Heliyon

ELSEVIER

Received:
8 April 2017
Revised:
15 December 2017
Accepted:
18 February 2018

Cite as:
George J. Besseris. Taguchi-generalized regression neural network micro-screening for physical and sensory characteristics of bread. Heliyon 4 (2018) e00551.
doi: 10.1016/j.heliyon.2018.e00551

Taguchi-generalized regression neural network micro-screening for physical and sensory characteristics of bread

George J. Besseris [a,b,*]

[a] *Advanced Industrial & Manufacturing Systems Program, Mechanical Engineering Department, Piraeus University of Applied Sciences, Greece*

[b] *Kingston University, London, UK*

[*] Corresponding author.

E-mail address: besseris@puas.gr (G.J. Besseris).

Abstract

Generalized regression neural networks (GRNN) may act as crowdsourcing cognitive agents to screen small, dense and complex datasets. The concurrent screening and optimization of several complex physical and sensory traits of bread is developed using a structured Taguchi-type micro-mining technique. A novel product outlook is offered to industrial operations to cover separate aspects of smart product design, engineering and marketing. Four controlling factors were selected to be modulated directly on a modern production line: 1) the dough weight, 2) the proofing time, 3) the baking time, and 4) the oven zone temperatures. Concentrated experimental recipes were programmed using the Taguchi-type $L_9(3^4)$ OA-sampler to detect potentially non-linear multi-response tendencies. The fused behavior of the master-ranked bread characteristics behavior was smart sampled with GRNN-crowdsourcing and robust analysis. It was found that the combination of the oven zone temperatures to play a highly influential role in all investigated scenarios. Moreover, the oven zone temperatures and the dough weight appeared to be instrumental when attempting

https://doi.org/10.1016/j.heliyon.2018.e00551

to synchronously adjusting all four physical characteristics. The optimal oven-zone temperature setting for concurrent screening-and-optimization was found to be 270−240 °C. The optimized (median) responses for loaf weight, moisture, height, width, color, flavor, crumb structure, softness, and elasticity are: 782 g, 34.8 %, 9.36 cm, 10.41 cm, 6.6, 7.2, 7.6, 7.3, and 7.0, respectively.

Keywords: Industrial engineering, Food science

1. Introduction

1.1. The need of neural networks in small scale experimentation

Probabilistic Neural Networks (NN) have provided a useful solver platform for a wide range of applications in data mining and knowledge discovery (Murphy, 2012). Still, there are crucial fields where NNs' capabilities have not been explored in as much depth yet. One such field deals with deciphering structured experimentation outcomes (Skrjanc, 2015; Tortum et al., 2007). It is of high demand in industrial product development and process improvement as well as in discovery projects. Practicality, timeliness, programmability and economics are the primary drivers that call for structured trial recipes. A great proponent of quick-and-easy "micro-mining" in production operations has been the internationally-known quality-guru Dr Genichi Taguchi (Taguchi et al., 2000, 2004). The great payback for engaging optimally organized trials in Japanese manufacturing has been well acknowledged. Taguchi methods promote rapid combinatorics plans deployed through Design of Experiments (DOE) aiming at shortening the knowledge discovery cycle under real working conditions and tight economic constraints. Taguchi proposed a series of compact-structure data-arrangements to accelerate the capturing of complex input-output relationships. Those predefined trial plans, known as orthogonal arrays (OAs), were to dramatically reduce the cost and time to collect meaningful data while offering a well-balanced view of the investigated physical phenomena. Searching the Scopus database, one will come across several thousand articles and reviews that refer to tools and techniques adopting Taguchi methods. However, reaching to solid judgements from analyzing standardized (OA-programmed) datasets has been an on-going topic for several decades. Remarkably, more than a hundred statistical methods have been published with the ultimate objective to robustify OA-solver predictability. Substantial hassle has been geared toward interpreting saturated Taguchi OA-datasets since OAs maximize the potential usability of the information content for the given data collection effort. It is pragmatic needs in operations that boost demand for screening non-linear, multi-parameter, multi-response, multi-data-type, saturated Taguchi-type OA-datasets. Surprisingly, there is a paradox with respect to the compatibility of Taguchi methods with NN-based solvers. Although Taguchi methods have been profitably used in the past to optimize neural network

https://doi.org/10.1016/j.heliyon.2018.e00551

performance, the reverse has not been proven to be true. Indeed, neural networks have not found a niche yet in screening complex Taguchi-type datasets (Cohn, 1996; Hering and Simandl, 2010; Issanchou, and Gauchi, 2008).

1.2. Collective intelligence and neural networks in small data

In this work, we introduce the idea of collective intelligence to be gathered from a crowd which is formed by participants that are machine learning entities (Abrahamson et al., 2013; Howe, 2009; Surowiecki, 2005). Each of the crowd's NN-members is capable of delivering a "private" decision in screening Taguchi-type input-output relationships. We show how the aggregated "learning from a crowd" may be mapped to reliable conjoint (pooled) judgments from a horde of "RankBots". As a "RankBot", we define a machine-learning solution, which is extracted from manipulating Taguchi-type OA-input/output datasets. At the core of the "supervised learning" of the RankBots rests a conventional NN-solver strategically selected to convert "small data". Consequently, the generalized regression neural networks (GRNN) are plausibly picked to be the standard NN-engine to represent the converting capacity for each of the individual members in the "RankBot community". The end deliverable of a RankBot is the strength of the hierarchy results of the screened effects as they are gleaned from the GRNN-solver sensitivity-analysis report.

Full-array DOE screening requires a fuzzy solver with high aptitude in the ability to handle: 1) "scarce data" and 2) "uncertain uncertainty". The inherent data scarcity which is unavoidably encountered in Taguchi-type OA datasets is a restricting condition that opposes the potential usefulness behind the adoption of the typical machine learning philosophy. However, it might be reasonably remedied by employing GRNNs. The GRNNs are propped up for stringent function approximation and classification problems. Additionally, GRNNs manage to sustain network performance when there is a demand for exclusively handling of small datasets (Specht, 1995). GRNNs have demonstrated a great efficiency in translating mini-datasets, sometimes 200,000 faster than backpropagating NNs while tolerating erroneous samples (Specht, 1995). The ultra-high speed of training is related to the special built-in feature that permits the parallel GRNN-processing of the sparse dataset. GRNN-generated decision surfaces approach the Bayesian optimal, a feature pivotal for robust decision making. The GRNN is an one-pass learning algorithm that is well-tested and known to be suitable for implementation when the sparse dataset is not anticipated to follow linearity, hence making it perfectly ideal to fit non-linear Taguchi-type OA datasets (Ozyildirim and Avci, 2013, 2016). Optimal utilization of non-linear OAs is achieved when a selected OA plan is loaded up with the maximum number of examined effects that it is designed to carry, i.e. it becomes saturated (Atkinson et al., 2007). Saturated OA-datasets are simple input-output relationships where no degrees of freedom can be spared for estimating an experimental uncertainty. This

https://doi.org/10.1016/j.heliyon.2018.e00551

is because all degrees of freedom are absorbed by the investigated parameter settings. This becomes a major source for spawning "uncertain uncertainty". In practice, the condition of "uncertain uncertainty" short-circuits all mainstream pure statistical data treatments, like ANOVA/MANOVA etc. Consequently, it strips off the inferential character of the outcomes of ordinary multivariate techniques downgrading them to mere (subjective) descriptive statistics. Still, a single supervised GRNN run is not expected to furnish a single terminal solution. This is not possible for two reasons: 1) "uncertain uncertainty" remains unresolved from a single GRNN-based solution and 2) the GRNN piles up additional uncertainty on the prediction since subsetting reserves part of the OA-dataset for testing/verification. Because of the latter complication, the compactness of an OA-dataset is bound to be destroyed as the unique trial recipes need to be redistributed in two more phases of NN-processing following the initial NN-training phase. The destruction of the orthogonality of the OA planner magnifies the mysterious influence of "uncertain uncertainty" in the GRNN data modelling process. Training in a single GRNN run has to proceed while shedding an unknown amount of information which was gained from the strict regimented OA trials and was intended for model fitting. We note that when we refer to "training in a single GRNN run" it is meant equivalently to "training a single RankBot".

For a crowd to be wise, it needs to conform to four fundamental criteria. The four perfect conditions are: 1) diversity of opinion, 2) independence, 3) decentralization, and 4) aggregation. All four elements are met here in our approach. Diversity of opinion is automatically enforced in the Rankbot crowd because the trial (input/output) data that GRNN is fed with for training are only a subset of the uniquely executed OA-trial recipes. The training dataset for each individual RankBot is randomly determined and hence the maximum of diversity of opinion is attained. The second condition – independence – also automatically holds since each RankBot is by design an independent solution with no ties or interactions to be allowed to influence or to be influenced by the results of other RankBots. Since each trial recipe offers only a unique piece of information, the combination of recipes that eventually comprise the training data subset for each individual RankBot "specializes on the local knowledge" of those particular RankBots involved. Therefore, the condition of decentralization is also present. Finally, we present the collective decision of the RankBot crowd using simple robust depictions of box-plotted results which is a preferred mechanism to ciphering "private" RankBot judgements. Thus, the condition of the fourth element – aggregation of RankBot opinions – is met.

It may be construed as an inherent impediment in our approach the fact that there is actually no formal mechanism that generates a (prescribed) crowd in order to extract wisdom from. However, a virtual crowd is assembled "on the way" as groups of RankBots complete a series of individual astute data screenings. The convenient way to accomplish this is to start analyzing a prudent pilot sample and based on that sample then to predict a wise (terminal) crowd. Collecting sufficient individual

https://doi.org/10.1016/j.heliyon.2018.e00551

RankBot opinions should drive to the converging of their aggregated outcomes. The practicality to reach to intelligent decisions by actually approximating a crowd, i.e. surveying only an adequate subset of the crowd (a subcrowd), is a subject that has been previously addressed (Ertekin et al., 2014). Contrary to the traditional notion that it is imperative that a crowd must exist before subsetting it, our approach reverses this idea by resorting to the novel, yet distinct, feature of "building a crowd through". In a nutshell, the availability of a presumptive crowd base is not required beforehand and the final sample base is tailored to the requisites of the specific problem. Our crowdsourcing plan permits a simple collective assessment and prediction which does not impose any comparison against a "golden standard" (Malone and Bernstein, 2015). A key advantage is that it expects uniform attentiveness across the members of the "RankBot subcrowd". Consequently, latency across its members is also uniform and its minimization is sustained through the particular selection of the fast GRNN solver. Moreover, the overall methodology receives a boost in overall execution because the RankBots do not have to undergo assessment on "gold-standard" datasets. No RankBot judgements will have to be dismissed because they failed to meet a minimum "gold-standard" specified performance. Such negation would either demand opinion replacement by soliciting extra RankBot judgements or, in a forgiving scenario, opting to have sub-par RankBot opinion weighted and downplayed. Consequently, either scenario would lead in delayed or compromised decision-making. Our method exploits the redundancy of RankBot judgments which materializes in a majority voting where there are several nominations. The data generated from RankBot annotations are democratic in nature and hence amenable to translation by ordinary robust statistical inference methods.

In our methodology, we claim no intent to ameliorate the prediction accuracy of the GRNN model by directly tweaking in some manner its inner workings. Instead, our strategy is to ensnare and quantify all unknown and unknowable uncertainty that is generated from the unique partnering of the collected OA-planned trials as surfaced after a barrage of GRNN model-fitting attempts. Practically, this is achieved by permitting an army of RankBots to perturb the OA dataset structure from a number of angles that are deemed sufficient only after the terminal RankBot subcrowd has been statistically determined. It should be highlighted the fact that our approach reconciles the two cultures of statistical modeling (Breiman, 2001). This is because collective intelligence is needed to be gathered from small, dense and complex data using the primitive machine learning capabilities of a GRNN application. This is true because we are constrained to convert information well outside the realms of large-data theory, where one would expect undeniably most machine learning algorithms are bound to thrive. Nevertheless, the machine learning culture is present in our approach. However, blurred information prevails, since there is no much data to sift through. Thus, machine learning is compelled to appear to act dumb. Henceforth, the stochastic precision from the aggregation of the RankBot judgements is the only

https://doi.org/10.1016/j.heliyon.2018.e00551

compass that the RankBot views significantly converge. To stamp significance on the effect hierarchy status, we rely on simple robust (boxplot) theory and thus the second culture has been honored too.

1.3. The bread as a popular highly-complex product

Bread has been among the most prevalent food preparations in the world since antiquity (Mondal and Datta, 2008). The bread-making procedure has evolved over the centuries from crafting loaves in home-made and artisan wood-fired ovens to mass producing on large automated systems (Cauvain and Young, 2006; Singh and Heldman, 2013). This progress has necessitated standardization of the bread-making processes while in the meantime homogenizing terminal bread characteristics that suit sensory expectations for a broad consumer base (Decock and Cappelle, 2005; Heenan et al., 2009; Therdthai and Zhou, 2003). Published work that epitomizes techniques for improving bread quality from organized − physical and sensory − trial data is rather rare. Reaching meaningful conclusions for engineered bread formulations ordinarily demand a substantial volume of trials on large-scale operations as well as expert evaluation on sensory trait preferences (Gao et al., 2015). Such realizations may only be consummated through elaborate experiments on industrial facilities. Nevertheless, producers are usually wary about surrendering precious machinery availability for conducting intricate experiments. Perhaps, the notion of disturbing a busy production line may be considered a 'risky' venture after all. To remedy this dilemma, current literature on bread-baking processes attempts to explore difficult bread-making phenomena through simulations (Chhanwal et al., 2012; Feyissa et al., 2012; Purlis, 2011). Particularly scarce to retrieve is industrial bread-processing research that accomplishes synchronous harmonization of physical and sensory traits. The need for specialized knowledge and data-driven mining techniques to describe bread quality has been well exposed both in theory and practice (Della Valle et al., 2014; Hadiyanto et al., 2008; Liu and Scanlon, 2003; Parimala and Sudha, 2015; Zanoni et al., 1993, 1994; Zhang and Datta, 2006). Data-driven product development − supported by modern data-mining and knowledge discovery tools − is well within the broader future scope of food engineering in general as this field is called upon to make the most of the innovative information technology (Hubert et al., 2016; Saguy et al., 2013; Thakur et al., 2010).

Developing bread products implicates highly complex activities. Dough materials undergo vast physical and biological transformations (Rask, 1989; Scanlon and Zghal, 2001). After the mixture formulation has been determined, a barrage of convoluted processes are executed that involve: mixing, kneading, portioning, rounding, pre-molding, pre-proofing, molding, proofing, baking, cooling, slicing and packaging. The continuous interplay of bio-rheological and chemo-physical morphing immensely sensitizes a loaf of bread permitting any opportunity for a process

https://doi.org/10.1016/j.heliyon.2018.e00551

inefficiency to be imprinted on the final product characteristics (Autio et al., 2001; Dobraszczyk and Morgenstern, 2003; Dobraszczyk, 2004; Jefferson et al., 2006; Martinez and Gomez, 2017; Purlis and Salvadori, 2007; Sliwinski et al., 2004; Stojceska and Butler, 2012; Vanin et al., 2009). The design of bread products may perhaps be oblivious to complicating phenomena such as the bread collapse which often requires more sophisticated treatment (Rzigue et al., 2016). It may be contemplated that the propensities of a poly-mechanized modern bread product might be propitiously screened and analyzed by making effective use of flexible and robust techniques. Unknown and unknowable intrusions may interfere with bread-baking, hence rendering imperative the utilization of a rigorous profiler to carry out the task of the data analysis. To confidently interpret bread features, the selected statistical profiler should be capable to intelligently outmaneuver any sampled data oddities. It is a particularly welcomed attribute for an agile data-driven screening method to ensure guarding against opportunities of compromised integrity in the collected dataset.

To gather exploitable industrial data, it is vital that the selected trial programming to be cogent by adhering to a short, structured and balanced schedule. Taguchi-type orthogonal arrays (OAs) provide the sampling medium to economically organize the trial recipes such that to accelerate the overall experimental effort (Taguchi et al., 2000, 2004). OAs are also flexible by facilitating the simultaneous testing of numerical and categorical inputs. In general, Taguchi-type OA samplers have been well accepted in food engineering applications (Besseris, 2015; Das Mohapatra et al., 2009; Oztop et al., 2007; Pouliou and Besseris, 2013; Sharif et al., 2014; Tasirin et al., 2007). However, the subsequent task of data conversion requires intensive and careful manipulation when addressing to complex materials like breads. This is because the advantageous assumptions of data normality and linearity may not be suitable for probing phenomena which are associated with bread-making. The multi-phase processing of doughs elevates the chances that unknown and unknowable intrusions may fortuitously dart in at any moment during data collection and blur the trial observations. Modern fuzzy-based techniques appear to be more resilient in dealing with multifaceted uncertainty because they are more tolerant to the lack of exactness when defining a screening problem (Besseris, 2014a; Lamrini et al., 2012; Ndiaye et al., 2009; Rousu et al., 2003). Reasonably then, we are motivated to introduce intelligent sampling in our development as a preferred option to homogenize the various sources of uncertainty.

