att identifiera trender under perioden 1990-2045. Arbetet omfattar världens samtliga regioner, utom Latinamerika. Till skillnad från tidigare analyser av denna karaktär är metoden kvantitativt baserad, där ett urval indikatorer valts ut utifrån teoribildning om internationell politik. Arbetet syftar till att identifiera långsiktiga trender i ett 2045-perspektiv, men också att fånga upp nya tendenser i de olika regionerna. Den tematik som analyseras rör väpnade konflikter, terrorism, migration, ekonomisk och militär maktbalans, demografi, ekonomisk utveckling, demokrati, ekonomisk utveckling, handel, jämställdhet, *governance*, arbetslöshet, ekonomisk ojämlikhet samt klimatförändringar. Ett gemensamt ramverk har utvecklats och som därefter tillämpas på världens regioner. Delstudien om Afrika består av 54 sidor text, ett stort antal grafer, samt fördjupade analyser av Etiopien, Nigeria och Sydafrika. Fallstudierna är mer kvalitativa till sin karaktär, men utgår från en gemensam struktur. Angreppssättet är mycket ambitiöst i sitt kvantitativa anslag. Samtidigt finns potential att ytterligare utveckla och precisera de teoretiska perspektiven. Tyngdpunkten på kontinental nivå ligger på strukturella långsiktiga förändringsmönster, medan fallstudierna i högre grad beaktar länders interna politiska dynamik. Men även här finns möjlighet att utveckla faktorer i ländernas interna politiska dynamik, som maktövertagande, relation mellan regering och opposition, politiska partier, eller förekomst av elektoralt våld.

6.2.2 Analys

Efter att ha gått igenom Afrikaprojektets publikationer under perioden 2014-2019 kan vi dra vissa slutsatser.

Många studier rymmer prediktioner, även om det mer sällan uttrycks explicit som syfte. Få studier definierar vilken tidshorisont som gäller för studiernas utsagor om framtiden (t.ex. ett, tre eller fem års sikt).

Metoden för prediktioner är sällan explicit uttalad, utan får utläsas implicit genom studiernas design. Metoden är huvudsakligen kvalitativt baserad och består av studier av ett eller två fall. I den mån prediktion är ett mål för studier, finns här potential att tydliggöra den metodik som används.

Potential finns att kombinera kvalitativa studier med kvantitativa data för att sätta in studier i ett större perspektiv. På vad sätt utmärker sig det studerade landet? Är det extremt eller är det ett så kallat normalfall? Här finns möjligheter att i högre grad utnyttja tillgängliga databaser.

Målsättningen bör vara att utnyttja kvantitativa data som rymmer mer trögrörliga indikatorer och kombinera dessa med mer snabbrorliga dynamiska mått som indikerar förändring.

Den arbetsmetodik som använts inom SLO-projektet (Strategisk långsiktig omvärldsanalys) är ett lovande exempel på studie som kombinerar kvantitativ och kvalitativ undersökningsmetod för att identifiera långsiktiga trender och nya tendenser i regioners och länders utveckling.

6.3 Möjliga områden för framtagande av modeller för prediktion

Prediktion utgör typiskt sett inte huvudfokus för de studier vi har granskat, men icke desto mindre tillför denna form av analys ett betydande mervärde, som i dagsläget adderas med tydlig metod.

Vi kommer i detta avsnitt att argumentera för två huvudspår när det kommer till sakområden för prediktion inom FOI:s Afrikaanalys:

1. Konflikter, samt dessas intensitet och spridning
2. Konsekvenser av konflikter med fokus på migration

Argumentet för det första spåret är att det är tydligt att intresset hos beställande kund för konflikter är relativt stabilt över tid. Dels rör det sig om ett fokus på olika typer av inomstatliga konflikter, respektive olika sorters insatser för att komma till bukt med dessa. Dels handlar det om ett fokus på länder där Sverige har, tidigare har haft eller i framtiden kan komma att ha fredsfrämjande trupp, i den prioriteringsordningen. Antalet faktorer som inkluderas i analysen av dessa konflikter kan med fördel utvidgas med kvantitativa data. Däremot ter det sig osannolikt att fokus på väpnade inomstatliga konflikter, där svensk trupp kan komma att agera fredsfrämjare, skulle förändras markant under överskådlig tid. Framtagande av en modell för prediktion kan därför med fördel fokusera på dessa konflikter.

Studiet av inomstatliga konflikter har växt exponentiellt under det gångna decenniet, vilket reflekteras i utbudet av databaser som fokuserar på olika dimensioner av väpnad konflikt. Under tidigt 2000-tal dominerades fältet av studier som fokuserade på *orsaker* på inomstatliga konflikter, där analysnivån utgjordes av jämförelser av stater på makronivå över tid. En mängd studier försökte på detta sätt klarlägga om det var möjligheten att finansiera väpnade konflikter eller politisk repression (*greed or grievance*) som bäst förklarade uppkomsten till inomstatliga konflikter. Sedan dess har fältet vidareutvecklats och fokus har skiftat, dels till variationer på mikronivå inom enskilda konflikter respektive över tid, dels till andra faser av konfliktcykeln, med fokus på exempelvis konflikters varaktighet över tid, hur konflikter upphör, hur säkerhetsstyrkor bör reformeras, hur icke-statliga ex-kombattanter bör bemötas,

hur fredsfrämjande genomförs och fredsavtal utformas för att maximera möjligheten att upprätthålla fred.^{79, 80, 81}

Som ett komplement till detta föreslås att en modell för prediktion tas fram av FOI med fokus på händelser *under* konflikten, respektive *konsekvenser* av konflikten. Något som stödjer detta är att intresset hos FOI:s externa uppdragsgivare (Försvarsdepartementet) främst är inriktat på ett skede när en konflikt uppstått och en fredsfrämjande insats aktualiseras.⁸² Vidare finns det redan ett flertal prediktiva system (se kapitel 3, 4 och 5) som fokuserar på just sannolikheten för inbördeskrig under en given tidsperiod. Möjligheten för FOI att på ett avgörande sätt bidra till detta fält torde således vara begränsad.

Däremot är utbudet av datadrivna prediktioner avseende konsekvenser av en redan pågående konflikt betydligt mer begränsat. För den som söker svar på huruvida en konflikt håller på att spridas regionalt inom ett land, om konfliktintensiteten kommer att öka, risken för spridning över landsgränser, eller huruvida betydande migrationsflöden kommer uppstå, är det mycket svårt att hitta kvalificerad hjälp. Modeller som angriper denna typ av frågor torde samtidigt vara av betydligt större intresse för FOI:s uppdragsgivare än regionala modeller, eftersom denna typ av händelser kan ha betydande påverkan på exempelvis fredsfrämjande trupp i ett land.

Vad gäller val av land (eller länder) för test av ett prototypsystem föreslås att det görs mellan de länder som varit mest i fokus för tidigare studier, d.v.s. Somalia, Mali, Libyen och Nigeria. Starka skäl för och emot kan anföras för vart och ett av dessa länder. Den islamistiska terrorgruppen al-Shabaab i **Somalia** har vuxit sig starkare under de senaste åren och internationell närvaro finns genom Afrikanska Unionens FN-stödda och till stor del EU-finansierade insats AMISOM. Situationen i Somalia kommer troligen inte att förbättras på fem till tio års sikt. I **Mali** syns i dagsläget inte någon lösning på den instabila säkerhetspolitiska situationen, utan konfliktbilden har snarare blivit alltmer komplex med en spridning av konflikten i Västafrika, framför allt till grannländerna Burkina Faso och Niger. Flera insatser finns i regionen i form FN-styrkan UNAMSIL, den franskledda Operation Barkhane samt EU CAP Sahel

⁷⁹ Kalyvas (2006).

⁸⁰ Jonsson (2014).

⁸¹ Vervimp, Justino & Brück (2009).

⁸² Baserat på de studier som beställts är dessa primärt Somalia, Mali, Libyen och Nigeria, i den ordningen. Ytterligare länder kan givetvis vara av intresse, inklusive exempelvis Centralafrikanska Republiken, Sudan, Sydsudan eller Demokratiska Republiken Kongo.

Mali, EUTM Mali samt EUCAP Sahel Niger, vilket uttrycker den prioritet EU tillmäter regionen. Situationen i **Libyen** påverkar stabiliteten i Nordafrika och Europa, med stor påverkan på migrationsflöden, destabiliserande inverkan på grannländer genom spridning av t.ex. små och lätta vapen, samt en betydande risk för en högintensiv inomstatlig konflikt. **Nigeria** är med 190 miljoner invånare central för utvecklingen i hela Västafrika. I landets norra delar har rebellgruppen Boko Haram en destabiliserande inverkan på området kring Tchadsjön och påtaglig risk för att även andra konflikter inom landet eskalerar. Med en stor export av olja är landet även strategiskt viktigt för Europa.

Det andra förslaget är att öppna för möjligheten att inkludera konsekvenser av konflikter i form av migration, ett tema som i stor utsträckning har lyst med sin frånvaro i den tidigare analysen. I dag bedriver FOI tillsammans med Polismyndigheten visst arbete genom EU-projektet MIRROR⁸³, men genom denna inriktning skulle FOI potentiellt även kunna attrahera uppdragsgivare såsom Utrikesdepartementet, Migrationsverket eller organisationer inom biståndsvärlden. Migrationen inom och från Afrika kommer sannolikt att öka i framtiden mot bakgrund av befolkningstillväxt och klimatförändringar, som leder till att vissa regioner i Afrika blir omöjliga att bo i. Migrationen till Europa kommer mot bakgrund av detta inte att minska, utan snarare öka, i framtiden. Ökade krav kommer att ställas på intern solidaritet inom EU, där det inte är osannolikt att Sverige ombeds att delta i olika insatser som gäller såväl gränsövervakning som bistånd till migranter. Det är inte omöjligt att dessa frågor även kommer att behöva hanteras av Försvarsdepartementet, i samarbete med andra myndigheter.

