Microbial variation and shelf life of chicken breast fillets: a year-long study across two European producers

Birgitte Moen[1*], Beatriz Nunes Silva[2], Ingrid Måge[1], Paula Teixeira[2], Solveig Langsrud[1], Ingunn Berget[1], Merete Rusås Jensen[1], Mari Øvrum Gaarder[1], Ina Aksnes Lian[3], Dina Fernandes[4], Solfrid Bjørkøy[3], Diogo Ferreira[4] and Annette Fagerlund[1].

[1]Nofima – Norwegian Institute of Food, Fisheries and Aquaculture Research, 1433 Ås, Norway

[2]Universidade Católica Portuguesa, CBQF – Centro de Biotecnologia e Química Fina – Laboratório Associado, Escola Superior de Biotecnologia, Rua Diogo Botelho 1327, 4169-005 Porto, Portugal

[3]Norsk Kylling AS, Havneveien 43, 7300 Orkanger, Norway

[4]Lusiaves – Indústria e Comércio Agro-Alimentar, S.A., Rua da Levada nº 3860-497 Pardilhó, Estarreja, Portugal

[*]Corresponding author: Birgitte Moen, birgitte.moen@nofima.no

Abstract

Despite extensive research on poultry spoilage, variability across batches, seasons, and facilities remains poorly resolved. This limits the ability to determine whether microbiological data can be used to predict sensory shelf life at the batch level, and thereby support adaptive, data-driven shelf-life determination. This study examined how initial microbial status, growth dynamics, microbiota composition, and sensory changes varied in modified atmosphere packed (MAP) chicken breast fillets from two European producers (A, Norway; B, Portugal) sampled over one year. Fillets were analysed after production and during storage at 4 °C or under a 4 to 8 °C temperature shift. Initial total viable counts (TVC) were lower in fillets from Producer A than B, but both producers showed approximately 2-log variation between batches. Maximum growth rates were similar, but fillets from Producer B showed consistently shorter lag times, reaching 7 $\log_{10}$ CFU/g earlier. Lag time was the primary determinant of shelf-life variation, whereas initial aerobic TVC had limited predictive value. During late storage, microbiota communities converged toward producer-specific spoiler profiles that were stable across batches (A: *Lactobacillales*, *Hafnia_Obesumbacterium*; B: *Brochothrix*, *Yersiniaceae*). Sensory deterioration correlated with TVC and was accelerated but not qualitatively altered by elevated temperature. Overall, initial aerobic TVC was not a reliable predictor of batch-specific shelf life; instead, early measurements that enable estimation of batch-specific lag time (i.e., the fraction of bacteria able to grow in the product) may provide an actionable basis for adaptive shelf-life setting within producers.

1. Introduction

Poultry meat, particularly chicken, remains one of the most consumed animal proteins in Europe due to its affordability, nutritional profile, perceived sustainability advantages relative to other meats, and broad consumer acceptance. In 2024, the EU produced an estimated 14.1 million tons of poultry meat, and the production is projected to grow by about 1 million tons over the period 2025-2035 (EU, 2025; Eurostat, 2025). However, chicken products are highly perishable and susceptible to microbial spoilage, which contributes to economic losses, food waste, and consumer dissatisfaction.

The microbial composition of chicken meat is shaped by multiple factors, including the animal's microbiota and farm environment, transport and pre-slaughter conditions, slaughter and processing practices, seasonal variation, and post-processing factors such as cutting/handling, as well as storage conditions (e.g., temperature) (Rouger et al., 2017; Sequino et al., 2025). Spoilage organisms commonly associated with poultry meat include *Pseudomonas* spp., *Enterobacteriaceae*, *Brochothrix thermosphacta* and lactic acid bacteria (LAB), mainly *Lactobacillus* spp. (sensu lato; including taxa reclassified, *e.g.* *Latilactobacillus*), *Carnobacterium* spp., *Leuconostoc* spp. and *Lactococcus* spp. (Herbert et al., 2015; Lauritsen et al., 2019; Lee et al., 2017; Morales et al., 2016; Sequino et al., 2025). *Photobacterium* spp. have also been isolated from spoiled poultry meat (Chen et al., 2020; Dourou et al., 2021; Sequino et al., 2025). The growth rate and establishment of the microbiota are also strongly influenced by the packaging atmosphere (Kandeepan and Tahseen, 2022). Under aerobic conditions, the shelf life is typically shorter (five to six days), and modified atmosphere packaging (MAP) is therefore commonly used to extend it (Holck et al., 2014; Jimenez et al., 1997; Kandeepan and Tahseen, 2022).

MAP with elevated CO_2 concentrations suppresses aerobic spoilage organisms (e.g. *Pseudomonas* spp.) and promotes CO_2-tolerant facultative anaerobes, including LAB and *B. thermosphacta* (Holck et al., 2014; Marmion et al., 2021). Under MAP conditions, the relationship between initial total viable counts (TVC) and spoilage development becomes less direct, as total counts integrate both inhibited non-spoilage background microbiota and spoilage-associated taxa that dominate late shelf life. In such systems, spoilage is not solely determined by microbial load, but also by the metabolic activity and interactions of specific spoilage organisms (SSOs) within the microbial community (Hansen et al., 2023; Shao et al., 2021). Nevertheless, TVC is still commonly used as an indicator of microbiological acceptability, with a frequently applied threshold of 7 $\log_{10}$ CFU/g (Holl et al., 2016; Zhang et al., 2012). Noticeable odour changes are typically reported when TVC exceeds 8 $\log_{10}$ CFU/g (Naveena et al., 2017; Rouger et al., 2018; Wang et al., 2016), although certain species may cause spoilage at lower levels (Marmion et al., 2021). Sensory evaluation is therefore essential to link microbial dynamics with perceived spoilage and to establish meaningful shelf-life endpoints.

Food waste can result from a mismatch between a product's actual shelf life under real storage conditions and the shelf life declared through date marking, which may contribute to premature disposal. Notably, the European Commission's market study on date marking estimates that up to 10% of the 88 million tonnes of food waste generated annually in the EU are linked to date marking practices (Commission et al., 2018). While producers cannot control storage conditions across the supply chain, including retail, restaurant, catering, distributor, or consumer, they may be able to reduce unnecessary waste by adjusting the use-by-dates to reflect batch-to-batch variability. However, despite extensive research on spoilage microbiota composition and packaging technologies, there is still limited understanding of how spoilage varies across batches, seasons and facilities. To date, most shelf-life studies

have focused on the effects of different temperatures or packaging concepts and have typically been based on a single or a limited number of batches. This knowledge gap poses a challenge for evaluating the feasibility of adaptive, data-driven shelf-life labelling, a concept where shelf life is predicted and the date mark is adjusted based on available information about each unit or batch, rather than fixed durations. Such information can derive from batch-specific characteristics at production (e.g., initial TVC, microbiota, pH, atmosphere) and/or from time-temperature exposure recorded along the supply chain (e.g., time-temperature indicator or time–temperature dependent loggers) (Li and Wang, 2017). Adaptive labelling is only feasible when shelf life is predictably related to these factors and when the resulting effect sizes are large enough to justify changes in labelling decisions (i.e., beyond measurement noise and consumer-handling variability). When this condition holds, shelf life may be estimated using a spectrum of modelling approaches, ranging from predictive microbiology (growth/inactivation kinetics and temperature dependence), to statistical frameworks (e.g., regression, survival analysis, mixed-effects to capture producer/batch/day), and, where warranted by complexity, machine-learning approaches (Tarlak, 2023). If real-time prediction is not feasible, a potential pragmatic alternative is to adjust date marking by seasonal or within-day patterns (e.g., different shifts), provided these patterns are stable and validated.

To address the current knowledge gap, this study evaluated whether batch-to-batch variation in bacterial load, microbiota composition, and storage growth dynamics in MAP-packaged chicken breast fillets is sufficiently structured and predictable to enable adaptive, data-driven shelf-life decision-making. Data were collected over one year from two producers. Three parallel hypotheses were formulated: 1) Systematic and predictable differences exist between production days over one year in initial microbial status and/or microbiota composition (production-day effects); 2) Systematic and predictable differences exist between batches

produced early and late on the same day (within-day batch effects); and 3) Higher initial TVCs are associated with a shorter time to sensory spoilage and/or a shorter time to 7 $\log_{10}$ CFU/g (initial load effects). Evidence supporting these primary hypotheses would indicate that the corresponding factors could potentially be exploited to inform adaptive, data-driven shelf-life decisions; conversely, negligible, inconsistent, or poorly predictable differences would challenge the feasibility of adaptive shelf-life determination.

