Proceedings of the International Conference on Heat Exchanger Fouling and Cleaning - 2024 (Peer-reviewed)
April 21–26, 2024, Lisbon, Portugal; Editors: H.U. Zettler and E.M. Ishiyama

ANALYSIS OF COMPLEX INTERDEPENDENCIES BETWEEN SOIL-SPECIFIC PROPERTIES AND CLEANING BEHAVIOR IN THE FOOD INDUSTRY BY USING A DECISION TREE

***S. Kricke[1], C. Drechsel[2], C. Schmidt[2], S. Zahn[2], H. Rohm[2], J.-P. Majschak[1] and H. Köhler[1]**
[1] Chair of Processing Machines/Processing Technology, Institute of Natural Materials Technology, TUD Dresden University of Technology, Germany, sebastian.kricke@tu-dresden.de
[2] Chair of Food Engineering, Institute of Natural Materials Technology, TUD Dresden University of Technology, Germany

ABSTRACT

Plate heat exchangers are commonly used for food processing, but tend to foul rapidly. Frequent cleaning guarantees product safety and quality, but accounts for high consumption of water, chemicals, energy, and time. Targeted soil characterization is essential for cleaning cost reduction. The effects of soil properties on cleaning are difficult to estimate since the interdependencies are complex. Principal component analysis and a decision tree were used to comprehend this complexity. A variety of carbohydrates formed the basis of the statistical exploration (eight native and modified starches, guar gum, xanthan, pudding, and whey protein isolate). Method adjustments allowed the characterization of soils during contact with different cleaning fluids. Solubility, fluid binding capacity, swelling-induced thickness increase, rheological properties, and pull-off forces were characterized. The cleaning time was obtained from planar channel flow experiments. The decision tree revealed an ordered relevance of the soil properties on cleaning. The effects are largely determined by the active cleaning mechanisms, which were quantified to help explain the results.

INTRODUCTION

Cleaning is a crucial process step to ensure consumer safety and high product quality in the food industry. Different products are commonly processed on the same machine. Cleaning in between product changes becomes necessary more frequently since batch sizes decrease due to economic trends such as product individualization [1]. Especially plate heat exchangers tend to foul rapidly at their hot surfaces. A high consumption of water, chemicals, energy, and time is required for cleaning. For fear of not fulfilling hygiene requirements, cleaning processes are often oversized even further. Product-specific cleaning process adaptations are rarely realized, but hold great potential for resource savings [2].

The characterization of the physicochemical properties of the soil (product deposit) and knowledge of their effects on the cleaning process are key factors for cleaning process adaptations. It is well known, that solubility and swelling properties, as well as mechanical properties and the adhesive forces to the surface determine the cleaning behavior of soil layers. However, the associated effects on the cleaning behavior are difficult to estimate, since the interdependencies of the soil properties are very complex [3–5].

Anderson et al. [6] and Chen et al. [7] used centrifugation tests to characterize technofunctional properties of various food ingredients. They first defined the water solubility index and water binding capacity to support recipe development. Schmidt et al. [8] first determined the solubility and binding capacity of various starch soil powders in contact with water and sodium hydroxide at different temperatures and discussed relationships to the cleaning behavior.

The cleaning fluid penetrates the soil layer and acts physically and chemically during cleaning. Polymeric networks such as long-chain carbohydrates or proteins swell and the cleaning fluid therefore influences the mechanical strength [9–12]. Köhler et al. [13] and Kricke et al. [14] characterized the swelling process by measuring the soil layer thickness increase. Dependencies on the cleaning process were observed and discussed.

The rheological properties of soil layers can be evaluated by oscillatory shear rheometry and rotational rheometry (specifically for viscous and viscoplastic fluids). The cleaning behavior of viscoplastic fluids is substantially determined by the critical shear stress [15, 16]. However, the cleaning behavior can be associated with the loss factor and complex shear modulus for viscoelastic soil layers such as starches and thickeners [10, 14, 17]. One particular challenge is the alteration of the mechanical properties of the soil layers over soaking time, particularly due to swelling and dissolution.

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

Pull-off experiments are commonly used to investigate adhesive and cohesive interactions [11, 13, 18]. Tsai et al. [19] and Fernandes et al. [20] estimated the critical shear stress of viscoplastic soil layers from micromanipulation experiments and related it with the cleaning behavior. Köhler et al. [13] linked the measured adhesive strength to the water mass fraction at the soil-substrate interface, and thus predicted the cleaning time of ketchup soil layers. Kricke et al. [14] connected the pull-off forces with the rheological properties of various starch soils.

The cleaning behavior of various soils can be distinguished by four cleaning mechanisms [21]. *Diffusive dissolution* is characterized by the dissolving of soil components within the cleaning fluid. *Cohesive separation* occurs when the cohesive bonds break and soil fragments are removed. *Viscous shifting* describes the push-out of flowable soils due to the stress applied by the flow. *Adhesive detachment* occurs when the adhesive bonds are exceeded, but the cohesive bonds remain intact. Hence, large coherent soil patches are removed at once [3, 10, 13, 22]. Which cleaning mechanism dominates, depends on the substrate, cleaning fluid, flow, and various soil properties, which are defined by soil composition and structure [14]. The cleaning fluid highly influences the soil structure, and therefore can account for changing cleaning mechanisms. Kricke et al. [14] reported *cohesive separation* as single cleaning mechanism of a native waxy maize starch when deionized water was used as cleaning fluid. In comparison, *viscous shifting* was predominant with sodium hydroxide. This phenomenon was related to gelatinization effects and newly created bonds, which were initialized by the alkaline environment. Furthermore, multiple cleaning mechanisms can be present simultaneously as reported by Golla et al. [23], who investigated a pregelatinized waxy maize starch. The authors presented a method to identify the present cleaning mechanism by analyzing the cleaning progressions spatially and temporally with neural networks.

Carbohydrate-based thickeners, such as starches, gums and modified derivatives, are commonly used in numerous food products. Their cohesive molecular networks enhance stability, texture, and overall sensory appeal of many foods [24]. The same networks are also responsible for the persistence during cleaning [3]. Whey proteins are a natural component of many dairy products. In dried form, they are often added to various foods to enhance their nutritional profile and structural properties. Both carbohydrate- and whey-protein-based foods as well as their components have repeatedly been used as model substances in cleaning studies [25]. Custard or pudding (cream pudding, blancmange) combines thickeners and whey protein components within one single complex product and was therefore already used as a model soil for cleaning investigations [26, 27].