Migration drivs av en mängd olika faktorer som påverkar både viljan att migrera och målet för migrationen. Vad gäller migration från Afrika är en stor del av migrationen inomafrikansk, med rörelser mot kustregioner och större städer. Andra rutter går mot Mellanöstern eller mot Medelhavet och vidare till Europa. En stor del av migrationen i Västafrika är även arbetsrelaterad och följer skördesäsongen, där avsikten inte primärt är att ta sig vidare till Europa. Den migration som rör sig mot Europa påverkas av en mängd faktorer på land-, region- samt individnivå. Några av de faktorer som påverkar personers individuella förutsättningar att migrera är t.ex. inkomst, kön, ålder eller kontakter med personer som tidigare har migrerat. En stor diaspora i ”mållandet” kan med andra ord generera nya migranter i form av anhöriginvandring. Även säkerhet och infrastruktur i form av hotell, vatten och transport längs med olika

⁸³ ”MIRROR Project”, <https://h2020mirror.eu/>.

migrantrutter bör vägas in i analysen, eftersom det påverkar de vägval migranter gör.

Det kritiska momentet för denna typ av inriktning är tillgång till datainformation om migranter. Migrationsdata bör kompletteras med information som både tar hänsyn till socio-ekonomisk utvecklingsnivå i migranternas ursprungsländer, information om migrantrutter från ursprungsort till måldestination, information om eventuella hinder på vägen i form av stängda gränser, ökade kontroller etc.⁸⁴ Även här kan en lämplig väg att gå vara att fokusera på existerande konflikter i Mali, Somalia och Libyen. Migrationen från området kring Tchadsjön kan även öka framgent som en följd av klimatförändringar och pågående konflikt i området.⁸⁵ Även klimatrelaterad migration från Mellanöstern och Nordafrika tros öka under detta sekel.⁸⁶

För att kunna förutse migrationsströmmars omfattning och riktning behövs data eftersom det kommer att påverka de insatser som krävs från EU och våra europeiska partners. Det finns politiska argument för att inkludera migration i modellen, som är ett nytt område för FOI. Sveriges blivande EU-kommissionär kommer att få ansvar för migrationsfrågor. Om FOI kan leverera relevant och vederhäftig information inom detta område, kommer detta troligen komma till nytta för berörda myndigheter.

⁸⁴ I en studie av Mexiko studeras hur narkotikaligors vägval påverkas av faktorer på mikronivå, exempelvis politisk tillhörighet på borgmästare. I en annan studie av Indien visas hur kvinnors väg till universitet påverkas av inrapporterade sexuella trakasserier längs med olika färdtrutter, vilket får kvinnor att välja en längre men säkrare resväg. Se Dell, M. (2015); Borker, G. (2017).

⁸⁵ Vivekananda et al. (2019).

⁸⁶ Lelieveld et al (2015).

7 Förutsättningar att skapa prediktiva system för FOI:s behov

I detta kapitel diskuterar vi möjligheterna att konstruera ett prediktivt system på prototypnivå på FOI under åren 2020-2021. Det finns naturligtvis många aspekter som är relevanta för en sådan bedömning, men nedan har vi valt ut fem områden som vi menar har stor påverkan på vad det fortsatta projektet kan åstadkomma.

7.1 Utmaningar inom modellering

De modelleringstekniker som används i de studerade systemen är alla bekanta för flera forskare vid FOI, så det ligger väl inom myndighetens kapacitet att utforska angreppssättet med flera olika modelltyper, kombinerade i en ensemble. Med detta sagt så ligger kanske den största utmaningen (inom modelleringsområdet) i att skapa någon eller några mer teoridrivna modeller, och då handlar det typiskt om olika former av regressionsmodeller.

Det som kan göra detta extra svårt är om vi fokuserar på ett område där ingen av våra forskare ännu har fördjupat sig. För de mindre teoriberoende, men mycket datakrävande, maskininlärningsteknikerna är det kanske snarast just tillgången på data som är mest gränssättande (se nedan). Dock finns det en inlärningsperiod även med många maskininlärningstekniker för att hitta effektiva sätt att tillämpa dem.

7.2 Utmaningar avseende mjukvara

Utveckling av ny mjukvara kräver ofta stora arbetsinsatser. En strategi där FOI skrev all mjukvara ”från scratch” bedömer vi inte skulle vara realistisk om målet är att ha en prototyp under 2020. Det som gör läget mer gynnsamt är att ViEWS-systemets kod finns öppet tillgänglig, och även kan modifieras för nya syften⁸⁷. Ett första steg under 2020 skulle därför kunna vara att identifiera vilka delar av koden som har skrivits för ViEWS som skulle kunna användas i ett FOI-system.

⁸⁷ Hegre et al. (2019).

Därefter skulle återstående nödvändiga komponenter behöva utvecklas och integreras med ViEWS-komponenterna.

En särskild svårighet står att finna i frågan om händelsedata. I det fall prototypen bedöms behöva automatkodade händelsedata så står valet mellan att använda de öppna databaser som redan finns (t.ex. GDELT, ICEWS, Phoenix eller TERRIER) eller att själv utveckla ett system för automatisk inhämtning och kodning. Vår bedömning är att det senare skulle vara en förhållandevis svår uppgift, även om det finns många resultat och mjukvarukomponenter att stödja sig på, och något som svårligen skulle kunna åstadkommas i ett mindre pilotprojekt.

7.3 Utmaningar avseende data

Här ligger troligen projektets största utmaningar, beroende på vilket fokus prototypsystemet ska ha. Om systemet ska inriktas mot konfliktrisk så finns det goda exempel att följa i de system som har diskuterats i denna rapport. I det fall att systemet istället ska riktas mot andra fenomen, t.ex. migration som nämndes i kapitel 6, så är datafrågan mer utmanande. En avvägning kommer att behöva göras mellan prototypens möjligheter att tillföra något nytt och svårigheten att sätta samman en tillräckligt omfattande och kvalitativ datamängd.

Det är också tydligt att ju mer snabbreagerande system vi avser att bygga, desto mer arbete kommer insamlingen och beredningen av data att kräva. Något som skulle skapa stor flexibilitet framöver vore en egen förmåga till automatkodning av olika typer av högfrekventa data, men som indikerades i förra avsnittet så är detta en kapacitet som kräver betydande forsknings- och utvecklingsresurser.

7.4 Utmaningar avseende datakraft

Konstruktion av maskininlärningsmodeller och användning av simuleringsmodeller är potentiellt mycket beräkningsintensiva. Även om det är svårt att säga något på detta stadium om hur mycket beräkningskapacitet vi kommer att behöva så är detta något som kan bli gränssättande. ViEWS-projektet har i sitt arbete använt beräkningsresurser från Swedish National Infrastructure for Computing (SNIC) vid Uppsala Multidisciplinary Center for Advanced Computational Science (Uppmax), vilka är att betrakta som ”superdatorresurser”.

FOIs datorkluster i Kista och Grindsjön, och i synnerhet det nyanskaffade DGX-2 systemet (hårdvarutillverkaren NVIDIA:s specialdesignade dator för träning av djupinlärningsmodeller), är i princip tillräckligt kraftfulla för träning och körning av de ovannämnda modellerna. Dessa resurser tas naturligtvis i första hand i anspråk av de projekt som har finansierat och anskaffat dem, men här föreligger ändå möjligheter att lösa uppkomna beräkningsbehov inom myndigheten.

7.5 Utmaningar avseende utvecklingsresurser

Projektet kommer under 2020 att omfatta ca 6 personmånader, vilket är en betydligt mindre satsning än några av de projekt som studerades under 2019:

- ViEWS har en budget på ca 25 MSEK över 5 år, så ca 5 MSEK per år i snitt
- Det tyska projektet ”IT-stöd Krisförvarning”⁸⁸, där initiativet PREVIEW ingår, var prognostiserat att kosta ca 150 MSEK från 2017 fram till 30 juni 2019, så ca 50-75 MSEK per år⁸⁹
- Om uppgifterna i Wired är korrekta så kostade ICEWS 35 MUSD, eller ca 350 MSEK i dagens penningvärde över 4 år, i snitt knappt 90 MSEK per år (se kapitel 3)

Nu är naturligtvis ambitionen med ett pilotprojekt vid FOI betydligt mer begränsad än med dessa tre, och det finns också såväl mjukvarukomponenter som erfarenheter och forskningsresultat att dra nytta av. Det blir dock viktigt att tidigt prioritera vilka aktiviteter som skulle göra störst nytta för FOI i det fortsatta arbetet. Ska man tro den amerikanske händelsedataforskaren Philip Schrodtt så finns det gott om kvalificerade forskningsuppgifter inom området, uppgifter som skulle kunna vara fokus för större projekt eller program.⁹⁰

⁸⁸ På tyska heter projektet ”IT-Unterstützung Krisenfrüherkennung”.

⁸⁹ Uppgifterna kommer från en officiell tysk skriftväxling: ”Drucksache 19/7604”, se <http://dip21.bundestag.de/dip21/btd/19/076/1907604.pdf>.

⁹⁰ ”Seven current challenges in event data”, <https://asecondmouse.wordpress.com/2019/03/13/seven-current-challenges-in-event-data/>.

8 Slutsatser

Vi har i denna rapport gett en introduktion till modellbaserad prediktion av säkerhetspolitiska händelser genom att beskriva fyra existerande system: det amerikanska ICEWS, de nederländska systemen CRAM och PVRAM och det svenska ViEWS.

Utifrån resonemang kring denna typ av teknologi i relation till Afrikaprojektets arbete har vi kunnat konstatera att det finns goda förutsättningar för FOI att nyttiggöra modellbaserade prediktioner inom det säkerhetspolitiska området. Detta gäller såväl prediktioner från existerande system som från framtida system inriktade mot andra frågor.

Framtagning av ett prediktivt system på prototypnivå under 2020 är möjlig, även om det innebär en betydande utmaning. Ett av projektets mål för 2020 är därför också att söka samarbeten och uppdrag som skulle möjliggöra för FOI att fördjupa och vidmakthålla färdigheter och kunskaper inom detta område. En tidig uppskattning är att en arbetsinsats på 2-3 personår vore tillräckligt för att utveckla ett system som skulle vara användbart i skarp verksamhet.