To test these hypotheses, packaged chicken breast fillets were collected from different production batches from two geographically distinct producers using different MAP gas compositions: one in Norway (Producer A) and one in Portugal (Producer B). Shelf-life studies were then conducted beyond the use-by date at two temperatures (stable 4 °C and a temperature shift from 4 to 8 °C) to assess microbial and quality changes during extended storage. Sampling was performed over one year, and for each production day, two batches originating from different chicken flocks were collected.

2. Materials and Methods

Samples were consumer packages of boneless, skinless chicken breast fillets (Pectoralis major and Pectoralis minor muscles), obtained from two poultry producers (Producer A and Producer B; Table 1). Hereafter, these samples are referred to as “chicken fillets” or “fillets”. The MAP of Producer A is termed high-CO_2 MAP, and the MAP of Producer B high-O_2 MAP.

Table 1. Producer- and packaging-level parameters.

	Producer A	*Producer B*
Country	Norway	Portugal
No chickens processed daily	60 000-90 000	190 000
No shifts	1	2
Chicken breed	Hubbard	Ross Cobb
MAP gas	50-55% CO_2, 45-50% N_2	55-70% O_2, 20-30% CO_2, balance N_2
Packaging material	PET/PE	PET/PE
Package size (g)	1000 g ± 15-20 g*	459-809 g (average 653 g)
No fillets per package	4-5**	3 to 4
Shelf life (days)	20	7

** Allowable overfill of up to 20 g and underfill of up to 15 g.*

*** Occasionally 3 large fillets*

2.1. Experimental design

From February to December 2024, consumer packages of chicken fillets packed on the slaughter day (day 0) were collected on seven production days distributed across the year (production days 1–7) (Table 2). In addition, a pilot shelf-life experiment was performed at Producer A with chicken fillets in November 2023 (production day 0). On each production day, two different flocks of chicken were sampled: the first ("early") and the last or next to last ("late") flock processed in the shift. For the pilot shelf-life experiment, only products from the early flock were analysed. For Producer A, the slaughter day was either Tuesday (A0, A1, A2, A3 and A7) or Thursday (A4, A5, and A6) and for Producer B, all sampled production days were Fridays.

149 *Table 2. Overview of the sampling design for Producers A and B, including production day, slaughter and*
150 *packaging date, storage temperature, time of day (early or late), storage and sampling days, and total number of*
151 *packages analysed for each scenario.*

152

Producer	*Production day*	*Slaughter and packaging date*	*Storage temp.*	*Early batch*	*Late batch*	*Storage and sampling days*	*N packages*
A	0	31.10.2023	4 °C	✓		1, 7, 15, 17, 20, 22, 23, 27, 30	27
A	0		8 °C	✓		15, 17, 20, 23	12
A	1	20.02.2024	4 °C	✓	✓	1, 17, 20, 22, 24, 27	36
A	1		8 °C	✓	✓	17, 20, 22, 24	24
A	2	14.05.2024	4 °C	✓	✓	1, 7, 20, 24, 27	30
A	2		8 °C	✓	✓	10, 14, 17, 20	24
A	3	21.05.2024	4 °C	✓	✓	1, 7, 20, 24, 27	30
A	3		8 °C	✓	✓	10, 14, 17, 20	24
A	4	29.08.2024	4 °C	✓	✓	1, 7, 20, 25, 27	30
A	4		8 °C	✓	✓	15, 18, 20	18
A	5	12.09.2024	4 °C	✓	✓	1, 7, 18, 20, 25, 27	33
A	5		8 °C	✓	✓	15, 18, 20, 23	21
A	6	07.11.2024	4 °C	✓	✓	1, 7, 15, 18, 20, 25, 27	37
A	6		8 °C	✓	✓	11, 15, 18, 20	22
A	7	12.11.2024	4 °C	✓	✓	1, 7, 15, 17, 20, 24, 27	38
A	7		8 °C	✓	✓	10, 15, 17, 20	22
B	1	23.02.2024	4 °C	✓	✓	0, 7, 10, 12, 14, 17	33
B	1		8 °C	✓	✓	7, 10, 12, 14	24
B	2	31.05.2024	4 °C	✓	✓	0, 5, 7, 11, 14, 17	32
B	2		8 °C	✓	✓	4, 7, 11	16
B	3	14.06.2024	4 °C	✓	✓	0, 5, 7, 11, 14, 17	36
B	3		8 °C	✓	✓	4, 7, 11	18
B	4	05.07.2024	4 °C	✓*	✓	0, 5, 7, 11, 14, 17	21
B	4		8 °C		✓	5, 7, 11	9
B	5	12.07.2024	4 °C	✓	✓	0, 5, 7, 11, 14, 17	36
B	5		8 °C	✓	✓	4, 7, 11	18

B	6	20.09.2024	4 °C	✓	✓	0, 5, 7, 11, 14, 17	36
B	6		8 °C	✓	✓	4, 7, 11	18
B	7**	04.10.2024	4 °C	✓	✓	0	6

**Only sampling day 0 samples from early batch were analysed.*

***Due to a packaging error at this production day, only fresh (Day 0) samples could be analysed, as the products could not be stored for subsequent shelf-life analyses.*

2.2. Storage conditions and sampling in the shelf-life study

The packages were transported to the research labs at 4 °C with overnight delivery (Producer A) or same -day delivery (Producer B). Two storage scenarios were used. All products were stored at 4 °C until day 3 (Producer B) or day 7 (Producer A); at that point half of the packages were moved to 8 °C (**shift** from 4 °C to 8 °C), simulating moderate temperature abuse during retail and consumer stages of the food chain during the last two-thirds of the declared product shelf life. The remaining packages were kept at 4 °C throughout storage (**stable** 4 °C).

Three replicate packages were sampled for each production batch and sampling point. Packages were sampled on the day of arrival, corresponding to day 1 for Producer A and day 0 for Producer B. Additional sampling was performed during storage at predefined time points (Table 2). In total, 731 packages were analysed from the two producers.

2.3. MAP gas analysis

To ensure the integrity of the MAP and to detect any potential gas leakage or incorrect gas composition, headspace concentrations of CO_2 and O_2 in the chicken packages were analysed using a Dansensor CheckMate 9900 or CheckMate 3 headspace gas analyser (Ametek Mocon) with accuracies of ±0.01% O_2 and ≤ ±0.8% CO_2. Predefined exclusion criteria were applied to remove packages with headspace values indicative of compromised MAP conditions. For Producer A, packages with O_2 > 1% were considered non-compliant and

excluded, with one day-1 package retained despite elevated O_2 because microbial growth was assumed to be unaffected at this early stage. For Producer B, packages with both $O_2 < 25\%$ and $CO_2 < 15\%$ were considered indicative of loss of package integrity and excluded. All gas measurements are listed in Supplementary Table 1.

2.4. Microbiological enumeration

Two fillets from each package were placed in a sterile stomacher bag with filter, weighed, and 150 mL of peptone water (0.1% peptone, 0.85% NaCl; Oxoid) was added. The contents were hand-massaged for 1 minute. For packages containing large fillets, a single fillet was used, weighed and 100 mL of peptone water added.

To determine TVC and obtain colonies for MALDI-TOF MS identification, serial dilutions of the bag homogenates were prepared in peptone water and plated on standard plate count agar (PCA; Oxoid). Plates were incubated aerobically at 15 °C for 5-7 days.

All TVC results are reported as $\log_{10}$ CFU/g and listed in Supplementary Table 1.

2.5. MALDI-TOF MS colony identification

To further investigate the dominant microbiota, bacterial isolates were identified by MALDI-TOF MS following a common analytical framework applied by both the Norwegian and Portuguese teams, with minor differences related to colony handling procedures and analytical platform (instrumentation, software versions, and reference libraries). Colonies were picked from PCA plates for MALDI TOF MS identification as follows: After determining the TVC, up to three colonies were randomly selected from a defined sector of each plate from day 0/1 and use-by date for each product, under both storage conditions. Colony material was transferred to a MALDI target plate either by smearing a small amount of biomass as a thin film using a sterile wooden application stick or by transferring approximately 0.1 mg of cell material directly onto the target plate. Subsequently, 1 µL of

HCCA matrix solution (10 mg/mL α-cyano-4-hydroxycinnamic acid in 50% acetonitrile and 2.5% trifluoroacetic acid) was added to each spot, using a new pipette tip for each position to prevent cross contamination.

In Portugal, spectra were acquired using a MALDI-TOF Biotyper Sirius mass spectrometer (Bruker Daltonics) and analysed using MBT Compass HT software (version 5.1.450.98) with the MBT Compass HT reference library (version 2023) and FlexControl 3.4 software (Bruker Daltonics). In Norway, spectra were acquired using a Biotyper MALDI-TOF mass spectrometer (Bruker Daltonics, Germany) and processed using MBT Compass 4.1, FlexControl 3.4, and BioTyper 3.1 software, in combination with the MBT 9607 MSP Library (released November 2020) and an in-house reference library. In both cases, spectra were acquired according to the manufacturer's standard measurement procedures.

