

Dielectric microwave sensors with multivariate calibration

F. Daschner¹ and R. Knöchel¹

¹Microwave Group, Faculty of Engineering, University of Kiel, Germany

Abstract. The composition of materials generally has an influence on its dielectric properties. Hence it may be possible to determine their composition using microwave dielectric spectra. However with complex materials like foodstuffs it is difficult to devise a suitable dielectric model. Therefore multivariate calibration methods are discussed here. With these calibrations it is also possible to process the measured scattering parameters directly. In this case it is not necessary to calculate the permittivity and more degrees of freedom are available for the design of a sensor. The direct processing of the S-parameters is demonstrated with a new transmission-line sensor.

1 Introduction

Water is often added to foodstuffs during processing. This is inevitable when the raw materials has to be cleaned. Furthermore the nutritiousness is saved and surface oxidation is reduced when the raw materials are wet. Sometimes water is added consciously in order to increase the financial profit. For quality control purposes there is a need to measure the water content of the foodstuffs. In addition the determination of other constituents like protein, fat and salt is needed. There are many measurement systems available to measure the composition of foodstuffs, but they have the disadvantage that they can only measure one constituent and they are relatively expensive and time consuming. Generally these methods alter the materials or destroy them. In short there is a need for inexpensive, fast, non destructive and compact systems for measuring the composition of foodstuffs. The method followed here to realise such a system uses the fact that a change in the composition of foodstuffs also alters their dielectric spectra in the microwave region. The problem is to relate changes of the dielectric spectrum and the composition of the material. Kent et al. have already made investigations

in this area. Principle component analysis and regression are used for the calibration (Kent and Anderson, 1996), (Kent, 1999). In the work to be described here other calibration methods are used for comparison purposes. Firstly, a dielectric modelling is applied. Secondly, orthogonalisation of the measured data is carried out with partial least squares regression (PLSR) to improve the performance. Finally, non-linear processing of the data is described using artificial neural networks. The measurements were made using the open-ended coaxial line. In order to improve the usability an alternative transmission-line sensor is discussed.

2 Experiments

The water content of flesh can be increased by dipping it in polyphosphate and saline solution. The two treatment parameters are the dipping time and the concentration of the solution. Because they have a different influence on the dielectric spectra of the samples both parameters have to be taken into account. For example the ionic conductivity of the samples will increase with the concentration of the solution.

2.1 Preparation of the samples

In the framework of a project financed by the European Commission (FAIR CT97-3020) many experiments with poultry-, pork-, and fish samples were carried out. One of those test series was selected to demonstrate the methods applied here.

Prawns (*pandalus borealis*) were peeled and treated in a polyphosphate solution with different concentrations (0%, 0.5%, 1% and 2%) and for different times (0 h, 2.5 h, 24 h, 48 h). The temperature of the solution was 4–5°C. After treatment the prawns were sieved and the surface gently dried by dabbing with a paper towel. The amount of added liquid taken up by a sample was determined by weighting the sample before and after treatment. The samples were minced and divided into two parts. One part was used for the dielectric measurement while the other was sent to an external laboratory, which determined the composition of the samples using

Table 1. Results of the dielectric modelling procedure

Prawns	R^2	RMSEc	RMSEv
Liquid-uptake [% of m_0]	0,745	2,75	3,37
Fat [% absolute]	0,616	0,064	0,077
Protein [% absolute]	0,889	0,59	0,73
Water [% absolute]	0,852	0,61	0,70
Salt [% absolut]	0,771	0,047	0,061

the conventional methods mentioned above. These measurements served as reference for the calibration methods discussed here.

