# Wageningen UR Livestock Research

*Partner in livestock innovations*

Report 590

# Description of the calculations to transform measures into criterion for the Welfare Quality® assessment protocol for laying hens

May 2012

LIVESTOCK RESEARCH
WAGENINGEN UR

## Colophon

**Publisher**
Wageningen UR Livestock Research
P.O. Box 65, 8200 AB Lelystad
Telephone +31 320 - 238238
Fax +31 320 - 238050
E-mail info.livestockresearch@wur.nl
Internet http://www.livestockresearch.wur.nl

**Editing**
Communication Services



Wageningen UR Livestock Research and Central Veterinary Institute of Wageningen UR, both part of Stichting Dienst Landbouwkundig Onderzoek (DLO Foundation), together with the Department of Animal Sciences of Wageningen University comprises the Animal Sciences Group of Wageningen UR (University & Research centre).

Single numbers can be obtained from the website.




**Abstract**
Results of a study on the Welfare Quality® assessment protocol for laying hens. It reports the calculations for transformation of welfare measures into scores per criteria of the Welfare Quality® assessment protocol.

**Keywords**
laying hens, Welfare Quality, welfare assessment

**Reference**
ISSN 1570 - 8616

**Author(s)**
T.G.C.M. van Niekerk
J.T.N.M. Thissen
C.G. van Reenen

**Title**
Description of the calculations to transform measures into criterion for the Welfare Quality® assessment protocol for laying hens

Report 590

LIVESTOCK RESEARCH
WAGENINGEN UR

Report 590

# Description of the calculations to transform measures into criterion for the Welfare Quality® assessment protocol for laying hens

# Beschrijving van de berekeningen om metingen om te zetten naar criteria socres voor het Welfare Quality® protocol voor leghennen

T.G.C.M. van Niekerk
J.T.N.M. Thissen
C.G. van Reenen

May 2012

Ministerie van Economische Zaken,
Landbouw en Innovatie

**This study was financed by the Ministry of Economic Affairs, Agriculture and Innovation.**

## Preface

Between 2004 and 2009 a method was developed within the Welfare Quality® project to assess animal welfare on cattle, pig and poultry farms. The resulting Welfare Quality® assessment protocols, published in 2009, provide a detailed account of the necessary measurements, and how these can be combined to provide a single overall statement of the state of welfare on the farm assessed. The method has attracted a lot of interest from European and national policy makers, NGO's and the farming community, but has to date not been adopted in any commercial scheme nor used by individual farmers to improve animal welfare on their farm. The main drawback seems to be the amount of time required to carry out the measurements. In 2010 the Dutch ministry of Economic Affairs, Agriculture and Innovation commissioned Wageningen UR Livestock Research to do a series of studies aiming to simplify and shorten the original protocols. In collaboration with former Welfare Quality® partners and the Dutch laying hen sector, farm visits were organised and data collected between April and August 2011. The findings, along with an attempt to find simpler alternatives for the time consuming measurements are presented in report 589. For laying hens the model to combine individual measurements to scores per welfare criteria had not been developed, so prior to calculating the outcomes in report 589, a model has been developed using the same systematics as used to develop the model for the other animal species. This report represents the model for laying hens and the work done to develop it.
The results will be presented to the international Welfare Quality® Network, which is working on further improvement of the protocols. They will also be recommended to the Dutch Ministry, who can introduce them for improvement of farm animal welfare in collaboration with the Dutch laying hen sector.
Without these three stakeholder groups, the Dutch Ministry, representatives from the laying hen industry and the Welfare Quality® Network this work would not have been possible. On behalf of the project team I would like to thank Bart Crijns, Amanda Manten, and Léon Arnts (Ministry of Economic Affairs, Agriculture and Innovation) and Henk Hulsbergen (PPE) for their contributions to this work.

Paul Vriesekoop
June 2012

## Summary

The European Welfare Quality® project developed standardized ways of assessing animal welfare for different categories of farm animals, e.g. broiler chickens and laying hens, sows, growing pigs, veal calves and dairy cattle.
Welfare Quality® assessment protocols are based on the approach that welfare is a multidimensional concept, that comprises both physical as well as mental health. Within the Welfare Quality® project the same framework has been used to measure welfare of animals. Different measures of welfare, e.g. for laying hens the number of keel bone deformations, are integrated into a score for twelve independent welfare criteria. These criteria are integrated into four principle scores, and these are subsequently integrated into one overall score for a flock.
The measures for the welfare assessment protocol for laying hens has been described in the Welfare Quality® assessment protocol for poultry (Welfare Quality®, 2009). The protocol describes measures indicative of laying hen welfare on-farm.

To evaluate the welfare of a flock, measures need to be transformed into criteria and principles. The transformation into principles and the classification of flocks is standard for all animal species. The transformation of measures into criteria is species-specific and is based on a multi criteria analysis with the help of experts specialized in the type of animal involved.

This report contains the description of the models and methodology to calculate the criterion-scores for the laying hens. The general construction is common as used for all other animal species:

- *Short indication of the measures used to construct the criterion* (for detailed description of the measures see the NEN-report: Welfare Quality®, 2009)
- *Type of mathematical construction used*
- *Parameters to be defined* (list of the parameters needed to assess the criterion according to the construction previously proposed)
- *Adjustments of the mathematical construction* (including the questions asked to experts, their answers and the exact adjustments made)
- *Calculation* (definition of the variables, from raw data or obtained by calculation, and algorithm summarising the construction)

To transform the measures of the Welfare Quality© protocol for laying hens into scores per principle 5 experts have been asked to fill in tables for the multi criteria analysis. The 5 experts consulted are: Claire Weeks (UK), Marion Staack (D), Ute Knierim (D), Ingrid de Jong (NL) and Thea van Niekerk (NL).

## Samenvatting

Het Europese Welfare Quality® project heeft gestandaardiseerde methoden ontwikkeld om het welzijn te bepalen van verschillende categorieën van landbouwhuisdieren, b.v. vleeskuikens en leghennen, zeugen, vleesvarkens, kalveren en melkvee.
Welfare Quality® protocollen zijn gebaseerd op de benadering dat welzijn een multidimensionaal concept is, dat bestaat uit zowel fysieke als mentale gezondheid. Binnen het Welfare Quality® project is hetzelfde raamwerk gebruikt om het welzijn van dieren te meten. Verschillende metingen van welzijn, b.v. voor leghennen het percentage borstbeenvervormingen, worden geïntegreerd tot een score voor twaalf onafhankelijke welzijnscriteria. Deze criteria worden vervolgens geïntegreerd tot scores voor vier principes en deze worden vervolgens geïntegreerd tot een overall welzijnsscore per koppel.
De metingen voor het leghennenprotocol zijn beschreven in de uitgave "The Welfare Quality® assessment protocol for poultry" (Welfare Quality®, 2009). Het protocol beschrijft de metingen die indicatief zijn voor het welzijn van leghennen op het legbedrijf.

Om het welzijn van een koppel leghennen te meten, moeten de metingen getransformeerd worden in criteria en principes. De transformatie van criteria nar principes en de eind-classificatie van koppels is gelijk aan die voor alle diersoorten. De transformatie van metingen naar criteria is soort-specifiek en gebaseerd op een multi-criteria analyse met behulp van experts, gespecialiseerd in de betrokken diersoort.

Dit rapport bevat de beschrijving van de modellen en de methodologie om de criteria-scores te berekenen voor leghennen. De algemene constructie is gelijk aan die voor alle andere diersoorten:

- *Korte indicatie van de metingen die gebruikt zijn bij de constructie van het criterium* (een gedetailleerde beschrijving van de metingen staat in het NEN-rapport: Welfare Quality®, 2009).
- *Gebruikt type mathematische constructie* (lijst van parameters nodig om het criterium te bepalen volgens de eerder aangegeven constructie)
- *Aanpassingen aan de mathematische constructie* (inclusief de aan de experts gestelde vragen en hun antwoorden)
- *Berekeningen* (definitie van de variabelen, van ruwe data of verkregen door berekeningen, en de algoritmes die de constructie samenvatten)

Om de metingen van het Welfare Quality© protocol voor leghennen te transformeren in scores per principe, zijn vijf experts gevraagd de tabellen voor de multi-criteria analyse in te vullen. Deze vijf experts zijn: Claire Weeks (UK), Marion Staack (D), Ute Knierim (D), Ingrid de Jong (NL) en Thea van Niekerk (NL).

# Index

**Preface**

**Summary**

**Samenvatting**

**1 Introduction ..... 1**
1.1 Welfare Quality® assessment protocols ..... 1
1.2 Aim of the project ..... 1

**2 Calculation of criterion scores for laying hens ..... 2**
2.1 Criterion 1: Absence of prolongued hunger ..... 2
2.2 Criterion 2: Absence of prolongued thirst ..... 4
2.3 Criterion 3: Comfort around resting ..... 6
2.4 Criterion 4: Thermal comfort ..... 10
2.5 Criterion 5: Ease of Movement ..... 12
2.6 Criterion 6: Absence of injuries ..... 16
2.7 Criterion 7: Absence of disease ..... 26
2.8 Criterion 8: Absence of pain due to management procedures ..... 30
2.9 Criterion 9: Expression of social behaviours ..... 31
2.10 Criterion 10: Expression of other behaviours ..... 37
2.11 Criterion 11: Good human-animal relationship ..... 44
2.12 Criterion 12: Positive emotional state ..... 46

# 1 Introduction

## 1.1 Welfare Quality® assessment protocols

The European Welfare Quality® project developed standardized ways of assessing animal welfare for different categories of farm animals, e.g. broiler chickens and laying hens, sows, growing pigs, veal calves and dairy cattle. For laying hens, the measurements have been described in the Welfare Quality® assessment protocol for poultry (Welfare Quality®, 2009). One of the key characteristics of the Welfare Quality® assessment protocols is that it focuses more on animal based measures (such as injuries or behaviour) than on design or management criteria (such as flock size) (Blokhuis et al., 2010).

Welfare Quality® assessment protocols are based on the approach that welfare is a multidimensional concept, that comprises both physical as well as mental health. Within the Welfare Quality® project the same framework has been used to measure welfare of animals. Different measures of welfare, e.g. for laying hens the number of keel bone deformations, are integrated into a score for twelve independent welfare criteria. These criteria are integrated into four principle scores, and these are subsequently integrated into one overall score for a flock. In table 1 the twelve welfare criteria and the four principles are listed.

**Table 1** The principles and criteria that are the basis for the Welfare Quality® assessment protocols (Blokhuis et al., 2010).

| Welfare Quality® Principles | Welfare Quality® Criteria |
|---|---|
| Good feeding | 1 Absence of prolonged hunger |
| | 2 Absence of prolonged thirst |
| Good housing | 3 Comfort around resting |
| | 4 Thermal comfort |
| | 5 Ease of movement |
| Good health | 6 Absence of injuries |
| | 7 Absence of disease |
| | 8 Absence of pain induced by management procedures |
| Appropriate behaviour | 9 Expression of social behaviours |
| | 10 Expression of other behaviours |
| | 11 Good human-animal relationship |
| | 12 Positive emotional state |

The measures for the welfare assessment protocol for laying hens has been described in the Welfare Quality® assessment protocol for poultry (Welfare Quality®, 2009). The protocol describes measures indicative of laying hen welfare on-farm. Although the individual measures are described, the calculations for the integration into the overall flock score is missing. In this report we present this integration of measures into criteria score and the underlying calculations. The further integration from criteria scores into principle scores is independent of animal species and is described in other reports (Welfare Quality®, 2009).

## 1.2 Aim of the project

The aim of the project was to twofold: 1. to develop the integration of individual measures into a score per criterion and thus to be able to make an overall flock score; 2. to determine if there are possibilities to simplify the laying hen assessment protocol in order to reduce the time necessary to perform the measures. This will improve the practical applicability of the assessment protocol.
In this report the calculations for the integration of individual measures into a score per criterion are presented. In report 589 (Van Niekerk et al., 2012) the results of measurements on 122 flocks are presented as well as the results of the simplification of the protocol.

# 2 Calculation of criterion scores for laying hens

To evaluate the welfare of a flock, measures need to be transformed into criteria and principles. The transformation into principles and the classification of flocks is standard for all animal species. The transformation of measures into criteria is species-specific and is based on a multi criteria analysis with the help of experts specialized in the type of animal involved.

This report contains the description of the models and methodology to calculate the criterion-scores for the laying hens. The general construction is common as used for all other animal species:

- *Short indication of the measures used to construct the criterion* (for detailed description of the measures see the NEN-report: Welfare Quality®, 2009)
- *Type of mathematical construction used*
- *Parameters to be defined* (list of the parameters needed to assess the criterion according to the construction previously proposed)
- *Adjustments of the mathematical construction* (including the questions asked to experts, their answers and the exact adjustments made)
- *Calculation* (definition of the variables, from raw data or obtained by calculation, and algorithm summarising the construction)

To transform the measures of the Welfare Quality© protocol for laying hens into scores per principle 5 experts have been asked to fill in tables for the multi criteria analysis. The 5 experts consulted are: Claire Weeks (UK), Marion Staack (D), Ute Knierim (D), Ingrid de Jong (NL) and Thea van Niekerk (NL).

## 2.1 Criterion 1: Absence of prolongued hunger

| Measure | Nature | Level of observation |
|---|---|---|
| Feeder space | Cardinal | Flock |

For this criterion one measure is used: feeder space per bird (cm/hen). First experts have been asked to rate feeder space per bird on a scale of 0 - 100, where 100 is best. In table 2 the results are presented. From the expert consultation a spline function is calculated.

**Table 2** Results obtained from the experts asked

| Feeder space | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|
| (cm/hen) | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* | *Index* |
| 18 | 100 | 80 | 90 | 81 | 80 | 86.20 | 85.71 |
| 17 | 100 | 80 | 85 | 71 | 71 | 81.40 | 78.57 |
| 16 | 90 | 80 | 80 | 59 | 62 | 74.20 | 71.43 |
| 15 | 80 | 70 | 75 | 55 | 55 | 67.00 | 64.29 |
| 14 | 75 | 60 | 70 | 49 | 48 | 60.40 | 57.14 |
| 13 | 60 | 50 | 60 | 44 | 41 | 51.00 | 50.00 |
| 12 | 45 | 40 | 50 | 39 | 34 | 41.60 | 42.86 |
| 11 | 25 | 30 | 40 | 34 | 27 | 31.20 | 35.71 |
| 10 | 15 | 20 | 30 | 30 | 20 | 23.00 | 28.57 |
| 9 | 10 | 15 | 20 | 28 | 19 | 18.40 | 21.43 |
| 8 | 0 | 10 | 15 | 25 | 18 | 13.60 | 14.29 |

Feeder space per hen is calculated as:

$$\text{cm feeder space per bird} = p = \frac{\text{total cm feeder space}}{\text{total number of birds present}}$$

p is used to calculate index I: 100 * (p - 6) / (20 - 6)

with 6 being the lowest feeder space per bird resulting in a score 0

Then the index is transformed into a score with I-spline functions (figure 1) as follows:

When I≤50: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When I≥50: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for the criterion Absence of prolonged hunger:

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.5566630927 |
| c1 | 0.0148527208 |
| d1 | -0.0001170891 |
| a2 | -11.6815798795 |
| b2 | 1.2625858900 |
| c2 | 0.0005664485 |
| d2 | -0.0000201643 |

In figure 2 the distribution of criterion scores for criterion 1, Absence of prolongued hunger, is presented as calculated from the dataset of 122 flocks.

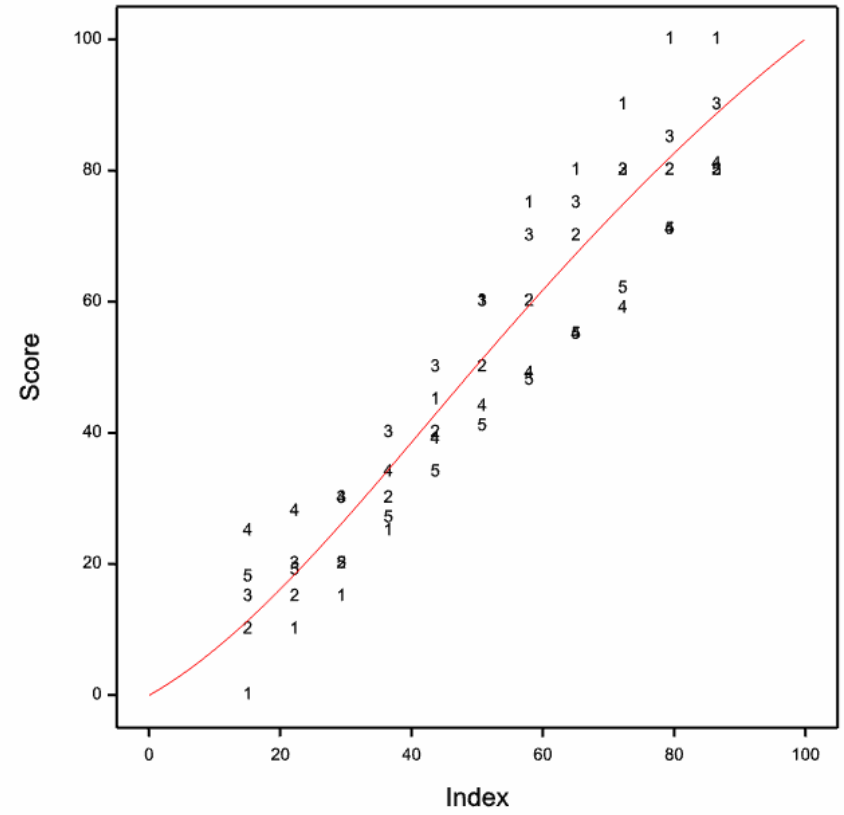


**Figure 1** Calculation of scores for criterion 1 (absence of prolongued hunger); spline with one interior knot at 50; (x axis, 100 * (Feeder space per bird - 6) / (20 - 6))

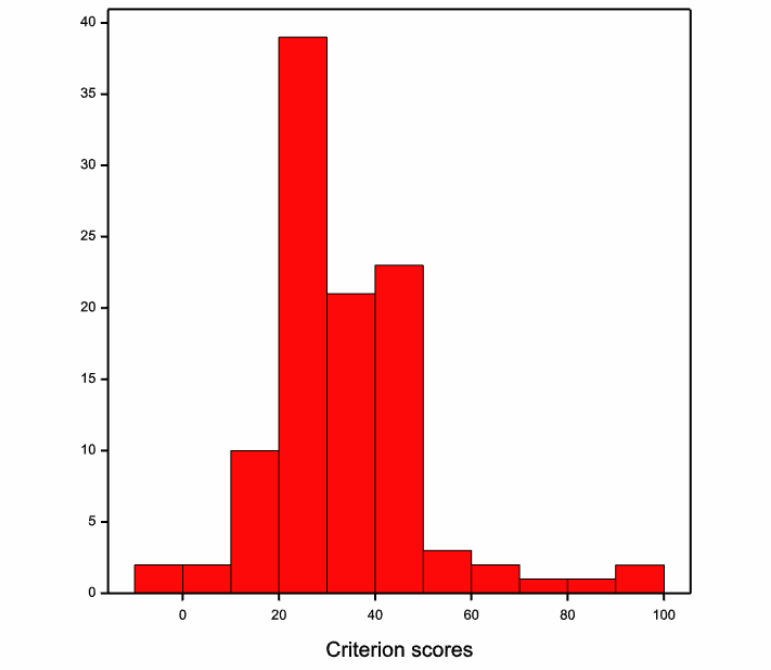


**Figure 2** Criterion 1 - Absence of prolonged hunger: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.2 Criterion 2: Absence of prolongued thirst

| Measure | Nature | Level of observation |
|---|---|---|
| Hens per nipple drinker | Cardinal | Flock |

For this criterion one measure is used: number of hens per nipple drinker. First experts have been asked to rate number of hens per nipple on a scale of 0 - 100, where 100 is best. In table 3 the results are presented. From the expert consultation a spline function is calculated.

