PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NORTHPOLE

NorthPoleInc.

Optimization of Eggnog using Design of Experiment and Machine Learning

Santa Claes[1,2]*, Nisse Ness E.[2,3], Jeremy Elf Tootoo[1], Dasher Sridhar[1,3], Tinsel Tian[3] and Laura Menorah[1,2,3]

[1]*University of NorthPole, Dept. Bioinformatics and Engineering, Arctic Avenue 1, Santa's Secret Village*

[2]*NorthPole Institute of Food Science and Technology, Candy Cane Lane 25, Santa's Secret Village*

[3]*University of NorthPole, Dept. Chemistry, Poinsettia Circle 12, Santa's Secret Village*

**Corresponding author: santa@dna20.com*

Received 28 June 2015, received in revised form 25 October 2015, accepted 24 December 2015

Keywords: Design of Experiment, Engineering, *Cinnamonum*, Alcohol

---

ABSTRACT

Eggnog has long been established as the key component of a peaceful winter holiday, as it induces somnolence in crazy uncles and sharply decreases in-laws' complaining. Despite the drink's importance, very little has been done to optimize the eggnog recipe. In this current research, DNA2.0 rational GPS engineering, Design of Experiment and Machine Learning was utilized to analyze and optimize key variables in the production of eggnog. The ability to capture multivariable interactions during optimization is critical. Six active variables were identified: dairy quantity, $C_{12}H_{22}O_{11}$ quantity, *Vanilla planifolia* quantity, *Cinnamonum verum* quantity, alcohol type, and alcohol quantity; with the goal of finding the optimal combination of all six variables. Due to the need to avoid excessive alcohol consumption and/or gluttony in the North Pole population, it was desirable to obtain statistically valid data using only a small fraction of available test combinations. Using a fractional-factorial design consisting of 18 individual trials with simultaneous multivariable combinations, a successful test program was completed. Internal validation test results of the optimum formula, with relative error of 0.48% were consistent with the theoretically predicted valued, demonstrating the reliability and success of the present study.

INTRODUCTION

Eggnog, also known as egg milk punch, is a chilled, sweetened, dairy beverage traditionally made with milk, sugar and whipped eggs (resulting in a frothy texture). Alcohol in the form of brandy, rum or bourbon is a standard addition in most protocols. The finished liquid is often garnished with ground cinnamon and/or nutmeg. Eggnog is frequently provided to elves, reindeer, and guests of the North Pole during the holiday season to encourage festive working conditions (Fig. 1).

Figure 1. Eggnog is routinely provided to North Pole residents and guests.

Eggnog is believed to have originated in East Anglia, England, developed from the warm posset, a medieval beverage made with hot milk and eggs. The term ‘nog’ may stem from the Middle English appellation ‘noggin’, a small, carved wooden mug used to serve alcohol. Alternatively, it may be derived from the Viking word ‘nog’ meaning ‘enough’. In 1066, Vikings had enough of the dreaded eggnog, left England and returned to Scandinavia [1].

Evidence shows the drink moved northward through Great Britain and was eventually carried to the North Pole in the early 17th century [2]. During the 18th century, the drink crossed the Atlantic to the British colonies. As brandy was heavily taxed in the colonies, the readily available Caribbean rum was often substituted. Kitchen records from Mt. Vernon indicate that George Washington served an eggnog which included brandy, rye whiskey, rum and sherry (Appendix A) [3]. When the supply of rum was reduced during the American Revolutionary War, Americans turned to domestically produced bourbon and moonshine as alternatives. The large number of Appalachian distillers soon made ethanol at various degrees of purity accessible to the American citizenry and which correlated with a marked increase in eggnog consumption rates [4].

The Great Eggnog Riot of 1826 occurred at the United States Military Academy at West Point on 23-25 December 1826. Whiskey was smuggled into the North Barracks of the academy to make eggnog for a Christmas party, culminating in inebriated cadets reeling through the barracks with swords, muskets and bayonets. The incident resulted in the court-martialing of 20 cadets and 1 enlisted soldier [5]. Eggnog has since been considered an enemy of the state by the Federal Bureau of Investigations (FBI) [6].