The relations between physicochemical properties and cleaning behavior have so far mainly been considered univariate. Kricke et al. [28] first used the principal component analysis (PCA) to investigate the interdependencies between multiple soil properties and their complex effects on the cleaning behavior. Different native starches from various botanical origins were investigated with regards to their chemical composition, swelling properties, and cleaning behavior. Kricke et al. [14] further extended the investigations by including modified starches and characterizing the pull-off behavior and rheological properties. The authors gained insights on the influence of the swelling process on the mechanical properties and their combined effect on the cleaning behavior, especially the occurring cleaning mechanism. PCA was helpful to reveal these interdependencies, but offers only limited transferability for industrial use due to its requirement of complete underlying datasets, and its complex comprehensibility.

In contrast, decision trees are commonly used for classification tasks and offer high interpretability due to their hierarchal top-down presentation. Each node represents an evaluation of the specific attribute, each branch provides a corresponding result of the evaluation, and each leaf node gives a final classification result. Decision trees handle both continuous and categorical data as well as slightly incomplete datasets. Hence, they are suitable to identify key attributes from datasets with unknown interdependencies and make them interpretable. Apart from their simplicity, decision trees prone to overfitting and are sensitive to small changes in the training data. Locally optimal decisions at each node might lead to overlooking globally optimal decisions [29, 30]. Decision trees have so far only been used to a limited extend in the food industry as to investigate the effects of ingredient properties on product quality when large datasets are available as discussed by Du and Sun [31] and Lin et al. [32]. Dębska and Guzowksa-Świder [33] used a decision tree to investigate the effect of various ingredient properties on the quality of beer beverages.

This study aims to introduce a decision tree approach to gain deeper understanding of the influence of different soil properties on the cleaning behavior. Furthermore, a PCA supported the results of the decision tree and its interpretation. A wide variety of carbohydrate-based thickeners, whey protein isolate, and commercially available vanilla pudding were analyzed in combination with different cleaning fluids. Solubility, fluid binding capacity, swelling-induced thickness increase, rheological properties, and pull-off properties were characterized. Scalar values of the properties allowed for statistical exploration with the principal component analysis and a decision tree.

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

MATERIALS AND METHODS

Investigated soils and cleaning fluids

Maize starch (MA), waxy maize starch (WMS), potato starch (PS), wheat starch (WS), and rice starch (RS) were used as native starches as in Kricke et al. [14] and Schmidt et al. [8]. The modified starches were an acetylated waxy maize distarch adipate (ADA-S), and a hydroxypropyl potato distarch phosphate (HDP-S) as used in Kricke et al. [14]. A pregelatinized waxy maize starch (SG), xanthan (Xa) (both Cargill Holding (Germany) GmbH, Düsseldorf, Germany), and guar gum (Gu) (C.E. Roeper GmbH, Hamburg, Germany) were used as further carbohydrate-based thickeners. A whey protein isolate (WPI) (Sachsenmilch Leppersdorf GmbH, Leppersdorf, Germany) and a commercially available vanilla crème pudding (VP) (K-Classic, Kaufland Dienstleistung GmbH & Co. KG, Neckarsulm, Germany) expanded the investigations towards proteins and commercially available products (mixtures of carbohydrate-based thickeners, proteins, and simple sugars).

Kricke et al. [14] described the sample preparation thoroughly. First, the soil substances were suspended in deionized water at defined concentrations as given in Kricke et al. [14] for the native and modified starches. Concentrations were adjusted for the additional soils in order to gain applicable soil suspensions: SG 6.0 g/100 g; Gu 2.5 g/100 g; Xa 3.0 g/100 g; and WPI 4.0 g/100 g. A luminescent, stabilized strontium aluminate crystal tracer was added (0.0267 g/100 g dry soil substance). The starch suspensions were gelatinized at 95 °C for 45 min and stirring with 1,000 rpm. A gelatinization of SG, Gu, Xa, and WPI was not necessary and these suspensions were stirred at 55 °C for 45 min at 1,000 rpm to obtain homogeneous mixtures.

The soil pastes were applied on stainless steel substrates (AISI 304, 2B finish, $Ra \leq 0.12$ µm) with a pipette. VP was applied on the substrates with a scraper blade and no further preparation. The substrates measured 40 mm in diameter for rheometry, 40 x 20 mm² for soil layer thickness and pull-off experiments, and 150 x 80 mm² for cleaning experiments. The samples had been dried at standard climate (23 °C, 50% relative humidity) for 18 h. The average soil mass coverage was $\overline{m''_{s,0}} = 50 \pm 5$ g/m² after drying. Soil powders were needed to determine solubility and fluid binding. They were obtained by removing the dried soil layers from the substrates and grinding them homogeneously as thoroughly described by Schmidt et al. [8].

For this study, the soil layers and soil powders were characterized during contact with the cleaning fluid. Five industrially relevant test conditions were equally investigated for all characterizations and cleaning experiments: deionized water at 25 °C and 55 °C, sodium hydroxide with a concentration of 1.0 g/100 g at 40 °C, and a concentration of 2.0 g/100 g at 25 °C and 55 °C (abbrev. T25H2O, T55H2O, T40N10, T25N20, T55N20).

Measurement of solubility and fluid binding

Soil powders were used to determine the solubility and binding properties during contact with the cleaning fluid as described in Schmidt et al. [8]. After soaking of 0.40 g soil powder in 20 g of the respective cleaning fluid for 10 min, the samples were centrifuged at 10,000 g for 10 min. The supernatant fluid was separated from the soil sediment. The solubility index *SI* (g/100 g initial dry matter) represents the ratio of the soil content within the supernatant fluid to the initial soil dry matter. The binding capacity *BC* (g fluid/g dry matter) refers to the mass of fluid that has penetrated the soil sediment related to the soil dry mass of the sediment.

