Journal Pre-proof

Data-driven categorization of on-farm animal welfare and slaughter outcomes in slow growing chickens with outdoor access: When data speak

Francesca Leone , Aran Ruiz-Ferreras , José Luis Lavín , Valentina Ferrante , Inma Estevez , Xavier Averós

PII: S0032-5791(26)01180-6
DOI: https://doi.org/10.1016/j.psj.2026.107547
Reference: PSJ 107547

To appear in: *Poultry Science*

Received date: 13 May 2026
Accepted date: 31 July 2026

Please cite this article as: Francesca Leone , Aran Ruiz-Ferreras , José Luis Lavín , Valentina Ferrante , Inma Estevez , Xavier Averós , Data-driven categorization of on-farm animal welfare and slaughter outcomes in slow growing chickens with outdoor access: When data speak, *Poultry Science* (2026), doi: https://doi.org/10.1016/j.psj.2026.107547

Data-driven categorization of on-farm animal welfare and slaughter outcomes in slow growing chickens with outdoor access: When data speak

Francesca Leone*, Aran Ruiz-Ferreras[†], José Luis Lavín[‡], Valentina Ferrante*, Inma Estevez[†§], Xavier Averós[†1]

* Department of Environmental Science and Policy, University of Milan, Via G. Celoria 10, 20146 Milan, Italy

[†] Department of Animal Science, NEIKER-Basque Institute for Agricultural Research and Development, Basque Research and Technology Alliance (BRTA), N-104, km. 355, 01192 Arkaute, Spain

[‡] Applied Mathematics Department, NEIKER-Basque Institute for Agricultural Research and Development, Basque Research and Technology Alliance (BRTA), Bizkaia Science and Technology Park 812L, 48160 Derio, Bizkaia, Spain

[§] IKERBASQUE, Basque Foundation for Science, Alameda Urquijo 36-5 Plaza Bizkaia, 48011 Bilbao, Spain

[1] Corresponding author: xaveros@neiker.eus (X. Averós)

Scientific section: Animal Well-Being and Behavior

ABSTRACT

Animal welfare is a crucial ethical and economic aspect in modern poultry production, even in alternative systems based on slow growing strains and outdoor access. Welfare assessment is commonly performed using standardized and quantitative protocols, such as the AWIN® assessment protocol, which provide objective animal-, resource-, and management-based indicators at flock level. However, the evaluation of multiple indicators requires trained assessors and time, and flock categorization depends on the application of pre-established thresholds. This study aimed to determine whether assessed on-farm welfare indicators allow the *a priori*, data-driven classification of commercial flocks according to their welfare, and whether such classification corresponds with slaughter outcomes. 98 slow growing chicken flocks from 22 farms in Northern Spain were assessed using animal-based, resource-based, and management-based indicators from the AWIN® protocol. Hierarchical clustering was used to group flocks according to these indicators. Generalized linear mixed models were used to characterize the welfare profile of the identified clusters, and to check whether on-farm welfare profiles corresponded with slaughter outcome differences. The hierarchical clustering identified 2 welfare profiles comprising 89 (Cluster 1) and 9 (Cluster 2) flocks. Cluster 2 exhibited the poorest welfare profile, characterized by a higher prevalence of lame, immobile, tail-damaged, terminally ill and dead birds, and the highest house NH_3 concentrations. Importantly, the poorer welfare profile of Cluster 2 was associated with lower slaughter performance, including a lower percentage of Category A carcasses, a higher percentage of Category 0 carcasses, and a trend towards higher total mortality. These findings demonstrate that raw on-farm welfare data can be used to classify slow growth chicken flocks into clusters reflecting distinct welfare profiles associated with slaughter performance and carcass quality. This data-driven pre-classification provides an objective framework for flock stratification, enabling the identification of flocks with welfare deficiencies and supporting targeted management interventions to improve animal

welfare and production outcomes, with potential application to other poultry production systems and livestock species.

Keywords: Animal Welfare; Broiler Chicken; Alternative System; Hierarchical Cluster Analysis; Slaughter Outcome.

INTRODUCTION

Animal welfare is a fundamental aspect of modern livestock production due to ethical and economic implications (Alonso et al., 2020). According to the World Organization for Animal Health (WOAH), welfare includes the physical and mental state of an animal in relation to its living conditions. Optimal welfare requires meeting species-specific physiological and behavioral needs and minimizing the negative influence of stressors on health and productivity (WOAH, 2008). Poor welfare conditions compromise animal health and performance, deteriorating production efficiency, profitability, and quality of the final product (Velarde and Dalmau, 2012). On-farm welfare problems are reflected at the slaughterhouse in the form of carcass downgrades or condemnations, which have direct implications for food quality, safety, and economic returns (Fernandes et al., 2021). Beyond its direct implications for animals, high welfare standards also contribute to the added value and market differentiation of animal-derived products. At the same time, animal welfare is intrinsically linked to public health, as poor welfare conditions can increase the risk of zoonotic disease transmission, consistent with the One Health perspective (Verkuijl et al., 2024). In the European Union, growing consumer awareness and ethical concerns regarding animals´ quality of life have further boosted the relevance of animal welfare as a pivotal aspect of animal production (Koknaroglu and Akunal, 2013), which highlights the importance of providing consumers with clear and objective information on which to base their purchasing decisions.

Broiler chickens are a clear example of production intensification, where fast growth rates and high stocking densities, among other issues, have raised consumers' welfare concerns (Meluzzi and Sirri, 2009). Public perception and societal expectations on livestock management practices have contributed to a shift towards production systems that prioritize high animal welfare standards (de Jonge et al., 2013). Alternative chicken production systems based on the use of slow growing breeds, the access to a free range, and the use of lower stocking densities are now well established within the poultry sector, largely driven by sustained market demand and welfare-oriented production standards (de Jonge et al., 2013; Nicol et al., 2024). These systems also have strong socio-economic implications, as they are often associated with small-scale or family farms, deeply embedded in the rural landscape and closely connected to local traditions and territorial identity. Their development aligns with a growing societal trend toward returning to rural life and re-establishing a closer relationship with the land (Boogaard, 2009). When integrated into short supply chains or direct marketing approaches, these systems facilitate the valorization of agroecological resources, contributing to rural development, preserving traditional practices, and producing *a priori* welfare-friendly, high-quality products (Costanza et al., 2021).

The determination of animal welfare requires objective assessment independently from the production system, and a broad spectrum of indicators to evaluate animal welfare are available to this aim (Ohl and van der Staay, 2012). Resource and management-based indicators have been widely used, and although informative, they only provide a partial view of animal welfare, as they evaluate the context in which birds live, but not their direct impact on the animals themselves (Blokhuis et al., 2010). Animal-based indicators are direct measures on the animals that assess how they cope with their production and management system (Burgstaller et al., 2022). In broiler chickens, these include, among others, injuries and lameness (de Jong et al., 2012; Marchewka et al., 2015), and flock records such as body weight, disease or mortality records, even if not directly measured on individual birds (EFSA, 2012). Animal-based indicators, when combined with resource- and management-based indicators and flock records, should provide a more comprehensive flock-level assessment of the welfare state of animals.

Assessment protocols using specific indicators have been proposed. The Welfare Quality project was the first to develop standardized protocols for welfare assessment of cattle, pigs, and poultry prioritizing animal-based indicators (Welfare Quality, 2009). Protocols are based on a scientifically grounded and well-established set of welfare principles and criteria (Botreau et al., 2007), and on different indicators providing information about each of them. The application of this assessment framework offered an exhaustive picture of animal welfare. Later, the AWIN® project, further building on the same principles and criteria, developed animal-based indicators and standardized assessment protocols for additional livestock species. This included sheep, goats, horses, donkeys, and turkeys, with the poultry welfare assessment method being first developed and tested on broilers. For turkeys and broilers, the transect method was proposed as a non-invasive, quantitative, and practical sampling approach to assess broiler and turkey welfare in production settings (Marchewka et al., 2013; 2015). This approach can be used to assess welfare in commercial poultry farms by recording welfare indicators while walking predefined transects set using fixed house devices, such as feeder and drinker lines, as visual references. Protocols from both projects are obviously relevant from a scientific perspective and are currently used within animal welfare certification schemes (e.g. the Welfair® certification scheme; https://www.animalwelfair.com/en/).

