



OPEN

DATA DESCRIPTOR

# A dataset of a proxy variable for the poultry trade flows in China

Lingcai Kong<sup>1,2</sup>, Xuan Zhou<sup>1,2</sup>, Mengwei Duan<sup>1,2</sup>, Xinyi Yang<sup>1,2</sup>, Caifeng Zhou<sup>1,2</sup> & Yilan Liao<sup>3</sup>✉

Understanding the intercity poultry trading network is crucial for assessing the risk of avian influenza prevalence. Unfortunately, the poultry trading network in China has rarely been described. Here, with a modified radiation model, we obtain values for a proxy variable for poultry trade flows among 318 prefecture-level cities in China in 2015 utilizing the product capacity and demand quantity of poultry of the cities. The results are validated by comparing the proxy variable values with the trade volumes investigated in the literature, and it is found that the modified radiation model can accurately predict the main poultry trade flows among cities. This is the first dataset on China's poultry trade pattern, and it can be used to analyze the production and consumption structure of poultry in prefecture-level cities within China. The dataset can be a tool for avian influenza epidemic risk assessment as well as a basis to develop prevention and control measures during an epidemic.

## Background & Summary

The poultry trade plays a key role in the food supply, as poultry is one of the main sources of animal protein, including milk, meat and eggs. However, it is also believed to play an extremely important role in the spread of avian influenza, particularly over long distances and across different regions<sup>1–3</sup>. Therefore, understanding the trade patterns is of great help to develop avian influenza control strategies. Unfortunately, poultry trade flows in China have rarely been reported, except for a few local surveys<sup>2,4,5</sup>.

To estimate interregional trade flows, survey, non-survey, and hybrid methods have been widely used<sup>6</sup>. The survey-based approach can provide precise results but is not always available due to the large workload and high costs. Non-survey approaches, such as gravity models, entropy and information theory models, neural network models, and behavior-based models, provide alternative practical methods in the estimation of trade flows, despite some loss in accuracy<sup>7</sup>. Among the non-survey-based methods, the most commonly used is the gravity models, emerging from mainstream modeling frameworks in economics, for trade pattern analysis and empirical research in economic studies<sup>8,9</sup>. The gravity model belongs to the so-called spatial interaction models, which predict and explain the movement of population, traffic, commodities and information between origins and destinations in geographical space<sup>10</sup>. Despite its widespread use, to estimate bilateral trades with gravity model, previous trade flow data is necessary to fit parameters in the model formula<sup>11</sup>. The radiation model was one alternative model with no empirical data required. It was proposed to predict the population mobility and migration patterns, requiring only the population data in each location<sup>11</sup>. Due to its parameter-free nature, the radiation model can be applied in cases lacking previous flow data.

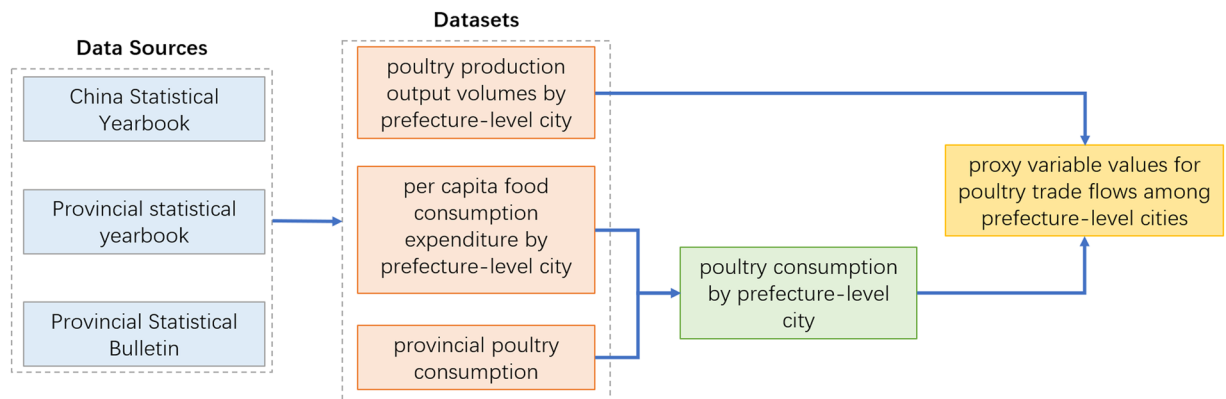
To fill the dataset gap in the field of poultry flows in mainland China, we adapted the radiation model, which was originally used to predict human mobility, and considered the production and consumption in each location the input to predict the interregional poultry trade flows. The whole data process framework is shown in Fig. 1. The results were validated with local survey data in the literature.

