**The ISCRI : A new real-time index**

**to track Inter-State Conflictual Relations**

The recent months have underlined how geopolitical tensions between countries can influence trade and capital flows and, more broadly, world GDP. To help its stakeholders to get a timely comprehensive geopolitical picture, SKEMA Business School has decided to create a new index aiming at tracking in real-time the conflictual relationships between any countries in the world, especially those critical in the world economy system. The bilateral **Inter-State Conflictual Relationships Index (ISCRI)** is based on textual information generated by English sources (e.g. as *Agence France Presse*, *Associated Press*, *Xinhua*, *BBC*, *Washington Post*, *New York Times*, *Google News*), and recorded in the Global Data on Events, Location, and Tone (GDELT) database.

We initiate the first year of our index in 1991, based on GDELT historical “reduced event” database for the period 1991-2013 and then we update the index using daily data files from 2014 onwards. The ISCRI can be understood as the cumulative outcomes of past interstate conflictual events: a stock fed by flows of conflictual events. We produce a Total ISCRI which is composed of a **Verbal ISCRI** and a **Material ISCRI**. A **verbal conflict** is a spokencriticism, threat, or accusation often related to past or future potential acts of material conflict. A **material conflict** is a physical act of a conflictual nature implying for example the exhibition of force posture, reduction of relations, coercion, assault, fight or the use of mass violence.

The interpretation of the index is simple but must be contextualised. If total ISCRI increases, it means that interstate relations are becoming more conflictual. If total ISCRI does not change, the conflictual nature of interstate relations is likely to be stable. If total ISCRI decreases, it implies that interstate relations are becoming less conflictual. A greater understanding of the severity of conflict escalation can be achieved by monitoring both verbal and material ISCRIs. Note that **we do not recommend using the indicator to compare ISCRI between countries** as each interstate relation is context-specific, with events not having the same weight or meaning across country pairs, and some interstate relations may receive more media attention than others. The real informational value of our indicator lies in its **time-series variation, which ought to serve as an early warning system**.

As an illustration of the ISCRI, we examine the evolution of Iran-Saudi Arabia conflictual relations between 2015 and 2023 in the Figure below (*we provide a more elaborated case study CHN-USA after the methodological note*). We take here the average of the directional ISCRI values for this pair. **Our indicator appears to perform well, capturing periods of (verbal and material) escalation and periods of de-escalation**. For example in September 2016, after several months of tensions, Saudi Arabia strongly criticises Iran for creating regional chaos by `exporting revolution’. In September 2019, a Saudi Arabian petrol refinery undergoes a drone attack by Yemen fighters allegedly supported by Iran. On the other hand, since mid-2020, diplomatic relations have become more cordial, with notably a bilateral meeting in September 2021.



Overall, **the ISCR indexes are available from 2014 onwards for virtually all country-pairs in the world**. Its methodology is simple and in line with best practices (*see methodological note below*). It is updated each month but has a near real-time frequency. It can be used in a descriptive manner to assess diplomatic tensions or as an explanatory variable in econometric models investigating, for instance, the determinants of international trade.

This methodological note describes the choices guiding the construction of the ISCRI indicator developed by Rodolphe Desbordes, Laurent Ferrara, Frédéric Munier and Francis Wolinski.

Following [Anderton & Carter (2011)](https://www.tandfonline.com/doi/full/10.1080/10242694.2010.491677), interstate relations are characterised by various phases. In Figure 1 below, durable peace may be interrupted by episodes of conflicts whose hostility intensity and duration vary with time.

A diagram of peace and conflict

Description automatically generated with low confidence

Figure 1: Source: Anderton & Carter (2011)

**Each phase of the conflict life cycle is associated with a sequence of observable events reflecting the evolution of interstate relations**. For example, a latent conflict (identified by an event e.g. threatening language) may become a sub-war (identified by an event e.g. imposing sanctions) or degrades to a stable peace situation through conflict prevention (identified by an event e.g. an apology). In the vein of [Rodrik et al. (2004)](https://link.springer.com/article/10.1023/b:joeg.0000031425.72248.85) in their distinction between institutions (stock) and policies (flow), it then makes sense to think of **interstate conflictual relations (ISCR) as a stock variable**, which is the cumulative outcome of past interstate conflictual events (ISCE). Hence, the stock of interstate relations can be written as and the evolution of ISCR over time is , where is the impact of inter-state event *i* on inter-state relations and is the rate at which ISCR decay without any additional conflictual events.