1.4. The purpose of this study

The purpose of this study is to synthesize an intelligent instrument to aid the 'smart-and-robust' data analysis of complex processes. Small data is at the crux of the study as a strategy for figuring out rapidly the strength of screened effects. The concepts of

https://doi.org/10.1016/j.heliyon.2018.e00551

GRNN and Wisdom of Crowds complement each other in an attempt to wrap-up strong influences from structured DOE mini-data. The Wisdom of Crowds may be actually represented by a subcrowd that carries the satisfactory information content. Technically, a subcrowd is and may be also referred equivalently to as smart sample. We are motivated to showcase a complex and yet a much familiar case which is drawn from the breadmaking process since bread is an every-day consumed product to billions of people around the globe. The stochastic interpretation is based on terminal bread properties. Relying on systematically collected observations, all trials are executed on massive operations. The data reduction process involves a strategy to concurrently decipher in an 'intelligently robust' fashion the multi-response multi-factorial outcomes of non-linear product/process screening. The suggested approach is agile since it is apt to accommodating a wide spectrum of complex characteristics that may be expressed in numerical and categorical forms. In this work, we diversify bread-baking screening in three distinct scenarios by reflecting upon the product from three crucial stand points: 1) product engineering, 2) product marketing and 3) product design. Regardless the scope of the screening, we propose a fast multi-response multi-factorial profiler which is equipped to resolve potential non-linearity in the examined effects while maintaining a distribution-free probabilistic framework. Non-linear Taguchi-type OA-samplers swiftly program and adeptly compile the compact dataset that includes multiple physical bread properties along with a comprehensive sensory performance. The investigated white pan bread data that will be illustrated in the case study have been exclusively accumulated from line operations in a large baked-goods enterprise. The master-ranking transformation concept is demonstrated to provide the homogenization medium to simultaneously treat a score of responses from various origins. Moreover, the master-ranking tactic consolidates the differentiated groups of weighted physical and sensory responses in order to simplify the concurrent smart-and-robust profiling effort (Besseris, 2012, 2013a).

2. Materials and methods

2.1. White pan bread materials

The investigated product is a white pan-bread brand. All experiments were conducted on production-line machinery and equipment in the premises of a world-class enterprise specializing in bakery goods. Dough formulation ingredients were not allowed to be altered as they constituted proprietary information. The collected dataset was to assist management in gleaning information from three distinct viewpoints. The first aspect regards terminal product profiling to be solely based on product engineering concerns. Therefore, the four physical responses of immediate impact that were conferred upon for researching were: 1) the bread weight (BW), 2) the bread moisture (M), 3) the bread height (H) and 4) the bread width (W). BW and M are characteristics

https://doi.org/10.1016/j.heliyon.2018.e00551

that are routinely monitored in conjunction to federal regulations. The four physical characteristics belong in the 'nominal-the-best' category (Taguchi et al., 2000). Product specification limits were designated as follows: 1) BW: 765 ± 20 g, 2) M: 34.5 ± 1.5%, 3) H: 9.70 ± 0.30 cm, and 4) W: 10.5 ± 0.5 cm. Furthermore, priority (importance) weight contribution for each characteristic was assigned based on the kaizen team deliberations: 1) BW: 40%, 2) M: 30%. 3) H: 20%, and 4) W: 10%.

The marketing performance of the white pan bread product relies heavily on its sensory characteristics. To cover separately this aspect, the five tracked components were: 1) color (CL), 2) flavor (FL), 3) crumb structure (CR), 4) softness (SF), and 5) elasticity (EL). A ten-grade Likert scale was used to score the performance for each trait (minimum rating = 1 to maximum rating = 10). All five sensory traits identify to the 'larger-is-better' category (Taguchi et al., 2000). A five-member expert panel representing the aforementioned departments was surveyed on sampled bread loaves. Each member recorded their marks for each conducted trial run, separately. The scores were accumulated to a total rating for each individual repetition/replicate run. The corresponding priority weights for the sensory traits were allocated as follows based on past experience: 1) CL: 25%, 2) FL: 20%, 3) CR: 10%, 4) SF: 20%, and EL: 25%.

The third aspect relates to the overall product design. It incorporates information by joining physical and sensory product performances from the preceding two scenarios. The priority weights have been allotted in this case such that the physical to sensory ratio to be 30/70.

The kaizen improvement team reasoned that the less explored − and hence less understood − controlling (process) factors should be investigated. The final list was: 1) the dough weight (DW) in g, 2) the proofing time (PT) in min, 3) the baking time (BT) in min, and 4) the oven zone temperatures (BTP) in °C. This decision was reached after realizing the lack of any previous integrated research in trade and scientific literature that might involve the examination of those four controlling factors. The restriction to proceed with testing merely four factors was balanced by the amount of experimentation which was permitted to replace operational availability. It should also be noted that while the first three controlling factors are continuous numerical variables, BTP will be treated as a categorical variable because it expresses empirically the selected oven temperature settings in pairs for both upper and lower zones. The collected dataset which is segregated in terms of physical characteristics (Table 1) and sensory traits (Table 2) may be accessed in Athanasiadou (2010). The same source details the scoring which is awarded by each expert panel member for each executed repetition/replicate trial run.

Because of the limited access to materials and machinery time, the study has been constrained to a minimal sampling effort in order to survey repeatability and reproducibility. Thus, repeats and replicates were only duplicated.

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 1. Original white pan bread data for the four physical characteristics (Athanasiadou, 2010).

Run #	BW 1A	BW 1B	BW 2A	BW 2B	M 1A	M 1B	M 2A	M 2B	H 1A	H 1B	H 2A	H 2B	W 1A	W 1B	W 2A	W 2B
1	755	759	757	763	34.8	34.7	34.6	34.9	9.65	9.13	9.3	9.53	10.19	10.33	10.37	10.58
2	759	755	752	754	35	34.2	34.4	34.8	9.32	9.35	9.49	9.26	10.73	10.68	10.67	10.66
3	751	753	749	753	35	33.9	33.9	34.4	9.62	9.28	9.42	9.08	10.05	10.24	10.41	10.47
4	784	780	781	789	35.4	35.2	35	34.9	9.54	9.49	9.52	9.45	10.11	10.21	10.2	10.26
5	760	762	754	761	33.6	33.9	33.3	34.7	9.5	9.53	9.37	9.59	10.31	10.54	10.45	10.35
6	783	782	777	785	35.7	35.4	34.3	34.9	9.44	9.44	9.39	9.36	10.33	10.41	10.41	10.53
7	791	794	790	782	35.1	34.8	34.3	34.7	9.28	9.49	9.35	9.34	10.26	10.29	10.34	10.26
8	807	804	799	801	36	35	34.8	35.6	9.59	9.4	9.49	9.51	10.05	10.2	10.17	10.27
9	783	786	783	787	33.3	33.9	34.4	34.7	9.5	9.51	9.49	9.49	10.25	10.47	10.57	10.51

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 2. Original white pan bread data for the five sensory traits (Athanasiadou, 2010).

Run #	CL 1A	CL 1B	CL 2A	CL 2B	FL 1A	FL 1B	FL 2A	FL 2B	CR 1A	CR 1B	CR 2A	CR 2B	SF 1A	SF 1B	SF 2A	SF 2B	EL 1A	EL 1B	EL 2A	EL 2B
1	25	27	31	30	32	28	35	31	32	38	37	39	31	33	30	31	33	28	31	31
2	33	31	34	32	36	35	36	35	38	43	40	38	37	39	40	39	34	33	38	34
3	29	34	31	31	33	31	27	29	37	39	35	39	33	33	32	30	36	34	34	33
4	34	31	30	35	38	34	35	36	42	39	39	36	38	35	38	35	37	36	33	33
5	25	26	25	26	32	31	28	30	38	36	31	34	26	32	27	30	25	28	28	28
6	29	35	33	33	37	36	36	36	34	35	39	35	37	36	36	38	33	37	33	35
7	32	35	33	32	33	37	35	35	36	40	38	39	33	36	36	36	35	36	37	35
8	32	33	34	33	40	39	37	36	39	41	39	41	41	38	37	38	38	33	32	35
9	29	29	29	30	32	30	31	33	34	35	33	36	36	35	33	35	36	34	34	35

https://doi.org/10.1016/j.heliyon.2018.e00551

2.2. Non-linear Taguchi-type OA-sampling

It is imperative to minimize the non-productive consumption of machinery availability on large-scale operations. Therefore, any experimentation on production line should call for a short schedule of trials to be rapidly performed. To program structured and balanced experiments that pack multi-effect variability information, the Taguchi-type OA-samplers furnish convenience in planning the required trials (Besseris, 2013b; Taguchi et al., 2000; Taguchi et al., 2004). Moreover, additional savings in time and materials are anticipated when profiling simultaneously for possible effect non-linearity (Besseris, 2014b). An appropriate non-linear OA-sampler that tracks down each examined effect on at least three predetermined settings while compounding variation from all four controlling factors on each measurement is the $L_9(3^4)$ OA. The resulting nine recipes that combine in saturated mode the four input settings have been listed in Table 3.

2.3. Smart-sampling (subcrowding) the condensed multiple responses

For each of the three screening scenarios, there is a procedure that involves the reduction of: 1) the number of trial repetitions, 2) the number of replicates, and 3) the number of weighted responses to form a single cumulative response. The condensation process relies on a combination of ranking and fusing the dataset columns and it is outlined in Section 3 (Theory/Calculation). The sequential transmutation of the originally replicated multi-response dataset terminates to a single homogenized 'unreplicated-saturated' response (Besseris, 2014a; Milliken and Johnson, 1989, 2009). This is repeated for each of the three respective scenarios separately. For each scenario, we resort to an intelligent engine to de-fuzzify the

Table 3. Taguchi-type OA-sampling ($L_9(3^4)$ OA) schedule for the white pan bread trials (Athanasiadou, 2010).

Run #	DW	PT	BT	BTP
1	880	50	34	280−250
2	880	55	37	270−240
3	880	60	40	260−230
4	900	50	37	260−230
5	900	55	40	280−250
6	900	60	34	270−240
7	920	50	40	270−240
8	920	55	34	260−230
9	920	60	37	280−250

https://doi.org/10.1016/j.heliyon.2018.e00551

inherent messiness in the transformed non-linear OA-dataset. The general-regression neural-network (GRNN) processor is employed to properly deal with the precarious smallness which is evidenced in all three versions of the condensed data (Murphy, 2012; Specht, 1990, 1991). In Fig. 1, we depict the implemented GRNN topology that relates each of the four considered inputs to the respective transformed output vector according to the type of profiling: 1) the weighted physical-characteristic screening, 2) the weighted sensory-trait screening or 3) the weighted synchronous screening of physical and sensory features. The three respective vectors that capture in each screening phase the total response fluctuation are: the weighted sum of squared ranks for physical characteristics (wSSRp), the weighted sum of squared ranks for sensory traits (wSSRs) and the weighted sum of squared master ranks (wSSMR). During the smart sampling (subcrowding) phase, Rankbot opinions (repeated independent GRNN runs) generate an effect-hierarchy list (H_j: j = 1, 2, 3…N; N = total number of Rankbot opinions). This list is created by sequentially appending the output rankings of the effects from the GRNN sensitivity analysis report (Besseris, 2015). Therefore, the input for each RankBot is the OA arrangement of Table 3 and any one of the three vectors, wSSRp, wSSRs or wSSMR. This means that three separate subcrowds might need to be formed to delineate the three different situations. The information resulting from collecting all RankBot subcrowd opinions are then bucketed separately for the three cases to be analyzed robustly with the approach in Section 2.4. The idea of crowdsourcing the OA dataset with RankBots and the smart sample processing is portrayed in Fig. 2.

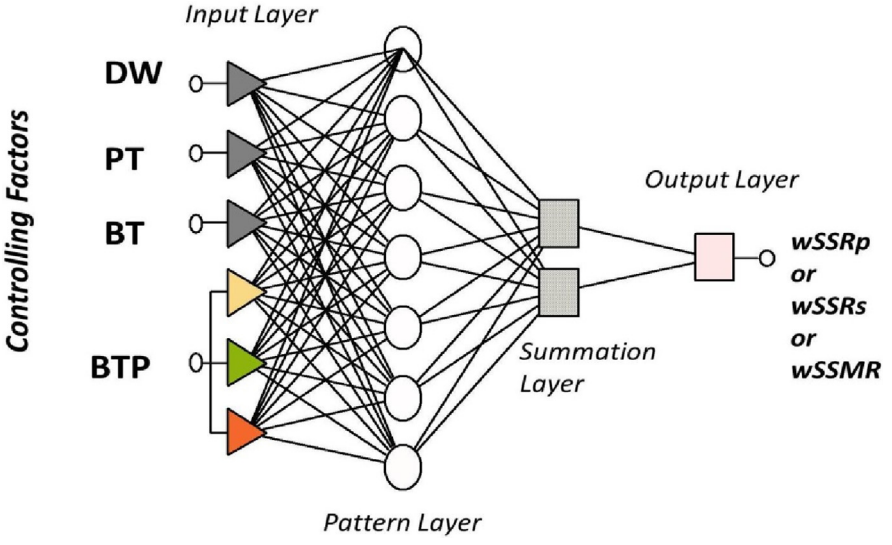


Fig. 1. The GRNN topology for creating smart samples (subcrowding process) − for all three investigated scenarios.

https://doi.org/10.1016/j.heliyon.2018.e00551

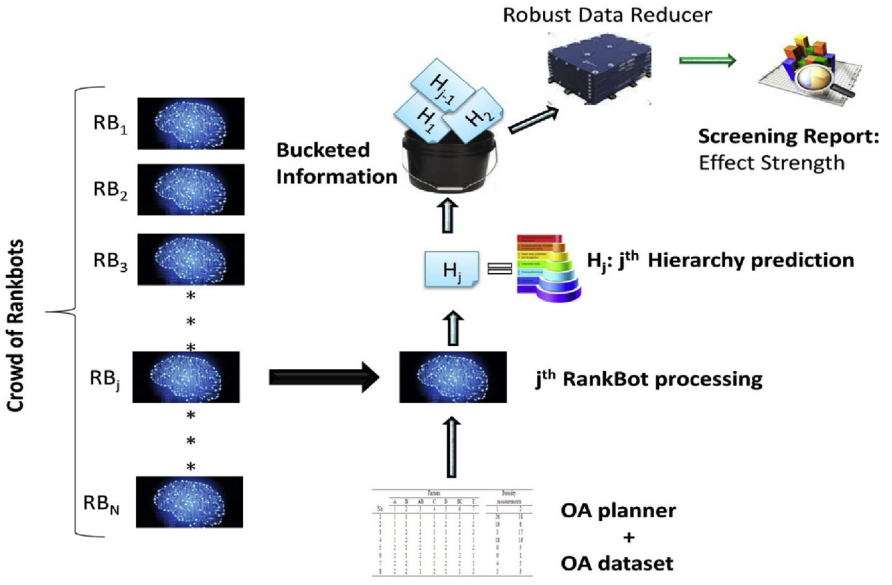


Fig. 2. Crowdsourcing the OA-dataset with RankBots and effect-hierarchy smart-sample processing.

To initialize a smart-sample dataset, the GRNN module is run thirty consecutive times (Besseris, 2014a). The adequacy of the smart-sample count is approximated by a cycle of checking, predicting and adjusting the smart-sample size until convergence to a final count value is achieved. Each time, the statistical estimator that feeds the test cycle is the largest of the standard deviation values of the examined effects. To rank statistically each effect, the margin of error – at a confidence interval of 95% – is designated at an absolute value of 0.5. This limiting value indicates a practical boundary which denotes the traversing to an adjacent rank position. When the series of prescribed transformations on the OA-dataset has been completed, the smart sample is robustly summarized by estimating the (distribution-free) central tendencies of the effects. The medians of the individual effects, at a 95%-confidence interval, are computed using the (one-sample) Wilcoxon's signed-rank test (Wilcox, 2010).

2.4. Robust and intelligent data analysis toolbox

For each of the three screening scenarios, the condensed $L_9(3^4)$ OA dataset as represented by the architecture in Fig. 1 is analyzed by the '*Intelligent Problem Solver*' (IPS) module. The IPS module is accessed from the submenu '*Neural Networks*' of the professional software Statistica 7.0 (StatSoft). The selected network-type option is '*GRNN*' with a maximum limit of tested networks set at 10,000. The preferred criterion for retaining networks is: 'Balance error against diversity'. Seven data entries are randomly allocated at the beginning for training and the balance is shared for selection and testing. Only the predicted effect hierarchy list – expressed in ranks – is

https://doi.org/10.1016/j.heliyon.2018.e00551

retained from the GRNN sensitivity analysis report. The rank ordering is initiated according to the largest GRNN network-error ratio value. Smart-sample adequacy is verified by iteratively comparing − and augmenting (if necessary) − the sample size with the predicted sample size (Besseris, 2015). This is accomplished by recursively calculating and determining each time the maximum standard deviation value among the examined effects. By inputting the maximum standard deviation value in the module '*Power and Sample Size for 1-sample t-test*' (MINITAB 17.0), the maximum smart-sample count is approximated. The median and its associated confidence interval estimations for the final smart sample are individually predicted for each effect using the '*1-sample Wilcoxon*' module (MINITAB 17.0).

For illustrational purposes, main-effects graphs and box-plot depictions are appended to further support the findings which are derived from the proposed screening method. Graphical portrayals for checking repetition consistency for the four physical characteristics have been prepared using linear regression fittings. Cross-correlations between the rank-condensed physical responses to detect possible associations between characteristics are computed using the Spearman's ρ correlation test. Linear cross-regressions between sensory traits have been carried out to inspect possible relationships on their cumulative values, i.e. by compounding both their repeats and replicates. To check the stability of the replicates and repeats of the sensory traits, the Spearman's ρ correlation test has been used. All required linear regression fittings and correlation estimations along with the main-effects graphs and box-plots have been obtained using the software package MINITAB 17.0.

