Better Training for Safer Food Initiative

Helmut Hartmann

Better Training for Safer Food is an initiative of the European Commission aimed at organising an EU training strategy in the areas of food law, feed law, animal health and animal welfare rules, as well as plant health rules.

2nd Quarter 2017
Information systems - Practical implementation of specific functionalities

Session 4.4

- Systematic development & testing
- Traceability functionality
- Risk analysis
- On-spot control and cross-compliance based on the working concept of the German animal identification, registration and traceability system (HIT)

Q&A Session
Systematic development & testing – part 1
Interaction of stakeholders during development process
Systematic development & testing – part 2

- Systematic testing of every aspect of the business logic
  - own test framework as HITP-client to communicate with appl. server
  - design test scenarios with data and interaction workflow for every user case
  - define expectations and proof assertions
  - some 10,000 test cases
Online-Demonstration

Live-Cycle of an animal based on the working concept of the German animal identification, registration and traceability system (HIT)
Traceability – part 1: at the small scale

- Animal live cycle
Traceability – part 2: at the large scale

- Epidemiologic research: trace back / trace forward

B = level backward

F = level forward

- Primary outbreak
- holder with contact
- indirect – consider
- indirect – ignore
- Ignored holder

© - 2014- European Union. All rights reserved. Certain parts are licensed under the conditions to the EU
Traceability – part 3: technical historiography

- Implicit technical historiography
  - Timestamp of storage – “valid from” (SYS_VON)
  - Timestamp of revocation – “valid until” (SYS_BIS)
  - No physical DELETE – only STORNO (like in book-keeping)
  - No physical UPDATE – only STORNO + INSERT
  - Responsible user for storage or revocation (MELD_BNR/MBN)

- Example: Sequence to correct business begin (DTYP_VON)
Traceability – part 4: Using technical history

- Delta transfer
  - Possible to retrieve only new or modified data, e.g. for replication
  - Store query and execution timestamp for each user

- Point in time retrieval
  - Generate reports at specific “knowledge time” (e.g. grass-land & diary cow premium)
  - Time reproducibility of data queries (reports & statistics)

- Example: retrieve data as it looked like on 1. Jan. 2010 12:00h
Traceability – part 5: business history

- Logical business historiography
  - Timestamp begin of effect – “effective since”
  - Timestamp end of effect – “effective until”
  - Distinguish between
    - Correction of information – it was stored but wrong
    - Change of information – it was correct in a certain period

- Example: cattle holder (Type: 1) switched to pig production (Type: 31)
Risk analysis – part 1 legal basis

- Integrated risk analysis and common on spot control regime
  - veterinary and IACS / CC, according to regulation (EC) 2630/97, regulation (EC) 2419/2001
  - common data basis, risk parameters and calculation model
  - consistent documentation through identical control reports

- Guidelines and provisions for parameter selection
  - see article 2 paragraph 4 regulation (EC) 1082/2003
  - explicitly mentioned
    - amount of livestock, changes in stock compared to previous year
    - aspects of public health
    - aspects of animal health, protection and welfare
    - control results and findings from previous years
Risk analysis – part 2 legal basis

- Amount of entities to control on spot
  - according to vet. requirements: formerly 5% now 3% of animal holders
  - according to IACS: 1% of subsidy applicants
  - in case of a certain percentage of irregularity, percentage has to be increased

- Requirement to draw comparison group by random selection
  - according to IACS: 20% - 25% within total amount
  - according to vet. resort under discussion
    - formerly we used 20%
    - according to DG AGRI one should use 0% (no random) for bovine and carpine
Risk analysis – former, classical approach

- Calculation of individual risk for holding via weight or point system
  - within different criteria a calculation formula or grouping scheme is established to assign a certain amount of “risk points” to each holding
  - points for different criteria weighted and combined as “score value”

- Drawing the required amount of holdings for on spot control
  - “risk proportional” – by random, probability derived from score
  - “worst first” – take holdings in descending succeeding order

