

SIPRI Arms Trade Data Analysis – Markov Chain Modeling

Jack Joy, Andrew Pavlak, Jeffrey Zhang, Mauricio Torres,
Sammy Aridi, Steve Ferenzi, and Weston Williams

1. Data Description

Describe the data clearly -- particularly any missing data that might impact your analysis -- and the provenance of your dataset. Who collected the data and why? (10/100 pts)

Provenance & intent: The Stockholm International Peace Research Institute (SIPRI) is a widely used, public, research-oriented source. The primary dataset is the SIPRI Arms Transfers database (<https://www.sipri.org/databases/armstransfers>) which documents deliveries of major conventional arms since 1950. SIPRI created the database for analysts, researchers, policymakers, and the media to better understand the arms trade and how it has changed over time. It uses a volume metric called the Trend-Indicator Value (TIV) which measures military capability/volume, not USD, to make weapons comparable across types.

Relevant variables: recipient country, supplier country, delivery year, item name/weapon description, number/quantity, TIV for delivered weapons

Data issues: We did not find any missing data in the dataset that would affect our specific research question. However, various columns and rows were not useful. For example, the dataset initially contained three empty columns that appeared to be placeholder fields, likely created when the data was exported or read into pandas. These columns had no meaningful values or headers and consisted mainly of missing or placeholder entries (“?” or NaN), so we removed them to ensure a cleaner and more efficient dataset for analysis. The same thing occurred with two rows, which we dropped. For nuanced analysis of this dataset, we identified some issues that would require deeper cleaning and formatting:

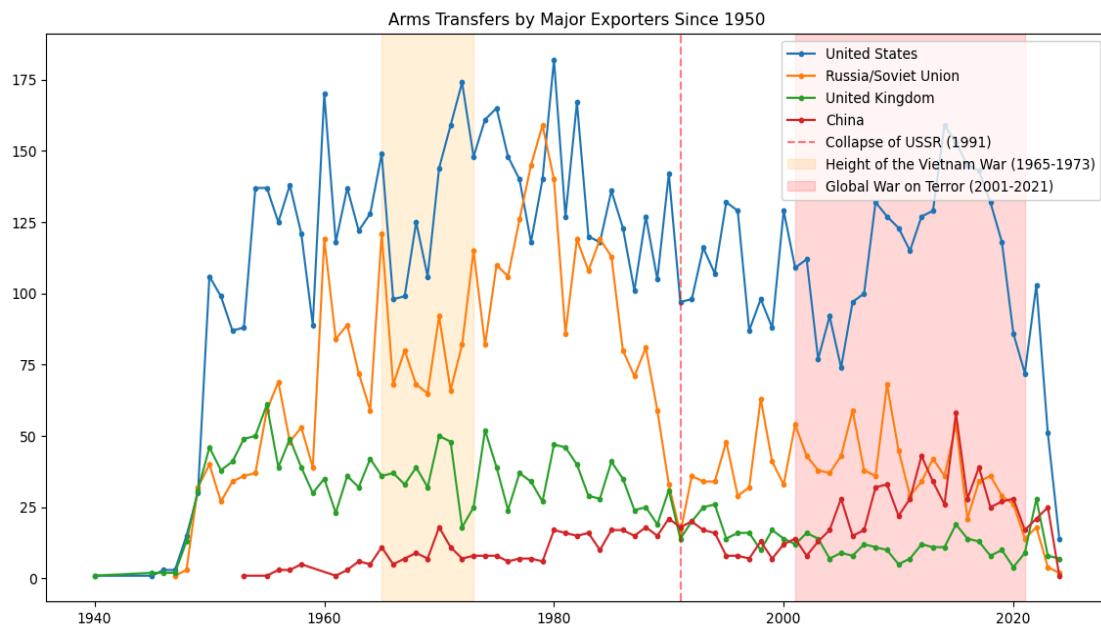
- Country name variants (e.g., “United States”, “United States of America”, “U.S.A.”, “PRC”, “People’s Republic of China”) require normalization before filtering by country.
- Inconsistent weapons descriptions require mapping to convert raw descriptions into broad categories like Aircraft, Ships, Missiles, Armored vehicles, Sensors, Artillery, Parts, Other.

