RHEOLOGY IN FIBRE FORMATION FOR MEAT-ANALOGUES
DATA ANALYSIS OF PROTEIN MELT RHEOLOGY DATA

Sandra Barman[1], Erik Kaunisto[1] and Mats Stading[1,2]

[1]RISE Research Institutes of Sweden, Gothenburg, Sweden
[2]Chalmers University of Technology, Gothenburg, Sweden

ABSTRACT

Global meat consumption increased four-fold during the last fifty years, while population doubled[1]. Even if the increase in European meat consumption has slowed (currently 80 kg per capita, twice the world average), it is forecasted to increase by 10% more to 2030[2–4]. The increase in meat eating is also nutritionally alarming as excessive consumption has been linked to health problems, such as coronary heart disease and certain cancers[5].

Fibrous, meat-like analogues are today commercially produced from soy, pea and wheat, utilizing an extruder to form a protein melt at high moisture content, high temperature and high pressure with subsequent active cooling on exit. A common denominator for the fibre formation in meat-analogues and plastics is that it is known **how** to produce the fibres but not exactly **why** they are formed. Consequently, it is still difficult to utilize the full potential of these techniques.

The current hypothesis on the mechanisms responsible for the fibre formation contribute to understanding but are not sufficient to fully describe the formation and cannot be used to predict fibre formation ability of protein melts thus hampering the use of more sustainable protein sources. Overall, the hypotheses range from "physical"[6–8], describing mechanisms in terms of fluid dynamics, heat transfer and phase separation, to "chemical" emphasizing the chemical interactions between protein chains or polymer crystallites.

This contribution will focus on rheology of the protein melts, and especially on how to use state-of-the-art statistical analysis to determine the influence of temperature, protein and moisture content on rheological properties of the melts.

INTRODUCTION

The food chain is responsible for 25% of global greenhouse gas emissions, with meat production alone accounting for 14.5%[9]. Food production heavily relies on natural resources, and 75% of arable land in Europe and North America is dedicated to meat production. Between 1961 and 2009, global meat consumption doubled, and with the growing population, it is expected to continue rising until at least 2050[2]. This trend is concerning from a nutritional standpoint, as excessive meat consumption has been linked to health issues such as coronary heart disease and certain cancers[5,10].

In the past five years, sales of plant-based meat analogues have surged, despite production methods being established since the early 1990s[11]. Today, fibrous analogues are commercially produced from soy, pea protein, and wheat gluten protein using an extruder to create a protein melt at high moisture content, temperature, and pressure, followed by active cooling upon exit[12–14]. These products are commonly known as High-Moisture Meat Analogues (HMMA). While the production process for HMMA fibres is understood, the exact mechanisms behind their formation remain unclear, limiting the full potential of these techniques.

This contribution will focus on the evaluation of rheological data of plant-protein melts[15] in order to determine the contribution of temperature, protein and moisture content using state-of-the-art statistical analysis.

MATERIAL AND METHODS

Sample preparation

Pea protein isolate with 86% protein content, (Roquette Pisane M9, Lestrem, France) was mixed with 15% pea fibre (Cosucra Swelite Warcoing, Hainaut, Belgium) and fed into an extruder (Brabender TwinLab-F 20/40, Duisburg, Germany) where water was added to 60-68% of the total weight. The mixture was heated and sheared in the extruder to 100°C to form a melt and extruded into a 13 mm cylinder which was air cooled instead of going into the cooling die, se **Fig. 1**. The same mixture is known to form a fibrous structure when heated to 150°C with subsequent cooling to ~80° in the cooling die.