[Gerner et al. (1994)](https://academic.oup.com/isq/article-abstract/38/1/91/1785243?login=false) defines an event as *an interaction, associated with a specific point in time, that can be described in a natural language sentence that has as its subject and object an element of a set of actors and as its verb an element of a set of actions, the contents of which are transitive verb* (p.95). In other words, who (source actor) did what to (action on a direct object) whom (target actor) can be identified in media reports (natural language sentence) at a given time. While these coding of events used to be done by human coders, machine coding of events (using for example TABARI software) is now widely employed. A good introduction to automated coding of very large scale political event data can be found in [Schrodt and Yonamine (2012)](https://projects.iq.harvard.edu/ptr/files/schrodt_yonamine_automated.pdf). For example, if the lead of a newswire is

President Reagan has threatened further action against the Soviet Union in an international television program beamed by satellite to more than 50 countries on 10 November 1987

the source actor is the president of the United States (USAGOV), the target actor is the Soviet Union (USR), and the action is “to threaten” (code 130 in the [CAMEO ontology](http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf)). The output of the automatic coding is then 19871110 (date) USA (source country) GOV (primary actor) USR (target country) … (likely no identified primary actor) 130 (CAMEO code).

The events generated by automatic coding [must be aggregated](https://www.allazimuth.com/2017/07/22/a-guide-to-event-data-past-present-and-future/) along three dimensions: actors, actions, temporality. Given our focus on ISR, the following criteria are applied: **i) the source and target countries are different and one of the actors involved in the event belongs to the government; ii) considered actions are purely hostile; iii) daily events are aggregated by month, to reduce noise and capture the likely “diplomatic sequences” involving back-and-forth bilateral actions**. **Note that our definition of ISR as a stock means that past diplomatic relations are taken into account, and by extension, are not acontextual.** For example, the occurrence of an on-going war which is increasingly less reported in news report due to media fatigue will not lead to a sharp drop in our measure of ISR; through the past stock, a `memory’ of the conflict persists.

Some scales, such as the [Goldstein scale](https://journals.sagepub.com/doi/abs/10.1177/0022002792036002007), have been suggested to grade events according to their hostility levels. However, [Schrodt and Gerner (2012)](https://parusanalytics.com/eventdata/books.html) and [Yonamine (2013)](https://etda.libraries.psu.edu/files/final_submissions/8666) consider that this kind of scaling can be contentious since weights may be perceived as arbitrary, anachronistic, and lead to dubious equivalences between a few low-intensity events and one high-intensity event. They advise simply using the counts of events, classified in the four broad categories suggested [by Duval and Thompson (1980)](https://www.jstor.org/stable/2110830?casa_token=cuVQzTGFjZ4AAAAA:WQnW4kGCJ6ZVeR0Wy-VtiGfdWwOugWAAZo1p5MMd8iVzXRZqyn2oOMZ6tgjYlgb11ZPi0U8zcjWF3nJbT0jPWMPJl8aLqoxwX6xvAknmX0eh4LPsogZH): verbal cooperation, material cooperation, **verbal conflict, material conflict**.

More precisely, verbal conflict is a **spoken** criticism, threat, or accusation often related to past or future potential acts of material conflict. and corresponds to CAMEO codes 10-14 (demand; disapprove; reject; threaten; protest) while material conflict is a **physical act** of a conflictual nature and corresponds to CAMEO codes 15-20 (exhibit force posture; reduce relations; coerce; assault; fight; use unconventional mass violence). **We adopt this categorisation of events in verbal and material conflicts.**