Once protocols are applied, the assessor has information on many indicators and can get a broad picture of the welfare status of animals. However, interpretation of multiple indicators and classification of animal welfare into meaningful categories from on-farm raw data may be challenging, especially for farmers. Thus, it would be relevant to investigate whether all collected indicators could provide an *a priori* classification of farms according to welfare profiles that could base final assessment scores. Once a solid welfare categorization is constructed based on all indicators, a next step might be the detection of reliable iceberg indicators (key welfare indicators that can reflect, or are closely correlated with, a range of other welfare indicators (Farm Animal Welfare Council, 2009)) that achieve a similar classification performance using smaller assessment effort. This would facilitate the practical flock management by farmers, and animal welfare risk assessment with a minimal time investment, thus facilitating management decisions to ensure welfare. This would be a practical advantage regarding complete Welfare Quality and AWIN® protocols.

Slaughter outcomes can provide retrospective information about the cumulative effects of farm rearing conditions, health, and management on final product quality. Therefore, it would be highly relevant to determine the correspondence between any farm welfare pre-classification and slaughter outcomes. This would be highly relevant for the industry, evidencing the interconnexion between good on-farm welfare and system productivity, and highlighting the importance of farm welfare self-assessments. Such information would also be useful to improve welfare monitoring and to undertake preventive actions during future production cycles.

The aim of this study was to determine whether commercial, slow growth chicken flocks cluster according to indicators collected during on-farm welfare assessments providing a practical flock pre-classification, and whether welfare profiles associated with clusters have some correspondence with slaughter outcomes. We hypothesized that hierarchical clustering analysis of on-farm data collected during welfare assessments would identify flocks with different on-farm welfare profiles. We also hypothesized that the identified on-farm welfare profiles would correspond to different patterns of slaughter outcomes. For this, slow growing flocks raised with outdoor access were assessed using the AWIN® chicken welfare assessment protocol, which was selected based on its faster on-farm applicability regarding other existing protocols.

MATERIALS AND METHODS

This was an observational study conducted in commercial farm conditions that did not involve animal manipulation. All assessed flocks complied with the requirements of the Spanish legislation on the protection of chickens produced for meat purposes (Real Decreto 692/2010).

Farms and Birds

This study involved the assessment of 98 commercial slow growing chicken flocks, understood as the group of chickens housed in a single house, from 22 farms with outdoor access (each farm had 2 to 3 independent houses, each of them housing about 1,200 birds. Therefore, each farm had 2,400 to 3,600 birds). All farms were in Northern Spain and belonged to the same production company. Flocks were all females from the Sasso XL451A or RB57N strains. Females were purchased at a commercial hatchery and transported to the farms where they were housed at an initial density of 12 birds/m^2, and 25 kg/m^2 at approximately 81 days of rearing. Birds were allowed to access the outdoor area (independent area for each house) from 28 days of age. Most houses use natural ventilation and lighting provided through windows, and floors were covered with wood shavings, straw or sawdust as litter material. Free access to fresh water was guaranteed through cup drinkers, bell drinkers or nipple drinkers, and birds were fed using principally pan feeders.

Data collection

Farm data. At around 70 days of age, the welfare of birds was assessed following the AWIN® chicken welfare assessment protocol, which focuses on animal-based, resource and environmental indicators, with mental state not directly assessed. Assessments were carried out between 10:00 and 12:00 am to minimize the influence of circadian patterns on behavior. A laser distance meter (GLM 250 VF Professional, Robert Bosch GmbH, Germany) was used to measure the total length and width of the house. Given the strong reaction of birds to the red laser pointer, a manual tape measure was employed to determine the width of each transect, defined as the house area comprised between 2 visual references (normally lines of feeders and drinkers). Once defined, the observer walks along the transect collecting the number of animals, within transect limits, showing any of the predetermined welfare indicators, which are expressed relative to the total number of birds assumed to be present within the assessed transect. 2 transects were randomly selected for the assessment of animal-based indicators and environmental parameters: one located at the center of the house and the other by a side wall. To minimize the risk of double counting, at least 1 transect was left unassessed between assessed transects (Marchewka et al., 2013; BenSassi et al., 2019a). An average of 1,200 birds were assessed per flock, 250 to 300 per transect. Assessed animal-based indicators per transect included the number of chickens immobile, lame, small, featherless on the head, back, and tail, (3 featherless areas were added into a single indicator named featherless total), with head, back or tail wounds, sick, terminally ill, and dead. These indicators were proposed by the AWIN® protocol to cover all welfare principles and criteria as defined in the Welfare Quality project and were assessed as defined by Marchewka et al. (2013) and BenSassi et al. (2019b). In addition, cumulative mortality on the assessment day (%) was also registered. 3 additional animal-based indicators were also collected to reflect specific welfare aspects of the production system. The number of birds expressing active pecking to other birds and the number of birds displaying head aggression were collected to capture additional information on feather pecking and aggression, which are specific welfare problems related to the assessed bird strains and production system. The number of birds performing dust bathing was also collected as a measure of positive behavioral expression, as we expected that specific production conditions, particularly the access to outdoor areas, would promote this. The occurrence of each indicator in the house was expressed as a percentage relative to the estimated number of chickens present in each assessed transect. Final house percentages were the average value for the 2 assessed transects. Once indoor assessment was performed, assessment of birds remaining outdoors was carried out. For this the animal-based indicators were also assessed on all birds present in the outdoor area and expressed as the percentage of animals present outdoors. Final flock occurrence of each indicator was calculated by adding up house and outdoor percentages, with no weighing being applied.

Resource- and management-based indicators were collectively considered as environmental parameters in this study, and included those defined in the AWIN® auditing protocol (available upon request), namely the outdoor range area available to birds (m^2), initial density (birds/m^2), and litter quality. Temperature (T, °C), relative humidity (RH, %), and the concentration of CO_2 and NH_3 (ppm) were also collected at birds' level as spot measurements. No additional indicators were considered for inclusion in the protocol nor assessed. Litter quality was expressed as the average of 6 assessments (3 per transect, collected at the beginning, middle and end of each transect) carried out using the WQ five-point scale (Welfare Quality, 2009). At the same 6 points of litter evaluation, T, RH, and the concentration of CO_2 and NH_3 were measured at bird level, and mean values per flock were calculated. These parameters were collected with specific sensors (AZ 7755, AZ Instrument Corp., Taiwan for T, RH, and CO_2; GasAlert NH_3 Extreme, BW Technologies, Canada for NH_3).

Slaughterhouse data. The slaughterhouse parameters for each flock were obtained from the company records. Production parameters included the total mortality (number of birds dead before slaughter (either on farm and/or during transport) relative to the flock initial number of birds; %), feed conversion ratio (FCR; average flock feed consumption (kg) relative to average flock carcass weight (kg)). Carcass quality classification (% of category A, B, and 0 carcasses) based on conformation, defects, and sanitary status of carcasses, were also collected. Category A included carcasses accomplishing the regulatory scheme requirements (Hazi, 2019), i.e. carcasses with optimal body conformation, free of bruises, lesions, or feather residues, and were destined for whole-bird retail. Category B included carcasses not accomplishing with the regulatory scheme requirements (Hazi, 2019), i.e. carcasses that presented minor aesthetic defects, or deviations in shape and weight, but remained suitable for commercial distribution as cuts or processed products. Category 0 included carcasses not accomplishing with the regulatory scheme requirements (Hazi, 2019), showing significant defects, like bone fractures or extensive bruises that excluded them from the fresh meat chain yet directed to the processed meat sector. Carcass rejections included carcasses not valid for human consumption due to sanitary problems. Category A, B, and 0 carcasses, and carcass rejections were expressed relative to the total number of carcasses per flock (%).

