Conceptual Framework for an online Early Warning System for Food Fraud Detection

Vahid Mojtahed[a], Hans Marvin[b], Yamine Bouzembrak[b], Rabin Neslo[c], Rolf Mader[d]

[a] Fera Science Ltd, Sand Hutton, England
[b] RIKILT Wageningen UR, Wageningen, the Netherlands
[c] Universitair Medisch Centrum Utrecht, the Netherlands
[d] FiBL, Switzerland

6th April, 2016 • Prague , The Czech Republic

Intro

- The costs of an adulteration incident averages between 2-15 (4-12) percent of a company's yearly revenues. This translates into a $400 million loss for a $10 billion company or as much as $60 million for a $500 million company (GMA, 2010).
- A sudden change in the status quo (economic, regulations or climate, etc.) could provide a susceptible environment to fraudulent activities and crimes. (Fraud/Crime Triangle)

Donald Cressey (1940s)

- There is a need for proactive methods (early warning systems) in contrast to reactive (detection and interventions) methods that are costly and time consuming for mitigating food fraud risks.

Controller and Auditor general New Zealand- 2011

FOOD INTEGRITY

Early Warning Systems

- Early Warning system: is a major element of disaster risk reduction. It prevents loss of life and reduces the economic impacts of the disasters.
- To be effective, the EWS needs to actively involve communities at risk, facilitate public awareness, disseminate alerts, and warning and ensure there is a constant level of preparedness.
- Four main functions: (UN-ISDR, 2014; Nuwan, 2010)

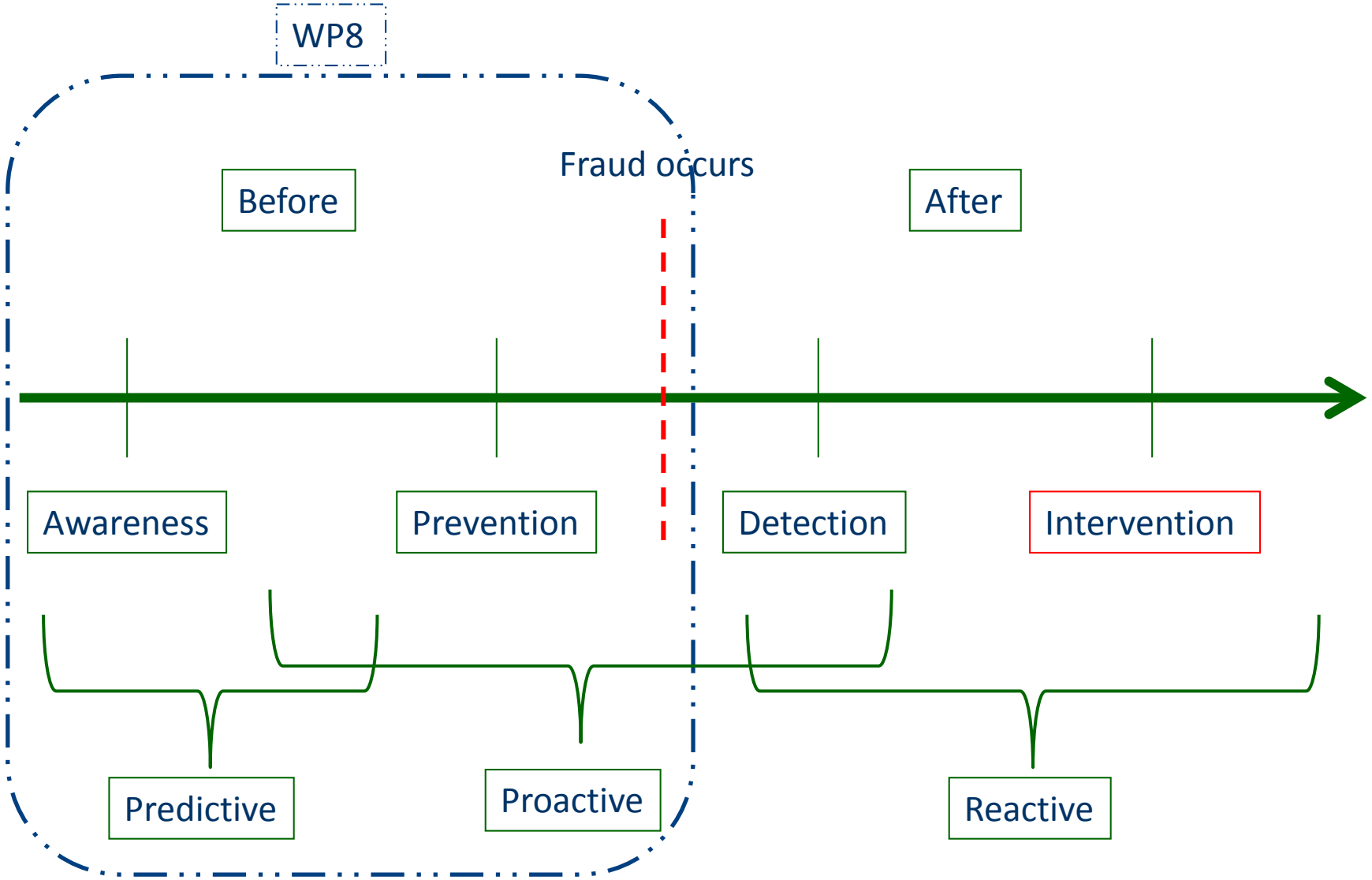
1) Risk analysis: **systematically collecting data** and undertaking **risk assessment of hazards and vulnerabilities**.