As argued by [Gasiorowski (1986)](https://academic.oup.com/isq/article-abstract/30/1/23/1790466), [Lowe (2006)](https://link.springer.com/chapter/10.1007/1-4020-4390-2_12) and [Desbordes (2010)](https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0343.2009.00353.x), it is crucial to use the average number of counts, rather than the sum of counts over a given period, to control for variation in media interest and event sources**. We thus calculate for each day, the fraction of events which are conflictual and then take the monthly average of these daily scores**. The monthly interstate events value varies therefore between 0 and 1. We could have directly calculated, for each month, a similar fraction but, in doing so, we would have diluted the influence of particularly conflictual days. More pragmatically, a day is a natural human unit which can be used in any temporal aggregation, e.g. a 5-week moving average as in [Anderton & Carter (2011)](https://www.tandfonline.com/doi/full/10.1080/10242694.2010.491677). As a **depreciation parameter**, we use , meaning that, in the absence of conflicts, the indicator decreases by 50% and more after 7 months. **Our ISCR indicator (ISCRI)** for month *m* of year *y* between source country S and target country T is then

where is the number of days in a given month, *C* is the daily number of verbal or material conflict interstate events, and *E* is the daily total number of cooperative and conflictual interstate events. If there is no conflictual event or no events on a given day , . We can also decompose this indicator into a `verbal conflict’ ISCR and a `material conflict’ ISCR. This indicator is directional ().

While our approach is grounded in the events data literature, it displays strong similarities with the construction of the *Geopolitical Risk (GPR) Index* proposed by [Caldara and Iacovello (2022)](https://www.aeaweb.org/articles?id=10.1257/aer.20191823). They calculate the share of articles mentioning adverse geopolitical events, based on their own `geopolitical’ dictionary, in leading newspapers (10) published in the United States (6), the United Kingdom (3), and Canada (1). It is time-specific but is not country-specific. While a country-specific GPR index can be calculated, it seems to capture a mix of exposure of countries to global risks and domestic conflictual events important enough to be reported in the press. **In contrast with the GPR index, our approach relies on more sources, makes use of a validated dictionary, is country-pair specific and focuses on interstate relations.** The two indexes are likely to be complementary.

**Our data source is** **the Global Data on Events, Location, and Tone** ([GDELT](https://www.gdeltproject.org/)), described [in Leetaru and Schrodt (2013)](http://data.gdeltproject.org/documentation/ISA.2013.GDELT.pdf). Text sources include sources written in English such as *Agence France Presse*, *Associated Press*, *Xinhua*, *BBC*, *Washington Post*, *New York Times*, *Google News*. Filtering involves avoiding sport or financial news which contain languages similar to those used to describe violent political conflicts as well as reducing the number of duplicate reports of events. [Saz-Carranza et al. (2021)](https://www.globe-project.eu/the-empirical-use-of-gdelt-big-data-in-academic-research_13809.pdf) provides a review of published academic papers which have used GDELT. Other recent examples are [Agarwal and Golley (2022)](https://onlinelibrary.wiley.com/doi/full/10.1111/twec.13252) and [Hinz (2022)](https://link.springer.com/article/10.1007/s10290-022-00461-6).

We use the [GDELT 1.0 Events database](https://www.gdeltproject.org/data.html), which is updated daily and provides worldwide coverage. We only consider the **events mentioned in the lead paragraph** as they are likely to be `true’ events and they are described in a way suitable for automatic interpretation ([King and Lowe, 2003](https://www.cambridge.org/core/journals/international-organization/article/abs/an-automated-information-extraction-tool-for-international-conflict-data-with-performance-as-good-as-human-coders-a-rare-events-evaluation-design/7BF83BD64D633189EBFD600EB05E7F2E)). **We initiate the first year of our indicator in 1991**, based on GDELT historical “reduced event” database for the period 1991-2013 and we then use the daily data files for 2014 onwards; unfortunately a distinction between `lead’ and `other’ events are is not available in the historical database; this is a minor issue given that we simply use the 1991-2013 data to create a starting point for our 2014 onwards ISR values and that **it is the time-series variation of the latter that truly matters.**