Statistical analysis

Prior to analysis, the dataset underwent a pre-processing pipeline. First, non-numeric and identification variables were excluded to ensure that the analysis exclusively focused on welfare and slaughter indicators. To enhance the stability of the hierarchical clustering algorithm, variables with zero variance were removed to avoid the inclusion of features that offered no discriminative power (nearZeroVar function, caret package; Kuhn, 2008). Remaining variables were subsequently centered and scaled (z-score standardization) to ensure equal contribution of each variable to similarity measurements regardless of their unit of measure. The inter-correlation structure of variables was examined using a correlation matrix heatmap (corrplot package; Wei and Simko, 2017), which allowed for the identification of highly collinear indicators. This ensured that redundant variables did not disproportionately influence the cluster identification process.

Hierarchical clustering analysis (HCA) was then performed using the Ward.D2 linkage method, via the R-packages Factoextra (Kassambara and Mundt, 2020) and FactoMineR (Lê et al., 2008), chosen for its ability to minimize total within-cluster variance and produce homogeneous groups. To ensure the robustness and optimal classification of these clusters, a comprehensive multi-criteria validation framework was implemented. This included the evaluation of cluster cohesion and separation through the Silhouette index, the Gap Statistic, and the Within-Cluster

Sum of Squares (WSS) elbow method using the factoextra (Kassambara and Mundt, 2020) and cluster (Maechler et al., 2026) R-packages. Furthermore, to assess the structural reliability of the partitions, internal validation criteria were computed, including the Cophenetic Correlation (to evaluate the preservation of original distance matrices), the Dunn index, and the Davies-Bouldin and Calinski-Harabasz indexes. Finally, cluster stability was statistically verified through bootstrap resampling (100 iterations) calculating Jaccard similarity coefficients (fpc R-package; Hennig, 2024), ensuring that identified clusters were statistically reliable.

Once validated, each observation (flock) was assigned to its corresponding cluster. Subsequently, differences in farm and slaughter indicators across clusters were analyzed using the GLIMMIX procedure in SAS 9.4 (SAS Institute, NC, USA). A normal distribution was assumed for each indicator, although a *priori* suitable, alternative distributions were also tested for variables showing skewed distributions (lognormal, Gamma distributions), and for counts containing high proportions of zero values (Poisson, negative binomial distributions), such as the case of animal-based indicators, in which case the estimated number of birds per transect was used as an offset in models. Nevertheless, visual inspection of Studentized residuals (histograms, normal quantile-quantile graphics, box plots, and residual vs. predicted value graphics) showed no relevant improvements in the distribution of residuals when using alternative distributions. The cluster was treated as a fixed effect, and the farm was included as a random effect to account for nested flock assessments. Statistical significance was set at $P < 0.05$, with P-values adjusted for multiple comparisons using Tukey's range tests.

RESULTS

Farm data results

Table 1 shows the mean, standard deviation (SD), median and interquartile range for on-farm animal-based indicators and environmental parameters (initial candidates and those used during hierarchical clustering analysis, i.e. those remaining in the study once variables with zero variance were removed). The mean value of cumulative mortality and the prevalence of the assessed animal-based indicators were generally low, except for the percentage of birds with total feathering problems, which was on average high and with a large flock variability. Total feathering problems were almost exclusively due to featherless areas in the tail. Outdoor range area had considerable variability between flocks, like environmental concentrations of NH_3, CO_2 and RH. Mean NH_3 values were on average high.

Unsupervised data analysis

The preliminary analysis of variable correlation structure revealed no instances of extreme multi-collinearity, allowing all welfare indicators to be retained for the hierarchical clustering analysis (HCA). Following the Ward.D2 linkage method, the dataset was partitioned into 2 distinct clusters. This K=2 solution was supported by a comprehensive multi-criteria validation framework: the Silhouette index peaked at K=2 (Figure 1), indicating optimal separation and cohesion. Furthermore, the structural validity of this partition was confirmed by a Cophenetic correlation of 0.694, a Dunn index of 0.264, a Davies-Bouldin index of 2.078, and a Calinski-Harabasz index of 12.502, all of which collectively substantiate the robustness of the identified groups.

The stability of this 2-cluster solution was further reinforced by bootstrap resampling (100 iterations), yielding a mean Jaccard similarity coefficient of 0.834 and confirming high cluster reliability. Finally, the Principal Component Analysis (PCA) projection visually demonstrated 2 well-defined and differentiated clusters in the multidimensional space, consistently capturing the variation in welfare and slaughter outcomes across the studied population (Figure 2).

Figure 3 shows the 2-dimensional dendrogram resulting from the hierarchical clustering analysis using Ward's minimum variance method. This analysis identified 2 clusters of flocks: Cluster 1 included most flocks (89 out of 98 total flocks), while Cluster 2 comprised 9 flocks (out of 98 flocks). Flocks classified in Cluster 2 corresponded to 5 farms, and only 1 farm had all its flocks classified in this cluster.

To provide a descriptive characterization of on-farm welfare profiles according to clusters, Table 2 presents the animal-based indicators mean values (± SE) for each cluster, as well as statistical comparisons between clusters. Cluster 2 was characterized by higher percentages of immobile, lame, total featherless, featherless in tail, tail wounds, terminally ill and dead birds, Cluster 2 also tended to have higher percentages of small birds.

With the same aim as Table 2, environmental parameters are reported in Table 3. The outdoor range area and NH_3 levels were significantly higher for Cluster 2.

Slaughterhouse data results

External validation of on-farm welfare results was sought through determining whether on-farm welfare clusters had any correspondence with slaughter outcomes. As shown in Table 4, Cluster 2 had smaller percentages of Category A carcasses and higher percentages of Category 0 carcasses. Cluster 2 also tended to have higher total mortality percentages.

DISCUSSION

Public awareness and consumer demand for higher welfare have promoted alternative poultry systems, typically based on slow growing strains with outdoor access to encourage natural behavior. While these systems are characterized by lower indoor stocking densities, longer production cycles, and promote exploration and movement (EFSA AHAW Panel, 2023), welfare issues can still arise (Relić et al., 2019; Alonso et al., 2020). Different protocols to monitor and certify animal welfare exist, but their practical application is often hindered by complexity, high time investment, and training requirements. This study found that welfare indicators from the AWIN® protocol classify flocks into distinguishable groups according to welfare profiles, and that these on-farm welfare profiles correspond to relevant differences in slaughter outcomes. This confirms our hypotheses and highlights the potential of data-driven approaches for on-farm animal welfare assessment. Animal welfare is inherently multidimensional (Botreau et al., 2007), and the simultaneous evaluation of different animal-, resource- and environment-based indicators may provide a more comprehensive representation of flock welfare than considering individual measures in isolation (Welfare Quality, 2009). In this context, the strength of hierarchical clustering as an unsupervised machine learning approach is that, by exploring data intrinsic structure, it allows the identification of welfare profiles as they naturally occur among commercial flocks rather than relying on predefined welfare categories. From a practical perspective, this approach provides an objective framework for flock stratification, allowing to identify groups of flocks sharing similar welfare characteristics, thereby facilitating targeted management decisions. Importantly, the identified welfare profiles were associated with differences in slaughter outcomes, suggesting that the resulting clusters capture biologically meaningful variation in on-farm welfare rather than simply reflecting pure statistical data partitioning. This capacity to summarize complex welfare information into interpretable welfare profiles aligns with the broader objectives of Precision Livestock Farming, where multivariate data are increasingly used to support evidence-based management decisions (Berckmans, 2014, 2017; Tullo et al., 2019). Beyond identifying welfare profiles, data-driven approaches may also contribute to simplifying welfare assessment protocols. Machine-learning algorithms have recently been used to identify subsets of welfare indicators capable of maintaining the predictive performance of complete assessment protocols

(Verós et al., 2024). Simplified protocols based on selected indicators could facilitate routine welfare monitoring by producers, enabling the detection of welfare deficiencies and supporting corrective interventions.