Exploratory Data Analysis (EDA)

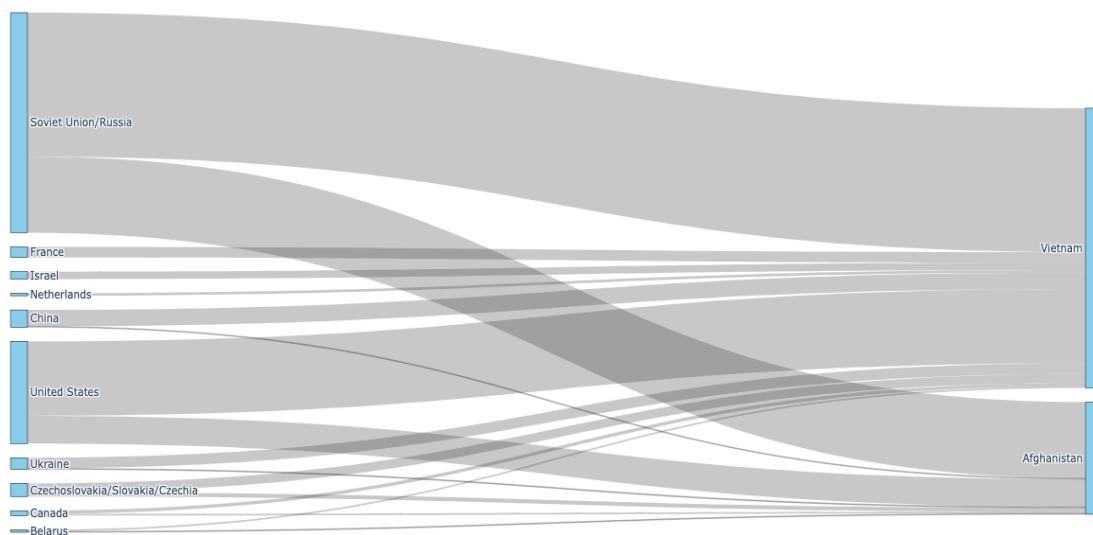
- Framing. Our study examines global arms transfers patterns, focusing on how weapons flow between states as instruments of geopolitical influence. These transfers are not isolated commercial transactions but components of broader strategic relationships shaped by politics, security concerns, and economic interests. Arms flows often reveal or reinforce shifts in international power structures—whether during the Cold War, the post-9/11 era, or emerging “Cold War 2.0” strategic rivalries over maritime chokepoints and regional security. To contextualize these dynamics, our EDA first maps

supplier–recipient relationships to reveal how arms transfers align with alliances and conflicts across time.

- **Key Observations.** The initial step of the EDA involved identifying which nations dominate global exports and tracing where those weapons went. Countries such as the United States, the Soviet Union/Russia, and Western European powers have historically accounted for the majority of global arms transfers, with the U.S. alone comprising roughly 30% of total exports. The data also highlight how recipients correspond to major conflict zones—Egypt during the Arab-Israeli wars, Iraq during the conflicts of the 2000s, and Ukraine in the post-2014 and 2022 crises. These spatial and temporal patterns suggest that arms transfers are both a reflection and a predictor of geopolitical tension.



Top 10 Exporters to Afghanistan and Vietnam (All Years)



- Temporal Context and Geopolitical Dynamics. Clear trends emerge when aligning transfer data with major historical events: arms flows surged during the height of the Cold War (1960s–1970s), declined after the fall of the USSR, and rose again during the Global War on Terror after 2001. U.S. exports reveal a wide distribution across 177 recipients, indicating a global strategy of broad-based military engagement rather than concentration in a few allies. In contrast, USSR and later Russian transfers clustered around specific client states and ideological allies such as East Germany and Syria. These temporal and structural contrasts provide a foundation for modeling of arms-trade behavior using nonparametric techniques such as Markov chains or kernel density estimation.

