Faculty of Agricultural and Nutritional Science

Christian-Albrechts-University Kiel
Institute of Animal Breeding and Husbandry

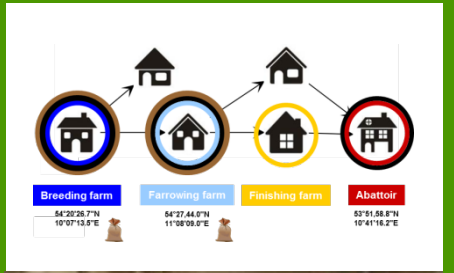
Impact of information load on the centrality parameters of a pig trade network in Northern Germany

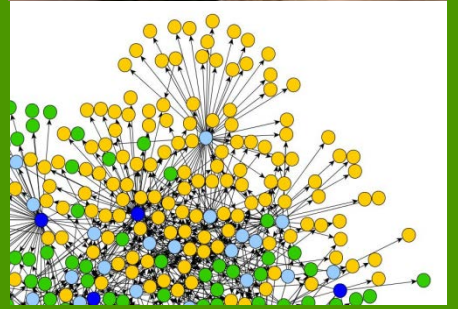
Kathrin Büttner and Joachim Krieter

Institute of Animal Breeding and Husbandry
Christian-Albrechts-University, Kiel, Germany

67th Annual EAAP Meeting Belfast, UK
August 29th to September 2nd, 2016

Session 64, Abstract number 23513, kbuettner@tierzucht.uni-kiel.de





Introduction

- **Useful application of network analysis for agricultural sciences**

  → **Behavioural research**

  – Social structure of animal groups (friendships, aggressions)

  – Abnormal behaviour (feather pecking, tail biting)

  → **Epidemiological studies**

  – Prediction of disease transmission

  – Implementation of appropriate control measures

Introduction

- **Useful application of network analysis for agricultural sciences**

  → **Behavioural research**

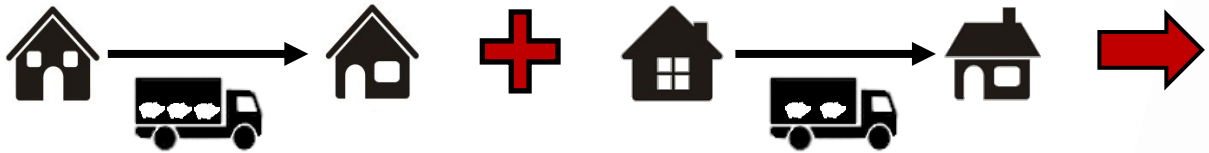
  – Social structure of animal groups (friendships, aggressions)

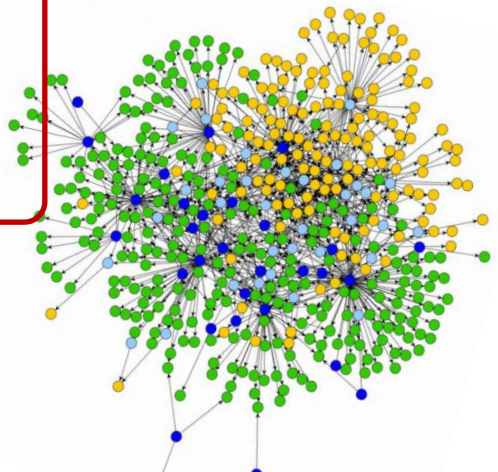
  – Abnormal behaviour (feather pecking, tail biting)