Flock general welfare situation

The starting population characteristics were described to contextualize the welfare profiles later identified through clustering. It is relevant to note that from the 5 farms belonging to Cluster 2, only in 1 case all assessed flocks were classified in that cluster. Given this, and the fact of having considered the farm as a random effect in our models, we may discard the farm effect as a factor affecting the obtained clustering pattern. The overall low prevalence of lameness and immobility in the study population may be explained by the superior skeletal integrity of slow growing genotypes (Castellini et al., 2016), in addition to the lower densities and outdoor access that promote bird movement and expression of more species-specific behaviors, and therefore bone strength (Relić et al., 2019). The prevalence of dust bathing was limited during assessments. Data on the prevalence of dust bathing in slow growth chickens with outdoor access is scarce. Values in the study of Rodriguez-Aurrekoetxea et al. (2015) were very small; but in commercial broilers offered different substrates prevalence varied from 0.5 to 28% (Baxter et al., 2018), higher than that observed in our case. This highly motivated behavior, considered a key positive welfare indicator (Papageorgiou et al., 2023), is essential for feather maintenance and overall health (Relić et al., 2019). However, its expression also depends on environmental conditions, and particularly on the availability and characteristics of substrate (Baxter et al., 2018) as well as to human presence (EFSA AHAW Panel, 2023). In addition, the transect method relies on spot observations, which inherently reduces the probability of observing behaviors that are infrequent or subjected to circadian patterns (Marchewka et al., 2013).

On-farm cumulative mortality remained below the 5% benchmark reported for intensive systems with fast growth lines (Manning et al., 2007). Flocks were characterized by a low prevalences of sick, terminally ill and dead birds. Overall, this reflects the greater robustness typically described for slow growing strains reared in outdoor systems (Castellini et al., 2016; Sarica et al., 2020; Nicol et al., 2024).

Despite the low incidence of welfare issues, some problems associated with feather pecking were detected. On average, 3.57% of birds were affected by feather damage located almost exclusively on the tail area, most likely caused by feather pecking. Feather pecking is difficult to prevent even in flocks reared at low densities and with access to outdoor areas (Relić et al., 2019). Its occurrence has been associated with social stress, limited foraging opportunities, and suboptimal litter quality (Dixon, 2008; Rodenburg et al., 2013), none of which are likely to occur in alternative production systems with reduced flock size, low density and plenty of well-designed outdoor areas. Genetic predisposition (Meuser et al., 2021) and longer production cycles may be risk factors affecting chances of developing this problem. Therefore, tail feather loss represents a meaningful welfare indicator to be considered for early risk identification of feather pecking.

Indoor environmental conditions were relevant since houses depended on natural ventilation and lighting. Temperature, RH, CO_2 concentration and litter quality remained within acceptable ranges (< 3,000 ppm for CO_2). Contrarily, NH_3 concentrations showed a large variability (13.16±10.73), with some flocks being exposed to inadequate air quality (> 20 ppm, Council Directive 2007/43/EC) which might have impacted on welfare status, although the nature of our study does not allow to establish causality. In addition, environmental measurements in the present study were punctually carried out during welfare assessments, thus it is unknown how prolonged flock exposure to elevated NH_3 levels was. However, Drake et al. (2010) indicated that NH_3 levels above 20 ppm can cause physiological stress and trigger the expression of stress-related behaviors (Miles et al., 2004). Continuous environmental measurements would help to better determine the actual exposure of birds to elevated NH_3 conditions, and its impact on welfare.

Flock classification

After describing the starting population, a hierarchical clustering analysis was performed to determine whether different on-farm welfare profiles could be determined based on collected data before any formal welfare categorization. This methodological approach was powerful as it allowed data itself to define the inner structure of welfare heterogeneity rather than imposing predefined categories (Blokhuis et al., 2010), although it must be acknowledged that final clusters depend on the chosen clustering methodology. The analysis identified 2 independent clusters differing in size, as shown in Figure 3, a number of clusters supported by the Silhouette index (Figure 1). Most flocks were concentrated in a dense plot region corresponding to Cluster 1, with 8 out of the 9 flocks comprising Cluster 2 remaining distant from those of Cluster 1 (Figure 2). Cluster differentiation was due to variations in both animal-based and environmental indicators. Understanding the welfare basis of this separation requires a closer examination of the indicators that likely drove flock clustering.

Cluster 1 grouped flocks with low prevalence of problems as measured through animal-based indicators. Prevalences of featherless and lame chickens were lower than those of Cluster 2, while immobile, terminally ill, dead birds, and birds with tail wounds were nearly absent. The percentage of small birds also tended to be smaller in Cluster 1. This pattern suggests that some animal-based indicators might have high discriminant power. Featherless birds and birds with tail wounds appear to be particularly sensitive, as they sharply distinguish both clusters and reflect behavioral issues such as feather pecking. Despite their low prevalence immobile, lame, terminally ill and dead birds also contribute to cluster differentiation because they occurred at disproportionate rates in Cluster 2. Conversely, small birds marginally contributed to the differentiation of clusters. This structured separation according to many of the studied indicators explains why the cluster analysis could successfully classify flocks *a priori* and suggests the potential to define a subset of animal-based indicators that should indeed suffice to reveal differences in flock welfare. A systematic definition of which subset of welfare indicators would best predict the outcomes of on-farm welfare assessments would therefore be relevant and the natural next step but is out of the scope of the present study.

Aligned with this, the environmental parameters, basically NH_3, confirmed Cluster 1 as that with the best environmental conditions. This exhibited the most favorable environmental profile, with NH_3 concentrations averaging 11.60 ppm, well below the 20-ppm limit established by Council Directive 2007/43/EC. Although, as previously mentioned, the observational nature of our study does not allow to establish causality between animal-based and environmental indicators, it must be acknowledged that prolonged exposure to elevated NH_3 concentrations may compromise respiratory function, heighten susceptibility to infectious diseases, and induce stress-related behaviors (Relić et al., 2019; Miles et al., 2004).

Environmental parameters were based on spot measurements taken during welfare assessments; thus, they did not capture the full temporal air quality dynamics. However, association between elevated NH_3 concentrations and increased prevalence of animal-based indicators allows us to hypothesize that environmental values of Cluster 2 were not transient fluctuations. High NH_3 levels might have been due to management and/or underlying structural ventilation differences among farms that likely persisted throughout the production cycle, and the trends observed for litter quality and RH would support this. They might also be due to microclimate variations associated with farm locations. Nevertheless, these aspects were not considered in our study, and so strong conclusions cannot be drawn. In naturally ventilated systems, indoor air exchange is strongly determined by wind speed and the thermal gradient between indoor and outdoor air, making ventilation efficiency highly sensitive to the geographical location and house orientation (Rojano et al., 2019). Under unfavorable conditions, these factors may hinder effective gas removal, resulting in stagnant air zones in which NH_3 produced through microbial decomposition of uric acid in the litter accumulates (Seedorf et al., 1998). Although non-significant differences between clusters were detected, the slightly higher RH recorded in Cluster 2 might have exacerbated this process, as litter moisture promotes anaerobic microenvironments that accelerate NH_3 volatilization. Interestingly, it has been recently found that temporal dynamics of NH_3 and CO_2 within poultry houses differ (Heck et al., 2026), which might explain discrepancies between NH_3 and CO_2 levels. In naturally ventilated houses, especially those with unfavorable orientations to seasonal winds, natural airflow may be insufficient to disperse accumulated gases efficiently, even when animals are kept at low densities (Weaver and Meijerhof, 1990). In fact, achieving sufficient control over adverse environmental conditions can be inherently challenging under natural ventilation.

Regarding the outdoor range, results suggest that total available space was a less decisive factor than indoor environmental conditions, which is not surprising as all farms provide a minimum of 2 m^2/bird. Given the results of Cluster 2, it seems that any potential positive influence of the additional outdoor space was overridden by indoor conditions (Campbell et al., 2025). The use of the outdoor area depends on many factors like the presence of shelters and shade, or weather conditions as birds are known to prefer warmer, less windy conditions (Rodriguez-Aurrekoetxea et al., 2015; Rodriguez-Aurrekoetxea, 2016; Sztandarski et al., 2021).