3. Theory/calculation

3.1. Screening physical characteristics

The group of the investigated physical characteristics is homogenized twice in order to accrue stratified information from the two separate data layers − associated with trial repeats and replicates. Next, our tactic entails the weighted non-parametric merging of the multiple homogenized responses. This leads to shaping into a new vector quantity which contains the concentrated information of the studied influences as well as accounting for their repeatability and reproducibility. A generalized arrangement of the saturated $L_9(3^4)$ OA-dataset structure for the four controlling factors, *DW*, *PT*, *BT* and *BTP* is depicted in Fig. 3. The uniform representation of the three-level factor-settings is maintained for convenience in the generic formalism {1, 2, 3} (Taguchi et al., 2004). The conducted replication and repetition rounds are denoted as *R* and *r*, respectively. The matrix elements for the four physical responses are symbolized as (Eqs. 1, 2, 3, and (4)):

https://doi.org/10.1016/j.heliyon.2018.e00551

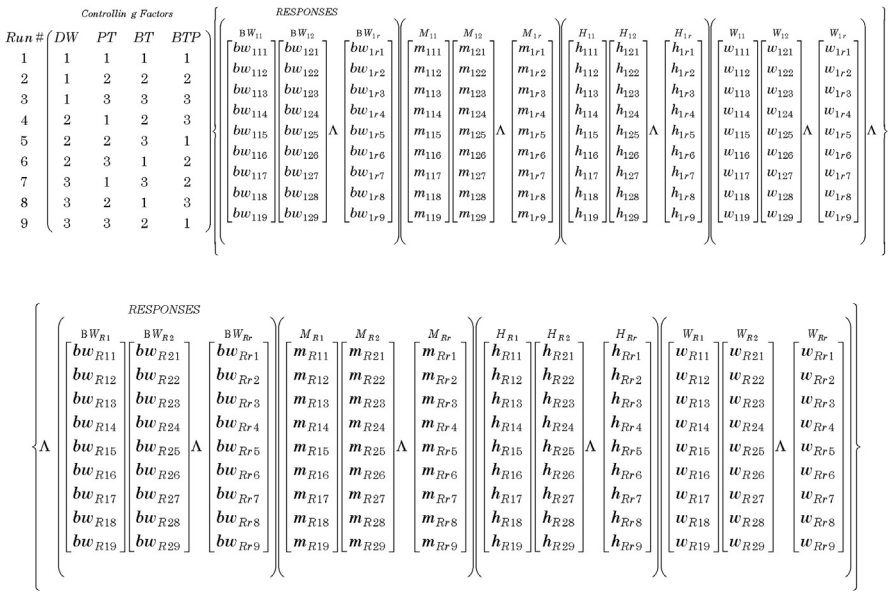


Fig. 3. The replicated saturated $L_9(3^4)$ OA dataset arrangement for the multi-response screening of the physical characteristics of the white pan bread.

$$BW_{ij} = \{bw_{ijk} | i = 1, 2, \ldots, R;\ j = 1, 2, \ldots, r;\ k = 1, 2, \ldots, 9\} \tag{1}$$

$$M_{ij} = \{m_{ijk} | i = 1, 2, \ldots, R;\ j = 1, 2, \ldots, r;\ k = 1, 2, \ldots, 9\} \tag{2}$$

$$H_{ij} = \{h_{ijk} | i = 1, 2, \ldots, R;\ j = 1, 2, \ldots, r;\ k = 1, 2, \ldots, 9\} \tag{3}$$

$$W_{ij} = \{w_{ijk} | i = 1, 2, \ldots, R;\ j = 1, 2, \ldots, r;\ k = 1, 2, \ldots, 9\} \tag{4}$$

For each of the four physical responses, the data processing is initiated by quantifying their absolute discrepancies from their corresponding target values (T_{BW}, T_M, T_H, T_W). It is followed by two successive phases of ranking and consolidation in order to compress both the repeats and the replicates. In all cases, by default, ordering precedence always awards rank '1' to the quantity possessing the smallest magnitude. In compact form, the sequence of the data manipulation steps culminates to the respective compressed vectors, bw'_k, m'_k, h'_k, w'_k (k = 1, 2,…, 9):

$$\begin{aligned} bw_{ijk} \rightarrow \left| bw_{ijk} - T_{BW} \right| = dbw_{ijk} \rightarrow rdbw_{i(jk)'} \rightarrow \sum_{j'=1}^{r} rdbw_{i(jk)'} \\ = srdbw_{ik'} \rightarrow rsrdbw_{(ik')'} \rightarrow \sum_{i'=1}^{R} rsrdbw_{(ik')'} = bw'_k \end{aligned} \tag{5}$$

https://doi.org/10.1016/j.heliyon.2018.e00551

$$m_{ijk} \to \left| m_{ijk} - T_M \right| = dm_{ijk} \to rdm_{i(jk)'} \to \sum_{j'=1}^{r} rdm_{i(jk)'}$$
$$= srdm_{ik'} \to rsrdm_{(ik')'} \to \sum_{i'=1}^{R} rsrdm_{(ik')'} = m'_k \tag{6}$$

$$h_{ijk} \to \left| h_{ijk} - T_H \right| = dh_{ijk} \to rdh_{i(jk)'} \to \sum_{j'=1}^{r} rdh_{i(jk)'}$$
$$= srdh_{ik'} \to rsrdh_{(ik')'} \to \sum_{i'=1}^{R} rsrdh_{(ik')'} = h'_k \tag{7}$$

$$w_{ijk} \to \left| w_{ijk} - T_W \right| = dw_{ijk} \to rdw_{i(jk)'} \to \sum_{j'=1}^{r} rdw_{i(jk)'}$$
$$= srdw_{ik'} \to rsrdw_{(ik')'} \to \sum_{i'=1}^{R} rsrdw_{(ik')'} = w'_k \tag{8}$$

The four homogenized vectors (Eqs. 5, 6, 7, and (8)) are rank-ordered once more to form the respective vectors: rbw'_k, rm'_k, rh'_k, rw'_k ($k = 1, 2, \ldots, 9$). The process of weighted compounding generates the condensed vector $wSSRp_k$ ($k = 1, 2, \ldots, 9$):

$$\left.\begin{matrix} bw'_k \to rbw'_k \\ m'_k \to rm'_k \\ h'_k \to rh'_k \\ w'_k \to rw'_k \end{matrix}\right\} \to \beta_{bw} \cdot rbw'^2_k + \beta_m \cdot rm'^2_k + \beta_h \cdot rh'^2_k + \beta_w \cdot rw'^2_k \to wSSRp_k \tag{9}$$

The weights (Eq. (9)), β_{dw}, β_m, β_h, and β_w, obey the normalization restriction: $\beta_{dw} + \beta_m + \beta_h + \beta_w = 1$. Thus, the finalized OA configuration that is fed to the intelligent processor is given in Fig. 4.

$$\begin{matrix} & & & \textit{Controllin g Factors} & & \\ \textit{Run \#} & & \textit{DW} & \textit{PT} & \textit{BT} & \textit{BTP} \end{matrix} \quad \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \textit{wSSRp} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{matrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 3 & 3 & 3 \\ 2 & 1 & 2 & 3 \\ 2 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{pmatrix} \begin{bmatrix} wSSRp_1 \\ wSSRp_2 \\ wSSRp_3 \\ wSSRp_4 \\ wSSRp_5 \\ wSSRp_6 \\ wSSRp_7 \\ wSSRp_8 \\ wSSRp_9 \end{bmatrix}$$

Fig. 4. The condensed (saturated) $L_9(3^4)$ OA dataset arrangement for the concurrent screening of the four physical characteristics.

https://doi.org/10.1016/j.heliyon.2018.e00551

3.2. Screening sensory traits

A generalized arrangement for the saturated $L_9(3^4)$ OA sensory dataset structure with the four controlling factors, *DW*, *PT*, *BT* and *BTP,* is depicted in Fig. 5. Following a similar rationale as in Section 3.1, we define the matrix elements for the five sensory traits (Eqs. 10, 11, 12, 13, and (14)):

$$CL_{ij} = \left\{ cl_{ijk} | i = 1,2,\ldots,R;\ j = 1,2,\ldots,r;\ k = 1,2,\ldots,9 \right\} \tag{10}$$

$$FL_{ij} = \left\{ fl_{ijk} | i = 1,2,\ldots,R;\ j = 1,2,\ldots,r;\ k = 1,2,\ldots,9 \right\} \tag{11}$$

$$CR_{ij} = \left\{ cr_{ijk} | i = 1,2,\ldots,R;\ j = 1,2,\ldots,r;\ k = 1,2,\ldots,9 \right\} \tag{12}$$

$$SF_{ij} = \left\{ sf_{ijk} | i = 1,2,\ldots,R;\ j = 1,2,\ldots,r;\ k = 1,2,\ldots,9 \right\} \tag{13}$$

$$EL_{ij} = \left\{ el_{ijk} | i = 1,2,\ldots,R;\ j = 1,2,\ldots,r;\ k = 1,2,\ldots,9 \right\} \tag{14}$$

Sequentially, repeats and replicates are aggregated to form their respective single sensory vectors:

$$cl_{ijk} \rightarrow \sum_{j=1}^{r} cl_{ijk} = cl'_{ik} \rightarrow \sum_{i=1}^{R} cl'_{ik} = cl''_{k} \tag{15}$$

$$fl_{ijk} \rightarrow \sum_{j=1}^{r} fl_{ijk} = fl'_{ik} \rightarrow \sum_{i=1}^{R} fl'_{ik} = fl''_{k} \tag{16}$$

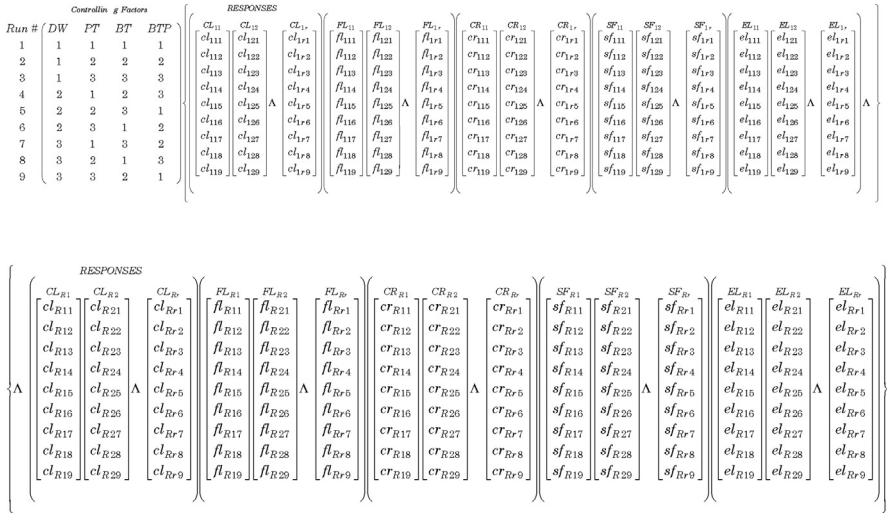


Fig. 5. A generalized $L_9(3^4)$ OA dataset arrangement for the concurrent screening of the five sensory traits.

https://doi.org/10.1016/j.heliyon.2018.e00551

$$cr_{ijk} \rightarrow \sum_{j=1}^{r} cr_{ijk} = cr'_{ik} \rightarrow \sum_{i=1}^{R} cr'_{ik} = cr''_{k} \tag{17}$$

$$sf_{ijk} \rightarrow \sum_{j=1}^{r} sf_{ijk} = sf'_{ik} \rightarrow \sum_{i=1}^{R} sf'_{ik} = sf''_{k} \tag{18}$$

$$el_{ijk} \rightarrow \sum_{j=1}^{r} el_{ijk} = el'_{ik} \rightarrow \sum_{i=1}^{R} el'_{ik} = el''_{k} \tag{19}$$

The five aggregate vectors (Eqs. 15, 16, 17, 18, and (19)) are rank-ordered to form rcl_k, rfl_k, rcr_k, rsf_k, and rel_k ($k = 1, 2,\ldots, 9$). In all cases, by default, ordering precedence always awards rank '1' to the quantity possessing the maximum magnitude. In turn, their weighted compounding generates the condensed vector $wSSRs_k$ ($k = 1, 2, \ldots, 9$):

$$\left.\begin{array}{l} cl''_k \rightarrow rcl_k \\ fl''_k \rightarrow rfl_k \\ cr''_k \rightarrow rcr_k \\ sf''_k \rightarrow rsf_k \\ el''_k \rightarrow rel_k \end{array}\right\} \rightarrow \beta_{cl} \cdot rcl_k^2 + \beta_{fl} \cdot rfl_k^2 + \beta_{cr} \cdot rcr_k^2 + \beta_{sf} \cdot rsf_k^2 + \beta_{el} \cdot rel_k^2 \rightarrow wSSRs_k \tag{20}$$

The weights (Eq. (20)) β_{cl}, β_{fl}, β_{cr}, β_{sf}, and β_{el} are normalized accordingly: $\beta_{cl} + \beta_{fl} + \beta_{cr} + \beta_{sf} + \beta_{el} = 1$. The finalized OA configuration that is fed to the intelligent processor is given in Fig. 6.

$$\begin{array}{c} \\ Run\,\# \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{array} \begin{array}{c} \textit{Controllin g Factors} \\ \begin{pmatrix} DW & PT & BT & BTP \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 3 & 3 & 3 \\ 2 & 1 & 2 & 3 \\ 2 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{pmatrix} \end{array} \begin{array}{c} wSSRs \\ \begin{bmatrix} wSSRs_1 \\ wSSRs_2 \\ wSSRs_3 \\ wSSRs_4 \\ wSSRs_5 \\ wSSRs_6 \\ wSSRs_7 \\ wSSRs_8 \\ wSSRs_9 \end{bmatrix} \end{array}$$

Fig. 6. The condensed saturated $L_9(3^4)$ OA dataset arrangement for the concurrent screening of the five sensory traits.

https://doi.org/10.1016/j.heliyon.2018.e00551

3.3. Concurrent weighted physical and sensory screening

Acquiring joint information from both types of responses becomes indispensable for improving overall product designing. It is a plausible extension to attempt streamlining physical feature integration with sensory performance. The two consolidated vectors, $wSSRp_k$ and $wSSRs_k$, carry the required information for completing this integration. However, they need to be rank-ordered such that to be aligned in the same scale. This is because the formation of the two vectors has been derived from compiling dissimilar response group sizes with variant weight criteria.

The corresponding master responses, MRp_k and MRs_k are prepared and their weighted squared rank summation generates the terminal vector $wSSMR_k$ (Besseris, 2012, 2013a):

$$\left.\begin{matrix} wSSRp_k \rightarrow MRp_k \\ wSSRs_k \rightarrow MRs_k \end{matrix}\right\} \rightarrow \beta_{MRp} \cdot MRp_k^2 + \beta_{MRs} \cdot MRs_k^2 \rightarrow wSSMR_k \tag{21}$$

The weights β_{MRp} and β_{MRs} (Eq. (21)) are normalized to form the constraint for the joint screening attempt: $\beta_{MRp} + \beta_{MRs} = 1$. The final OA configuration that is fed to the intelligent processor is given in Fig. 7. The complete sequence of data manipulation and decision steps have been outlined in Fig. 8.