## Methods

**Data sources.** To predict poultry trade flows, poultry production output volumes and consumption were needed. First, we used the slaughtered poultry volume for each prefecture-level city, collected from the statistical yearbook or bulletin (2016) of each province, autonomous region and municipality, as the proxy variable of the poultry production output volumes. Then, for poultry consumption, the per capita consumption of poultry by household at the province level was collected from the China Statistical Yearbook 2016<sup>12</sup>. To obtain the poultry

<sup>1</sup>North China Electric Power University, Department of Mathematics and Physics, Baoding, 071003, China. <sup>2</sup>Hebei Key Laboratory of Physics and Energy Technology, North China Electric Power University, Baoding, 071003, China.

<sup>3</sup>Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, State Key Laboratory of Resources and Environmental Information System, Beijing, 100101, China. ✉e-mail: [liaoYL@reis.ac.cn](mailto:liaoYL@reis.ac.cn)



**Fig. 1** The whole data processing framework for poultry trade flow dataset production.

consumption of each prefecture-level city, a downscaling method was applied, implemented by multiplying the provincial per capita poultry consumption of households by a weight, to reflect the differences between cities. The weight for city  $i$  was set as the ratio of the per capita food consumption expenditure of the city to the mean of the values in the province to which city  $i$  belongs. Finally, by multiplying the population size, the total poultry consumption of each city was obtained. The above indicators of cities were collected from the statistical yearbook or bulletin (2016) of each province.

**Original radiation model.** The original radiation model was first presented by Simini *et al.* to predict population mobility and migration patterns<sup>11</sup>, which were assumed determined by a job selection process that consists two steps: (1) job seeking, assuming that the number of employment opportunities offered in each location is in proportion to the resident population there; (2) job selecting, individuals choose the closest job that offers benefits higher than the next-best offer available in the home location. Thus, distance has priority over benefits, as individuals are willing to accept more distant jobs only if there is a high enough benefit<sup>13</sup>.

Simini *et al.* also provided an equivalent description for the radiation model, using an analogy with the radiation emission and absorption processes in physical sciences<sup>11</sup>. The probability that  $T_{ij}$  particles emitted from location  $i$  (with population  $m_i$ ) are absorbed by another location  $j$  (with population  $n_j$ ) is a binomial distribution with average

$$\langle T_{ij} \rangle = T_i p_{ij} = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})} \quad (1)$$

and variance  $T_i p_{ij}(1 - p_{ij})$ , where  $T_i = \sum_{j \neq i} T_{ij}$  is the total particles emitted by location  $i$ ,

$$p_{ij} = \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})}$$

is the probability that a particle emitted from location  $i$  with population  $m_i$  is absorbed by another location  $j$  with population  $n_j$ , given that  $s_{ij}$  is the total population in all locations (except  $i$  and  $j$ ) within a circle of radius  $r_{ij}$  (the distance between  $i$  and  $j$ ) centered at  $i$ .

**Modified radiation model.** Here, as another analogy, we consider suppliers of poultry products in search of consumers as job-searchers, they choose closer consumers unless more distant consumers are willing to pay a sufficiently high price. In other words, we regard poultry been consumed as they obtained ‘work opportunities’. In this way, we can predict the poultry trade flows by modifying the original radiation model. In detail, we consider two locations  $i$  and  $j$  with production  $m_i$ ,  $m_j$  and consumption  $n_i$ ,  $n_j$ , respectively, at distance  $r_{ij}$  from each other and denote with  $s_{ij}$  the total consumption in the circle of radius  $r_{ij}$  centered at  $i$  (excluding  $i$  and  $j$ ). The average flow  $T_{ij}$  from  $i$  to  $j$  can be predicted by the radiation model Eq.(1) with  $T_i = \sum_{j \neq i} T_{ij} = k_i m_i$ , the total poultry production output to other regions from location  $i$ , where  $k_i$  is the proportion of exports to production in location  $i$ . In this study, we assume a constant  $k = 1$  for all cities. In fact, any positive value for it provides proportional predictions of the proxy variable for poultry trade flows.