**Automatic event coding based on media reports is a noisy process**. A dataset is comprised of rows (records) and columns (variables) with information contained in cells. Following [Biemer (2016)](https://textbook.coleridgeinitiative.org/chap-errors.html#sec:10-2), the total error in the data can then be heuristically expressed as Total error = Row error + Column error + Cell error. A row error may be due to omissions, duplications, and erroneous inclusions. A column error may be due to wrong or erroneous labelling. A cell error may be due to inaccurate, wrong or missing data. As highlighted by [Wang et al. (2016)](file:///C:\Users\Rodolphe%20Desbordes\Dropbox\Risk\Methodological%20note\science.sciencemag.org\content\353\6307\1502), [Demarest and Langer (2022)](https://journals.sagepub.com/doi/10.1177/0049124119882453) and [Raleigh et al. (2023)](https://www.nature.com/articles/s41599-023-01559-4), automatic event coding are particularly prone to row and cell errors, such as false positives with associated mis-coding (events that never happened), duplicates (reporting of a single event numerous times), false negatives (events happened but were not reported), inaccurate information (e.g. geocoded location). The growing presence of (ironically often automated) “fake news” is also a concern.

All these issues do not invalidate the use of event data as long as the objective is clearly defined. As argued by ([King and Lowe, 2003](https://www.cambridge.org/core/journals/international-organization/article/abs/an-automated-information-extraction-tool-for-international-conflict-data-with-performance-as-good-as-human-coders-a-rare-events-evaluation-design/7BF83BD64D633189EBFD600EB05E7F2E), quantitative coding helps to identify, without any specific international knowledge, **potential** interstate conflictual relations which then need to be subject to a thorough qualitative investigation. In our case, this **early warning system** may be less noisy than focusing, for instance, on the occurrence of local political protests, because **we are solely interested in state actors whose conflictual actions are likely to be expressed in a clear diplomatic language or through symbolic events and to be widely reported in the international press**. **We do not recommend using the indicator to compare ISCRI between countries** as each interstate relation is context-specific, with events not having the same weight or meaning across country pairs, and some interstate relations may receive more media attention than others (a systematic row error). The real informational value of our indicator lies in its **time-series variation**.

This example shows how well the ISCRI tracks the conflictual relations between China and the United States

****

The adequacy between quantitative and qualitative interpretations is classified into 4 categories:  **Very Consistent / Consistent / Surprising / Very Surprising.**

**2015**: **very consistent**. The indicator shows a peak of tension in the summer of 2015 which stems from Washington's questioning in May of Beijing's occupation of islands in the South China Sea.

**. January [1.57]**: **consistent**

**. February [1.56]**: **consistent**

**. March [1.59]**: **consistent**

**. April [1.62]**: **consistent**

**. May [1,65]**: **very consistent** May 30, 2015, Secretary of Defense Ashton Carter asks China to cease its claims to island territories in the China Sea.

**. June [1,7]**: **very consistent**

**. July [1.78]**: **very consistent**

**. August [1.76]**: **very consistent**

**. September [1.67]**: **consistent**

**. October [1.61]**: **consistent**

**. November [1.55]**: **consistent**

**. December [1.54]**: **consistent**

**2016: consistent**. A year without almost any hiccups, marked by good relations between Presidents Obama and Xi. The election in November of Donald Trump changes the game with the "Trump-Taiwan Call"

**. January [1.51]**: **consistent**

**. February [1.51]**: **consistent**

**. March [1.54]**: **consistent**

**. April [1.61]**: **consistent**

**. May [1.71]**: **consistent**

**. June [1.72]**: **Consistent**

**. July [1.72]**: **consistent**

**. August [1.65]**: **consistent**

**. September [1.62]**: **consistent**

**. October [1.70]**: **consistent**

**. November [1.66]**: **consistent** Election of Donald Trump

**. December [1,72]**: **very consistent** Phone call between Donald Trump and Mrs. Tsai, President of Taiwan. The indicator shows a small peak, without much consequence. It must be said that China itself minimized the incident by blaming it on a ruse by Taiwan rather than a mistake by the elected president.

**2017**: **very consistent.** The indicator clearly shows that the incident of December 2017 caused tensions quickly eased by Trump himself.