It should be noted that small, dense and complex data will cause most of the machine learning algorithms to act really dumb. Unfortunately, this is true. We picked GRNN because will reasonably resist to express dumbness at small dataset requirements. By blending the responses, we reduce by 1/n (n = number of characteristics) the opportunities for the ML algorithm to be wandering across multiple blurred surfaces.

$$\begin{array}{c} \textit{Controllin g Factors} \\ \begin{array}{c} Run\,\# \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{matrix} \end{array} \begin{pmatrix} DW & PT & BT & BTP \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 3 & 3 & 3 \\ 2 & 1 & 2 & 3 \\ 2 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{pmatrix} \begin{array}{c} wSSMR \\ \begin{bmatrix} wSSMR_1 \\ wSSMR_2 \\ wSSMR_3 \\ wSSMR_4 \\ wSSMR_5 \\ wSSMR_6 \\ wSSMR_7 \\ wSSMR_8 \\ wSSMR_9 \end{bmatrix} \end{array} \end{array}$$

Fig. 7. The arrangement of the saturated $L_9(3^4)$ OA dataset for the concurrent (weighted) screening of physical and sensory bread responses.

https://doi.org/10.1016/j.heliyon.2018.e00551

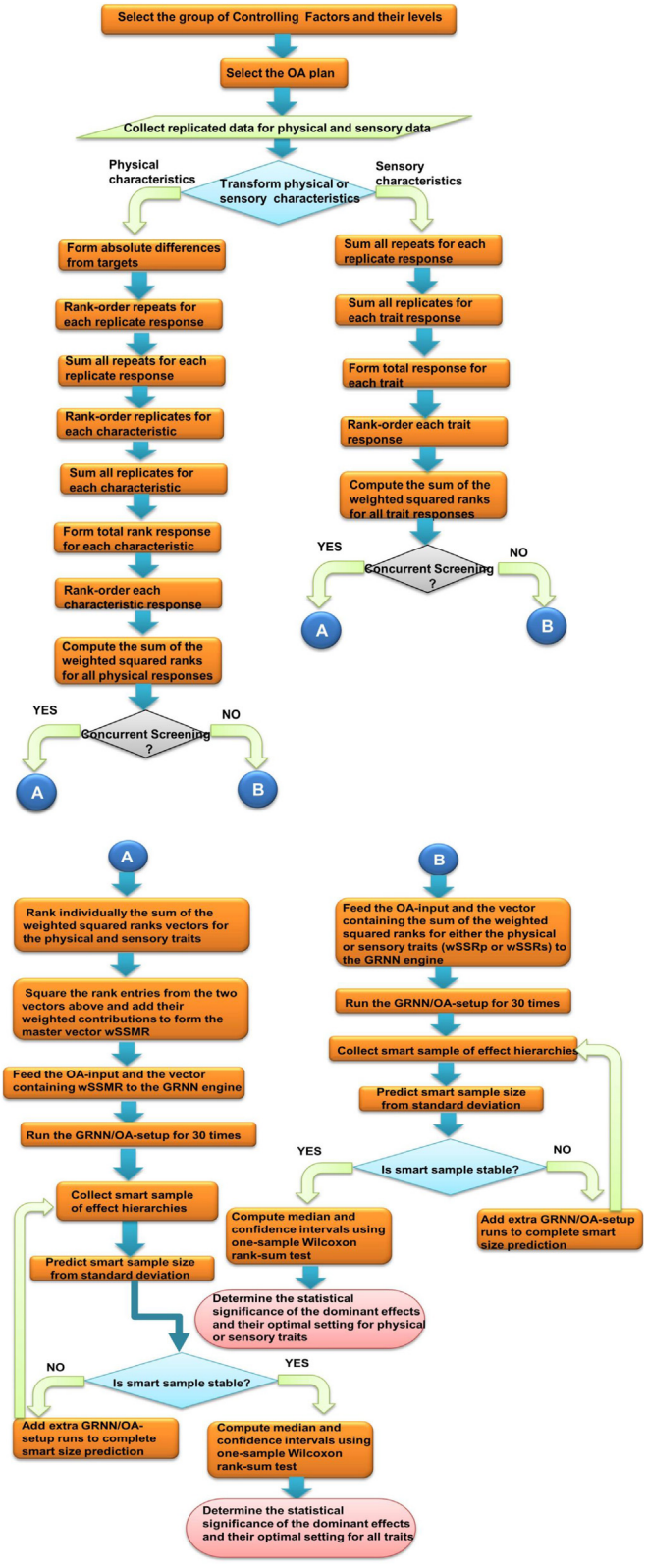


Fig. 8. Comprehensive outline that describes the data conversion strategy.

https://doi.org/10.1016/j.heliyon.2018.e00551

Ranks in this situation are always proper because they provide strong protection to super reduced-size structured data similar to our paradigm where multi-micro datasets are generated using Taguchi-type OA. For example, the experimental recipe requirement which directly influences the total trial volume has been reduced by 9 times in the illustrated $L(3^4)$ OA. Ranks are associated to promoting robust median statistics which carry the highly-desired property of maximum achievable breakdown point at a rate of 50% (Wilcox, 2010). Resolving the ever-present data messiness that lurks in non-linear systems and feeds on the imposing small data constraint supersedes any elusive efficiency concerns associated with normality.

After rank ordering the various response replicates, the scale across characteristics now becomes uniform, easing the strain on the NN. To quantify effects across different multiple groups, usually *squaring ranks* becomes a reasonable operation according to Kruskal-Wallis theory of multi-level effects (Wilcox, 2010). Therefore, blending naturally consolidates the data fusion process because it reduces: 1) the overall time of execution and 2) the stress from the NN to work with many different small data groups associated with the various characteristics. This decongests the NN as it would otherwise strive to fit little data to many different surfaces and stupefy more the NN output process. Not a pleasant situation for NN regression. On the other hand, data compression of trial replicates is central to classical Taguchi methods (data means or signal-to-noise ratios). So, replicates will always need to be reduced to an unreplicated form.

4. Results

4.1. Screening physical characteristics

4.1.1. Data prescreening

The collected raw data of the physical characteristics for the white pan bread experiments are tabulated in Table 1. A prescreening phase to assess the repeatability status of the four replicated datasets is useful before proceeding with the data reduction steps. The minimum requirement was to record two measurements (repetitions) per replicate. In Figs. 9, 10, 11, 12, we provide a view of the repeatability tendencies after performing a linear regression for each of the two characteristic replicates. The index number − 1 or 2 − next to the response symbol denotes the replication round.

Likewise, the capital letter − A or B − denotes the repetition sequence. The BW fittings are judged as satisfactory; both slopes are higher than 90% and their associated coefficients of determination also match at values over 90%. The fittings of W, M and H are less consistent. Their notably departing trends indicate an insidious destabilization when compared to the BW fittings. In particular, the fittings of M and H reveal declining slope magnitudes well below 50%. Ostensibly, such output instability will

https://doi.org/10.1016/j.heliyon.2018.e00551

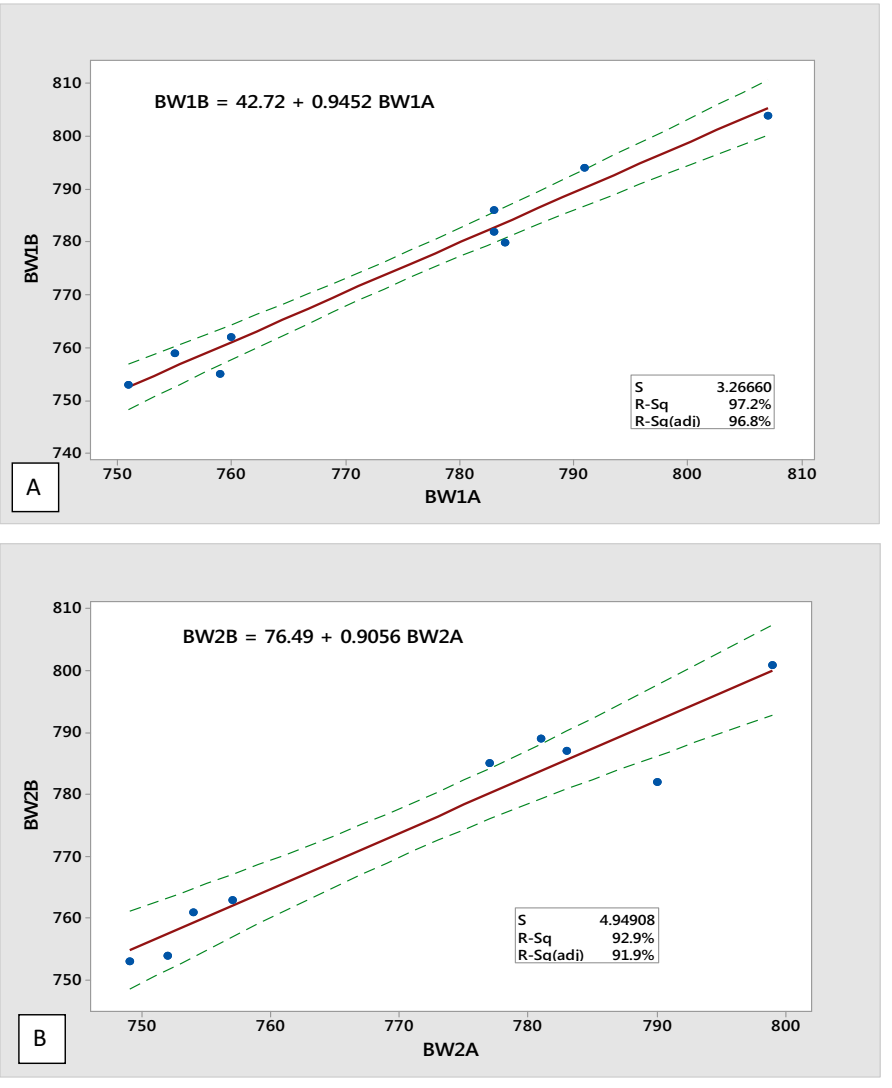


Fig. 9. Regression of repeatability trials of bread (95% CIs) for weight (BW) – replicate 1 (A) and 2 (B).

impose substantial blurring when attempting to filter out any weak effects. Furthermore, there is a persistent presence of outlier points in all eight graphs justifying the necessity for implementing a resilient data analyzer. However, this propensity is in accord with the complex physics that harness the modeling of bread properties. We conclude that messiness arguably rules the basic data structure. Hence, product characterization would be more fruitful if it fosters a distribution-free framework.

In Table 4, we list the transformed responses of the physical characteristics from Table 1. Results have been obtained after a sequential data manipulation for each of the four physical responses. The tabulation commences with entering the differenced response magnitudes from their respective target values (Section 3.1). Intermediate computed quantities are also exhibited to facilitate transparency of the

https://doi.org/10.1016/j.heliyon.2018.e00551

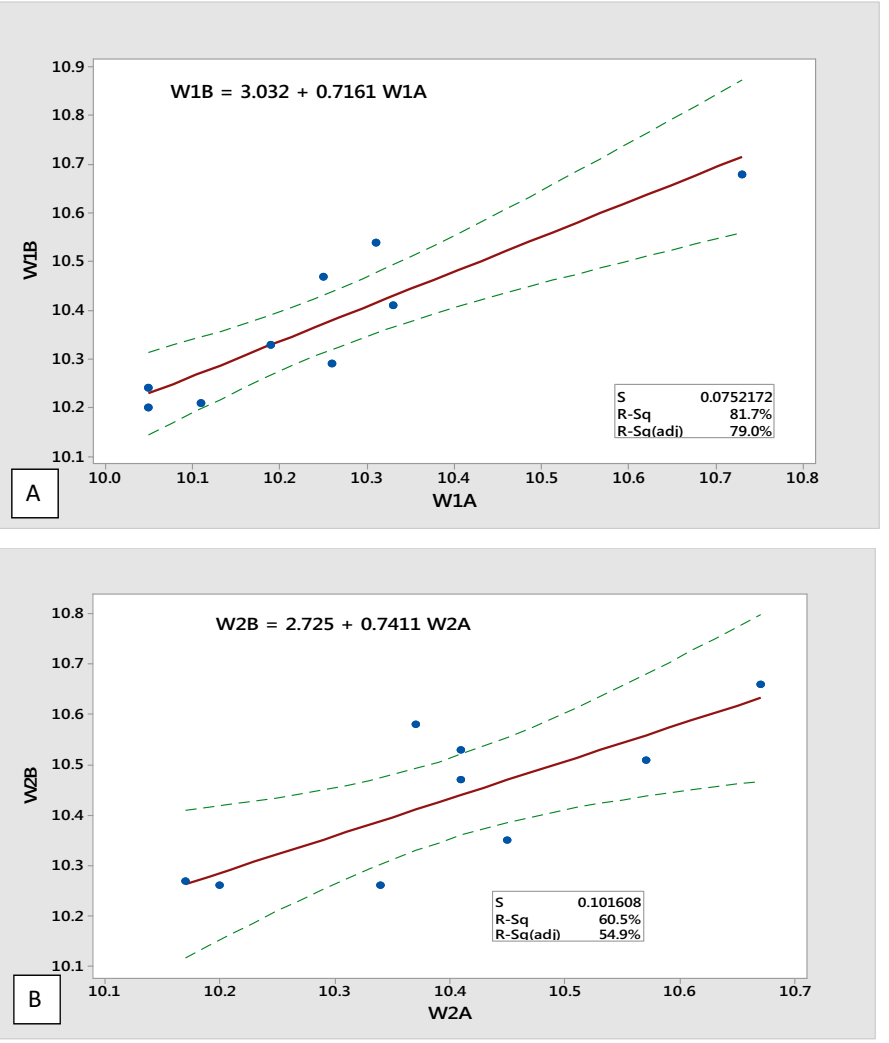


Fig. 10. Regression of repeatability trials of bread (95% CIs) for width(W) − replicate 1 (A) and 2 (B).

conversion process. Sequential rank ordering and compounding in two layers − first for repeats and then for replicates − lead to the condensed and homogenized vectors: BW′, M′, W′ and H′. Finally, the computed (two-way) cross-correlations (Spearman's ρ test) of BW′, M′, W′ and H′ are listed in Table 5. The outcomes suggest no direct relationships between any of the physical responses at a significance level of 0.05. Hence, there is no obvious reason for excluding any of the four examined responses from the concurrent screening procedure.

4.1.2. Concurrent screening of the four physical responses

Ranking individually the four homogenized responses according to prescription of Section 3 and subsequently fusing them together using the weight distribution as stated in Section 2, we contrive the single condensed response, wSSRp (Table 6). We feed the

https://doi.org/10.1016/j.heliyon.2018.e00551

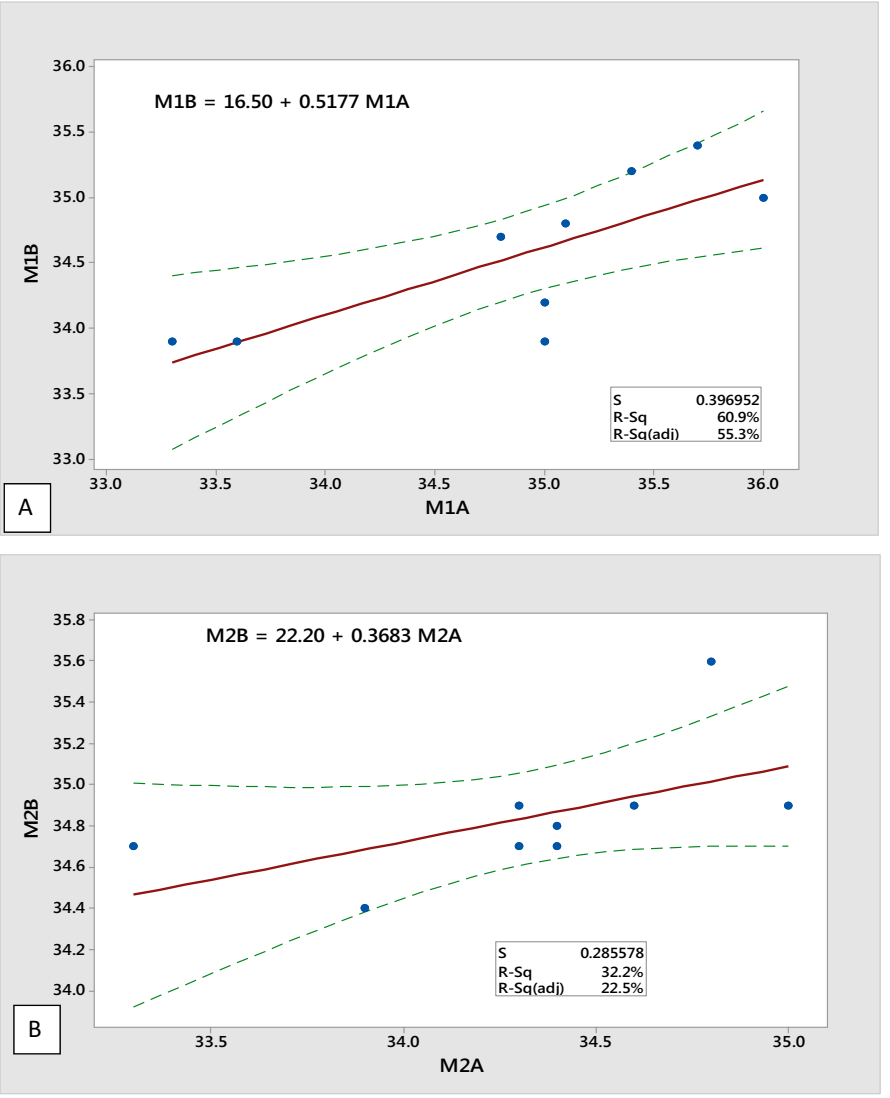


Fig. 11. Regression of repeatability trials of bread (95% CIs) for moisture(M) – replicate 1 (A) and 2 (B).