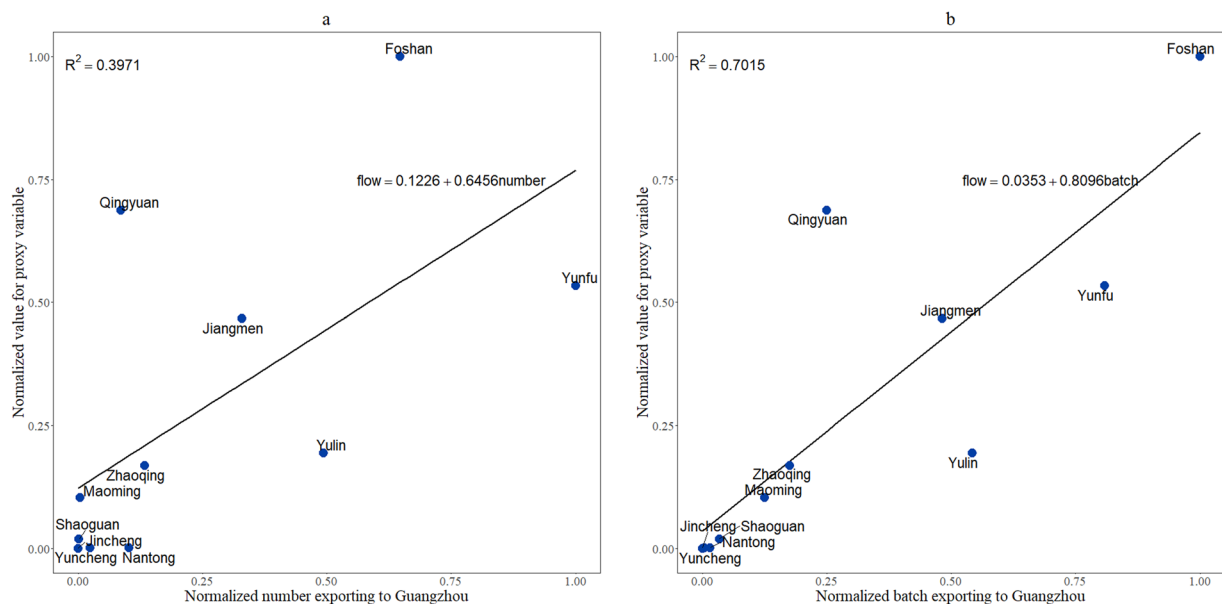
### Data Records

We produced two data products: the poultry production output and consumption data for each prefecture-level city in 2015 and the values for the proxy variable estimated for poultry trade flows among them. The dataset can be found in figshare<sup>14</sup>.

The data are stored in comma-separated values (CSV) format. The dataset on production and consumption contains the following fields:

**city** – The city included in the calculation.

**city\_code** – The administrative area codes of the city.



**Fig. 2** The regression results of the proxy variable for poultry trade flows with investigated number and batches from main sources exporting to Guangzhou.

**province** – The province to which the city belongs to.

**lon** – The longitude of the city.

**lat** – The latitude of the city.

**pop** – The population in the city in 2015.

**consumption** – The estimated total poultry consumption (value) in the city in 2015.

**product** – The estimated total poultry production (value) in the city in 2015.

The modified radiation model predicts the values for the proxy variable for poultry trade flows among each pair of cities. The dataset on these values contains the following fields:

**Origin** – The origin city of the flow.

**O.city\_code** – The administrative area code of the origin city.

**O.Prov** – The province the origin city belongs to.

**Destination** – The destination city of the flow.

**D.city\_code** – The administrative area code of the destination city.

**D.Prov** – The province the destination city belongs to.

**Flow** – The predicted values for the proxy variable for the poultry trade volume with the presented model from the origin city to the destination city.

**Prop** – The contribution rate of the proxy variable for poultry trade volume from the origin city to the total imports for the destination city.

**CumlProp** – The cumulative contribution rate of the proxy variable for poultry trade volume from the origin city for the destination city, calculated from largest to smallest for each origin city.

## Technical Validation

Few investigations of the poultry trade have been publicly released. Even so, we validated the estimation using investigated flows or statistical results published in available studies.