**Table 3** Results obtained from the experts asked

| hens / nipple | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** | ***Index*** |
| 14 | 0 | 10 | 10 | 20 | 7 | 9.40 | 12.50 |
| 12 | 10 | 15 | 15 | 25 | 15 | 16.00 | 25.00 |
| 10 | 20 | 20 | 20 | 30 | 23 | 22.60 | 37.50 |
| 9 | 25 | 30 | 30 | 35 | 31 | 30.20 | 43.75 |
| 8 | 35 | 40 | 40 | 40 | 39 | 38.80 | 50.00 |
| 7 | 45 | 50 | 50 | 45 | 47 | 47.40 | 56.25 |
| 6 | 55 | 60 | 65 | 55 | 55 | 58.00 | 62.50 |
| 5 | 75 | 70 | 75 | 60 | 63 | 68.60 | 68.75 |
| 4 | 85 | 80 | 80 | 65 | 71 | 76.20 | 75.00 |
| 3 | 100 | 80 | 90 | 70 | 80 | 84.00 | 81.25 |
| 2 | 100 | 80 | 95 | 85 | 90 | 90.00 | 87.50 |

Hens per nipple is calculated as:

$$\text{number of hens per nipple drinker} = p = \frac{\text{total number of birds present}}{\text{total number nipple drinkers}}$$

p is used to calculate index I: 100 * (1- (p - 0) / (16 - 0))

with 16 being the highest number of hens resulting in a score 0

Then the index is transformed into a score with I-spline functions (figure 3) as follows:

When I≤60: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When I≥60: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for the criterion Absence of prolonged thirst:

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.4843455745 |
| c1 | 0.0032201509 |
| d1 | 0.0000548403 |
| a2 | 121.3433693264 |
| b2 | -5.5791743420 |
| c2 | 0.1041776816 |
| d2 | -0.0005051922 |

In figure 4 the distribution of criterion scores for criterion 2, Absence of prolonged thirst, is presented as calculated from the dataset of 122 flocks.

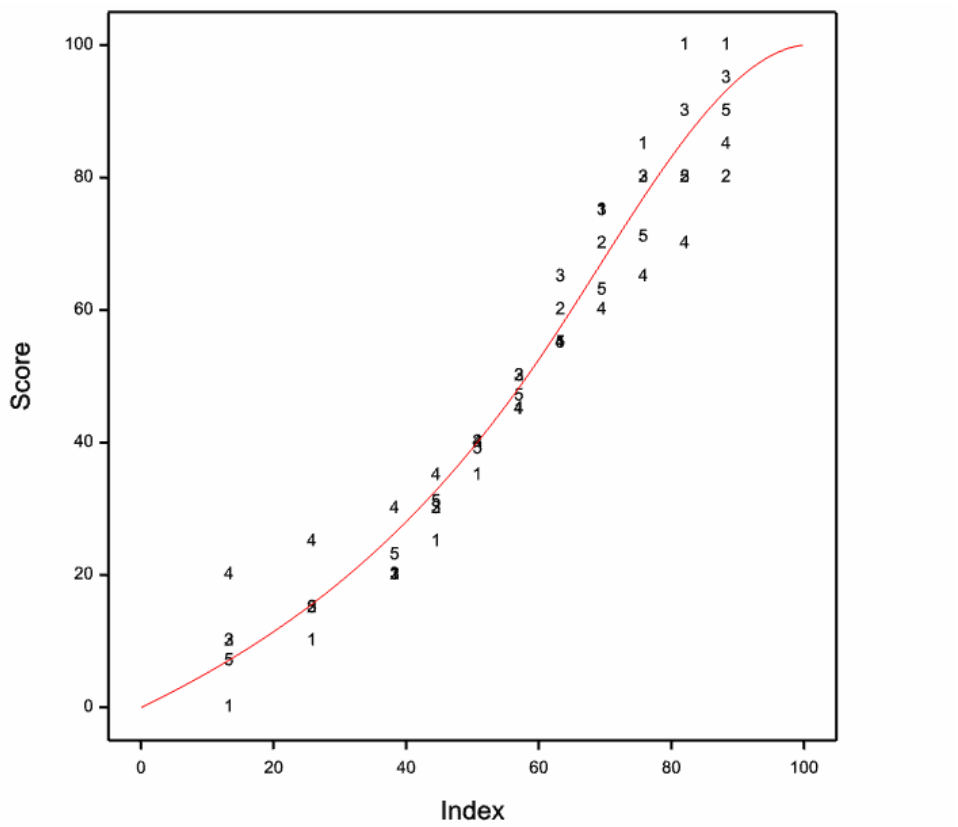


**Figure 3** Calculation of scores for criterion 2 (absence of prolonged thirst); spline with one interior knot at 60; (x axis, 100 * (1- (birds/nipple - 0) / (16 - 0)))

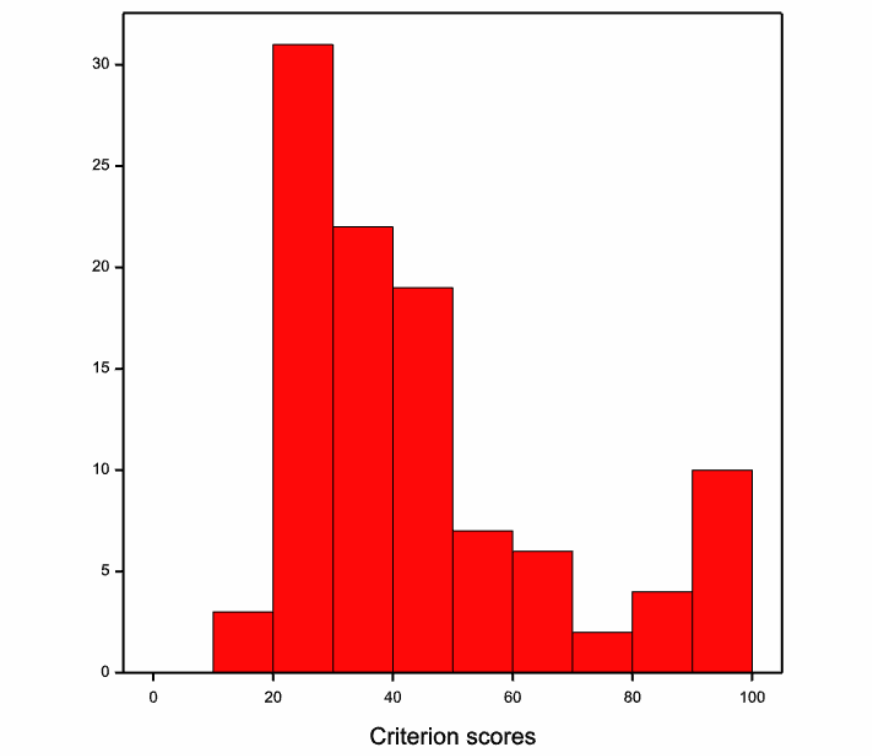


**Figure 4**

### 2.3 Criterion 3: Comfort around resting

| Measure | Nature | Level of observation |
|---|---|---|
| Perch positioning | Ordinal | Flock |
| Perch length | Cardinal | Individual |
| Evidence of red mites | Ordinal | Flock |
| Dust | Ordinal | Flock |

*First stage*

Perch position and available perch length per bird were combined and experts were asked to rank the various combinations by giving them a score (table 4). In case no perches were present the score was set on 0.
For evidence of red mites and dust, the experts had to assign a score to the different possible scores (Tables 5 and 6).

**Table 4** Results obtained from the experts asked to rank and score the various combinations of perch position and perch length.

| **Positioning perches (% perches in resting zone)** | **Perch length** | ***Score per expert*** | | | | | |
|---|---|---|---|---|---|---|---|
| | | ***Expert 1*** | ***Expert 2*** | ***Expert 3*** | ***Expert 4*** | ***Expert 5*** | ***Mean*** |
| >50% | >25 cm/hen | 100 | 100 | 90 | 70 | 100 | 92 |
| >50% | 21-25 cm/hen | 90 | 80 | 85 | 80 | 100 | 87 |
| >50% | 16-20 cm/hen | 40 | 50 | 50 | 60 | 80 | 56 |
| >50% | 10-15 cm/hen | 15 | 17 | 20 | 45 | 55 | 30 |
| >50% | <10 cm/hen | 0 | 10 | 15 | 30 | 19 | 15 |
| <50% | >25 cm/hen | 80 | 80 | 85 | 45 | 90 | 76 |
| <50% | 21-25 cm/hen | 70 | 60 | 80 | 50 | 90 | 70 |
| <50% | 16-20 cm/hen | 20 | 40 | 40 | 35 | 70 | 41 |
| <50% | 10-15 cm/hen | 5 | 13 | 15 | 25 | 40 | 20 |
| <50% | <10 cm/hen | 0 | 5 | 10 | 20 | 15 | 10 |

**Table 5** Results obtained from the experts asked to score the 3 possible scores for evidence of red mites

| **Evidence of red mites** | ***Score per expert*** | | | | | |
|---|---|---|---|---|---|---|
| | ***Expert 1*** | ***Expert 2*** | ***Expert 3*** | ***Expert 4*** | ***Expert 5*** | ***Mean*** |
| 0 | 100 | 100 | 100 | 100 | 100 | 100 |
| 1 | 20 | 50 | 30 | 15 | 40 | 40 |
| 2 | 0 | 15 | 5 | 0 | 15 | 15 |

**Table 5** Results obtained from the experts asked to score the 3 possible scores for dust

| **Dust** | ***Score per expert*** | | | | | |
|---|---|---|---|---|---|---|
| | ***Expert 1*** | ***Expert 2*** | ***Expert 3*** | ***Expert 4*** | ***Expert 5*** | ***Mean*** |
| 0 | 100 | 100 | 90 | 60 | 90 | 90 |
| 1 | 40 | 50 | 40 | 30 | 40 | 40 |
| 2 | 0 | 15 | 10 | 0 | 15 | 15 |

*Second stage*

To aggregate the three sub-scores into the criterion-scores the experts were asked to assign relative importance to the three measures, while having the possibility to focus on worst scores. To understand their way of reasoning, we presented to them a dataset composed of 13 virtual farms characterised by their three sub-scores (Table 6).

**Table 6** Results obtained from the experts asked to score13 imaginary farms for the criterion Comfort around resting

| Flock | Perching | Red mites | Dust | Score per expert Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 | Mean |
|---|---|---|---|---|---|---|---|---|---|
| Flock 1 | 25 | 50 | 75 | 25 | 40 | 40 | 40 | 25 | 34 |
| Flock 2 | 25 | 75 | 50 | 35 | 50 | 45 | 60 | 70 | 52 |
| Flock 3 | 40 | 50 | 60 | 40 | 50 | 55 | 45 | 60 | 50 |
| Flock 4 | 40 | 60 | 50 | 55 | 55 | 60 | 50 | 70 | 58 |
| Flock 5 | 50 | 25 | 75 | 30 | 40 | 30 | 45 | 25 | 34 |
| Flock 6 | 50 | 40 | 60 | 40 | 50 | 55 | 55 | 45 | 49 |
| Flock 7 | 50 | 50 | 50 | 55 | 50 | 50 | 50 | 55 | 52 |
| Flock 8 | 50 | 60 | 40 | 60 | 55 | 60 | 55 | 65 | 59 |
| Flock 9 | 50 | 75 | 25 | 65 | 50 | 35 | 55 | 60 | 53 |
| Flock 10 | 60 | 40 | 50 | 50 | 50 | 50 | 40 | 40 | 46 |
| Flock 11 | 60 | 50 | 40 | 60 | 55 | 50 | 50 | 50 | 53 |
| Flock 12 | 75 | 25 | 50 | 65 | 55 | 35 | 35 | 30 | 44 |
| Flock 13 | 75 | 50 | 25 | 75 | 60 | 40 | 50 | 45 | 54 |

The three partial scores, 1=Perching, 2=Red mites, 3=Dust, are combined to a score for criterion 3 using a Choquet integral with the following parameters:

| Parameter | Estimate | Parameter | Estimate |
|---|---|---|---|
| μ 1 | 0.3925 | μ 12 | 0.7320 |
| μ 2 | 0.5503 | μ 13 | 0.3928 |
| μ 3 | 0 | μ 23 | 0.5504 |

In figure 5 the distribution of criterion scores for criterion 3, Comfort around resting, is presented as calculated from the dataset of 122 flocks.

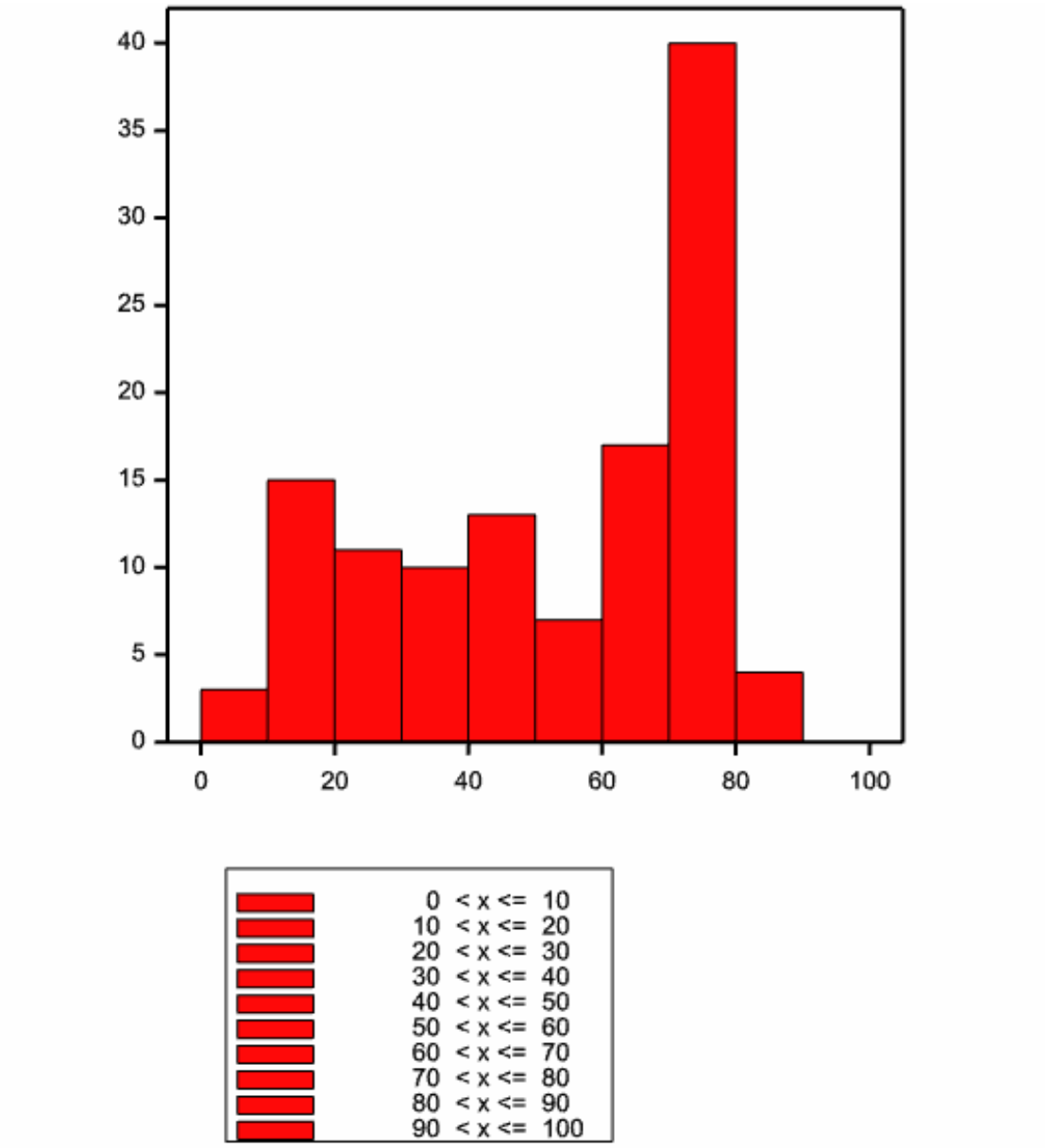


**Figure 5** Criterion 3 - Comfort around resting: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.4 Criterion 4: Thermal comfort

| Measure | Nature | Level of observation |
|---|---|---|
| Panting | Cardinal | Flock |
| Huddling | Cardinal | Flock |

The two measures (panting and huddling) are complementary: birds pant when they are too warm and huddle when they are too cold. However, there may be large variation in a henhouse, causing a part of the flock to be cold and another part of the flock to be hot. So both measures can theoretically be assessed on the same day in the same flock.

*First stage*

The measures are assessed at flock level as a percentage of the birds. The experts have been asked to evaluate 7 imaginary flocks with varying levels of percentages birds panting and huddling. The results are presented in tables 7 and 8.

**Table 7** Results from experts' evaluation of the percentage of birds panting

| Flock | % Panting | Score per expert | | | | | |
|---|---|---|---|---|---|---|---|
| | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* |
| Flock 1 | 0 | 100 | 100 | 100 | 40 | 100 | 70 |
| Flock 2 | 5 | 80 | 40 | 30 | 60 | 50 | 63 |
| Flock 3 | 10 | 65 | 20 | 20 | 45 | 25 | 35 |
| Flock 4 | 20 | 45 | 18 | 15 | 50 | 19 | 35 |
| Flock 5 | 30 | 15 | 15 | 10 | 45 | 15 | 30 |
| Flock 6 | 40 | 10 | 10 | 10 | 55 | 10 | 33 |
| Flock 7 | 50 | 0 | 5 | 5 | 50 | 5 | 28 |

**Table 8** Results from experts' evaluation of the percentage of birds huddling

| Flock | % Panting | Score per expert | | | | | |
|---|---|---|---|---|---|---|---|
| | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* |
| Flock 1 | 0 | 100 | 100 | 100 | 80 | 100 | 96 |
| Flock 2 | 5 | 80 | 40 | 35 | 65 | 25 | 49 |
| Flock 3 | 10 | 65 | 20 | 25 | 55 | 19 | 37 |
| Flock 4 | 20 | 50 | 18 | 20 | 50 | 5 | 29 |
| Flock 5 | 30 | 40 | 15 | 15 | 45 | 0 | 23 |
| Flock 6 | 40 | 20 | 10 | 10 | 40 | 0 | 16 |
| Flock 7 | 50 | 15 | 5 | 5 | 30 | 0 | 11 |

*Second stage*

To aggregate the two measures into the criterion-scores, the minimum of both measures is used.

In figure 6 the distribution of criterion scores for criterion 4, Thermal comfort, is presented as calculated from the dataset of 122 flocks.

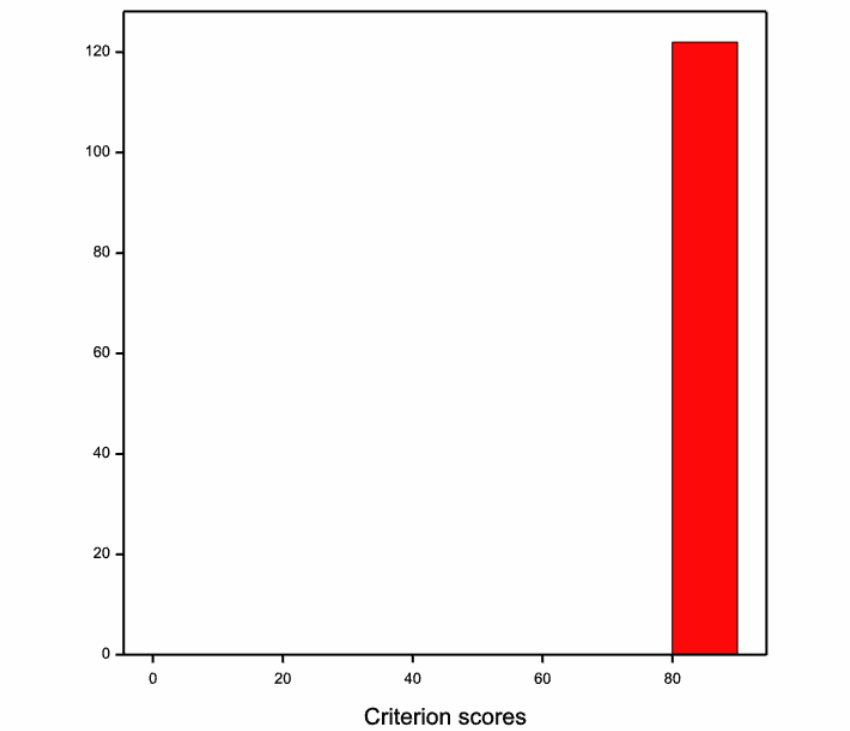


**Figure 6** Criterion 4 - Thermal comfort: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.5 Criterion 5: Ease of Movement

| Measure | Nature | Level of observation |
|---|---|---|
| Stocking density (cm2/bird) | Cardinal | Flock |
| Percentage perforated floor | Cardinal | Flock |

Although stocking density is expressed on individual level, it is assessed at flock level by dividing the total amount of available space by the number of hens present.
For identifying the percentage of perforated floor the total amount of perforated floor

*First stage*

The experts have been asked to evaluate 9 imaginary flocks with varying stocking densities. The results are presented in table 10.