Soil layer thickness measurements during swelling

A method for the investigation of the soil layer thickness increase during swelling was established by Köhler et al. [13] and Kricke et al. [14]. The samples were placed horizontally, with the soil layer upwards, in a transparent and heatable basin. A camera imaged the probes from the side and allowed the time resolved measurement of the thickness increase $\Delta h(t_{\text{soak}}) = h(t_{\text{soak}}) - h(t_{\text{soak}} = 0)$ during swelling within the static cleaning fluid for 1800 s. Kricke et al. [14] discussed the relevance of the initial thickness increase for the cleaning process. Hence, $\Delta h(t_{\text{soak}} = 60\,\text{s})$ was used as a scalar value for further statistical analyses.

Rheological analyses

For this study, the viscoelastic behavior of the soil layers was evaluated by an ARES-G2 rheometer (TA Instruments–Waters LLC, New Castle, USA) with a 25 mm upper plate as adapted from Helbig et al. [10] and Kricke et al. [14]. In contrast to these studies, exchangeable stainless steel substrates of 40 mm diameter were used for the rheometrical experiments. This allowed for the examination of a wider range of samples in a shorter time. Furthermore, sample preparation and drying conditions were consistent with pull-off, thickness increase, and cleaning experiments. The soiled substrates were fixed inside a sample cup and acted as lower plate. The tempered cleaning fluid was directly applied onto the soil layer and the rheometer was set to the same temperature. After soaking for 2 min, dynamic time sweeps (0.5% strain, 1 rad/s, 5 min duration) with constant normal forces of 1 N (starches, VP, WPI) or 3.75 N (Gu, Xa) were conducted. The absolute value of the complex shear modulus $|G^*|$ and the loss factor $\tan\delta$ were evaluated 200 s after first addition of the cleaning fluid. The complex shear modulus G^* describes the deformation resistance of the soil and is generally given as logarithmic expression $\log(|G^*|)$. The loss factor $\tan\delta$ indicates the ratio of viscous to elastic

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

proportions, whereas a value of 0 describes an ideal elastic solid and a value of 1 indicates the transition from elastic to viscous behavior.

Pull-off experiments

Helbig et al. [10] and Kricke et al. [14] established and used a micromanipulation device to determine the strength of soil layers during their pull-off from the substrate after soaking within static cleaning fluid. For this study, the soaking time was set to $t_{soak} = 60\,s$ and the samples were positioned at a gap distance of $100 \pm 10\,\mu m$ between the scraper blade and the substrate surface after soaking. The scraper blade moved through the soaked soil layer at a constant speed of 2.6 mm/s and pulled it off. The required pull-off force was measured. It contains adhesive or cohesive binding forces, depending on the failure mode, that break when the soil layer is deformed and displaced [18, 19]. Normalization of the pull-off forces with respect to substrate width B of 20 mm gave the normalized pull-off force F/B and averaging these over the substrate length L of 40 mm let to the averaged pull-off force $\bar{F}/B$ as described by Köhler et al. [13].

Cleaning experiments

Joppa et al. [22] exhaustively described a test setup, which was used to investigate cleaning behavior in a planar channel flow. The 150 x 80 mm² samples were inserted in a 78 x 5 mm² cross-sectional planar flow area and were cleaned by the cleaning fluid at a bulk velocity of 1 m/s for 1,800 s (Reynolds numbers ranged from 10,300 to 18,500 depending on fluid temperature). UV lamps excited the luminescent tracer within the soil through the transparent cover of the test section. A monochrome camera continuously captured the remaining soil during the cleaning process. The measured gray values were averaged for each captured time step within a centered 40 x 40 mm² range of interest. Normalization of the intensity $I_{raw}(t)$ with respect to the initial intensity $I_{raw}(t = 0)$ ensured the comparability of the different cleaning experiments.

Joppa et al. [34] introduced a method to account for the swelling induced intensity increase of the soil during the cleaning process. Swelling experiments were conducted accordingly. Some soils were found to exhibit further intensity affecting processes other than just the expected intensity increase, such as continuous folding of the guar gum soil layer within the static cleaning fluid. Hence, the actual cleaning progress could not be completely distinguished from interfering swelling processes. The cleaning time $t_{c,80}$, at which 80% of the initial intensity was reduced, was found to be an practical scalar target value to compare the different cleaning experiments.

Köhler et al. [21] discussed that *diffusive dissolution* and *cohesive separation* could not be differentiated by macroscopic observation, and therefore were both treated as *cohesive separation*., A method for quantification of the occurring cleaning mechanisms *cohesive separation*, *viscous shifting*, and *adhesive detachment* was developed for this study. It referred to the cleaning mechanism detection method developed by Golla et al. [23]. Visual inspection of the captured cleaning process in combination with the corresponding intensity curve allowed the time-resolved differentiation of the cleaning mechanisms. *Cohesive separation* (*coh*) exhibited a characteristic continuous decrease in intensity. The intensity also declined continuously when *viscous shifting* (*visc*) occurred, but a bright wave-like soil agglomeration was observable at the cleaning front. A sudden drop in intensity was observed for *adhesive detachment* (*adh*). For this study, the cleaning state *no cleaning* (*no clean*) was additionally used to account for remaining soil at the end of the cleaning experiment. In this way, the occurring cleaning mechanisms could be quantified for each cleaning experiment as their proportion of the overall cleaning process as depicted in figure 1.

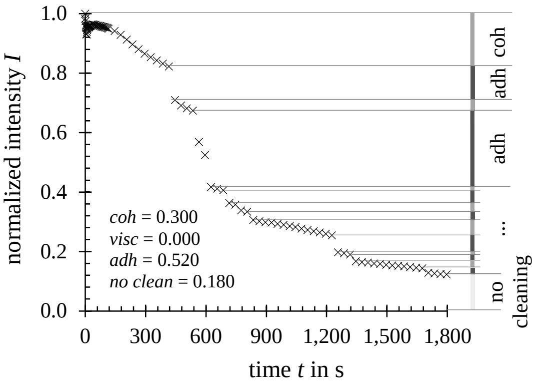

Fig. 1. Representative evaluation of the occurring cleaning mechanisms from a cleaning progress of wheat starch and T55N20.