Overall, results confirm that multivariate data collected during on-farm welfare assessments contain sufficient inherent structure to allow an *a priori* categorization of flocks according to different welfare profiles. Clusters did not arise from predefined thresholds or imposed criteria but directly emerged from the natural data variability of animal-based welfare indicators and environmental parameters. In this sense, Cluster 2 concentrated both animal-based and environmental issues. These results support the notion that collected animal-based welfare indicators and environmental parameters could be used to pre-classify farms according to on-farm welfare.

Another objective of the study was to assess whether the resulting on-farm welfare profiles had some correspondence with slaughter outcomes. Given that slaughter outcomes were not used during the hierarchical clustering analysis, this served as an external validation for the on-farm clustering approach. Results confirm this correspondence and validate the detected on-farm welfare profiles, since analyses revealed that slaughter differences across clusters are fully consistent with these. Total mortality tended to be higher for Cluster 2, indicating that poorer on-farm welfare conditions tend to be associated with higher losses at the end of the production cycle, which agrees with available literature (EFSA AHAW Panel, 2023). Carcass classification further reinforced the connection between welfare and slaughter outcomes, as the proportion of category A carcasses was lowest and category 0 highest in Cluster 2, suggesting that sub-optimal on-farm welfare produces cumulative effects that persist throughout the production cycle and affect final product quality, as already described (Wigham et al., 2019). This fact validates the results of the hierarchical clustering approach taken with the on-farm welfare indicators, which has strong practical implications. Nevertheless, we must acknowledge that our results are contextualized within the production and management specificities of the company we worked with, and so they should not be globally extrapolated. Increasing the volume of information with flocks raised and assessed under different regional and productive conditions should broaden the applicability of our results to a more global context.

Convergence of results among animal-based indicators, environmental parameters, and slaughterhouse outcomes suggests that signs of impaired welfare during rearing can predict negative consequences at the slaughterhouse. This highlights the importance of integrating welfare assessment into routine monitoring systems to identify different welfare profiles, and to trigger targeted measures to improve on-farm welfare that will benefit slaughter productive efficiency. However, the question of when such on-farm interventions can be effectively implemented is critical. If flocks are evaluated at around 70 days, risk factors are detected only late in the cycle, but we are currently lacking the data to determine when the AWIN® protocol should be first used on slow growth flocks to trigger corrective measures in case early welfare issues are detected. Data from fast growth flocks have shown that animal-based indicators with a potentially high discriminative power should be monitored at earlier stages (BenSassi et al., 2019b, c) to permit a timely identification of flocks at higher risk of welfare degradation and enable interventions to prevent further deterioration. Integrating these indicators into routine flock monitoring frameworks for slow growing broilers may enhance the capacity to detect emerging problems, supporting the notion of a more proactive, effective flock management.

CONCLUSION

The present study demonstrates that welfare indicators collected during on-farm welfare assessments can be used to obtain a data-driven pre-classification of slow growing chicken flocks with outdoor access according to differing welfare profiles. Hierarchical clustering allowed the identification of 2 clusters with varying on-farm welfare levels as measured through animal-based indicators and environmental parameters. This clustering indicated that the information collected during welfare assessments already contains sufficient structure to discriminate flocks with different welfare profiles, even before the application of any formal welfare score to data. Importantly, on-farm welfare profiles were consistently associated with slaughter outcomes, with poor on-farm welfare flocks also showing a lower proportion of top-quality carcasses and tending to have higher total mortality. This reinforces the relevance of good on-farm welfare not only for the animals, but also for its economic implications along the production chain. Overall, the results support the use of hierarchical clustering analysis as a complementary tool to welfare assessment protocols, providing with science-based arguments to identify flocks with welfare issues and to support targeted, corrective actions to improve on-farm welfare.

ACKNOWLEDGMENTS

We acknowledge Lumagorri, their technical staff and all involved farmers for allowing access to their farms for on-farm data collection, for access to their slaughter records, and for trust and continuous, positive collaboration. We also acknowledge the MCIN/AEI/https://doi.org/10.13039/501100011033/ and the EU NextGenerationEU/PRTR for funding the D-AWARE project (R+D+I Project PDC2022-133081-C21). We also acknowledge the PhD School in Environmental Sciences of the University of Milan for supporting the stay of Francesca Leone at NEIKER-Basque Institute for Agricultural Research and Development.

DECLARATION OF COMPETING INTERESTS

All authors declare that there are no conflicts of interest regarding the publication of this article.

CRediT AUTHORSHIP CONTRIBUTION STATEMENT

Francesca Leone: Conceptualization, Methodology, Formal analysis, Investigation, Data Curation, Writing – Original Draft. **Aran Ruiz-Ferreras:** Investigation, Data Curation, Writing – Review & Editing. **Jose Luis Lavín:** Conceptualization, Methodology, Software, Formal analysis, Writing- Review & Editing. **Valentina Ferrante:** Conceptualization, Writing –

Review & Editing. **Inma Estevez:** Conceptualization, Writing – Review & Editing, Funding acquisition. **Xavier Averós:** Conceptualization, Methodology, Formal analysis, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

REFERENCES

Alonso, M. E., J. R. González-Montaña, and J. M. Lomillos. 2020. Consumers' Concerns and Perceptions of Farm Animal Welfare. Animals. 10:385. doi:10.3390/ani10030385

Averós, X., J. L. Lavín, J. L., and I. Estevez, I. 2024. The potential of decision trees as a tool to simplify broiler chicken welfare assessments. Scientific Reports 14 : 22943. https://doi.org/10.1038/s41598-024-74260-4

Baxter, M., C. L. Baillie, and N. E. O'Connell. 2018. An evaluation of potential dustbathing substrates for commercial broiler chickens. Animal 12:1933-1941. doi: 10.1017/S1751731117003408

BenSassi, N., X. Averós, and I. Estevez. 2019a. Broiler chickens on-farm welfare assessment: estimating the robustness of the transect sampling method. Frontiers in Veterinary Science 6:236. doi: 10.3389/fvets.2019.00236

BenSassi, N., J. Vas, G. Vasdal, X. Averós, I. Estevez, and R. C. Newberry. 2019b. On-farm broiler chicken welfare assessment using transect sampling reflects environmental inputs and production outcomes. PLoS ONE. 14:e0214070. doi:10.1371/journal.pone.0214070

BenSassi, N., X. Averós, and I. Estevez. 2019c. The potential of the transect method for early detection of welfare problems in broiler chickens. Poultry Science 98:522–532. doi:10.3382/ps/pey374

Berckmans, D. 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. Revue Scientifique et Technique 33:189–196. http://dx.doi.org/10.20506/rst.33.1.2273

Berckmans, D. 2017. General introduction to precision livestock farming. Animal Frontiers 7:6–11. https://doi.org/10.2527/af.2017.0102

Blokhuis, H. J., I. Veissier, M. Miele, and B. Jones. 2010. The Welfare Quality® project and beyond: Safeguarding farm animal well-being. Acta Agriculturae Scandinavica, Section A - Animal Science. 60:129–140. doi:10.1080/09064702.2010.523480

Boogaard, B. K. 2009. The socio-cultural sustainability of animal farming : an inquiry into social perceptions of dairy farming in the Netherlands and Norway. PhD Diss. Wageningen University.