2. Phenomenon to Model

What phenomenon are you modeling? Provide a brief background on the topic, including definitions and details that are relevant to your analysis. Clearly describe its main features, and support those claims with data where appropriate. (10/100 pts)

Research question: How does U.S. involvement in foreign wars impact the imports and distribution of weapons and arms trade before, during, and after conflict? Specifically, how might patterns in model-based predictions derived from the Vietnam War (1970s) be used to analyze similarities and differences in the War on Terror in Afghanistan (2001-2021)?

Key assumption: Historical delivery patterns (1950–latest year) are informative about near-term futures. We can use past arms-transfer data to forecast the next few years if no exogenous geopolitical shocks occur (war, sanctions, regime change, major diplomatic realignments) that would drastically change who buys or sells weapons.

Strategic Context: SIPRI-relevant U.S. military arms support to South Vietnam grew from modest weapons and material replacement supplies pre-1955, into a very large and platform-diverse conventional arms relationship during the Vietnam War period of 1955–1975 (air, armor, artillery, naval craft, helicopters, ammunition, logistics). Post-1975 transfers essentially ceased until diplomatic normalization decades later. The scale and character of transfers were tightly coupled to Cold War geopolitics and shifting U.S. domestic politics after the Soviet Union collapsed in 1991.

(1) Pre-War (prior to 1955)

What was supplied: the United States began subsidizing and equipping French and associated Vietnamese forces in the late 1940s and early 1950s with major conventional equipment and replacement material (small arms and ammunition, transport and utility equipment, replacement ordnance for French inventories, etc.)

Political context: in the immediate post-WWII Cold War environment the U.S. prioritized stemming communist expansion in Asia. Washington's objective was to sustain anti-communist French and associated Vietnamese forces without large direct combat commitments (initially via grants, sales, and equipment transfers routed through the U.S. Military Assistance Advisory Group (MAAG) and French authorities). The result was an

early, programmatic flow of conventional materiel intended to keep anti-communist forces viable while avoiding large U.S. combat deployments.

(2) Vietnam War period (1955–1975)

What was supplied: this period represents the largest, most diversified U.S. conventional arms transfers to the Republic of Vietnam (ARVN, RVN Navy, RVN Air Force).

Deliveries escalated after the early 1960s as the US moved from advisory to large-scale materiel support. Major categories include:

- Aircraft and aviation: transport and combat aircraft, a rapid build-up of helicopters (e.g., UH-1 Hueys for mobility and medevac), and fixed-wing support/attack platforms
- Armored vehicles and tanks: M48 Patton tanks and a range of APCs and tracked vehicles
- Artillery and air-defense equipment: towed/howitzers, fire-control and anti-air weapons systems.
- Naval equipment: patrol boats, coastal craft and small naval platforms to build coastal interdiction and riverine capability.
- Large quantities of ammunition, spare parts, logistics vehicles, and radar/communications

Political context: domestic U.S. politics and Cold War dynamics both increased escalation in Vietnam. The domino-theory held sway among policymakers (Eisenhower, Kennedy, Johnson), leading to progressively larger aid packages and, eventually, direct military intervention. U.S. policy moved from “training and equipping” (MAAG era) to full-scale material sustainment for an expanded ARVN and extensive air/sea operations against communist North Vietnam. As a result, conventional-weapons transfers and replenishment (aircraft, helicopters, tanks, naval craft, artillery) formed the backbone of recorded U.S. support. Public opinion in the United States shifted sharply during the 1960s, as the growing anti-war movement and mounting U.S. casualties led to decreased domestic support for continued escalation. President Johnson, facing both the war’s unpopularity and pressure to prioritize his Great Society social programs, began de-escalating U.S. involvement. After years of “Vietnamization” under President Richard Nixon (a policy aimed at shifting combat responsibility to South Vietnamese forces), Saigon fell to North Vietnam in April 1975.