2) Monitoring and warning: studying the factors that indicate a disaster in imminent and methods that are used to **detect these factors** in a clear and understandable way.

3) Dissemination and communication: communicating the risk and warnings to those who are in danger

4) Response capability: building national and community response plan to testing of plan and promotion of readiness to ensure people know how to response.

Timeline of fraudelent event management



Our Objectives

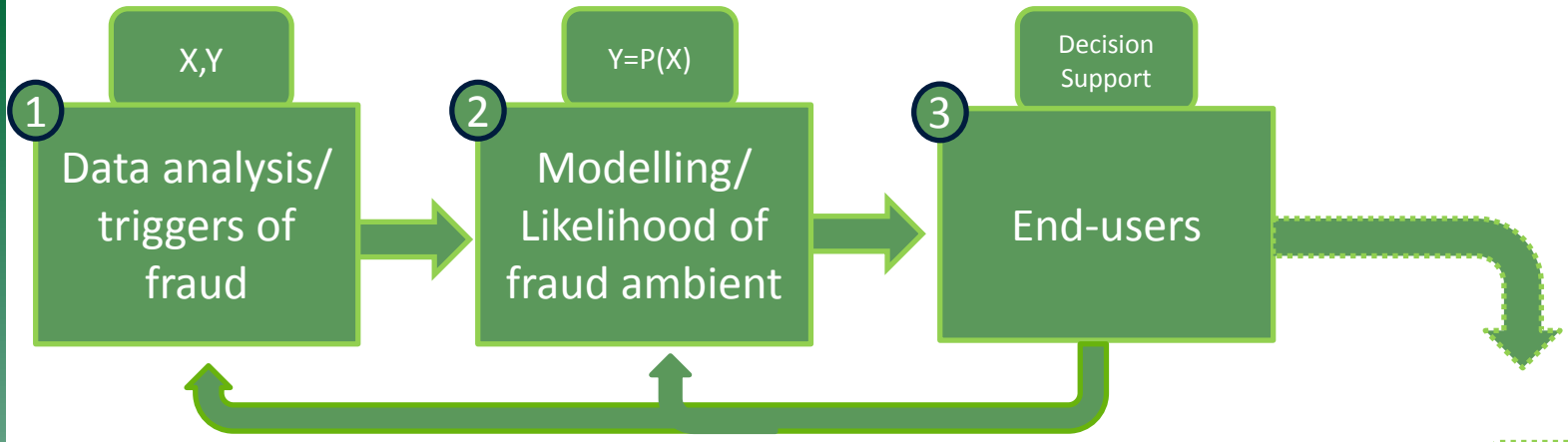
O1. Develop a structured approach for collecting and analysing information regarding potential drivers of the EU food chain fraud events and frequency of fraud incidents for commodities.

O2. Build innovative approaches and models to quantify the likelihood of fraud susceptible environment based on the information gathered in O1.

O3. Estimate the possible EU food chain fraud risks in some case studies based on likelihood of fraud (O2) and impact assessment (health and socio-economic).