**. January [1.82]**: **very consistent** Start of Donald Trump's term

**. February [1.86]**: **very consistent** Trump reassures China, indicating that he still adheres to the one-China doctrine. The incident is deflated

**. March [1.81]**: **very consistent**

**. April [1,74]**: **consistent** Trump welcomes President Xi to Mar-a-Lago in Florida. The two countries declare the meeting fruitful

**. May [1.77]**: **consistent**

**. June [1.70]**: **consistent**

**. July [1.71]**: **consistent**

**. August [1.79]**: **consistent**

**. September [1.73]**: **consistent**

**. October [1.71]**: **consistent**

**. November [1.63]**: **consistent**

**. December [1.65]**: **consistent**

**2018: very consistent**. 2018 marks the rise in bilateral tensions, driven by a US policy clearly hostile to China and marked in particular by the launch of a trade war. The indicator takes off until the end of 2020 and the election of Joe Biden to the White House. In this sense, its reliability is remarkable.

**. January [1,7]**: **consistent**

**. February [1.75]**: **consistent**

**. Mars [1.76]**: **very consistent**. Beginning of the "Trade Wars" launched by Donald Trump against China. Increase in US tariffs by $50 billion on goods imported from China. Interesting to note: the "Material risk" curve marks an inflection that becomes continuous from this period.

**. April [1.77]**: **consistent**

**. May [1,8]**: **consistent**

**. June [1.84]**: **consistent**

**. July [1.87]**: **very consistent**. Escalation of the Trade Wars with new US tariffs that give rise to retaliatory tariffs from Beijing.

**. August [1.97]**: **consistent**

**. September [2.05]**: **consistent**

**. October [2,11]**: **very consistent**. Mike Pence's speech on a hardening of the position. Pence accuses China of intellectual property theft and interference in U.S. domestic politics. China counters that these claims are baseless.

**. November [2.06]**: **consistent**

**. December [2.09]**: **consistent**. Canada arrests on behalf of the United States Meng Wanzhou, Vice Chairman of the Board and Chief Financial Officer of Huawei.

**2019: very consistent**

**. January [2.09]**: **consistent**

**. February [2.02]**: **consistent**

**. March [2.06]**: **consistent**. China is suing the US, which is banning the use of Huawei equipment for US companies.

**. April [2.05]**: **consistent**

**. May [2.12]**: **consistent** New U.S. tariff increases. Further escalation.

**. June [2.15]**: **consistent**

**. July [2.26]**: **consistent**

**. August [2.34]**: **very consistent**. The United States declares that China manipulates its currency ("currency manipulator"). A first since 1994. Beijing counters that this accusation could destabilize financial markets.

**. September [2,3]**: **consistent**

**. October [2.28]**: **consistent**

**. November [2,29]**: **consistent**. Trump signs law officially supporting Hong Kong protesters against Beijing. Beijing imposes sanctions on US groups operating in China.

**. December [2.33]**: **consistent**

**2020: very consistent**. The indicator captures the rise in tensions and its peak in the summer of 2020 with Mike Pompeo's speech.

**. January [2.28]**: **very consistent**

* + Signing of the Phase 1 agreement in a climate of tension that remains high since a few days before the signing, Washington reiterated its accusations against China manipulating its currency ("currency manipulator"). It should be noted that despite the agreement maintains high tariffs on many Chinese imports.
  + Beginning of the Covid-19 epidemic. The United States prohibits non-U.S. citizens from entering its territory if they have recently visited China. The United States criticizes Beijing, Trump keeps talking about the "Chinese virus" while a Chinese official says it was Americans who brought the virus to China.

**. February [2.29]**: **consistent**

**. March [2.35]**: **very consistent** China expels 13 American journalists (New York Times, Wall Street Journal, and Washington Post) in response to the limitation of Chinese journalists in the US.

**. April [2,48]**: **very consistent** Beginning of Sino-American cooperation around the epidemic in a climate of suspicion. Trump says the WHO is biased and too pro-China.

**. May [2.66]**: **consistent**

**. June [2,7]**: **very consistent** Trump signs legislation to punish officials and companies that

infringe Hong Kong's freedoms and autonomy. China counters by saying that this is foreign interference in internal affairs. The US once again declares that China's claims in the South China Sea are unfounded.