entries from the four non-linear inputs, in the $L_9(3^4)$ OA format (Table 3), along with the corresponding output values of the condensed response wSSRp into the GRNN processor according to the layout of Fig. 1 to form the proper smart sample. We gather the start-up effect-hierarchy scores which are generated through the sensitivity-analysis reports after 30 consecutive rounds of GRNN-module runs (Table 7). Checking for adequacy, the largest estimated standard deviation value was attributed to BT at 0.828 min. Thus, a prediction at a margin of error of 0.5 (confidence interval at 95%) returns a minimum sample base of fourteen runs. The estimated test power for the thirty-run sample was at least 90% which strengthens the credibility of the sampling process. Hence, we conclude that the initial intelligent sampling effort (subcrowding) is rendered sufficient and capable to predict any weak influences. Nevertheless, we observe that out of the possible twenty-four permutation outcomes (hierarchy sequences), only nine made

https://doi.org/10.1016/j.heliyon.2018.e00551

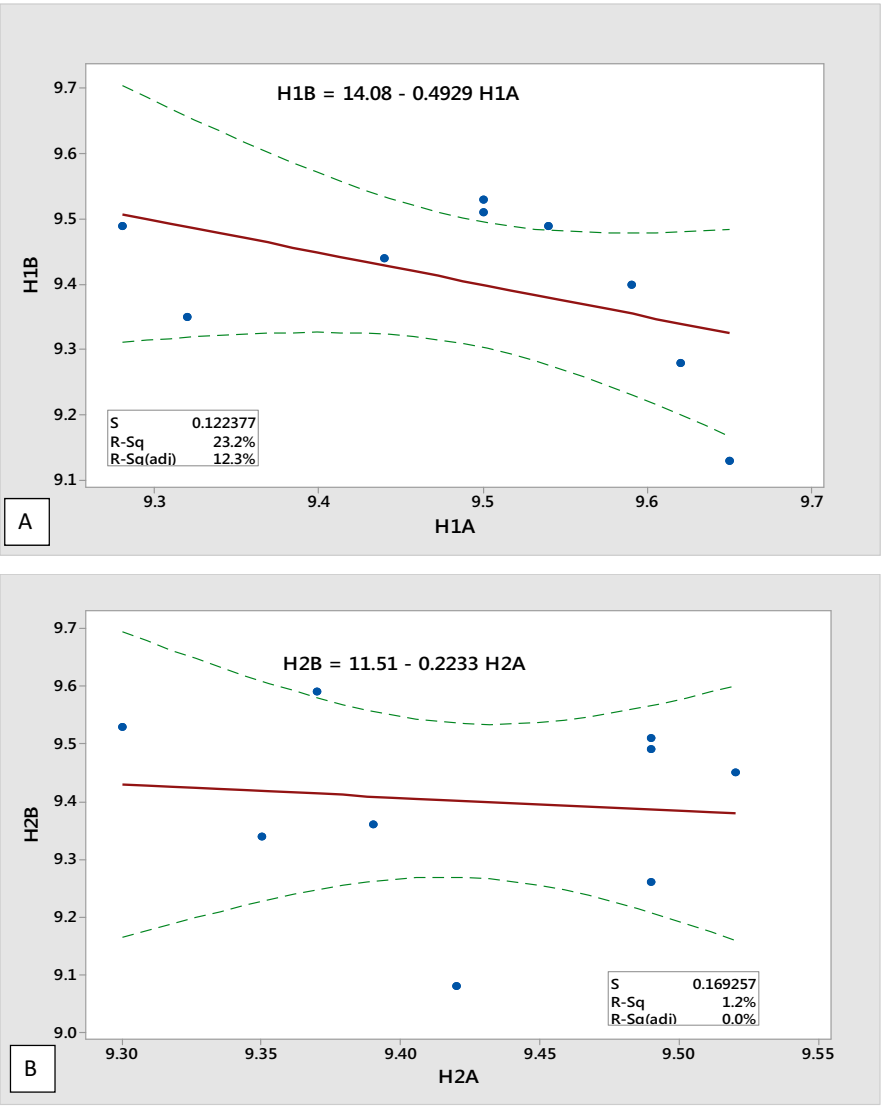


Fig. 12. Regression of repeatability trials of bread (95% CIs) for height (H) – replicate 1 (A) and 2 (B).

an appearance at least once. Leading the hierarchy is the BTP effect in the 26 out of the 30 intelligent runs. Next in order is the DW effect with 18 occurrences to claim the second place. The χ^2-test statistic value of 130 was highly significant ($p < 0.001$) when checking equality of proportions across the twenty-four possible permutations. The subcrowding median estimations and the 95%-confidence intervals from Wilcoxon's one-sample (rank-sum) test are listed in Table 8. Clearly, BTP is the predominant effect which registers no variation at all at its predictive interval since location and dispersion values all collapse to the top (rank) performance of '1'. Next, DW retains securely the sole second positioning. The estimated median value of '2' coincides with the second available position in the hierarchy which is also well-confined within the margin of error with respect to the upper boundary (2.5). Effects BT and PT are rather statistically

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 4. Physical characteristics OA-dataset transformations.

A. Bread weight(BW) transformations

Run #	dBW 1A	dBW 1B	dBW 2A	dBW 2B	rdBW 1A	rdBW 1B	rdBW 2A	rdBW 2B	srBW 1	srBW 2	rsrBW 1	rsrBW 2	BW′
1	10	6	8	2	5.5	3.5	3	1	9	4	4.5	2	6.5
2	6	10	13	11	3.5	5.5	8	4.5	9	12.5	4.5	6	10.5
3	14	12	16	12	8	7	9.5	6.5	15	16	7	8	15
4	19	15	16	24	13	9	9.5	15	22	24.5	11	12	23
5	5	3	11	4	2	1	4.5	2	3	6.5	1	3	4
6	18	17	12	20	11.5	10	6.5	13	21.5	19.5	10	9	19
7	26	29	25	17	15	16	16	11	31	27	16	15	31
8	42	39	34	36	18	17	17	18	35	35	17.5	17.5	35
9	18	21	18	22	11.5	14	12	14	25.5	26	13	14	27

B. Moisture(M) transformations

Run #	dM 1A	dM 1B	dM 2A	dM 2B	rdM 1A	rdM 1B	rdM 2A	rdM 2B	srM 1	srM 2	rsrM 1	rsrM 2	M′
1	0.3	0.2	0.1	0.4	3	1	2.5	13	4	15.5	1.0	7.5	8.5
2	0.5	0.3	0.1	0.3	6	3	2.5	10.5	9	13	2.0	5.0	7.0
3	0.5	0.6	0.6	0.1	6	9.5	16	2.5	15.5	18.5	7.5	9.0	16.5
4	0.9	0.7	0.5	0.4	14	12	15	13	26	28	14.5	17.0	31.5

(*continued on next page*)

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 4. (*Continued*)

B. Moisture(M) transformations

Run #	dM 1A	dM 1B	dM 2A	dM 2B	rdM 1A	rdM 1B	rdM 2A	rdM 2B	srM 1	srM 2	rsrM 1	rsrM 2	M′
5	0.9	0.6	1.2	0.2	14	9.5	18	7	23.5	25	11.0	13.0	24.0
6	1.2	0.9	0.2	0.4	16.5	14	7	13	30.5	20	18.0	10.0	28.0
7	0.6	0.3	0.2	0.2	9.5	3	7	7	12.5	14	4.0	6.0	10.0
8	1.5	0.5	0.3	1.1	18	6	10.5	17	24	27.5	12.0	16.0	28.0
9	1.2	0.6	0.1	0.2	16.5	9.5	2.5	7	26	9.5	14.5	3.0	17.5

C. Bread (loaf) height(H) transformations

Run #	dH 1A	dH 1B	dH 2A	dH 2B	rdH 1A	rdH 1B	rdH 2A	rdH 2B	srdH 1	srdH 2	rsrdH 1	rsrdH 2	H′
1	0.05	0.57	0.4	0.17	1	18	16	2	19	18	11.0	9.0	20.0
2	0.38	0.35	0.21	0.44	15	14	6.5	17	29	23.5	17.5	13.0	30.5
3	0.08	0.42	0.28	0.62	2	16.5	10	18	18.5	28	10.0	16.0	26.0
4	0.16	0.21	0.18	0.25	4	9.5	3	9	13.5	12	6.5	2.0	8.5
5	0.2	0.17	0.33	0.11	7.5	5	12	1	12.5	13	3.0	4.5	7.5
6	0.26	0.26	0.31	0.34	11.5	11.5	11	13	23	24	12.0	14.0	26.0
7	0.42	0.21	0.35	0.36	16.5	9.5	14	15	26	29	15.0	17.5	32.5
8	0.11	0.3	0.21	0.19	3	13	6.5	4	16	10.5	8.0	1.0	9.0
9	0.2	0.19	0.21	0.21	7.5	6	6.5	6.5	13.5	13	6.5	4.5	11.0

https://doi.org/10.1016/j.heliyon.2018.e00551

D. Bread (loaf) width(W) transformations

Run #	dW 1A	dW 1B	dW 2A	dW 2B	rdW 1A	rdW 1B	rdW 2A	rdW 2B	srdW 1	srdW 2	rsrdW 1	rsrdW 2	W′
1	0.31	0.17	0.13	0.08	15	4.5	9	6	19.5	15	11.0	8.5	19.5
2	0.23	0.18	0.17	0.16	9	6	13	11.5	15	24.5	8.5	12.0	20.5
3	0.45	0.26	0.09	0.03	17.5	12	7.5	2.5	29.5	10	15.0	4.5	19.5
4	0.39	0.29	0.3	0.24	16	13	17	15.5	29	32.5	14.0	18.0	32.0
5	0.19	0.04	0.05	0.15	7	2	4	10	9	14	3.0	7.0	10.0
6	0.17	0.09	0.09	0.03	4.5	3	7.5	2.5	7.5	10	2.0	4.5	6.5
7	0.24	0.21	0.16	0.24	10	8	11.5	15.5	18	27	10.0	13.0	23.0
8	0.45	0.3	0.33	0.23	17.5	14	18	14	31.5	32	16.0	17.0	33.0
9	0.25	0.03	0.07	0.01	11	1	5	1	12	6	6.0	1.0	7.0

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 5. Spearman's ρ cross-correlations for the four physical responses and their associated p-values (in parenthesis).

	M′	H′	W′
H′	−0.651		
	(0.057)		
W′	0.067	−0.029	
	(0.864)	(0.940)	
BW′	0.368	0.126	0.452
	(0.330)	(0.748)	(0.222)

Table 6. The weighted rank condensed vector wSSRp for intelligent processing.

Run #	rM′	rH′	rW′	rBW′	wSSRp
1	2	5	4.5	2	9.83
2	1	8	6	3	20.30
3	4	6.5	4.5	4	21.68
4	9	2	8	6	45.90
5	6	1	3	1	12.30
6	7.5	6.5	1	5	35.43
7	3	9	7	8	49.40
8	7.5	3	9	9	59.18
9	5	4	2	7	30.70

indistinguishable since they share the likelihood to coexist in the third position obscuring each other's potency. Summarizing, it is BTP and DW that outrightly control synchronously the four examined physical responses.

4.2. Screening sensory traits

4.2.1. Data prescreening

The collected raw data for the five sensory traits are listed in Table 2. In Table 9, we computed the cross-correlations for the two layers of datasets for all five sensory responses, i.e. for: 1) repeats (within each replicate), and 2) replicates (after pooling the repeats). It appears that FL and SF exhibit strong repeatability while the remaining three traits appear less distinguishing. Additional sampling would perhaps ameliorate the need for higher resolution. Since the measurement scale is by definition uniform for all five sensory traits, it is meaningful to aggregate the scores on the replicate level. In Table 10, we pool the repeats for each response replicate separately. With the exception of EL′, the rest of the four responses seem to show more convincing reproducibility (Table 9). CR' would be benefited with supplying extra data to refine further its outcome on a significance level of 0.05. The overall

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 7. Smart (subcrowd) samples for the three screening scenarios.

GRNN Run #	Weighted Physical Characteristics				Weighted Sensory Traits				Weighted Combination of both groups			
	DW	PT	BT	BTP	DW	PT	BT	BTP	DW	PT	BT	BTP
1	2	3	4	1	4	2	3	1	2	3	4	1
2	2	4	3	1	4	3	2	1	4	2	3	1
3	2	4	3	1	2	3	4	1	3	4	2	1
4	3	4	2	1	4	3	2	1	2	3	4	1
5	3	4	2	1	4	2	3	1	4	3	2	1
6	2	4	3	1	3	2	4	1	3	4	2	1
7	3	4	2	1	4	3	2	1	4	3	2	1
8	2	3	4	1	3	4	2	1	2	4	3	1
9	1	3	4	2	4	3	2	1	4	3	2	1
10	2	3	4	1	3	2	4	1	3	4	2	1
11	2	1	3	4	3	1	4	2	4	2	3	1
12	2	4	3	1	3	2	4	1	2	3	4	1
13	2	3	4	1	3	4	2	1	3	2	1	4
14	2	3	4	1	3	4	2	1	2	3	4	1
15	2	4	3	1	3	4	2	1	3	4	2	1
16	4	3	2	1	3	2	4	1	2	3	4	1
17	3	4	2	1	4	2	3	1	3	4	2	1
18	3	4	2	1	3	4	2	1	3	4	2	1
19	1	3	2	4	3	4	2	1	2	4	3	1
20	2	3	4	1	3	4	2	1	2	3	4	1
21	3	2	4	1	3	2	4	1	3	2	4	1
22	3	4	2	1	3	2	4	1	2	3	4	1
23	1	4	3	2	4	3	2	1	3	4	2	1
24	2	3	4	1	4	3	2	1	3	4	2	1
25	2	3	4	1	4	3	2	1	3	1	4	2
26	2	3	4	1	3	2	4	1	4	3	2	1
27	2	4	3	1	2	4	3	1	2	3	4	1
28	2	4	3	1	4	3	2	1	3	4	2	1
29	3	4	2	1	3	2	4	1	3	4	2	1
30	2	4	3	1	4	3	2	1	2	4	3	1

(mixed) sensory performance may not be considered unusual because the organoleptic process itself carries a certain degree of subjectiveness. A second round of data aggregation that involves the pooling of the sensory replicates generates the single-vector responses that are listed in Table 11. Before proceeding with the concurrent screening procedure, the five sensory responses are tested for cross-

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 8. Effect screening results from robust subcrowding sample analysis.

			GRNN-Results		
			Confidence Interval (95%)		
	Factors	**N**	**Median**	**Lower**	**Upper**
Physical	DW	30	2.0	2.0	2.5
Characteristics	PT	30	3.5	3.5	3.5
	BT	30	3.0	3.0	3.5
	BTP	30	1.0	1.0	1.0
Sensory Traits	DW	30	3.5	3.0	3.5
	PT	30	3.0	2.5	3.0
	BT	30	3.0	2.5	3.0
	BTP	30	1.0	1.0	1.0
Physical	DW	30	3.0	2.5	3.0
Characteristics	PT	30	3.5	3.0	3.5
And	BT	30	3.0	2.5	3.0
Sensory Traits	BTP	30	1.0	1.0	1.0

correlation. In Figs. 13, 14, 15, 16, 17, we provide linear regression plots for all ten pairs. We infer that there is no strong evidence for excluding any of the five sensory traits due to correlation from the concurrent screening procedure. We notice that the adjusted coefficient of determination never exceeds 80% (EL″ vs CL″ graph), while it may dip as low as 16% (EL′ vs CR″ graph). We remark that such disparate manifestations may not be foreign to complex materials like breads. In turn, this may encourage that a more relaxed screening approach might be advisable; one that relies less on parametric distributions in modeling the effects.

Table 9. Spearman's ρ cross-correlations for the five sensory responses and their associated p-values for repeats and replicates (pooled repeats).

	Replicate #	**Repetitions Spearman's *P***	**p-value**	**Replicates Spearman's ρ**	**p-value**
CL	1	0.418	0.263	0.758	0.018
	2	0.551	0.124		
FL	1	0.782	0.013	0.828	0.006
	2	0.830	0.006		
CR	1	0.568	0.111	0.619	0.076
	2	0.234	0.545		
SF	1	0.745	0.021	0.962	0.001>
	2	0.826	0.006		
EL	1	0.295	0.441	0.300	0.432
	2	0.503	0.168		

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 10. Pooling of the repeats for the five sensory traits – combined data in replicates.

Run #	CL′1	CL′2	FL′1	FL′2	CR′1	CR′2	SF′1	SF′2	EL′1	EL′2
1	52	61	60	66	70	76	64	61	61	62
2	64	66	71	71	81	78	76	79	67	72
3	63	62	64	56	76	74	66	62	70	67
4	65	65	72	71	81	75	73	73	73	66
5	51	51	63	58	74	65	58	57	53	56
6	64	66	73	72	69	74	73	74	70	68
7	67	65	70	70	76	77	69	72	71	72
8	65	67	79	73	80	80	79	75	71	67
9	58	59	62	64	69	69	71	68	70	69

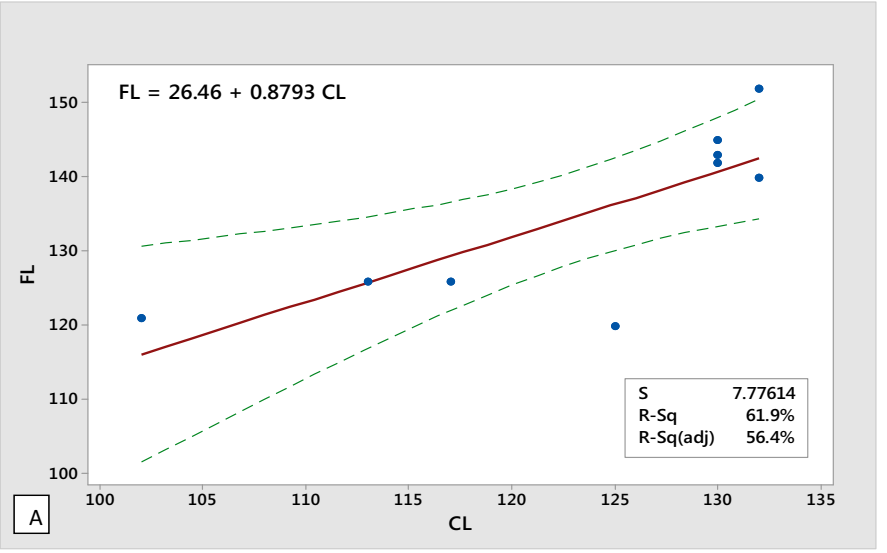
4.2.2. Concurrent screening of the five sensory responses

In Table 11, we progress from homogenizing the ranked sensory quantities, rCL to rEL, to their weighed compounding such that to form the condensed vector wSSRs. We repeat the same intelligent processing for the wSSRs vector using the same procedure as with the physical characteristics. The hierarchy outcomes for the effects are gathered from the subcrowding GRNN sensitivity analysis. The output after completing thirty consecutive GRNN module runs is tabulated in Table 7. Checking for subcrowd sampling adequacy, the largest estimated standard deviation value was attributed to BT at 0.925 min. This predicts a minimum subcrowd sampling effort of sixteen module runs at a margin of error of 0.5 (confidence interval at 95%). The estimated test power for the thirty-run smart sample was at least 82%. The start-up intelligent sampling is sufficient to filter-out dormant effects. Only seven out of the

Table 11. Pooled replicates and weighted consolidation of the ranked sensory responses.