**Table 10** Results from experts' evaluation of stocking density ($cm^2$/hen)

| cm2/hen | Score per expert | | | | | |
|---|---|---|---|---|---|---|
| | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* |
| 400 | 0 | 0 | 0 | 18 | 0 | 3.60 |
| 600 | 0 | 5 | 5 | 25 | 15 | 10.00 |
| 800 | 10 | 10 | 10 | 35 | 20 | 17.00 |
| 1000 | 20 | 15 | 15 | 40 | 40 | 26.00 |
| 1200 | 30 | 30 | 40 | 50 | 60 | 42.00 |
| 1400 | 45 | 40 | 60 | 55 | 80 | 56.00 |
| 1600 | 55 | 60 | 75 | 60 | 90 | 68.00 |
| 1800 | 75 | 80 | 80 | 75 | 100 | 82.00 |
| 2000 | 80 | 90 | 90 | 85 | 100 | 89.00 |

Stocking density is calculated as:

$$cm^2 \text{ per hen} = p = \frac{\text{total available space } (cm^2)}{\text{total number of birds present}}$$

p is used to calculate index I: 100 * (1 - (2000 - p) / (2000 - 400))
with 400 being the lowest space per hens resulting in a score 0

Then the index is transformed into a score with I-spline functions (figure 7) as follows:

When $I \leq 30$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $I \geq 30$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for the stocking density ($cm^2$/hen):

| COEF | COEF1 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.7164389632 |
| c1 | -0.0025580518 |
| d1 | 0.0000955685 |
| a2 | 3.2771454904 |
| b2 | 0.3887244142 |
| c2 | 0.0083657665 |
| d2 | -0.0000258073 |

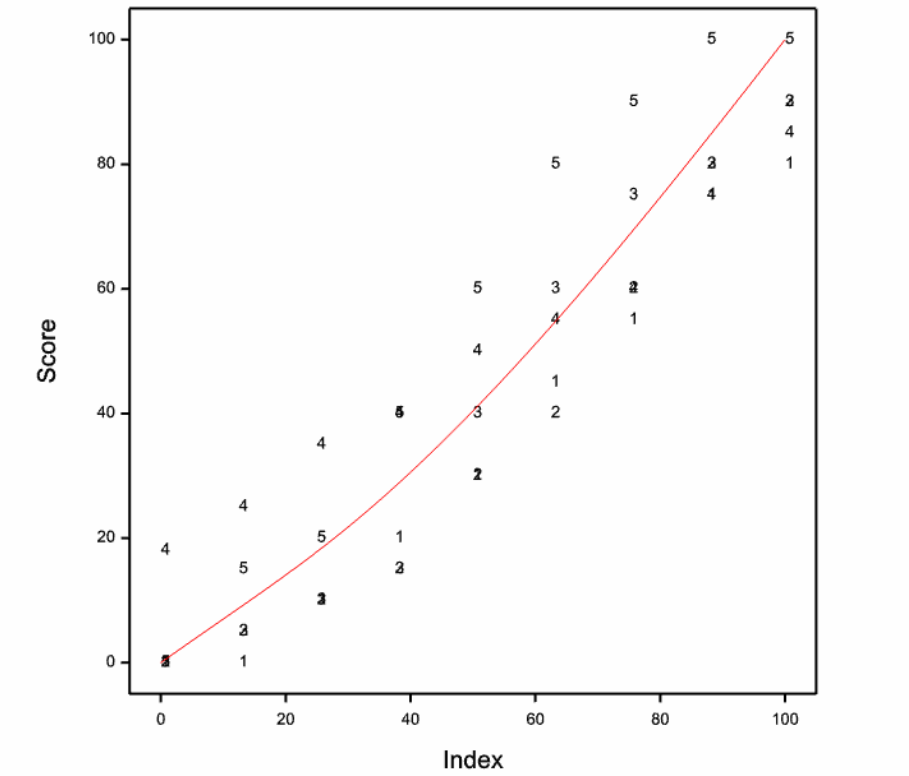


**Figure 7** Calculation of scores for stocking density ($cm^2$/hen); spline with one interior knot at 30; (x axis, 100 * (1 - (2000 - $cm^2$/hen) / (2000 - 400)))

The experts have been asked to evaluate 11 imaginary flocks with varying percentages of perforated floor. The results are presented in table 11.

**Table 11** Results from experts' evaluation of % perforated floor

| % Perforated floor | *Score per expert* | | | | | |
|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** |
| 0 | 100 | 10 | 100 | 55 | 100 | 73 |
| 10 | 80 | 10 | 90 | 60 | 100 | 68 |
| 20 | 75 | 10 | 85 | 60 | 100 | 66 |
| 30 | 55 | 50 | 80 | 70 | 90 | 69 |
| 40 | 40 | 50 | 70 | 50 | 70 | 56 |
| 50 | 20 | 55 | 60 | 45 | 50 | 46 |
| 60 | 15 | 35 | 40 | 40 | 30 | 32 |
| 70 | 10 | 20 | 15 | 35 | 15 | 19 |
| 80 | 0 | 10 | 10 | 30 | 5 | 11 |
| 90 | 0 | 0 | 5 | 25 | 0 | 6 |
| 100 | 0 | 0 | 0 | 20 | 0 | 4 |

Percentage perforated floor is calculated as:

$$\% \text{ perforated floor} = p = \frac{\text{total available area perforated floor } (cm^2)}{\text{total available space } (cm^2)}$$

p is used to calculate index I: 100 - p

Then the index is transformed into a score with I-spline functions (figure 8) as follows:

When I≤70: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$

When I≥70: I = (a2 +b2*I +c2*$I^2$ + d2*$I^3$)

The following coefficients are used for the percentage perforated floor:

| COEF | COEF2 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.1568389736 |
| c1 | 0.0251972531 |
| d1 | -0.0002114463 |
| a2 | -556.2933684725 |
| b2 | 23.9979833768 |
| c2 | -0.3153905246 |
| d2 | 0.0014104003 |

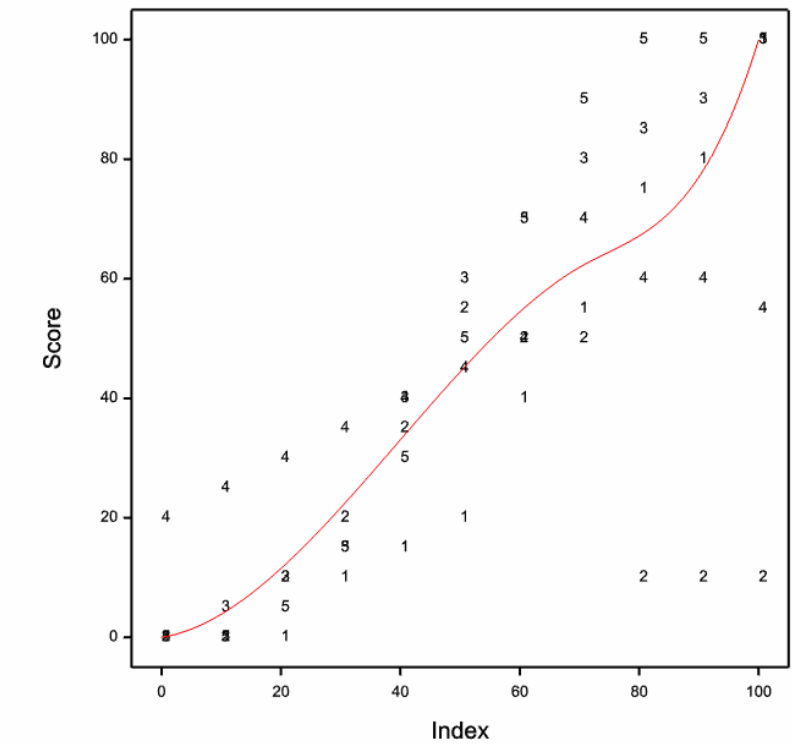


**Figure 8** Calculation of scores for percentage perforated floor; spline with one interior knot at 70; (x axis, 100 - % perforated floor)

*Second stage*

To aggregate the three sub-scores into the criterion-scores the experts were asked to assign relative importance to the three measures, while having the possibility to focus on worst scores. To understand their way of reasoning, we presented to them a dataset composed of 5 virtual farms characterised by their three sub-scores (Table 12).

**Table 12** Results obtained from the experts asked to score 13 imaginary farms for the criterion Ease of movement

| Flock | Score for stocking density | Score for perforated floor | *Score per expert* | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** |
| Flock 1 | 25 | 75 | 35 | 30 | 30 | 25 | 30 | 34 |
| Flock 2 | 40 | 60 | 45 | 45 | 45 | 40 | 45 | 52 |
| Flock 3 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Flock 4 | 60 | 40 | 55 | 50 | 50 | 55 | 50 | 58 |
| Flock 5 | 75 | 25 | 65 | 55 | 35 | 70 | 45 | 34 |

The two partial scores, 1=stocking density, 2=perforated floor, are combined to a score for criterion 5 using a Choquet integral with the following parameters:

| Parameter | estimate |
|---|---|
| μ 1 | 0.5828 |
| μ 2 | 0.1138 |

In figure 13 the distribution of criterion scores for criterion 5, Ease of movement, is presented as calculated from the dataset of 122 flocks.

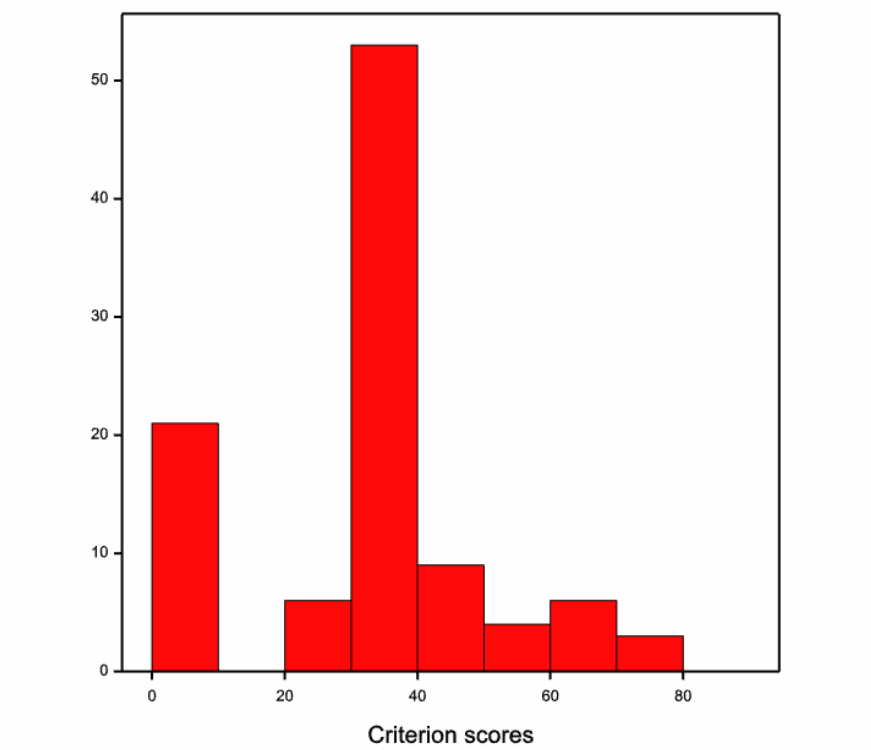


**Figure 13** Criterion 5 - Ease of movement: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.6 Criterion 6: Absence of injuries

| Measure | Nature | Level of observation |
|---|---|---|
| Keel bone damage | Ordinal | Flock |
| Skin lesions | Ordinal | Flock |
| Food pad lesions | Ordinal | Flock |
| Toe damage | Ordinal | Flock |

Keel bone damage is scored on a 2-point scale on 100 birds per flock. Skin lesions and food pad lesions are scored on the same 100 birds on a 3-point scale. Although these are cardinal measures, they are expressed as ordinal measure by expressing then as percentage of the flock with score 0, score 1 or score 2.
Toe damage is scored on the 100 individually examined birds and is expressed as percentage hens with toe damage.

*First stage*

The experts have been asked to evaluate 18 imaginary flocks with varying levels of percentages birds with keel bone damage. The results are presented in table 13.

**Table 13:** Results from experts' evaluation of Keel bone damage (% hens with keel bone damage)

| % hens with keel bone damage | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** | ***Index*** |
| 0 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 2 | 100 | 80 | 90 | 22 | 90 | 76.4 | 98 |
| 5 | 100 | 70 | 80 | 20 | 75 | 69.0 | 95 |
| 10 | 80 | 50 | 70 | 18 | 50 | 53.6 | 90 |
| 15 | 75 | 40 | 55 | 16 | 19 | 41.0 | 85 |
| 20 | 65 | 25 | 50 | 14 | 15 | 33.8 | 80 |
| 25 | 55 | 18 | 40 | 12 | 10 | 27.0 | 75 |
| 30 | 50 | 15 | 35 | 11 | 5 | 23.2 | 70 |
| 40 | 40 | 10 | 20 | 10 | 3 | 16.6 | 60 |
| 50 | 20 | 9 | 15 | 9 | 0 | 10.6 | 50 |
| 60 | 15 | 8 | 10 | 8 | 0 | 8.2 | 40 |
| 70 | 15 | 7 | 5 | 7 | 0 | 6.8 | 30 |
| 75 | 10 | 6 | 0 | 6 | 0 | 4.4 | 25 |
| 80 | 10 | 5 | 0 | 5 | 0 | 4.0 | 20 |
| 90 | 0 | 0 | 0 | 4 | 0 | 0.8 | 10 |
| 95 | 0 | 0 | 0 | 3 | 0 | 0.6 | 5 |
| 98 | 0 | 0 | 0 | 2 | 0 | 0.4 | 2 |
| 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Keel bone damage is calculated as:

$$\% \text{ hens with keel bone damage} = p = \frac{\text{number of birds with keel bone score 2}}{\text{total number of birds examined}}$$

p is used to calculate index I: 100 - p

Then the index is transformed into a score with I-spline functions (figure 14) as follows:

When $I \leq 80$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $I \geq 80$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for keel bone damage:

| COEF | COEF1 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.1898672160 |
| c1 | -0.0023733402 |
| d1 | 0.0000651008 |
| a2 | -2537.3344655810 |
| b2 | 95.3399111060 |
| c2 | -1.1917489052 |
| d2 | 0.0050208324 |

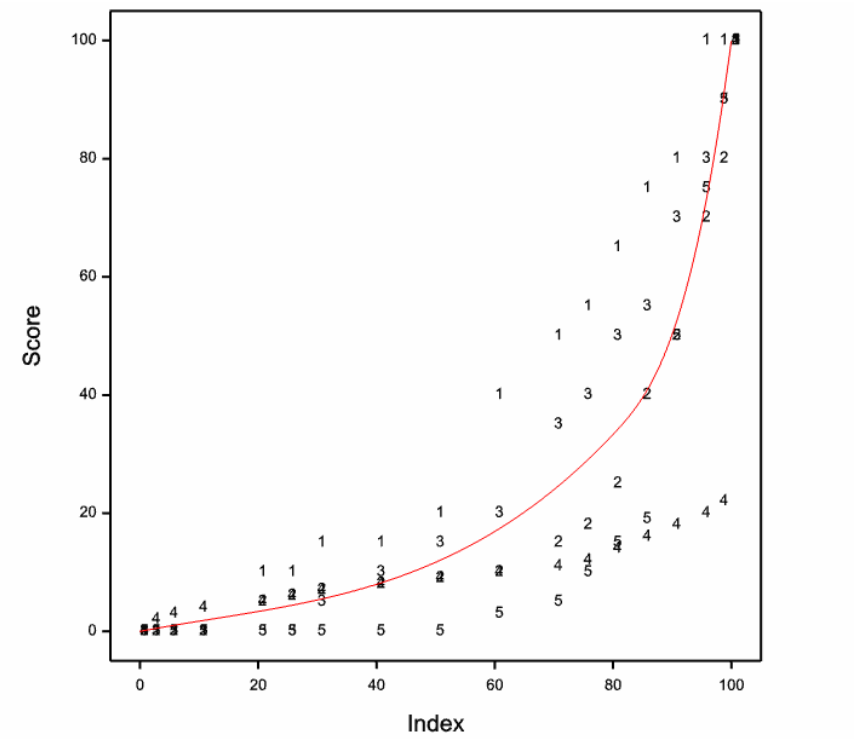


**Figure 14** Calculation of scores for keel bone damage; spline with one interior knot at 80; (x axis, 100 - % keel bone damage)

In figure 15 the distribution of percentage keel bone damage is presented as calculated from the dataset of 122 flocks.