Although many attempts at determining the ideal quantities to be used in the production of eggnog have been made, all have suffered from the One Factor at a Time Optimization (OFAT) strategy [7], with less advantageous final outcomes. We instead here use Design of Experiment (DoE) Optimization, which relies on systematic variance to explore the optimization of multiple variables simultaneously (Fig. 2) [8].

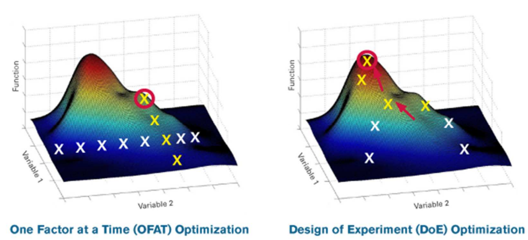

Figure 2. Example for optimizing conditions: the first set of experiments alters the temperature (X axis), the second set alters the pH (Y axis) of the reaction. The Z-dimension indicates the resulting enzymatic activity. First sampling (white) and second sampling (yellow) show how this 2-dimensional space is navigated using OFAT or DoE with each final outcome achieved (red circle). Now imagine an optimization in 5, 10, or even 50 dimensions. The DoE process is faster, more efficient, captures multivariable interactions, and finds the best solution.

In this current research, five active ingredients (6 variables) were identified with the goal to find the optimum combination of all five components. Optimal quantities of milk, sugar, vanilla, and cinnamon, as well as both alcohol type and quantity are determined and analyzed using a DoE methodology combined with machine learning. Using DoE allows for the accurate determination of the relative effects of the many different factors that influence the quality of the product.

Variable Selection

Before initiating a DoE optimization, there is often a need to identify a good starting point. The North Pole Institute of Food Science and Technology (NoPoIFoST) has developed a standardized process to uniformly and accurately sample the phylogenetic Christmas tree of one or more recipe families. The sampled sequence space is derived from public domain EggNOG v4.5 [9] database including orthologous recipes from 190,000 unique cookbooks and 2,383 different cuisines, including Mrs. Claus' previously published protocol (Appendix B: Eggnog materials and preparation guidelines).

Variable	Level 1	Level 2	Level 3
Dairy	200 ml	400 ml	600 ml
$C_{12}H_{22}O_{11}$	20g	30g	40g
Vanilla	3 ml	5 ml	7 ml
Cinnamonum	4g	6g	8g
Alcohol Type	Rum	Whiskey	Brandy
Alcohol Quantity	25ml	100ml	200ml

Table 1. Variables and levels tested in eggnog optimization.

Using these starting points, Santa's GMP eggnog production facility (Santa's Village, North Pole) was able to build a complete alignment of all homologs of the eggnog family and identify available diversity. This led to the identification of 6 key variables in the eggnog optimization: dairy (quantity), $C_{12}H_{22}O_{11}$ sweetener (quantity), *Vanilla planifolia* (quantity), *Cinnamonum verum* (quantity) and alcohol (type and quantity). Target values and levels of deviation were also determined and assigned (Table 1).

Design of Experiment

In order to test all possible combinations of the variables (full-factorial experimentation) we would need to produce and assay 729 batches of eggnog (3^6). Due to health risks (alcoholism, obesity, excessive caroling when inebriated) [10], this was deemed impractical. Using DoE methodology we were able to systematically select an axiomatic subset (orthogonal arrays) of combinations. Orthogonal arrays for the parameter design indicating the number of and conditions for each experiment were then developed. The selection of orthogonal arrays is based on the number of parameters and the levels of variation for each parameter. An orthogonal array for 6 parameters with 3 levels each requires an L18 array (Fig. 3) [11].

