From networks to flows: Using flow maps to understand mobility patterns in cattle trade

Keywords: Origin-Destination Data, Flow Maps, Vector Fields, Cattle Epidemic Modelling, Cattle Trade Networks

Extended Abstract

The study of mobility patterns is important in a variety of research areas [1, 2], and it has a direct impact on policymaking in urban environments, analyses of disease transmission, and epidemic control [3, 4]. There are many approaches grounded in Network Science to model cattle trade in various regions [5, 6]. While the power of network approaches to model disease spread is undeniable, they suffer from shortcomings. Very few approaches have used large cattle trade datasets such as the ones available in Brazil or the USA [7] where the level of uncertainty about the trade is quite high [8]. For example, Brazil is the second-largest producer of beef in the world, with an annual production of 10.3 million metric tons [9]. Yet, Brazil is known to have a less-than-perfect tracking system, leading to uncertainty in the cattle trade network [8]. Network-based approaches often ignore relevant spatial information; when cattle herds are transported between locations, areas in-between are affected. In fact, they might have higher spatial centrality because ground transportation is used in most cattle movements. This information can be easily lost in networks because they only capture origin-destination pairs.

Hence, we propose to convert the network of movements into a vector-field representation, as it complements the benefits of network analyses because it can capture dynamic mechanisms that would be hard to do using networks. It is a general approach that could be applied to other types of mobility dataset (e.g. human mobility, bird migrations). The first step is to set divisions of the region being investigated–a granularity for the study–and project the network in this space. We can look at all the outgoing edges for each cell and combine them into a resulting vector using a relevant method; this resulting vector represents the general flow direction for that cell. We then use an interpolation method to complete the field [10]. There are so-called critical points that need to be analysed. It is our opinion that *sinks* and *sources* play a major role in epidemiological research as destinations and starting points of infection, respectively. An analysis of sink-source dynamics in mobility trades is a valuable approach that is captured in moving from network methods to flow methods. In order to understand the global behaviour of the trade patterns, it is essential to analyse the dynamics of the critical points on the vector fields of networks of different months.

Furthermore, to look at the proposed methodology in a real scenario, we use the cattle mobility dataset from the state of Minas Gerais in Brazil. Minas Gerais produces more cattle heads than many countries around the world. Vector fields are generated using the monthly trades for a spatial subdivision (micro-regions). Figure 1-A is a toy example of the process of how trades can be used to produce an output vector for each micro-region.

A cosine similarity analysis is then conducted during the year of study (2013 in the case here). The cosine similarity values are shown in Figure 1B-1 for all micro-regions. Different colours in Figure 1B-2 indicate different clusters based on the cosine similarity. Looking at Figure 1B-1 we can see that the vector direction for the micro-regions remains nearly unchanged over a year, indicating a predictable direction for cattle trading from those regions. The cosine similarity measure shows how a micro-region's vector behaviour has changed overtime; being in a range from steadiness or dynamic. As well as comparing the behaviour of one spatial area with others, which can lead to the division of larger spatial areas into predictable and unpredictable regions. Figure 1C shows sinks and sources, for four separate months. One can clearly see that the sinks and sources change for different periods, which indicates that high dynamics of cattle trades in Minas Gerais is present. When comparing Figure 1B-2 with Figure 1C, most of the areas of the largest cluster of micro-regions are not critical points in the vector field. Comparatively, the areas with high sink (red areas in Figure 1C-1) and source (dark blue areas in Figire1C-2) density correspond to those that belong to the cluster of parts whose behaviour is not steady. In other words, there are some risky parts in the system that act as sources (or sinks), and we cannot estimate where the epidemic will spread based on their direction of flow. Even though we cannot predict the exact direction in which diseases spread, we can restrict trade by knowing where the source points of flow are for each period of time.

References

1. Gonzalez MC, Hidalgo CA, and Barabasi AL. Understanding individual human mobility patterns. nature 2008; 453:779–82
2. Barbosa H, Barthelemy M, Ghoshal G, James CR, Lenormand M, Louail T, Menezes R, Ramasco JJ, Simini F, and Tomasini M. Human mobility: Models and applications. Physics Reports 2018; 734:1–74
3. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, and Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the National Academy of Sciences 2009; 106:21484–9
4. Arenas A, Cota W, Gómez-Gardeñes J, Gómez S, Granell C, Matamalas JT, Soriano-Paños D, and Steinegger B. Modeling the spatiotemporal epidemic spreading of COVID-19 and the impact of mobility and social distancing interventions. Physical Review X 2020; 10:041055
5. Bajardi P, Barrat A, Natale F, Savini L, and Colizza V. Dynamical patterns of cattle trade movements. PloS one 2011; 6:e19869
6. Motta P, Porphyre T, Handel I, Hamman SM, Ngu Ngwa V, Tanya V, Morgan K, Christley R, and Bronsvoort BM. Implications of the cattle trade network in Cameroon for regional disease prevention and control. Scientific reports 2017; 7:1–13
7. Farokhnejad S, Cardoso D, Rocha C, Mata AS da, and Menezes R. A Data-driven Approach to Cattle Epidemic Modelling under Uncertainty. *CompleNet*. to appear. Springer Nature, 2022
8. The Laundering of Cattle (In Portuguese). `https://piaui.folha.uol.com.br/materia/lavagem-da-boiada/`. (last accessed: 21-11-2022). 2022
9. The Biggest Producers of Beef in the world. `https://agroninja.com/countries-where-beef-production-is-defining/`. (last accessed: 21-11-2022). 2022
10. Farokhnejad S, Oliveira M, Ribeiro E, and Menezes R. Using Vector Fields in the Modelling of Movements as Flows: A Case Study with Cattle Trade Networks. *CompleNet*. to appear. Springer Nature, 2023