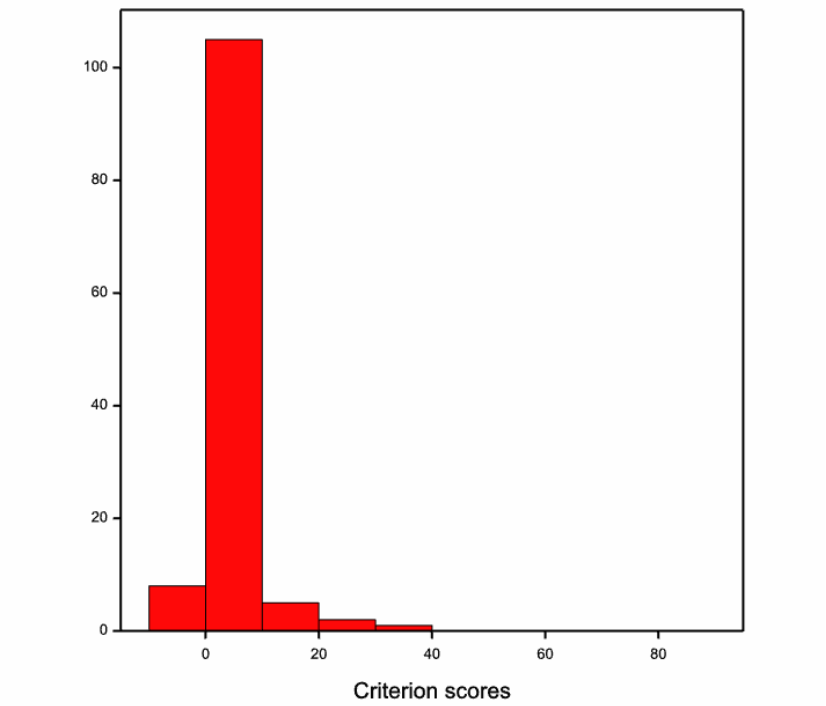


**Figure 15** Percentage of keel bone damage: Distribution of scores as calculated from the dataset of 122 flocks

The experts have been asked to evaluate 18 imaginary flocks with varying levels of percentages birds with skin lesions. The results are presented in table 14

**Table 14** Results from experts' evaluation of Skin lesions (per score classes *)

| Flock | Score 0 | Score 1 | Score 2 | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* | *Index* |
| Flock 1 | 100 | 0 | 0 | 100 | 100 | 100 | 100 | 100 | 100.0 | 100.0 |
| Flock 2 | 95 | 5 | 0 | 90 | 90 | 90 | 85 | 80 | 87.0 | 97.5 |
| Flock 3 | 95 | 0 | 5 | 75 | 50 | 50 | 50 | 50 | 59.0 | 95.0 |
| Flock 4 | 90 | 10 | 0 | 80 | 40 | 50 | 70 | 55 | 55.0 | 95.0 |
| Flock 5 | 90 | 5 | 5 | 75 | 35 | 40 | 45 | 20 | 43.0 | 92.5 |
| Flock 6 | 80 | 20 | 0 | 65 | 20 | 30 | 35 | 20 | 34.0 | 90.0 |
| Flock 7 | 80 | 15 | 5 | 60 | 15 | 20 | 30 | 15 | 28.0 | 87.5 |
| Flock 8 | 80 | 10 | 10 | 60 | 10 | 15 | 20 | 0 | 26.0 | 87.5 |
| Flock 9 | 75 | 25 | 0 | 55 | 20 | 25 | 25 | 5 | 21.0 | 85.0 |
| Flock 10 | 75 | 20 | 5 | 50 | 15 | 15 | 21 | 0 | 20.2 | 85.0 |
| Flock 11 | 75 | 15 | 10 | 35 | 10 | 10 | 15 | 0 | 19.8 | 85.0 |
| Flock 12 | 70 | 30 | 0 | 45 | 15 | 15 | 24 | 0 | 15.0 | 82.5 |
| Flock 13 | 70 | 25 | 5 | 35 | 10 | 10 | 20 | 0 | 14.0 | 82.5 |
| Flock 14 | 70 | 20 | 10 | 25 | 5 | 5 | 12 | 0 | 9.8 | 80.0 |
| Flock 15 | 60 | 40 | 0 | 15 | 4 | 10 | 20 | 0 | 9.4 | 80.0 |
| Flock 16 | 60 | 35 | 5 | 10 | 3 | 5 | 22 | 0 | 8.0 | 77.5 |
| Flock 17 | 60 | 30 | 10 | 10 | 2 | 5 | 11 | 0 | 5.6 | 75.0 |
| Flock 18 | 50 | 40 | 10 | 5 | 0 | 5 | 10 | 0 | 4.0 | 70.0 |

** Score 0= hens without skin lesions; score 1=hens with minor lesions; score 2=hens with more/larger lesions*

Skin lesions are calculated as:

$$\% \text{ hens with score 1} = p1 = \frac{\text{number of hens with score 1}}{\text{total number of birds examined}}$$

$$\% \text{ hens with score 2} = p2 = \frac{\text{number of hens with score 2}}{\text{total number of birds examined}}$$

p1 and p2 are used to calculate index I: 100 - ((1*p1 + 2*p2)/2)

Then the index is transformed into a score with I-spline functions (figure 16) as follows:

When $I \leq 72$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $72 \leq I \leq 95$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$
When $I \geq 95$: $I = (a3 + b3*I + c3*I^2 + d3*I^3)$

The following coefficients are used for skin lesions:

| COEF | COEF2 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.0000004965 |
| c1 | -0.0000049820 |
| d1 | 0.0000166068 |
| a2 | -1439.4936246929 |
| b2 | 59.9789015160 |
| c2 | -0.8330452738 |
| d2 | 0.0038732748 |
| a3 | 6918.6018461844 |
| b3 | -203.9609565524 |
| c3 | 1.9452690330 |
| d3 | -0.0058751965 |

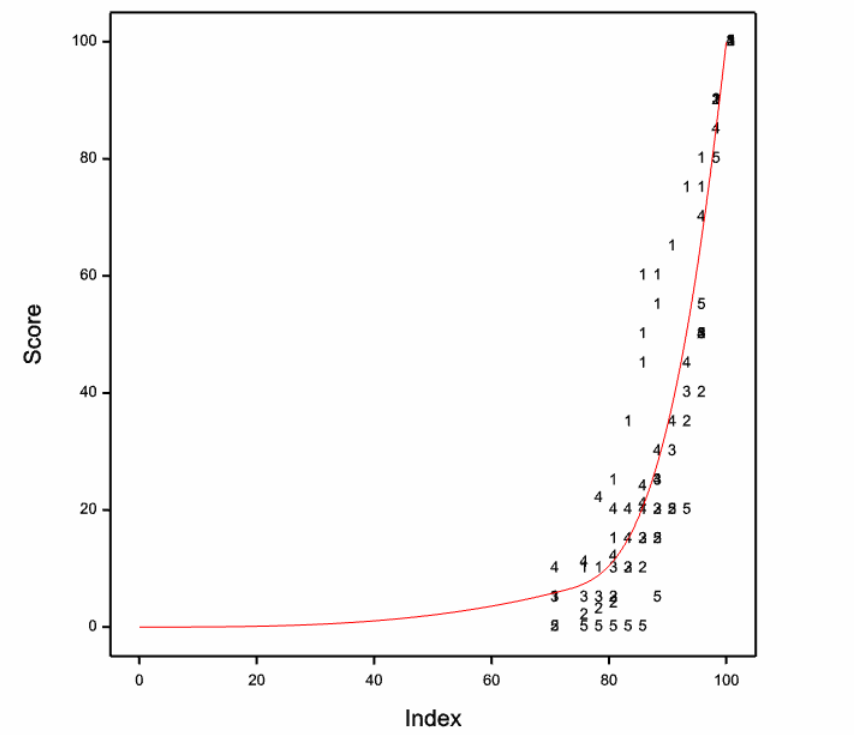


**Figure 16** Calculation of scores for skin lesions; spline with two interior knots at 72 and 95; (x axis, 100 - ((1*score1 + 2*score2)/2))

In figure 17 the distribution of skin lesion score is presented as calculated from the dataset of 122 flocks.

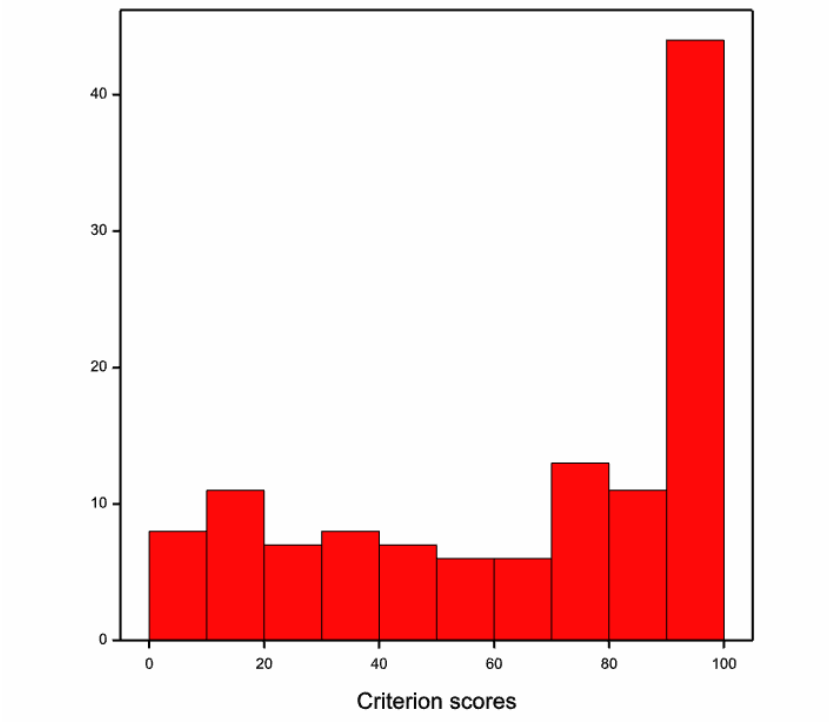


**Figure 17** Skin lesions: Distribution of scores as calculated from the dataset of 122 flocks

The experts have been asked to evaluate 18 imaginary flocks with varying levels of percentages birds with foot pad lesions. The results are presented in table 15.

**Table 15** Results from experts' evaluation of Foot pad lesions *

| Flock | Score 0 | Score 1 | Score 2 | Score per expert | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* |
| Flock 1 | 100 | 0 | 0 | 100 | 100 | 100 | 100 | 100 | 100.0 |
| Flock 2 | 90 | 10 | 0 | 80 | 80 | 70 | 80 | 75 | 75.0 |
| Flock 3 | 80 | 20 | 0 | 75 | 50 | 60 | 45 | 70 | 70.0 |
| Flock 4 | 80 | 15 | 5 | 45 | 40 | 50 | 25 | 19 | 19.0 |
| Flock 5 | 75 | 25 | 0 | 65 | 45 | 50 | 40 | 40 | 40.0 |
| Flock 6 | 75 | 20 | 5 | 45 | 35 | 40 | 22 | 18 | 18.0 |
| Flock 7 | 60 | 40 | 0 | 50 | 20 | 35 | 25 | 15 | 15.0 |
| Flock 8 | 60 | 35 | 5 | 40 | 15 | 20 | 21 | 10 | 10.0 |
| Flock 9 | 50 | 50 | 0 | 40 | 20 | 15 | 22 | 0 | 0.0 |
| Flock 10 | 50 | 45 | 5 | 30 | 15 | 10 | 20 | 0 | 0.0 |
| Flock 11 | 50 | 40 | 10 | 20 | 10 | 5 | 18 | 0 | 0.0 |
| Flock 12 | 45 | 50 | 5 | 30 | 10 | 10 | 20 | 0 | 14.0 |
| Flock 13 | 25 | 75 | 0 | 20 | 5 | 10 | 20 | 0 | 11.0 |
| Flock 14 | 25 | 70 | 5 | 15 | 0 | 5 | 19 | 0 | 7.8 |
| Flock 15 | 25 | 65 | 10 | 10 | 0 | 5 | 12 | 0 | 5.4 |
| Flock 16 | 0 | 100 | 0 | 0 | 0 | 10 | 18 | 0 | 5.6 |
| Flock 17 | 0 | 95 | 5 | 0 | 0 | 5 | 15 | 0 | 4.0 |
| Flock 18 | 0 | 90 | 10 | 0 | 0 | 5 | 10 | 0 | 3.0 |

** Score 0= hens without skin lesions; score 1=hens with minor lesions; score 2=hens with more/larger lesions*

Food pad lesions are calculated as:

$$\%\ \text{hens with score 1} = p1 = \frac{\text{number of hens with score 1}}{\text{total number of birds examined}}$$

$$\%\ \text{hens with score 2} = p2 = \frac{\text{number of hens with score 2}}{\text{total number of birds examined}}$$

p1 and p2 are used to calculate index I: 100 - ((2*p1 + 7*p2)/7)

Then the index is transformed into a score with I-spline functions (figure 18) as follows:

When I≤72: I = (a1 + b1*I + c1*$I^2$ + d1*$I^3$)
When 72≤I≤95: I = (a2 +b2*I +c2*$I^2$ + d2*$I^3$)
When I≥95: I = (a3 +b3*I +c3*$I^2$ + d3*$I^3$)

The following coefficients are used for food pad lesions:

| COEF | COEF2 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.0000005325 |
| c1 | -0.0000053436 |
| d1 | 0.0000178121 |
| a2 | -1431.0301007264 |
| b2 | 59.6262547598 |
| c2 | -0.8281477638 |
| d2 | 0.0038518047 |
| a3 | 11755.3030986256 |
| b3 | -356.7842710343 |
| c3 | 3.5551209670 |
| d3 | -0.0115280857 |

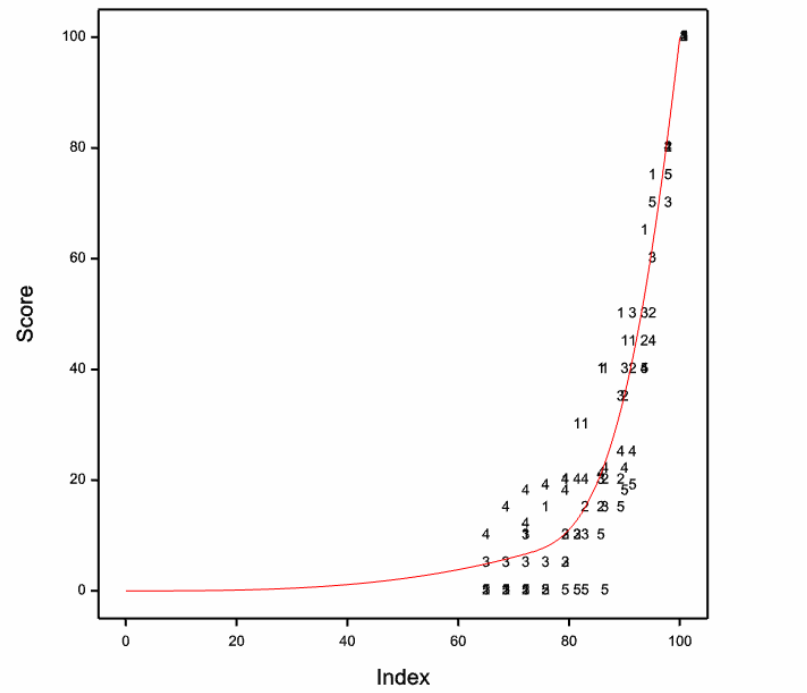


**Figure 18** Calculation of scores for food pad lesions; spline with one interior knot at 80; (x axis, 100 - ((2*score1 + 7*score2)/7))

In figure 19 the distribution of percentage food pad lesions is presented as calculated from the dataset of 122 flocks.

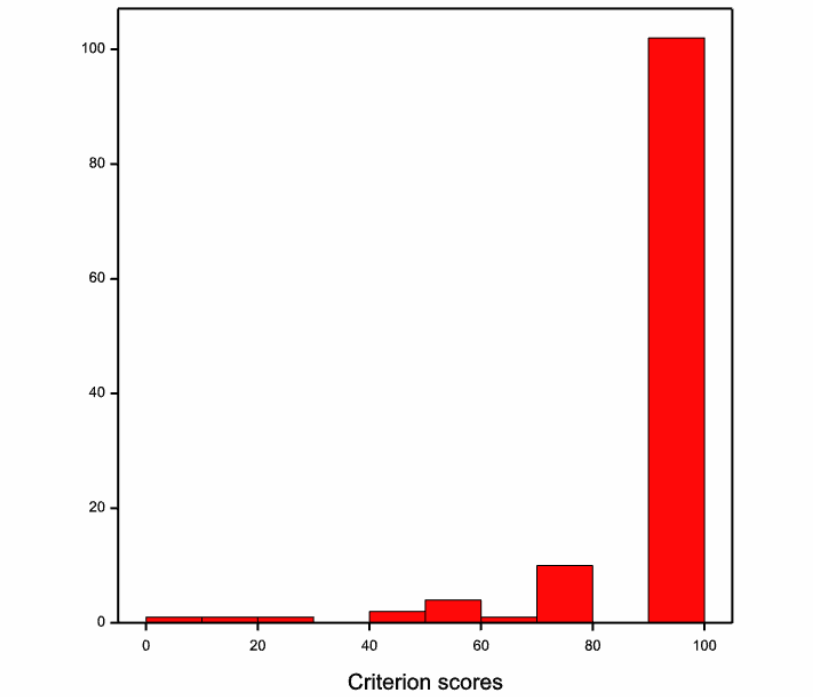


**Figure 19** Food pad lesions; Distribution of scores as calculated from the dataset of 122 flocks

The experts have been asked to evaluate 16 imaginary flocks with varying levels of percentages birds with toe damage. The results are presented in table 16.

**Table 16** Results from experts' evaluation of Toe damage (% hens with toe damage)

| % hens with toe damage | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** | ***Index*** |
| 0 | 100 | 100 | 100 | 100 | 100 | 100.00 | 100.00 |
| 1 | 80 | 90 | 80 | 35 | 20 | 99.00 | 61.00 |
| 2 | 75 | 80 | 60 | 30 | 10 | 98.00 | 51.00 |
| 3 | 55 | 60 | 50 | 25 | 5 | 97.00 | 39.00 |
| 4 | 50 | 50 | 40 | 20 | 0 | 96.00 | 32.00 |
| 5 | 40 | 40 | 35 | 15 | 0 | 95.00 | 26.00 |
| 6 | 20 | 30 | 30 | 12 | 0 | 94.00 | 18.40 |
| 7 | 15 | 20 | 25 | 9 | 0 | 93.00 | 13.80 |
| 8 | 15 | 18 | 20 | 7 | 0 | 92.00 | 12.00 |
| 9 | 15 | 16 | 15 | 5 | 0 | 91.00 | 10.20 |
| 10 | 15 | 14 | 10 | 5 | 0 | 90.00 | 8.80 |
| 20 | 10 | 12 | 5 | 0 | 0 | 80.00 | 5.40 |
| 40 | 0 | 10 | 5 | 0 | 0 | 60.00 | 3.00 |
| 60 | 0 | 5 | 5 | 0 | 0 | 40.00 | 2.00 |
| 80 | 0 | 0 | 5 | 0 | 0 | 20.00 | 1.00 |
| 100 | 0 | 0 | 0 | 0 | 0 | 0.00 | 0.00 |

Toe damage is calculated as:

$$\text{percentage toe damage} = p = \frac{\text{number of hens with toe damage}}{\text{number of hens examined}}$$

p is used to calculate index I: 100 - p

Then the index is transformed into a score with I-spline functions (figure 20) as follows:

When $I \leq 50$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $50 \leq I \leq 90$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$
When $I \geq 90$: $I = (a3 + b3*I + c3*I^2 + d3*I^3)$

The following coefficients are used for toe damage:

| COEF | COEF4 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.0473985825 |
| c1 | -0.0009479717 |
| d1 | 0.0000195478 |
| a2 | 2.7904245378 |
| b2 | -0.1200268897 |
| c2 | 0.0024005378 |
| d2 | -0.0000027756 |
| a3 | -64136.4092672520 |
| b3 | 2137.8533028245 |
| c3 | -23.7528587522 |
| d3 | 0.0879796665 |

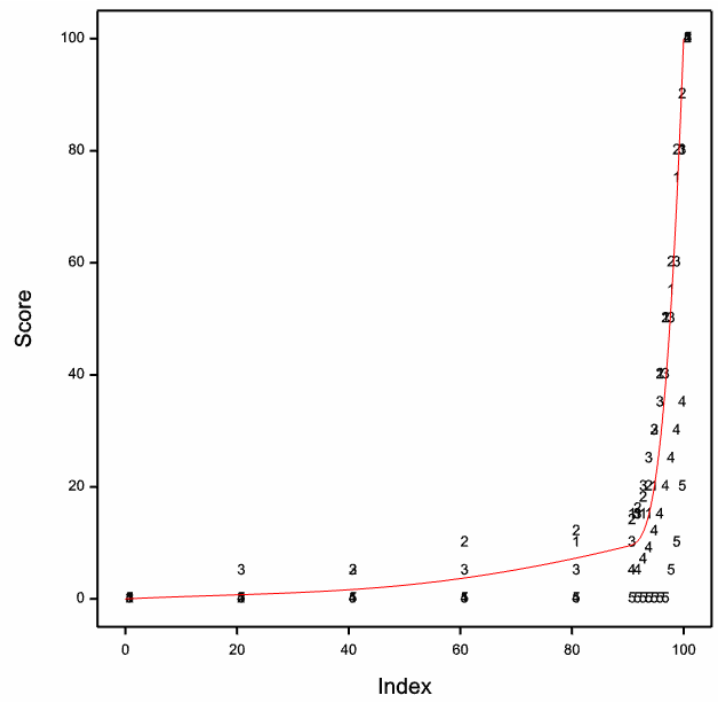


**Figure 20** Calculation of scores for percentage toe damage; spline with two interior knots at 50 and 90; (x axis, 100 - % toe damage)

In figure 21 the distribution of percentage toe damage is presented as calculated from the dataset of 122 flocks.