  → **Epidemiological studies**

  – Prediction of disease transmission

  – Implementation of appropriate control measures

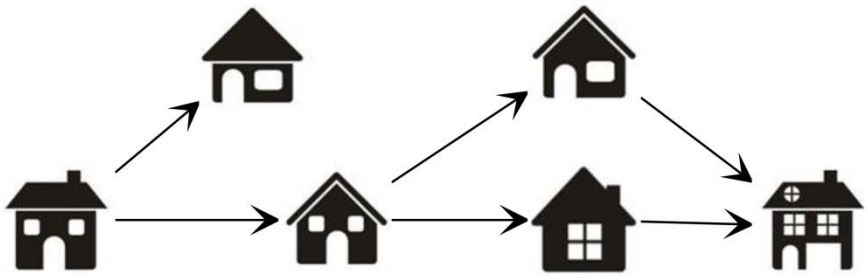




Introduction

- **Challenge of network analysis:**

  Incomplete data sets & various information loads

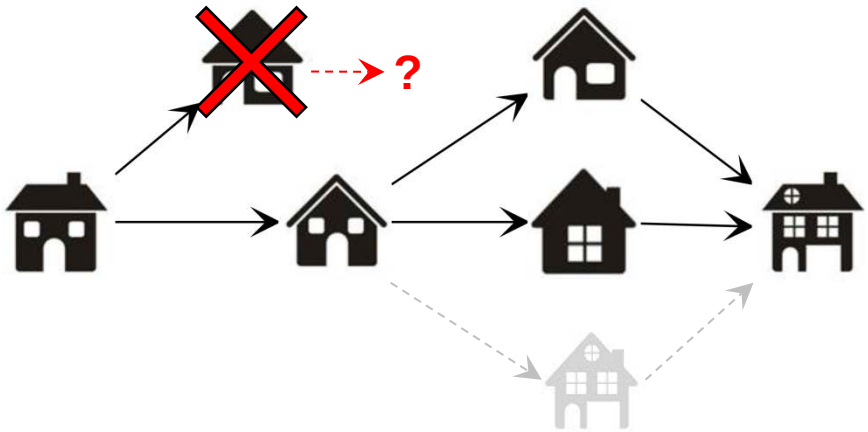


Introduction

- **Challenge of network analysis:**

  Incomplete data sets & various information loads

  → Missing or false positive nodes or edges



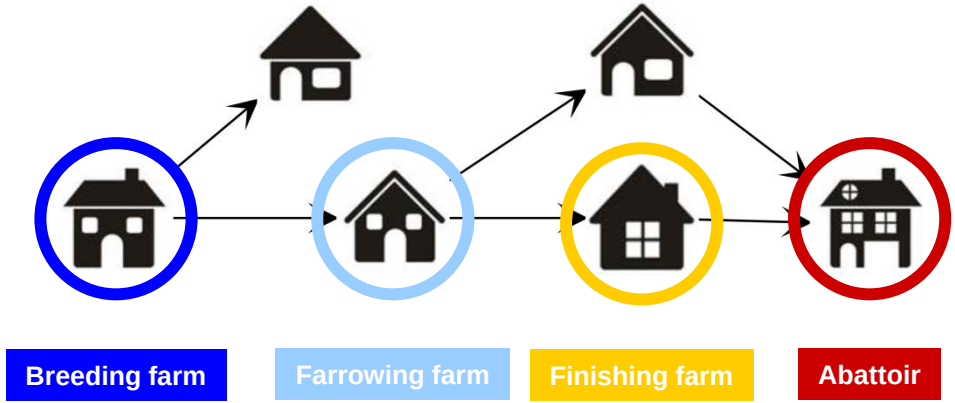
Introduction

- **Challenge of network analysis:**

  Incomplete data sets & various information loads

  → Missing or false positive nodes or edges

  → Missing information about some of their attributes



Farm type

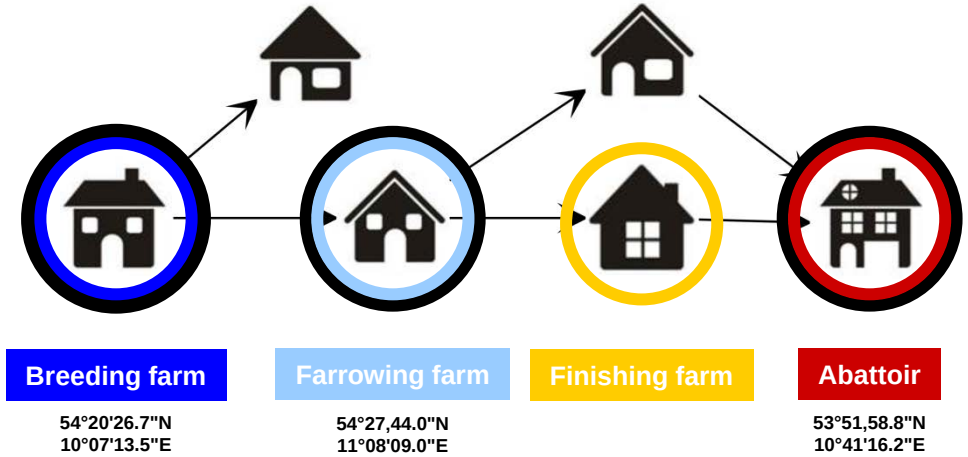
Introduction

- **Challenge of network analysis:**

  Incomplete data sets & various information loads

  → Missing or false positive nodes or edges

  → Missing information about some of their attributes



Location

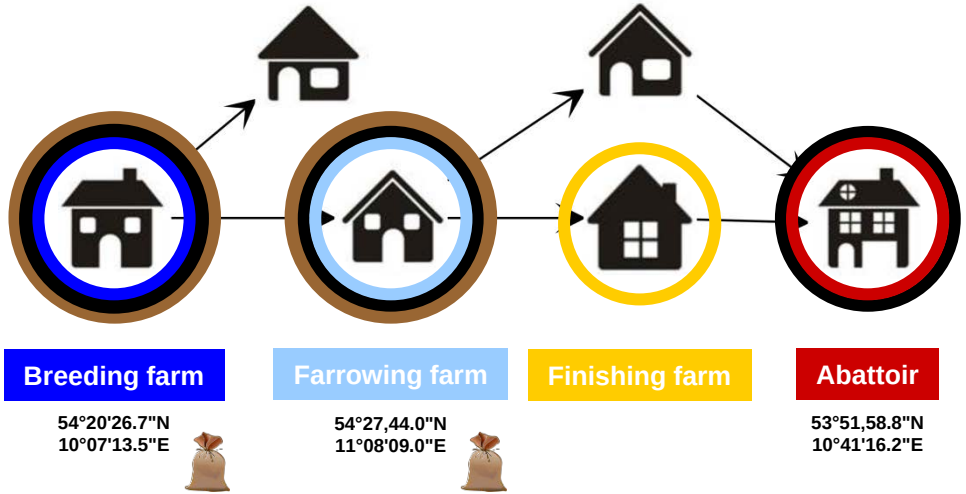
Introduction