En bieffekt av detta projekt är också att fokus har riktats på användningen av stora datamängder inom det säkerhetspolitiska området, också i fall där ambitionen inte är prediktiv. Även nuläge eller historik som byggs upp av data förefaller initialt kunna vara till stor nytta i FOI:s arbete.

Inom både HCSS och PREVIEW har man lagt betydande arbete på att bygga upp sin förmåga till datainhämtning och att etablera egna datalager.⁹¹ Dessa lager kan sedan ligga till grund för mängder av olika tillämpningar, såväl prediktiva som andra. Här finns alltså ytterligare ett uppslag till en strategisk satsning för FOI att överväga: en långsiktig uppbyggnad av kunskap och infrastruktur som kan utgöra grunden för framtida datadrivna stöd till analysverksamheten.

⁹¹ Korrespondens med ViEWS och HCSS.

Referenser

Allison, P.D. (2010) "Missing Data" ur Millsap, R.E. & Maydeu-Olivares, A. (red.) *The SAGE Handbook of Quantitative Methods in Psychology* (SAGE Publications).

Beger, A., Dorff, C.L. & Ward, M.D. (2014) "Ensemble forecasting of irregular leadership change", *Research and Politics* Oct-Dec:1-7.

Benedetti, R. (2010) "Scoring Rules for Forecast Verification", *Monthly Weather Review* 138(1):203-211.

Bergenwall, S. & Gunnarson, C. (red.) (2018) "FOI:s uppdaterade underlag inför nästa försvarspolitiska inriktningsbeslut", FOI-2016-1544.

Bergenwall, S. & Eriksson, M. (2014) "Mellanöstern och Nordafrika i ett 5-10-årsperspektiv: stöd till LEDS INRIS:s omvärldsanalys", FOI-R--4008--SE.

Borker, G. (2017) "Safety First: Perceived Risk of Street Harassment and Educational Choices of Women", Job Market Paper, Department of Economics: Brown University, USA.

Brier, G.W. (1950) "Verification of Forecasts Expressed in Terms of Probability", *Monthly Weather Review* 78(1):1-3.

Bussman, M., Dorussen, H. & Gleditsch, N.P. (2014) "Against All Odds: 2013 Richardson Award to Mats Hammarström and Peter Wallensteen", *Peace Economics, Peace Science and Public Policy* 20(2):235-243.

Dell, M. (2015) "Trafficking Networks and the Mexican Drug War", *The American Economic Review*, 105 (6): 1738-1779.

Dickovick & Eastwood (2016) *Comparative Politics: Integrating Theories, Methods, and Cases* (New York, Oxford: Oxford University Press).

Diehl, P.F. (1992) "The Correlates of War Project: A Bibliographic Essay, 1961-1991", *International Studies Notes* 17(2):21-33.

Engvall, J. (2018) "OSCE and Military Confidence-Building in Conflicts: Lessons from Georgia and Ukraine", FOI-R--4750--SE.

Engvall, J. & Berglund, C. (2017) "Utvärderingsmetod: Rapport från workshop 20 november 2017", FOI Memo 6342.

Eriksson, M. (2014) "Udda fåglar? Algeriet och Marocko i ljuset av de arabiska revolterna", FOI Memo 4868.

Eriksson, M. (2015) "A fratricidal Libya and its second civil war", FOI-R--4177--SE.

Gasinska, K. (red.), Hagström Frisell, E., Hedenskog, J., Hellström, J., Sjökvist, E., Lindström, M., Sundberg, A., Bergenwall, S., Gunnarson, C. (2019) "Foreign military bases and installations in Africa", FOI-R--4658--SE.

Gerner, D.J., Schrod, P.A., Yilmaz, O. & Abu-Jabr, R. (2002) "Conflict and mediation event observations (CAMEO): A new event data framework for the analysis of foreign policy interactions", *International Studies Association, New Orleans*.

Good, I.J. (1952) "Rational Decisions", *Journal of the Royal Statistical Society Series B* 14(1):107-114.

Hegre, H., Allansson, M., Basedau, M., Colaresi, M., Croicu, M., Fjelde, H., Hoyle, F., Hultman, L., Höglad, S., Jansen, R., Mouleb, N., Muhammad, S.A., Nilsson, D., Nygård, H.M., Olafsdottir, G., Petrova, K., Randahl, D., Rød, E.G., Schneider, G., von Uexküll, N. & Vestby, J. (2019) "ViEWS: A political early-warning system", *Journal of Peace Research* 56(2):155-174.

Hoff, P.D. & Ward, M.D. (2004) "Modeling Dependencies in International Relations Networks", *Political Analysis* 12(2):160-175.

Honaker, J., King, G. & Blackwell, M. (2011) "Amelia II: A Program for Missing Data", *Journal of Statistical Software* 45(7):1-47.

Hull Wiklund, C. & Ingerstad, G. (2015) "The Regionalisation of Peace Operations in Africa. Advantages, challenges and the way ahead", FOI-R--4031-SE.

Jonsson, M. (2014) *A Farewell to Arms. Motivational Change and Divergence Inside FARC-EP 2002-2010* (Uppsala: Acta Universitatis Upsaliensis).

Kalyvas, S.N. (2006) *The Logic of Violence in Civil War* (New York: Cambridge University Press).

Leetaru, K. & Schrodt, P.A. (2013) "GDELT: Global data on events, location, and tone, 1979–2012", *ISA annual convention* 2(4):1-49.

Lelieveld, J., Proestos, Y., Hadjinicolau, P., Tanarthe, M., Tyrllis, E. & Zittis, G. (2016) "Strongly increasing heat extremes in the Middle East and North Africa (MENA) in the 21st Century", *Climate Change* 137(1-2):245-260.

Lins de Albuquerque, A. (2014) "Explaining the 2012 Tuareg rebellion in Mali and the lack thereof in Niger", FOI Memo 5099.

Lins de Albuquerque, A. (2017) "Terrorism in Africa: a quantitative analysis", FOI-R--4398--SE.

Lustick, I. (2002) "PS-I: A User-Friendly Agent-Based Modeling Platform for Testing Theories of Political Identity and Political Stability", *Journal of Artificial Societies and Social Simulation* 5(3).

Mahoney, S., Comstock, E., deBlois, B. & Darcy S. (2011) "Aggregating Forecasts Using a Learned Bayesian Network", *Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference*.

Montgomery, J.M., Hollenbach, F.M. & Ward, M.D. (2012) "Improving Predictions Using Ensemble Bayesian Model Averaging", *Political Analysis* 20(3):271-291.

O'Brien, S.P. (2010) "Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research", *International Studies Review* 12:87-104.

O'Brien, S. P. (2013) "A multi-method approach for near real time conflict and crisis early warning" ur Subrahmanian, V.S. (red.) *Handbook of Computational Approaches to Counterterrorism* (New York, NY: Springer).

Pettersson, T. & Eck, K. (2018) "Organized violence, 1989-2017", *Journal of Peace Research* 55(4):535-547.

Rõõs, H., Usanov, A., Farnham, N. & Sweijis, T. (2018) "Improving the Early Warning Function of Civil War Onset Models Using Automated Event Data", Working Paper, <https://dx.doi.org/10.2139/ssrn.2948293>.

Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys* (New York: Wiley).

Schrodt, P.A., Davis, S.G. & Weddle, J.L. (1994) "Political Science: KEDS-A Program for the Machine Coding of Event Data", *Social Science Computer Review* 12(4):561-587.

Schrodt, P.A. (2010) "Automated production of high-volume, real-time political event data", *APSA 2010 Annual Meeting Paper*.

Tetlock, P. (2005) *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, N.J.: Princeton University Press).

Tollefsen, A.F., Strand, H. & Buhaug, H. (2012) "PRIO-GRID: A unified spatial data structure", *Journal of Peace Research* 49(2):363-374.

Torbjörnsson, D. (2016) "Managing Communal Conflict in Africa", FOI-R--4226--SE.

Torbjörnsson, D. (2017) "Explaining the differences in al-Shabaab expansion into Ethiopia and Kenya", FOI Memo 6061.

Torbjörnsson, D. & Jonsson, M. (2017) "Boko Haram. On the verge of defeat or a long-term threat?", FOI-R--4488--SE.

Toth, Z. & Kalnay, E. (1997) "Ensemble forecasting at NCEP and the breeding method", *Monthly Weather Review* 125(12):3297-3319.

Vervimp, P., Justino, P. & Brück, T. (2009) "The Analysis of Conflict: A Micro-Level Perspective", *Journal of Peace Research* 46(3):307-314.

Vivekananda, J., Wall, M., Sylvestre, F. & Nagarajan, C. (2019) *Shoring up Stability. Addressing Climate and Fragility Risks in the Lake Chad Region*. (Berlin: adelphi).

Ward, M.D. & Gleditsch, K.S. (2002) "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace", *Political Analysis* 10(3):244-260.

Ward, M.D., Metternich, N.W., Dorff, C.L., Gallop, M., Hollenbach, F.M., Schultz, A. & Weschle, S. (2013a) "Learning from the Past and Stepping into the Future: Toward a New Generation of Conflict Prediction", *International Studies Review* 15:473-490.

Ward, M.D., Beger, A., Cutler, J., Dickenson, M., Dorff, C. & Radford, B. (2013b) "Comparing GDELT and ICEWS Event Data", *Analysis* 21(1):267-297.

Ward, M.D. (2016) “Can We Predict Politics? Toward What End?”, *Journal of Global Security Studies* 1(1):80-91.

Wedgwood, J., Ruvinsky, A. & Siedlecki, T. (2012) “What lies beneath: Forecast transparency to foster understanding and trust in forecast models” ur Nicholson, D.M. & Schmorow, D.D. (red.) *Advances in Design for Cross-Cultural Activities – Part II* (Boca Raton, FL: CRC Press).