Botreau, R., I. Veissier, A. Butterworth, M. B. M. Bracke, and L. J. Keeling. 2007. Definition of criteria for overall assessment of animal welfare. Animal Welfare. 16:225. doi:10.1017/S0962728600031390

Burgstaller, J., T. Wittek, N. Sudhaus-Jörn, and B. Conrady. 2022. Associations between Animal Welfare Indicators and Animal-Related Factors of Slaughter Cattle in Austria. Animals. 12:659. doi:10.3390/ani12050659

Campbell, Y. L., L. L. Walker, B. M. Bartz, J. O. Eckberg, and A. N. Pullin. 2025. Outdoor access versus conventional broiler chicken production: Updated review of animal welfare, food safety, and meat quality. Poult. Sci. 104:104906. doi:10.1016/j.psj.2025.104906

Castellini, C., C. Mugnai, L. Moscati, S. Mattioli, M. Guarino Amato, A. Cartoni Mancinelli, and A. Dal Bosco. 2016. Adaptation to organic rearing system of eight different chicken genotypes: behaviour, welfare and performance. Italian Journal of Animal Science. 15:37–46. doi:10.1080/1828051X.2015.1131893

Costanza, M., Sannino, M., Ciani, E., Crecchio, C., and Marotta, C. 2021. Poultry farming diversification into short supply chains: a sustainable strategy to enhance rural development and consumer trust. Sustainability, 13:8565. https://doi.org/10.3390/su13158565

de Jong, I., C. Berg, A. Butterworth, and I. Estevez. 2012. Scientific report updating the EFSA opinions on the welfare of broilers and broiler breeders. EFS3. 9. doi:10.2903/sp.efsa.2012.EN-295

De Jonge, J., and H. C. M. Van Trijp. 2013. The impact of broiler production system practices on consumer perceptions of animal welfare. Poultry Science 92:3080–3095. doi:10.3382/ps.2013-03334

Dixon, L. M. 2008. Feather Pecking Behaviour and associated Welfare issues in Laying Hens. Avian Biology Research. 1:73–87. doi:10.3184/175815508X363251

Drake, K. A., C. A. Donnelly, and M. S. Dawkins, M.S. 2010. Influence of rearing and lay risk factors on propensity for feather damage in laying hens. British Poultry Science 51:725-733. doi: 10.1080/00071668.2010.528751

EFSA Panel on Animal Health and Welfare (AHAW). 2012. Scientific Opinion on the use of animal-based measures to assess welfare of broilers. EFS2. 10. doi:10.2903/j.efsa.2012.2774

EFSA AHAW Panel (EFSA Panel on Animal Health and Welfare). 2023. Welfare of broilers on farm. EFS2. 21. doi:10.2903/j.efsa.2023.7788

European Council. 2007. Council Directive 2007/43/EC of 28 June 2007 laying down minimum rules for the protection of chickens kept for meat production. Official Journal of the European Union. L 182:19–28.

Farm Animal Welfare Council. 2009. Farm animal welfare in Great Britain: Past, present and future. Farm Animal Welfare Council, London, 57 pp.

Fernandes, J. N., P. H. Hemsworth, G. J. Coleman, and A. J. Tilbrook. 2021. Costs and Benefits of Improving Farm Animal Welfare. Agriculture. 11:104. doi:10.3390/agriculture11020104

Heck, J., E. Rauch, A. Piller, B. Schiller, M. Erhard, P. Schmidt, and H. Louton. 2026. Levels of two pollutant gases, temperature, and relative humidity and their influence on broiler welfare. Journal of Applied Poultry Research 35:100694. https://doi.org/10.1016/j.japr.2026.100694

Hennig, C. 2024. fpc: Flexible Procedures for Clustering. R package version 2.2-13. https://CRAN.R-project.org/package=fpc

Kassambara, A., and F. Mundt. 2020. Factoextra: extract and visualize the results of multivariate data analyses. R Package Version 1.0.7. https://CRAN.R-project.org/package=factoextra

Koknaroglu, H., and T. Akunal. 2013. Animal welfare: An animal science approach. Meat Science. 95:821–827. doi:10.1016/j.meatsci.2013.04.030

Kuhn, M. 2008. Building Predictive Models in R Using the caret Package. Journal of Statistical Software. 28:1–26. https://doi.org/10.18637/jss.v028.i05

Lê, S., J. Josse, and F. Husson. 2008. FactoMineR: a package for multivariate analysis. Journal of Statistical Software. 25:1–18. doi:10.18637/jss.v025.i01

Maechler, M., P. Rousseeuw, A. Struyf,, M. Hubert, and K. Hornik, K. 2026. cluster: Cluster Analysis Basics and Extensions (R package version 2.1.8.2). https://CRAN.R-project.org/package=cluster

Manning, L., S. A. Chadd, and R. N. Baines. 2007. Key health and welfare indicators for broiler production. World's Poultry Science Journal. 63:46–62. doi:10.1017/S0043933907001262

Marchewka, J., T. T. N. Watanabe, V. Ferrante, and I. Estevez. 2013. Welfare assessment in broiler farms: Transect walks versus individual scoring. Poult. Sci. 92:2588–2599. doi:10.3382/ps.2013-03229

Marchewka, J., I. Estevez, G. Vezzoli, V. Ferrante, and M. M. Makagon. 2015. The transect method: a novel approach to on-farm welfare assessment of commercial turkeys. Poult. Sci. 94:7–16. doi:10.3382/ps/peu026

Meluzzi, A., and F. Sirri. 2009. Welfare of broiler chickens. Italian Journal of Animal Science. 8:161–173. doi:10.4081/ijas.2009.s1.161

Meuser, V., L. Weinhold, S. Hillemacher, and I. Tiemann. 2021. Welfare-Related Behaviours in Chickens: Characterization of Fear and Exploration in Local and Commercial Chicken Strains. Animals. 11:679. doi:10.3390/ani11030679

Miles, D. M., Branton, S. L., and Lott, B. D. 2004. Atmospheric ammonia is detrimental to the performance of modern commercial broilers. Poult. Sci. 83:1650–1655. doi: 10.1093/ps/83.10.1650

Nicol, C. J., S. M. Abeyesinghe, and Y.-M. Chang. 2024. An analysis of the welfare of fast-growing and slower-growing strains of broiler chicken. Front. Anim. Sci. 5:1374609. doi:10.3389/fanim.2024.1374609

Ohl, F., and F. J. Van Der Staay. 2012. Animal welfare: At the interface between science and society. The Veterinary Journal. 192:13–19. doi:10.1016/j.tvjl.2011.05.019

Papageorgiou, M., M. Goliomytis, O. Tzamaloukas, D. Miltiadou, and P. Simitzis. 2023. Positive Welfare Indicators and Their Association with Sustainable Management Systems in Poultry. Sustainability. 15:10890. doi:10.3390/su151410890

Real Decreto 692/2010. 2010. Por el que se establecen las normas mínimas para la protección de los pollos destinados a la producción de carne y se modifica el Real Decreto 1047/1994, de 20 de mayo, relativo a las normas mínimas para la protección de terneros. BOE de 3 de junio 135:47986–47995. https://www.boe.es/eli/es/rd/2010/05/20/692

Relić, R., E. SossiDou, A. Dedousi, L. Perić, İ. Božičković, and M. Đukić-Stojčić. 2019. Behavioral and health problems of poultry related to rearing systems. Ankara Üniversitesi Veteriner Fakültesi Dergisi. 66:423–428. doi:10.33988/auvfd.597496

Rodenburg, T. B., M. M. Van Krimpen, I. C. De Jong, E. N. De Haas, M. S. Kops, B. J. Riedstra, R. E. Nordquist, J. P. Wagenaar, M. Bestman, and C. J. Nicol. 2013. The prevention and control of feather pecking in laying hens: identifying the underlying principles. World's Poultry Science Journal. 69:361–374. doi:10.1017/S0043933913000354

Rodriguez-Aurrekoetxea, A., E. H. Leone, and I. Estevez. 2015. Effects of panels and perches on the behaviour of commercial slow-growing free-range meat chickens. Applied Animal Behaviour Science 173: 111–122. doi:10.1016/j.applanim.2015.02.004

Rodriguez-Aurrekoetxea, A. 2016. The behaviour and use of space of the domestic fowl in alternative meat and egg production systems. PhD Thesis. Universidad del País Vasco/Euskal Herriko Unibertsitatea, Spain.