Instrument calibration was carried out using the Bacterial Test Standard (Bruker Daltonics). Identification scores ≥2.0 were considered reliable for species level identification, while scores between 1.7 and 2.0 were accepted for genus level identification. Scores below 1.7, in the absence of Sanger sequencing confirmation, were considered unreliable and recorded as "NoID".

MALDI-TOF MS was used to support the 16S rRNA gene amplicon (MiSeq) results, particularly by providing species-level identification for taxa that were not resolved to species level by amplicon sequencing (e.g., *Lactobacillales* = *Carnobacterium divergens* and *Yersiniaceae* = *Serratia proteamaculans*). In addition to MALDI-TOF, BLAST searches of representative sequences were used to confirm taxa.

2.6. DNA extraction and 16S rRNA gene amplicon sequencing

For DNA analysis, 2-10 mL homogenate per sample were collected immediately after hand-massaging and pelleted by centrifugation at 13,000 *g* for 5 min. The pellet was washed in

0.5 mL 1×PBS, centrifuged again at 13,000 *g* for 5 min, and stored at -20 °C until analysis. DNA extraction from pellets was performed with the DNeasy PowerLyser PowerSoil kit (Qiagen) as previously described (Moen et al., 2026).

16S rRNA gene PCR (V4 region) and paired-end sequencing (2×150 bp) using MiSeq reagent v3 kits on a MiSeq instrument (Illumina) were performed following the EMP 16S Illumina Amplicon Protocol v.2 with modifications (Caporaso et al., 2023) and protocol by Caporaso *et al*. (Caporaso et al., 2012) as previously described (Møretrø et al., 2021). Six MiSeq runs were performed, and the resulting demultiplexed raw reads have been deposited in the NCBI archives under BioProject accession no. PRJNA1310835. BioSample IDs for each sample are reported in Supplementary Table 1.

2.7. Microbial diversity analysis

The sequences were processed in QIIME 2 (qiime2-2023.5) (Bolyen et al., 2019). Briefly, the data processing was: demultiplexed using demux, paired ends were joined using vsearch, quality filtered based on a q-score above 30, denoised using deblur 16S, and taxonomy was achieved using classify-sklearn with the SILVA database (Silva 138 99% OTUs from the 515F/806R region) (Amir et al., 2017; Bokulich et al., 2013; Bolyen et al., 2019; Hamady and Knight, 2009; Hamady et al., 2008; McDonald et al., 2012; Rognes et al., 2016). Mitochondrial and chloroplast sequences were filtered out. The sOTU level taxonomy table was exported to text files and further processed as described below. A selection of representative sequences was queried against the NCBI BLAST nucleotide database (BLASTN 2.17.0) to obtain additional information about possible species (accessed 04.05.2026).

Counts exported from QIIME 2 were merged and further processed in Python. In total, 720 samples were analysed and, altogether, 3104 sOTUs were detected from a total of 24 420901

sequences after removing singletons. The mean number of sequences obtained per sample was 33918 (max 237 735 and min 1882). The mean number of sequences per sOTU was 7868 (max 8 236081 and min 2).

Alpha diversity was assessed using the Shannon, Simpson and Chao1 indices, as well as observed richness (S_{obs}), calculated using the scikit-bio Python package (v0.7.0). beta diversity was calculated from counts transformed to relative abundances. For beta diversity analyses on data subsets, only bacteria present in >95% of samples and with a maximum abundance >1% were retained. Filtered taxa were aggregated into a variable named "other" to maintain row sums of 100%. Average abundances were transformed by centred log ratio (CLR) prior to statistical analysis to account for the compositional nature of the data (Gloor et al., 2017). Prior to the CLR transform, zero values were replaced by 0.65 times the lowest value in the data set, as suggested by Lubbe *et al.*, (Lubbe et al., 2021), and the "other" column was adjusted accordingly to keep the row sum constant.

2.8. Bacterial growth modelling

For each production batch and storage scenario, the modified Gompertz model (1) was fitted to the data (Zwietering et al., 1990). Zwietering parameterisation was used as it yields directly interpretable parameters (lag time, maximum growth rate, and asymptotic/maximum population levels). The model is given by:

$$y(t) = y_0 + (y_{max} - y_0) * exp\left\{ -exp\left[\frac{\mu_{max} e}{(y_{max} - y_0)}(\lambda - t) + 1\right]\right\} \quad (1)$$

where $y(t)$ is the population level at time t ($\log_{10}$ CFU/g), y_0 is the initial value, y_{max} is the upper asymptote, which represents the maximum population, λ is the lag time, μ_{max} is the maximum specific growth-rate parameter and e is Euler's number. Parameters y_0, y_{max}, λ and μ_{max} were estimated by fitting aerobic TVC data ($\log_{10}$ CFU/g) for each individual

combination of production batch and storage scenario, using bounded nonlinear least squares with the Trust Region Reflective algorithm (curve_fit, SciPy, v1.16.1). Since the data was sparse, initial parameter guesses were derived from pooled data across all batches within the same producer and storage scenario, while parameter estimation was performed separately for each batch and storage scenario. To operationalise shelf life, the fitted Gompertz models were used to estimate the time required for TVC to reach 7 $\log_{10}$ CFU/g (Holl et al., 2016; Zhang et al., 2012); this predicted time was used as an estimate of shelf life for each batch × storage scenario combination.

2.9. Sensory evaluation of chicken fillets

Samples from all timepoints described above were included in the sensory analysis. Sensory evaluations were conducted using different panel configurations across the study. For fillets from Producer A, samples from production day 1 were evaluated in two sessions (early and late) by a semi-trained panel (9 members), whereas all subsequent sessions were evaluated by a trained professional panel (10 members). Trained professional assessors undergo regular training and performance monitoring in accordance with ISO 8586:2012, and sensory testing was carried out in facilities compliant with ISO 8589:2007. Samples from Producer B were evaluated by an in-house semi-trained panel trained for this purpose (15 members). Assessors selected for the semi-trained panels (Norway and Portugal) were required to successfully complete a basic taste screening prior to participation. Following selection, the semi trained panels received structured training in the Attribute Intensity Measurement method (AIM) applied in the study. The AIM method builds upon Quantitative Descriptive Analysis (QDA) according to ISO 13299:2016 General Guidance for establishing a sensory profile and employs the same 15 cm unstructured line scale where intensities are assessed from 1 (no intensity) to 9 (high intensity) for each attribute. In the AIM-method, a smaller number of attributes are assessed than in a traditional QDA (Lawless and Heymann, 2010). Training

comprised instruction and practice in the use of intensity scales, with particular emphasis on consistent application of the 1–9 scale. Assessors were further trained in odour recognition and descriptive terminology, including targeted training in identifying odour attributes relevant to raw chicken. In addition, assessors received instruction in standard procedures and appropriate conduct for participation as assessors in sensory analysis.

Fillets to be used in sensory analysis were vacuum packed (<99.0%) and stored in the dark at -40 °C (producer A) or -20 °C (Producer B). Before evaluation, frozen fillets were thawed overnight at 4 °C. On the day of evaluation, all samples were removed from cold storage 2.5 h prior to serving. Fillets were cut and prepared 30 min before serving. Chicken fillets were prepared according to a standardized procedure applied throughout the study. The ends of each fillet were removed, and the remaining portion was cut transversely into three equal pieces. Each piece was then split longitudinally, resulting in six pieces per fillet. Fillet pieces were placed in white plastic cups and covered with stainless-steel lids.

Panel calibration was conducted prior to the main study through a pre-test. During this calibration phase, samples were evaluated under red light to mask visual differences. Samples were served in a fixed sequence without validation testing. Following individual assessments, the assessors participated in a joint review session in the discussion room, where the stainless-steel lids were replaced with lids labelled BEST–MID–WORST to enable an informed comparative discussion. This calibration procedure was used to align understanding of attributes, intensity scaling, and use of reference points across assessors.

Assessors were provided with a fork and instructed to lift each chicken piece and evaluate odour both from the underside of the fillet and from the headspace within the cup. Samples were evaluated at room temperature and labelled with random three-digit codes. Each assessor was served one piece of fillet (width 2-3 cm, length 3-4 cm, thickness 1-2 cm) per sample.

Odour was evaluated using AIM-method on a 15 cm unstructured line scale from 1 to 9 as described above. Attribute selection was based on prior knowledge of storage related changes in poultry products, preliminary screening of representative samples, and panel consensus discussions to ensure relevance. Eleven odour attributes were included: total intensity, sour/fermented, sweet, cloying, yeasty, ammonia/burning/pungent, sulphur, process, burnt, metallic, and rancid.