Yin (2014) *Case Study Research: Design and Methods* (London: Sage).

Appendix A

Strictly proper scoring rules

Proper scoring rules^{1, 2} har, likt många andra kraftfulla verktyg inom prediktion, sitt ursprung inom meteorologin. En proper scoring rule har egenskapen att ge ett incitament till att vara sanningsenlig. Den åstadkommer detta genom egenskapen, s.k. *propriety*, att den bästa *förväntade* poängen uppnås genom att ange (och få poängsatt) den sannolikhetsfördelning som svarar mot vad man faktiskt tror. Om en poängsättningsregel har egenskapen att det *enda* sättet att maximera sin förväntade poäng är att ange en sanningsenlig fördelning, då kallas den *strictly proper*. Notationen i det som följer är tagen ifrån Tödter (2011).

Brierpoäng

f^t = den bedömda sannolikheten i fall t , $f^t \in [0,1]$

$\mathbf{o}^t$ = utfallet i fall t ; 1 om händelsen inträffade, annars 0

$\mathbf{f}^t$ = vektorn av bedömda sannolikheter i fall t (av dimension n för n olika hypoteser)

$\mathbf{o}^t$ = vektorn av utfall i fall t ; 1 för den hypotes som realiserades, 0 för övriga (även denna av dimension n)

Poäng vid en enskild bedömning:

$$PS^t = (\mathbf{f}^t - \mathbf{o}^t)^2$$

Vid flera hypoteser (vektorform), snitt-PS (probability score) över T tillfällen:

$$\overline{PS} = \frac{1}{T} \sum_t (\mathbf{f}^t - \mathbf{o}^t)^2$$

¹ Ett försök till svensk översättning skulle vara "sunda poängsättningsregler" eller "korrekta poängsättningsregler".

² Se t.ex. Winkler & Murphy (1968).

Detta ger ett tal mellan 0 och 2, där 0 är bäst och 2 sämst.

Binära fallet, konventionell snitt-Brierpoäng:

$$\overline{BS} = \frac{\overline{PS}}{2} = \frac{1}{2T} \sum_t 2(f_1^t - o_1^t)^2$$

Här har man traditionellt delat med 2, och får då ett tal mellan 0 och 1, där 0 är bäst och 1 sämst. Det nedre indexet (ettan) är kategorin *inträffat*, och den bedömda sannolikheten f_2^t för *icke-inträffat* i fall t kan skrivas $1 - f_1^t$, vilket leder till att summan ser ut som den gör.

Logaritmiska poäng

Ett alternativ till Brierpoäng är logaritmiska poäng, eller *negativa logaritmen av den bedömda sannolikheten för det inträffade* (se t.ex. Good, 1952).

Logaritmiska poäng, ibland kallade *ignorance score*, är besläktade med Brierpoäng på så sätt de senare är en approximation av de tidigare (Benedetti, 2010). De skiljer sig dock åt på två viktiga punkter:

1. Den bästa logaritmiska poängen är 0, samma som för Brierpoäng, men den sämsta är ∞ (oändlighet), till skillnad mot Brierpoängens 2 (eller 1, se ovan).
2. Logaritmiska poäng är *lokala* vilket innebär att poängen enbart beror på sannolikheten för inträffade händelser, något som inte gäller för Brierpoäng.

Liksom Brierpoäng är logaritmiska poäng *strictly proper*.

$$IGN^t = -\log_2 f_j^t = \sum_k^J o_k^t \log_2 \frac{o_k^t}{f_k^t}$$

J är antalet hypoteser. Trots den till synes krångliga summan till höger så kan poängen beskrivas (förhållandevis) enkelt som *negativa logaritmen av den bedömda sannolikheten för det inträffade*.

Uppdelning i komponenter

Låt oss kalla Brierpoäng för PS (probability score) och logaritmiska poäng för IGN (ignorance score). Om vi tar en uppsättning prediktioner och beräknar snittpoängen $\overline{PS}$ och $\overline{IGN}$ så kan dessa delas upp i komponenter som berättar om

olika aspekter av prediktionen: *reliabilitet*, *resolution* och *osäkerhet*.³ Dessa ges ofta beteckningarna REL (*reliability*), RES (*resolution*) och UNC (*uncertainty*). I båda fallen ges snittpoängen av följande uttryck, och de kan alltså delas upp på samma sätt:

$$\overline{PS} = REL - RES + UNC$$

$$\overline{IGN} = REL - RES + UNC$$

Hög reliabilitet innebär att snittfrekvensen hos predicerade händelser ligger nära deras bedömda sannolikhet. Ett annat sätt att uttrycka det är att händelser som bedöms ha en sannolikhet på t.ex. 0.2 också inträffar 2 gånger av 10 i snitt. Ett sätt att uppnå hög reliabilitet för händelsetyper där det finns bra historiska data är att ständigt ange den s.k. *basnivån* eller historiska relativa snittfrekvensen. Låt oss anta att det på en viss plats har regnat 36% av dagarna de senaste 100 åren. En prediktionsstrategi kan då vara att ange sannolikhet 0.36 för regn *varje dag*, något som troligen skulle ge god reliabilitet. Det skulle dock vara högst otillfredsställande, vilket leder oss till resolution.

Hög resolution innebär att prediktionerna varierar utifrån vad det faktiska utfallet blev. Bäst är det förstås om man anger högre sannolikheter för händelser som till slut inträffar och lägre för sådant som inte inträffar, men även det omvända kan ge hög resolution. Dock påverkas reliabiliteten negativt om man skulle vara ”omvänt kalibrerad” i det avseendet.

Den komponent som ofta kallas *osäkerhet*, men som vi också skulle kunna kalla *basnivåpoäng*, anger den poäng som man får vid en ren ”basnivåstrategi”, d.v.s. att alltid ange den historiska relativa frekvensen i varje prediktion. I det fallet blir både resolutions- och reliabilitetskomponenterna 0, vilket lämnar UNC som enda bidragande faktor.

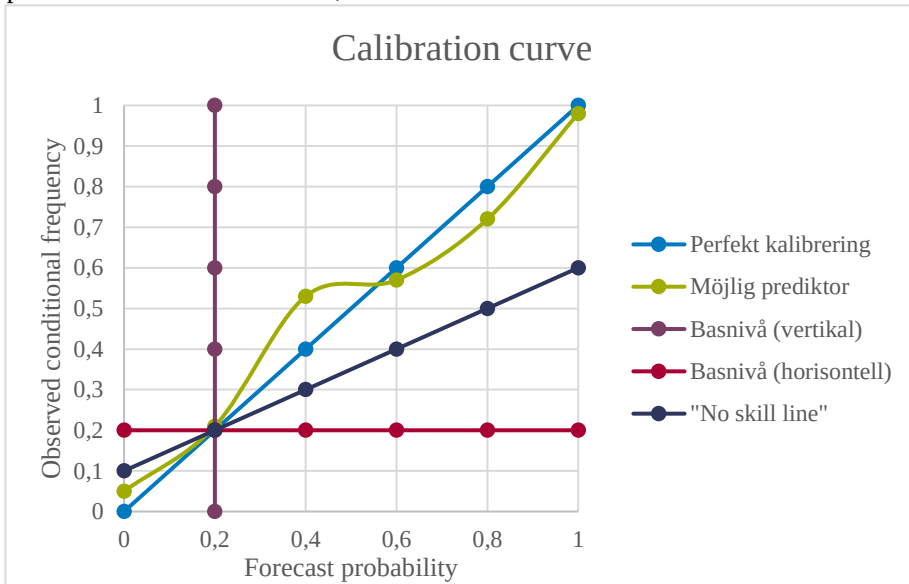
Kalibreringskurva

Ett kompakt sätt att mäta (och visa) prestation vid prediktioner är *kalibreringskurvan*⁴. Den har likheter med den probabilistiska *confusion matrix* som introducerades tidigare (se avsnitt 3.4.1) då den ställer bedömd sannolikhet mot observerade frekvenser.

³ Se t.ex. Tödter (2011).

⁴ Den kallas, utöver *Calibration Curve*, ibland också *Reliability Diagram* eller *Attributes Diagram*.

Diagrammet är mycket informationsrikt och visar flera av de koncept som har diskuterats tidigare. Den ljusblå referenskurvan visar hur en perfekt kalibrerad prediktor skulle se ut över tid, d.v.s. att händelser som tillskrivs en viss



Figur 1 Kalibreringskurva (Tödter, 2011)

sannolikhet också observeras ha motsvarande snittfrekvens. Den gröna kurvan exemplifierar hur en faktisk prediktor typiskt avviker från perfekt kalibrering. De områden där den gröna kurvan ligger ovanför den ljusblå är sådana där prediktorn är för "försiktig", d.v.s. där den faktiska snittfrekvensen ligger över prediktorns sannolikhet. När den gröna kurvan ligger under den ljusblå har vi den omvända situationen: prediktorn är för "aggressiv" och ansätter högre sannolikheter än händelsernas snittfrekvens.

Om man ritar in den s.k. *basnivåstrategin*, d.v.s. att alltid ange basnivån som sin valda sannolikhet, så får man inte en linje, utan en enskild punkt där de två basnivålinjerna korsas (vid $\{0,2;0,2\}$ i figuren). Punkten ligger förstås på den perfekta kalibreringskurvan, men illustrerar också frånvaron av resolution genom att inte ha någon utsträckning i x-led.

Slutligen har vi den mörkblå kurvan: *no skill line*. Så länge en punkt på kalibreringskurvan ligger mellan *no skill line* och den vertikala (vinröda) basnivålinjen så ger den ett fördelaktigt bidrag till prediktorns Brierpoäng eller sannolikhetspoäng. Konditionala frekvenser utanför detta område försämrar alltså totalpoängen.