- **Challenge of network analysis:**

  Incomplete data sets & various information loads

  → Missing or false positive nodes or edges

  → Missing information about some of their attributes



Feed supply

Aim of the study

1. **Evaluating the variation of the centrality parameters between different network versions based on various information loads**
2. **Assessing the network robustness, meaning the point at which incomplete data sets may influence the centrality parameters**

Materials & Methods

Three network versions

Network A

Contains all trade contacts with information about the supplier, the purchaser as well as the truck

978 nodes

2.280 edges

Network B

Only those trade contacts stayed in the data set with full geographical location

866 nodes

1.884 edges

Network C

Only trade contacts with additional information about the farms, e.g. farm type, stayed in the data set

188 nodes

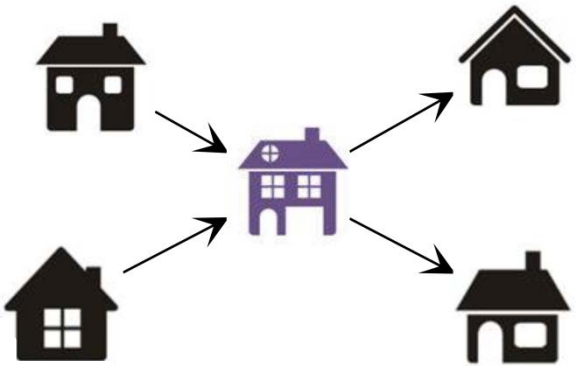
625 edges

Materials & Methods

Centrality parameters - “What characterizes a central or important farm?”