2.2 Dielectric measurements

The dielectric measurements were carried out using an open-ended coaxial line with a diameter of 3 mm. This sensor had been well investigated and has become a standard for measuring the dielectric spectra of liquids and soft materials (Gadja and Stuchly, 1983), (Stuchly et al., 1982), (Marsland and Evans, 1987). The dielectric probe kit HP85070 in combination with a HP8510 automatic network analyser was used for the measurements. The dielectric spectra of the samples were measured in the frequency range from 200 MHz to 12 GHz at 31 frequency points with logarithmic separation. Each measurement was repeated five times. Outliers caused by bad contact between the probe and the sample were removed. The means of the permittivity values were calculated and used in order to remove the influence of remaining inhomogeneities of the samples. A total number of 91 data sets were available for the data processing.

3 Processing of the data

3.1 Dielectric modelling

Foods are generally complex materials. First of all flesh is inhomogeneous. It consists of different kinds of tissues like muscles, blood vessels, connective and fatty tissue. The main constituent of flesh is water. A part of it is within the cells while the other part is in the extra cellular space. But Flesh also consists of many other organic as well as inorganic substances. Some of these ingredients have complex interactions. A part of the water is bound by hydrophilic groups of organic molecules. Another fraction of the water is rotationally hindered by other molecules. The dielectric dispersion of water is in the range of microwave frequencies depending on its temperature. Smaller peptides and amino acids also have relaxation frequencies in the microwave region (Grant et al., 1978).

However, the microwave dielectric spectrum is mainly formed by water. due to the complex interaction it is difficult to create a sufficiently accurate model for the dielectric behaviour of flesh. On the other hand it should be possible to determine the existence of constituents which do not

have a relaxation frequency in the measured frequency range – provided that they have an interaction with water. The first method discussed here tries to apply a dielectric model of the flesh. The applied model consists only of three components: The free water (γ -Dispersion), the bound water (δ -Dispersion) and the salt content.

$$\epsilon_r(\omega) = \epsilon_\infty + \frac{\epsilon_{s\gamma} - \epsilon_\infty}{1 + (j\omega\tau_\gamma)^{(1-\alpha_\gamma)}} + \frac{\epsilon_{s\delta} - \epsilon_\infty}{1 + (j\omega\tau_\delta)^{(1-\alpha_\delta)}} - j \frac{\sigma}{\epsilon_0\omega}. \quad (1)$$

Where ϵ_∞ is the permittivity at infinite Frequency, ϵ_s is the static permittivity, ω is the angular frequency, τ is the mean relaxation time, α describes the distribution of the relaxation times and σ is the conductivity. Both the free and bound water are described by Cole-Cole-relaxations while the salt content is considered in the 3rd term. ϵ_∞ is fixed to the value of 4.6 for both dispersions.

The seven model parameters are calculated from the measured dielectric spectra using a non-linear curve fitting procedure. This procedure is described in (Grant et al., 1978) and based upon the Gauss-Newton-method. The calculation effort of the curve fitting is relatively high ($\approx 10^6$ floating point operations for one curve fitting). After the calculation of the model parameters the samples were divided randomly into a calibration (2/3) and a validation group (1/3). The model parameters of the calibration group are used as regressors in a multiple linear regression (MLR) (Johnson and Wichern, 1992). A composition value (e.g. water content) is described as linear combination of the model parameters. The results for the prediction of the composition values using dielectric modelling are shown in Table 1.

By means of the coefficient of determination R^2 one can see that the performance of this method is disappointing. With the prawn samples only the water and protein content is predicted with moderate accuracy ($R^2 > 80\%$). The root mean square errors (RMSE) of the calibration and validation group are also given in this table.

Obviously the poor performance can be explained by the simplicity of the model used which considers only three components. One could expand the model now to improve the accuracy, but the measurement accuracy is not high enough to fit such a detailed model with any confidence. The dominant source of error with the open ended coaxial line is the contact between the probe and the sample. Furthermore the material is not homogeneous enough. Hence it is not reasonable to expand the model with other components. On the other hand the use of too primitive a model can cause a loss of information during the curve fitting procedure.