**Very interesting: the "material risk" indicator marks a peak; It captures sanctions.**

**. July [2.86]**: **very consistent**

* + July 22: The United States orders China to close its consulate in Houston, alleging it is a site of espionage and theft of intellectual property. In response, China closed the consulate in Chengdu.
  + July 23: Secretary of State Mike Pompeo's speech titled "Communist China and the Future of the Free World," which marks a profound shift in U.S. policy. He said the era of engagement with the Chinese Communist Party is over, condemning its unfair trade practices, intellectual property theft, human rights abuses in Xinjiang and Hong Kong, and aggressive manoeuvres in the East China Sea and South China Sea. It calls on Chinese citizens and democracies around the world to pressure Beijing to change its behaviour and respect the rules-based international order.

**. August [3]**: **consistent**

**. September [3.1]**: **consistent**

**. October [3.1]**: **consistent**

**. November [3,1]**: **very consistent** In the context of the US presidential elections, the Trump administration is keeping up the pressure on China. John Ratcliffe, the director of intelligence, calls China "the greatest threat to the United States today." The Commerce Department blacklists a large number of Chinese companies. The State Department is tightening visa requirements for the CCP's 9 million members. The US prohibits investment by US companies in Chinese companies working with the People's Army of China. China is adopting retaliatory measures.

**. December [3.1]**: **consistent**

.**2021: very consistent**. Acme and gradual deflation of the sharpest tensions with the coming to power of Joe Biden.

**. January [3,2]**: **very consistent. This is the highest point of the indicator. It corresponds to the last days of the Trump administration when Mike Pompeo declared that Beijing was committing "genocide" and a "crime against humanity" against the Uighurs.**

**. February [3.08]**: **consistent**

**. March [2.99]**: **consistent** The Biden administration maintains the tariffs decided by the Trump administration and continues to blacklist US companies. Imports from Xinjiang are banned.

**. April [2.9]**: **consistent** In his State of the Union address, Joe Biden insists on the need for the US to continue the technological competition against China.

**. May [2.9]**: **consistent**

**. June [2,9]**: **consistent** NATO Summit where China is designated as a "Security Challenge". The statement sought by Washington states: "China's stated ambitions and assertive behaviour present systemic challenges to the rules-based international order and to areas relevant to alliance security".

**. July [2.99]**: **consistent**

**. August [2.99]**: **consistent**

**. September [2.9]**: **consistent**

**. October [2.8]**: **consistent**

**. November [2.75]**: **consistent** Slight decrease in the indicator.

* + 10 November: Signing of a climate collaboration agreement between Washington and Beijing on climate
  + November 15: First official meeting between Biden and Xi, by video. Discussion about safeguards to avoid conflict. Xi says the US is "playing with fire" over its support for Taiwan.

**. December [2.82]**: **consistent**

**2022: very consistent**. Biden's election defuses the Trump administration's most aggressive language elements, but the climate remains high, especially given the war in Ukraine.

**. January [2.73]**: **consistent**

**. February [2,7]**: **consistent** The US declares a "diplomatic boycott" regarding the Beijing Winter Olympics. Beginning of the invasion of Ukraine by Russia. China refuses to condemn the invasion. Chinese media are spreading the news that the US is developing biological weapons in laboratories in Ukraine. Xi criticizes Western sanctions against Russia.

**. March [2.66]**: **consistent**

**. April [2,62]**: **consistent** speech by Anthony Blinken who designates China as the "most serious long-term challenge to the international order".

**. May [2.58]**: **consistent**

**. June [2.5]**: **consistent**

**. July [2.55]**: **consistent**

**. August [2.54]**: **consistent** Nancy Pelosi's Coherent Visit to Taiwan, Condemned by Beijing

**. September [2.52]**: **consistent**

**. October [2.48]**: **consistent** US restrictions on Chinese processors.

**. November [2,4]**: **consistent** First physical meeting between Biden and Xi in Indonesia; the two men seek to re-establish a relationship of dialogue. Biden says the US will play the game of competition with Beijing but that he is not seeking conflict.

**. December [2.43]**: **consistent**

**2023: consistent (until now)**

**. January [2.41]**: **consistent**

**. February [2.48]**: **consistent** The United States destroys a Chinese balloon it suspects is a spy balloon.

**. March [2.54]**: **consistent**

**. April [2.52]**: **consistent**