- Improve and adjust by retrospection of the control results

- Main disadvantages
  - depends highly on expert appreciation
  - difficult to improve systematically
Risk analysis with statistical methodology

paradigm shift towards mathematical and statistical methods

- According to EU bodies in announcements and control visits
  - it’s secondary which parameter you use
  - of primary importance is to use effective parameters
  - you have to monitor and adjust your parameter selection
  - the covered risk must be significantly higher in the risk drawn population compared to the randomly drawn comparison group

- Best practice is to use statistical methodology
Risk analysis statistical approach

Year N - 1
2012

- Animal & holding data I & R system
- Aid application data
- Various other data
- Result off on spot control (risc drawn)
- Results (random)

Year N
2013

- Animal & holding data I & R system
- Aid application data
- Various other data
- Result off on spot control (risc drawn)
- Results (random)

Year N + 1
beginning
2014

- Risc score for every holder
- Holders drawn by maximum risc
- H. drawn by random

Year N + 1
end of year
2014

- Result off on spot control (risc drawn)
- Results (random)
Risk analysis – details, step by step

The following slides show the process of the model building and model application very detailed - step by step
Risk analysis – step 1: data preparation

- Data basis for risk analysis
  - extract data from AIT system
  - receive data from paying agencies and vet. admin off all states

- Plausibility check for potential parameters
  - descriptive statistic (sample next slide) for overview and check
  - disqualification of parameters
    - insufficient amount of observations
    - unusable – to little variation, date values

- Decisions to be taken
  - which scope? (bovine, carpine, first pillar aid ...)
  - do we need combination/integration of different scopes?
  - what is the prediction objective?
    (breach yes/no, sanction height – absolute or relative ...)
## Risk analysis – step 1: preparation, descriptive statistic

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>name</em></td>
<td>nummer</td>
<td>CA01_FLA</td>
<td>CA02_FLA</td>
<td>CA03_FLA</td>
<td>CA04_FLA</td>
<td>CA05_FLA</td>
<td>CA06_FLA</td>
<td>CA07_FLA</td>
<td>CA08_FLA</td>
<td>CA100_BTKA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>1</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td>119980</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>PCT_POP</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>NMISS</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>MIN</td>
<td>4</td>
<td>0</td>
<td>-9999</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>MEAN</td>
<td>5</td>
<td>26.85</td>
<td>0.98</td>
<td>17.27</td>
<td>10.53</td>
<td>0.86</td>
<td>0.86</td>
<td>1.1</td>
<td>1.3</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>MAX</td>
<td>6</td>
<td>2837.6</td>
<td>100</td>
<td>1743.95</td>
<td>1227.65</td>
<td>647.34</td>
<td>642.38</td>
<td>272.47</td>
<td>547.46</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>P1</td>
<td>7</td>
<td>0.38</td>
<td>-37</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>P5</td>
<td>8</td>
<td>1.2</td>
<td>-9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>P10</td>
<td>9</td>
<td>2.58</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>P25</td>
<td>10</td>
<td>6.25</td>
<td>0</td>
<td>0.48</td>
<td>1.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>P50</td>
<td>11</td>
<td>16.78</td>
<td>0</td>
<td>7.63</td>
<td>5.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>P75</td>
<td>12</td>
<td>36.26</td>
<td>0</td>
<td>21.27</td>
<td>14.62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>P90</td>
<td>13</td>
<td>61.98</td>
<td>3</td>
<td>45.89</td>
<td>27.14</td>
<td>1.49</td>
<td>0.88</td>
<td>2.19</td>
<td>2.47</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>P95</td>
<td>14</td>
<td>84.89</td>
<td>14</td>
<td>67.01</td>
<td>37.15</td>
<td>3.69</td>
<td>3.8</td>
<td>5.86</td>
<td>5.9</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>P99</td>
<td>15</td>
<td>151.75</td>
<td>100</td>
<td>129.25</td>
<td>63.05</td>
<td>13.53</td>
<td>19.35</td>
<td>21.32</td>
<td>23.04</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Generated via “Better Means”-Macro by Myra A. Oltsik and Peter Crawford (Paper 059-31 / SUGI 31)
Risk analysis – step 2: choosing parameter candidates