Activities Workflow

1. Constructing databases with economic and non economic drivers as well as fraud incidences for analysis and identification of fraud triggers and tipping points. [Awareness]
2. Developing quantitative (probabilistic) predictive models of fraud linking causes to effects. [Predictive/Proactive]
3. Validating and generalizing fraud models by end-users (regulators, producers, retailers, etc.).
4. Estimating economic and health risks of fraud.



- Y: Effect
  - Fraud databases (RASFF, HorizonScan™, USP, etc.)
- X: Causes
  - Trade data (shocks, volume, etc.)
  - Relative global prices
  - Change in prices of substitutes/complementary
  - Durability of good
  - Profit margin of industry
  - Consumer behaviour (%income spent on that good)
  - Supply chain structure
  - Climate (temp, precipitation)
  - Pesticide usage for that product
  - Past events
  - Penalties to fraud
  - Testing difficulty/frequency
  - Change in regulations
  - Origin country

A quantitative model will be built for detecting triggers of fraud for some commodities where data are available and they are perceived as high risk. This model is not going to rely on the past events only but rather tries to identify emerging hazards by analysing available data and could be updated using Bayes rule to incorporate new information that becomes available through time.

This model could also capture subjective information provided by the experts and/or stakeholders for which no data is available.

The end-users could be regulators, retailers, or producers who will look at the likelihood of fraud to decide whether to test the product for using in the production or for allowing importing of it.

Risk = P x £

4 Economic/ Health Risk

Using the probability of the fraud (Hazard) and the data that describes the exposure to hazard (# customers lost, damages to business, loss of lives, etc.) we attempt to monetize the direct and indirect costs of fraud.

FoodIntegrity Framework

1. Constructing and analysing databases

I. Data mining/statistical analysis on the market and non market-based triggers of the food fraud including prices and trades of commodities.

- Early detection methods: Based on anomaly and breakout detection in the data and on identifying outliers in the time series of prices, trade, production, etc.
- Used by IT companies, insurance and credit card companies for detecting unusual patterns (traffics) in servers' metrics or financial fraud.

II. Constructing a data-base of food fraud frequency for different commodities

- Text mining activities (Coordinated with RIKILT)
- HorizonScan™

Taxonomy of fraud triggers in the literature

Socio-economic Characteristics

- Relative prices of ingredients/inputs
- Price of final product
- Consumers' behaviour (Preferences, attitude, consumption pattern, etc.)
- Demand characteristics
- Excess demand / short supply with lags (previous year demand or current year supply)
- Economic downturn (Dennis&Kelly, 2013)
- Tariffs, taxes, and custom duties – country of origin mislabelling (China) for avoiding anti dumping duties

Industries' or sectors' Characteristics

- Profit margin (Gross value added as a proxy)
  - Low profit margin driving fraud in high volume products
  - High profit margin driving fraud in high premium products
- Supply chains length and complexity
- Production capacities

Commodities' Characteristics

- Durability (Dennis&Kelly, 2013)
- Substitutability – pomegranate juice with apple/grape juice

Climate/Pests/Health

- Climate variation: addition of Sudan dye when paprika has mould
- Vulnerability to pests/antibiotics residues
- Visits to vets

Regulations and Standards

- Penalties
- Change in regulations and standards – Ochratoxin A from 30 µg/kg to 20 µg/kg for spices