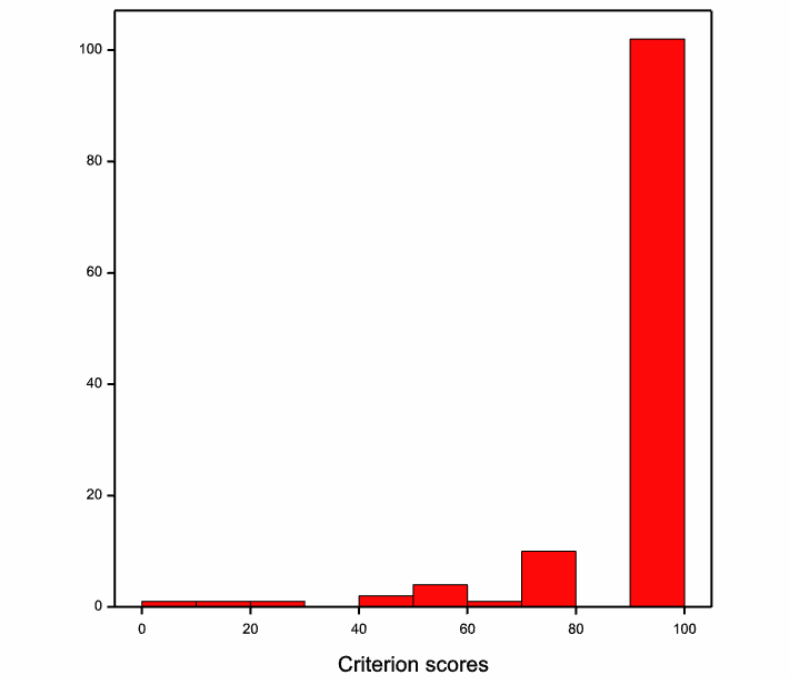


**Figure 21** Percentage of toe damage: Distribution of scores as calculated from the dataset of 122 flocks

*Second stage*

The four partial scores, 1=Keel bone damage, 2=Skin lesions, 3=Food pad lesions, 4=Toe damage are combined to a score for criterion 6 using a Choquet integral with the following parameters:

| Parameter | Estimate | Parameter | Estimate | Parameter | Estimate |
|---|---|---|---|---|---|
| μ 1 | 0.2752 | μ 12 | 0.3333 | μ 123 | 0.6304 |
| μ 2 | 0.1885 | μ 13 | 0.2752 | μ 124 | 0.3570 |
| μ 3 | 0.1385 | μ 14 | 0.2752 | μ 134 | 0.3870 |
| μ 4 | 0.1185 | μ 23 | 0.1885 | μ 234 | 0.5004 |
| | | μ 24 | 0.1885 | | |
| | | μ 34 | 0.1385 | | |

In figure 22 the distribution of criterion scores for criterion 6, Absence of injuries, is presented as calculated from the dataset of 122 flocks.

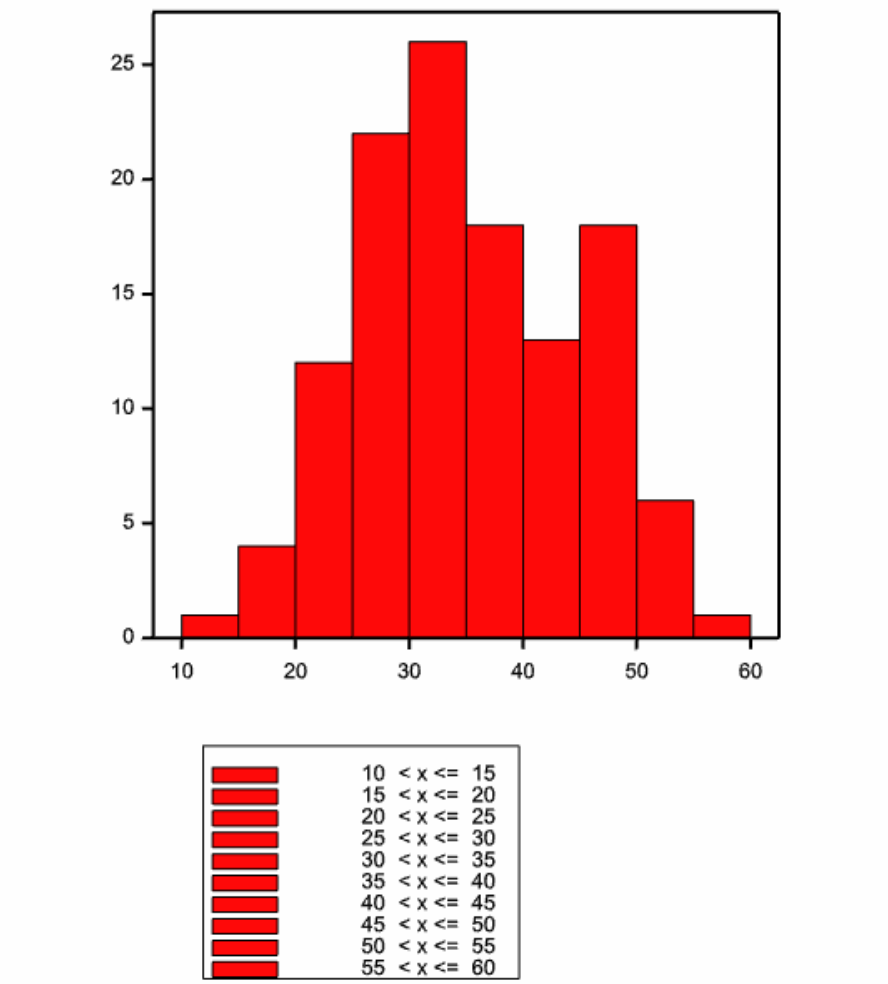


**Figure 22** Criterion 6 - Absence of injuries: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.7 Criterion 7: Absence of disease

| Measure | Nature | Level of observation |
|---|---|---|
| Mortality | Cardinal | Flock |
| Culls | Cardinal | Flock |
| Enlarged crops | Ordinal | Flock |
| Eye pathology | Ordinal | Flock |
| Respiratory infections | Ordinal | Flock |
| Enteritis | Ordinal | Flock |
| Comb abnormality | Ordinal | Flock |
| Parasites and flies | Ordinal | Flock |

*First stage*

To set an alarm threshold for percentage mortality the experts have been consulted (table 17). Warning signal is set on 50% of the alarm signal.
The health problems enlarged crops, eye pathology, respiratory infections, enteritis and comb abnormalities are scored on a 3-point scale. For all these health issues a score 1 has been regarded as warning signal and a score 2 as an alarm (see table 18).

For evidence for parasites and flies scoring has been done on a 2-point scale (present or not present). Table 19 provided the conversion to warning signals and alarms.

**Table 17** Results from experts' evaluation on alarm threshold for Mortality (%)

| **% hens with keel bone damage** | ***Score per expert*** | | | | | |
|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** |
| 0 | 10 | 10 | 4 | 10 | 10 | 8.8 |

**Table 18** Conversion of scores to warning and alarm thresholds for Enlarged crops, eye pathology, respiratory infections, enteritis and comb abnormalities.

| **Number of birds affected** | **Scoring on 3-point scale** | **Conversion to thresholds** |
|---|---|---|
| less than 3 birds | 0 | 0 |
| 3 to 25 birds | 1 | warning |
| 25 or more birds | 2 | alarm |

**Table 19** Conversion of scores to warning and alarm thresholds for parasites and flies

| | | **Flea faeces** | |
|---|---|---|---|
| | | **0 (= not present)** | **1 (= present)** |
| **Parasites** | **0 (= not present)** | 0 | warning |
| | **1 (= present)** | alarm | alarm |

Second stage

The experts have been asked to evaluate 20 imaginary flocks with varying numbers of warnings and alarms on various disease issues. The results are presented in table 20.

**Table 20** Results from experts' evaluation of Absence of disease

| Flock | Areas Normal | Areas with 1 or more warnings | Areas with 1 or more alarms | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* | *Index* |
| Flock 1 | 7 | 0 | 0 | 100 | 100 | 100 | 100 | 100 | 100.00 | 100.00 |
| Flock 2 | 5 | 2 | 0 | 75 | 50 | 40 | 80 | 20 | 53.00 | 80.95 |
| Flock 3 | 5 | 0 | 2 | 45 | 40 | 15 | 50 | 15 | 39.80 | 71.43 |
| Flock 4 | 4 | 3 | 0 | 55 | 25 | 30 | 70 | 19 | 33.00 | 71.43 |
| Flock 5 | 4 | 0 | 3 | 30 | 20 | 10 | 45 | 10 | 31.00 | 61.90 |
| Flock 6 | 3 | 4 | 0 | 50 | 15 | 20 | 55 | 15 | 23.40 | 52.38 |
| Flock 7 | 3 | 0 | 4 | 25 | 10 | 10 | 35 | 0 | 23.00 | 57.14 |
| Flock 8 | 3 | 2 | 2 | 40 | 12 | 10 | 45 | 10 | 19.00 | 52.38 |
| Flock 9 | 2 | 5 | 0 | 20 | 10 | 15 | 45 | 5 | 16.00 | 42.86 |
| Flock 10 | 2 | 0 | 5 | 15 | 8 | 10 | 30 | 0 | 13.80 | 42.86 |
| Flock 11 | 1 | 6 | 0 | 15 | 7 | 10 | 35 | 2 | 12.60 | 28.57 |
| Flock 12 | 1 | 0 | 6 | 10 | 5 | 5 | 25 | 0 | 11.80 | 33.33 |
| Flock 13 | 0 | 7 | 0 | 15 | 4 | 10 | 30 | 0 | 10.20 | 28.57 |
| Flock 14 | 0 | 6 | 1 | 10 | 3 | 10 | 28 | 0 | 9.40 | 23.81 |
| Flock 15 | 0 | 5 | 2 | 10 | 2 | 10 | 25 | 0 | 9.00 | 14.29 |
| Flock 16 | 0 | 4 | 3 | 5 | 1 | 5 | 24 | 0 | 7.00 | 19.05 |
| Flock 17 | 0 | 3 | 4 | 5 | 0 | 5 | 23 | 0 | 6.60 | 14.29 |
| Flock 18 | 0 | 2 | 5 | 0 | 0 | 5 | 22 | 0 | 5.40 | 9.52 |
| Flock 19 | 0 | 1 | 6 | 0 | 0 | 5 | 21 | 0 | 5.20 | 4.76 |
| Flock 20 | 0 | 0 | 7 | 0 | 0 | 0 | 20 | 0 | 4.00 | 0.00 |

Absence of disease is expressed as number of warnings and alarms.

absence of disease = p

p is used to calculate index I: 100 * (1 - (2 * #Warnings + 3 * #Alarms)/ (3 * Total possible alarms))

Then the index is transformed into a score with I-spline functions (figure 23) as follows:

When $I \leq 25$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $I \geq 25$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for the Absence of diseases:

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.7479647348 |
| c1 | -0.0199985563 |
| d1 | 0.0002620602 |
| a2 | 1.3655962489 |
| b2 | 0.5840931849 |
| c2 | -0.0134436943 |
| d2 | 0.0001746620 |

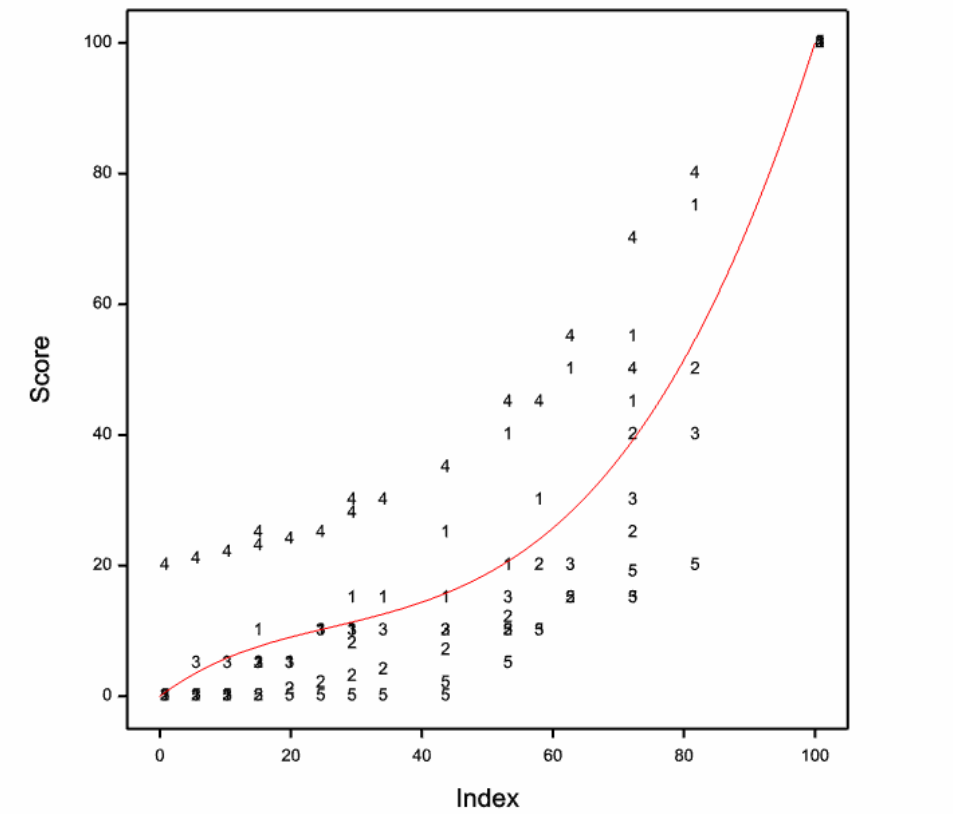


**Figure 23** Calculation of criteria scores for absence of disease; spline with one interior knot at 25; (x axis, 100 * (1 - (2 * #Warnings + 3 * #Alarms)/ (3 * Total possible alarms)))

In figure 24 the distribution of absence of disease is presented as calculated from the dataset of 122 flocks.

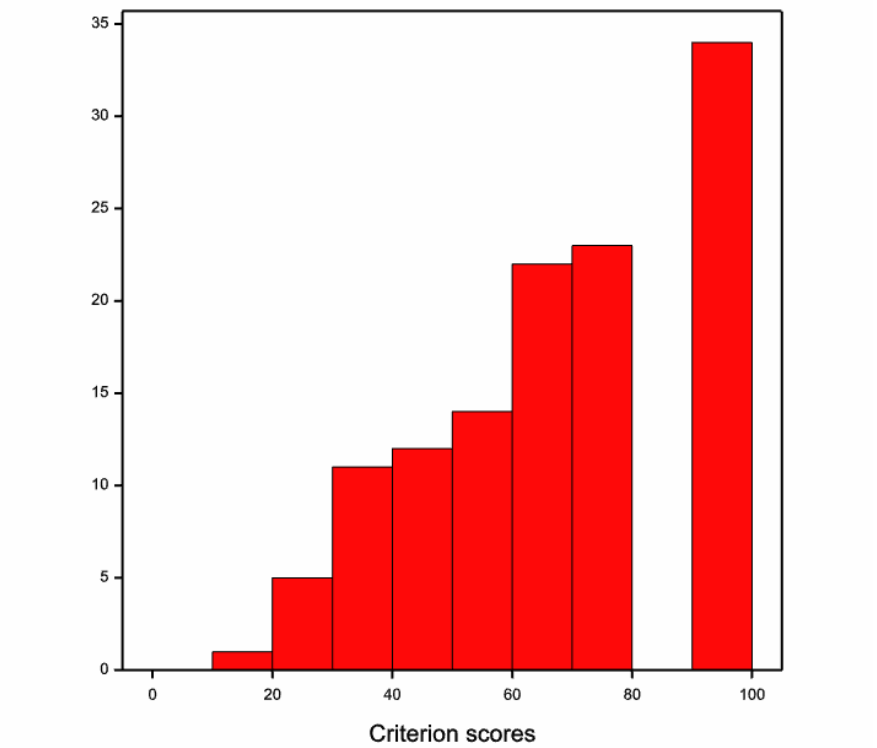


**Figure 24** Criterion 7 - Absence of disease: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.8 Criterion 8: Absence of pain due to management procedures

| Measure | Nature | Level of observation |
|---|---|---|
| Beak treatment | Cardinal | Flock |

The experts have been asked to evaluate 18 imaginary flocks with varying levels of percentages birds with keel bone damage. The results are presented in table 21.

**Table 21** Results from experts' evaluation of beak treatment

| ***Beak treatment **** | ***Beaks with score 2*** | ***Score per expert*** | | | | | |
|---|---|---|---|---|---|---|---|
| | | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** |
| Score =0 | 0% | 100 | 100 | 100 | 100 | 100 | 100.00 |
| *Score= 1 ** | *0%* | *77* | *70* | *70* | *5* | *87* | *61.80* |
| Score =1 or 2 | 1-10% | 55 | 40 | 40 | 1 | 75 | 42.20 |
| Score =1 or 2 | 11-20% | 35 | 15 | 20 | 0 | 55 | 25.00 |
| Score =1 or 2 | 21-30% | 20 | 10 | 15 | 0 | 35 | 16.00 |
| Score =1 or 2 | 31-40% | 15 | 5 | 15 | 0 | 15 | 10.00 |
| Score =1 or 2 | 41-50% | 10 | 0 | 10 | 0 | 0 | 4.00 |
| Score =1 or 2 | 51-60% | 10 | 0 | 10 | 0 | 0 | 4.00 |
| Score =1 or 2 | 61-70% | 5 | 0 | 5 | 0 | 0 | 2.00 |
| Score =1 or 2 | 71-80% | 5 | 0 | 5 | 0 | 0 | 2.00 |
| Score =1 or 2 | 81-90% | 0 | 0 | 5 | 0 | 0 | 1.00 |
| Score =1 or 2 | 91-100% | 0 | 0 | 0 | 0 | 0 | 0.00 |

** Score 0 = not trimmed; score 1 = moderate treatment, no abnormalities; score 2= severe treatment and/or abnormalities*

*** After the MCE was carried out, it appeared that one option was missing, being score 1 without any birds with score 2. As this option was needed to carry out the calculations, estimates have been calculated as average between the adjacent scores (except for expert 4, where a low score in line with the other scores is chosen).*

No further transformation is carried out. Depending on the category a beak score falls in it is converted into the mean score for that category.
Beak scores equal scores for criterion 8 - Absence of pain induced by management procedures.

In figure 25 the distribution of scores for criterion 8 is presented as calculated from the dataset of 122 flocks.

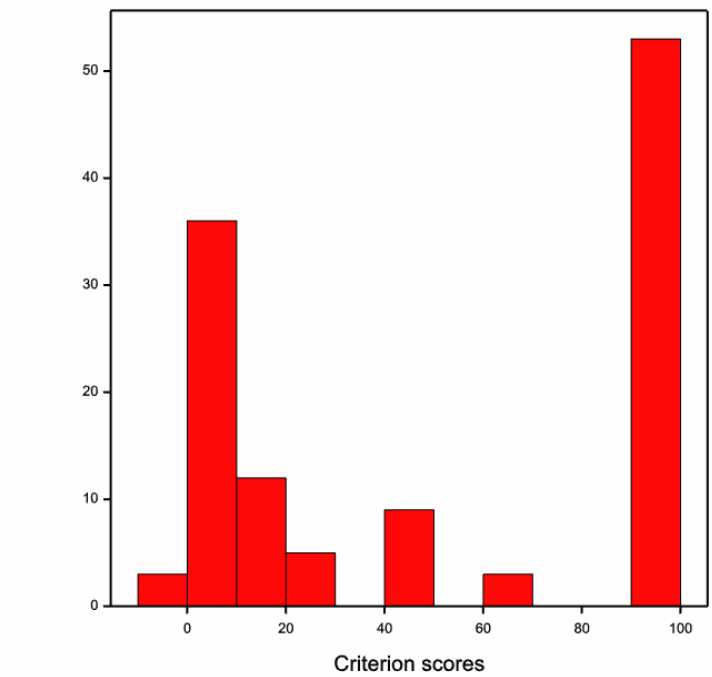


**Figure 25** Criterion 8 - Absence of pain induced by management procedures: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.9 Criterion 9: Expression of social behaviours

| Measure | Nature | Level of observation |
|---|---|---|
| Aggression | Cardinal | Flock |
| Plumage score | Ordinal | Flock |
| Comb pecks | Ordinal | Flock |

*First stage*

Aggression has been scored on a 2-point scale: present or not present. To convert this to a scale from 0 to 100, the experts have been asked to give a score to presence of aggression (table 26). The mean score is assigned to presence of aggression. If no aggression was present a score of 100 has been given.