3.2 Partial least squares regression

Now the measurement problem is generalized and no model has to be found but rather a function which describes the relationship between the dielectric spectrum and the composition values. The direct use of the dielectric data in a MLR has the

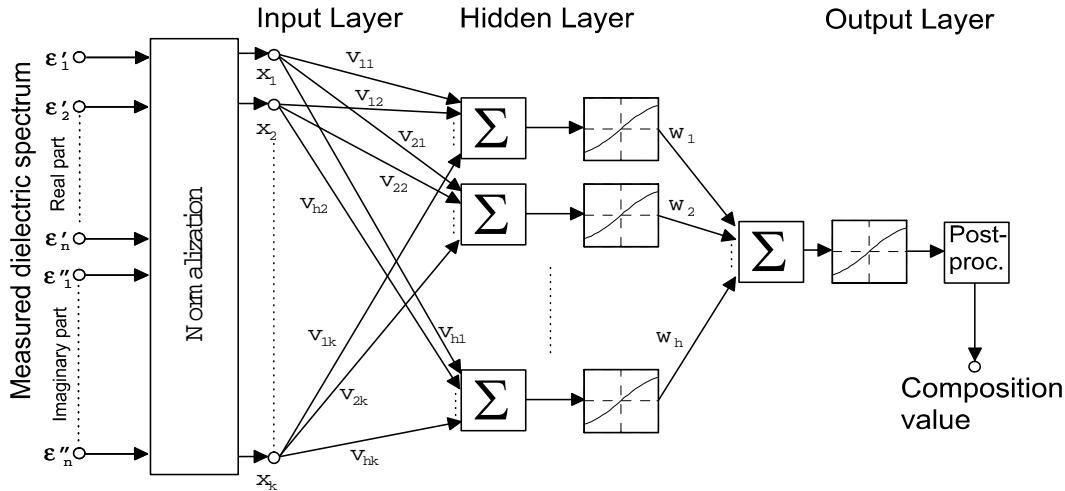


Fig. 1. Architecture of the used MLFF-ANN.

Table 2. Results obtained with PLSR

Prawns	R^2	RMSEc	RMSEv
Liquid-uptake [% of m_0]	0,744	2,76	3,27
Fat [% absolute]	0,436	0,078	0,086
Protein [% absolute]	0,936	0,45	0,66
Water [% absolute]	0,942	0,38	0,57
Salt [% absolute]	0,845	0,039	0,052

Table 3. Results obtained with the used MLFF-ANN

Prawns	R^2	RMSEc	RMSEv
Liquid-uptake [% of m_0]	0,919	1,55	2,59
Fat [% absolute]	0,800	0,046	0,076
Protein [% absolute]	0,939	0,43	0,53
Water [% absolute]	0,933	0,41	0,46
Salt [% absolute]	0,792	0,045	0,047

disadvantage that the data are too collinear. Hence the calculation of the coefficients of the calibration equation may be numerically unstable and overfitting occurs.

For this reason Kent and Anderson proposed the use of principal component analysis (PCA) and principal component regression (PCR) for the processing of dielectric data. The principal components are an orthogonal linear transformation of the measured data. Hence the collinearity is removed completely. The only criterion for the orthogonalisation is to maximise the variance of the new data.

The partial least squares regression (PLSR) performs the calibration more directly, the composition values being considered during the orthonormalisation. Originally the PLSR was developed to process economic data, but Martens used this technique also for near infrared spectroscopy (Marten and Naes, 1989). (Archibald et al., 1998) PLSR has also been

applied to the determination of the water content of wheat using microwave transmission measurements. PLSR reduces the data to a new set of data which is called *hidden path variables*. The PLSR algorithm used here is called *PLS1* and is described in (Marten and Naes, 1989).

The only degree of freedom is the number of hidden path variables to be calculated. If too many of those are used overfitting occurs. The optimal number of factors used has to be determined empirically for each composition value (here: ≈ 12). The results obtained with PLSR are shown in Table 2. They are better than those obtained with the dielectric modelling.