Because sensory profiling provides quantitative measurements within a session but only semi-quantitative comparability between sessions (Lawless and Heymann, 2010), absolute intensity scores were not considered directly comparable across sessions. This limitation was particularly relevant for fillets from Producer A, for which two different panels were used. Sensory data were therefore interpreted primarily as within-session and within-producer patterns over storage time and temperature, rather than as directly comparable absolute scores across all sessions.

2.10. Statistical analyses

All statistical analyses were performed separately for fillets from each producer, as they differed markedly in initial microbiota, dominant spoilage organisms, and expected shelf life.

For microbial data, the differences between experimental factors such as storage scenario (temperature) and production days, were assessed using multifactorial analysis of variance (ANOVA). Univariate responses ($\log_{10}$ CFU/g and Gompertz growth parameters) were analysed using univariate ANOVA, while multivariate differences in alpha- and beta diversity were assessed using ANOVA-Simultaneous Component Analysis (ASCA) (Jansen et al., 2005; Khomich et al., 2021). Effect sizes are reported as explained variance (%).

Pairwise correlations among the estimated growth parameters were calculated separately for each producer and storage scenario, using the parameter estimates obtained from Gompertz

model fits to each individual batch × storage-scenario combination. Correlation coefficients were calculated in MATLAB using the corr function with default settings, corresponding to Pearson correlation coefficients, and visualised as heatmaps. In contrast, comparisons of pooled batch means for growth parameters were obtained from additive ANOVA models including storage scenario and batch as fixed factors. For these models, pooled batch estimates were compared using the Tukey-Kramer honestly significant difference procedure (multcompare in MATLAB). Because each batch contributed one estimate at 4 °C and one at 8 °C, the pooled batch point estimates are equivalent to the arithmetic mean of the two scenario-specific estimates, whereas the comparison intervals depend on the ANOVA residual variance and the Tukey-Kramer adjustment.

Sensory data were analysed at the session level. Principal Component Analysis (PCA) was performed separately for each sensory session to assess the dimensionality of sensory change during storage. Attribute intensities were also ranked descriptively within each session. LOESS regression (R function loess with span 1) was used to visualise total intensity as a function of storage day and, separately, as a function of TVC. To evaluate the relationship between sensory changes and microbiology, Pearson and Spearman rank correlations were calculated between sensory attributes and bacterial load ($\log_{10}$ CFU/g), both overall and by sensory session.

3. Results

3.1. Initial microbial load and composition of fresh chicken

Analyses of fresh chicken breast fillets on the day of slaughter (day 0; Producer B) or the day after slaughter (day 1; Producer A) were performed to determine the initial TVC and to characterise the initial bacterial community composition. Statistical analyses then examined variation between production days and sampling times within the same day.

Mean initial TVC were lower at Producer A (2.4 $\log_{10}$ CFU/g; range 1.3-3.0) than at Producer B (4.3 $\log_{10}$ CFU/g; range 3.3-5.4) (Fig. 1).

Both producers showed an approximately 2-log difference between the batches with the lowest and highest TVC. For both producers, initial TVC tended to be higher during the warmer months. For Producer A, the highest initial loads were observed in batches A2 to A4, corresponding to production in May and August (Table 2). For Producer B, initial TVC increased progressively from batches B1 to B5, spanning production from February to July (Table 2).

The microbiota of fresh chicken breast fillets was noticeably different between the two producers, with a higher abundance of *Bacillaceae* (typically not associated with spoilage) and *Lactobacillus* spp. in fillets from Producer A, and *Pseudomonas* spp., *Yersiniaceae* and *Brochothrix* spp. in fillets from Producer B (Fig. 2).

Overall, the ten most prevalent bacterial taxa in fresh chicken fillets (mean relative abundance across both producers) were, at mixed taxonomic rank (families and genera): *Bacillaceae* (14.7%), *Pseudomonas* (12.6%), *Lactobacillus* (12.2%), *Yersiniaceae* (8.8%), *Acinetobacter* (6.3%), *Brochothrix* (2.8%), *Chryseobacterium* (1.9%), *Faecalibacterium* (1.9%), *Bacteroides* (1.8%), and *Psychrobacter* (1.7%).

Between-day differences were the major source of variation in both TVC and microbiota in fresh chicken fillets. For Producer A, production day explained 39% of the variance in TVC, 48% in alpha diversity and 47% in beta diversity (all $p < 0.001$), exceeding the main effect of time of day (early vs. late: 16%, 9%, and 10%, respectively). The day × time interaction was also substantial (21%, 17%, 19%; $p \leq 0.004$), indicating that early-late differences depend on production day. For Producer B, the dominance of production day was even stronger, explaining 86% of TVC and 58% and 54% of alpha- and beta diversity, respectively (all

391 $p < 0.001$), whereas the main effect of time of day was small (TVC 2%, $p = 0.032$; α 6%, $p =$
392 0.002; β 9%, $p < 0.001$). The interaction was negligible for TVC (3%, $p = 0.26$) but remained
393 evident for the microbiota (21% for α and 17% for β; both $p < 0.001$) (Supplementary Tables
394 2 and 3). Consistent with these results, the ASCA score plots indicated day-to-day structuring
395 in both datasets. Clustering by production day was strong for Producer B, whereas several
396 production days overlapped for Producer A, indicating more moderate between-day
397 separation. Early-late shifts within a production day were present in both datasets but were
398 more pronounced for Producer A (Supplementary Figs. 1, 2, and 3). Loadings indicated that
399 separation along PC1 reflected overall alpha diversity, with all metrics loading positively.

400 For Producer A, several late batches (A1, A2, A4, and A5) were characterised by higher
401 relative abundances of *Bacillaceae* and were accompanied by lower alpha diversity than the
402 corresponding early batches (Supplementary Figs. 1 and 2). Production day A6 showed high
403 *Bacillaceae* abundance and low alpha diversity in both early and late samples and was the
404 only production day where the late batch had a higher TVC than the early batch. In contrast,
405 production day A3 showed low *Bacillaceae* abundance and high alpha diversity in both early
406 and late samples (Supplementary Figs. 1 and 2).

407 For Producer B, consistent with the stronger production day effect, individual days showed
408 distinct microbial profiles: production day B1 combined lower TVC with high relative
409 abundances of *Pseudomonas* and *Brochothrix*; the late batch from B1 also exhibited higher
410 alpha diversity than the other batches (Fig. 1, Fig. 2, Supplementary Figs. 1 and 3).
411 Production day B2 was marked by higher abundances of *Lactobacillus*, whereas production
412 day B5 stood out with a markedly elevated initial TVC relative to all other production days
413 (Fig. 1B), together with low alpha diversity, and an increased relative abundance of
414 *Yersiniaceae* (Fig. 2, Supplementary Figs. 1 and 2).

Overall, fresh chicken fillets showed clear producer-specific differences in both initial TVC and microbiota composition, with production day emerging as the dominant source of variability within each producer.

3.2. Microbial growth dynamics during storage

Storage experiments were conducted under two temperature scenarios: stable storage at 4 °C and a temperature shift from 4 °C to 8 °C (see Materials and Methods). At predefined time points, TVC was quantified and microbiota composition (alpha- and beta diversity) were characterized to capture temporal changes during storage. The products from the two producers appeared to differ markedly in bacterial growth dynamics during storage (Fig. 3 and Fig. 4).

Maximum specific growth rates (μ_{max}) were comparable between producers, whereas products from Producer B tended to exhibit shorter lag times (λ) and a higher estimated upper asymptote (y_{max}) than those from Producer A. As expected, μ_{max} values were higher under the temperature-shift scenario for both producers. In contrast, lag times did not differ significantly between storage scenarios (Supplementary Tables 4 and 5), likely because all products were stored at 4 °C during the initial phase of storage (7 days for Producer A; 3 days for Producer B), and these periods closely matched the estimated lag times, such that the temperature increase occurred largely after lag completion.

Accordingly, the predicted time required to reach 7 $\log_{10}$ CFU/g differed significantly between storage scenarios and was consistently longer for Producer A than for Producer B (Fig. 5, Supplementary Tables 4 and 5). For Producer A, this time ranged from 16 to 29 days at stable 4 °C (mean 21 days), and from 12 to 20 days under temperature shift to 8 °C (mean 14 days). For Producer B, the corresponding times were 8 to 17 days at stable 4 °C (mean 11 days) and 6 to 10 days under temperature shift to 8 °C (mean 7 days).

Variation in time to reach 7 $\log_{10}$ CFU/g was driven by both storage scenario and production batch (Supplementary Tables 4 and 5). Because missing data resulted in an unbalanced design, production day and time of day were not analysed separately for the storage data but were combined into the single factor batch. For Producer A, storage scenario explained 43% of the variance and batch explained 53%. For Producer B, the corresponding effects were 49% for scenario and 44% for batch (Supplementary Tables 4 and 5). Visually, these batch differences appeared to align more strongly with production day than with early-late differences within the same production day (Fig. 5).