Statistical analyses

All statistical analyses were performed with IBM SPSS Statistics 28.0 (IBM Corp., Armonk, NY). The obtained properties and cleaning behavior were gathered within so called objects. An object is generally defined as a unit of investigation and combines the essential properties for the statistical exploration. The cleaning system is defined as the triad of the specific soil, the cleaning fluid, and the substrate [21]. In the present case, the substrate was kept constant, and therefore the objects were formed by the specific combination of the present soil and cleaning fluid. The dataset consisted of 59 objects comprising 55 objects for training and induction of the decision tree, and four objects for its validation. A sufficiently large data set was needed for decision tree induction and furthermore, vanilla pudding (VP) combined several ingredient components. Hence, it was used for validation only and not tested with T40N10. The dataset was tested for multicollinearities regarding Pearson-correlation, and the properties and cleaning behavior were found to be suitable for the statistical analyses.

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

Table 1: Dataset fragment containing representative objects with mean values of their properties, cleaning behavior, and cleaning category as well as distribution of training and test data.

object			soil properties						cleaning behavior					cleaning category	
	soil	fluid	*BC*	*SI*	Δh	log(|*G**|)	*tan δ*	$\overline{F}/B$	$t_{c,80}$	*coh*	*visc*	*adh*	*no clean*		
			g/g	g/100 g	mm			N/m	s						
1	WMS	T25H2O	21.80	65.25	0.11	1.76	0.22	3.79	1800	0.000	0.000	0.000	1.000	not cleaned	training data
2	WMS	T55H2O	21.16	76.22	0.18	1.40	0.32	0.91	744	0.977	0.023	0.000	0.000	medium	training data
3	WMS	T40N10	9.64	58.69	0.22	1.01	1.06	0.91	436	0.383	0.617	0.000	0.000	fast	training data
4	WMS	T25N20	5.14	27.23	0.13	1.23	1.38	1.56	1800	0.740	0.000	0.000	0.260	not cleaned	training data
5	WMS	T55N20	4.25	32.60	0.23	0.94	1.67	0.60	322	0.240	0.760	0.000	0.000	fast	training data
…															training data
40	SG	T55N20	2.15	23.03	0.23	0.93	4.61	0.51	167	0.700	0.233	0.000	0.067	fast	training data
41	Gu	T25H2O	43.23	7.45	0.30	4.11	0.22	66.14	833	0.000	0.000	0.880	0.120	medium	training data
42	Gu	T55H2O	47.04	6.17	0.18	3.84	0.34	17.99	323	0.000	0.000	0.983	0.017	fast	training data
…															training data
55	WPI	T55N20	0.79	0.00	0.25	not measurable		1.15	27	0.000	0.000	0.960	0.040	fast	training data
56	VP	T25H2O	5.71	57.42	0.36	3.11	0.20	3.11	1800	0.455	0.000	0.000	0.545	not cleaned	test data
57	VP	T55H2O	4.86	13.49	0.39	3.76	0.40	0.58	446	0.800	0.000	0.200	0.000	fast	test data
58	VP	T25N20	8.18	70.22	0.33	3.50	0.26	1.71	156	0.757	0.000	0.243	0.017	fast	test data
59	VP	T55N20	4.97	30.37	0.41	3.73	0.25	0.37	19	0.820	0.000	0.173	0.007	fast	test data

Table 1 shows a dataset fragment for some exemplary objects. Each object carried the mean values of the different properties and cleaning behavior values. They therefore contained the six properties *BC*, *SI*, Δh, log(|*G**|), *tan δ*, and $\overline{F}/B$ as well as the five cleaning behavior values $t_{c,80}$, the cleaning mechanism proportions *coh*, *visc*, and *adh*, as well as *no clean*. The cleaning time $t_{c,80}$ was set to 1800 s when no cleaning occurred, or only less than 80% of the initial soil amount could be removed during experimental period. Furthermore, the cleaning time was segmented into four target cleaning categories to allow for induction of the decision tree. The cleaning category 'fast' was attributed to an object when $t_{c,80} < 600$ s. The cleaning category 'medium' was given to objects with $600\,\text{s} \leq t_{c,80} < 1200\,\text{s}$. $1200\,\text{s} \leq t_{c,80} < 1800\,\text{s}$ marked the thresholds for the cleaning category 'slow'. Objects with $t_{c,80} \geq 1800$ s were attributed with the cleaning category 'not cleaned'. In contrast to the other soils, the rheological properties of WPI were not measurable since the soil layer was already destroyed at first contact with the upper plate regardless of the cleaning fluid used for soaking. This indicates a very low strength of the soaked soil layer, but could not be proved for this study. Hence, the category 'not measurable' was attributed to log(|*G**|) and *tan δ* for the five WPI objects to categorize this phenomenon.

Only continuous variables can be analyzed for interdependencies with PCA, and therefore the WPI objects could not be included. Consequently, 50 objects with six properties and five cleaning behavior values were examined by PCA. In contrast, the induction of a decision tree is also possible when categorical and continuous variables are present concurrently since the continuous variables are categorized in equally distributed intervals for induction of the decision tree.

Multiple decision trees were built iteratively with varying sets of statistical parameters to obtain the decision tree with the highest informative value for the dataset. The decision trees were compared with regard to various factors. The accuracy was taken into account since it gives the ratio of correctly assigned objects to the total number of objects. A weighted assessment of the incorrectly assigned objects provided further information on the quality of the decision tree. The objects were evaluated based on their proximity to the target cleaning category, receiving a rating ranging from 1 to 0. Objects correctly assigned to the final nodes were given a rating of 1. If the assigned object differed by one category, a rating of 0.66 was given. A difference of two categories resulted in a rating of 0.33, and 0 was given for the maximum distance of three categories between 'fast' and 'no clean'. The weighted assessment is calculated as the ratio of the sum of ratings to the total number of objects. The obtained decision trees were additionally evaluated qualitatively. The extent of overfitting was examined for each decision tree, the resulting final nodes were inspected with regards to a clear ordinal structure of the target cleaning categories, and the physical sense of the results was accounted.

The CHAID algorithm uses a Chi-square test to split the attributes at the certain nodes and to reduce the information entropy of the dataset. It is commonly used and best suitable for datasets with both categorical and continuous variables and was therefore used as splitting method for the decision tree inductions. It allowed for variation of the significance value for splitting. Furthermore, the number of categorization intervals of the continuous variables was varied to obtain various decision trees. The tendency to overfitting was countered by varying the minimum subgroup size, a method also known as pruning. Promising decision trees were

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

further refined. The threshold values for splitting at the nodes were slightly adjusted to increase the score of the weighted assessment.