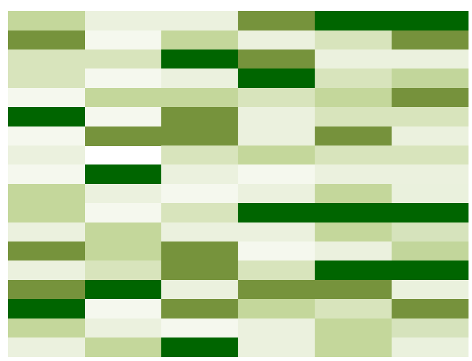

Figure 3. Heat Map for L18 Orthogonal Array comprised of 6 variables and 3 levels.

Outcomes

In accordance with IUPAC standards [12], eggnog was assessed based on 4 key sensory characteristics including color, odor, consistency and taste by trained professional testers from the NoPoIFoST (Fig. 4). Testers were asked to assess the attributes with the following weights: specific volume 10% (volume/weight, mL/g), surface texture 5%, appearance 5%, color 10%, structure 10%, viscosity 10%, acidity 10%, flavor 30%, and odor 10%. Palates were cleansed between samples using mistletoe infused snow as per NoPoIFoST protocols [13]. Samples were analyzed under normal Christmas lights conditions [14].

Figure 4. Trained professional testers from NoPoIFoST assess eggnog samples for this study.

It should be noted that there was a small, yet vocal, portion (n=3) of the testing population who preferred dairy quantity = 0ml, but as that was determined to no longer technically constitute eggnog, these data points were excluded from the final results.

Sequence−function correlation models were built using a portfolio of Santa's supervised machine learning algorithms essentially as described [15]. Briefly, machine learning algorithms were used to build non-linear models of the dataset by calculating a 6-dimensional weight vector w, where the activity Y_j of a variant X_j is estimated as $Y_j = (\sum_j=1..6 w_j x_{i,j})$, providing a quantitative measure of the relative effect of the j-th substitution on eggnog (Table 2). Bootstrapping techniques [16] were used to create 1,000 data sets containing a training subset and test subsets by randomly splitting 20% of infolog sequences (x_i) into the test set and the rest into the training set. We interpret the weight distribution as an indirect measure of eggnog variable epistasis [17].

Supervised machine learning algorithms were used to select values for w_j that resulted in the highest correlation between measured and predicted activities for test subsets. Five unique eggnog recipes predicted to significantly outperform the best recipes in the training set were calculated and the corresponding eggnogs were made in larger volumes and assessed by the general North Pole population as well as DNA2.0 employees.

To further examine the effects of alcohol quantity in eggnog in the elf population, several behavioral parameters were evaluated. The elves daily productivity and activity levels were tracked and recorded. Figure 5 displays a significant decrease in working behavior as evidenced by quantitated toy output with a simultaneous increase in leisure behavior seen in caroling [18] and participation in reindeer games.

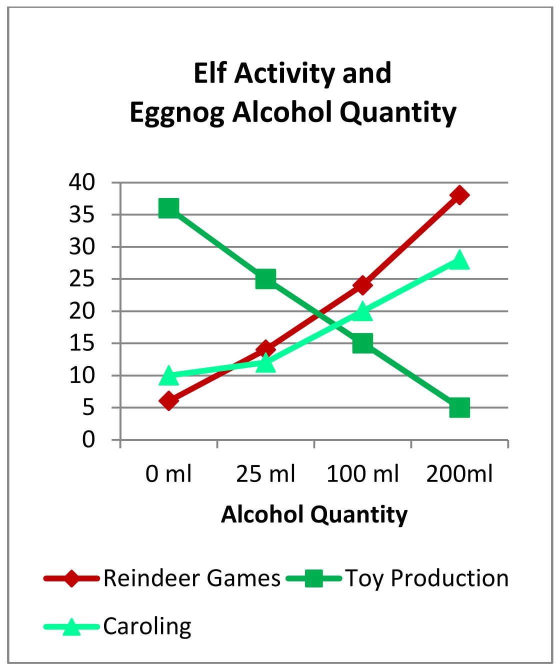

Figure 5. Assessments of productive work and enjoyment in response to alcohol quantity in eggnog, as measured by elves participation in caroling, games, and toy manufacturing.