**Table 22** Results from experts' evaluation on presence of aggression

| Presence of aggression | *Score per expert* | | | | | |
|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** |
| 0 | 20 | 20 | 10 | 40 | 30 | 24 |

The experts have been asked to evaluate 18 imaginary flocks with varying levels of percentages birds with plumage damage. The results are presented in table 23.

**Table 23** Results from experts' evaluation of Plumage damage *****

| Flock | % hens with Score 0 | % hens with Score 1 | % hens with Score 2 | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* | *Index* |
| Flock 1 | 100 | 0 | 0 | 100 | 100 | 100 | 100 | 100 | 100.00 | 100.00 |
| Flock 2 | 95 | 5 | 0 | 80 | 90 | 90 | 85 | 95 | 88.00 | 98.33 |
| Flock 3 | 95 | 0 | 5 | 60 | 50 | 50 | 65 | 85 | 80.00 | 96.67 |
| Flock 4 | 90 | 10 | 0 | 75 | 70 | 85 | 80 | 90 | 65.00 | 93.33 |
| Flock 5 | 90 | 5 | 5 | 60 | 40 | 50 | 60 | 80 | 62.00 | 95.00 |
| Flock 6 | 80 | 20 | 0 | 65 | 50 | 70 | 60 | 80 | 58.00 | 93.33 |
| Flock 7 | 80 | 15 | 5 | 60 | 30 | 45 | 55 | 70 | 56.00 | 91.67 |
| Flock 8 | 80 | 10 | 10 | 50 | 25 | 30 | 50 | 60 | 52.00 | 90.00 |
| Flock 9 | 75 | 25 | 0 | 55 | 35 | 60 | 55 | 75 | 51.00 | 90.00 |
| Flock 10 | 75 | 20 | 5 | 50 | 20 | 40 | 55 | 65 | 46.00 | 88.33 |
| Flock 11 | 75 | 15 | 10 | 40 | 18 | 25 | 50 | 55 | 43.00 | 86.67 |
| Flock 12 | 70 | 30 | 0 | 50 | 25 | 55 | 55 | 70 | 43.00 | 86.67 |
| Flock 13 | 70 | 25 | 5 | 45 | 20 | 40 | 55 | 55 | 38.00 | 86.67 |
| Flock 14 | 70 | 20 | 10 | 15 | 15 | 25 | 50 | 50 | 37.60 | 85.00 |
| Flock 15 | 60 | 40 | 0 | 20 | 10 | 50 | 50 | 60 | 31.00 | 83.33 |
| Flock 16 | 60 | 35 | 5 | 15 | 5 | 35 | 50 | 50 | 31.00 | 83.33 |
| Flock 17 | 60 | 30 | 10 | 15 | 2 | 20 | 45 | 40 | 24.40 | 80.00 |
| Flock 18 | 50 | 40 | 10 | 15 | 1 | 15 | 40 | 30 | 20.20 | 76.67 |

** score 0 = hens without plumage damage; score 1 = hens with minor plumage damage; score 2 =hens with severe plumage damage and naked areas*

The percentage plumage damage is calculated as: index I: 100 - (1*%score1 + 3*%score2)/3

Then the index is transformed into a score with I-spline functions (figure 26) as follows:

When I≤77: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When 77≤I≤95: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$
When I≥95: $I = (a3 + b3*I + c3*I^2 + d3*I^3)$

The following coefficients are used for plumage damage:

| COEF | COEF2 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.0000016787 |
| c1 | -0.0000168414 |
| d1 | 0.0000561385 |
| a2 | -1642.7309864806 |
| b2 | 64.0025076390 |
| c2 | -0.8312182175 |
| d2 | 0.0036544129 |
| a3 | -3354.1612190858 |
| b3 | 118.0476737478 |
| c3 | -1.4001147118 |
| d3 | 0.0056505410 |

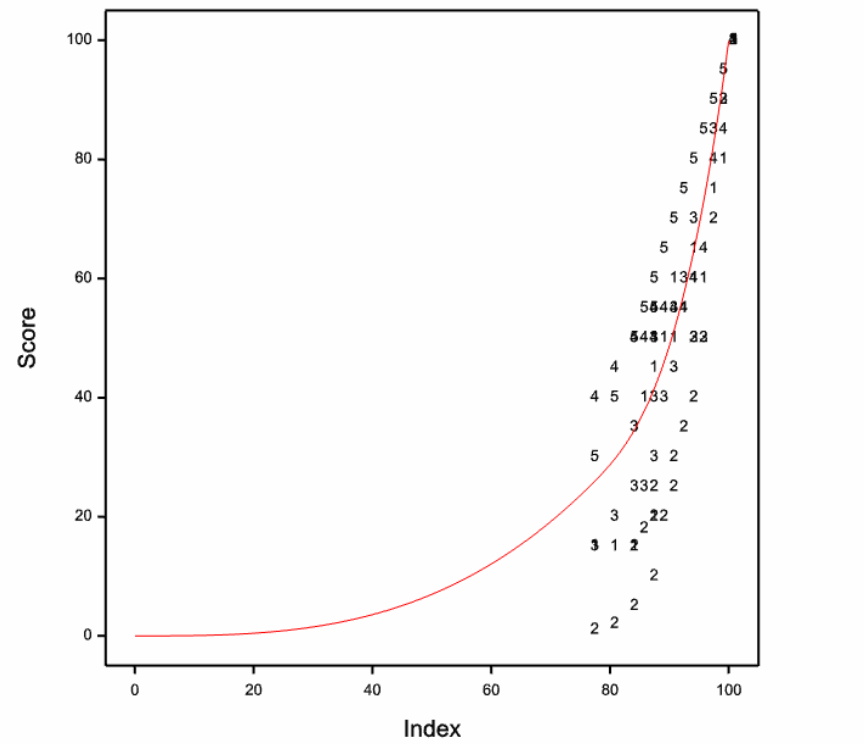


**Figure 26**: Calculation of scores for plumage condition; spline with two interior knots at 77 and 95; (x axis, 100 - (1*%score1 + 3*%score2)/3)

In figure 27 the distribution of percentage plumage damage is presented as calculated from the dataset of 122 flocks.

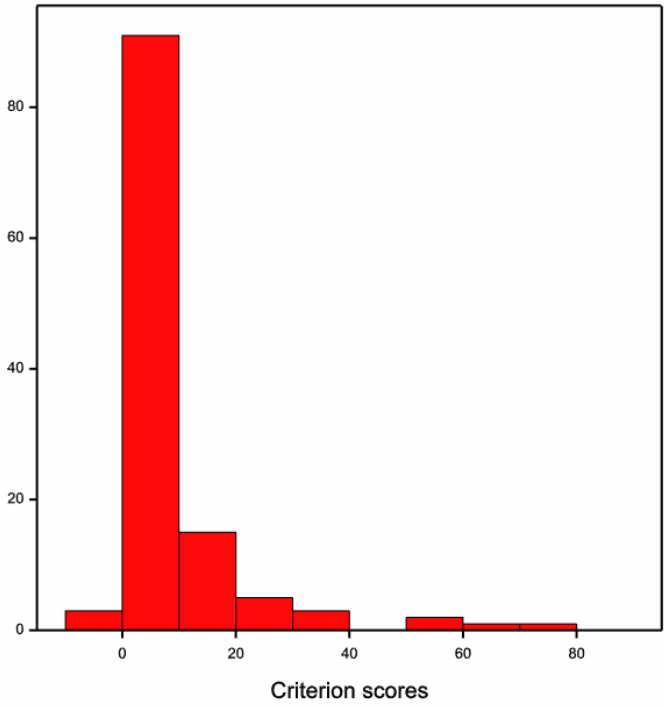


**Figure 27** Plumage damage: Distribution of scores as calculated from the dataset of 122 flocks

The experts have been asked to evaluate 21 imaginary flocks with varying levels of percentages birds with comb pecks. The results are presented in table 24.

**Table 24** Results from experts' evaluation of Comb pecks

| Flock | % hens with Score 0 | % hens with Score 1 | % hens with Score 2 | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* | *Index* |
| Flock 1 | 100 | 0 | 0 | 100 | 100 | 100 | 100 | 100 | 100.00 | 100.00 |
| Flock 2 | 90 | 10 | 0 | 90 | 90 | 80 | 80 | 80 | 84.00 | 94.00 |
| Flock 3 | 90 | 0 | 10 | 75 | 70 | 50 | 20 | 60 | 64.00 | 88.00 |
| Flock 4 | 80 | 20 | 0 | 80 | 80 | 60 | 40 | 60 | 55.00 | 90.00 |
| Flock 5 | 80 | 0 | 20 | 60 | 60 | 30 | 18 | 20 | 52.00 | 82.00 |
| Flock 6 | 70 | 30 | 0 | 70 | 65 | 50 | 35 | 40 | 42.00 | 76.00 |
| Flock 7 | 70 | 15 | 15 | 55 | 55 | 30 | 19 | 10 | 37.60 | 80.00 |
| Flock 8 | 70 | 0 | 30 | 35 | 45 | 20 | 10 | 0 | 34.00 | 70.00 |
| Flock 9 | 60 | 40 | 0 | 60 | 55 | 45 | 30 | 20 | 33.80 | 76.00 |
| Flock 10 | 60 | 0 | 40 | 20 | 50 | 15 | 8 | 0 | 27.80 | 64.00 |
| Flock 11 | 50 | 50 | 0 | 55 | 50 | 40 | 25 | 0 | 22.00 | 70.00 |
| Flock 12 | 50 | 0 | 50 | 20 | 40 | 15 | 5 | 0 | 20.60 | 58.00 |
| Flock 13 | 40 | 60 | 0 | 40 | 45 | 35 | 19 | 0 | 18.60 | 60.00 |
| Flock 14 | 40 | 0 | 60 | 10 | 30 | 10 | 5 | 0 | 16.00 | 50.00 |
| Flock 15 | 30 | 70 | 0 | 20 | 35 | 30 | 18 | 0 | 15.20 | 52.00 |
| Flock 16 | 30 | 35 | 35 | 15 | 20 | 15 | 7 | 0 | 11.40 | 44.00 |
| Flock 17 | 30 | 0 | 70 | 0 | 15 | 5 | 3 | 0 | 11.00 | 46.00 |
| Flock 18 | 20 | 80 | 0 | 15 | 20 | 25 | 16 | 0 | 11.00 | 40.00 |
| Flock 19 | 20 | 40 | 40 | 10 | 15 | 10 | 17 | 0 | 10.40 | 36.00 |
| Flock 20 | 10 | 90 | 0 | 15 | 10 | 15 | 15 | 0 | 6.00 | 36.00 |
| Flock 21 | 0 | 90 | 10 | 5 | 5 | 10 | 10 | 0 | 4.60 | 30.00 |

*score 0 = no comb pecks; score 1 = hens with less than 3 comb pecks; score 2 =hens with 3 or more comb pecks*

The percentage comb pecks is calculated as: index I: 100 - (3*%score1 + 5*%score2)/5

Then the index is transformed into a score with I-spline functions (figure 28) as follows:

When I≤75: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When I≥75: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for comb pecks:

| COEF | COEF2 |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.1971201438 |
| c1 | -0.0026247697 |
| d1 | 0.0000866983 |
| a2 | -534.0030577086 |
| b2 | 21.5584170490 |
| c2 | -0.2874681222 |
| d2 | 0.0013528426 |

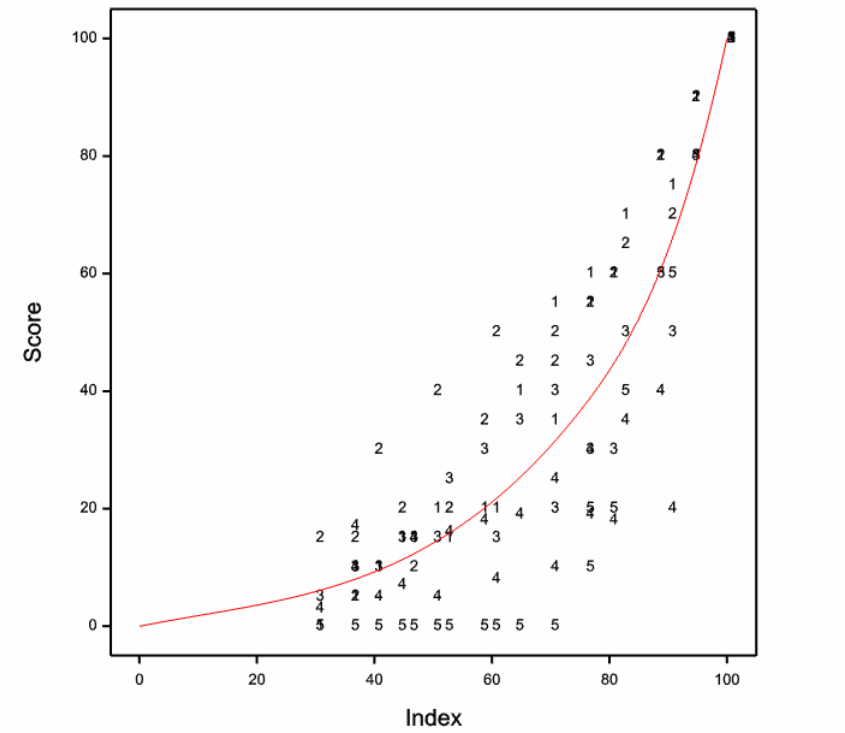


**Figure 28** Calculation of scores for % comb pecks; spline with one interior knot at 75; (x axis, 100 - (3*%score1 + 5*%score2)/5)

In figure 29 the distribution of percentage comb pecks is presented as calculated from the dataset of 122 flocks.

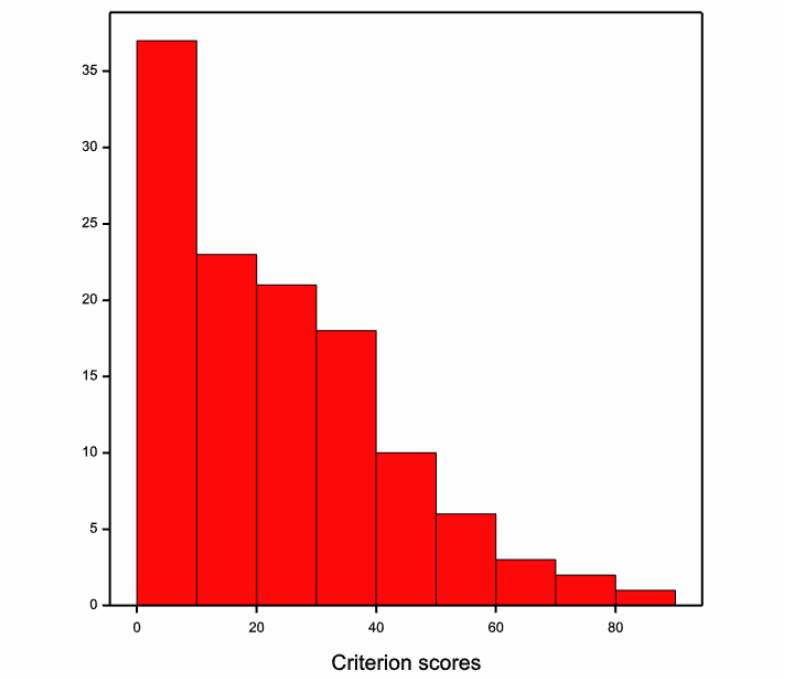


**Figure 29** % Comb pecks: Distribution of scores as calculated from the dataset of 122 flocks

*Second stage*

To aggregate the three sub-scores into the criterion-scores the experts were asked to assign relative importance to the three measures, while having the possibility to focus on worst scores. To

understand their way of reasoning, we presented to them a dataset composed of 5 virtual farms characterised by their three sub-scores (Table 25).

**Table 25** Results obtained from the experts asked to score13 imaginary farms for the criterion Expression of social behaviour

| Flock | Score for aggression | Score for feather damage | Score for comb pecks | Score per expert Exp. 1 | Exp. 2 * | Exp. 3 | Exp. 4 | Exp. 5 | Mean |
|---|---|---|---|---|---|---|---|---|---|
| Flock 1 | 25 | 50 | 75 | 40 | | 45 | 45 | 60 | 47.50 |
| Flock 2 | 25 | 75 | 50 | 65 | | 50 | 55 | 60 | 57.50 |
| Flock 3 | 40 | 50 | 60 | 50 | | 50 | 50 | 55 | 51.25 |
| Flock 4 | 40 | 60 | 50 | 55 | | 50 | 55 | 55 | 53.75 |
| Flock 5 | 50 | 25 | 75 | 30 | | 30 | 28 | 45 | 33.25 |
| Flock 6 | 50 | 40 | 60 | 45 | | 45 | 40 | 50 | 45.00 |
| Flock 7 | 50 | 50 | 50 | 50 | | 50 | 50 | 50 | 50.00 |
| Flock 8 | 50 | 60 | 40 | 55 | | 50 | 50 | 50 | 51.25 |
| Flock 9 | 50 | 75 | 25 | 60 | | 50 | 50 | 45 | 51.25 |
| Flock 10 | 60 | 40 | 50 | 40 | | 45 | 40 | 45 | 42.50 |
| Flock 11 | 60 | 50 | 40 | 50 | | 50 | 50 | 45 | 48.75 |
| Flock 12 | 75 | 25 | 50 | 35 | | 30 | 30 | 35 | 32.50 |
| Flock 13 | 75 | 50 | 25 | 50 | | 40 | 35 | 35 | 40.00 |

** expert 2 did not feel comfortable with the measures falling under this criterion and therefore did not fill in the table.*

The three partial scores, 1=Aggression, 2=Plumage score, 3=Comb pecks, are combined to a score for criterion 9 using a Choquet integral with the following parameters:

| Parameter | Estimate | Parameter | Estimate |
|---|---|---|---|
| μ 1 | 0.09750 | μ 12 | 0.5376 |
| μ 2 | 0.5257 | μ 13 | 0.1985 |
| μ 3 | 0.1521 | μ 23 | 0.7818 |

In figure 30 the distribution of criterion scores for criterion 9, Expression of social behaviour, is presented as calculated from the dataset of 122 flocks.