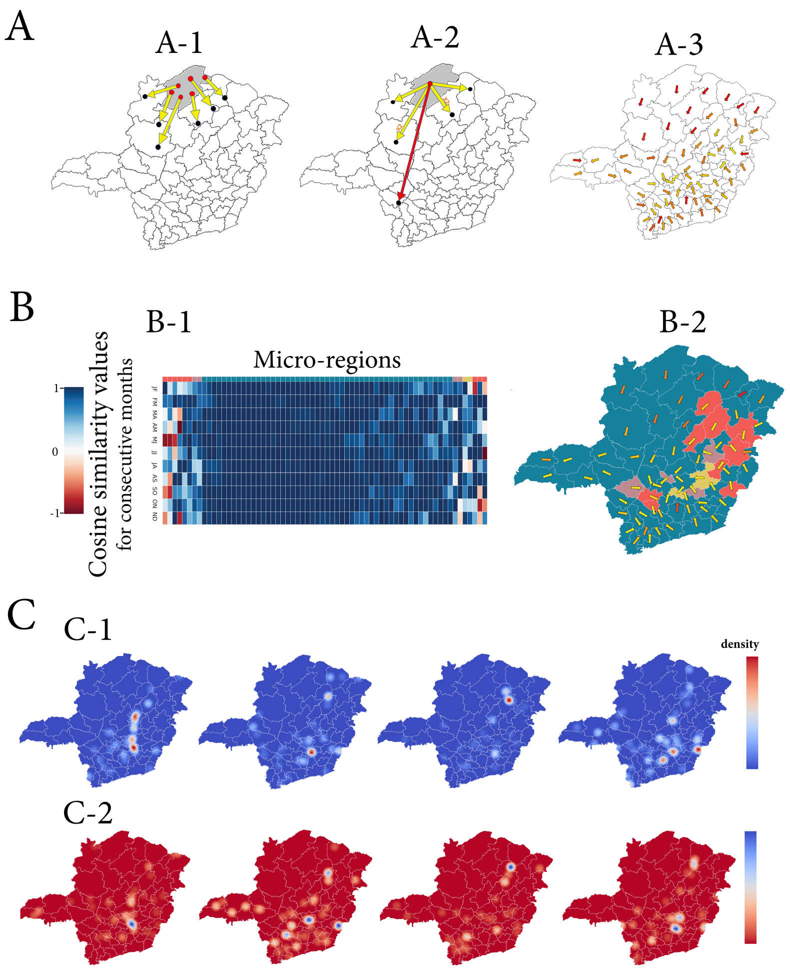

Figure 1: **(A) An illustration of how to generate a vector field from a cattle mobility network. (A-1)** Take the grey part as an example and concentrate on the trades happening between it and other parts. **(A-2)** The trades are captured with a vector starting from the centre of the location (in grey) and ending at the centre of the destinations (yellow vectors). All vectors are then aggregated using a relevant method. Here, we add the vectors into a final vector (red vector), which we assign to the grey location. **(A-3)** Visualisation of what we call the flow map. For a better visualisation, all vectors are displayed with the same length. The colour of each vector corresponds to a range that shows the size of the vector. **(B) Cosine similarity between vectors of micro-regions for eleven consecutive months of 2013. (B-1)** As an example, two flow maps of two consecutive months Jan-Feb are used to calculate the similarity of the angle between two vectors associated with each city. The result of the cosine similarity calculation for each city is a set of values ranging from -1 to 1. Using k-means clustering, micro-regions are grouped according to their similarity values (a feature vector). Each group is illustrated with the same colour in **(B-2)**. The largest cluster of micro-regions **(B-2)** contains the group that does not exhibit a significant change in direction of the vectors.**(C) Critical points in the cattle trade vector field.** Sinks **(C-1)** and sources **(C-2)** are shown in vector field for four months. Initial vector fields are generated based on trades happening within specific months. A triangle-based interpolation method is used to create a vector field from the initial vectors. We found critical points in this vector field and identified sinks and sources. In **(C-1)**, red areas indicate locations with a high density of sinks in the vector fields. Dark blue shows the areas with zero density of sink points. In **(C-2)**, the emphasis is on sources. Dark blue areas indicate high-density of sources, and dark red areas are the ones without source points.