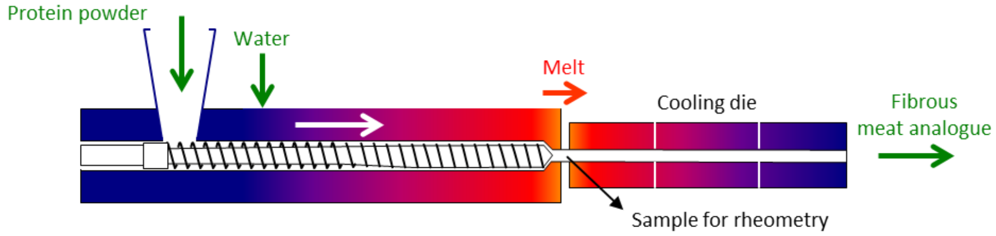

FIGURE 1: Principle of meat HMMA extrusion and sample preparation.

Rheometry

An HR 30 rheometer (TA Instruments, New Castle, DE, USA) was used for Small Amplitude Oscillatory Shear (SAOS) analysis, equipped with a 25 mm-diameter parallel plate system. Both plates were temperature controlled, and the measuring system was enclosed in a solvent-trap enclosure. Slices 13 mm in diameter and 2 mm thick were cut using a vacuum holder[16] and placed in the measuring gap. The gap was actively adopting to changes in samples volume with temperature. Mechanical spectra at 0.1-30 Hz were recorded during heating at 10°C intervals to give complex modulus as a function of angular frequency.

Statistical modelling

A variety of machine and statistical learning methods are available for analysing and interpreting data of the type presented in this study[17,18]. These range from classical methods such as linear regression and ANOVA, to more contemporary methods including regularized regression, random forests and neural networks.

We will use models from the generalized additive mixed model (GAMM) framework[19,20]. The specific formulation of model components within this framework depends on whether a frequentist or Bayesian approach is taken. In essence, however, the GAMM framework allows for including smooth components as functions of predictor variables. These can be both parametric and non-parametric, and linear and non-linear. Regularization of smooth, non-linear effects ensure that models are not overfit to data. The framework also accommodates non-Gaussian outcome distributions and supports modelling hierarchical and time-dependent structures using random components. The GAMM framework functions as an extension of classical statistical techniques, such as linear regression and ANOVA, encompassing these as special cases[21,22].

GAMM offer greater flexibility than classical statistical regression models, enabling the extraction of complex patterns within data. There is however a trade-off between flexibility and interpretability[17,23]. Compared to highly flexible machine learning methods such as random forests or neural networks, GAMMs tend to be more interpretable while still capturing nonlinear relationships. While highly flexible models like the mentioned machine learning methods may be preferable for large datasets due to their predictive power, for medium sized datasets such as the one presented here, GAMMs strike a good balance between flexibility and interpretability.

We formulate two models to capture the influence of temperature (T), moisture content (MC), protein concentration (PC) and experiment (E). Both models are formulated related to $1/T$ and $1/MC$ due to connections with physically motivated Arrhenius equations. The first model includes non-linear smooth effects of $1/T$ and $1/MC$ and random effects of PC and E, formulated as

$$\log({G_i}^*) = f_T(1/T_i) + f_{MC}\big(1/MC_{E_i}\big) + \mu_{PC_i} + \mu_{E_i} + \epsilon_i. \qquad (M1)$$

Each individual observation is indexed by $i = 1, \dots, 252$. For every observation i, the corresponding complex modulus G_i^* is recorded, along with the temperature T_i at which the measurement was made. Additionally, each observation is linked to one of 42 experiments, denoted by an experiment identifier $E_i \in\ (1, \dots, 42)$. Associated with each experiment are two variables: moisture content MC_{E_i} and protein concentration PC_{E_i}. In the model, f_T and f_{MC} represent smooth functions of the inverse temperature and inverse moisture content, respectively. The terms μ_{PC_i} and μ_{E_i} denote random effects, capturing variation by fitting individual intercepts for each protein concentration level and each experiment. Residual errors ϵ_i are assumed to follow a normal distribution. Note that without the components f_T and f_{MC}, model (M1) reduces to a classic ANOVA model.