→ **In-degree:** Number of direct ingoing trade contacts

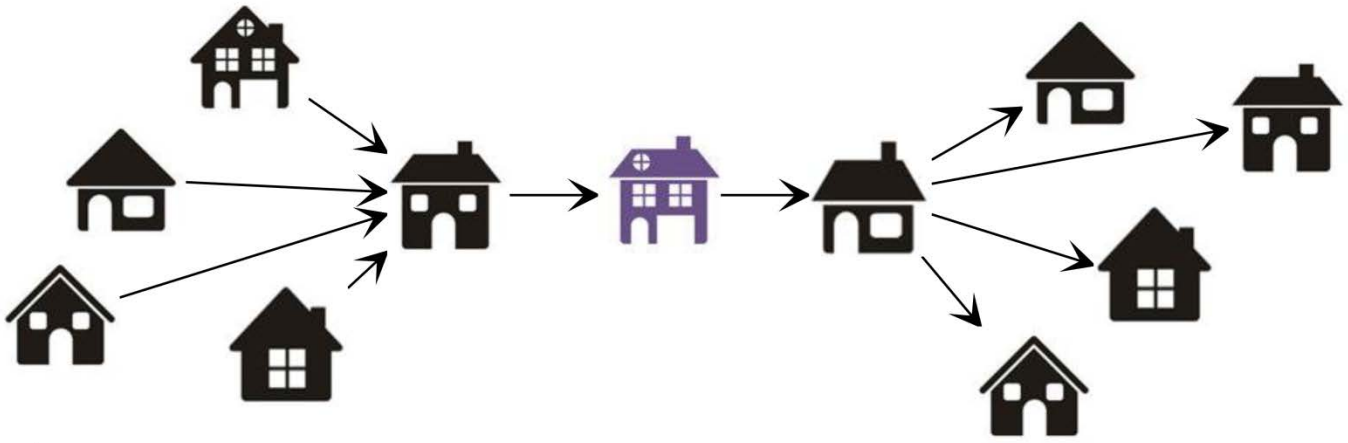
→ **Out-degree:** Number of direct outgoing trade contacts



Materials & Methods

Centrality parameters - “What characterizes a central or important farm?”

→ **Betweenness:** Number of shortest paths a farm lies on

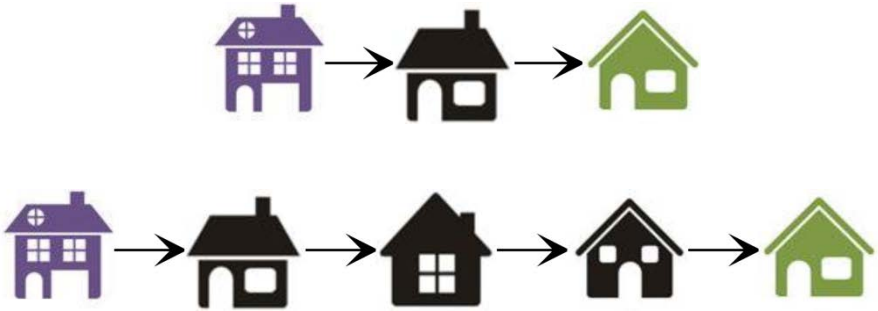


Materials & Methods

Centrality parameters - “What characterizes a central or important farm?”