RESULTS AND DISCUSSION

Principal component analysis

Figure 2 presents the results of the PCA and gives an insight into the interdependencies between the properties and cleaning behavior of the objects. The included properties and cleaning behavior loaded on both principal components (PC), whereas PC1 explained 34.4% and PC2 explained 26.1% of the overall variance. PC3 only explained another 9.7%, and is therefore not presented.

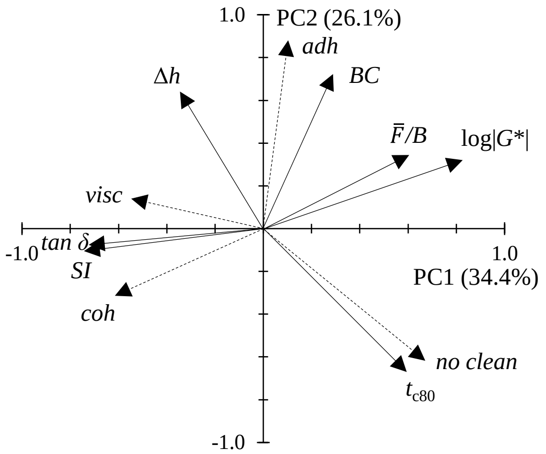

Fig. 2. Factor plot of the principal component analysis of all determined properties and cleaning behavior (dashed lines) of the investigated objects.

Dependencies were found for the mechanical properties as log(|G*|), *tan δ*, and $\overline{F}/B$ strongly loaded on PC1. A high log(|G*|) also resulted in high $\overline{F}/B$, and therefore a high resistance against pull-off since both log(|G*|) and $\overline{F}/B$ correlated positively with PC1. The opposing direction of *tan δ* indicated that objects with a high inner strength possessed a lower *tan δ* and therefore a pronounced elastic behavior. The solubility *SI* also showed large negative correlations with PC1 indicating connections with rheological properties. Objects with a high solubility tended to form weaker layers. Both *BC* and Δ*h* only slightly loaded on PC1. A moderate positive association was found between *BC* and log(|G*|) as well as $\overline{F}/B$. Objects which bound a lot of cleaning fluid might have tended to develop a high inner strength. Only slight connections were found for *SI* and *BC* as well as Δ*h*. This indicated that the solubility of the objects did not necessarily affect the binding capacity or result in major changes of layer thickness during swelling. Instead, *BC* and Δ*h* had a strong correlation with PC2, and therefore showed a dependency.

For the cleaning behavior, *cohesive separation* and *viscous shifting* loaded strongly negative on PC1. *Cohesive separation* was mainly found as predominant cleaning mechanism for objects with a high solubility *SI* and low inner strength log(|G*|) and pull-off force $\overline{F}/B$. The orientation of *viscous shifting* indicated a strong positive connection with *tan δ* as well as a negative connection with both log(|G*|) and $\overline{F}/B$. Dependencies of the cleaning mechanism *adhesive detachment* were mainly found for *BC* since both loaded on PC2. $t_{c,80}$ and the cleaning state *no cleaning* were found to strongly load on both PC1 and PC2, and point into the same direction. This indicated a correlation and aligned with logical assumptions. Furthermore, objects with low Δ*h* were hard to clean as indicated by the opposing orientation to *no cleaning*. The orientation of a*dhesive detachment* and *viscous shifting* alongside PC1 and PC2 indicated a negative correlation with $t_{c,80}$ suggesting quick cleaning processes once these cleaning mechanisms appeared. Furthermore, the positive orientation of log(|G*|), $\overline{F}/B$, $t_{c,80}$, and *no cleaning* alongside PC1 showed that objects with a high inner strength were generally hard to clean. The discussed findings are in good agreement with Kricke et al. [14].

Decision tree

Figure 3 shows the decision tree which was found to possess the highest informative value for the dataset. Rectangular boxes represent the decision nodes and the values at the branches give the found thresholds for splitting of the dataset. The final nodes are presented as rounded boxes and contain the predicted cleaning category. The assigned objects are given underneath the final nodes. The information on the actual cleaning category is given for each object. Furthermore, the quantified cleaning mechanisms are represented by the bar charts for each object. The diagram at the second decision node refers to the threshold adjustment and shows the original and adapted threshold as well as the passed on objects.

The presented decision tree was developed setting the significance value for splitting at each node to $p \leq 0.05$. The presented decision tree was furthermore obtained by using ten categorization intervals for the continuous values and setting the minimum subgroup size to three to reduce tendency to overfitting. The accuracy was 76.4% and was further increased to 86.1% when taking the weighted assessment into account.

The decision tree split the objects with regard to their log(|G*|) at the first node. Ten objects were classified as ‘fast cleaning’ when either log(|G*|) was comparably low at a value below 1.015 or not measurable as it was the case for WPI. Mainly objects with *viscous shifting* as the predominant cleaning mechanism, such as WMS and SG within alkaline environment, were assigned to the first final node alongside WPI. WMS and SG possessed a comparably high value of *tan δ* confirming the assumptions made from PCA.