We first compare the origin cities exporting chicken to Guangzhou City with the investigation undertaken by Lin Hong *et al.*<sup>4</sup>. According to the investigation, chickens were the dominant avian species of live bird trade with a wide source, including 122 prefecture-level cities of 16 provinces of China. In addition to Guangzhou city itself, Yunfu, Foshan, Jiangmen, Zhaoqing, Qingyuan, Maoming and Shaoguan from Guangdong Province; Yulin from Guangxi Zhuang Autonomous Region; Nantong from Jiangsu Province; and Yuncheng and Jincheng from Shanxi Province were found to be the main sources exporting chickens to Guangzhou<sup>4</sup>. Obviously, most of them are located near Guangzhou, except Nantong, Yuncheng and Jincheng. We did regression analyses of the proxy variable for poultry trade flows with publicly available data on number and batches of chicken exporting to Guangzhou. Considering the different scales, we first normalized them to the range [0,1] with the formula

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}}.$$

Origin	Values for proxy variable for trade flows predicted	Proportion to total columns to Guangzhou	Cumulative proportion to total columns to Guangzhou	Reported as sources by Lin <i>et al.</i>
Foshan	114268.7	0.223071	0.223071	TRUE
Qingyuan	78539.65	0.153322	0.376393	TRUE
Yunfu	60975.59	0.119034	0.495428	TRUE
Jiangmen	53338.82	0.104126	0.599554	TRUE
Nanping	40481.61	0.079027	0.678581	FALSE
Yulin-Jx	22127.38	0.043196	0.721777	TRUE
Huizhou	20305.16	0.039639	0.761416	TRUE
Zhaoqing	19249.77	0.037579	0.798995	TRUE
Maoming	11754.45	0.022947	0.821941	TRUE
Guilin	8503.421	0.0166	0.838541	TRUE
Heyuan	7992.58	0.015603	0.854144	TRUE
Dongguan	7894.495	0.015411	0.869555	TRUE
Chongqing	6889.456	0.013449	0.883005	FALSE
Meizhou	5871.328	0.011462	0.894467	TRUE
Qinzhou	5600.758	0.010934	0.9054	TRUE
Hengyang	5173.651	0.0101	0.9155	TRUE
Ganzhou	3856.188	0.007528	0.923028	TRUE
Shanwei	3852.58	0.007521	0.930549	TRUE
Jieyang	2875.017	0.005613	0.936161	TRUE
Longyan	2462.79	0.004808	0.940969	TRUE
Zhanjiang	2311.45	0.004512	0.945481	TRUE
Shaoguan	2133.053	0.004164	0.949645	TRUE
Nanning	1839.922	0.003592	0.953237	TRUE
Shantou	1584.008	0.003092	0.95633	FALSE
Ji-an	1215.081	0.002372	0.958702	TRUE
Changde	1157.966	0.002261	0.960962	TRUE
Chenzhou	961.7035	0.001877	0.96284	FALSE
Guigang	867.1939	0.001693	0.964532	TRUE
Changchun	858.2947	0.001676	0.966208	FALSE
Zhongshan	702.2188	0.001371	0.967579	TRUE