→ **Ingoing closeness:** Average distance from all other reachable farms

→ **Outgoing closeness:** Average distance to all other reachable farms



Comparison

1. Between different network versions
2. Within different network versions

Comparison

1. **Between different network versions**

   Influence of various information loads on the outcome of network analysis

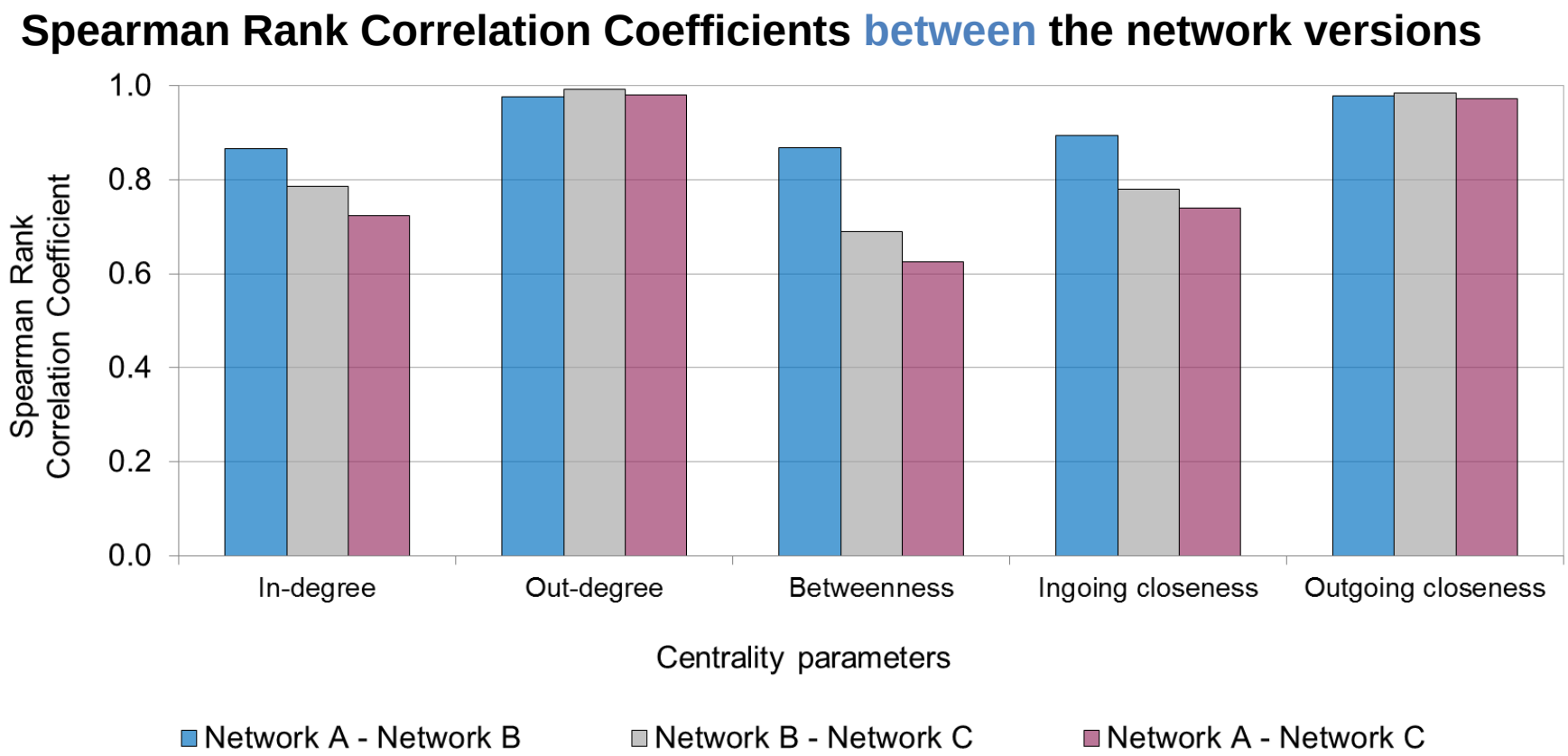
Materials & Methods

1. **Comparison between different network versions**

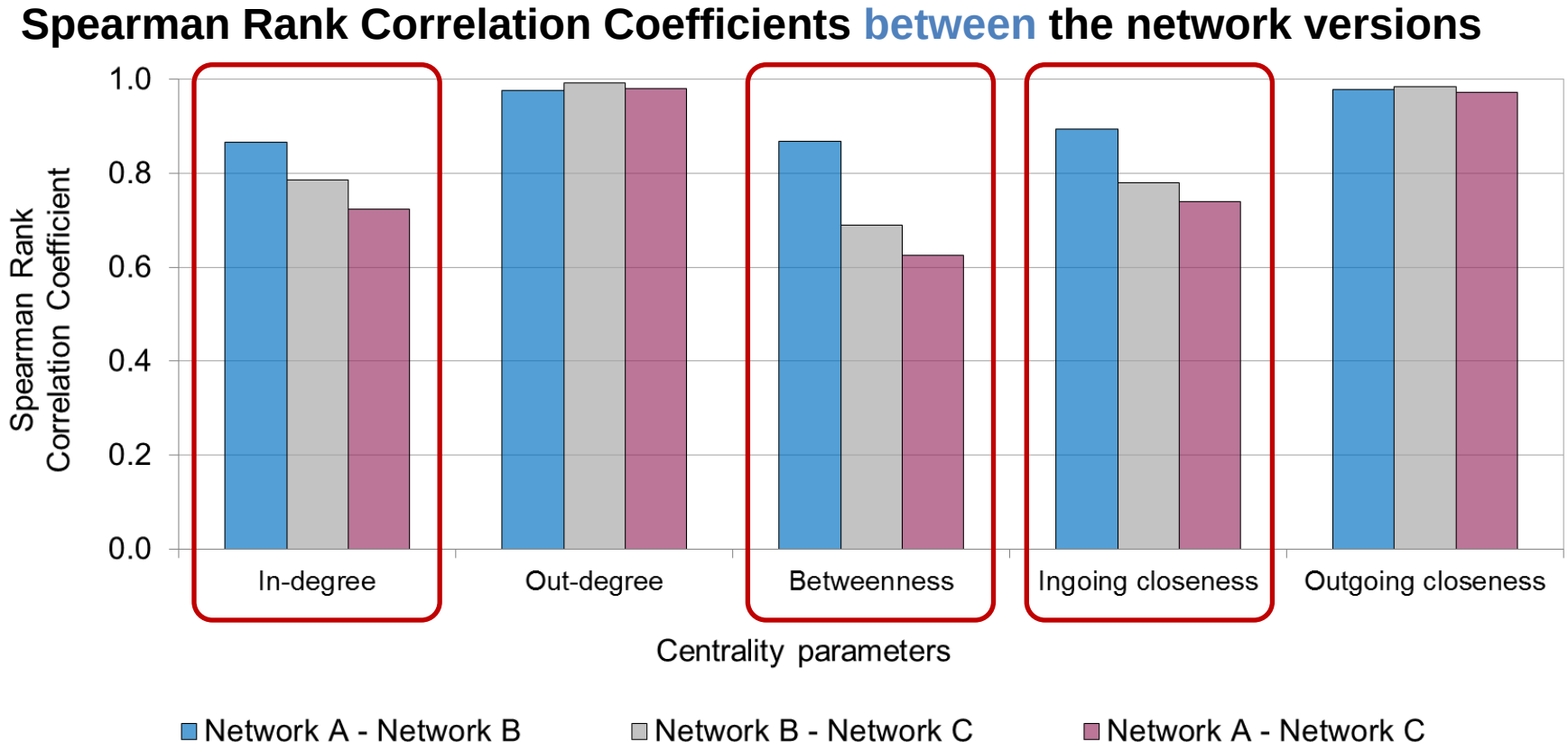
   → Pairwise calculation of the Spearman Rank Correlation Coefficients of the centrality parameters for all network versions

   - Network A – Network B
   - Network B – Network C
   - Network A – Network C

Results – 1. Between different network versions



Results – 1. Between different network versions



Discussion – 1. Between different network versions

Discussion

1. **Comparison between different network versions**

   → **Most robust results for out-degree and outgoing closeness**

   - Highly right-skewed distribution of out-degree and outgoing closeness
   - In-degree, ingoing closeness and betweenness had a smaller range
   - It is more likely to remove a node with a relatively high in-degree than a node with a high out-degree

Comparison

2. Within different network versions

Assumption: Network elements (nodes & edges) which appeared less frequently did not really belong to the studied producer community