- Prerequisites and primary selection
  - more than 230 candidates, some 100,000 observations
  - stepwise „forward-selection“ combined with „backward-elimination“
  - inclusion or exclusion by experts decision possible

- Analysis of interdependences (correlation analysis)
  - two different statistical methods (see sample)
    - simple cross correlation matrix
    - variance inflation factor – a measure for multiple inter correlation
  - often highly correlated parameters, one of group selected by expert
  - advantage of reduction
    - avoiding model bias
    - easier to handle
Risk analysis – step 2: choosing ..., correlation analysis

**Variance inflation**
- Exclude > 10

**Cross correlation matrix**
- Attention if corr.value > 0.4 and < -0.4
Risk analysis – step 3: model building

- Iterative model generation
  - build model from pre selected parameter candidates
  - calculate model quality measures and prediction certainty
  - exclude parameters with weak or uncertain influence
  - iterate to improve

- Model assessment by means of quality measures
  - Akaike's information criteria (AIC)
  - p-values for model and parameters
  - convergence

- Used statistical techniques
  - logistic regression for binary forecast objects
  - linear or regression for continual forecast objects
Risk analysis – step 3: model building / example

model example:

\[
Risk = (-1.9176 + \frac{0.8709}{(-0.00004)} * R11_HB_ST1) + (0.0252 * R120_FWA2) + (-0.00074 * R24_BF_FFN) + (0.00544 * R27_FL_EI) + (0.0180 * R33_FL_NAL) + (0.0141 * R40_FL_AFS) + (-1.5001 * R43_FL_SFA) + (0.0547 * R46_FL_DK) + (-0.00123 * R54_DSB_R) + (-0.0316 * R59_DSB_PF) + (0.8572 * R73_AB_EGF) + (-0.8200 * R77_S_ELE1) + (1.2814 * R78_S_ELE2))
\]

transformation:

\[
Scoring = \frac{e^{Risk}}{1 + e^{Risk}}
\]
**Risk analysis – step 4: forecast, scoring, drawing**

- Apply model to data of current year to forecast
  - calculate formula for parameter values of each holder
  - no interpretation possible for resulting objective value at this stage

- Transformation according to used statistical techniques
  - for binary models: probability or breach / irregularity
  - for continuous models: estimated height of sanction

- Drawing of holders for on spot control
  - different percentage according to regulation requirements
  - complex additional logic if integration of different scopes
Risk analysis – step 5: assessment of effectiveness

- All control report results are completely entered into CDB
- Effectiveness calculated by comparing risk drawn to random
  - assessment in regard to different objectives
  - statement whether better or not AND whether discrepancy is significant
Thank you for your attention!
Better Training for Safer Food is an initiative of the European Commission aimed at organising an EU training strategy in the areas of food law, feed law, animal health and animal welfare rules, as well as plant health rules.

**ADT Projekt Gesellschaft der Arbeitgemeinschaft Deutscher Tierzüchter mbH**

Adenauerallee 174  
53113 Bonn, Germany  
Internet: [www.btsf.adtprojekt.net](http://www.btsf.adtprojekt.net)  
[www.adtprojekt.de](http://www.adtprojekt.de)  
E-Mail: [2013-96-03@adt.de](mailto:2013-96-03@adt.de)  
[adtprojekt@adt.de](mailto:adtprojekt@adt.de)  
Tel.: +49 (228) 91447 - 30  
Fax: +49 (228) 91447 - 31

---

**Better Training for Safer Food**  
**BTSF**  
European Commission  
Consumers, Health and Food Executive Agency  
DRB A3/042  
L-2920 Luxembourg