Across producers and storage scenarios, time to reach 7 $\log_{10}$ CFU/g increased with lag time (r = 0.69-0.88), indicating that longer lag phases were associated with longer times to reach this threshold (Supplementary Fig. 7). By contrast, the fitted initial TVC (y_0) showed little association with time to 7 $\log_{10}$ CFU/g overall and was slightly negative for fillets from Producer B (r = -0.31 and -0.24), indicating that initial counts mattered less than subsequent growth dynamics. The association with the upper asymptote (y_{max}) varied by producer and scenario: it was negative for Producer A in both scenarios (r = -0.86 and -0.76) and for Producer B under temperature shift (r = -0.78), but moderately positive for Producer B at stable 4 °C (r = 0.44) (Supplementary Fig. 7).

For Producer A, time to 7 $\log_{10}$ CFU/g showed only weak correlations with day 1 aerobic TVC (r = 0.18 and 0.16), similar to those observed for the fitted initial levels (y_0). However, higher initial counts on plates incubated under product-like MAP were associated with shorter shelf life (r = -0.38 and -0.55) (Supplementary Fig. 7). This pattern is consistent with product-like MAP incubation better capturing CO_2-tolerant bacteria relevant to late shelf life (Moen et al., 2026). No corresponding initial MAP counts were available for Producer B. Associations between time to 7 $\log_{10}$ CFU/g and TVC measured later in storage were stronger and inverse: for Producer A, TVC at day 7 was strongly inversely correlated with time to 7 $\log_{10}$ CFU/g (r

464 = −0.91 and −0.87), whereas for Producer B the corresponding correlations were moderate
465 under stable storage at day 5 (r = −0.59 and −0.56) and stronger under temperature shift at day
466 4 (r = −0.84 and −0.81) (Supplementary Fig. 7). These inverse correlations were expected to
467 be stronger for later TVC measurements, because these samples were taken closer to the end
468 of lag phase or after the onset of faster growth, when differences in remaining time before the
469 7 $\log_{10}$ CFU/g threshold become more directly reflected in the observed counts.

470 Together, these results indicate that variation in time to 7 $\log_{10}$ CFU/g was most strongly
471 associated with lag time. In contrast, initial TVC (fitted or aerobic) showed little association,
472 and relationship with upper asymptote was inconsistent across producers and storage
473 scenarios.

474 3.3. Microbiota at late shelf life

475 To identify the bacterial groups dominating near the end of shelf life, the microbiota of
476 samples with TVC ≥ 7 $\log_{10}$ CFU/g was analysed. Restricting the analysis to this threshold
477 ensured that communities were compared at a similar late-stage spoilage level across batches
478 and storage scenarios.

479 Although the initial microbiota of fresh chicken fillets from both producers was highly
480 diverse, storage led to the emergence of distinct, producer-specific communities dominated by
481 only a few bacterial groups. At this stage, microbiota composition grouped primarily by
482 producer, with no systematic separation between early and late batches (Supplementary Fig. 8
483 and 9). Samples were therefore pooled across time of day for presentation in Fig. 6.

484 Stored chicken fillets from Producer A were dominated by *Lactobacillales* and
485 *Hafnia_Obesumbacterium*, while fillets from Producer B showed high abundances of
486 *Brochothrix* and *Yersiniaceae* (Fig. 6). The sOTUs identified as *Lactobacillales* and
487 *Yersiniaceae* can theoretically encompass a wide range of genera, however, complementary

species-level identification by MALDI-TOF MS and BLAST analysis of representative sequences suggested that these sOTU most likely corresponded to *C. divergens* and *S. proteamaculans*, respectively. Similarly, the sOTUs assigned to *Hafnia_ Obesumbacterium* and *Brochothrix* most likely corresponded to *H. alvei and B. thermosphacta*.

Under stable storage at 4 °C, CO_2-tolerant Gram-positive spoilage-associated taxa dominated the microbiota, with *Lactobacillales* (*C. divergens*) dominating fillets from Producer A and *Brochothrix* (*B. thermosphacta*) dominating fillets from Producer B (Fig. 6). In both cases, the temperature shift from 4 °C to 8 °C was associated with increased relative abundances of *Enterobacterales*, namely *Hafnia* (*H. alvei*) for Producer A and *Yersiniaceae* (*S. proteamaculans*) for Producer B. However, the magnitude of this shift differed between producers: microbiota composition in fillets from Producer A was more strongly affected, consistent with storage scenario explaining a larger proportion of the beta diversity variation than in Producer B (16% vs. 5%; Supplementary Tables 4 and 5). This producer difference co-occurred with lower initial TVC in Producer A.

As shown in Fig. 6, the late-shelf-life microbiota was largely consistent across production days within each producer, with no systematic day-to-day shifts in dominant taxa. Minor deviations were observed for individual production days, but these did not alter the overall producer-specific community structure. These patterns were consistent with multivariate analyses (ASCA) presented in Supplementary Figs. 8 and 9, confirming the overall producer-specific community structure at late shelf life.

3.4. Sensory development during storage

Sensory data were analysed to describe changes in odour intensity and key descriptive attributes during storage under the two temperature scenarios and to assess whether these changes tracked microbial growth within each producer. Because samples were assessed by

different panels/sessions (Materials and Methods), comparisons focus on within-session and within-producer patterns over time and temperature rather than absolute intensity levels across sessions. Details of the sensory sessions, panels, and evaluated odour attributes are provided in Supplementary Tables 6 and 7.

PCA showed that sensory changes during storage were effectively one-dimensional, with PC1 explaining 80-97% of the variance across sessions for Producer A and >95% for Producer B (Supplementary Figs. 10 and 11). This indicates that, within each producer, storage changed the sensory profile in broadly the same way at both temperatures. Higher temperature mainly accelerated this development, rather than producing a different pattern of odour attributes. All odour attributes increased during storage, and for fillets from both producers, total intensity, cloying, and sour-fermented were among the three most intense attributes. The strong correlation among all attributes suggested that total intensity captured most of the sensory variation and provided a practical metric for linking sensory and microbial development. This pattern was observed despite clear differences in the dominant late-shelf-life microbiota between fillets from the two producers.

LOESS regression showed that total intensity increased with storage time, and that higher temperature accelerated this development (Fig. 7). The relationship with bacterial load was then examined by correlation analysis of sensory attributes against TVC (Table 3, and Supplementary Tables 6 and 7). Across all sensory sessions, the Spearman correlations between total intensity and TVC were $\rho = 0.69$ for fillets from Producer A and $\rho = 0.71$ for fillets from Producer B. Among the individual odour attributes, the highest Spearman correlation was observed for sour-fermented for fillets from Producer A and for ammonia/burning/pungent for fillets from Producer B ($\rho = 0.75$ and 0.74, respectively) (

Table 3). When examined within individual sensory sessions, correlations between total intensity and TVC varied but remained consistently moderate to high, ranging from $\rho = 0.57\text{-}0.96$ for fillets from Producer A and $\rho = 0.58\text{-}0.93$ for Producer B (Supplementary Tables 6 and 7). The relationship with TVC was further visualized using LOESS regression, which showed that total intensity increased with TVC in a similar pattern at both temperatures, with higher temperature mainly accelerating progression rather than altering the overall pattern (Fig. 8).

Overall, the sensory results supported the microbial data by showing that odour changes, with total intensity as a practical summary attribute, closely tracked TVC within each producer across sensory sessions and storage scenarios.

Table 3. Pearson and Spearman correlations between each sensory attribute and mean observed TVC (log_{10} CFU/g) overall.

	Producer A		**Producer B**	
Sensory attribute	Pearson	Spearman	Pearson	Spearman
Total intensity	0.61	0.69	0.73	0.71
Ammonia/burning/pungent	0.53	0.58	0.72	0.74
Burnt	0.24	0.38	0.38	0.42
Cloying	0.50	0.57	0.71	0.73
Metal	0.12	0.11	0.64	0.64
Process	0.17	0.23	0.24	0.24
Rancid	0.14	0.04	0.38	0.38
Sour fermented	0.73	0.75	0.69	0.71
Sulphur	0.42	0.60	0.31	0.44
Sweet	0.24	0.12	0.56	0.58
Yeasty	0.37	0.31	0.63	0.68

4. Discussion

Adaptive, data-driven shelf-life determination requires systematic, predictable variation so that microbiological measurements taken at production or early during storage can support robust labelling decisions. In this context, the present results showed that production-day effects were the clearest and most useful source of variation in chicken fillets from the two producers, whereas differences between early and late batches sampled on the same production day were smaller and less consistent. Accordingly, the data supported the first hypothesis and only partly supported the second hypothesis. By contrast, the hypothesis that higher initial TVC would predict shorter shelf life was not supported when fillets from the two producers were analysed separately. Instead, variation in time to 7 $\log_{10}$ CFU/g, used as a practical microbiological benchmark for comparing shelf life across batches, was more closely linked to lag time than to initial TVC. This indicates that, for each producer, differences in shelf life were driven mainly by the initial number of spoilage bacteria and their early growth dynamics rather than by total bacterial load at the start of storage.