MCC och F-poäng

Matthews korrelationskoefficient (MCC) är ett värde som ibland används för att sammanfatta en binär (2 x 2) förvirringsmatris (se avsnitt 3.4.1). MCC beräknas utifrån alla de fyra ”fälten” – TP, TN, FP och FN – och antar värdet 1 för en perfekt prediktor, 0 för en som är jämförbar med slump, och -1 för en prediktor som ständigt ger felaktiga prediktioner⁵. Genom att MCC beräknas utifrån hela matrisen så undviker man en del märkliga och missvisande resultat som annars kan uppstå, t.ex. höga *accuracy*-siffror för triviala modeller när det finns en stor obalans mellan de faktiska antalen positiva och negativa.

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

F-poäng är ett annat sammanfattande mått för en förvirringsmatris och anges typiskt med en siffra, såsom F1, F2 eller F0,5. Denna siffra (parameter) återfinns som β i formeln nedan.

Måttet är det s.k. *harmoniska medelvärde* mellan *precision* och *recall*. Siffran β anger det relativa viktförhållandet mellan komponenterna: en siffra över 1 ger större vikt till *recall* medan en siffra under 1 ger mer genomslag åt *precision*.

Eftersom F-poäng beräknas ifrån *precision* och *recall* tar måttet inte hänsyn till mängden sanna negativa, TN, vilket kan göra det olämpligt i vissa sammanhang.

$$F_{\beta} = (1 + \beta^2) \frac{\text{precision} \times \text{recall}}{(\beta^2 \times \text{precision}) + \text{recall}}$$

Poängen kan också skrivas på följande sätt:

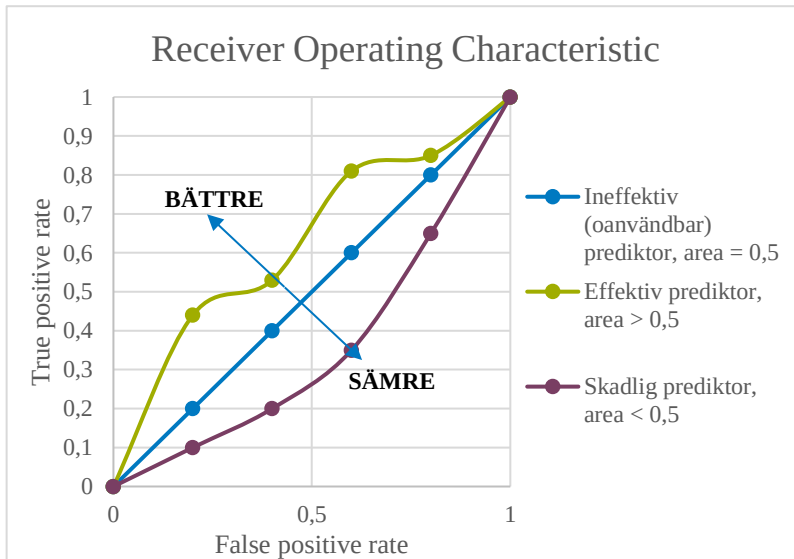
$$F_{\beta} = \frac{(1 + \beta^2) \times TP}{(1 + \beta^2) \times TP + \beta^2 \times FN + FP}$$

⁵ Och därmed är perfekt ”omvänt” kalibrerad.

Areamått

En annan grupp av vanligt förekommande prestationsmått fås genom att ta fram kurvor över förhållanden mellan utvalda nyckeltal och dels studera kurvan, dels beräkna arean under dessa kurvor. Det sistnämnda ger, i de fall det är meningsfullt, ett sammanfattande nyckeltal för förhållandet i grafen.

Ett av dessa nyckeltal kallas *Area Under Curve – Receiver Operating characteristic* (AUC-RO) eller *Area Under Receiver Operating Characteristic* (AUROC). Den s.k. ROC-kurvan (se figur 2) är en graf över TPR mot FPR, d.v.s. *true positive rate* mot *false positive rate*. Anledningen till att just denna kurva är intressant är att det finns en inneboende konflikt mellan att maximera antalet sanna positiva och att minimera antalet falska positiva⁶.



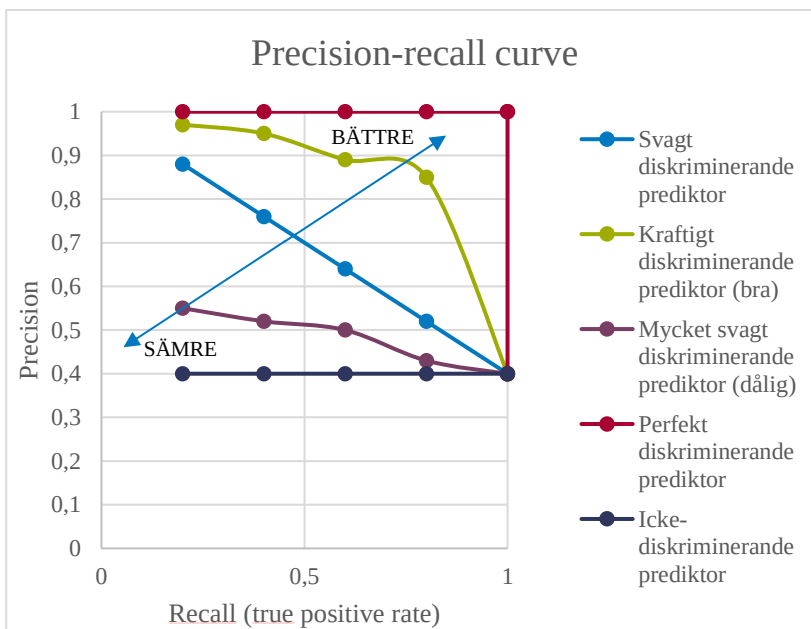
Figur 2 ROC-kurva

Kurvan visar hur väl en given prediktor klarar den avvägningen. När man sedan tar arean under kurvan så kan den tolkas som sannolikheten att en slumpmässigt

⁶ Detta går att se genom att överväga två extrema strategier som är varandras spegelbilder: (1) att klassificera alla fall som positiva och (2) att klassificera alla fall som negativa. Den första ger trivialt ett maximalt antal sanna positiva, vilket är bra, men det kommer till priset av ett maximalt antal falska positiva. Den andra strategin ger noll falska positiva, men här kommer det istället till priset av noll sant positiva. Realistiskt, om än inte teoretiskt (det kan finnas en perfekt, icke trivial strategi), kommer varje annan strategi att innebära en avvägning mellan sannolikheten att få sanna positiva (som man vill ha hög) och sannolikheten att få falska positiva (som man vill ha låg).

vald positiv får en högre sannolikhet för att vara just positiv än en slumpmässigt vald negativ. Värden nära 1 är positiva, värden nära 0,5 indikerar att vinsten med att använda klassificeraren är liten jämfört med att bara slumpmässigt tilldela klasser, och ju längre under 0,5 arean är desto mer missledande är klassificeraren (i snitt).

Den andra vanligt förekommande areaberäkningen är *Area Under Curve – Precision Recall* (AUC-PR) eller *Area Under Precision Recall Curve* (AUPRC). Ibland används också förkortningen AUPR. Här är det alltså andelen korrekta positiva bedömningar (Pr) som ställs emot *recall*, som bara är ett annat namn för TPR. Avvägningen står mellan sannolikheten att klassificera ett positivt fall som sådant (TPR) och sannolikheten att en klassificering som positiv faktiskt är korrekt (Pr). Arean under PR-kurvan (se figur 3) är snittprecisionen tagen över flera (alla) nivåer på TPR. En stark prediktor (i detta avseende) har en PR-kurva som ligger nära fältets övre och högra gränser, medan en svag ligger nära den s.k. basnivån, vilket är andelen positiva i den population man har studerat.



Figur 3 PR-kurva

En fördel som ibland lyfts fram för AUPRC relativt AUROC är att den förra är bättre på att belysa skillnaden mellan prediktorer (eller klassificerare) när andelen positiva i en studerad mängd är liten. Ett klassiskt exempel är just studier av större väpnade konflikter; konflikter är förhållandevis ovanliga, så det finns betydligt fler negativa än positiva. I sådana sammanhang är det fullt möjligt att uppnå ett högt AUROC-värde, trots att AUPRC-värdet är lågt.

Referenser

Tödter, J. (2011) *New Aspects of Information Theory in Probabilistic Forecast Verification*, Master Thesis in Physics (Frankfurt/Main: Goethe University).

Winkler, R.L. & Murphy, A.H. (1968) “Good Probability Assessors”, *Journal of Applied Meteorology* 7:751-758.

Appendix B

Random forest

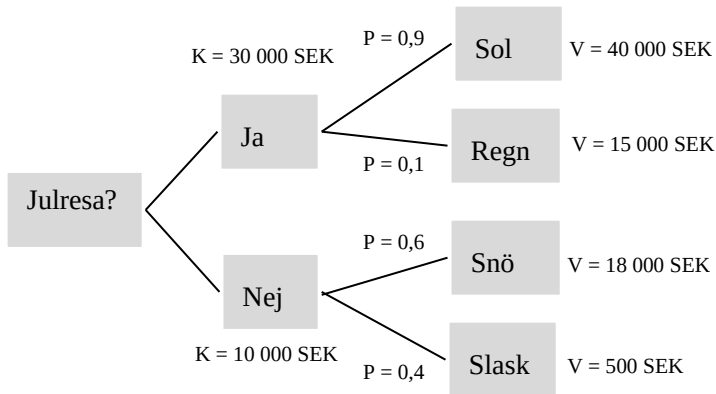
*Random forest*¹ är en ensemble av s.k. *beslutsträd* och vi inleder därför med att förklara detta begrepp.

Beslutsträd som beslutsstöd

Ett beslutsträd är ett enkelt men ändå kraftfullt beslutsstödsverktyg som består av en hierarkisk trädstruktur. Trädet utgår från en startnod (så kallad *rot*) som via utgående grenar är kopplade till flera noder (så kallade *barn*). Varje nod är eventuellt i sin tur kopplade till sina respektive barn tills man kommer till noder som saknar barn; dessa slutnoder brukar kallas *löv*. Trädstrukturen är en modell för hela processen, d.v.s. beslutskedjan som undersöks.