Rojano, F., P. E. Bournet, M. Hassouna, P. Robin, M. Kacira, and C. Y. Choi. 2019. Modelling the impact of air discharges caused by natural ventilation in a poultry house. Biosystems Engineering 180:168–181. doi:10.1016/j.biosystemseng.2019.02.001

Sarica, M., U. S. Yamak, M. A. Boz, K. Erensoy, E. Cilavdaroglu, and M. Noubandiguim. 2020. Performance of fast, medium and slow growing broilers in indoor and free-range production systems. SA J. An. Sci. 49:1127–1138. doi:10.4314/sajas.v49i6.16

Seedorf, J., J. Hartung, M. Schröder, K. H. Linkert, S. Pedersen, H. Takai, J. O. Johnsen, J. H. M. Metz, P. W. G. Groot Koerkamp, G. H. Uenk, V. R. Phillips, M. R. Holden, R. W. Sneath, J. L. L. Short, R. P. White, and C. M. Wathes. 1998. A Survey of Ventilation Rates in Livestock Buildings in Northern Europe. Journal of Agricultural Engineering Research 70:39–47. doi:10.1006/jaer.1997.0274

Sztandarski, P., J. Marchewka, F. Wojciechowski, A. B. Riber, S. Gunnarsson, and J. O. Horbańczuk. 2021. Associations between weather conditions and individual range use by commercial and heritage chickens. Poult. Sci. 100:101265. doi:10.1016/j.psj.2021.101265

Velarde, A., and A. Dalmau. 2012. Animal welfare assessment at slaughter in Europe: Moving from inputs to outputs. Meat Science. 92:244–251. doi:10.1016/j.meatsci.2012.04.009

Verkuijl, C., J. Smit, J. M. H. Green, R. E. Nordquist, J. Sebo, M. N. Hayek, and M. J. Hötzel. 2024. Climate change, public health, and animal welfare: towards a One Health approach to reducing animal agriculture's climate footprint. Front. Anim. Sci. 5:1281450. doi:10.3389/fanim.2024.1281450

Weaver, W. D., and R. Meijerhof. 1991. The Effect of Different Levels of Relative Humidity and Air Movement on Litter Conditions, Ammonia Levels, Growth, and Carcass Quality for Broiler Chickens. Poultry Science 70:746–755. doi:10.3382/ps.0700746

Wei, T., and V. Simko, V. 2017. R package "corrplot": Visualization of a Correlation Matrix. https://github.com/taiyun/corrplot

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

Wigham, E., A. Grist, S. Mullan, S. Wotton, and A. Butterworth. 2019. The influence of welfare training on bird welfare and carcass quality in two commercial poultry primary processing plants. Animals. 9:584. doi:10.3390/ani9080584

SITOGRAPHY

European Commission, 7th Framework Programme. 2011-2015. AWIN - Animal Welfare Indicators. Accessed April 2025. https://cordis.europa.eu/project/id/266213

WOAH - World Organisation for Animal Health. Animal welfare. Accessed March 2025. https://www.woah.org/en/what-we-do/animal-health-and-welfare/animal-welfare/#:~:text=According%20to%20the%20Terrestrial%20Code,socio%2Deconomic%20and%20ecological%20systems

Table 1. Summary descriptive statistics (mean, standard deviation (SD), median, and interquartile range) for the candidate animal-based, resource-based and environmental indicators. Indicators marked with an asterisk were included in the clustering analysis.

Indicators	Mean	SD	Median	Interquartile range
Animal-based indicators				
Cumulative mortality* (%)	1.96	1.33	1.58	1.25
Immobile* (%)	0.02	0.06	0.00	0.00
Lame* (%)	0.16	0.35	0.00	0.19
Small* (%)	0.09	0.26	0.00	0.00
Featherless total* (%)	3.58	5.95	1.87	3.07
Featherless in head (%)	0.00	0.00	0.00	0.00
Featherless in back (%)	0.02	0.10	0.00	0.00
Featherless in tail (%)	3.55	5.96	1.85	3.12
Back wounds* (%)	0.00	0.02	0.00	0.00
Tail wounds* (%)	0.03	0.17	0.00	0.00
Head aggression (%)	0.00	0.02	0.00	0.00
Active feather pecking (%)	0.00	0.03	0.00	0.00
Sick* (%)	0.02	0.08	0.00	0.00
Terminally ill* (%)	0.00	0.02	0.00	0.00
Dead* (%)	0.01	0.03	0.00	0.00
Dust bathing* (%)	0.16	0.83	0.00	0.00
Resource-based and environmental indicators				
Outdoor range area (m^2)	3,242.00	1,704.00	2,696.00	544.00
Initial density* (n/m^2)	12.60	2.65	12.14	2.10
Litter quality* (WQ 5-point score)	2.29	0.66	2.33	1.00
Temperature* (T; °C)	15.28	3.46	14.53	3.69
Relative humidity* (RH; %)	67.78	10.80	69.11	12.34
Ammonia* (NH_3; ppm)	13.16	10.73	10.33	10.41
Carbon dioxide* (CO_2; ppm)	1,649.60	622.79	1,580.67	796.42

Table 2. Mean values (± SE) of animal-based indicators for each cluster.

Animal-based indicators	Cluster 1		Cluster 2		P-value
	Mean	*SE*	*Mean*	*SE*	
Cumulative mortality (%)	1.98	0.21	2.72	0.49	0.1220
Immobile (%)	0.00	0.01	0.13	0.02	<0.0001
Lame (%)	0.08	0.04	0.90	0.10	<0.0001
Small (%)	0.07	0.03	0.25	0.09	0.0677
Featherless total (%)	3.32	1.09	6.21	1.68	0.0413
Featherless in head (%)	0.00	0.00	0.00	0.00	0.7523
Featherless in back (%)	0.02	0.01	0.00	0.04	0.5149
Featherless in tail (%)	3.30	1.09	6.22	1.68	0.0388
Back wounds (%)	0.00	0.00	0.00	0.01	0.5778
Tail wounds (%)	0.01	0.02	0.26	0.05	<0.0001
Head aggression (%)	0.00	0.00	0.00	0.06	0.7589
Active feather pecking (%)	0.00	0.00	0.00	0.01	0.7655
Sick (%)	0.02	0.01	0.02	0.03	0.9525
Terminally ill (%)	0.00	0.00	0.02	0.01	0.0044
Dead (%)	0.00	0.00	0.05	0.01	<0.0001
Dust bathing (%)	0.19	0.12	0.10	0.32	0.7700

Table 3. Mean values (± SE) for resource-based and environmental indicators for each cluster.

Environmental/resource indicators	**Cluster 1**		**Cluster 2**		**P-value**
	Mean	*SE*	*Mean*	*SE*	
Outdoor range area (m^2)	3,066.00	201.00	5,107.00	580.00	0.0010
Initial density (n/m^2)	13.00	0.59	12.60	0.83	0.5739
Litter quality (WQ 5-point score)	2.27	0.08	2.50	0.24	0.3551
Temperature (T; °C)	15.30	0.52	15.00	1.28	0.8692
Relative humidity (RH; %)	67.30	1.36	71.30	3.93	0.3268
Ammonia (NH_3; ppm)	11.60	1.42	27.80	3.78	<0.0001
Carbon dioxide (CO_2; ppm)	1,633.00	88.60	1,751.00	233.00	0.6194

Table 4. Mean values (± SE) for slaughterhouse parameters for each cluster.