**Table 1.** The top 30 sources exporting poultry to Guangzhou predicted by the presented modified radiation model.

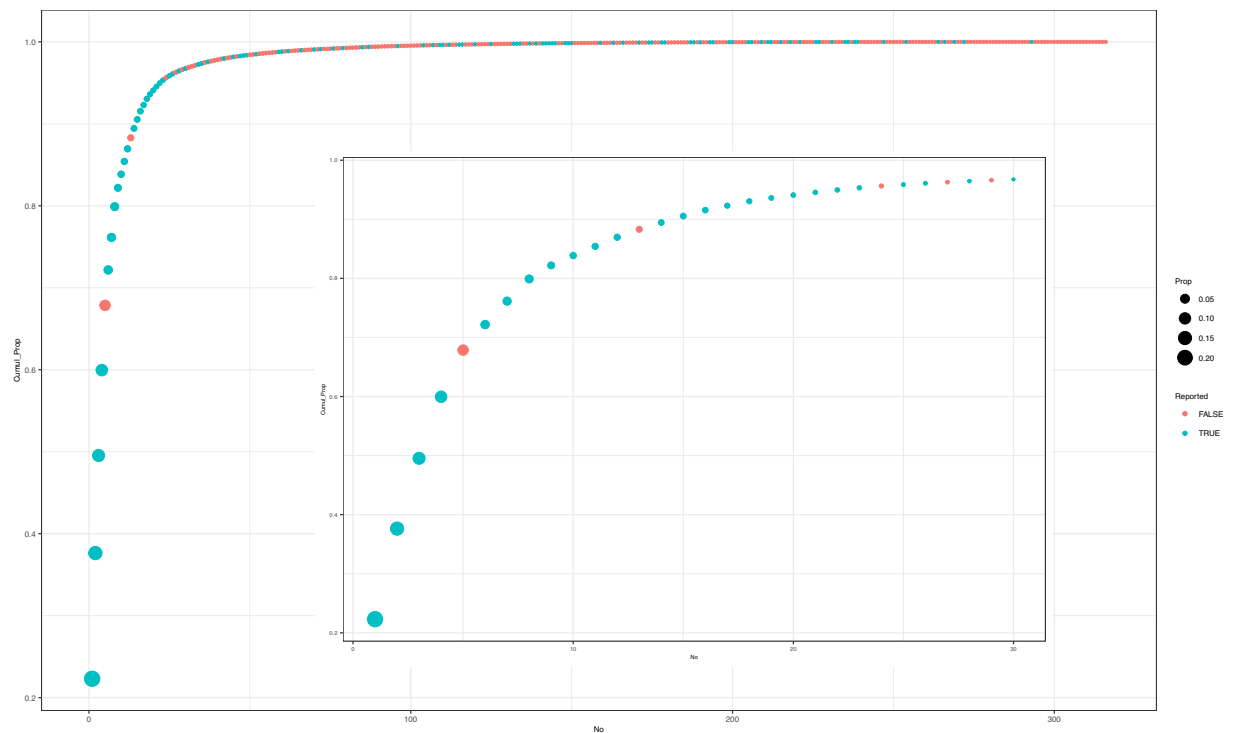
The regression coefficients for the predicted proxy variable to the number and batches were 0.6456 and 0.8096, and with p-values of 0.0377 and 0.0013, respectively (Fig. 2). These results show that the actual flow data were positively associated with the proxy variable.

Now we compare the predicted sources of chicken exporting to Guangzhou with that had been reported<sup>4</sup>. First, we sorted the origin cities in descending order by the values for the proxy variable for poultry trade flows predicted with the presented radiation model. Then, the contribution rates and cumulative contribution rates for the total proxy variable value exported to Guangzhou were calculated. Table 1 shows the top 30 sources exporting poultry to Guangzhou according to our prediction. All the main chicken sources near Guangzhou were accurately predicted as the main sources exporting poultry to Guangzhou, except for Shaoguan, which was not included in our estimation because of missing data. Figure 3 shows the cumulative proportion of the values for the proxy variable for poultry trade flows exported to Guangzhou. Among the top 30 sources, 25 (83.3%) have been reported as chicken sources<sup>4</sup>. Therefore, the presented modified radiation model predicted the main sources exporting poultry to Guangzhou.

