Measurements of distance in international trade of goods

Distance is one of the fundamental geographical concepts, alongside space and region. It can be defined in numerous ways, depending on the context. The most commonly distinguished types are physical distance, which is the empirically measured line between locations; time distance, which measures how long it takes to travel a given section; and economic distance, understood as the cost of transport in space. This work focuses on physical distance, as other types require separate studies. Distance is a key element of various models and economic analyses, including those related to trade. One of the most commonly used models is the gravity model of trade. Its popularity stems from highly successful empirical analyses, despite its simplicity.

The gravity model of trade assumes that trade exchange between two objects can be described in a manner analogous to gravitational attraction, i.e., by the size of both objects and the distance between them. The first relationship is positive, while the second one is a negative dependency. The gravity model of trade is applied in many different analyses, such as measuring the potential volume of exchange between countries, estimating the effects of free trade agreements or sanctions on trade.

Unfortunately, there is a lack of uniformity in the measurements of distance used in these types of analyses. Therefore, it is important to standardize this issue and understand the differences between various distance measures and their impact on the analyses.

The earliest studies in this field used the physical distance between points, either along a straight line, a road, or a railway track. However, determining the distance between countries (spatial objects) becomes more complicated. It is assumed that the distances between countries can be defined as the distance between their major centers of population and economic activity. The two most popular measurements are the distance between national capitals or the largest cities. Among the papers read by author on this topic, more than half (53%) used the first method, while more than one-fifth (22%) used the second method. In most countries, the capital is the most important city; however, there are exceptions in 28 countries (India, Pakistan, Vietnam, China, Kazakhstan, Brazil, Nigeria, New Zealand, Bolivia, Tanzania, Turkey, Canada, the United States, Myanmar, Australia, Ecuador, Ivory Coast, Cameroon, the United Arab Emirates, Switzerland, Morocco, Belize, South Africa, Israel, Belgium, Benin, Sudan, and the Philippines). A third simple method (but less popular) is measuring the distance between the centroids of a country's land area. Another approach involves calculating the distance between major sea ports, though this is limited to maritime countries. These few methods are easy to understand and calculate, so their popularity is not surprising. The questions are: what are the differences in estimations when using these various methods, and are there better methods that use more complex measurements?

Distance can be calculated in many other ways, based on average distances between multiple cities, centers of gravity, or equidistance. In countries with multiple centers,

such as the United States, it is reasonable to calculate the distance to the nearest one, rather than always using Washington, D.C. For example, in the case of trade with China or Australia, the most important city seems to be Los Angeles, rather than New York. A similar situation occurs in other large countries like Canada, Brazil, Russia, India, China, and Australia. Going one step further, in the next two methods, the distance is calculated between the closest cities with populations above one million and one hundred thousand people, respectively.

Another group of measurements is based on central points of human activity in a country. It is difficult to precisely calculate the center of population or economic activity for every country worldwide due to the lack of data. A compromise may be to calculate the gravitational center of major cities in a country, i.e., the point that is, on average, the least distant from all cities. An improvement to this method is to take into account the population of these cities as a weight. Using these cartographic methods, it is possible to calculate the location of the new center points, between which the distance may be measured.

An additional field for measuring distance is provided by remote sensing methods, which allow obtaining data for the entire world from satellites, without concerns about data quality and repeatability. The main approach is to identify places with human infrastructure and, based on this, calculate the center of human buildings. Alternatively, there is great potential in using data on light emissions at night because they are closely correlated with actual economic activity and can be more easily identified.

This work develops a gravity model of trade using various distance measures, based on a literature review and cartographic or remote sensing methods. In summary, there are twelve different methods for measuring distance between countries:

1. Distance between capitals
2. Distance between the largest cities
3. Distance between geographical centroids
4. Distance between major ports
5. Distance between the largest cities with special treatment for large multicenter countries
6. Distance between the nearest cities with populations over one million
7. Distance between the nearest cities with populations over one hundred thousand
8. Distance between the gravitational centers of a group of cities
9. Distance between the gravitational centers of a group of cities with populations
10. Distance between the centers of human buildings based on remote sensing data
11. Distance between the centers of nighttime light emissions based on remote sensing data
12. Complex distance measurement from CEPII's GeoDist database

Then, using the information criteria (Akaike's AIC and Schwarz's Bayesian BIC), it is determined which of them gives the best estimates. The search is for minimal values. The gravity model of trade has a form: ln trAB = ln GDPA + ln popA + ln GDPB + ln popB + ln disAB i + seaAB + borAB + hisAB + cons; where trAB means trade exchange between countries A and B, GDPA means Gross Domestic Product of country A, popA means the total population of country A, disAB i means i measurement of distance between countries A and B, seaA is dummy variable meaning access to sea of country A, borAB means that there is physical border between countries A and B, hisAB means that in the past countries A and B was one country, and cons means constant. Between 12 models only data about distance are changing. As an estimator is used Poisson Pseudo Maximum Likelihood (PPML). Data are collected for 182 countries and for time period of 25 years (between 1998 and 2023). Sources of the date are UN comtrade database, World Development Indicators and CEPII's GeoDist database.

In conclusion, more complex cartographic methods provide statistically better results. The best method was the one that calculated distance based on the gravitational centers of a group of cities, taking their population into account (model 9). The second-best methods were those measuring the distance between the nearest cities (models 6 and 7), which were very similar. However, the worst measurements came from remote sensing data (models 10 and 11), possibly due to errors in pixel classification.

Another important finding is that there were no major differences between the most popular measures based on capitals and the largest cities (models 1 and 2), and their results were only slightly worse than those of the top models.

After ranking the models from best to worst, the final order looks as follows: 9, 7, 6, 12, 8, 5, 2, 1, 4, 3, 11, 10. There are small gaps between the models from 9 to 1, and a large gap between models 1 and 4, and then between models 3 and 11. The comparison of the methods according to their complexity and actual usefulness will be part of future projects. To sum up, there are differences in estimations using different measurements of distance between countries. The more advanced methods provide statistically better results.