Materials & Methods

2. Comparison within different network versions

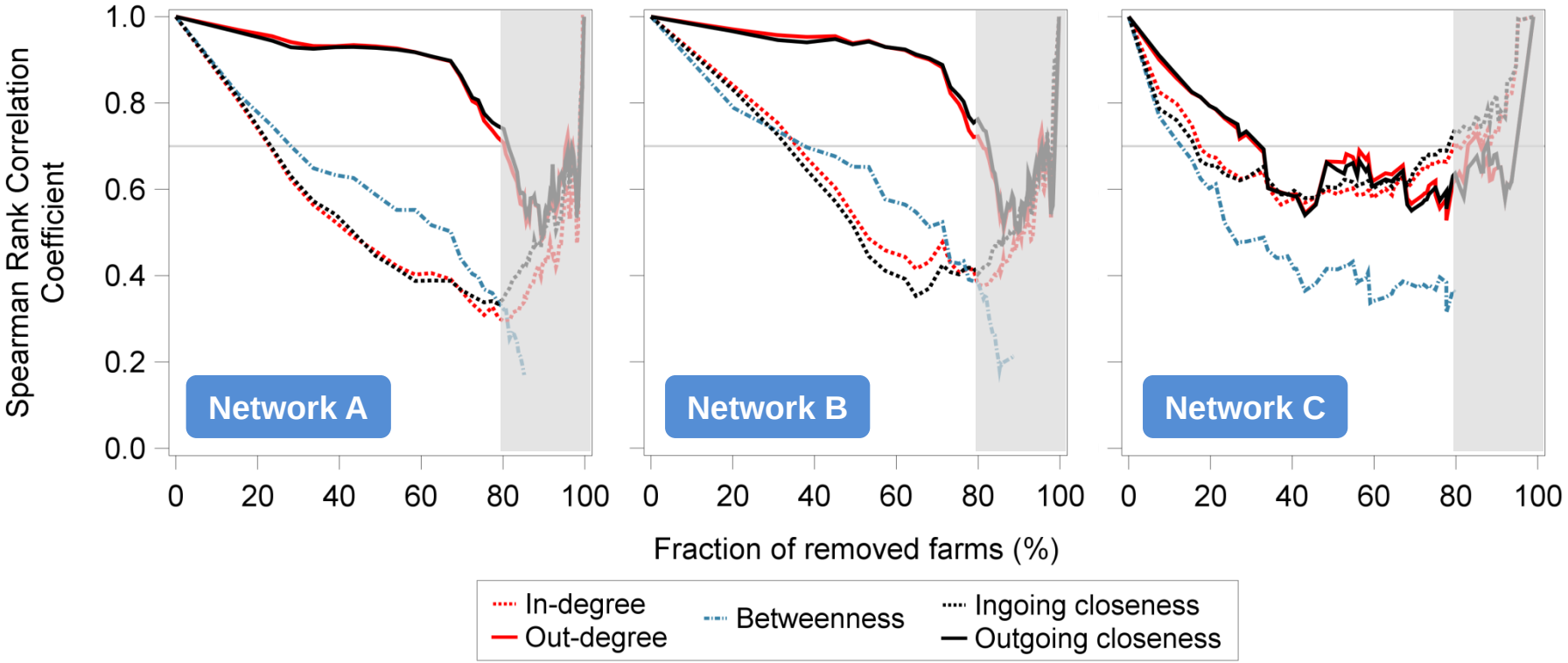
→ Two removal scenarios

- **Removal scenario 1** (Removal of edges according to their frequency of appearance)
- **Removal scenario 2** (Removal of nodes according to their frequency of appearance)

→ Calculation of the Spearman Rank Correlation Coefficient between the original network version and each removal step