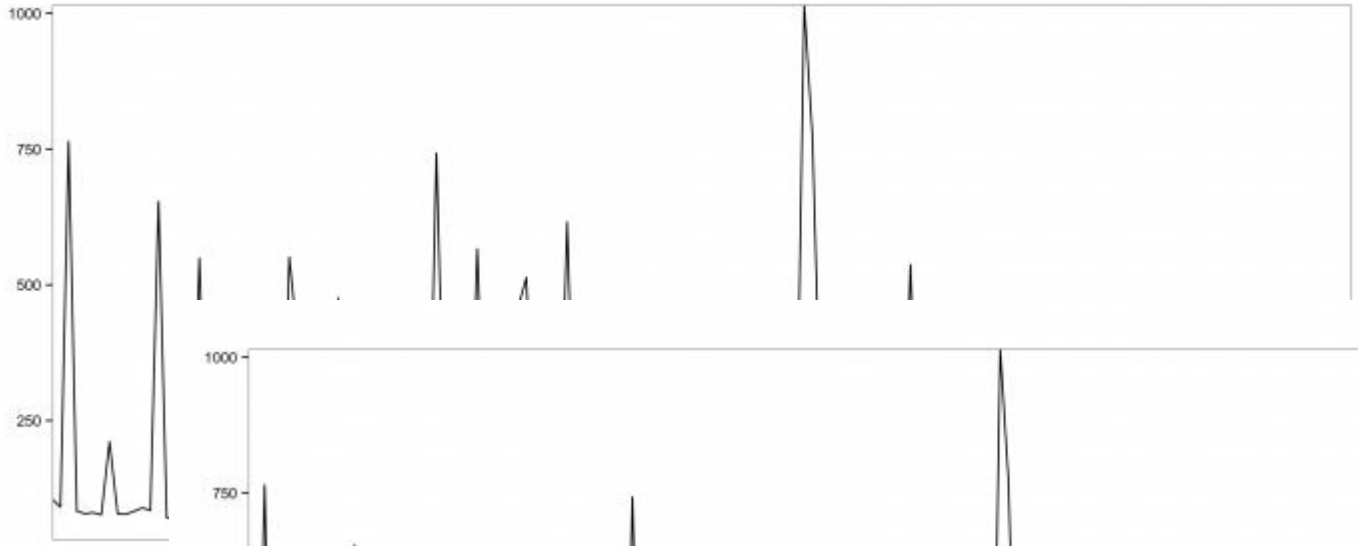
FOOD INTEGRITY

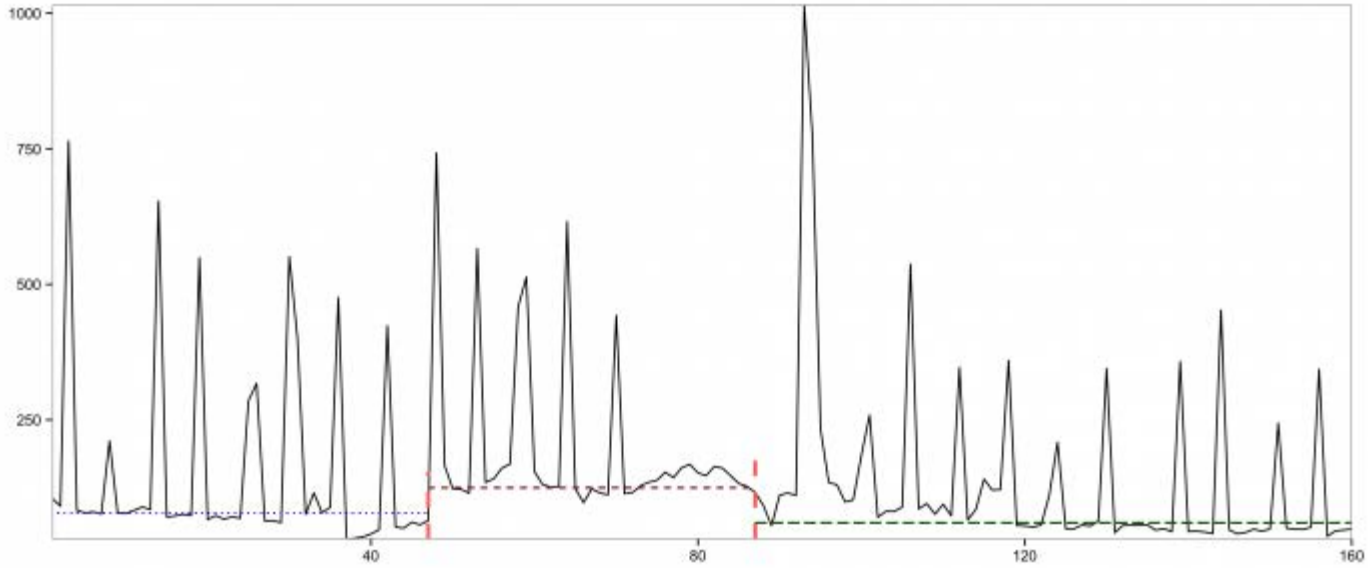
Breakouts and Anomalies

- Anomalies are **point-in-time** anomalous data points.
- Breakouts are characterized by a tilt from one **steady state** to another.
- Some techniques in anomaly detection are not applicable in the context of economic data because of the inherent **seasonal and trend** components.
- Also, anomalies are **contextual** in nature and hence, techniques developed for anomaly detection in one domain cannot be easily used in another domain.