The second model instead includes linear effects of $1/T$ and $1/MC$,

$$\log({G_i}^*) = \beta_T/T_i + \beta_{MC}/MC_{E_i} + \mu_{PC_i} + \mu_{E_i} + \epsilon_i. \qquad (M2)$$

Variable ranges used in the experiments are $T = 40$, 50, 60, 70, 80, 90°C for temperature and $PC = 75\%$, 80%, 85% for protein concentration. Five levels of added water were recorded, 60, 62, 64, 66 and 68%, and moisture content MC was recorded for each experiment.

This setup of using models of increasing complexity is a standard method for evaluating how much complexity is needed in a model[24]. A third model formulation is possible, where both linear components $1/T$ and $1/MC$ and non-linear components $f_T(1/T)$ and $f_{MC}(1/MC)$. We denote this model $(M3)$. Model $(M1)$ and $(M3)$ are in practice the same. Model formulation $(M3)$ has the advantage that the model $(M3)$ is nested within (M3), which simplifies assessment of whether the inclusion of non-linear components is warranted.

The models are fitted using the R package *mgcv*[25,26]. The adjusted R^2_{adj} is used as a measure of model fit. To quantify the contribution from each predictor, a model with (full model) and without (sub-model) the predictor is fitted, and the difference in R^2_{adj} between the two models is used to quantify the contribution from the predictor. The difference is denoted δR^2_{adj}. More details are available in Tsegaye et al[15].

RESULTS AND DISCUSSION

Statistical analysis

The statistical analysis allows us to quantify the size and statistical significance of effects of individual variables, the non-linear nature of each effect, and potential interactions between variables. We also assess the importance of allowing for non-linear effects compared to using linear effects.

Table 1 and Table 2 shows the importance (R^2_{adj}) and statistical significance (p-value) of each variable in model (M1) and model (M2), respectively (Table 1 and Table 2 are reproduced from Tsegaye et al.[15]).

Evaluating if the non-linear smooth effects are actually needed is in part a subjective judgement. The non-linear components of model (M3) are statistically significant (p-value $<$ $1e-16$), but in general statistical significance does not necessarily imply practical significance. Comparing the model with non-linear components (model (M1)) and the model with linear components (model (M2)) in time and moisture content, model (M1) has R^2_{adj} 99.2% while model (M2) has R^2_{adj} 98.3%. Thus, there is definitely a non-linear component in the fitted relationships, but we leave to the reader the decision of whether the effect of the non-linear component is considerable enough to warrant choosing model (M1) over the simpler model (M2). Fig. 1a) and b) provides more detail about the fitted non-linear effects and how they differ from linear effects (Fig. 1 is reproduced from the Supplementary material of Tsegaye et al.[15].).

The non-linear effects in Fig. 1 show that the effect of temperature, $f_T(1/T) \in [-0.48{,}0.61]$, is roughly three times the size of the effect of moisture content, $f_{MC}(1/MC) \in [-0.17{,}0.11]$). The size of the effect of experiment, $\mu_{E_i} \in [-0.18{,}0.27]$), is between moisture content and temperature. This also fits with the ordering shown in Table 1.

This means that effects we have not measured have a relatively large impact on the variation of the complex modulus G^* between different experiments. When controlling for these effects using μ_{E_i}, however, the complex modulus is explained almost fully by temperature and moisture content. Note that the random effect μ_{E_i} serves the same function as random effects in ANOVA, which controls for grouping / hierarchical structure in the data is not explicitly modelled using covariates.

This indicates that unmeasured factors significantly contribute to the variability observed in the variation of rheological behaviour across experiments. However, once these influences are accounted for through the inclusion of the random term μ_{E_i}, the variation can be largely explained by temperature and moisture content as indicated by the high R^2_{adj}-values. Importantly, μ_{E_i} functions similarly to random effects in ANOVA, taking into account grouped or hierarchical nature of the data when that structure is not directly captured through explicitly included variables like temperature and moisture content.