Run #	CL′	FL′	CR′	SF′	EL′	rCL	rFL	rCR	rSF	rEL	wSSRs
1	113	126	146	125	123	8	6.5	6	8	8	56.85
2	130	142	159	155	139	4	4	2	1	3	10.05
3	125	120	150	128	137	6	9	5	7	7	49.75
4	130	143	156	146	139	4	3	3	4	3	12.15
5	102	121	139	115	109	9	8	8	9	9	75.90
6	130	145	143	147	138	4	2	7	3	5.5	19.06
7	132	140	153	141	143	1.5	5	4	5	1	12.41
8	132	152	160	154	138	1.5	1	1	2	5.5	9.23
9	117	126	138	139	139	7	6.5	9	6	3	38.25

https://doi.org/10.1016/j.heliyon.2018.e00551



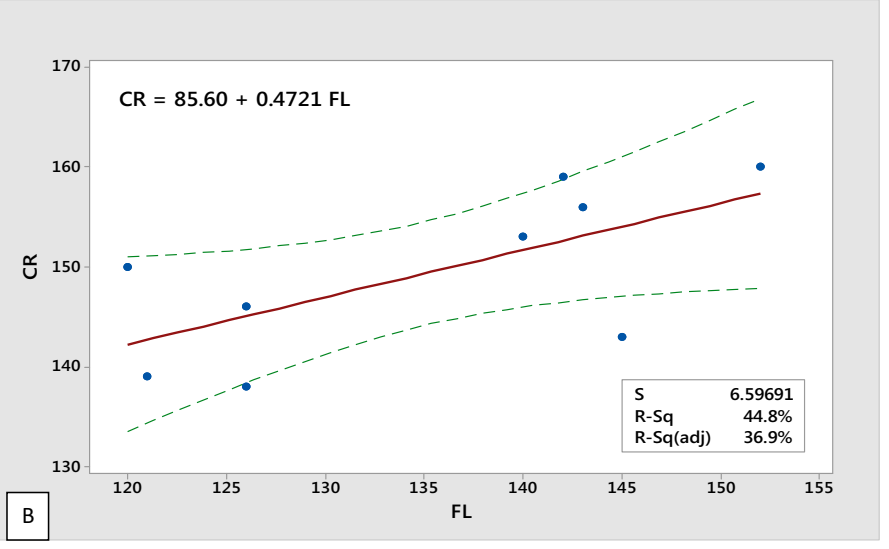


Fig. 13. Cross-regressions of pooled sensory-trait data: A) flavor vs color, B) crumb structure vs flavor.

twenty four possible permutations of the effect orderings made an appearance at least once in the GRNN smart sample. The BTP effect comes out superior in 29 out of the 30 intelligent runs. The DW effect follows next with 16 occurrences at the second place. The χ^2-test statistic value of 134 was highly significant ($p < 0.001$), thus rejecting the equality of the 24 proportions. The subcrowding median estimations and their assorting 95%-confidence intervals from Wilcoxon's one-sample (rank-sum) test are listed in Table 8. Again, the BTP effect retains its top placement as evidenced from a diminishing variability around its median estimation. The rest of the effects, DW, BT and PT, appear to share the third positioning.

https://doi.org/10.1016/j.heliyon.2018.e00551

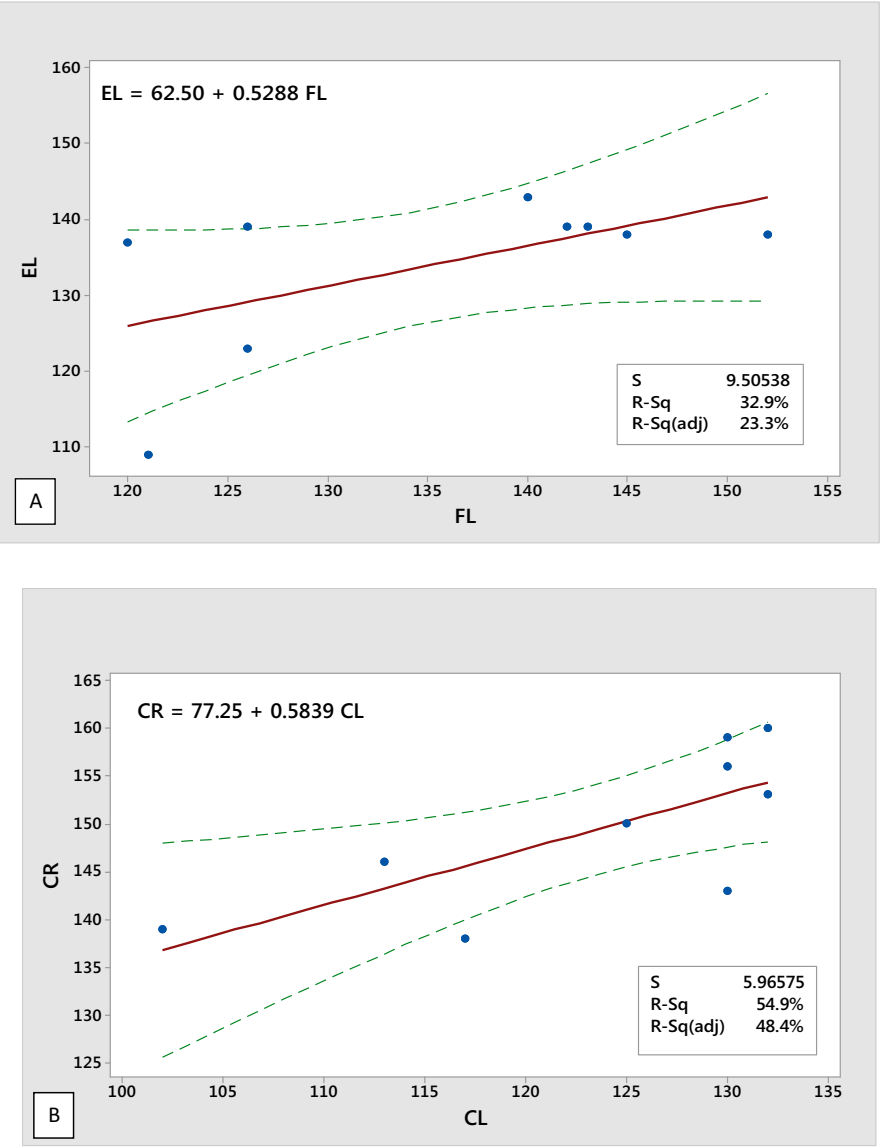


Fig. 14. Cross-regressions of pooled sensory-trait data: A) elasticity vs flavor, B) crumb structure vs color.

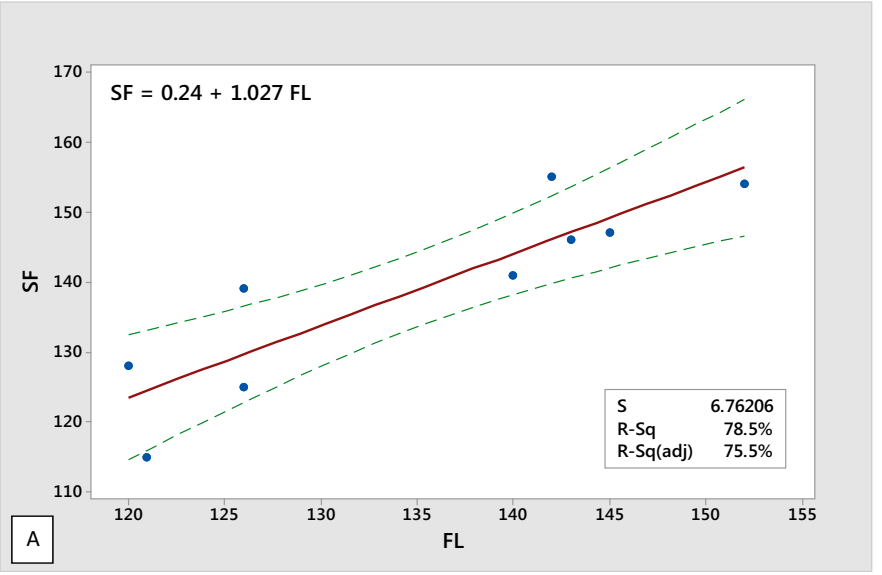
This occurrence confounds any chance for declaring the three effects as contributing. Summarizing, it is the sole influence of BTP that controls in a synchronous manner the five sensory responses.

4.3. Screening physical and sensory traits

4.3.1. Data prescreening

The weighted and condensed vectors, wSSRp and wSSRs, which have been computed in Tables 6 and 11, respectively, are inspected for possible correlation. In Fig. 18, we display their linear regression fitting. We discern no relationship

https://doi.org/10.1016/j.heliyon.2018.e00551



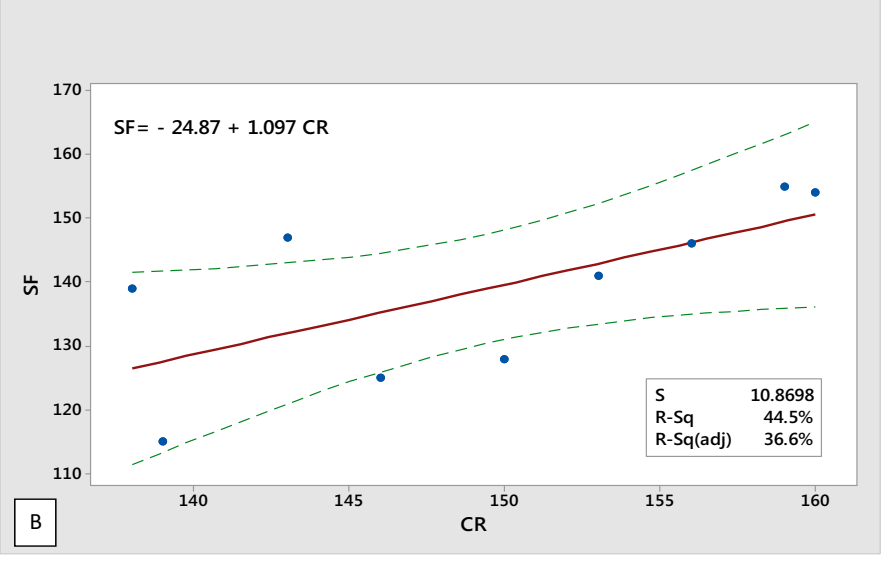


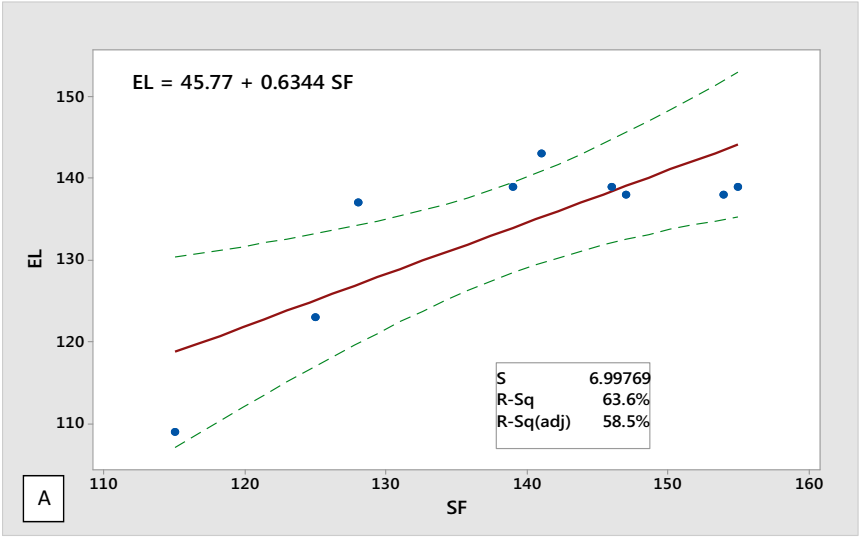
Fig. 15. Cross-regressions of pooled sensory-trait data: A) softness vs flavor, B) softness vs crumb structure.

between the two master responses as the fitted slope (0.54) and the adjusted coefficient of determination (54%) do not fare strongly. Moreover, glancing Fig. 18, the persistence of outlier points emerges as frequent enough to justify a robust data treatment.

4.3.2. Concurrent screening of physical and sensory responses

In Table 12, we create the master-rank vectors MRp and MRs by rank-ordering independently the two corresponding condensed vectors, wSSRp and wSSRs. The

https://doi.org/10.1016/j.heliyon.2018.e00551



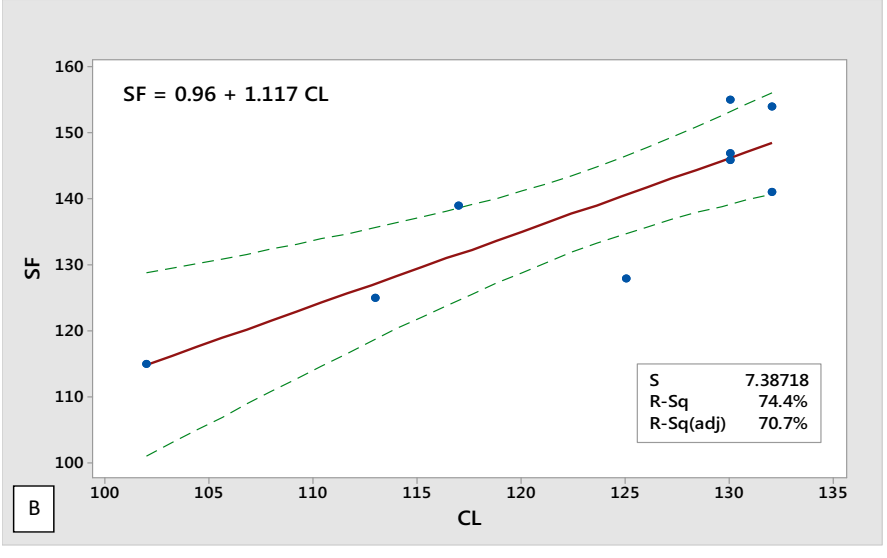
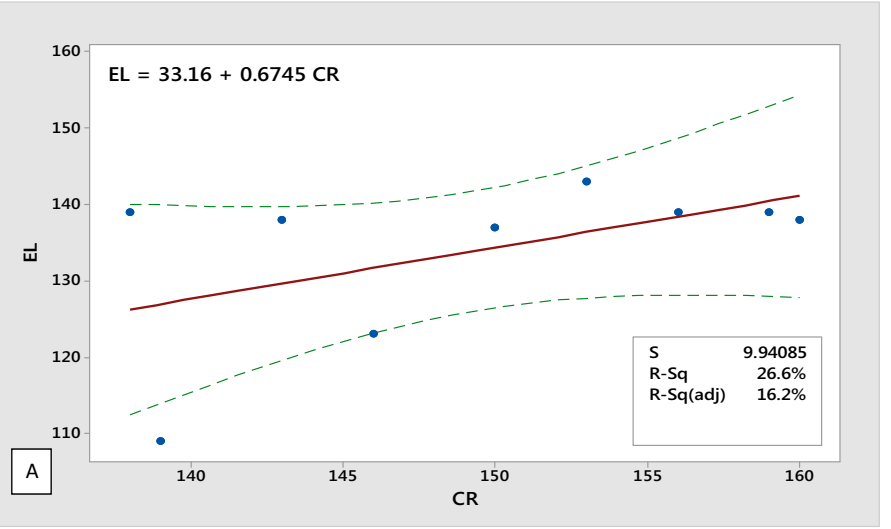


Fig. 16. Cross-regressions of pooled sensory-trait data: A) elasticity vs softness, and B) softness vs color.

weighted-squared sum of the two vectors returns the terminal wSSMR vector that fuses together internal information from all nine monitored responses. We repeat once more the intelligent processing on the wSSMR vector as in the previous two scenarios. The effect-hierarchy outcomes of crowdsourcing are tabulated in Table 7. Checking for subcrowd sampling adequacy, the largest estimated standard deviation value was attributed to BT at 0.961 min. This predicts a minimum smart sampling effort of seventeen GRNN module runs at a margin of error of 0.5 (confidence interval at 95%). The estimated test power for the thirty-run subcrowd sample was at least 79%. Only seven out of the twenty four possible permutations of the effect orderings made an appearance at least once in the GRNN output. The BTP effect

https://doi.org/10.1016/j.heliyon.2018.e00551



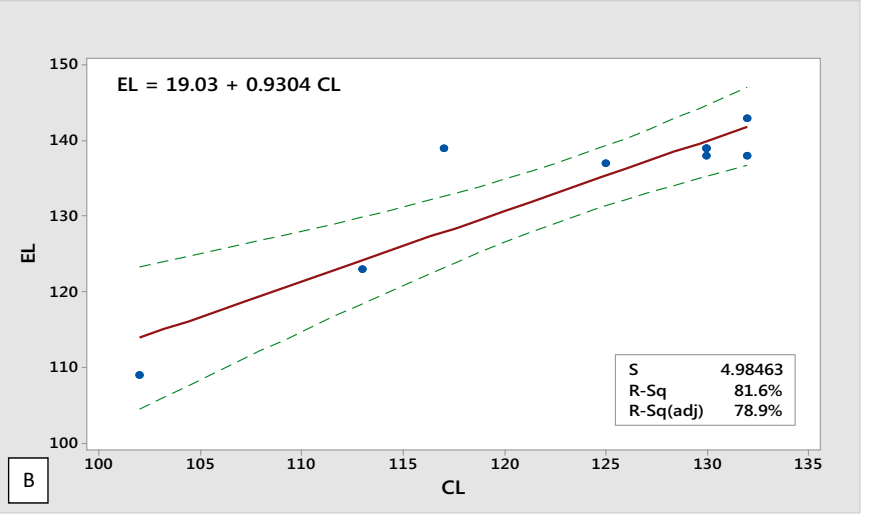


Fig. 17. Cross-regressions of pooled sensory-trait data: A) elasticity vs crumb structure, and B) elasticity vs color.