(3) Post-war (1975 to current)

After the fall of Saigon, U.S. military assistance to the unified communist Vietnam dropped to zero for many years because of the post-war break in bilateral relations and U.S. embargoes on military and most economic assistance. SIPRI-style records therefore show near-absence of U.S. major conventional arms transfers to Vietnam during the Cold War’s final decades. Normalization of relations in the 1990s led to gradual non-military engagement and later limited security cooperation in the 2000s–2010s, but major conventional arms transfers from the U.S. to Vietnam remain extremely limited compared with the 1955–75 build-up.

Main Features

- Variables
 1. Recipient: the country that receives arms as an importer. These are the countries that we are analyzing.
 2. Year of order: the year that the transaction occurred. We are using this to define different periods of war and non-war.
 3. SIPRI TIV for total order: a numerical value created by SIPRI that quantifies the volume of international transfers of major conventional weapons from one country to another.
 4. Support: a categorical variable that we created from SIPRI TIV for total order by dividing them into 3 different buckets: *Low*, *Medium*, and *High*. These buckets define the discrete states of the Markov Chain.
 - Low: Bottom 40% of SIPRI TIV for total order values.
 - Medium: 40% - 80% of SIPRI TIV for total order values.
 - High: Top 20% of SIPRI TIV for total order values.
- Transition Matrix

	Low	Medium	High
Low	$P(L \text{ to } L)$	$P(L \text{ to } M)$	$P(L \text{ to } H)$
Medium	$P(M \text{ to } L)$	$P(M \text{ to } M)$	$P(M \text{ to } H)$
High	$P(H \text{ to } L)$	$P(H \text{ to } M)$	$P(H \text{ to } H)$

- Time Periods: Pre-War, During War, and Post-War. Each time period generates its own matrix allowing us to compare how the counties' imports evolve temporally.

3. Non-Parametric Model Description

Describe your non-parametric model (empirical cumulative distribution functions, kernel density function, local constant least squares regression, Markov transition models). How are you fitting your model to the phenomenon to get realistic properties of the data? What challenges did you have to overcome? (15/100 pts)

Arms transfers do not usually happen on a regular schedule. They tend to come in unpredictable waves, with long stretches of no activity followed by a few major deals. These ups and downs are driven by a mix of politics, wars, shifting alliances, and economic changes that do not follow any clear-cut pattern. This complexity makes nonparametric methods like Markov chains, kernel density estimation, and bootstrapping work well. They let us pick up on real trends across different categories and time periods without trying to squeeze the data into a rigid formula.