Slaughterhouse parameters	Cluster 1		Cluster 2		P-value
	Mean	*SE*	*Mean*	*SE*	
Total mortality (%)	2.33	0.29	3.26	0.57	0.0874
Feed Conversion Ratio	3.53	0.03	3.62	0.06	0.1183
Category A carcasses (%)	80.10	0.72	75.40	1.72	0.0069
Category B carcasses (%)	16.80	0.65	17.50	1.50	0.6345
Category 0 carcasses (%)	2.96	0.49	7.12	1.35	0.0036
Carcass rejections (%)	0.07	0.01	0.09	0.04	0.5881

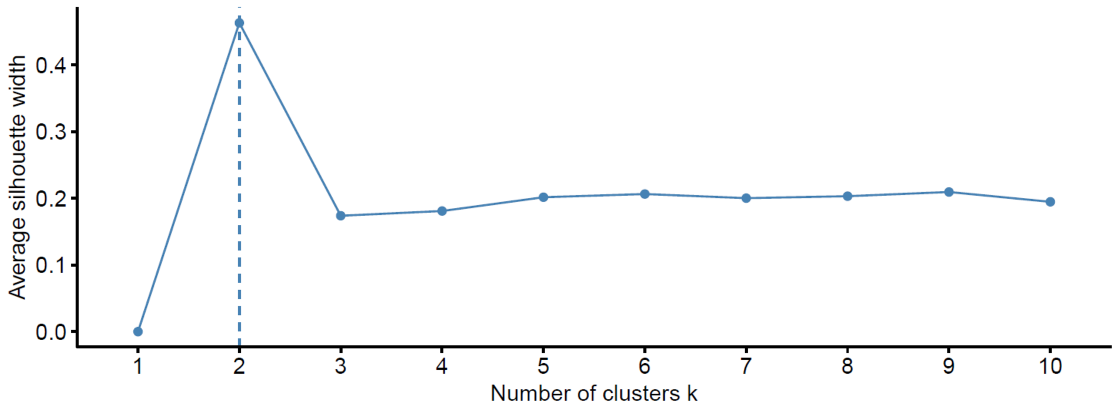

Figure 1. Silhouette analysis for determining the optimal number of clusters.

Caption: average silhouette width obtained for clustering solutions ranging from k = 1 to k = 10. The highest average silhouette width was observed for k = 2 (dashed vertical line), indicating that a 2-cluster solution provided the best separation and cohesion among flock welfare profiles.

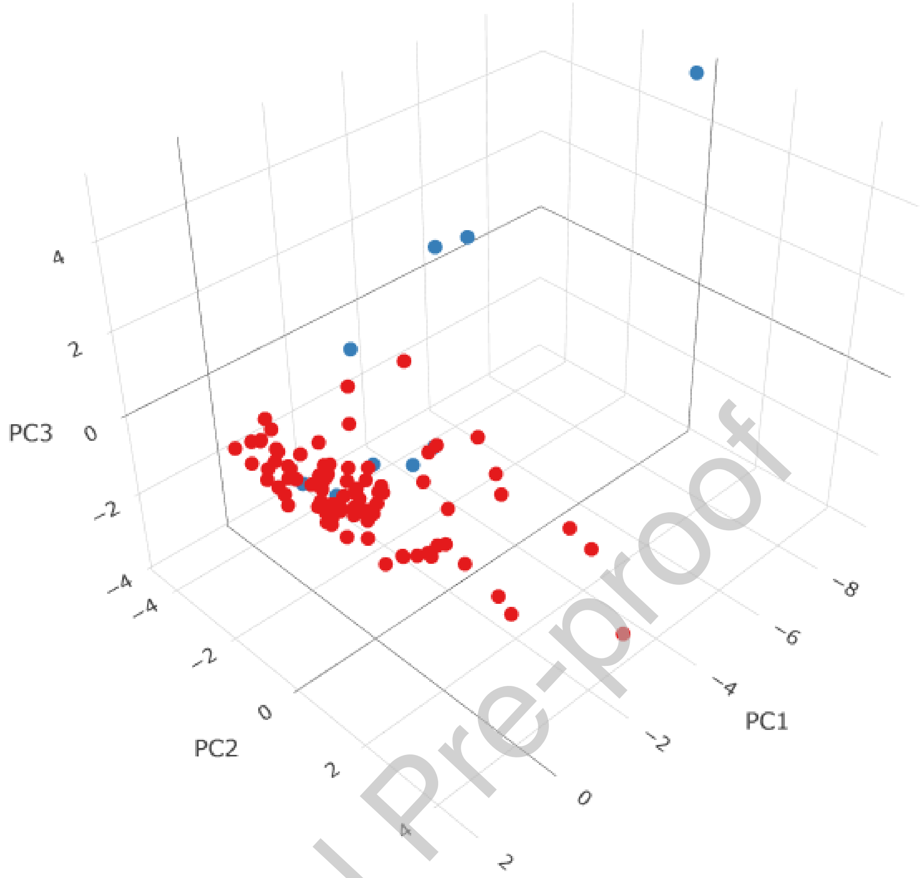

Figure 2: Three-dimensional principal component representation of the hierarchical clustering results according to the first 3 principal components (PC1, PC2, and PC3).

Caption: each point represents 1 flock. Cluster 1 flocks are colored red, and Cluster 2 flocks are colored blue. Principal component analysis was used to visualize the multivariate separation between the 2 clusters identified by hierarchical clustering.

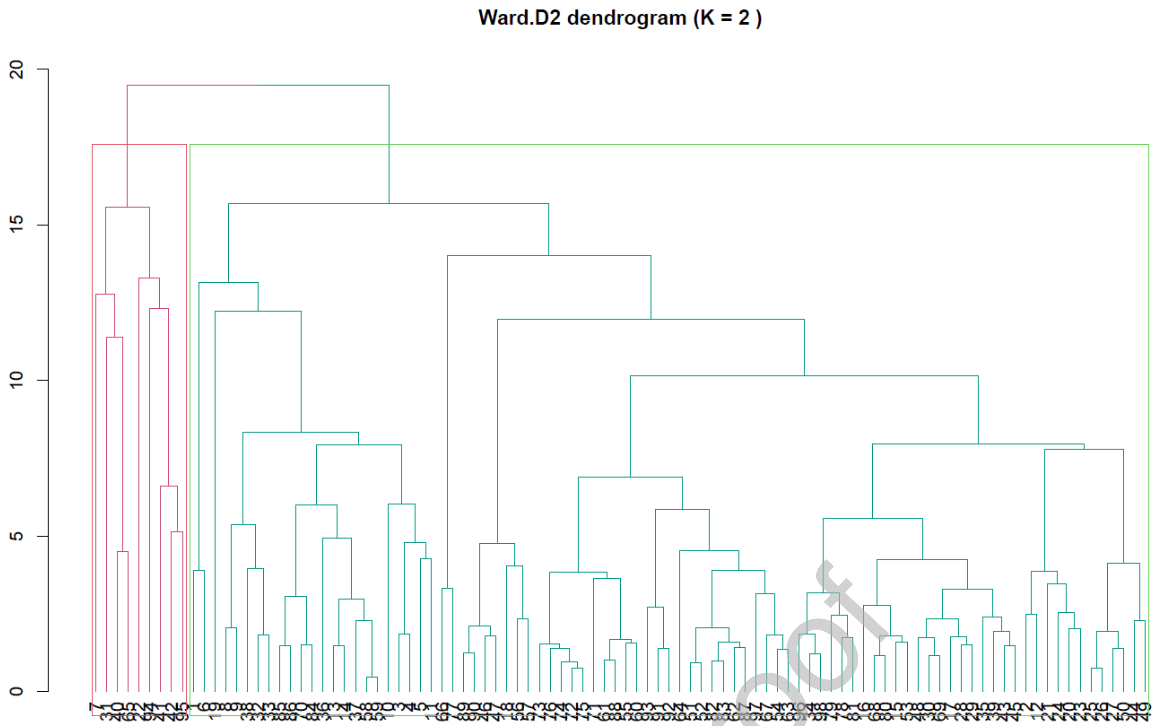

Figure 3. Hierarchical clustering two-dimensional dendrogram, produced using the Ward.D2 agglomerative clustering method, showing how observations (flocks) are distributed among clusters.

Caption: each value in the X axis corresponds to a flock. Y axis shows the linkage height. Branches in green correspond to flocks of Cluster 1, and branches in red correspond to flocks of Cluster 2.

Declaration of interests

We, the Authors of the original research manuscript entitled *'Data-driven categorization of farm animal welfare and slaughter outcomes in slow growing chickens with outdoor access: When data speak'*, submitted for consideration in *Poultry Science*, declare that there are no conflicts of interest regarding the publication of this article.