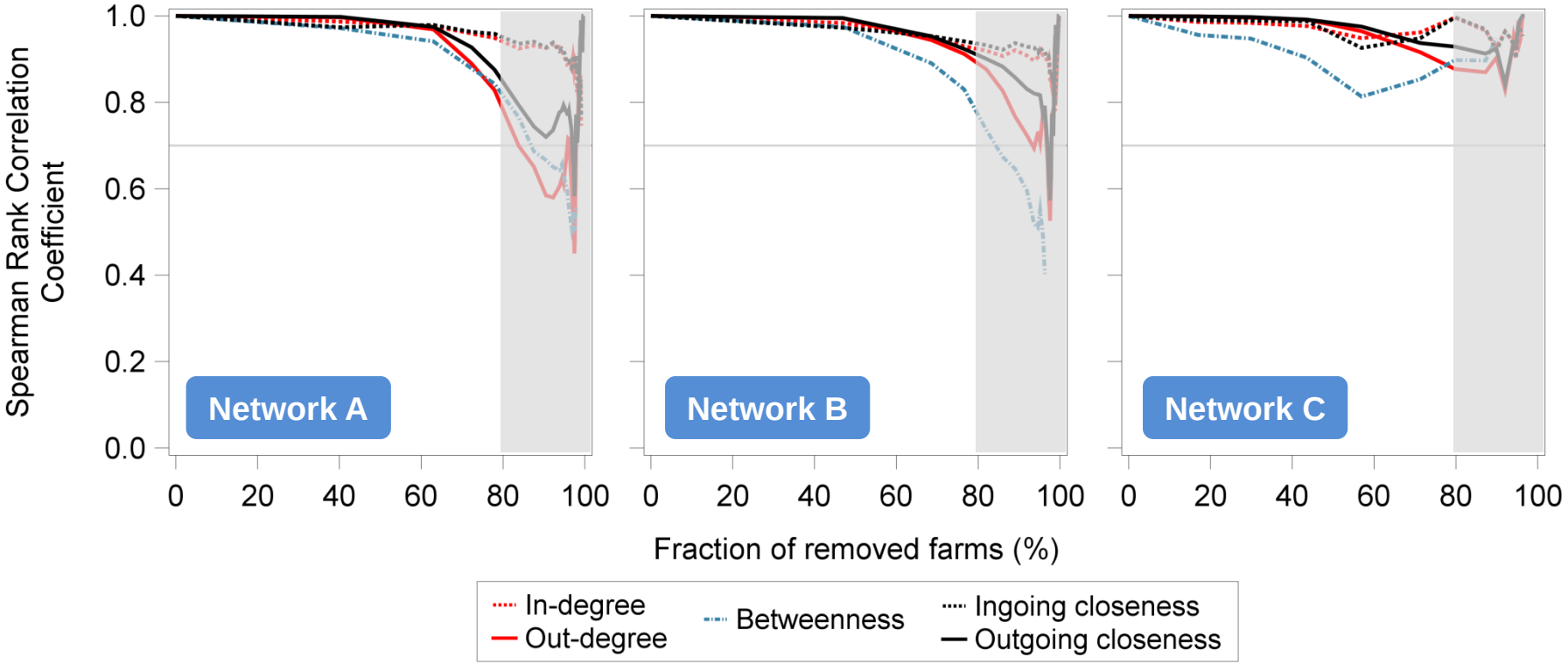
Results – 2. Within different network versions

Removal scenario 1 **Removal of edges according to their frequency of appearance**



Results – 2. Within different network versions

Removal scenario 2 **Removal of nodes according to their frequency of appearance**



Discussion – 2. Within different network versions

Discussion

2. Comparison within different network versions

→ **Removal of nodes had less impact than removal of edges**

- Explanation: Topology of the pork supply chain
- Farms with a low frequency are located at the margins of the network
- Removal of nodes only trims the margins of the network
- Edges with a low frequency can appear at every position in the network
- Removal of edges leads to a higher fragmentation of the network

Conclusion

- **Out-degree and outgoing closeness**: Stable ranking of the most central nodes in both comparisons
- **Reliable results** even if there are missing or false information in the data set
- Important if these centrality parameters are used for the **prediction of disease transmission** or the **implementation of disease control measures**

Thank you for your attention!

Deutsche Forschungsgemeinschaft

This project is kindly financed by the DFG.