Statistical analysis of eggnog optimization was performed by the GoogElf Technology Group. The modeling was shown quite suitable for the prediction of sensory evaluation valued of eggnog, with *P* value < 0.05, *P* value of lack of fit >0.05, and R^2 value of 0.9430.

X_1	**X_2**		
	0	**1**	**2**
0	αi (y_{00})	βk (y_{01})	γ j (y_{02})
1	β j (y_{10})	γi (y_{11})	αk (y_{12})
2	γk (y_{20})	αj (y_{21})	βi (y_{22})

Table 2. Graeco-Latin Square. Factor A and B Combinations (x1 denotes the levels of factor A and x2 denotes the levels of factor B). (α,β, γ) forms a Latin Square and (i, j,k) forms another Latin Square. (α,β, γ) and (i, j,k) jointly form a Graeco-Latin Square. This implies that SS for (α,β, γ) and SS for (i, j,k) are orthogonal. $SS_{AB}{}^2 = 3n[(\bar{y}_i - \bar{y}_\cdot)^2 + (\bar{y}_j - \bar{y}_\cdot)^2 + (\bar{y}_k - \bar{y}_\cdot)^2]$.

Conclusions

The validation test results of the optimum formula were consistent with the theoretically predicted values, demonstrating the reliability and effectiveness of DoE and machine learning. Internal ratings showed a 99.286% approval rating for the optimized formula (Figure 6).

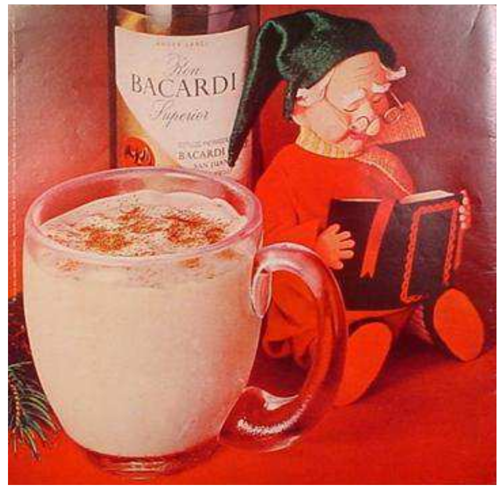

Figure 6. North Pole resident enjoying the final DoE optimized eggnog formula and protocol.

APPENDIX A: Mount Vernon Eggnog receipt

One quart cream, one quart milk, one dozen tablespoons sugar, one pint brandy, ½ pint rye whiskey, ½ pint Jamaica rum, ¼ pint sherry—mix liquor first, then separate yolks and whites of eggs, add sugar to beaten yolks, mix well. Add milk and cream, slowly beating. Beat whites of eggs until stiff and fold slowly into mixture. Let set in cool place for several days. Taste plenty and frequently.
Note: 1 quart = 946.35 ml, 1 pint = 473.17 ml, 1 tablespoon = 14.78 ml.

APPENDIX B: Mrs. Claus' Eggnog reagents and protocol

2 *Gallus gallus* domesticus Eggs, full complement of albumen and vitellus (Charles River)
30.0 grams glucose and fructose, $C_{12}H_{22}O_{11}$ (Sigma)
2.15 grams NaCl (Sigma)
4.9 ml *Vanilla planifolia* derivative, $C_8H_8O_3$ (Madagascar Laboratory Supplies)
472 ml *Bovinae bos taurus* (preferably Holstein) organic lactation product (Berkeley Farms)
25-100 ml distilled alcohol 80 proof (Bacardi)
4 grams ground *Cinnamomum verum* (Penzeys Spices)
4 grams ground *Myristicacea fragrans* (Nutmeg) (Penzeys Spices)

Protocol: Place egg with shell removed into 1L beaker, gradually add $C_{12}H_{22}O_{11}$, and NaCl. Stir thoroughly without application of temperature gradient (magnetic stirrer preferred at U North Pole, but not required). Add $C_8H_8O_3$ and lactation product, mix gently. Add desired volume distilled alcohol. Decant into personal size sterilized glassware. Dust with ground *Cinnamomum* and *Myristica*. Imbibe.