Varje nod representerar en händelse (t.ex. en beslutspunkt) i processen och de utgående grenarna representerar olika konsekvenser (t.ex. resurskostnader) som följer händelsen. Konsekvenserna kan vara antingen deterministiska eller stokastiska. Genom att jämföra olika löv kan man välja det mest fördelaktiga alternativet och identifiera den strategi (sekvens av beslut) som med största sannolikheten leder från roten till det valda lövet. Figur 1 visar ett hypotetiskt beslutsträd för en eventuell julresa. För att kunna fatta ett beslut behöver man även definiera en funktion som översätter eländet som en grå julafton hemma och en regnig semesterresa orsakar till kronor, d.v.s. löven behöver kunna omsättas till någon form av värde som anger hur lycklig eller olycklig man blir om faktiskt hamnar i dem.

¹ Breiman (2001).



Figur 1: Beslutsträd för en julresa. "K" står för kostnad, "P" för sannolikhet och "V" för utfallets subjektiva värde.

Beslutsträd som prediktiv modell

Som prediktiv modell är beslutsträd en modell som kan tränas med maskininlärningsalgoritmer och förutsäger värdet på en målvariabel (beroende variabel) baserad på flera inputvariabler (oberoende variabler). Träningen, eller beslutsträdsinlärningen, går ut på att träningsdata² successivt delas in i delmängder baserat på något kriterium som på "bästa" sätt fördelar datapunkterna. Denna process upprepas för varje delmängd på ett rekursivt sätt tills alla element i en nod har samma värden för målvariabeln, eller när delning inte längre förbättrar värdet till prediktionerna. Det färdigbyggda trädet används sedan för att förutsäga målvariabeln utifrån givna oberoende variabler.

Beslutsträd kan användas både för klassificering där den beroende variabeln är en uppsättning av diskreta variabler och för regression där den beroende variabeln antar kontinuerliga värden. I det första fallet kallas beslutsträdet *klassificeringsträd* och i det senare fallet *regressionsträd*. Det finns flera olika algoritmer för att producera beslutsträd; en bra startpunkt för att lära sig mer om beslutsträd är Quinlan (1986).

² En samling datapunkter där både oberoende och beroende variabler antar kända värden.

Ensembler av beslutsträd

En *random forest* består alltså av en uppsättning beslutsträd som konstrueras under träning och som vid prediktion anger typvärdet (vid klassificering) eller medelvärdet (vid regression) av prediktionerna från beslutsträden. En fördel med *random forest* jämfört med enskilda beslutsträd är att dessa ensembler i en mycket mindre grad drabbas av överanpassning³ till träningsdata.

Random forest är en s.k. ”bagging”-metod och ska inte förväxlas med ”boosting”-tekniker som diskuteras i nästa avsnitt.

XGBoost och GBM

Innan vi går in på XGBoost (*eXtreme Gradient Boosting*) och GBM (*Gradient Boosted Machine*) förklarar vi begreppet *boosting* som är grunden för både XGBoost och GBM.

Boosting

Boosting⁴ är en metod för att omvandla s.k. *svaga* modeller (d.v.s. modeller som bara är något bättre än slumpen) till *starka* modeller, vilket har bevisats vara möjligt.⁵ Det finns flera olika *boosting*-algoritmer, men den mest intuitiva algoritmen går ut på att man tränar och kombinerar en uppsättning av svaga modeller av samma typ genom slumpmässig provtagning. Man skapar N stycken träningsuppsättningar genom slumpmässig provtagning med återläggning från den ursprungliga datauppsättningen. Det som skiljer *boosting* från *bagging* (som *random forest* är ett exempel på) är att man i *boosting* ger datapunkterna olika sannolikhetsvikter, något som innebär att vissa av dem kommer med oftare än andra i de nya träningsdatauppsättningarna (se figur 2).

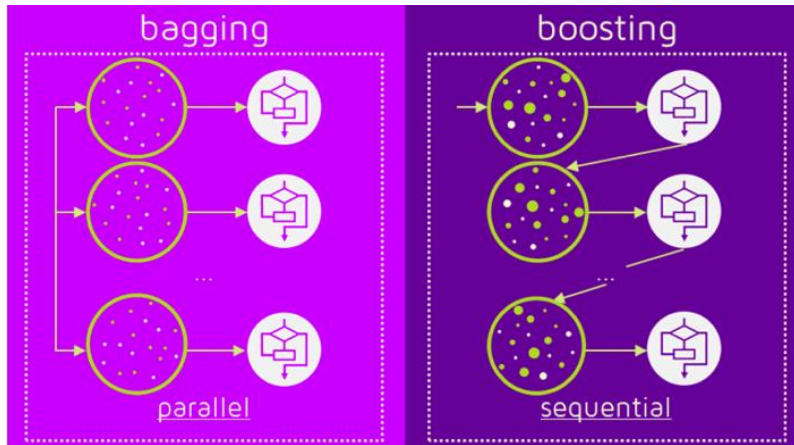
I *boosting* byggs modellerna på ett sekventiellt sätt och vid träning av varje modell tar man hänsyn till den tidigare modellens framgång. Efter varje träningssteg fördelas sannolikhetsvikterna igen och de data som modellen tidigare har misslyckats med kommer att få ökade vikter för att betona de svåraste fallen. På så sätt kommer efterföljande modeller att fokusera på dem

³ Inom statistik innebär överanpassning (eng. *overfitting*) att man producerat en modell som överensstämmer alltför nära med en viss datauppsättning och därför fungerar illa på nya data.

⁴ Kearns, M. (1988) ”Thoughts on Hypothesis Boosting”, opublicerat manuskript, <https://www.cis.upenn.edu/~mkearns/papers/boostnote.pdf>.

⁵ Schapire (1990).

under träningen. Vid själva prediktionen använder man sedan alla modellerna samtidigt, viktar de enskilda förutsägelseerna och kombinerar dem.⁶



Figur 2: Jämförelse mellan *bagging* och *boosting*. *Bagging* sker parallellt med jämnt viktade data. *Boosting* sker istället seriellt med ojämnt viktade data.

Gradient boosting

Både XGBoost⁷ och GBM⁸ är *gradient boosting*-metoder, vilket i sin tur är en speciell typ av *boosting*. Det finns dock vissa prestandapåverkande skillnader i algoritmerna, men dessa ligger utanför denna rapports ramar. Den intresserade läsaren hänvisas till Chen & Guestrin (2016).

Den största skillnaden mellan *gradient boosting*-algoritmer och andra typer av *boosting* (t.ex. den metod som beskrevs ovan) är hur de två algoritmerna identifierar bristerna hos de svaga modellerna (t.ex. beslutsträd). Medan man i "vanlig" *boosting* fokuserar på datapunkter så arbetar gradient boosting-algoritmer med lutningen hos den s.k. *förlustfunktionen*. Förlustfunktionen, vilken varierar beroende på tillämpningsområde, är ett mått på hur bra modellen (med de nuvarande parametrarna) är på den avsedda prediktionen.

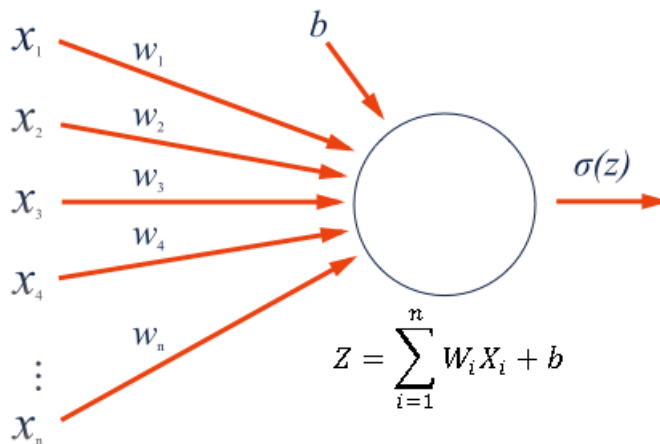
⁶ Jämför med *random forest* där man räknar typvärde eller medelvärde.

⁷ "Scalable and Flexible Gradient Boosting", <https://xgboost.ai/>.

⁸ Friedman (2001).

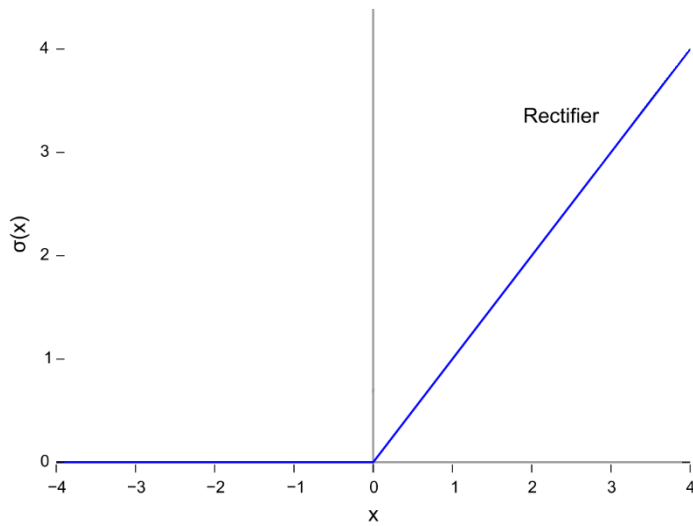
Artificiella neuronnät

Ett artificiellt neuronnät (ANN) är en funktionsapproximationsmetod som mappar en uppsättning indata (oberoende variabler) till en uppsättning utdata (beroende variabler). Grundbyggstenen i ett neuronnät är artificiella neuroner (kallas också noder) som mappar flera ingångsvärden till ett utgångsvärde (se figur 3).