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

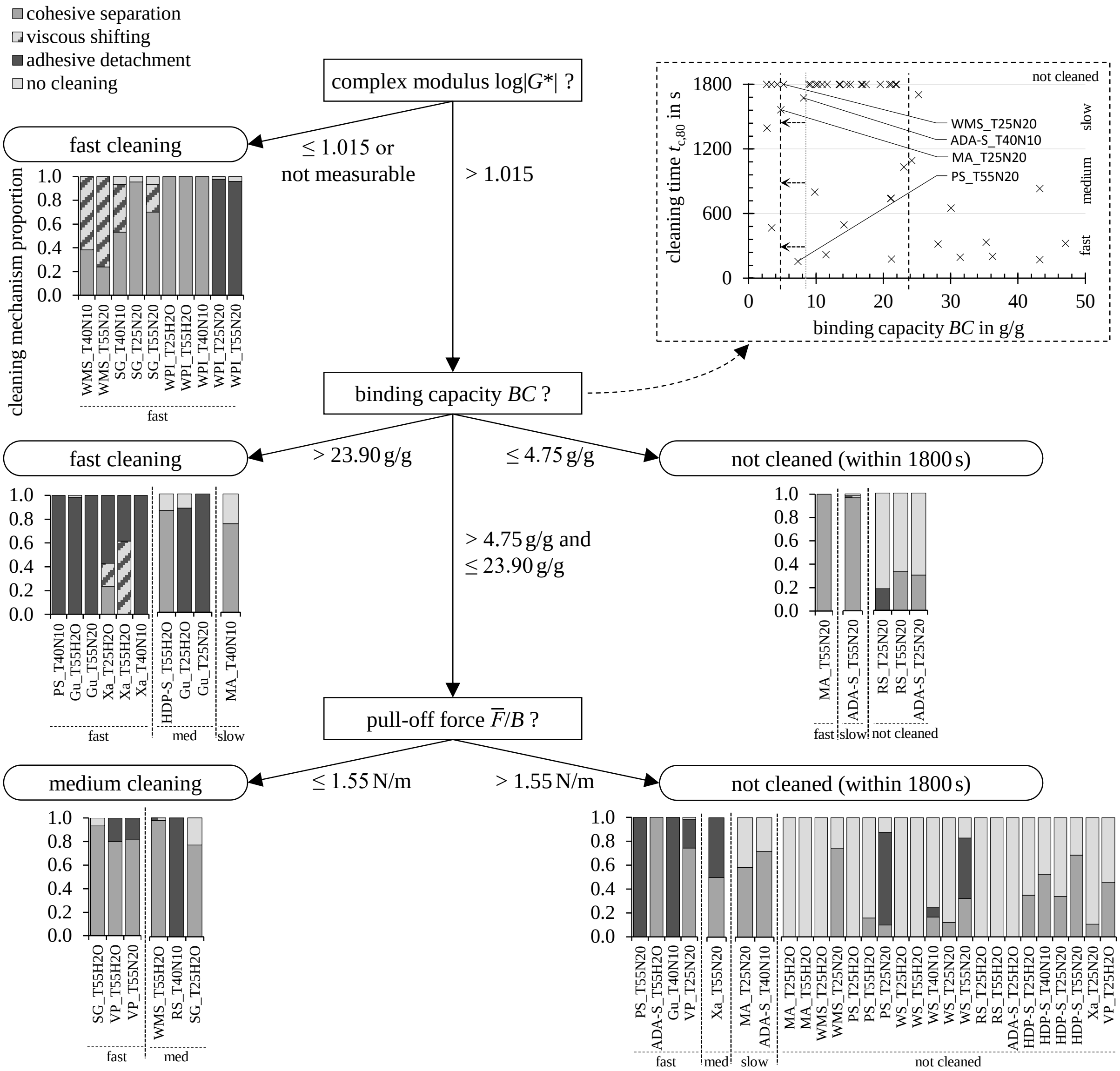

Fig. 3. Decision tree built for analysis of relevant properties. Decision nodes in rectangular boxes, thresholds alongside branches, and final nodes indicated by rounded boxes. Each assigned object is given with its observed cleaning behavior and actual cleaning category underneath final nodes. Diagram at decision node *BC* shows $t_{c,80}$ and cleaning category over *BC* for individual objects as well as threshold adjustment.

The binding capacity was found to best split the objects at the second node of the decision tree. Most objects with $BC > 23.90$ g/g possessed the cleaning category ‘fast cleaning’ and were correctly assigned to the respective final node. Few objects with high *BC* actually possessed the cleaning categories ‘medium cleaning’ and ‘slow cleaning’, and were therefore incorrectly assigned. Nevertheless, most of the objects gathered within this final node showed *adhesive detachment* as predominating cleaning mechanism supporting the findings from PCA. It is assumed that cohesive bonds were strengthened due to the absorbed fluid while the adhesive binding forces were reduced. Objects with low $BC \leq 4.75$ g/g were mainly not cleaned, and therefore assigned to the respective third final node.

Objects with a medium *BC* in between these thresholds were passed on to the third decision node. Final splitting of the objects was done according to $\overline{F}/B$. A comparably low threshold of $\overline{F}/B \leq 1.55$ N/m was found for objects with mainly ‘medium cleaning’ and only SG_T55H2O was assigned slightly incorrect with the actual cleaning category ‘fast cleaning’. All remaining objects with $\overline{F}/B > 1.55$ N/m were assigned to the final node ‘not cleaned’. The majority of the objects of the dataset gathered within this final node. Most objects were assigned correctly, but comparatively many objects with the cleaning category ‘fast cleaning’ were also assigned to that final node. These objects showed mainly *adhesive detachment* as predominating cleaning mechanism but their *BC* was comparatively

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

low. This implies that BC alone is not sufficient to separate objects from the dataset that are fast cleaned and possessed *adhesive detachment* as predominant cleaning mechanism.

VP was used for validation of the decision tree. All four objects possessed $\log(|G^*|) > 1.015$ and $4.75\,\text{g/g} < BC < 23.90\,\text{g/g}$, and were therefore passed on to the third decision node. VP_T55H2O and VP_T55N20 showed comparatively low $\overline{F}/B$ and were assigned to the final node ‘medium cleaning’ although their actual cleaning category was ‘fast cleaning’. Both objects were therefore assigned slightly incorrect. VP_T25H2O possessed high $\overline{F}/B$ and was not cleaned completely during 1800 s in the cleaning experiments. The object was assigned correctly to the respective final node ‘not cleaned’. VP_T25N20 was assigned to the same final node although its actual cleaning category was ‘fast cleaning’. *Adhesive detachment* was found as predominating cleaning mechanism supporting the previous findings. The weighted assessment, which accounted for the proximity to the actual cleaning category, resulted in an accuracy of 58 % for the validation data.

The presented decision tree aligned well with the findings from PCA, which therefore supported the interpretation of the underlying interdependencies. Mechanical and binding properties led to a physically comprehensible categorization of the expected cleaning time. The remaining properties Δh, SI, and $\tan\delta$ were not implemented as decision nodes since knowledge on $\log(|G^*|)$, BC, and $\overline{F}/B$ seemed sufficient for the current dataset distribution. In future analysis, it may be possible to reduce the effort required for soil characterization. However, multiple objects accumulated in the last final node ‘not cleaned’, including relatively many incorrectly assigned objects with *adhesive detachment*. This indicates that BC alone is not sufficient to explain this cleaning mechanism, and that the dataset was still too small for clear separations at previous nodes. Furthermore, it is to be stated that the object distribution might be unfavorable with regards to the target category since only few objects showed ‘slow cleaning’ or ‘medium cleaning’. A larger dataset for induction of the decision tree as well as a further differentiation of the cleaning target category ‘not cleaned’ beyond 1800 s could most likely lead to a more even distribution of the objects. It is also conceivable that different soil properties would be chosen for splitting at the decision nodes.