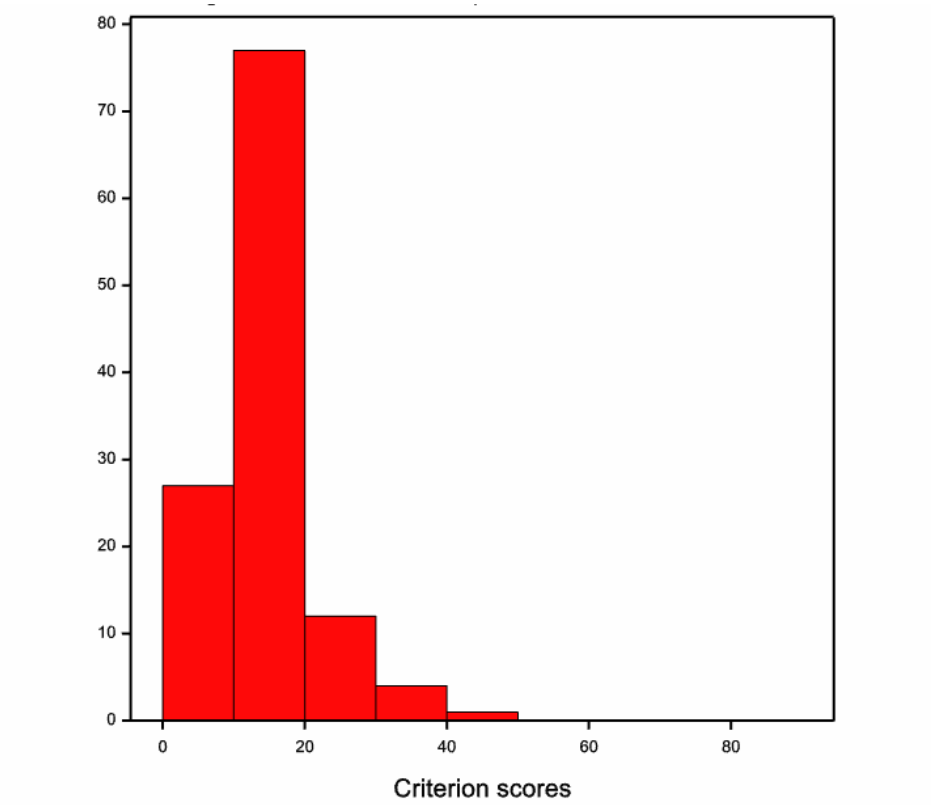


**Figure 30** Criterion 9 - Expression of social behaviour: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.10 Criterion 10: Expression of other behaviours

| Measure | Nature | Level of observation |
|---|---|---|
| Nestbox use | Cardinal | Flock |
| Nest space per hen | Ordinal | Individual |
| Environmental enrichment | Cardinal | Individual |

*First stage*

The experts have been asked to evaluate the 9 possible situations regarding nestbox use. The results are presented in table 26.

**Table 26** Results obtained from the experts asked to score the 9 possible situations regarding nestbox use.

| Nestboxes available | Distribution of nest boxes | Distribution of eggs within nest rows | Distribution of eggs between nest rows | Score per expert Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 | Mean |
|---|---|---|---|---|---|---|---|---|---|
| Nests | Even | Even | Even | 100 | 100 | 90 | 100 | 100 | 98.00 |
| Nests | Even | Even | Not even | 80 | 70 | 40 | 75 | 90 | 71.00 |
| Nests | Even | Not even | Even | 80 | 55 | 40 | 75 | 70 | 64.00 |
| Nests | Even | Not even | Not even | 70 | 50 | 35 | 50 | 50 | 51.00 |
| Nests | Not even | Even | Even | 70 | 90 | 90 | 100 | 60 | 82.00 |
| Nests | Not even | Even | Not even | 60 | 50 | 40 | 75 | 30 | 51.00 |
| Nests | Not even | Not even | Even | 60 | 30 | 40 | 75 | 10 | 43.00 |
| Nests | Not even | Not even | Not even | 50 | 20 | 35 | 50 | 5 | 32.00 |
| No nests | n.a. | n.a. | n.a. | 0 | 0 | 0 | 0 | 0 | 0.00 |

No further transformation is carried out. Depending on the category a score for nestbox use falls in, it is converted into the mean score for that category.

The experts have been asked to evaluate 11 imaginary flocks with varying nest space per hen. The results are presented in table 27.

**Table 27** Results from experts' evaluation of Nest space ($cm^2$/hen)

| Nest space (cm2/hen) | Score per expert Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 | Exp. 5 | Mean | Index |
|---|---|---|---|---|---|---|---|
| 160 | 80 | 80 | 80 | 80 | 100 | 84.00 | 100.00 |
| 140 | 80 | 70 | 75 | 60 | 90 | 75.00 | 87.50 |
| 120 | 70 | 60 | 70 | 55 | 70 | 65.00 | 75.00 |
| 100 | 55 | 50 | 65 | 50 | 60 | 56.00 | 62.50 |
| 95 | 45 | 40 | 55 | 45 | 50 | 47.00 | 59.38 |
| 90 | 35 | 30 | 50 | 40 | 20 | 35.00 | 56.25 |
| 85 | 25 | 20 | 30 | 25 | 15 | 23.00 | 53.12 |
| 80 | 20 | 15 | 20 | 20 | 10 | 17.00 | 50.00 |
| 60 | 15 | 10 | 15 | 18 | 5 | 12.60 | 37.50 |
| 40 | 10 | 5 | 10 | 15 | 0 | 8.00 | 25.00 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0.00 | 0.00 |

Nest space is calculated as:

$$cm^2 \text{ per hen} = p = \frac{\text{total available space } (cm^2)}{\text{total number of birds present}}$$

p is used to calculate index I: 100 * p / 160

with 160 being the highest space per hens resulting in a score100

Then the index is transformed into a score with I-spline functions (figure 31) as follows:

When I≤50: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When I≥50: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for the nest space ($cm^2$/hen):

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.0000000000 |
| c1 | 0.0068928967 |
| d1 | 0.0000730466 |
| a2 | 41.9755393835 |
| b2 | -2.5185323631 |
| c2 | 0.0572635440 |
| d2 | -0.0002627577 |

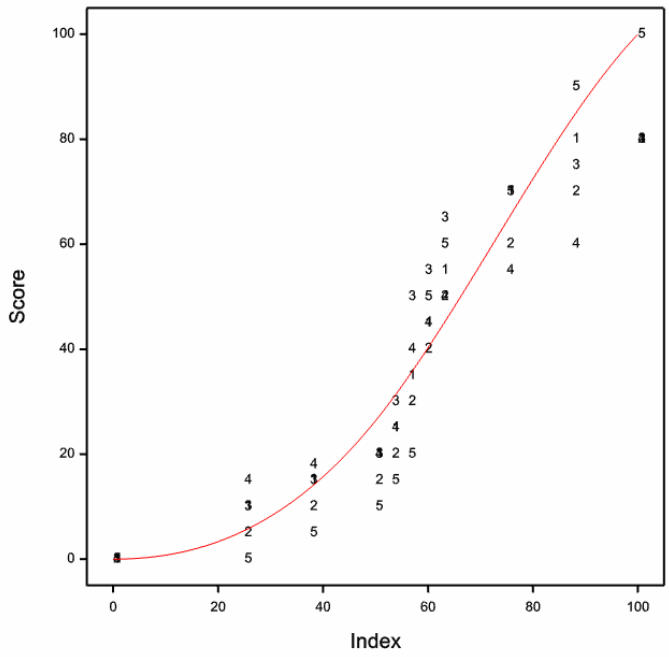


**Figure 31** Calculation of scores for nest space ($cm^2$/hen); spline with one interior knot at 50; (x axis, 100 * nestspace /160)

In figure 32 the distribution of nest space is presented as calculated from the dataset of 122 flocks.

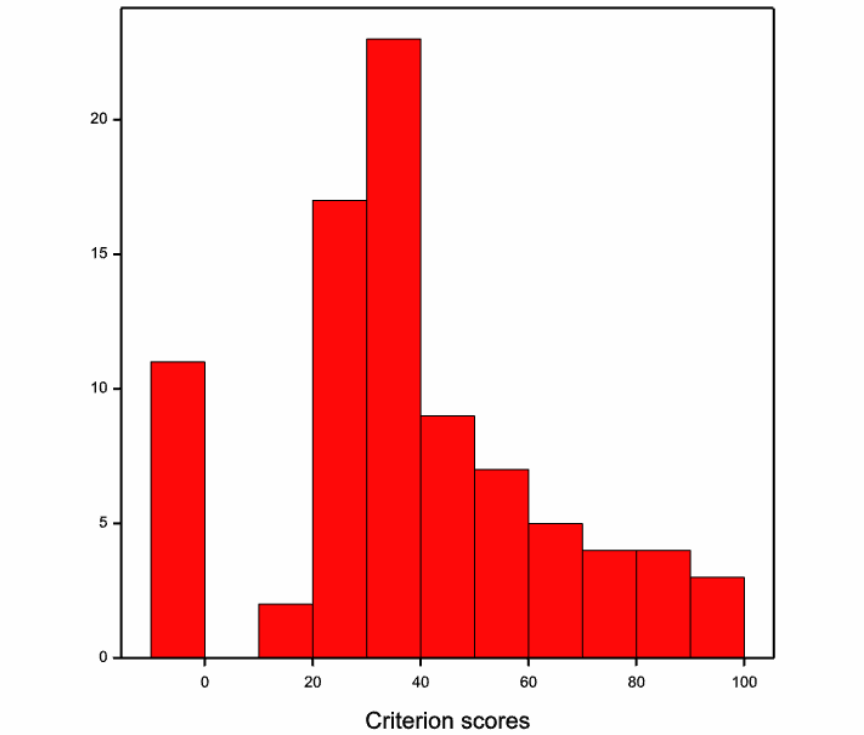


**Figure 4.32** Nest space (cm2/hen): Distribution of criterion scores as calculated from the dataset of 122 flocks

Use of environmental enrichments has been scored on 5 aspects:

| | Score 0 | Score 1 | Score 2 |
|---|---|---|---|
| a. Use of litter | good | moderate | bad |
| b. Use of Enrichments | good | moderate | bad |
| c. Use of free range | good | moderate | bad |
| d. Cover on range | good = >5% | moderate = ≤ 5% | bad = n.a or 0% |
| e. Use of covered veranda | good = 50-100% | moderate = < 50% | bad = n.a. |

For each aspect score 1 equals a warning and score 2 equals an alarm. The experts have been asked to evaluate 16 imaginary flocks with varying number of warnings and alarms. The results are presented in table 28.

**Table 28** Results obtained from the experts asked to score the 16 imaginary flocks with varying number of warnings and alarms regarding use of environmental enrichment.

| Flock | Normal areas | Areas with warning | Areas with alarm | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Exp. 1* | *Exp. 2* | *Exp. 3* | *Exp. 4* | *Exp. 5* | *Mean* | *Index* |
| Flock 1 | 5 | 0 | 0 | 100 | 100 | 100 | 80 | 100 | 96.00 | 100.00 |
| Flock 2 | 4 | 1 | 0 | 80 | 90 | 80 | 70 | 80 | 80.00 | 86.67 |
| Flock 3 | 4 | 0 | 1 | 70 | 60 | 50 | 55 | 70 | 63.00 | 73.33 |
| Flock 4 | 3 | 2 | 0 | 70 | 60 | 60 | 65 | 60 | 61.00 | 80.00 |
| Flock 5 | 3 | 0 | 2 | 55 | 40 | 40 | 45 | 40 | 47.00 | 60.00 |
| Flock 6 | 2 | 3 | 0 | 65 | 30 | 50 | 60 | 30 | 44.00 | 60.00 |
| Flock 7 | 2 | 0 | 3 | 40 | 20 | 15 | 40 | 10 | 37.00 | 46.67 |
| Flock 8 | 1 | 4 | 0 | 50 | 25 | 40 | 55 | 15 | 25.00 | 40.00 |
| Flock 9 | 1 | 0 | 4 | 15 | 15 | 10 | 35 | 5 | 23.60 | 33.33 |
| Flock 10 | 1 | 2 | 2 | 20 | 18 | 30 | 40 | 10 | 22.00 | 33.33 |
| Flock 11 | 0 | 5 | 0 | 15 | 20 | 35 | 40 | 0 | 16.00 | 20.00 |
| Flock 12 | 0 | 4 | 1 | 10 | 15 | 20 | 35 | 0 | 16.00 | 26.67 |
| Flock 13 | 0 | 3 | 2 | 10 | 13 | 15 | 30 | 0 | 13.60 | 20.00 |
| Flock 14 | 0 | 2 | 3 | 5 | 10 | 10 | 20 | 0 | 9.00 | 13.33 |
| Flock 15 | 0 | 1 | 4 | 0 | 5 | 5 | 19 | 0 | 5.80 | 6.67 |
| Flock 16 | 0 | 0 | 5 | 0 | 0 | 0 | 17 | 0 | 3.40 | 0.00 |

The score for environmental enrichment is calculate with the following formula:

index I: 100 * (1 - (2 * #warnings + 3 *#alarms)/ (3 * total number of possible warnings))

Then the index is transformed into a score with I-spline functions (figure 33) as follows:

When I≤25: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When I≥25: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for environmental enrichment:

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 0.7593934044 |
| c1 | -0.0048562286 |
| d1 | 0.0000838304 |
| a2 | 0.4150899805 |
| b2 | 0.7095826068 |
| c2 | -0.0028637967 |
| d2 | 0.0000572646 |

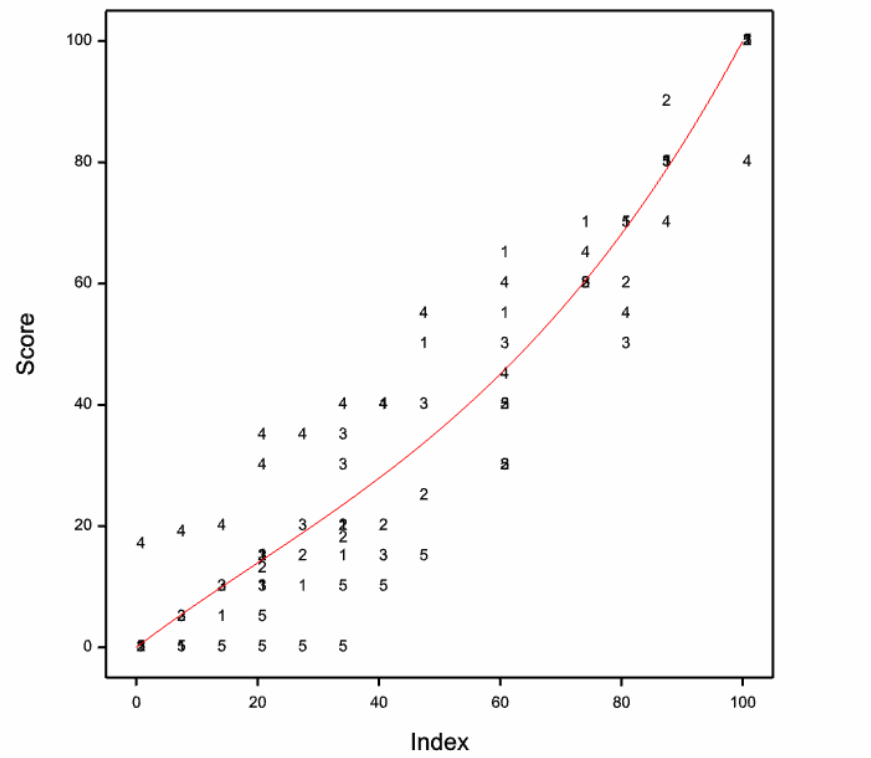


**Figure 33** Calculation of scores for environmental enrichment; spline with one interior knot at 25; (x axis, 100 * (1 - (2 * #warnings + 3 *#alarms)/ (3 * total number of possible warnings)))

In figure 34 the distribution of environmental enrichment is presented as calculated from the dataset of 122 flocks.

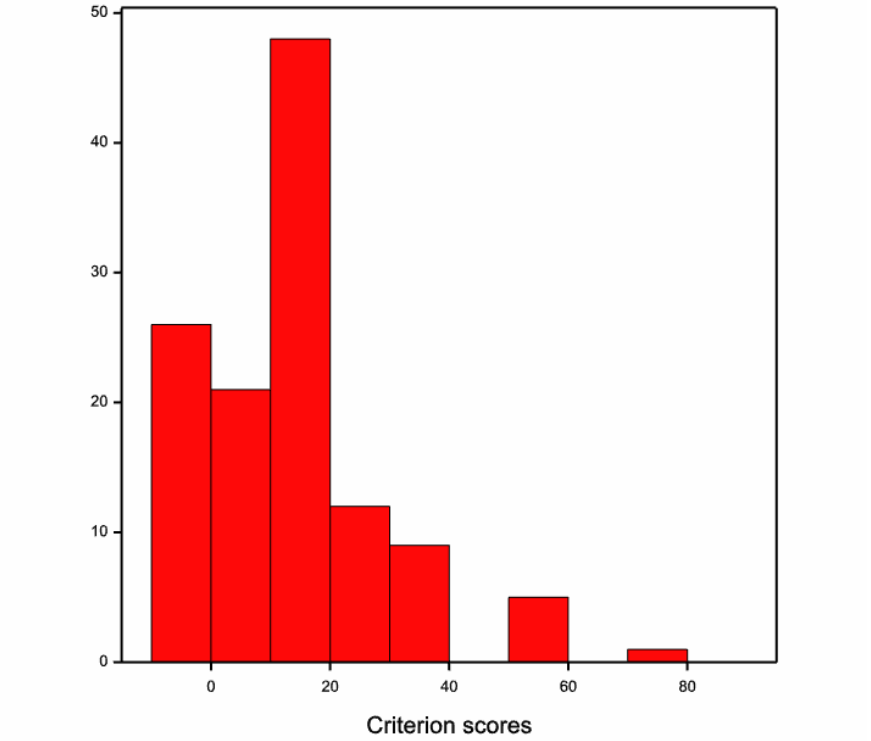


**Figure 34** Environmental enrichment: Distribution of criterion scores as calculated from the dataset of 122 flocks

*Second stage*

To aggregate the three sub-scores into the criterion-scores the experts were asked to assign relative importance to the three measures, while having the possibility to focus on worst scores. To understand their way of reasoning, we presented to them a dataset composed of 5 virtual farms characterised by their three sub-scores (Table 29).

**Table 29** Results obtained from the experts asked to score12 imaginary farms for the criterion Expression of social behaviour

| Flock | Use of nestboxes | Nest space | Use of environment | Score per expert Exp. 1 | Exp. 2 * | Exp. 3 | Exp. 4 | Exp. 5 | Mean |
|---|---|---|---|---|---|---|---|---|---|
| Flock 1 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100.00 |
| Flock 2 | 100 | 25 | 75 | 80 | 55 | 80 | 75 | 50 | 68.00 |
| Flock 3 | 100 | 50 | 50 | 60 | 50 | 70 | 50 | 70 | 60.00 |
| Flock 4 | 100 | 75 | 25 | 55 | 40 | 60 | 35 | 50 | 48.00 |
| Flock 5 | 100 | 100 | 0 | 10 | 18 | 15 | 15 | 15 | 14.60 |
| Flock 6 | 75 | 50 | 50 | 50 | 60 | 55 | 58 | 60 | 56.60 |
| Flock 7 | 50 | 75 | 50 | 50 | 55 | 55 | 50 | 55 | 53.00 |
| Flock 8 | 50 | 50 | 50 | 45 | 50 | 50 | 50 | 50 | 49.00 |
| Flock 9 | 50 | 50 | 25 | 35 | 30 | 40 | 30 | 40 | 35.00 |
| Flock 10 | 25 | 50 | 25 | 25 | 28 | 30 | 25 | 25 | 26.60 |
| Flock 11 | 25 | 25 | 50 | 25 | 35 | 35 | 35 | 30 | 32.00 |
| Flock 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |

The three partial scores, 1=Nestbox use, 2=Nest space per hen, 3=Environmental enrichment, are combined to a score for criterion 10 using a Choquet integral with the following parameters:

| Parameter | Estimate | Parameter | Estimate |
|---|---|---|---|
| μ 1 | 0.2665 | μ 12 | 0.2665 |
| μ 2 | 0.09200 | μ 13 | 0.7268 |
| μ 3 | 0.01000 | μ 23 | 0.5000 |

In figure 35 the distribution of criterion scores for criterion 10, Expression of other behaviour, is presented as calculated from the dataset of 122 flocks.