Production day explained more of the observed variability than within-day batch difference. This likely reflects that production day captures the combined influence of multiple factors that vary from one day to the next, such as weather conditions and changes in the processing plant environment (e.g., cleaning efficacy or pre-production drying) (Rouger et al., 2017; Sequino et al., 2025). By contrast, differences between early and late batches sampled within the same production day were smaller and not directionally consistent across production days. We assessed this within-day (time-of-day) hypothesis because time-of-day patterns have been reported in salmon processing (early vs later fillets differing in spoilage bacteria), and poultry studies document processing-environment reservoirs and cross-contamination that could plausibly generate shift-dependent patterns (Chen et al., 2020; Lauritsen et al., 2019; Moretro

and Langsrud, 2017). However, within-day differences were non-systematic, consistent with a mixture of flock-to-flock variation and shift-specific processing effects, suggesting that production day is a more useful grouping factor than sampling time for adaptive shelf-life decisions.

Within each producer, shelf life was more closely linked to the model-derived lag time parameter (λ) than to initial TVC. This should not be interpreted as shelf life was determined by an adaptation phase of the whole bacterial community. Under MAP, the apparent lag more plausibly reflects the time required for the spoilage-associated bacteria to increase from low initial abundance until they dominate total counts. In that context, TVC measures at production may be a poor indicator of batch-specific shelf life, because it combines the background microbiota with the smaller fraction of spoilage-associated bacteria that later shape product deterioration (Holl et al., 2016; Kolbeck et al., 2020; Lauritsen et al., 2019; Marmion et al., 2021). This helps explain why initial TVC showed little relation to time to 7 $\log_{10}$ CFU/g within either producer, even though there were clear differences in both initial counts and shelf life. The point is therefore not simply that initial TVC was a poor indicator of shelf life in our study, but why it had limited predictive value under these packaging conditions.

The interpretation that initial presence and growth of specific spoilage bacteria are more important than initial aerobic TVC in determining shelf life under MAP is supported by several observations and aligns with the specific spoilage organism (SSO) framework emphasised in EFSA guidance, which recommends identifying the relevant spoilage microorganisms and assessing their growth under the intended packaging and storage conditions rather than relying on total counts alone (Hazards et al., 2020). Time to 7 $\log_{10}$ CFU/g showed robust associations with later TVC measurements (e.g., day 7 TVC for Producer A), indicating that the useful signal becomes clearer once spoilage-associated

bacteria have already gained ground in the product. Consistent with this, correlations were stronger when initial TVC was measured under product-like MAP conditions rather than aerobically. This also aligns with earlier observations that MAP-matched incubation more effectively recovers CO_2-tolerant bacteria relevant to late shelf life (Moen et al., 2026).

Nevertheless, lag should not be overinterpreted as an early lever for shelf-life setting. In this study it was estimated from Gompertz fits rather than measured directly, and the precision of those estimates was constrained by limited sampling during the first days of storage. The practical implication is that prediction of shelf life for products under MAP may require either denser early sampling than was currently performed or measures that target spoilage-associated bacteria more specifically than a general TVC at production. Furthermore, since lag is not directly observable at the time of packaging, its strong association with shelf life is not immediately actionable for adaptive date marking. In our study, variability was primarily structured at the production-day level, implying that any practical shelf-life adjustment, if implemented, would need to be made at the production-day level using information available before product release (e.g., raw-material and process/environmental monitoring data), rather than by estimating λ itself. A lag-based prediction could, however, still be communicated later to downstream actors in the food chain through digital information systems, for example via QR codes, apps, or other barcode-linked data layers (Li and Wang, 2017), once sufficient storage information has accumulated.

Although 7 $\log_{10}$ CFU/g does not by itself define sensory rejection, it was used as a comparative microbiological benchmark in the present study because it provided a common late-storage reference point across batches and storage scenarios. Given potential panel and session effects, sensory results were interpreted in terms of relative development over storage time and temperature within each producer, rather than as absolute differences in intensity across sessions or producers. Within each producer, odour development was correlated with

increasing TVC, while higher temperature mainly accelerated sensory change rather than producing a qualitatively different odour pattern. Therefore, time to 7 $\log_{10}$ CFU/g functioned as a practical marker of spoilage progression, even if it should not be interpreted as a universal sensory cut-off. This interpretation is also consistent with previous poultry studies in which spoilage has been linked to microbial levels around 7 $\log_{10}$ CFU/g in some product-specific settings, while more noticeable odour changes in raw chicken have often been observed at higher counts, typically above 8 $\log_{10}$ CFU/g (Katiyo et al., 2020; Spyrelli et al., 2021). At the same time, recent work on cooked MAP chicken shows that sensory differentiation does not necessarily follow a universal TVC threshold, reinforcing that microbial counts should be interpreted in relation to product type, storage conditions and the mode of sensory evaluation (Soderqvist et al., 2024). Differences in spoilage-associated community composition may further shift the relationship between bacterial counts and sensory rejection (Marmion et al., 2021).

Producer-specific spoilage trajectories were one of the clearest ecological findings in this study. Although the microbiota in fresh fillets varied between production days, storage led to late-shelf-life communities that were more consistent within each producer, with *Lactobacillales* (likely *C. divergens*) and *Hafnia* dominating in fillets from Producer A and *Brochothrix* and *Yersiniaceae* (likely *S. proteamaculans*) for Producer B. Similar convergence towards a limited set of dominant spoilage taxa in MAP poultry has been reported previously, although the taxa involved depend on the packaging atmosphere and storage conditions (Holl et al., 2016; Lauritsen et al., 2019). The results are also consistent with the general MAP literature, where elevated CO_2 suppresses aerobic spoilers such as *Pseudomonas* and promotes facultative anaerobes and other CO_2-tolerant bacteria, including lactic acid bacteria and *Brochothrix* (Holl et al., 2016; Jimenez et al., 1997; Lauritsen et al., 2019; Soderqvist et al., 2024). However, the present study was not designed to separate the

effects of producer ecology and packaging atmosphere, because the two producers used different MAPs. Consequently, observed producer differences reflect both the applied MAP conditions and broader production-environment factors. Estimating the independent contribution of MAP composition would require an experimental design with controlled manipulation or standardisation of gas composition.

The temperature shift from 4 °C to 8 °C modified the microbiota development, most notably by increasing *Enterobacterales* (*Hafnia* and *Yersiniaceae*) at late shelf life. This temperature shift-related change in the microbiota was more pronounced in fillets from Producer A than Producer B, consistent with storage scenario accounting for a larger share of beta diversity variation in Producer A. One plausible explanation is that fillets from Producer A had lower TVC when moved to 8 °C, making the microbiota composition more responsive to temperature shift. The combination of higher initial TVC including CO_2 tolerant bacteria may explain why Producer B fillets exhibited shorter lag phases and reached 7 $\log_{10}$ CFU/g sooner than Producer A fillets. The CO_2-containing high-O_2 MAP used by Producer B may have favoured the increase and late shelf-life dominance of *Brochothrix* consistent with previous studies (Sarfraz et al., 2021). However, *Brochothrix* was already present at substantial relative abundance at packing, indicating that packaging composition alone is unlikely to explain its dominance.

Interestingly, although fillets from the two producers differed in lag time (λ), upper asymptote ($y_{\max}$), and late shelf-life microbiota, the estimated maximum growth rate (μ_{max}) was similar. This suggests that, under the conditions studied here, growth rate was influenced more by the shared temperature regime and product matrix than by the identity of the taxa that later dominated spoilage. This is consistent with temperature as a dominant external driver of microbial growth kinetics and shelf-life loss in refrigerated foods (Albrecht et al., 2020; Labuza et al., 1997). In practical terms, this makes time-temperature indicators (TTIs) or other

673 temperature-history tools relevant for adaptive, data-driven shelf-life management. At the
674 same time, the present results indicate that temperature alone is unlikely to explain all relevant
675 variation, since fillets produced on different production days still differed in lag time and time
676 to 7 $\log_{10}$ CFU/g. Temperature-history tools may therefore be most useful when combined
677 with producer or production day level adjustment, rather than used as stand-alone
678 package-level solutions. A key next step will be to identify measurable indicators for
679 spoilage-associated taxa or early growth behaviour that can be assessed at production and
680 integrated with temperature-history information, in order to support robust and practical
681 adaptive shelf-life determination strategies. However, even where microbiologically feasible,
682 adaptive shelf-life labelling may face practical barriers, including implementation costs,
683 integration with downstream systems, and the requirement that remaining variability is large
684 enough to yield meaningful waste reduction once consumer handling is considered.