3.3 Nonlinear data processing using artificial neural networks

The calibration equation obtained with PLSR is a linear combination of the measured dielectric data, but the unknown function may be non linear. One method to approximate a non linear function is the use of artificial neural networks (ANN). Suitable architectures for function approximation are multi layer feed forward (MLFF) networks (Patterson, 1997). The architecture of the MLFF-ANN used is shown in Fig. 1. Many degrees of freedom are left in this architecture. Unfortunately using the only guidelines known all these parameters would be chosen. Therefore intuition must be invoked in designing such a network by trial and error. The configuration of the displayed network proves very well suited for the application discussed here. It has one hidden layer, which contains 5–10 neurons. The activation functions of the neurons in the hidden layer are non-linear (*tansig*-function) while those of the output layer are linear.

After a fundamental architecture has been chosen the ANN has to be trained. That requires the determination of the weights (w and v) of the connections between the inputs and neurons in order to minimise the error of the function approximation. The artificial neural networks discussed here are trained with back-propagation which is a descent gradi-

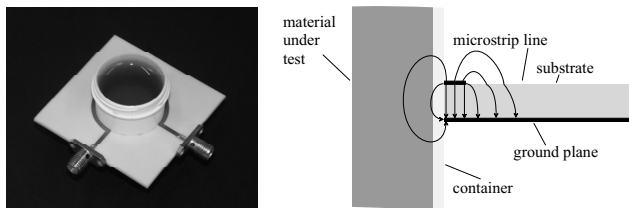


Fig. 2. Left hand side: Photograph of the transmission-line sensor. Right hand side: Cross section of the transmission-line.

ent algorithm. It can happen that the training algorithm converges only to a local minimum or it may not converge at all. Another problem is the requirement of random numbers to be used as starting values of the weights. Hence the success of a training exercise depends on those starting values, and a bad training has to be repeated. A further disadvantage is that ANN's have a tendency to overfit. That means the network imprints only the training data and loses its required properties of generalisation. The more complex the network is the higher is the risk of over-fitting. To avoid overfitting the method of *early stopping* (Mathworks, 2000) was used.

The results obtained using the best configuration found are in Table 3. A distinct improvement of the RMSE_v is observable. Especially with the estimated liquid uptake the accuracy is conspicuously higher in comparison to the other methods discussed here. Unfortunately, in comparison to them more calibration data are necessary. The calculation effort during the training is some order of magnitude higher. In short the training of a MLFF-ANN is more expensive in computer time and requires more calibration data sets. The higher effort however, leads to a better prediction of the composition values.

It can be seen that the prediction of the fat and salt content is relatively poor. This is explainable by the small range of these composition values within all samples of the test series. With other test series where a higher variation obtains (e.g. fat content in herring) better results are found.

4 Transmission-line sensor

The open-ended coaxial line has some disadvantages in practical use. For example, the sensitive area is relatively small, and the sensor is soiled during each measurement. The dielectric data are only an intermediate result for the determination of the composition values. They are calculated from the measured reflection factors at the aperture of the open-ended coaxial line. Because PLSR and the ANN do not need any physically modelling it is also possible to use the measured S-parameters directly as input variables. This means that there are more possibilities for the design of a sensor.

The new sensor discussed now is designed with regard to a reduction of the disadvantages of the open-ended coaxial line. The material under test is filled in cheap containers which have a -slightly conical-, cylindrical shape and a volume of 10 ml. They are made of polypropylene with a thick-

ness of ~ 1 mm. Hence electromagnetic fields can penetrate the walls of the containers. The transmission line, which is constructed in microstrip, surrounds the container. The container is placed into a hole in the board. A photograph of the design is shown in Fig. 2.