CONCLUSION

A wide range of carbohydrates, WPI and commercially available vanilla pudding were extensively characterized within this study. Respective methods had been refined to allow for soil characterization during contact with the cleaning fluid. Thus, a large dataset was established comprising solubility, binding, swelling and mechanical properties as well as extensive information on the cleaning behavior of 59 objects. PCA revealed underlying dependencies of the cleaning behavior on the solubility and binding properties as well as between the mechanical properties and the cleaning behavior. The currently characterized soil properties are all relevant for cleaning and none of the properties alone is sufficient for cleaning prediction as cleaning remains a multidimensional problem.

A decision tree was used for the first time to interpret these multidimensional interdependencies in a comprehensible way. Purposive decision nodes accumulated objects with mainly the same cleaning category. Rheological properties were confirmed to be mainly responsible for the cleaning mechanism *viscous shifting* and *cohesive separation*, and the binding capacity was found to be a key property for *adhesive detachment*, but not alone sufficient for its explanation. In general, the decision tree proved to be a suitable method for dependency explanation, but still needs an even larger dataset for a robust validation and extensive information reproduction.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Oleg Barteld, Chair of Processing Machines/Processing Technology, as well as Dennis Schab, Nadine Thamm, Meike Brunner, and Kristin Eichelberger, Chair of Food Technology, TUD Dresden University of Technology, by supporting the measurements. This research project is supported by the Industrievereinigung für Lebensmitteltechnologie und Verpackung e.V. (IVLV), the Arbeitsgemeinschaft industrieller Forschungsvereinigungen, Otto von Guericke‘ e.V. (AiF) and the Federal Ministry of Economic Affairs and Climate Action (AiF Project IGF 20861 BR).

NOMENCLATURE

Abbreviation

ADA-S	Acetylated distarch adipate
CHAID	Chi-square automatic interaction detectors
Gu	Guar gum
HDP-S	Hydroxypropyl distarch phosphate
MA	Maize starch
PC	Principal component
PCA	Principal component analysis
PS	Potato starch
RS	Rice starch
SG	Pregelatinized waxy maize starch
T25H2O	T = 25°C, deionized water
T55H2O	T = 55°C, deionized water
T40N10	T = 40°C, c_{NaOH} = 1.0 g/100 g
T25N20	T = 25°C, c_{NaOH} = 2.0 g/100 g
T55N20	T = 55°C, c_{NaOH} = 2.0 g/100 g
VP	Vanilla pudding
WMS	Waxy maize starch
WPI	Whey protein isolate

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

WS	Wheat starch
Xa	Xanthan

Latin symbols

adh	Proportion of adhesive detachment, -
B	Width, m
BC	Binding capacity, g/g
c	Concentration, g/100 g
coh	Proportion of cohesive separation, -
F	Force, N
*G**	Complex shear modulus, Pa
h	Height, mm
I	Intensity, -
L	Length, m
m''	Mass coverage, g/m^2
no clean	Proportion of no cleaning, -
p	Significance value, -
Ra	Average roughness, µm
SI	Solubility index, g/100 g
t	Time, s
T	Temperature, °C
visc	Proportion of viscous shifting, -

Greek symbols

tan δ	Loss factor, -

Subscript

0	initial
80	80% of initial intensity
c	cleaning
NaOH	sodium hydroxide
raw	original
s	soil
soak	soaking

REFERENCES

[1] Claassen, G. D. H., and Hendrix, E. M. T., On modelling approaches for planning and scheduling in food processing industry, in: Murgante, B., Misra, S., Rocha, A.M.A.C., Torre, C., et al. (Eds.), *Computational Science and Its Applications – ICCSA 2014*, Springer International Publishing, Cham 2014, pp. 47–59.

[2] Alvarez, N., Daufin, G., and Gésan-Guiziou, G., Recommendations for rationalizing cleaning-in-place in the dairy industry: Case study of an ultra-high temperature heat exchanger, *Journal of Dairy Science*, vol. 93, pp. 808–821, 2010.

[3] Fryer, P. J., and Asteriadou, K., A prototype cleaning map: A classification of industrial cleaning processes, *Trends in Food Science & Technology*, vol. 20, pp. 255–262, 2009.

[4] Herrera-Márquez, O., Serrano-Haro, M., Vicaria, J. M., Jurado, E., Fraatz-Leál, A. R., Zhang, Z. J., Fryer, P. J., and Avila-Sierra, A., Cleaning maps: A multi length-scale strategy to approach the cleaning of complex food deposits, *Journal of Cleaner Production*, vol. 261, pp. 121254, 2020.

[5] Agüeria, D. A., Libonatti, C., and Civit, D., Cleaning and disinfection programmes in food establishments: A literature review on verification procedures, *Journal of Applied Microbiology*, vol. 131, pp. 23–35, 2021.

[6] Anderson, R. A., Conway, H. F., and Peplinski, A. J., Gelatinization of corn grits by roll cooking, extrusion cooking and steaming, *Die Stärke*, vol. 22, 1970.

[7] Chen, J. Y., Piva, M., and Labuza, T. P., Evaluation of water binding capacity (WBC) of food fiber sources, *Journal of Food Science*, vol. 49, pp. 59–63, 1984.

[8] Schmidt, C., Brunner, M., Berger, C., Zahn, S., and Rohm, H., Solubility and swelling of soils from native starch, *International Journal of Food Science & Technology*, vol. 57, pp. 6755–6762, 2022.

[9] Liu, W., Fryer, P. J., Zhang, Z., Zhao, Q., and Liu, Y., Identification of cohesive and adhesive effects in the cleaning of food fouling deposits, *Innovative Food Science & Emerging Technologies*, vol. 7, pp. 263–269, 2006.

[10] Helbig, M., Zahn, S., Böttcher, K., Rohm, H., and Majschak, J.-P., Laboratory methods to predict the cleaning behaviour of egg yolk layers in a flow channel, *Food and Bioproducts Processing*, vol. 113, pp. 108–117, 2019.