685

686 5. Conclusions

687 This study examined whether batch-to-batch variation in the microbiological status and
688 growth dynamics of MAP-packaged chicken breast fillets is sufficiently structured and
689 predictable to support adaptive, data driven shelf-life determination. Variation was most
690 strongly structured at the production-day level: production day was the most consistent
691 grouping factor for both initial microbial status and shelf life, whereas differences between
692 batches produced on the same day were smaller and inconsistent. The results therefore
693 support the hypothesis that production-day effects are systematic and relevant for shelf-life
694 decisions, while providing only limited support for within-day batch effects as a basis for
695 shelf-life adjustment. Initial aerobic TVC did not explain within-producer variation in shelf
696 life. Instead, the time required to reach 7 $\log_{10}$ CFU/g was most strongly associated with the

model-derived lag parameter (λ), indicating that shelf-life variation was more closely related to early growth kinetics than to total aerobic TVC at production. Initial TVC therefore appears to add predictive value only when it reflects the starting abundance of spoilage-associated taxa that later develop under MAP. These findings suggest that adaptive shelf-life determination is more likely to succeed when based on structured production-day effects and early indicators of spoilage-community development, rather than on initial TVC.

Acknowledgements

The work was funded by European Union's Horizon Europe research and innovation programme under grant agreement N° 101136248, the Norwegian Fund for Research Fees for Agricultural Products: Micropack (NFR354136), the Norwegian Fund for Research Fees for Agricultural Products: Enable (NFR 354160) and the Scientific collaboration under the FCT project UID/50016/2025.

Declaration of generative AI use

During the preparation of this work the authors used Copilot to refine wording and improve structure. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Albrecht, A., Ibald, R., Raab, V., Reichstein, W., Haarer, D., Kreyenschmidt, J., 2020. Implementation of Time Temperature Indicators to Improve Temperature Monitoring and Support Dynamic Shelf Life in Meat Supply Chains. Journal of Packaging Technology and Research 4, 23–32.

Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Xu, Z.Z., Kightley, E.P., Thompson, L.R., Hyde, E.R., Gonzalez, A., Knight, R., 2017. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. Msystems 2.

Bokulich, N.A., Subramanian, S., Faith, J.J., Gevers, D., Gordon, J.I., Knight, R., Mills, D.A., Caporaso, J.G., 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nature Methods 10, 57–59.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodriguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J.R., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L.J., Kaehler, B.D., Bin Kang, K., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A.,

740 Turnbaugh, P.J., Ul-Hasan, S., vander Hooft, J.J.J., Vargas, F., Vazquez-Baeza, Y.,
741 Vogtmann, E., von Hippel, M., Walters, W., Wan, Y.H., Wang, M.X., Warren, J., Weber,
742 K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y.L., Zhu, Q.Y.,
743 Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible
744 microbiome data science using QIIME 2. Nature Biotechnology 37, 852–857.
745 Caporaso, J.G., Ackermann, G., Apprill, A., Bauer, M., Berg-Lyons, D., Betley, J., Fierer, N.,
746 Fraser, L., Fuhrman, J.A., Gilbert, J.A., Gormley, N., Humphrey, G., Huntley, J., Jansson,
747 J.K., Knight, R., Lauber, C.L., Lozupone, C.A., McNally, S., Needham, D.M., Owens, S.M.,
748 Parada, A.E., Parsons, R., Smith, G., Thompson, L.R., Thompson, L., Turnbaugh, P.J.,
749 Walters, W.A., Weber, L., 2023. EMP 16S Illumina Amplicon Protocol V.2 protocols.io,
750 https://www.protocols.io/view/emp-16s-illumina-amplicon-protocol-kqdg3dzzl25z/v2.
751 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens,
752 S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R.,
753 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq
754 platforms. Isme Journal 6, 1621–1624.
755 Chen, S.H., Fegan, N., Kocharunchitt, C., Bowman, J.P., Duffy, L.L., 2020. Impact of Poultry
756 Processing Operating Parameters on Bacterial Transmission and Persistence on Chicken
757 Carcasses and Their Shelf Life. Applied and Environmental Microbiology 86.
758 Commission, E., Health, D.-G.f., Safety, F., ICF, Anthesis, Lyndhurst, B., WRAP, 2018.
759 Market study on date marking and other information provided on food labels and food waste
760 prevention – Final report. Publications Office.
761 Dourou, D., Spyrelli, E.D., Doulgeraki, A.I., Argyri, A.A., Grounta, A., Nychas, G.J.E.,
762 Chorianopoulos, N.G., Tassou, C.C., 2021. Microbiota of Chicken Breast and Thigh Fillets
763 Stored under Different Refrigeration Temperatures Assessed by Next-Generation Sequencing.
764 Foods 10.

765 EU, 2025. European Commission: Directorate-General for Agriculture and Rural
766 Development. EU agricultural outlook 2025-2035. 2025. Available from:
767 doi/10.2762/4491268.
768 Eurostat, 2025. Poultry statistics. Statistics Explained. Available at:
769 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Poultry_statistics (Accessed:
770 20 February 2026).
771 Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome
772 Datasets Are Compositional: And This Is Not Optional. Frontiers in Microbiology 8, 2224.
773 Hamady, M., Knight, R., 2009. Microbial community profiling for human microbiome
774 projects: Tools, techniques, and challenges. Genome Research 19, 1141–1152.
775 Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., Knight, R., 2008. Error-correcting
776 barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods 5,
777 235–237.
778 Hansen, A.A., Langsrud, S., Carlehog, M., Haugen, J.E., Moen, B., 2023. CO2 packaging
779 increases shelf life through reduction of off-odor production by CO2 tolerant bacteria in
780 addition to growth inhibition of the spoilage bacteriota. Food Control 144.
781 Hazards, E.P.o.B., Koutsoumanis, K., Allende, A., Alvarez-Ordonez, A., Bolton, D., Bover-
782 Cid, S., Chemaly, M., Davies, R., De Cesare, A., Herman, L., Nauta, M., Peixe, L., Ru, G.,
783 Simmons, M., Skandamis, P., Suffredini, E., Jacxsens, L., Skjerdal, T., Da Silva Felicio,
784 M.T., Hempen, M., Messens, W., Lindqvist, R., 2020. Guidance on date marking and related
785 food information: part 1 (date marking). EFSA Journal 18, e06306.
786 Herbert, U., Albrecht, A., Kreyenschmidt, J., 2015. Definition of predictor variables for MAP
787 poultry filets stored under different temperature conditions. Poultry Science 94, 424–432.

Holck, A.L., Pettersen, M.K., Moen, M.H., Sorheim, O., 2014. Prolonged Shelf Life and Reduced Drip Loss of Chicken Filets by the Use of Carbon Dioxide Emitters and Modified Atmosphere Packaging. Journal of Food Protection 77, 1133–1141.

Holl, L., Behr, J., Vogel, R.F., 2016. Identification and growth dynamics of meat spoilage microorganisms in modified atmosphere packaged poultry meat by MALDI-TOF MS. Food Microbiology 60, 84–91.

Jansen, J.J., Hoefsloot, H.C.J., van der Greef, J., Timmerman, M.E., Westerhuis, J.A., Smilde, A.K., 2005. ASCA: analysis of multivariate data obtained from an experimental design. Journal of Chemometrics 19, 469–481.

Jimenez, S.M., Salsi, M.S., Tiburzi, M.C., Rafaghelli, R.C., Tessi, M.A., Coutaz, V.R., 1997. Spoilage microflora in fresh chicken breast stored at 4 degrees C: influence of packaging methods. Journal of Applied Microbiology 83, 613–618.

Kandeepan, G., Tahseen, A., 2022. Modified Atmosphere Packaging (MAP) of Meat and Meat Products: A Review. Journal of Packaging Technology and Research 6, 137–148.

Katiyo, W., de Kock, H.L., Coorey, R., Buys, E.M., 2020. Sensory implications of chicken meat spoilage in relation to microbial and physicochemical characteristics during refrigerated storage. Lwt-Food Science and Technology 128.

Khomich, M., Mage, I., Rud, I., Berget, I., 2021. Analysing microbiome intervention design studies: Comparison of alternative multivariate statistical methods. PLoS ONE 16, e0259973.

Kolbeck, S., Ludwig, C., Meng, C., Hilgarth, M., Vogel, R.F., 2020. Comparative Proteomics of Meat Spoilage Bacteria Predicts Drivers for Their Coexistence on Modified Atmosphere Packaged Meat. Frontiers in Microbiology 11.