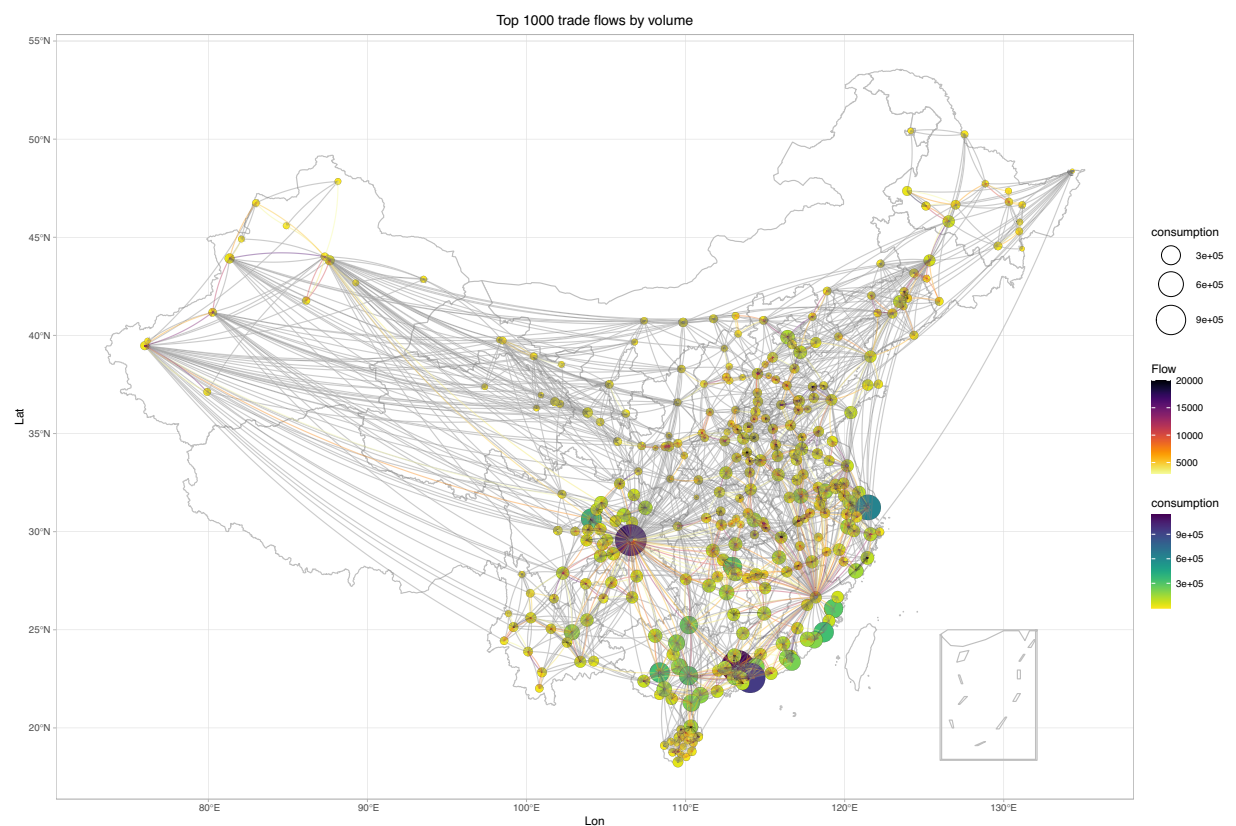
Since very few investigations on the poultry trade flows in 2015 are available, we compare some statistical results in the years around 2015 with our estimation. First, we compare our estimation with the movements of live broilers out of and within Guangxi in 2017 analyzed by Tang *et al.*<sup>5</sup>. According to their analysis of the movements out of Guangxi, Guangdong Province received the most movements (77.5%), followed by Guizhou (9.9%) and Hunan (8.6%). In our estimation, the corresponding proportions are 68.58%, 1.31% and 13.15%, respectively. That is, two of the three main exporting provinces from Guangxi were successfully predicted.

Another study that investigated the live poultry trading network around Changsha city, Hunan Province, was also used for comparison with our estimation<sup>2</sup>. According to the study, of the live poultry exported to Changsha, approximately 64% were from cities in Hunan Province, and the remaining 36% were from Hubei, Henan, Guangxi, Jiangxi, and Guangdong Provinces<sup>2</sup>. For our estimation, approximately 71% were from Hunan, and the other main sources included Jiangxi, Fujian, Chongqing, Hubei, Guangxi and Guangdong Provinces. Except for Fujian and Chongqing, the other provinces were reported to be the main source provinces. For the provinces importing live poultry from Changsha, Hunan Province itself accounted for approximately 79%, while





**Fig. 3** The cumulative rate of exportation to Guangzhou ordered by the volume predicted by the radiation model. The blue points represent cities that have been reported as exporting chicken sources to Guangzhou, while the opposite is true for the red points. The size of the points was determined by the predicted volume.



**Fig. 4** Poultry trade flows contributed 90% of the exports to each prefecture-level city predicted by the presented radiation model.

the remaining were exported to Hubei, Guangdong, Guangxi, Henan, Yunnan, Jiangxi, Zhejiang and Shanghai, although the last three provinces (city) accounted for only a small percentage<sup>2</sup>. In our estimation, in addition to Hunan Province, where Changsha is located, the aforementioned provinces were also the main destinations of poultry exported from Changsha, with ranks of 3, 4, 5, 8, 16, 2, 12 and 14, respectively, by the proxy variable values in descending order.

These comparisons show that the values calculated by the presented modified radiation model provide an accurate proxy for the main poultry flows among cities. What we need to do is to find a balance between keeping more flows and ensuring accuracy. Since a few sources contributed most of the poultry trade volumes<sup>5</sup>, we can use the contribution rates as threshold values to obtain the main sources for each city. Figure 4 shows the trade flows that cumulatively contributed to more than 90% of the total volumes for each city. These values can help analyze the production and consumption structure of poultry within China and can be used to assess avian influenza epidemic risk, which is the basis for developing prevention and control measures during an epidemic.