[11] Cuckston, G. L., Alam, Z., Goodwin, J., Ward, G., and Wilson, D. I., Quantifying the effect of solution formulation on the removal of soft solid food deposits from stainless steel substrates, *Journal of Food Engineering*, vol. 243, pp. 22–32, 2019.

[12] Yang, J., Kjellberg, K., Jensen, B. B. B., Nordkvist, M., Gernaey, K. V., and Krühne, U., Investigation of the cleaning of egg yolk deposits from tank surfaces using continuous and pulsed flows, *Food and Bioproducts Processing*, vol. 113, pp. 154–167, 2019.

[13] Köhler, H., Liebmann, V., Golla, C., Fröhlich, J., and Rüdiger, F., Modeling and CFD-simulation of cleaning process for adhesively detaching film-like soils with respect to industrial application, *Food and Bioproducts Processing*, vol. 129, pp. 157–167, 2021.

[14] Kricke, S., Berger, C., Zahn, S., Köhler, H., Rohm, H., and Majschak, J.-P., Influence of rheological properties and pull-off forces of native and modified starches on cleaning in plane channel flow, *Heat and Mass Transfer*, vol. 60, pp. 861-870, 2024.

[15] Palabiyik, I., Lopez-Quiroga, E., Robbins, P. T., Goode, K. R., and Fryer, P. J., Removal of yield-stress fluids from pipework using water, *AIChE Journal*, vol. 64, pp. 1517–1527, 2018.

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com

[16] Wilson, D. I., Atkinson, P., Köhler, H., Mauermann, M., Stoye, H., Suddaby, K., Wang, T., Davidson, J. F., and Majschak, J.-P., Cleaning of soft-solid soil layers on vertical and horizontal surfaces by stationary coherent impinging liquid jets, *Chemical Engineering Science*, vol. 109, pp. 183–196, 2014.

[17] Xin, H., Chen, X. D., and Özkan, N., Cleaning rate in the uniform cleaning stage for whey protein gel deposits, *Food and Bioproducts Processing*, vol. 80, pp. 240–246, 2002.

[18] Ali, A., de'Ath, D., Gibson, D., Parkin, J., Alam, Z., Ward, G., and Wilson, D. I., Development of a 'millimanipulation' device to study the removal of soft solid fouling layers from solid substrates and its application to cooked lard deposits, *Food and Bioproducts Processing*, vol. 93, pp. 256–268, 2015.

[19] Tsai, J.-H., Fernandes, R. R., and Wilson, D. I., Measurements and modelling of the 'millimanipulation' device to study the removal of soft solid layers from solid substrates, *Journal of Food Engineering*, vol. 285, pp. 110086, 2020.

[20] Fernandes, R. R., Suleiman, N., and Wilson, D. I., In-situ measurement of the critical stress of viscoplastic soil layers, *Journal of Food Engineering*, vol. 303, pp. 110568, 2021.

[21] Köhler, H., Liebmann, V., Joppa, M., Fröhlich, J., Majschak, J.-P., and Rüdiger, F., On the concept of computational fluid dynamics-based prediction of cleaning for film-like soils, *Heat Transfer Engineering*, vol. 43, pp. 1406–1415, 2022.

[22] Joppa, M., Köhler, H., Rüdiger, F., Majschak, J.-P., and Fröhlich, J., Experiments and simulations on the cleaning of a swellable soil in plane channel flow, *Heat Transfer Engineering*, vol. 38, pp. 786–795, 2017.

[23] Golla, C., Freiherr Marschall, W., Kricke, S., Rüdiger, F., Köhler, H., and Fröhlich, J., Identification of cleaning mechanism by using neural networks, *Food and Bioproducts Processing*, pp. 86–102, 2023.

[24] Himashree, P., Sengar, A. S., and Sunil, C. K., Food thickening agents: Sources, chemistry, properties and applications - A review, *Int. J. Gastron. Food Sci.*, vol. 27, 2022.

[25] Toure, Y., Mabon, N., and Sindic, M., Soil model systems used to assess fouling, soil adherence and surface cleanability in the laboratory. A review, *Biotechnology, Agronomy, Society and Environment*, vol. 17, pp. 527–539, 2013.

[26] Joppa, M., Hanisch, T., and Mauermann, M., Experimental study on the cleaning effect of discontinuous sprays, *Food and Bioproducts Processing*, vol. 136, pp. 123–129, 2022.

[27] Lelièvre, C., Legentilhomme, P., Gaucher, C., Legrand, J., Faille, C., and Bénézech, T., Cleaning in place: effect of local wall shear stress variation on bacterial removal from stainless steel equipment, *Chemical Engineering Science*, pp. 1287–1297, 2002.

[28] Kricke, S., Böttcher, K., Zahn, S., Majschak, J.-P., and Rohm, H., Effect of physicochemical properties of native starches on cleaning in falling film and plane channel flow experiments, *Heat Transfer Engineering*, vol. 43, pp. 1416–1425, 2022.

[29] Myles, A. J., Feudale, R. N., Liu, Y., Woody, N. A., and Brown, S. D., An introduction to decision tree modeling, *Journal of Chemometrics*, pp. 275–285, 2004.

[30] Kotsiantis, S. B., Decision trees: a recent overview, *Artif Intell Rev*, pp. 261–283, 2019.

[31] Du, C.-J., and Sun, D.-W., Learning techniques used in computer vision for food quality evaluation: A review, *Journal of Food Engineering*, pp. 39–55, 2006.

[32] Lin, Y., Ma, J., Wang, Q., and Sun, D.-W., Applications of machine learning techniques for enhancing nondestructive food quality and safety detection, *Critical Reviews in Food Science and Nutrition*, vol. 63, pp. 1649–1669, 2023.

[33] Debska, B., and Guzowska-Swider, B., Decision trees in selection of featured determined food quality, *Analytica Chimica Acta*, pp. 261–271, 2011.

[34] Joppa, M., Köhler, H., Rüdiger, F., Majschak, J.-P., and Fröhlich, J., Prediction of cleaning by means of computational fluid dynamics: Implication of the pre-wetting of a swellable soil, *Heat Transfer Engineering*, vol. 41, pp. 178–188, 2020.

ISBN: 978-0-9984188-3-4; Published online www.heatexchanger-fouling.com