

Optimizing Pistachio Roasting Process with INForm

Pistachio nuts have a high economic value for their producer companies, and a large percentage of pistachios are consumed in their roasted form. Roasting helps to develop the characteristic aroma, colour and taste, and also reduces the presence of aflatoxins. Getting the correct redness/greenness and blueness/yellowness ratios is important.

Control of the pistachio roasting process as yet is more of an art than a science, and in *LWT – Food Science and Technology* **41** (2008) 26-33, Talip Kahyaoglu has presented 13 experiments in which the temperature and time of roasting have been varied systematically, and the effect on the moisture content and colour of in-shell pistachios, pistachio kernels and ground pistachio kernels measured. For modelling, Kahyaoglu used Response Surface (statistical) methods and Gene Expression Programming. In all, 13 different outputs were measured, with a separate model developed for each property.

Kahyaoglu used RSM to determine the optimum conditions for the roasting process, but this could not involve all of the properties because there were so many of them. Getting an optimum from

INForm integrates two different modelling techniques – neural networks and Gene Expression Programming – with visualization and optimization capabilities that make it much easier to handle this problem. In this note, the modelling and optimization of Kahyaoglu's data has been undertaken.

Modelling properties of roasted pistachios

Two separate sets of models were developed for each property, one using neural networks, and the other using GEP

which has been implemented in **INForm** v4. The **INForm** default parameters were used in all cases. 10% of the data (one record) was withheld to validate the neural network model, and to facilitate comparison, the same record was also withheld for the GEP modelling.

As expected, it took longer to derive the GEP models, but this was still achieved in about twenty minutes on a Dell Inspiron Core2 Duo laptop.

Property	R ² neural network model	R ² GEP model
Moisture content	0.996	0.990
Darkness in shell	0.979	0.922
Greenness in shell	0.903	0.915
Yellowness in shell	0.823	0.855
Yellowness index in shell	0.870	0.897
Darkness in kernel	0.851	0.859
Greenness in kernel	0.910	0.929
Yellowness in kernel	0.951	0.872
Yellowness index in kernel	0.976	0.857
Darkness in ground kernel	0.911	0.912
Greenness in ground kernel	0.926	0.914
Yellowness in ground kernel	0.975	0.912
Yellowness index in ground kernel	0.397	0.900

Table 1. ANOVA R² values to assess model quality for neural network and GEP models

The results in Table 1 give the ANOVA R² values for the different models. For this

data set, both modelling methods give quite similar values, and it is clear that both have been able to develop good models. The values compare well with the results from statistical Response Surface Methods, as given by Kahyaoglu. The R^2 values from the second order statistical models are given in Table 2, and comparison with Table 1 shows that in general both the GEP and neural network models are better.

Property	R^2 statistical model
Moisture content	0.988
Darkness in shell	0.859
Greenness in shell	0.787
Yellowness in shell	0.833
Yellowness index in shell	0.632
Darkness in kernel	0.774
Greenness in kernel	0.862
Yellowness in kernel	0.714
Yellowness index in kernel	0.721
Darkness in ground kernel	0.812
Greenness in ground kernel	-
Yellowness in ground kernel	0.938
Yellowness index in ground kernel	0.709

Table 2. R^2 values for statistical models (from Kahyaoglu's work)

Optimizing the Roasting Process

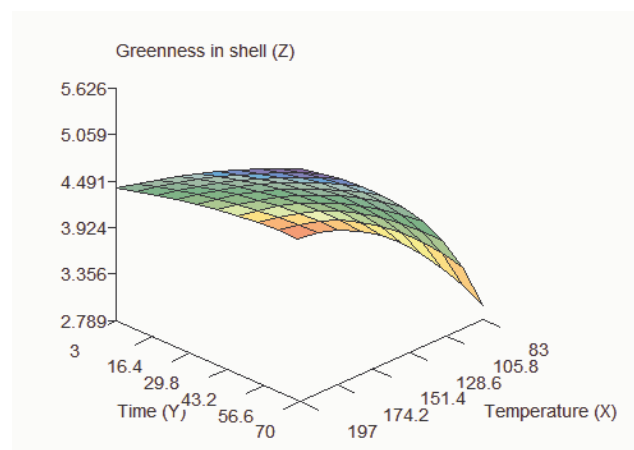
Kahyaoglu carried out an optimization of the moisture content, the darkness/whiteness in shell, and the greenness/redness and yellowness/blueness in shell, using response surface methods. The requirements were that moisture content should be between 25 and 35 g/kg, the darkness parameter should be at a maximum, greenness should be at a minimum (which actually according to Kahyaoglu's definition means the minimum redness in the pistachio), and yellowness should be in the range of 10-14.

Because optimization using response surfaces involves examining surfaces of one response compared to two inputs, finding an optimum for a large number of properties can be a challenge. Therefore Kahyaoglu has concentrated only on a few important ones.

Using **INForm**, and the GEP models which **INForm** developed, the required values for each property were specified as 'desirability

functions', with the desirabilities for any properties whose values were not specified by Kahyaoglu taken as 'flat', so that any value within the range was acceptable. All properties were given equal weighting in the optimization.

The optimum conditions from the **INForm** study suggest a temperature of 107 degrees, for 70 minutes. This is consistent with the findings of Kahyaoglu. It should be noted that 70 minutes is the maximum time in the design space, and that only one experiment corresponded to this time. This suggests that perhaps investigating longer times would be useful, if this is not inconsistent with other processing goals (e.g. cost).



Response surface for 'Greenness in shell' using INForm's GEP models

Conclusions

Both neural networks and Gene Expression Programming develop better models than the second order statistical treatment, as evidenced by the ANOVA R^2 values for each model.

Despite the fact that the GEP models developed by Kahyaoglu were better than the corresponding statistical models, the lack of integrated optimization meant that the poorer RSM models were used to determine the optimum conditions for the process. This highlights the need for integrating visualization and optimization alongside modelling, as is done in **INForm**.

Copyright 2010 Intelligensys Ltd. All rights reserved.