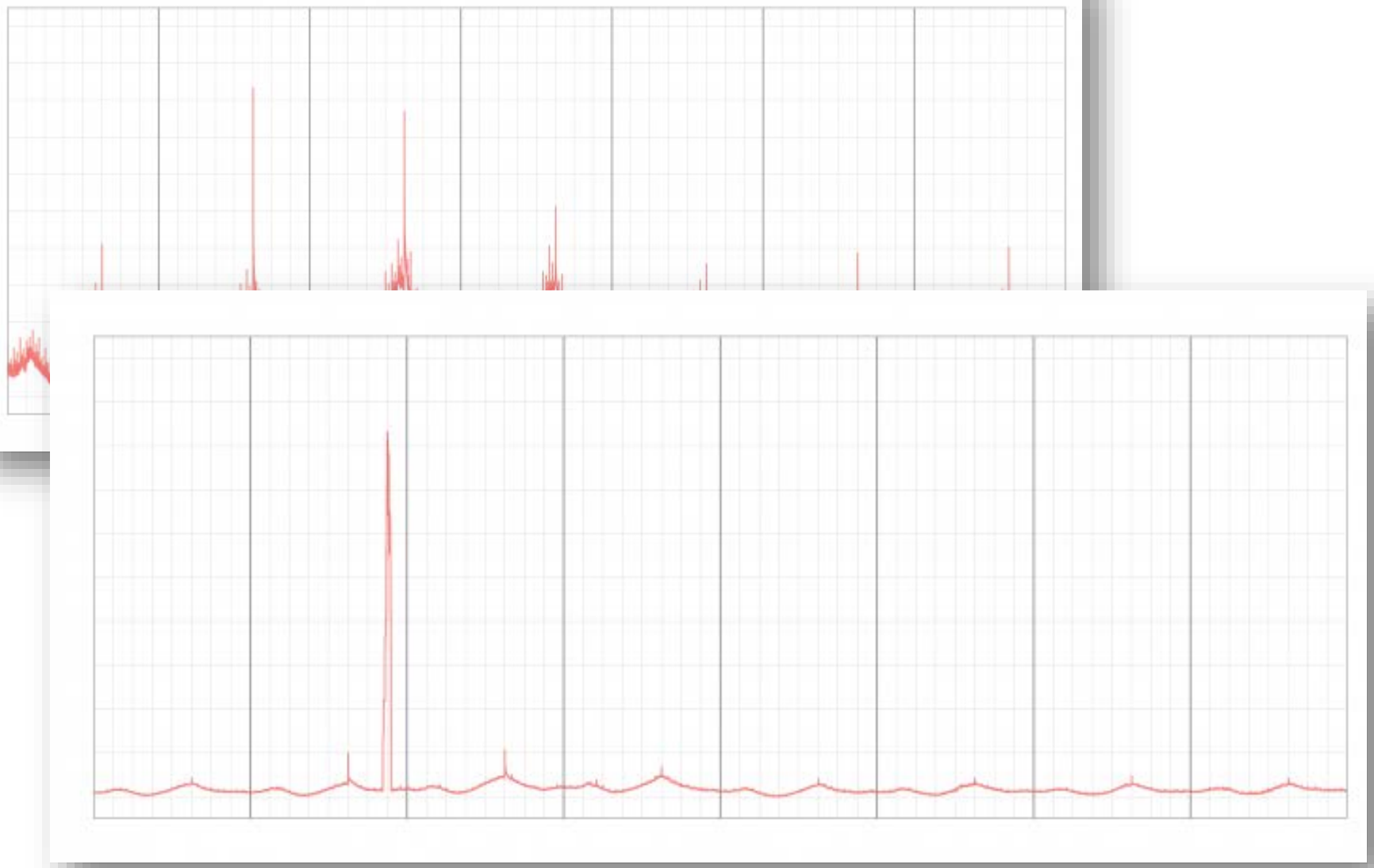
FOOD INTEGRITY

Breakout Vs Anomaly detection



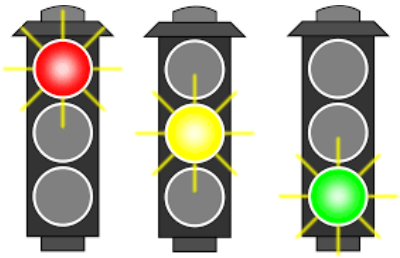


Breakout Vs Anomaly detection

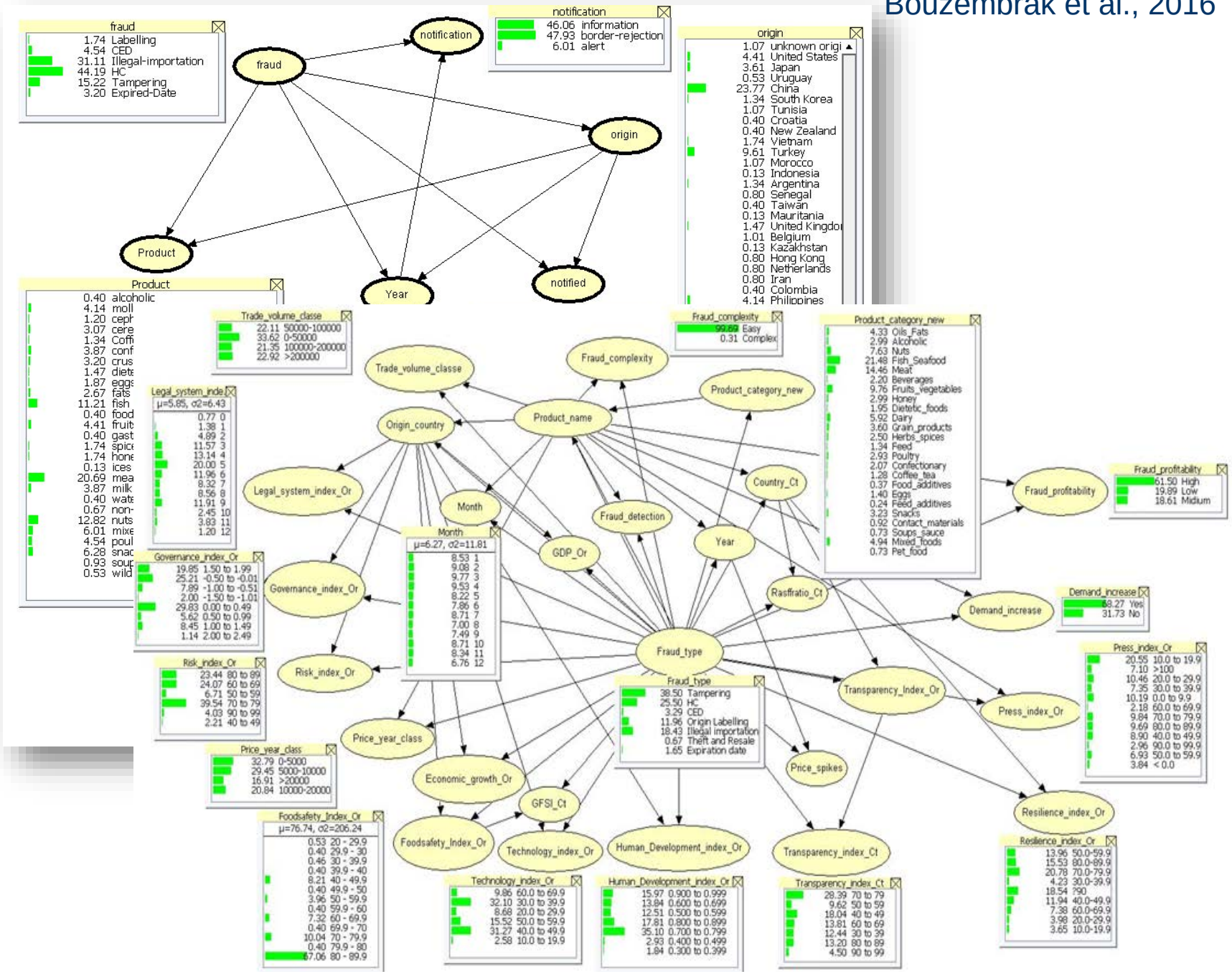


2. Estimating the likelihood of fraud susceptible environment

- A Bayesian Network of Belief has been developed based on data from EMM, RASFF, etc.
- We will attempt to build quantitative models that could provide us with the likelihood of the environment prone to food fraud incidents.
- The parameters of the model will be estimated and they identify the weights of each trigger contributing to the probability of fraud.
- This model will be online meaning that it is going to be a real-time indicator of the likelihood of fraud with the parameters of it constantly changing as new data arrives.