## Code availability

The codes are fully operational under R 4.1.3. The package `stplanr` v0.4.0<sup>15</sup> was referenced to implement the modified radiation; the visualization of the trade flows was implemented with the package `sf` v1.0-4<sup>16</sup>. The codes for the modified radiation model and the estimation are available on Github: <https://github.com/MathGISer/Poultry-Trade-Network>.

Received: 15 August 2022; Accepted: 20 October 2022;

Published online: 11 November 2022

## References

- Li, R. *et al.* Live Poultry Trading Drives China's H7N9 Viral Evolution and Geographical Network Propagation. *Frontiers in Public Health* **6**, <https://doi.org/10.3389/fpubh.2018.00210> (2018).
- Zhang, R. *et al.* Live poultry feeding and trading network and the transmission of avian influenza A(H5N6) virus in a large city in China, 2014–2015. *Int J Infect Dis* **108**, 72–80, <https://doi.org/10.1016/j.ijid.2021.05.022> (2021).
- Zhou, X. *et al.* The role of live poultry movement and live bird market biosecurity in the epidemiology of influenza A (H7N9): A cross-sectional observational study in four eastern China provinces. *Journal of Infection* **71**, 470–479, <https://doi.org/10.1016/j.jinf.2015.06.012> (2015).
- Lin Hong *et al.* Market Chain Analysis Based on Different Avian Species in the Wholesale Live Bird Markets in Guangzhou City (in Chinese). *China Animal Health Inspection* **34**, 8–13, <https://doi.org/10.3969/j.issn.1005-944X.2017.01.003> (2017).
- Tang, H. *et al.* Analysis of the movement of live broilers in Guangxi, China and implications for avian influenza control. *Transboundary and Emerging Diseases* **69**, e775–e787, <https://doi.org/10.1111/tbed.14351> (2021).
- Liu, W., Li, X., Liu, H., Tang, Z. & Guan, D. Estimating inter-regional trade flows in China: A sector-specific statistical model. *Journal of Geographical Sciences* **25**, 1247–1263, <https://doi.org/10.1007/s11442-015-1231-6> (2015).
- Sargento, A. L. M., Ramos, P. N. & Hewings, G. J. D. Inter-regional trade flow estimation through non-survey models: an empirical assessment. *Economic Systems Research* **24**, 173–193, <https://doi.org/10.1080/09535314.2011.574609> (2012).
- Anderson, J. E. in *Annual Review of Economics*, Vol 3 Vol. 3 *Annual Review of Economics* (eds K. J. Arrow & T. F. Bresnahan) 133–160 (2011).
- Head, K. & Mayer, T. in *Handbook of International Economics* Vol. 4 (eds G., Gopinath, E., Helpman, & K., Rogoff) 131–195 (Elsevier, 2014).
- Roy, J. R. & Thill, J.-C. in *Fifty Years of Regional Science* (eds Raymond J. G. M. Florax & David A. Plane) 339–361 (Springer Berlin Heidelberg 2004).
- Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96, <https://doi.org/10.1038/nature10856> (2012).
- National Bureau of Statistics of China. *China Statistical Yearbook - 2018*. (China Statistics Press 2018).
- Simini, F., Maritan, A. & Neda, Z. Human Mobility in a Continuum Approach. *PLOS ONE* **8**, e60069, <https://doi.org/10.1371/journal.pone.0060069> (2013).
- Kong, L. *et al.* A dataset of proxy variables for poultry trade flows in China, *figshare*, <https://doi.org/10.6084/m9.figshare.c.6146979.v1> (2022).
- Lovelace, R. & Ellison, R. `stplanr`: A Package for Transport Planning. *The R Journal* **10**, 7–23, <https://doi.org/10.32614/RJ-2018-053> (2018).
- Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* **10**, 439–446, <https://doi.org/10.32614/RJ-2018-009> (2018).

## Acknowledgements

The study was supported by the National Natural Science Foundation of China (Grant No: 42171419), the Fundamental Research Funds for the Central Universities (No. 2021MS117) and the Natural Science Funds of Hebei (Grant No. D2022502001). The funders had no role in study design, decision to publish, or preparation of the manuscript.

## Author contributions

Y.L.L. and L.C.K. designed and supervised the study. X.Z., M.W.D., X.Y.Y. and C.F.Z. participated in data collection. L.C.K. developed the model and analyzed the model outputs. L.C.K. and Y.L.L. prepared the first draft of the manuscript. All authors approved the final manuscript as submitted.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Y.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022