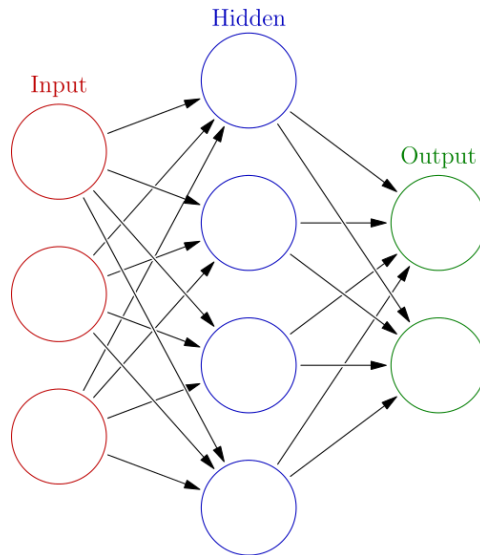
Figur 3 En enskild neuron, med n stycken ingångsvariabler (x), vikter (w) och bias (b). Innan aktiveringsfunktionen appliceras beräknas den viktade summan och bias läggs till.

Mappningsfunktionen består av en viktad summa av indata samt en biasterm. Vikter och bias är från början okända parametrar som beräknas under träningen. Den resulterade summan bildar inputvärdet till en *aktiveringsfunktion* som genererar neuronens slutliga utdata. Det finns flera olika aktiveringsfunktioner, och de väljs typiskt i samband med typen av neuronnät. En mycket vanligt förekommande aktiveringsfunktion är en *rectifier* (se figur 4).



Figur 4 Rectifierfunktion. Värden av funktionen är lika med 0 för negativa x och lika med x annars.

Ett neuronnät består normalt av flera lager av neuroner: ett in-lager, ett ut-lager och ett eller flera dolda lager (se figur 5). Neuronnätet sägs att vara "djupt" om det finns fler än ett dolt lager, men numera används regelmässigt neuronnät med stora antal dolda lager.



Figur 5 Ett neuronnät med ett dolt lager.⁹

Träningen av ett nätverk innebär att bestämma alla parametrar (vikter) i nätverket. I början av träningsprocessen sätts parametrarna till slumpmässiga tal och nätverket matas med träningsdata, d.v.s. kända kombinationer av indata och (önskade) utdata. För varje inmatat träningsdatum beräknas differensen mellan producerade utdata och önskade utdata. En iterativ algoritm ser till att successivt modifiera vikterna i rätt riktning så att denna differens minimeras. Om träningen går bra kan nätverket inte bara producera korrekta utdata för träningsdata utan även för indata som det inte har exponerats för tidigare. En modell som kan skapa korrekt utdata för icke observerad data sägs att kunna generalisera.

Neuronnät kan användas både för regression (funktionsapproximation) och klassificeringsproblem. I det senare fallet, består utdata av neuronnätet av sannolikheten för att indata tillhör varje klass.

Beskrivningen här handlar om den enklaste formen av artificiella neuronnät s.k. *framåtmatande* (eng. *feedforward*). I dessa finns inga *cykler*, vilket kan sägas motsvara att inga data passerar ”bakåt” i nätverket. I framåtmatande neuronnät

⁹ Glosser.ca. Denna bild är licensierad under Attribution-ShareAlike 3.0 Unported (<https://creativecommons.org/licenses/by-sa/3.0/legalcode>).

matas indata in i det första lagret och passerar varje nod högst en gång innan resultaten når ut-lagret.

Det finns många olika neuronnäsarkitekturer. Dessa skiljer sig åt i hur neuronerna är kopplade till varandra, antalet dolda lager, huruvida det finns en återkoppling mellan noderna, typer av aktiveringsfunktioner och mycket annat. Den intresserade hänvisas till Gurney (1997).

Referenser

Breiman, L. (2001) "Random Forests", *Machine Learning* 45:5-32.

Chen, T. & Guestrin, C. (2016) "XGBoost: A Scalable Tree Boosting System", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Friedman, J.H. (2001) "Greedy Function Approximation: A Gradient Boosting Machine", *The Annals of Statistics* 29(5):1189-1232.

Gurney, K. (1997) *An Introduction to Neural Networks* (London: UCL Press), https://www.inf.ed.ac.uk/teaching/courses/nlu/assets/reading/Gurney_et_al.pdf.

Quinlan, J.R. (1986) "Induction of Decision Trees", *Machine Learning* 1:81-106.

Schapire, R.E. (1990) "The Strength of Weak Learnability", *Machine Learning* 5(2):197-227.

Appendix C

Nedan återfinns ViEWS-projektets publicerade prestationsmått.

Tabell 1 Prestationsmått för ViEWS¹

Modell	AUROC	Brierpoäng	AUPR
CM level, SB baseline	0,59465	0,17544	0,34507
CM level, SB ensemble	0,93066	0,09248	0,79742
CM level, OS baseline	0,62301	0,13593	0,31912
CM level, OS ensemble	0,92341	0,07582	0,78521
CM level, NS baseline	0,65748	0,09701	0,32645
CM level, NS ensemble	0,92338	0,06157	0,67972
PGM level, SB baseline	0,53370	0,00632	0,00513
PGM level, SB ensemble	0,95198	0,00589	0,24656
PGM level, OS baseline	0,56548	0,00523	0,00468
PGM level, OS ensemble	0,95484	0,00503	0,19899
PGM level, NS baseline	0,53694	0,00394	0,00250
PGM level, NS ensemble	0,89649	0,00393	0,04929

CM level = country-month level

PGM level = PRIO-grid-month level

SB = state-based violence

OS = one-sided violence

NS = non-state violence

¹ "How well do we do?", <https://www.pcr.uu.se/research/views/current-forecasts/how-well-do-we-do/>.

Appendix D

Vidare reflektioner kring användning av kvantitativa data inom Afrikaprojektets analyser

Med tanke på den tematik som behandlas inom Afrikaprojektet – konflikter, rebeller, terrorism, inbördeskrig – finns stora möjligheter att ta hjälp av existerande databaser. Det finns också ett värde i att använda existerande databaser för att sätta in enfallsstudier i ett större regionalt sammanhang. Detta är särskilt önskvärt i regioner där konflikter har en regional dynamik och spridits utanför konflikternas ursprungsländer.

Användning av kvantitativa data för att förbättra enfallsstudier

Den kanske enklaste och mest användbara metoden inom Afrikaprojektets studier, och då i synnerhet sådana med prediktiv ansats, är att i större omfattning nyttja kvantitativa databaser inom ramen för kvalitativt inriktade studier. Detta gäller i synnerhet inom R-rapporterna, där mer än 90 procent av studierna saknar substantiella kvantitativa inslag, vilket gör detta till ”långt hängande frukt” analytiskt.

Exakt vilka databaser som är användbara varierar beroende på studiernas inriktning och syfte. Några illustrativa exempel kan ge en fingervisning om hur databaserna med enkla medel kan vara användbara. Inom studier av väpnade grupper och inbördeskrig, kan databaser illustrera konfliktintensitet, regional spridning och eventuell spridning över gränser ge tydlig beskrivning av konfliktens utveckling över tid. I Torbjörnsson och Jonsson (2017) används exempelvis *Global Terrorism Database* (GTD) för att illustrera antalet terroristattentat utförda av Boko Haram samt antalet offer mellan 2009 och 2016. Utöver detta kan analysen även disaggregeras på regional nivå, exempelvis utifrån typ av attacker, händelser eller måltavlor. Givet att förövare bakom attacker har rapporterats, kan dataseten även ge en fingervisning om vilken väpnad grupp som utgör störst hot i ett visst land, vilket kan fungera som en indikator på regional spridning av en konflikt. Data från ACLED (se avsnitt 2.4) kan användas på liknande sätt.

Inom studier av enskilda länder där analysen rymmer en bedömning av säkerhetsläget i landet och risken för en ny konflikt skulle prediktioner från ViEWS eller ”vanliga” data från UCDP:s konfliktdatabas kunna användas för att uppmärksamma eskalering av olika typer av våld. Denna information hade exempelvis kunnat användas i Gunnarson & Gasinska (2019) där en del fråga i studien var att identifiera säkerhetspolitiska utmaningar i Niger och i

Sahelregionen. Källorna ovan hade här kunnat användas för att uppmärksamma eskalering av våld mellan stater och rebellgrupper vid Nigers gränser till Mali, Burkina Faso, Nigeria och Libyen. De hade också kunna användas för att beskriva inslag av ökat våld mot civila, framför allt vid gränsen till Mali och Nigeria, vilket lett till ett ökat antal internflyktingar i Niger.

Händelsedatabaser ger en relativ precision i bedömning av exempelvis konfliktintensitet och geografisk koncentration, på ett sätt som är svår eller omöjlig att uppnå med enbart kvalitativa data eller egen datainsamling.¹ Därmed ger de även en mer robust empirisk bas utifrån vilken det är möjligt att, exempelvis, göra prediktioner om konfliktintensitet, spridning osv. Däremot kan data inte ersätta behovet av teoriansknytning, resonemang om vilka variabler som *förklarar* ett visst utfall, eller tolkning av data.

Även för studier som analyserar eller utvärderar fredsfrämjande insatser kan denna typ av data bidra till analysen. Exempelvis kan konfliktintensitet, regional spridning av våldet, respektive spridning av våldet vara viktiga kontextvariabler som är helt avgörande för om en insats har förutsättningar att lyckas med sitt uppdrag eller ej. Exempel på kontextuella faktorer som påverkar insatsers förutsättningar är om konflikten eskalerat eller deeskalerat sedan en insats inleddes; huruvida säkerhetsläget i regionen där insatsen genomförs är gott, dåligt, eller riskerar att försämrats på grund av spridning av konflikten; samt huruvida en regionalisering av fredsfrämjande kan förklaras av att allt fler konflikter sprids regionalt över landsgränser. På motsvarande sätt kan data om institutioners kvalitet eller sårbarhet, regional utvecklingsnivå och jämlikhet bidra med viktig bakgrundsinformation som påverkar konflikters varaktighet över tid.