To conclude, we show that regression within the GAMM framework, which extends classical ANOVA and linear regression, can reveal non-linear, physically motivated relationships between experimental variables and rheological behaviour. We find that relationships involving temperature T and moisture content MC approximate Arrhenius type behaviour, i.e. they are log-linear with respect to $1/T$ and $1/MC$, even when non-linear effects are permitted. The strength of the effect is found to follow temperature > moisture content > protein concentration, with protein concentration showing no statistically significant effect.

Table 1 Non-linear mixed models (model (M1)).

Model component	δR^2_{adj}	p-value
Temperature	88.0 %	$< 10^{-16}$
Moisture content	3.6 %	0.0002
Protein content	0.1 %	0.79
Experiment	7.8 %	$< 10^{-16}$

Table 2 Linear mixed models (model (M2)).

Model component	δR^2_{adj}	p-value
Temperature	87.9 %	$< 10^{-16}$
Moisture content	3.6 %	0.00005
Protein content	0.1 %	0.82
Experiment	9.4 %	$< 10^{-16}$

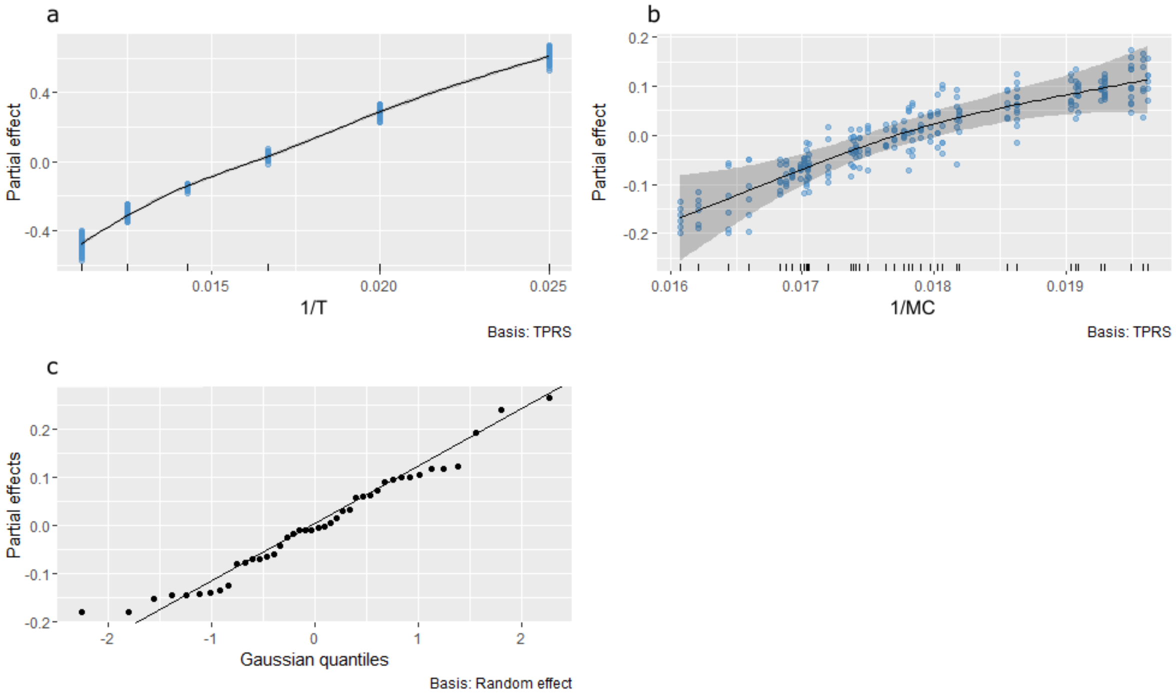

FIGURE 1: Fitted non-linear effects for (a) temperature, $f_T(1/T)$ and (b) moisture content, $f_{MC}(1/MC)$,, and quantile-quantile plot of (c) the random effect μ_{E_i} of experiment. Panels (a) and (b) show partial residuals (blue dots) and confidence bounds for the fitted non-linear effect. The confidence bounds are so close to the fitted non-linear effect in (a) that they are not visible. Fitted effects for protein concentration are excluded from the visualization as it was not significant. The figure was created using the *draw*-function from the *gratia*-package in R[27].