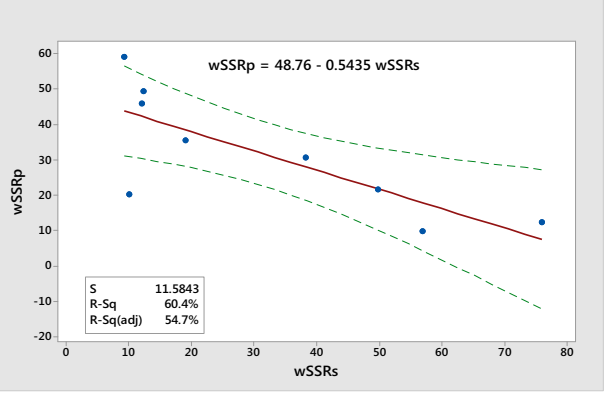


Fig. 18. Linear regression fit of wSSRp versus wSSRs (with 95% confidence interval).

https://doi.org/10.1016/j.heliyon.2018.e00551

Table 12. Weighted consolidation of the master physical and sensory responses.

Run #	MRp	MRs	wSSMR
1	1	8	45.1
2	3	2	5.5
3	4	7	39.1
4	7	3	21.0
5	2	9	57.9
6	6	5	28.3
7	8	4	30.4
8	9	1	25.0
9	5	6	32.7

comes up superior in 28 out of the 30 intelligent runs. The BT effect follows next with 14 occurrences at the second place. The χ^2-test statistic value of 126.8 was highly significant ($p < 0.001$), thus rejecting the equality of the 24 proportions. The subcrowding median estimations and the 95%-confidence intervals from Wilcoxon's one-sample (rank-sum) test are listed in Table 8. The BTP effect maintains its outstanding performance. The rest of the effects − DW, BT and PT − possess overlapping confidence intervals. Thus, their potency may not be contemplated any further. Summarizing, it is the sole influence BTP that synchronously adjusts all physical and sensory responses. The optimized (median) responses for BW, M, H, W, CL, FL, CR, SF, and EL are − with interquartile range in parenthesis: 782 (28.25) g, 34.8 (0.65) %, 9.36 (0.1) cm, 10.41 (0.34) cm, 6.6 (0.3), 7.2(0.2), 7.6(0.7), 7.3(0.5), and 7.0(0.5), respectively.

5. Discussion

A robust and intelligent screening scheme was implemented to profile white pan bread properties in three distinct scenarios that involved the probing of: 1) four physical characteristics, 2) five sensory traits and 3) the synchronous screening of the physical and sensory responses. The datatypes of the examined controlling factors are non-restrictive. It is worth mentioning that the categorical variable BTP was transmuted by piecing together − into a single effect − the pairs of (numerical) ranges that signified the relevant oven-zone baking temperatures. This was a simplification trick that aided in restraining the total number of tested parameters down to four. Otherwise, complementing the experimental design with two separate temperature controls would double up the total volume of the projected trials. This is because trial runs would have been programmed by the next larger OA planner, the 18-run sampler (L_{18} OA), which is primed to accommodate five or more nonlinear (three-level) controlling factors.

https://doi.org/10.1016/j.heliyon.2018.e00551

The study spotlighted several aspects that an engineer is bound to face in product design and development: 1) the data-driven decision-making approach for delineating the complex bread behavior on pragmatic mass-production conditions, 2) the urgency to make product performance forecasts relying on small samples due to economic and availability limitations, 3) the inevitability to handle small but dense datasets, 4) the inherent data messiness of complex materials, 5) the demand for smart and robust tools to model complex (food) properties on minimum assumptions and 6) the power of the Wisdom of NN Crowds to overcome the data conversion and mining of limited, messy and complex datasets. We note that the underlying data messiness is coped with early on during our data compression process. This stems from the fact that bread-making is dictated by a plethora of intricate activities constituting its supply chain. The mathematics that describe the complete product development cycle are not tractable. Consequently, the task of deterministically predicting bread properties in full through detailed modeling remains elusive. To this end, our strategy was conceived to diagnose any potent effects from a nominated list of controlling effects by rounding up all sources of error into a single "master"-uncertainty. It is advantageous that we bypass seeking thorough details of the mechanism that governs this master-uncertainty. We merely monitor and homogenize the master-uncertainty as a single entity such that to effectively control its perturbation on the stochastic landscape. The actual sources that blur the studied effects and contribute to the collective vagueness may include any remnant effects of unexplored parameters and other unknown and unknowable intrusions. Irrespective of their origin, all unresolved influences have been disorderly bunched together to form the master uncertainty in our formalism. A sizeable uncertainty when present interferes by overshadowing the profiling process, thus making any potent effects hard to discover. This is true when the comparative frame of reference expects the quantified uncertainty to be the measuring stick. Even strong effects are virtually downsized on comparison and they are rendered to appear deflated. So the strategy of direct contrasting of the estimated variances of the studied effects with a "lump sum" residual error may not be prudent. Ordinary multi-response multi-factorial treatments such as the multivariate analysis of variance (MANOVA) or the general linear model (GLM) are based on F-ratio testing. Thus, the preferred measure of the strength of the effects is set per a residual-error basis. It is this aspect that may be amenable to dubious diagnostics. Oppositely, the proposed method may have a promising feature to offer on that respect by countering uncertainty whereas averting its direct involvement when mediating the terminal effect hierarchy.

The full exploitation (saturation) of the small $L_9(3^4)$ OA planning scheme – for the economic and practical reasons discussed previously – along with the dramatic condensation of the dataset rendered impervious to checking the validity of basic assumptions such as: 1) normality, 2) heteroscedasticity and 3) sparsity. Such assumptions must hold in order for the standard treatments of MANOVA and GLM to be

https://doi.org/10.1016/j.heliyon.2018.e00551

fruitful. On the contrary, the new solver is not obstructed in extracting impartial information despite the fact that the validity of those assumptions may be absent.

Although the proposed methodology rests on a fully robust framework which employs an intelligent "meta-sampling" (crowdsourcing) approach, it is instructive to compare our results for agreement with other practical yet naïve techniques. In Figs. 19, 20, 21, we depict how the controlling factors fare by summarizing their corresponding spreads in box-plots for all three available scenarios. The inner drawn boxes indicate the 95% confidence interval for the estimated medians. It is discerned that the BTP is undisputedly the major source that regulates variation in the doubly-weighted nexus of physical and sensory responses (Fig. 21). This finding is in accord with the individual performance of the two weighted groups of characteristics. The scientific rationale for the physical characteristics is that increasing temperature influences all four responses. This is because BW and M decrease as water evaporates while H and W increase with the growth of the air bubbles as bread dough expands. Increasing temperature regulates bread texture and thus directly influences sensory-related responses. Elasticity and softness decreases since entanglement density of high-polymer gluten molecules increases with water evaporation. Darkening of the crust is anticipated with increased oven temperature. In the case of the weighted physical screening (Fig. 19), DW locks the second position in hierarchy as ascertained by its diminutive confidence interval estimation of its median. The box plots in Figs. 19, 20, 21 are also useful in illustrating the great diversity which is hidden in the meta-data distributions which have been generated by smart-sampling (sub-crowding). Moreover, they accentuate the great difficulty that the stochastic

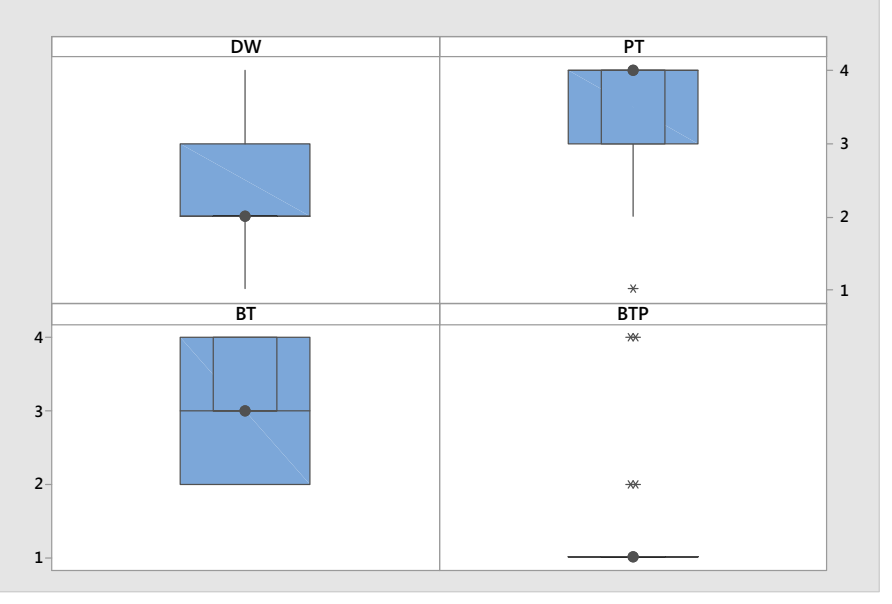


Fig. 19. Box-plot contrasting with 95%-confidence interval of median of the four effects for the weighted profiling of the physical characteristics.

https://doi.org/10.1016/j.heliyon.2018.e00551

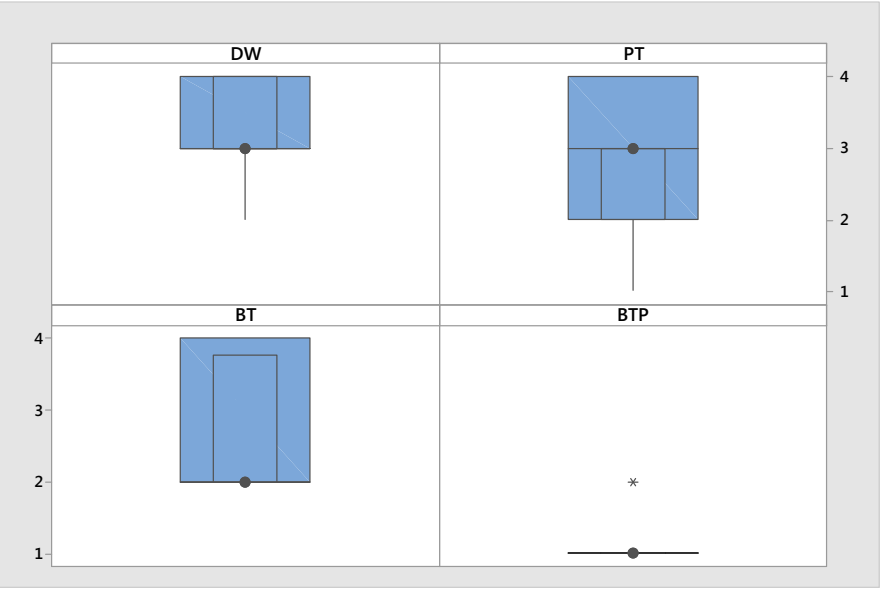


Fig. 20. Box-plot contrasting with 95%-confidence interval of median of the four effects for the weighted profiling of the sensory traits.

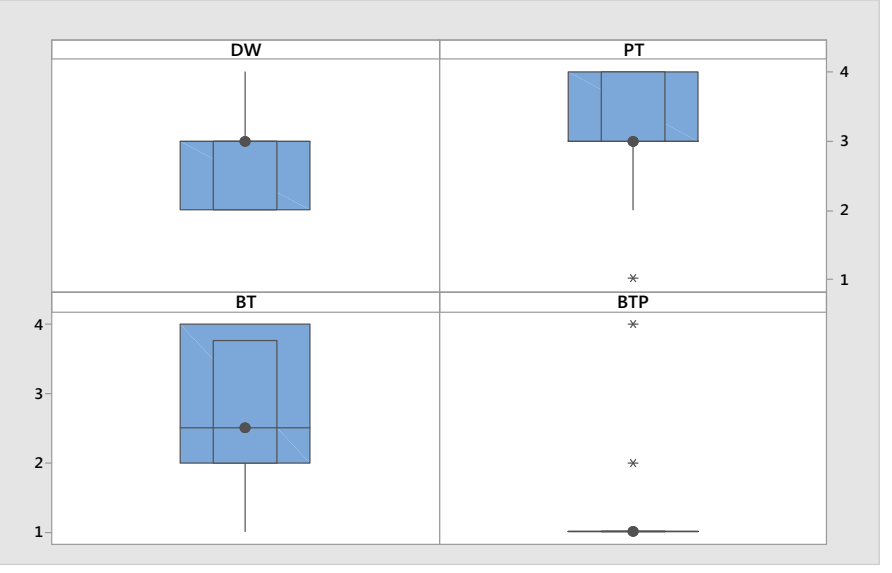


Fig. 21. Box-plot contrasting with 95%-confidence interval of median of the four effects for the weighted profiling of the combination of physical characteristics and sensory traits.

decoding process stumbles upon in uncovering and prioritizing the strength of the effects. The dicey asymmetry is prevalent around the median for all effects. The enigmatic disturbance rhythm looms even on the weak performing effects. Such data idiosyncrasies substantiate our strategy to resort to a robust and intelligent data processing.

https://doi.org/10.1016/j.heliyon.2018.e00551

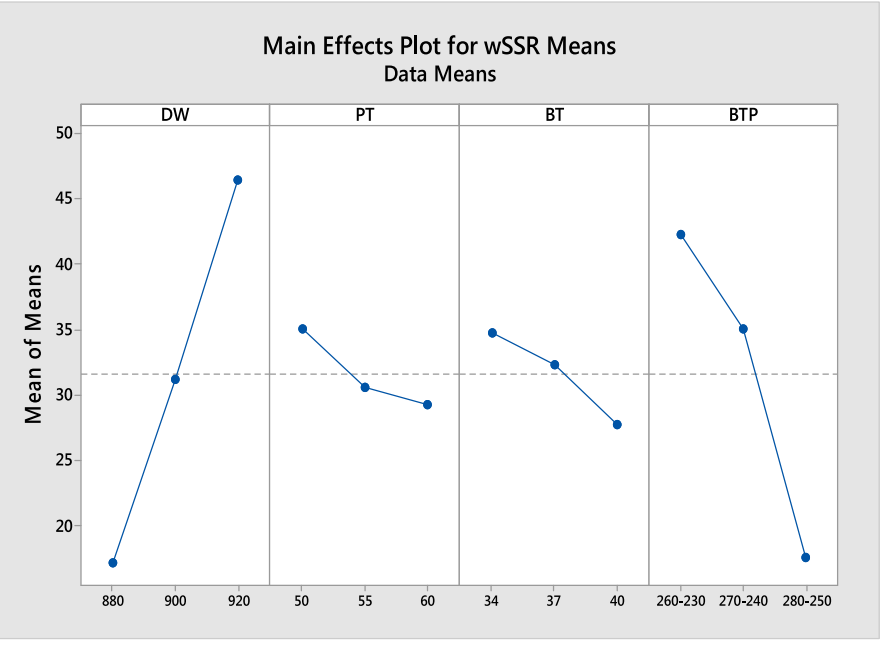


Fig. 22. Main effects plots for wSSRp.