Vårt argument är att kvantitativa data förbättrar prediktion, vare sig den är kvalitativ eller kvantitativ, eftersom det är svårt, rent av omöjligt, att förutspå en utveckling om man inte har relativt precisa mått på situationen i nuläget eller tillgång till data över hur en konfliktsituation utvecklats över tid. För prediktioner på kort till medellång sikt torde beskrivande kvantitativa data, i kombination med teoretiskt baserade resonemang om orsakssamband, kunna avsevärt förbättra de enfallsstudier som dominerar Afrikaprojektets produktion i dag. Teoretiskt grundade resonemang skulle även kunna bidra till mer avgränsade och ”vassare”

¹ Både GTD och ACLED har med rätta kritiserats för vissa brister. Typiska fel kan inkludera att samma incident inkluderats flera gånger i datasetet, saknade värden, samt att våld i perifera, rurala delar av landet har mindre sannolikhet att inkluderas i databasen. Dessa felkällor korrigeras dock svårligen av egen datainsamling från samma källor (huvudsakligen medierapportering), även om källkritisk granskning av dataseten är önskvärd.

studier som inriktas på relevanta variabler och skeenden som identifieras som särskilt kritiska i samhällsvetenskaplig forskning.

Det finns flera exempel på databaser som med fördel kan användas för att stärka Afrikaprojektets, men även övriga projekts, analyser. För analys av länders förutsättningar till långsiktig stabilitet finns ett stort antal databaser med information om regimstabilitet och institutioners kvalitet², statens legitimitet³, förekomst av korruption⁴ eller demokratis olika dimensioner.⁵ Om syftet är att undersöka socio-ekonomiska förutsättningar för fred finns tillgång till databaser om fattigdom, utbildning och hälsa.⁶ Om studien syftar till att studera genderaspekter finns tillgång till databaser som mäter ojämlikhet mellan könen eller kvinnors möjligheter till politiskt deltagande.⁷ För bedömningar av risken för klimatkonflikt, migration eller svält finns tillgång till klimatdatabaser.⁸ Dataset om migration även hos organisationer som International Organization of Migration (IOM) och FN:s flyktingorgan UNHCR.

Användning av kvantitativa data för att förbättra prediktion i regionala studier

Även i regionala eller tematiska studier kan existerande databaser bidra till att förbättra prediktion, givetvis beroende på studiens fokus och tillgång till data. I studier av regionala samarbeten som syftar till att begränsa eller förhindra en konflikts regionala spridning kan existerande databaser beskriva graden av spridning av en väpnad grupp bortom landets gränser. Genom ViEWS grafiska presentationer av sina prediktioner kan man se hur konflikten i Mali har bedömts sprida sig till grannländerna Niger och Burkina Faso. I diskussioner om regionala säkerhetssamarbeten kan kvantitativa data bidra med information om problemets omfattning och geografiska koncentration.

² Quality of Government, Freedom House.

³ Fund for Peace; Afrobarometer. Det finns även liknande databaser för andra regioner, med varierande tillgänglighet (Eurobarometer, Asiabarometer, Latinobarometro).

⁴ Transparency International, World Bank Indicators.

⁵ V-Dem, Freedom House, Fund for Peace: State Fragility Index.

⁶ World Bank: World Development Indicators, UNDP: Human Development Index eller OECD.

⁷ UNDP: Gender Equality Index; V-DEM: Women's Political Empowerment Index.

⁸ GISTEMP, the Global Surface Temperature Change data of the National Aeronautics and Space Administration Goddard Institute for Space Studies (NASA-GISS), EM-DAT, The International Disaster Database.

Rörande fördelar och nackdelar med regionala samarbeten inom fredsfrämjande kan data om konflikters intensitet över tid ge en fingervisning om utfallet av insatserna. Om studiens fokus är transnationella säkerhetshot kan en kombination av databaser ge en god överblick av problemets omfattning, geografiska koncentration, samt utveckling över tid.

Användning av kvantitativa data för långsiktiga, prediktiva studier

Som framgått av tidigare skrivningar är det få studier som har en explicit prediktiv ansats inom Afrikaprojektet. En studie av särskild betydelse är ”Säkerhetspolitiska trender i Afrika söder om Sahara” som tagits fram inom ramen för SLO-projektet (Strategisk långsiktig omvärldsanalys) på uppdrag av Försvarsmakten. Analysen kombinerar på ett intressant sätt kvantitativ analys av långsiktiga trender med ett urval fallstudier av Etiopien, Nigeria och Sydafrika. Den tematik som valts ut tar sitt avstamp i olika teorier om internationell politik, bland annat realism, liberalism, moderniseringsskolan, konstruktivism, gender och institutionell teori. Fokus är på långsiktiga trender med tonvikten på relativt trögrörliga indikatorer. Som tidigare nämnts saknas exempelvis faktorer i landanalyserna som lyfter fram länderna interna politiska dynamik, som maktövertagande, relationen mellan regering-opposition, politiska partier, politisk splittring eller förekomst av elektoralt våld.

Beroende på tidsperspektiv i prediktioner bör en metod utvecklas som tar hänsyn till både långsiktiga strukturella förändringar och trögrörliga indikatorer, men också mer dynamiska och snabbrorliga mått som kan fånga in skeenden, händelser och aktörer. Detta förefaller särskilt viktigt om syftet är att studera förändringar i konflikters intensitet eller konsekvenser, såsom föreslås här.

Exempel på strukturella och mer trögrörliga faktorer är t.ex. indikatorer som beskriver länders regimtyp, grad av demokrati i dess olika dimensioner, förekomsten av korruption och förutsättningar för god styrning, omfattningen av demokratiska värderingar hos befolkningen, ekonomisk utvecklingsnivå, långsiktiga klimatrelaterade förändringar (medeltemperatur), förekomst av parallella maktstrukturer i form av organiserad brottslighet eller data som belyser om landet är transitland för vapen, narkotika, människosmuggling eller andra destabiliserande faktorer.

Mer dynamiska variabler omfattar exempelvis faktorer som belyser förändringar i den politiska interaktionen mellan regering och opposition⁹, uppkomst och spridning av folkliga protester, allmänna val, förekomst av våld före, under och efter allmänna val. För att fånga in klimatrelaterade effekter kan man istället för att analysera medeltemperatur – som är ett trögrörligt mått – observera längden på torkperioder, översvämningar, stormar, minskade skördar, vattentillgång världsmarknadspriser på stapelvaror som ris eller vete. Dynamiska mått kan även beskriva migration som ökar, minskar, ändrar riktning eller ”fastnar” på olika platser beroende på ökade kontroller, stängda gränser eller andra problem som påverkar migranternas resrutt. Ökade eller minskade beslag på narkotika skulle kunna vara en möjlighet att belysa närvaro av organiserad brottslighet och eventuellt även kunna indikera att maktrelationerna mellan organiserad brottslighet och stat har förändrats. Dataset som belyser vapenhandel och försvarsutgifter är givetvis också intressanta.

Referenser

Gunnarson, C. & Gasinska, K. (2019) ”Niger i cyklonens öga – vad kan EU göra?”, FOI Memo 6782.

Torbjörnsson, D. & Jonsson, M. (2017) ”Boko Haram. On the verge of defeat or a long-term threat?”, FOI-R--4488--SE.

⁹ Denna aspekt undersöktes t.ex. i projektet CRISP (se Ward et al., 2013a), men resultaten är inte längre tillgängliga.

Appendix E

Kodning av publikationer inom Afrikaprojektet 2014-2019

A. Författare

B. Titel

C. Publikationsår

D. Typ av dokument

Brief

Rapport

”Ej tillgänglig” innebär att vissa dokument inte finns tillgängliga på webben.

E. Antal sidor

F. Region

Nordafrika

Sahel

Västafrika

Marocko, Algeriet, Tunisien, Libyen, Egypten
Mauretanien, Mali, Burkina Faso, Niger, Tchad
Nigeria, Benin, Togo, Ghana, Côte d’Ivoire,
Liberia, Sierra Leone, Guinea, Guinea-Bissau,
Senegal

Afrikas Horn

Djibouti, Eritrea, Etiopien, Somalia, Kenya,
Sudan, Sydsudan, Uganda, Kenya

Centralafrika

Kamerun, CAR, Ekvatorial-Guinea, Gabon,
Republiken Kongo, DRK, Burundi, Rwanda

Södra/östra Afrika

Tanzania, Zambia, Angola, Namibia, Botswana,
Zimbabwe, Mocambique, Sydafrika, Madagaskar

G. Länder

Kommentar: Saknas landspecifikation för vissa studier, t.ex. Afrikanska säkerhetsstrukturer, som omfattar flera länder. Dessa studier klassificeras ”ej specificerat”.

H. Huvudtematik

Skriv 3-4 nyckelord för vad studien handlar om.

I. Görs prediktion i studien?

Nej

Bakåtblickande analys som i huvudsak syftar till

att förstå nuet

I viss mån

Framtida utveckling diskuteras, t.ex. i slutsatsen

Ja	Uttrycklig ambition att försöka bedöma framtida utveckling uttrycks i frågeställning eller under egen rubrik
----	--

Insatsanalyser: I de flesta insatsanalyser görs policyrekommendationer vilka är framåtblickande. Dessa studier klassificeras ”i viss mån”.

J. Metod: Kvant/kval

Kvalitativ
Kvantitativ
Kombination kvant-kval

K. Teorianvändning

Saknas	Saknas referenser till teorier eller begrepp
I viss mån	Referenser görs till teori, men strukturerar inte analysen. Teorin kan vara implicit, t.ex. genom att teoretiska begrepp används
Stark	Teorin strukturerar analys, d.v.s. teorikonsumtion
Mycket stark	Teoriprovning

L. Metod: Beskrivande/förklarande

Beskrivande
Förklarande
Beskrivande och förklarande

M. Metod/Materialinsamling

Skrivbord	Om inte annan materialinsamling nämns
Intervjuer	
Utvärdering	

FOI är en huvudsakligen uppdragsfinansierad myndighet under Försvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.



FOI
Totalförsvarets forskningsinstitut
164 90 Stockholm

Tel: 08-55 50 30 00
Fax: 08-55 50 31 00

www.foi.se