REFERENCES

(1) Weis, T. SFSGEC - Meatification and the Madness of the Doubling Narrative. *Can. Food Stud. Rev. Can. Études Sur Aliment.* **2015**, *2* (2), 296–303. https://doi.org/10.15353/cfs-rcea.v2i2.105.

(2) Steinfeld, H. *Livestock's Long Shadow - Environmental Issues and Options*; Food and Agriculture Organization of the United Nations (FAO), 2006. http://www.fao.org/3/a0701e/a0701e00.htm (accessed 2021-03-29).

(3) Westhoek, H. *The Protein Puzzle - The Consumption of Meat, Dairy and Fish in the European Union*; PBL Netherlands Environmental Assessment, 2011.

(4) *FAOSTAT*. FAOSTAT. http://faostat.fao.org/ (accessed 2010-09-23).

(5) Tammi, R.; Kaartinen, N. E.; Harald, K.; Maukonen, M.; Tapanainen, H.; Smith-Warner, S. A.; Albanes, D.; Eriksson, J. G.; Jousilahti, P.; Koskinen, S.; Laaksonen, M. A.; Heikkinen, S.; Pitkäniemi, J.; Pajari, A.-M.; Männistö, S. Partial Substitution of Red Meat or Processed Meat with Plant-Based Foods and the Risk of Colorectal Cancer. *Eur. J. Epidemiol.* **2024**, *39* (4), 419–428. https://doi.org/10.1007/s10654-024-01096-7.

(6) Sägesser, C.; Mair, T.; Braun, A.; Dumpler, J.; Fischer, P.; Mathys, A. Application of a Shear Cell for the Simulation of Extrusion to Test the Structurability of Raw Materials. *Food Hydrocoll.* **2025**, *160*, 110736. https://doi.org/10.1016/j.foodhyd.2024.110736.

(7) Kaunisto, E.; Wassén, S.; Stading, M. A Thermodynamical Finite Element Model of the Fibre Formation Process during Extrusion of High-Moisture Meat Analogues. *J. Food Eng.* **2024**, *362*, 111760. https://doi.org/10.1016/j.jfoodeng.2023.111760.

(8) Sandoval Murillo, J. L.; Osen, R.; Hiermaier, S.; Ganzenmüller, G. Towards Understanding the Mechanism of Fibrous Texture Formation during High-Moisture Extrusion of Meat Substitutes. *J. Food Eng.* **2019**, *242*, 8–20. https://doi.org/10.1016/j.jfoodeng.2018.08.009.

(9) Gerber, P. J.; Steinfeld, H.; Henderson, B.; Mottet, A.; Opio, C.; Dijkman, J.; Falcucci, A.; Tempio, G. *Tackling Climate Change through Livestock: A Global Assessment of Emissions and Mitigation Opportunities.*; Food and Agriculture Organization of the United Nations (FAO): Rome, 2013. https://www.cabidigitallibrary.org/doi/full/10.5555/20133417883 (accessed 2025-04-22).

(10) Glenn, A. J.; Wang, F.; Tessier, A.-J.; Manson, J. E.; Rimm, E. B.; Mukamal, K. J.; Sun, Q.; Willett, W. C.; Rexrode, K. M.; Jenkins, D. J.; Hu, F. B. Dietary Plant-to-Animal Protein Ratio and Risk of Cardiovascular Disease in 3 Prospective Cohorts. *Am. J. Clin. Nutr.* **2024**, *120* (6), 1373–1386. https://doi.org/10.1016/j.ajcnut.2024.09.006.