Bouzembrak et al., 2016



3. Validation and generalization of the fraud model

- The outcome of Workflows 1&2, provides the end-users with a decision-making toolbox for predicting fraud along their supply chain.
- The activities surrounding statistical analysis and quantitative model estimation will be carried out for selected commodities identified by end-users which are among the most common fraudulent commodities.
- Once they are constructed, they are going to be validated, refined, and generalized by end-users in an iterative process.

4. Economic and health risk estimation of fraud

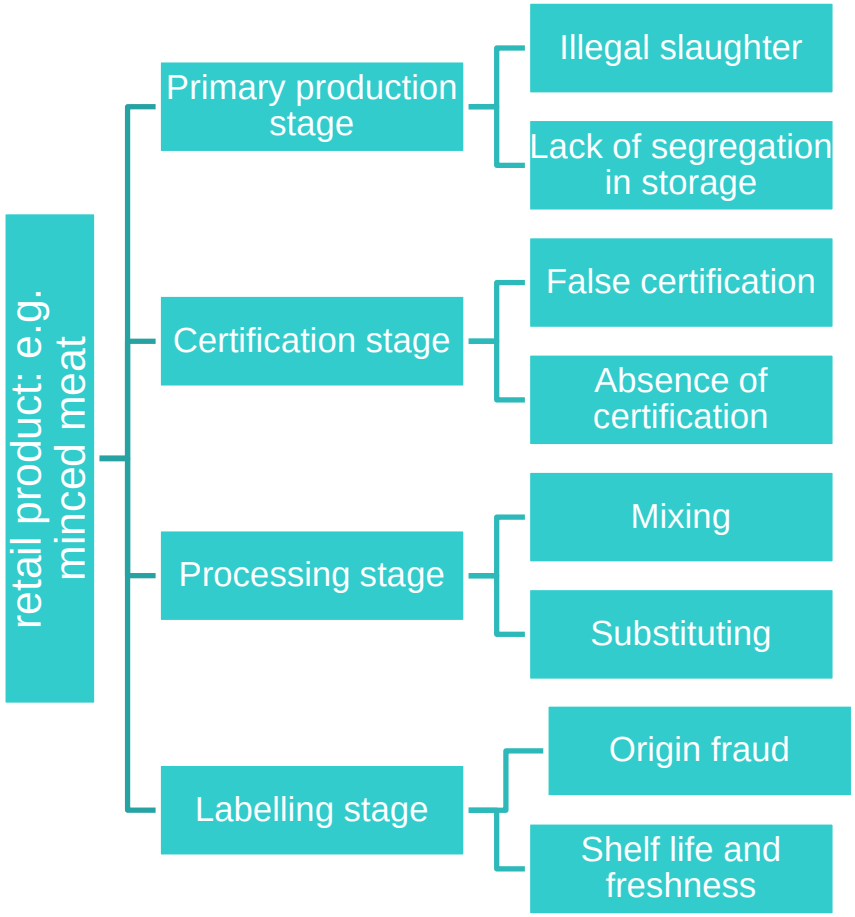
- The impacts' assessment is important for prioritising resources dedicated to detection, intervention, and prevention.
- A case-specific rather than holistic analysis will be carried out because:
  - the impacts' assessment will vary depending on the commodities types, and fraud types (safety or not)
  - The activity is very data intensive
  - The data is not publicly available or difficult to quantify
- Case studies with industries/businesses to assess the potential impacts of the fraud in their food supply chain.

Commodity: Beef

Impacts

- Health
  - Death/illness
  - Measured antibiotic resistance
- Economics
  - Consumer confidence loss
  - Reputational damage
  - Recalls
  - Adverse media
  - Organizational disruption
- Environmental
  - Waste disposal
  - Change in Carbon footprint

Fraud Type



Conclusions

- Changes in status quo makes the food supply chain more susceptible to food fraud incidents that are rooted in incentives for economic gains.
- We need a set of sensitivity analysis and data mining methods to detect subtle individual or collective changes in the triggering factors of the fraud events.
- Reliability, lead-time, and scope are three key metrics for evaluating any Early Warning System that needs further investigations.
- The impact assessment of food fraud while inherently important, it is difficult and complex in nature and should be tailored to businesses.