ACKNOWLEDGEMENT: We thank the entire DNA2.0 team for review and taste analyses of eggnog protocols.

Funding: The project was supported by Santa's slush-fund Grant no. 042315

Competing Interests: The authors declare competing financial interests: DNA2.0 provides protein engineering and expression, Vectorology, bioinformatics services, GeneGPS optimization and gene synthesis to a global customer base. Better, smarter and faster than anybody.

REFERENCES:

1. Bengtsson, F.G., *The Long Ships*, Stockholm, Sweden, Norstedts förlag, 1954.
2. Scrooge, E., B. Cratchit, and J. Marley, *Bah, Humbug.* Proc Natl Grinch Society, 1843. 76(13):22-6397.
3. Lender, M.E. and J.K. Martin, *Drinking in America: A History*, New York, New York, Free Press, 1987.
4. Snap, S., C. Crackle, and P. Pop, *Merrily snap, crackle and pop in a bowl of eggnog.* Kellogg's Rice Krispies, 1996. 34(38):1.
5. Agnew, J., *Eggnog Riot:The Christmas Mutiny at West Point*, San Rafael, CA, Presidio Press, 1979.
6. Weiner, T., *Enemies: A History of the FBI.* New York, New York, Random House, 2012.
7. Heinz, D. and T. Hunt, *Two tomatoes walk across the street: biological effects of the infinite ketchup battle.* J. Food Cosmet. Ind., 2002. 72:42-62.
8. Czitrom, V., *One-Factor-at-a-Time Versus Designed Experiments.* Am. Stat., 1999. 53(2):126-131.
9. Huerta-Cepas, J., et al., *eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences.* Nucleic Acids Res, 2015.
10. Claes, S., *et al.*, *Carolome: Functional Imprints of Culture Memes in Global Genome.* Proc Natl Acad Sci Northpole, 2010. 12:25-31.
11. Taguchi, G., *Introduction to quality engineering.*1986, Dearborn, MI: Asian Productivity Organization (Distributed by American Supplier Institute Inc).
12. Daniels, J., J. Beam, and J. Walker, *Fate of endogenous ^{14}C scotch, ^{235}U bloodymary and other tritiated compounds injected in laryngeal and pharyngeal pathways.* Clin. Alc. B.P.R. Soc., 1975. 89:35-78.
13. Coca, C. and K. Cola, *Gardening is the art that uses flowers and plants as paint, and the soil and sky as canvas.* J. of Electron Microscopy, 2001. 23(3):66-78.
14. Claes, S., *et al.*, *Metabolic engineering of Picea abies for receptor mediated induction of fluorescence and olfactory signaling.* Proc. Natl. Acad. Sci. NorthPole, 2008. 12:25-31.
15. Einstein, Z., D. Zweistein, and V. Dreistein, *Evaluation of different biological data and computational classification methods for use in protein interaction prediction.* Machine Learning and Knowledge Discovery in Databases, 2006. 6912: 453-469.
16. Sinatra, N., *These boots are made for walkin.* J. Shoe Manufacturing Science and Engineering, 1966. 2(4):69.
17. Alka, J.D. and M. Seltzer, *Mapping of various things for various reasons.* Hum. Genet., 2001. 32(33):33-44.
18. Garland, J., Torme, M., Jones, J., Luft, L., and Minelli, L., *The Christmas Show.* CBS Inc, The Judy Garland Show, 1963. 15(1).

Happy Holidays & Merry New 2016