It is also customary to comment on the insight which is gained by inspecting the classical main-effects plot, in spite of its known lack of statistical rigor. From Figs. 22, 23, 24, it visually stands out the dominance of BTP with the sharp non-linear modulations to be identified in at least the two of the three versions – including the

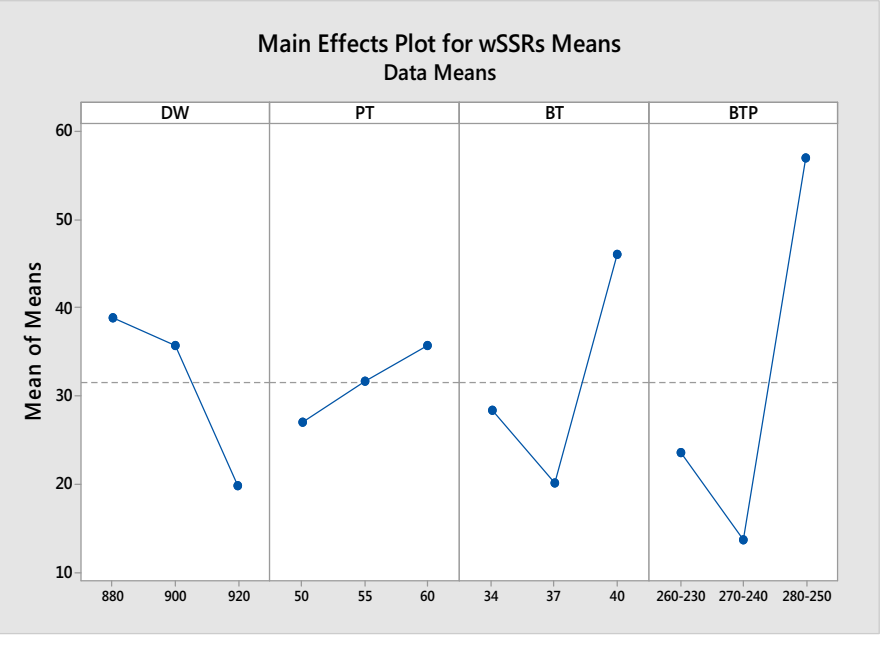


Fig. 23. Main effects plots for wSSRs.

https://doi.org/10.1016/j.heliyon.2018.e00551

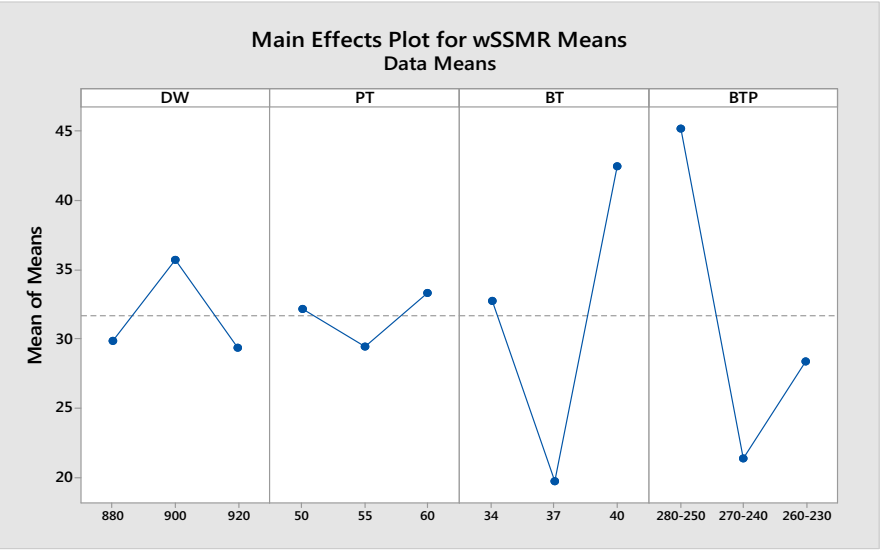


Fig. 24. Main effects plots for wSSMR.

master response. A favorable adjustment of the weighted physical responses coincides to the BTP setting of 280−250 °C (Fig. 22). However, the total weighted sensory performance (Fig. 23) as well as the outlook from the overall concurrent screening (Fig. 24) conjointly locate the optimal setting of BTP at the range of 270−240 °C. The appreciable altering of the quirky BTP tendency from a dipping monotonous profile in the physical screening to a convex-shaped curvature in the sensory screening is also noteworthy. Furthermore, DW seems to compete in importance with BTP in the physical screening scenario projecting an optimal adjustment at 880 g. It is the low-end material load that accommodates terminal specifications for the baked loaf critical dimensions and weight more closely. On the overall concurrent screening scenario, BT traces a convex curve with its optimal minimum setting to be pinpointed at 37 min.

6. Conclusions

A concurrent multi-response screening method was proposed to profile white pan bread properties. This was achieved by bringing together the GRNN intelligence and the concept of the Wisdom of Crowds. The main impetus of this proposal was to provide an agile medium for rapid and robust product design and development/improvement. The joint application of intelligent data homogenization with fast distribution-free smart microanalytics showed promise in dealing with the quirky morphologies of small, dense and diverse datasets, which plague the analysis of complex bread-making processes.

Four controlling factors were gauged for simultaneous potency and non-linearity. Three tested scenarios were elaborated to unravel accordingly product concerns

https://doi.org/10.1016/j.heliyon.2018.e00551

with regards to: quality, marketing and design. Dense datasets were collected through the non-linear $L_9(3^4)$ OA sampler to accommodate observations for as many as four physical characteristics and five sensory traits. Replication assessment demonstrated the omnipresent data messiness associated with bread-making. The profiled effects were strategically analyzed in two layers employing order statistics on subcrowding samples. Smart sampling was conducted by a GRNN engine on homogenized and fused responses. On the first layer the two groups of physical and sensory responses were separately analyzed to predict the presence of any statistically dominant effects. This was achieved by allotting importance weights to each of the responses in their respective groups. On the second layer, fused information for each of the two groups which was extracted from the first layer was weighted and aggregated once more. This action consolidated the total multi-response variation from all nine responses to a single master response. It was predicted that the combination of the oven zone temperatures to be highly influential in all three scenarios. The suggested optimal setting was 270−240 °C. Dough weight appeared to be instrumental in synchronously adjusting all four weighted physical characteristics at 880 g.

The simplicity and agility of treating in a comprehensive manner any combination of physical and categorical variables alike foreshadows the usefulness of our approach. Thus, it may be actually extended to many other combinations of effects and characteristics for other kinds of complex food goods. Complementary assortment of indicators with respect to issues of safety, quality, productivity and marketing offers great potential for future studies.

Declarations

Author contribution statement

George Besseris: Conceived and designed the analysis; Performed the experiments; Analyzed and interpreted the data; Contributed analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

https://doi.org/10.1016/j.heliyon.2018.e00551

References

Abrahamson, S., Ryder, P., Unterberg, B., 2013. Crowdstorm: the Future of Innovation, Ideas, and Problem Solving. John Wiley and Sons, Hoboken, N.J.

Athanasiadou, S., 2010. Screening of Factors Affecting Commercial White Pan Bread Quality Using the Taguchi Approach. MSc Thesis in Quality Management. The University of West of Scotland, Scotland, UK.

Atkinson, A., Donev, A., Tobias, R., 2007. Optimum Experimental Designs, With SAS. Oxford University Press.

Autio, K., Flander, L., Kinnunen, A., Heinonen, R., 2001. Bread quality relationship with rheological measurements of wheat flour dough. Cereal Chem. 78 (6), 654−657.

Besseris, G.J., 2012. Multi-response multi-factorial master ranking in non-linear replicated-saturated DOE for qualimetrics. Chemometr. Intell. Lab. Syst. 116 (1), 47−56.

Besseris, G.J., 2013a. Profiling effects in industrial data mining by non-parametric methods. Eur. J. Oper. Res. 220 (1), 147−161.

Besseris, G.J., 2013b. A distribution-free multi-factorial profiler for harvesting information from high-density screenings. PLoS One 8, e73275.

Besseris, G.J., 2014a. Multi-response non-parametric profiling using Taguchi's qualimetric engineering and neurocomputing methods: screening a foaming process in a solar collector assembly. Appl. Soft Comput. 22 (1), 222−237.

Besseris, G.J., 2014b. A fast-and-robust profiler for improving polymerase chain reaction diagnostics. PLoS One 9 (9), e108973.

Besseris, G.J., 2015. Profiling multiple static and transient puff-pastry characteristics with a robust-and-intelligent processor. J. Food Eng. 164, 40−54.

Breiman, L., 2001. Statistical modeling: the two cultures. Stat. Sci. 16, 199−231.

Cauvain, S.P., Young, L.S., 2006. Baked Products: Science, Technology and Practice. Wiley Online Library.

Chhanwal, N., Tank, A., Raghavarao, K.S.M.S., Anandharamakrishnan, C., 2012. Computational fluid Dynamics (CFD) modeling for bread baking process - a review. Food Bioprocess Technol. 5 (4), 1157−1172.

Cohn, D.A., 1996. Neural network exploration using optimal experimental design. Neural Networks 9, 1071−1083.

 https://doi.org/10.1016/j.heliyon.2018.e00551

Das Mohapatra, P.K., Maity, C., Rao, R.S., Pati, B.R., Mondal, K.C., 2009. Tannase production by Bacillus licheniformis KBR6: optimization of submerged culture conditions by Taguchi DOE methodology. Food Res. Int. 42, 430–435.

Decock, P., Cappelle, S., 2005. Bread technology and sourdough technology. Trends Food Sci. Technol. 16 (1-3), 113–120.

Della Valle, G., Chiron, H., Cicerelli, L., Kansou, K., Katina, K., Ndiaye, A., Whitworth, M., Poutanen, K., 2014. Basic knowledge models for the design of bread texture. Trends Food Sci. Technol. 36 (1), 5–14.

Dobraszczyk, B.J., Morgenstern, M.P., 2003. Rheology and the breadmaking process. J. Cereal Sci. 38, 229–245.

Dobraszczyk, B.J., 2004. The physics of baking: rheological and polymer molecular structure–function relationships in breadmaking. J. Non Newton. Fluid Mech. 124, 61–69.

Ertekin, S., Rudin, C., Hirch, H., 2014. Approximating the crowd. Data Min. Knowl. Discov. 28, 1189–1221.

Feyissa, A.H., Gernaey, K.V., Adler-Nissen, J., 2012. Uncertainty and sensitivity analysis: mathematical model of coupled heat and mass transfer for a contact baking process. J. Food Eng. 109, 281–290.

Gao, J., Wong, J.X., Lim, J.C.-S., Henry, J., Zhou, W., 2015. Influence of bread structure on human oral processing. J. Food Eng. 167, 147–155.

Hadiyanto, H., Esveld, D.C., Boom, R.M., van Straten, G., van Boxtel, A.J.B., 2008. Product quality driven design of bakery operations using dynamic optimization. J. Food Eng. 86, 399–413.

Heenan, S.P., Hamid, N., Dufour, J.-P., Harvey, W., Delahunty, C.M., 2009. Consumer freshness perceptions of breads, biscuits and cakes. Food Qual. Prefer. 20, 380–390.

Hering, P., Simandl, M., 2010. Sequential optimal experiment design for neural networks using multiple linearization. Neurocomputing 73, 3284–3290.

Howe, J., 2009. Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business. Three Rivers Press, New York.

Hubert, S., Helmers, T., Groß, F., Delgado, A., 2016. Data driven stochastic modelling and simulation of cooling demand within breweries. J. Food Eng. 176, 97–109.

Issanchou, S., Gauchi, J.-P., 2008. Computer-aided optimal designs for improving neural network generalization. Neural Networks 21, 945–950.

https://doi.org/10.1016/j.heliyon.2018.e00551

Jefferson, D.R., Lacey, A.A., Sadd, P.A., 2006. Understanding crust formation during baking. J. Food Eng. 75 (4), 515−521.

Lamrini, B., Della Valle, G., Trelea, I.C., Perrot, N., Trystrama, G., 2012. A new method for dynamic modelling of bread dough kneading based on artificial neural network. Food Control 26, 512−524.

Liu, Z., Scanlon, M.G., 2003. Predicting mechanical properties of bread crumb. Food Bioprod. Process. 81 (3), 224−238.

Malone, T.M., Bernstein, M.S., 2015. Handbook of Collective Intelligence. MIT Press, Cambridge, MA.

Martinez, M.M., Gomez, M., 2017. Rheological and microstructural evolution of the most common gluten-free flours and starches during bread fermentation and baking. J. Food Eng. 197, 78−86.

Milliken, G.A., Johnson, D.E., 1989. Analysis of Messy Data, Volume II: Nonreplicated Experiments. Chapman and Hall/CRC, Boca Raton, FL.

Milliken, G.A., Johnson, D.E., 2009. Analysis of Messy Data Volume I: Designed Experiments. Chapman and Hall/CRC, Boca Raton, FL.

Mondal, A., Datta, A.K., 2008. Bread baking - a review. J. Food Eng. 86 (4), 465−474.

Murphy, K.P., 2012. Machine Learning: a Probabilistic Perspective. MIT Press, Cambridge, MA.

Ndiaye, A., Valle, G.D., Roussel, P., 2009. Qualitative modelling of a multi-step process: the case of French breadmaking. Expert Syst. Appl. 36 (2), 1020−1038.

Oztop, M.H., Sahin, S., Sumnu, G., 2007. Optimization of microwave frying of potato slices by using Taguchi technique. J. Food Eng. 79, 83−91.

Ozyildirim, B.M., Avci, M., 2013. Generalized classifier neural network. Neural Networks 39, 18−26.

Ozyildirim, B.M., Avci, M., 2016. One pass learning for generalized classifier neural network. Neural Networks 73, 70−76.

Parimala, K.R., Sudha, M.L., 2015. Wheat-based traditional flat breads of India. Crit. Rev. Food Sci. Nutr. 55 (1), 67−81.

Pouliou, M.A., Besseris, G.J., 2013. Robust screening of cake product characteristics by the Taguchi method. Int. J. Ind. Syst. Eng. 14 (2), 207−229.

Purlis, E., Salvadori, V.O., 2007. Bread browning kinetics during baking. J. Food Eng. 80 (4), 1107−1115.

https://doi.org/10.1016/j.heliyon.2018.e00551

Purlis, E., 2011. Bread baking: technological considerations based on process modelling and simulation. J. Food Eng. 103, 92−102.

Rask, C., 1989. Thermal properties of dough and bakery products: a review of published data. J. Food Eng. 9 (3), 167−193.

Rousu, J., Flander, L., Suutarinen, M., Autio, K., Kontkanen, P., Rantanen, A., 2003. Novel computational tools in bakery process data analysis: a comparative study. J. Food Eng. 57, 45−56.

Rzigue, A., Monteau, J.-Y., Marmi, K., Le Bail, A., Chevallier, S., Reguerre, A.-L., Jury, V., 2016. Bread collapse: causes of the technological defect and impact of depanning time on bread quality. J. Food Eng. 182, 72−80.

Saguy, I.S., Singh, R.P., Johnson, T., Fryer, P.J., Sastry, S.K., 2013. Challenges facing food engineering. J. Food Eng. 119, 332−342.

Scanlon, M.G., Zghal, M.C., 2001. Bread properties and crumb structure. Food Res. Int. 34 (10), 841−864.

Sharif, K.M., Rahman, M.M., Azmir, J., Mohamed, A., Jahurul, M.H.A., Sahena, F., Zaidul, I.S.M., 2014. Experimental design of supercritical fluid extraction − a review. J. Food Eng. 124, 105−116.

Singh, R.P., Heldman, D.R., 2013. Introduction to Food Engineering, fifth ed. Academic Press, Waltham, MA.

Skrjanc, I., 2015. Evolving fuzzy-model-based design of experiments with supervised hierarchical clustering. IEEE Trans. Fuzzy Syst. 23, 861−871.

Sliwinski, E.L., Kolster, P., Van Vliet, T., 2004. On the relationship between large-deformation properties of wheat flour dough and baking quality. J Cereal Sci. 39 (2), 231−245.

Specht, D.F., 1990. Probabilistic neural networks. Neural Networks 3 (1), 109−118.

Specht, D.F., 1991. A general regression neural network. IEEE Trans. Neural Networks 2 (6), 568−576.

Specht, D.F., 1995. Probabilistic neural networks and general regression neural networks. In: Chen, C.H. (Ed.), Fuzzy Logic and Neural Network Handbook. McGraw-Hill, New York.

Stojceska, V., Butler, F., 2012. Investigation of reported correlation coefficients between rheological properties of the wheat bread doughs and baking performance of the corresponding wheat flours. Trends Food Sci. Technol. 24 (1), 13−18.

https://doi.org/10.1016/j.heliyon.2018.e00551

Surowiecki, J., 2005. The Wisdom of Crowds: Why the Many are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. Anchor, New York.

Taguchi, G., Chowdhury, S., Taguchi, S., 2000. Robust Engineering: Learn How to Boost Quality while Reducing Costs and Time to Market. McGraw-Hill, New York, NY.

Taguchi, G., Chowdhury, S., Wu, Y., 2004. Quality Engineering Handbook. Wiley-Interscience, Hoboken, NJ.

Tasirin, S.M., Kamarudin, S.K., Ghani, J.A., Lee, K.F., 2007. Optimization of drying parameters of bird's eye chilli in a fluidized bed dryer. J. Food Eng. 80, 695–700.

Thakur, M., Olafsson, S., Lee, J.-S., Hurburgh, C.R., 2010. Data mining for recognizing patterns in foodborne disease outbreaks. J. Food Eng. 97, 213–227.

Therdthai, N., Zhou, W., 2003. Recent advances in the studies of bread baking process and their impacts on the bread baking technology. Food Sci. Technol. Res. 9 (3), 219–226.

Tortum, A., Yayla, N., Celik, C., Gokdag, M., 2007. The investigation of model selection in artificial neural networks by the Taguchi method. Phys. A 386, 446–468.

Vanin, F.M., Lucas, T., Trystram, G., 2009. Crust formation and its role during bread baking. Trends Food Sci. Technol. 20 (8), 333–343.

Wilcox, R.R., 2010. Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy. Springer, London.

Zanoni, B., Peri, C., Pierucci, S., 1993. A study of the bread-baking process. I: a phenomenological model. J. Food Eng. 19 (4), 389–398.

Zanoni, B., Pierucci, S., Peri, C., 1994. Study of the bread baking process - II. Mathematical modelling. J. Food Eng. 23 (3), 321–336.

Zhang, J., Datta, A.K., 2006. Mathematical modeling of bread baking process. J. Food Eng. 75 (1), 78–89.

 https://doi.org/10.1016/j.heliyon.2018.e00551