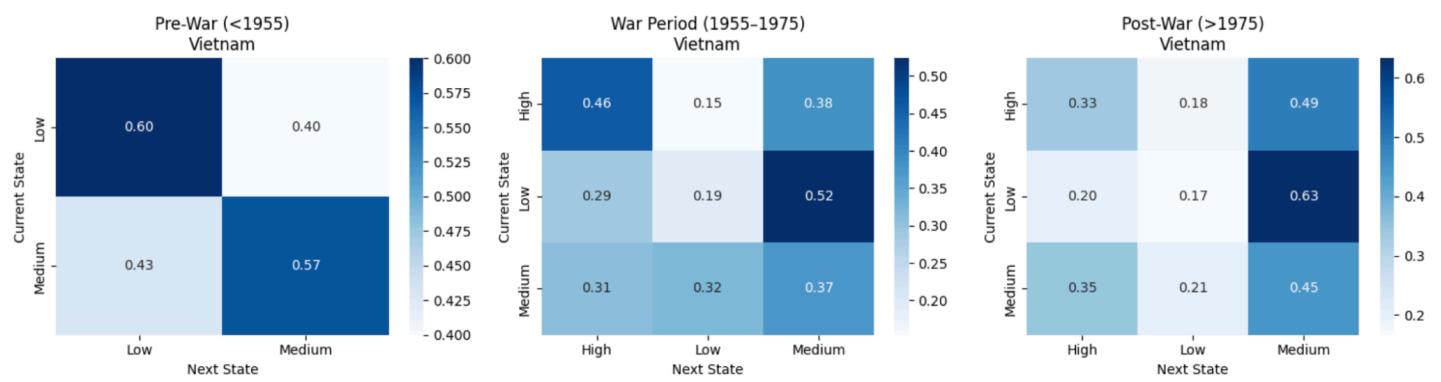
We chose a Markov chain because it effectively models transitions between states over time, allowing us to capture how arms trade activity evolves from one period to the next. Unlike other nonparametric models, it emphasizes the probabilistic relationships between sequential events rather than static outcomes. This made it ideal for analyzing changes in arms trade behavior before, during, and after conflict periods.

We encountered quite a few challenges while building the model, such as dealing with missing or incomplete data, figuring out what was causing certain patterns, and making sense of some of the model's more unpredictable results. One of the biggest issues was the imbalance in arms trade volumes. The U.S., for example, had way more transactions, so its trade patterns were easier to track and model over time. On the other hand, countries like Vietnam had very little data, which introduced more uncertainty and variance.

Our original plan was to study arms trades before, during, and after major wars. This worked out okay for the Vietnam War since both the U.S. and Vietnam had enough data for all three phases. But when we switched to looking at the Afghanistan War, Afghanistan did not have any trade data after 2021 when the Taliban overthrew the government. As such, the Markov chain results were not as expected, and we had to pivot and only focus on the pre-war and wartime periods.

Additionally, there was a discussion amongst our group about which Vietnam naming conventions should be aggregated and included in the model. There were four names that related to Vietnam: Viet Nam, South Vietnam, Viet Minh, and Viet Cong. We realized that our investigation did not want to investigate the imports of Viet Minh and Viet Cong as they were the enemy and receiving arms from China and Russia which creates an inaccurate model.