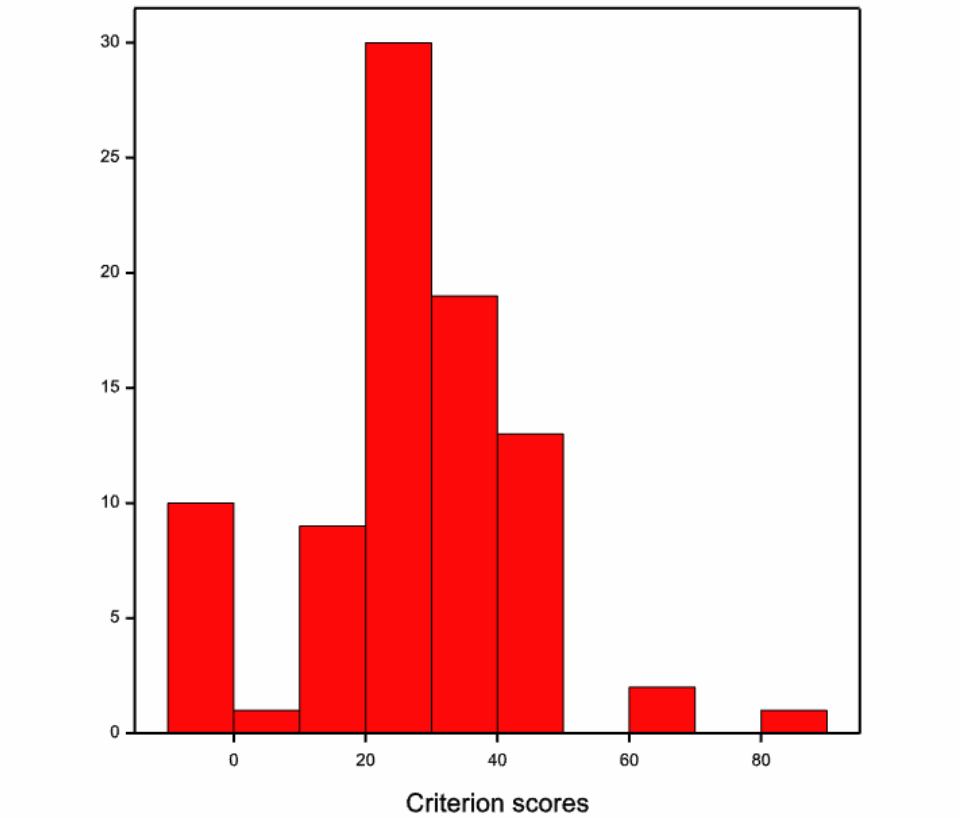


**Figure 35** Criterion 10 - Expression of other behaviour: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.11 Criterion 11: Good human-animal relationship

| Measure | Nature | Level of observation |
|---|---|---|
| Avoidance distance test | Ordinal | Flock |

The experts have been asked to evaluate 7 imaginary flocks with varying avoidance distances. The results are presented in table 30.

**Table 30** Results from experts' evaluation of avoidance distance

| Avoidance distance (cm) | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** | ***Index*** |
| 0 | 100 | 100 | 100 | 30 | 100 | 86.00 | 100.00 |
| 25 | 80 | 80 | 90 | 90 | 90 | 86.00 | 83.33 |
| 50 | 60 | 60 | 70 | 85 | 80 | 71.00 | 66.67 |
| 75 | 30 | 40 | 50 | 60 | 70 | 50.00 | 50.00 |
| 100 | 15 | 20 | 20 | 55 | 40 | 30.00 | 33.33 |
| 125 | 10 | 15 | 15 | 45 | 20 | 21.00 | 16.67 |
| 150 | 0 | 10 | 10 | 35 | 15 | 14.00 | 0.00 |

Avoidance distance is calculated as:

$$\text{average distance (cm)} = p = \frac{\text{sum of all measures}}{\text{total number of birds measured}}$$

p is used to calculate index I: 100 * (150 - p) / 150

with 150 being the longest distance resulting in a score 100

Then the index is transformed into a score with I-spline functions (figure 36) as follows:

When $I \leq 70$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $I \geq 70$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for the avoidance distance:

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 1.0631400519 |
| c1 | -0.0025548862 |
| d1 | 0.0000298844 |
| a2 | 135.2892601467 |
| b2 | -4.7349710998 |
| c2 | 0.0802752731 |
| d2 | -0.0003645449 |

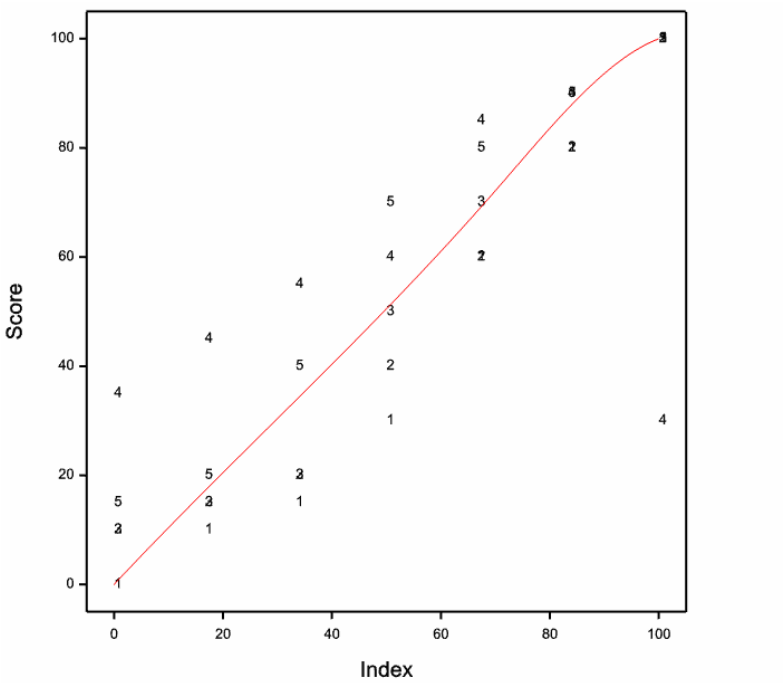


**Figure 36** Calculation of scores for avoidance distance; spline with one interior knot at 70; (x axis, 100 * (150 - avoidance distance) / 150)

In figure 37 the distribution of avoidance distance is presented as calculated from the dataset of 122 flocks. The score for avoidance distance equals the score for criterion 11 Good human-animal relation.

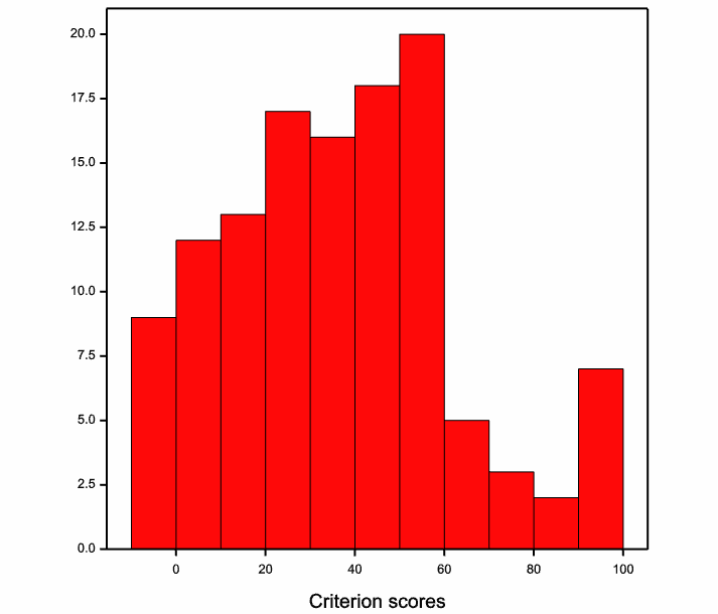


**Figure 37** Criterion 11 - Good human-animal relation: Distribution of criterion scores as calculated from the dataset of 122 flocks

### 2.12 Criterion 12: Positive emotional state

| Measure | Nature | Level of observation |
|---|---|---|
| Novel Object Test (NOT) | Ordinal | Flock |
| Qualitative Behaviour Assessment (QBA) | Ordinal | Flock |

*First stage*

The experts have been asked to evaluate 11 imaginary flocks with varying outcomes of the Novel Object test. The results are presented in table 31.

**Table 31** Results from experts' evaluation of Novel Object Test

| Novel Object Test | *Score per expert* | | | | | | |
|---|---|---|---|---|---|---|---|
| | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** | ***Index*** |
| 20 | 100 | 100 | 95 | 35 | 100 | 86.00 | 100.00 |
| 18 | 100 | 90 | 90 | 40 | 100 | 84.00 | 90.00 |
| 16 | 80 | 80 | 85 | 40 | 100 | 77.00 | 80.00 |
| 14 | 75 | 70 | 80 | 45 | 90 | 72.00 | 70.00 |
| 12 | 65 | 60 | 75 | 45 | 80 | 65.00 | 60.00 |
| 10 | 60 | 50 | 65 | 50 | 70 | 59.00 | 50.00 |
| 8 | 55 | 40 | 55 | 65 | 60 | 55.00 | 40.00 |
| 6 | 40 | 30 | 45 | 60 | 50 | 45.00 | 30.00 |
| 4 | 30 | 20 | 25 | 55 | 30 | 32.00 | 20.00 |
| 2 | 20 | 10 | 15 | 25 | 10 | 16.00 | 10.00 |
| 0 | 0 | 0 | 0 | 20 | 0 | 4.00 | 0.00 |

The outcome of the Novel Object Test (NOT) is calculated as:

$$NOT = p = \frac{\text{total number of hens within 30 cm of Novel Object}}{\text{total number of counts}}$$

p is used to calculate index I: 100 * p / 20

with 20 being the maximum number of hens resulting in a score 100

Then the index is transformed into a score with I-spline functions (figure 38) as follows:

When $I \leq 70$: $I = (a1 + b1*I + c1*I^2 + d1*I^3)$
When $I \geq 70$: $I = (a2 + b2*I + c2*I^2 + d2*I^3)$

The following coefficients are used for Novel Object count:

| COEF | COEF |
|---|---|
| a1 | 0.0000000000 |
| b1 | 1.9659146505 |
| c1 | -0.0191480328 |
| d1 | 0.0000801151 |
| a2 | -187.6819357043 |
| b2 | 10.0094261971 |
| c2 | -0.1340553408 |
| d2 | 0.0006272927 |

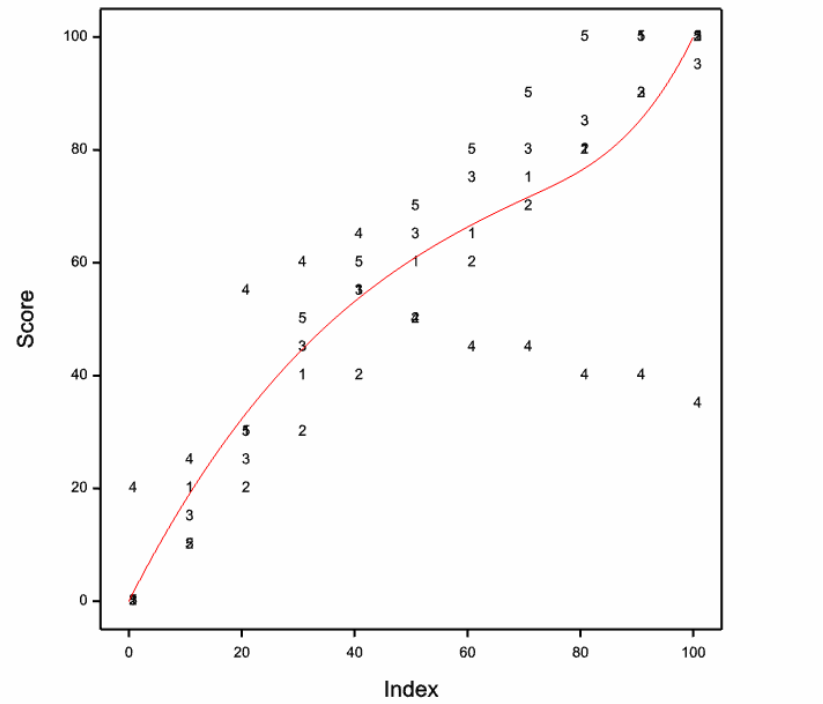


**Figure 38** Calculation of scores for novel object test; spline with one interior knot at 70; (x axis, 100 * not / 20)

In figure 39 the distribution of the outcome of the novel object test is presented as calculated from the dataset of 122 flocks.

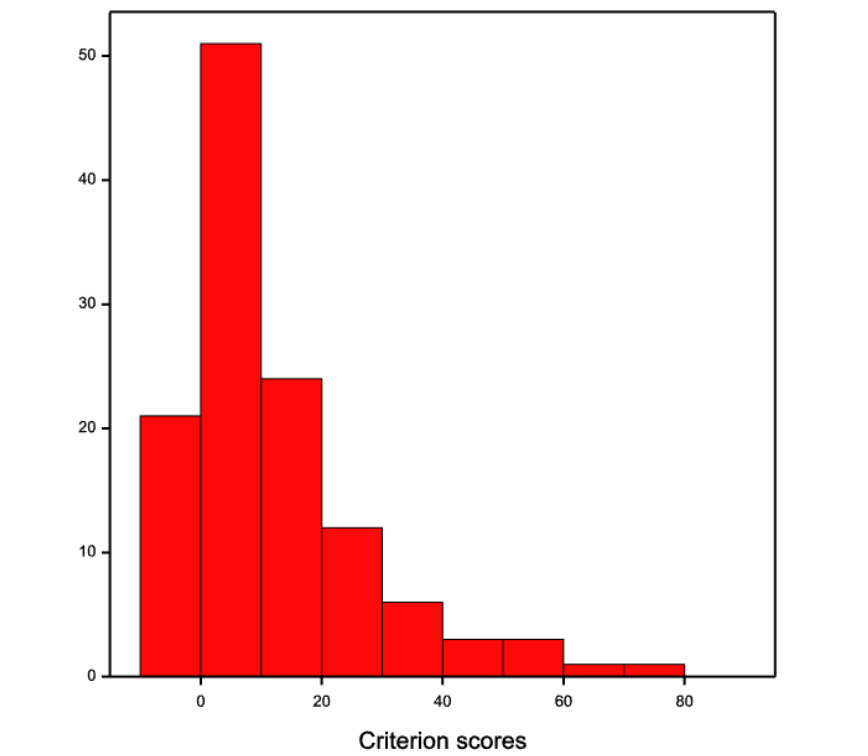


**Figure 39** Novel Object Test: Distribution of criterion scores as calculated from the dataset of 122 flocks

The scores for QBA have been transformed to a score from 0-100 with the use of an algorithm developed within the Welfare Quality® which describes the relationship between Index values (Figure 40, horizontal axis) and welfare scores (Figure 40, vertical axis). The inputs for the calculation of Index values are (i) values for the various terms comprising the QBA, and (ii) weights attributed to each term (derived from a Principal Component Analysis).

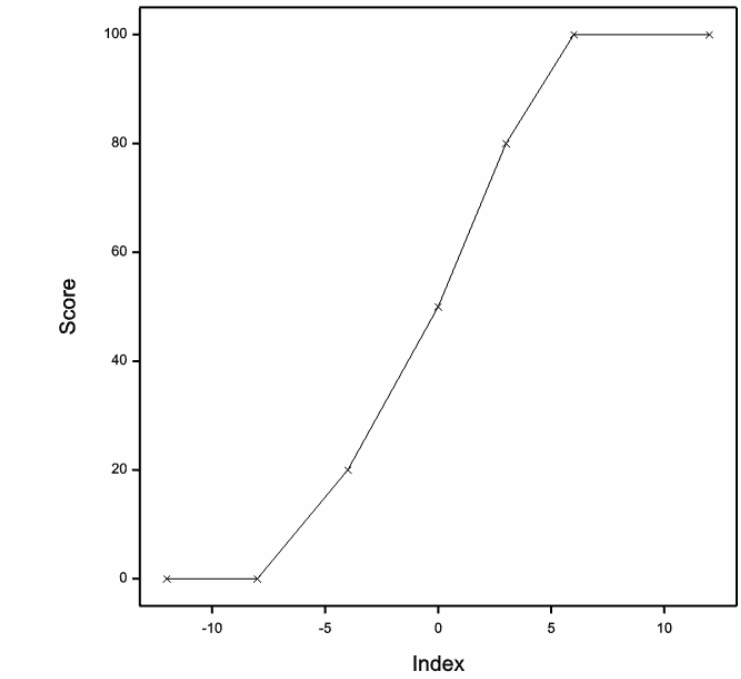


**Figure 40** Transformation of scores for QBA into a scale of 0-100;

*Second stage*

To aggregate the three sub-scores into the criterion-scores the experts were asked to assign relative importance to the three measures, while having the possibility to focus on worst scores. To understand their way of reasoning, we presented to them a dataset composed of 5 virtual farms characterised by their three sub-scores (Table 29).

**Table 29** Results obtained from the experts asked to score12 imaginary farms for the criterion Positive emotional state

| Flock | NOT | QBA | *Score per expert* | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | ***Exp. 1*** | ***Exp. 2*** | ***Exp. 3*** | ***Exp. 4*** | ***Exp. 5*** | ***Mean*** |
| Farm 1 | 25 | 75 | 60 | 55 | 60 | 30 | 30 | 47.00 |
| Farm 2 | 40 | 60 | 50 | 55 | 55 | 45 | 45 | 50.00 |
| Farm 3 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50.00 |
| Farm 4 | 60 | 40 | 50 | 45 | 45 | 50 | 55 | 49.00 |
| Farm 5 | 75 | 25 | 35 | 35 | 40 | 40 | 70 | 44.00 |

The two partial scores, 1=Novel Object Test, 2=Qualitative Behaviour Assessment, are combined to a score for criterion 12 using a Choquet integral with the following parameters:

| Parameter | Estimate |
|---|---|
| μ 1 | 0.3897 |
| μ 2 | 0.4483 |

In figure 41 the distribution of criterion scores for criterion 12, Positive emotional state, is presented as calculated from the dataset of 122 flocks.

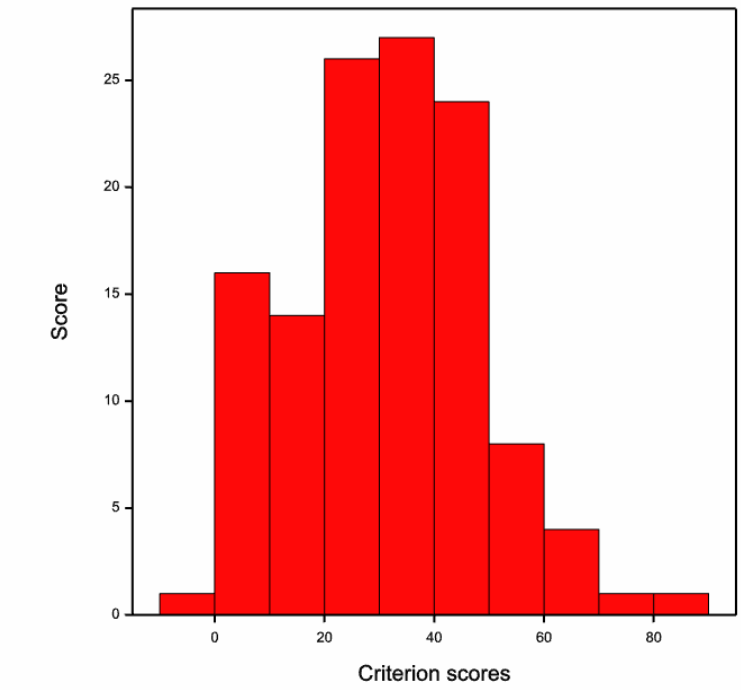


**Figure 41** Criterion 12 - Positive emotional state: Distribution of criterion scores as calculated from the dataset of 122 flocks

## Literature

Welfare Quality®, 2009. Welfare Quality® assessment protocol for poultry (broilers, laying hens). Welfare Quality® Consortium, Lelystad, Netherlands.

Van Niekerk, T.G.C.M., H. Gunnink & C.G. van Reenen, 2012. Welfare Quality® assessment protocol for laying hens, results assessment of 122 flocks. Report 589 Wageningen UR, Livestock Research. 63 pages.

Wageningen UR Livestock Research

Edelhertweg 15, 8219 PH Lelystad T 0320 238238 F 0320 238050

E info.livestockresearch@wur.nl I www.livestockresearch.wur.nl