(11) Cheftel, J. C.; Kitagawa, M.; Quéguiner, C. New Protein Texturization Processes by Extrusion Cooking at High Moisture Levels. *Food Rev. Int.* **1992**, *8* (2), 235–275. https://doi.org/10.1080/87559129209540940.

(12) Jones, O. G. Recent Advances in the Functionality of Non-Animal-Sourced Proteins Contributing to Their Use in Meat Analogs. *Curr. Opin. Food Sci.* **2016**, *7*, 7–13. https://doi.org/10.1016/j.cofs.2015.08.002.

(13) Kyriakopoulou, K.; Keppler, J. K.; Van Der Goot, A. J.; Boom, R. M. Alternatives to Meat and Dairy. *Annu. Rev. Food Sci. Technol.* **2021**, *12* (1), 29–50. https://doi.org/10.1146/annurev-food-062520-101850.

(14) Krona, A.; Klose, F. P.; Gold, J.; Kadar, R.; Stading, M. Developing Cultured Meat Scaffolds of Extruded Vegetable-Based Proteins. *Annu Trans. Nord. Rheol Soc* **2017**, *25*, 311–313.

(15) Tsegaye, B.; Barman, S.; Bovagne, L.; Ellwanger, F.; Kaunisto, E.; Lorén, N.; Kádár, R.; Stading, M. Rheological Properties of Pea Protein Melts Used for Producing Meat Analogues. *Appl. Rheol.* **2025**, *35* (1), 20250036. https://doi.org/10.1515/arh-2025-0036.

(16) Stading, M.; Langer, R. Mechanical Shear Properties of Cell-Polymer Cartilage Constructs. *Tissue Eng.* **1999**, *5* (3), 241–250. https://doi.org/10.1089/ten.1999.5.241.

(17) Hastie, T.; Friedman, J.; Tibshirani, R. *The Elements of Statistical Learning*; Springer Series in Statistics; Springer New York: New York, NY, 2001. https://doi.org/10.1007/978-0-387-21606-5.

(18) James, G.; Witten, D.; Hastie, T.; Tibshirani, R.; Taylor, J. *An Introduction to Statistical Learning: With Applications in Python*; Springer Texts in Statistics; Springer International Publishing: Cham, 2023. https://doi.org/10.1007/978-3-031-38747-0.

(19) Hastie, T.; Tibshirani, R. Generalized Additive Models. *Stat. Sci.* **1986**, *1* (3), 297–310.

(20) Ruppert, D.; Wand, M. P.; Carroll, R. J. *Semiparametric Regression*; Cambridge university press, 2003.

(21) Mundo, A. I.; Tipton, J. R.; Muldoon, T. J. Generalized Additive Models to Analyze Nonlinear Trends in Biomedical Longitudinal Data Using R: Beyond Repeated Measures ANOVA and Linear Mixed Models. *Stat. Med.* **2022**, *41* (21), 4266–4283. https://doi.org/10.1002/sim.9505.
(22) Yu, Z.; Guindani, M.; Grieco, S. F.; Chen, L.; Holmes, T. C.; Xu, X. Beyond t Test and ANOVA: Applications of Mixed-Effects Models for More Rigorous Statistical Analysis in Neuroscience Research. *Neuron* **2022**, *110* (1), 21–35. https://doi.org/10.1016/j.neuron.2021.10.030.
(23) Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; Elhadad, N. Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: Sydney NSW Australia, 2015; pp 1721–1730. https://doi.org/10.1145/2783258.2788613.
(24) Wood, S. N. *Generalized Additive Models: An Introduction with R*; chapman and hall/CRC, 2017.
(25) Wood, S. N. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2011**, *73* (1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x.
(26) Wood, S. N. Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models. *J. Am. Stat. Assoc.* **2004**, *99* (467), 673–686. https://doi.org/10.1198/016214504000000980.
(27) Simpson, G. L. Gratia: Graceful Ggplot-Based Graphics and Other Functions for GAMs Fitted Using Mgcv. *R Package Version 0100* **2024**.