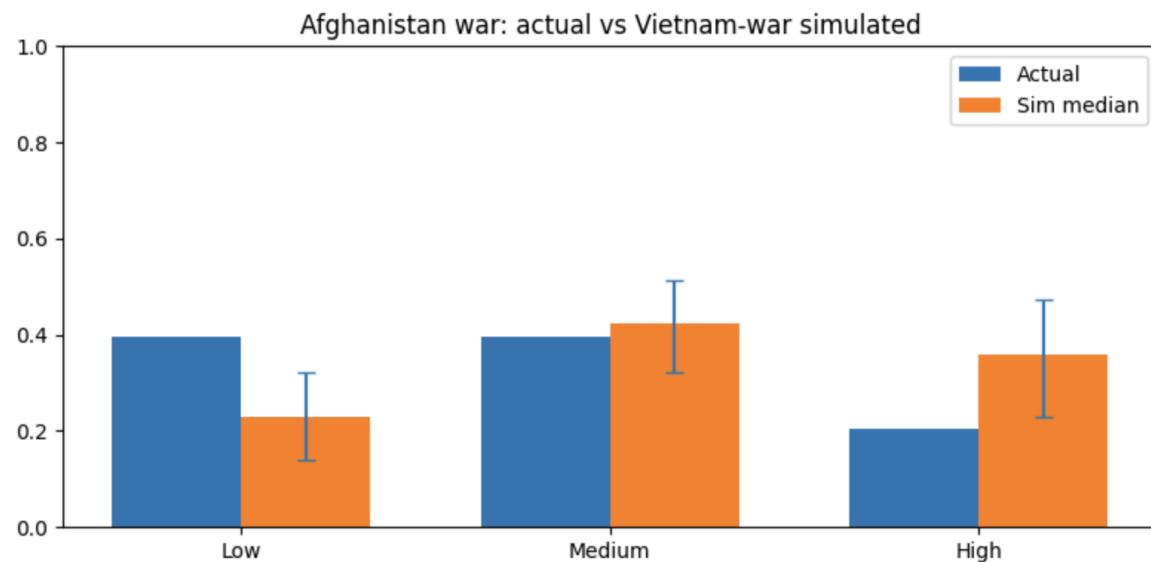
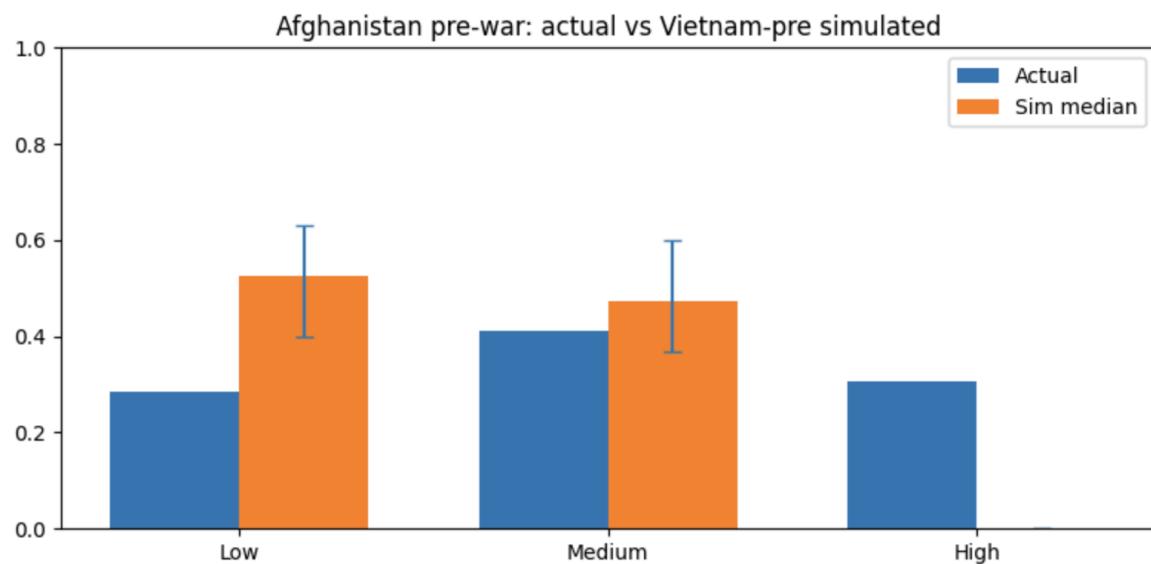
The Markov chain results show that Vietnam's arms trade activity was stable before the war, with most transitions remaining within low or medium levels. During the war period (1955–1975), trade patterns became much more volatile, reflecting rapid shifts in arms inflows as military aid and supply sources fluctuated. After 1975, Vietnam's trade activity stabilized again but at much lower levels, indicating reduced international arms exchanges in the post-war period. Overall, the model effectively captures how conflict intensity influenced the stability and magnitude of Vietnam's arms trade over time.



4. New Sequence Creation

Either use your model to create new sequences (if the model is more generative) or bootstrap a quantity of interest (if the model is more inferential). (15/100 pts)

Using the matrices derived from Vietnam, we simulated Afghanistan's trade patterns for both its pre-war (before 2001) and war (2001–2021) periods, generating thousands of synthetic sequences to approximate its state distribution. We then compared the simulated proportions of states to Afghanistan's actual data and evaluated model performance using confidence intervals and chi-square distance metrics. This approach allowed us to test whether Vietnam's historical transition behavior could credibly explain Afghanistan's arms trade dynamics.