Labuza, T., Sam Saguy, I., Taoukis, P., 1997. Kinetics of Food Deterioration and Shelf-Life Prediction, in: Valentas, K.J., Rotstein, E., Singh, R.P. (Eds.), Handbook of Food Engineering Practice, I CRC Press, pp. 361–404.

Lauritsen, C.V., Kjeldgaard, J., Ingmer, H., Bisgaard, M., Christensen, H., 2019. Microbiota encompassing putative spoilage bacteria in retail packaged broiler meat and commercial broiler abattoir. International Journal of Food Microbiology 300, 14–21.

Lawless, H.T., Heymann, H., 2010. Sensory Evaluation of Food: Principles and Practices, 2nd ed. Springer, New York.

Lee, H.S., Kwon, M., Heo, S., Kim, M.G., Kim, G.B., 2017. Characterization of the Biodiversity of the Spoilage Microbiota in Chicken Meat Using Next Generation Sequencing and Culture Dependent Approach. Korean Journal for Food Science of Animal Resources 37, 535–541.

Li, D., Wang, X.J., 2017. Dynamic supply chain decisions based on networked sensor data: an application in the chilled food retail chain. International Journal of Production Research 55, 5127–5141.

Lubbe, S., Filzmoser, P., Templ, M., 2021. Comparison of zero replacement strategies for compositional data with large numbers of zeros. Chemometrics and Intelligent Laboratory Systems 210.

Marmion, M., Ferone, M.T., Whyte, P., Scannell, A.G.M., 2021. The changing microbiome of poultry meat; from farm to fridge. Food Microbiology 99.

McDonald, D., Clemente, J.C., Kuczynski, J., Rideout, J.R., Stombaugh, J., Wendel, D., Wilke, A., Huse, S., Hufnagle, J., Meyer, F., Knight, R., Caporaso, J.G., 2012. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. Gigascience 1.

Moen, B., Fagerlund, A., Pursti, S.M., Jensen, M.R., Langsrud, S., 2026. Covariation between air chiller and chicken product microbiota in a poultry processing facility. PLoS ONE 21, e0344423.

837 Morales, P.A., Aguirre, J.S., Troncoso, M.R., Figueroa, G.O., 2016. Phenotypic and
838 genotypic characterization of Pseudomonas spp. present in spoiled poultry fillets sold in retail
839 settings. Lwt-Food Science and Technology 73, 609–614.
840 Moretro, T., Langsrud, S., 2017. Residential Bacteria on Surfaces in the Food Industry and
841 Their Implications for Food Safety and Quality. Comprehensive Reviews in Food Science and
842 Food Safety 16, 1022–1041.
843 Møretrø, T., Moen, B., Almli, V.L., Teixeira, P., Ferreira, V.B., Asli, A.W., Nilsen, C.,
844 Langsrud, S., 2021. Dishwashing sponges and brushes: Consumer practices and bacterial
845 growth and survival. International Journal of Food Microbiology 337, 108928.
846 Naveena, B.M., Khansole, P.S., Kumar, M.S., Krishnaiah, N., Kulkarni, V.V., Deepak, S.J.,
847 2017. Effect of sous vide processing on physicochemical, ultrastructural, microbial and
848 sensory changes in vacuum packaged chicken sausages. Food Science and Technology
849 International 23, 75–85.
850 Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahe, F., 2016. VSEARCH: a versatile open
851 source tool for metagenomics. Peerj 4, e2584.
852 Rouger, A., Moriceau, N., Prevost, H., Remenant, B., Zagorec, M., 2018. Diversity of
853 bacterial communities in French chicken cuts stored under modified atmosphere packaging.
854 Food Microbiology 70, 7–16.
855 Rouger, A., Tresse, O., Zagorec, M., 2017. Bacterial Contaminants of Poultry Meat: Sources,
856 Species, and Dynamics. Microorganisms 5.
857 Sarfraz, J., Hansen, A.Å., Haugen, J.E., Le, T.A., Nilsen, J., Skaret, J., Huynh, T.P., Pettersen,
858 M.K., 2021. Biodegradable Active Packaging as an Alternative to Conventional Packaging: A
859 Case Study with Chicken Fillets. Foods 10.
860 Sequino, G., Cobo-Diaz, J.F., Valentino, V., Tassou, C., Volpe, S., Torrieri, E., Nychas, G.J.,
861 Ordóñez, A.A., Ercolini, D., De Filippis, F., 2025. Environmental microbiome mapping in

poultry processing chain and assessment of microbial dynamics in response to different storage conditions. Food Microbiology 128.

Shao, L.T., Chen, S.S., Wang, H.D., Zhang, J.H., Xu, X.L., Wang, H.H., 2021. Advances in understanding the predominance, phenotypes, and mechanisms of bacteria related to meat spoilage. Trends in Food Science & Technology 118, 822–832.

Soderqvist, K., Peterson, M., Johansson, M., Olsson, V., Boqvist, S., 2024. A Microbiological and Sensory Evaluation of Modified Atmosphere-Packed (MAP) Chicken at Use-By Date and Beyond. Foods 13.

Spyrelli, E.D., Papachristou, C.K., Nychas, G.E., Panagou, E.Z., 2021. Microbiological Quality Assessment of Chicken Thigh Fillets Using Spectroscopic Sensors and Multivariate Data Analysis. Foods 10.

Tarlak, F., 2023. The Use of Predictive Microbiology for the Prediction of the Shelf Life of Food Products. Foods 12.

Wang, J.M., Zhuang, H., Hinton, A., Zhang, J.H., 2016. Influence of in-package cold plasma treatment on microbiological shelf life and appearance of fresh chicken breast fillets. Food Microbiology 60, 142–146.

Zhang, Q.Q., Han, Y.Q., Cao, J.X., Xu, X.L., Zhou, G.H., Zhang, W.Y., 2012. The spoilage of air-packaged broiler meat during storage at normal and fluctuating storage temperatures. Poultry Science 91, 208–214.

Zwietering, M.H., Jongenburger, I., Rombouts, F.M., van 't Riet, K., 1990. Modeling of the bacterial growth curve. Applied and Environmental Microbiology 56, 1875–1881.

Figure legends

Fig. 1. Total viable counts (TVC; log10 CFU/g) in MAP-packaged chicken fillets for each production day, shown separately for Producer A (left) and Producer B (right). Small points represent individual packages (raw measurements, jittered), and large points show batch means (early = grey; late = black). Error bars indicate mean ± SEM within batch. Early and late batches are shown side-by-side for each production day.

Fig. 2. Dominant microbiota in fresh chicken breast fillets. Stacked bar plots show the relative abundance of dominant bacterial taxa across production days for producers A and B. For each production day, early and late batches were combined (i.e., batch timing is not shown separately in this figure). Taxa belonging to the same genus are represented using different shades within the same colour palette (e.g., all Lactobacillus spp. sOTUs are shown in varying tones of blue). The category "other" comprises taxa below the dominance threshold.

Fig. 3. Bacterial growth during storage. Estimated average growth curves for each producer and storage scenario are given in bold; blue = stable 4 °C; red = shift from 4 to 8 °C. Dashed lines represent growth curves fitted to individual production batches. Estimated TVC growth curves for all batches are given in Supplementary Fig. 4.

Fig. 4. Estimated growth parameters obtained from Gompertz model fitting. Bold dots and lines indicate the means and corresponding 95% confidence intervals. Opaque dots represent parameter estimates for individual batches.

Fig. 5. Multiple comparisons of batch means for time to reach 7 log10 CFU/g (Time to 7 log) and lag time (λ) for Producer A (top row) and Producer B (bottom row). The 95% comparison intervals (horizontal lines) were calculated using the Tukey-Kramer honestly significant difference (HSD) method; thus, non-overlapping intervals indicate statistically significant differences between batch means. Estimates are pooled across storage scenarios (stable 4 °C and 4 to 8 °C temperature shift). Results for the remaining Gompertz parameters are provided in Supplementary Figs. 5 and 6.

Fig. 6. Relative abundances of dominant bacterial taxa in chicken fillets at late shelf life from Producers A and B under the two storage scenarios. Only samples with $\log_{10}$ CFU/g > 7 were included. Values represent mean relative abundances across samples within each producer, storage scenario, and production day.

Fig. 7. LOESS regression of total intensity against storage day for stable 4 °C (blue) and temperature shift from 4 °C to 8 °C (red), shown separately for Producer A (left) and Producer B (right). Symbols represent observed total intensity values across production days. Production day 1 for Producer A is excluded because it was evaluated by a different panel using a different part of the scale. Absolute intensity should not be compared directly between producers.

Fig. 8. LOESS regression of total intensity against mean observed TVC (log10 CFU/g) for both storage temperatures combined, shown separately for Producer A (left) and Producer B (right). Production day 1 for Producer A is excluded since the panels used scales differently. Absolute intensity should not be compared directly between producers.