5. Critical Evaluation

Critically evaluate your work in part 4. Do your sequences have the properties of the training data, and if not, why not? Are your estimates credible and reliable, or is there substantial uncertainty in your results? (15/100 pts)

The simulated sequences for Afghanistan only partially share the properties of the Vietnam training data. When using the Vietnam pre-war transition matrix, the generated Afghanistan pre-war sequences fail to reproduce the frequency of high-intensity import years because the Vietnam pre-war data itself contained almost no transitions into the “High” state, leading to underestimation of major import events. In contrast, sequences generated from the Vietnam war-period matrix more closely resemble Afghanistan's wartime dynamics, as both exhibit persistent medium-to-high import states and similar state proportions that almost fall within the simulated 95% confidence intervals. While the war-period estimates are credible, with a small chi-square distance indicating a good fit, there remains substantial uncertainty due to small sample sizes, wide confidence bands, and differences in geopolitical context between the two countries. Overall, the model captures the general wartime behavior of sustained arms inflows but lacks reliability for pre-war conditions, as the Vietnam data do not adequately represent Afghanistan's earlier arms trade patterns.

6. Conclusion

Write a conclusion that explains the limitations of your analysis and potential for future work on this topic. (10/100 pts)

Limitations and Caveats

1. **Markov assumption is not completely realistic.** Arms-provision depends on multi-year contracting processes, logistical availability, and political dynamics. A Markov chain assumes that the next state of a system depends only on its current state, not on the sequence of events that came before. As such, a first-order MC ignores these longer lags and would require higher-order Markov chains to more realistically model the complexity.
2. **Political and temporal incomparability.** Vietnam's Cold War context differs structurally from the post-2001 Afghanistan context. The geopolitical calculus of strategic competition between the US and USSR shaped supplier choices and volumes in ways that do not map neatly onto the multipolar 2001–2021 environment (NATO/coalition donors, different supplier mix). As such, transition probabilities learned from South Vietnam are not definitive.
3. **SIPRI coverage limits:** SIPRI focuses on major conventional arms and omits a large share of small arms/light weapons, covert transfers, security-assistance contract spending, logistics, and development support—all critical elements of comprehensive assistance to South Vietnam and Afghanistan. Relying only on SIPRI misses those channels, so combining SIPRI with other datasets (e.g., USAID and Foreign Military Financing

(FMF) records, congressional reports, etc.) would more accurately capture the full support picture.

4. **SIPRI naming conventions:** Inconsistent country naming within the SIPRI database introduces uncertainty into data accuracy and aggregation. For example, the dataset lists “Viet Nam,” “South Vietnam,” and “Viet Cong (South Vietnam)” as separate entities, which complicates efforts to capture a unified record of U.S. transfers to South Vietnam. These variations suggest possible discrepancies in how the original data were categorized or recorded, leading to potential duplication or omission during analysis.
5. **Limited sample size.** The South Vietnam era is one historical case, so the derived transition matrix risks overfitting and poor out-of-sample performance. Mitigation would require pooling training data across multiple comparable cases (U.S. support to other partner states in Cold War/post-Cold War settings).

Future Potential

Arms transfer data, such as that provided by the SIPRI database, enables non-parametric modeling of broader patterns of inter-state competition and conflict. Such analyses hold significant value for both policymakers and academic researchers. Three promising future applications include:

1. **Conflict Intensity Visualization** — Mapping historical global arms transfer patterns to identify potential future hotspots where military buildups may escalate into low-intensity conflict or open warfare.
2. **Conflict Stabilization Dynamics** — Modeling the trajectory of military activity in states currently engaged in, or at risk of entering, stabilization or counterinsurgency operations.
3. **Strategic Competition at Maritime & Land Chokepoints** — Predicting how states may use arms transfers to influence nations controlling key geographic chokepoints such as major canals, straits, and transit corridors.