In Fig. 2 a cross section of the transmission line is also shown (vertical to the propagation direction of the wave). The expected electric field is outlined too. One can see that there is no homogeneous field distribution. The electric field penetrates all 4 materials: air, the substrate of the board, the polypropylene of the container wall and finally the material under test. But only a part of the fringing field penetrates the latter. Hence the sensitivity of this sensor may be low. On the other hand the attenuation caused by the losses of the material may be limited. In short it is difficult to model the topology with analytical equations. The set-up was rather designed with improvement of practical usage being the main consideration.

Performance of the new sensor was verified with a further test series. Five kinds of honey were diluted in 10 grades up to an absolute water content of 22%. The frequency points between 2 GHz and 6.1 GHz were selected as input variables for the PLSR in order to determine the water content of the honey samples. Below 2 GHz the sensitivity of the sensor is too low. PLSR was also carried out for the measurements using the open-ended coaxial line for comparison purposes.

The results can be seen in Fig. 3. The prediction of the water-content is more accurate when the transmission line sensor is used. The RMSE of the validation group is only 0.48% absolute while the RMSE_v obtained with the open-ended coaxial line is 0.736%. The coefficient of determination R^2 is also better with the transmission-line sensor: it increases from 92.2% to 98.6%.

However further investigations have to be done in order to examine if it is also possible to determine the composition of more complex foodstuffs like prawns.

5 Conclusions

The determination of the composition of foodstuffs using conventional methods is time consuming, alters or destroys the samples and needs a lot of expensive devices. The approach suggested here is based upon the influence of the ingredients on the microwave dielectric spectrum. Several methods were investigated to extract the hidden information about the composition of the samples from the dielectric spectrum. The first method uses a simple dielectric model. However the performance of this method was not satisfactory and the computational effort was relatively high. The use of partial least squares regression leads to better results. The use of multi layer feed-forward artificial neural networks (MLFF-ANN) enables a non-linear processing of the data. The computational effort during calibration with respect to training is relatively high and more calibration data are needed, but with this method the best results were obtained.

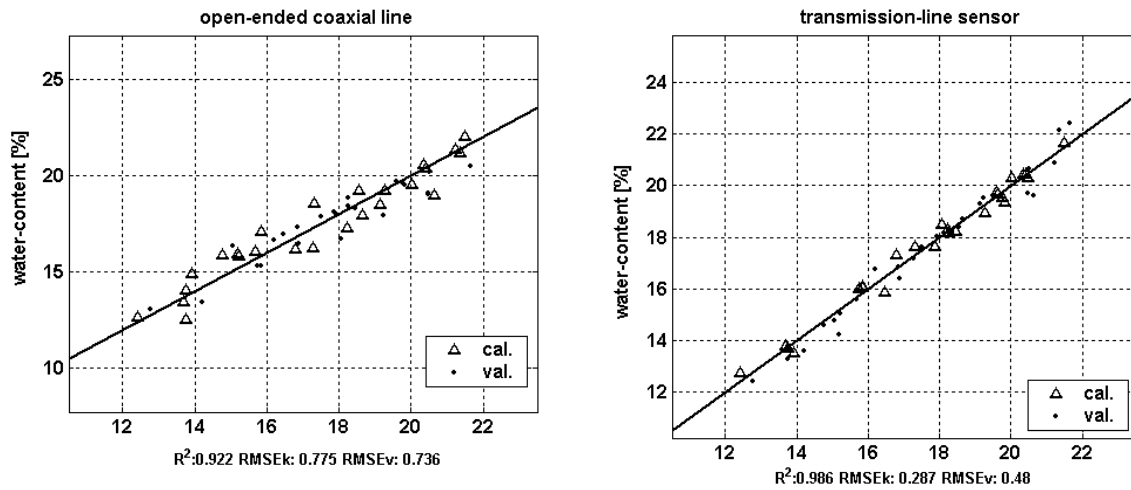


Fig. 3. Comparison of the results for the prediction of the water-content of the honeys samples. Left hand side: open-ended coaxial, right hand side: transmission line sensor.

The strategy to process directly the measured S-parameters leads to more degrees of freedom for the design of a probe. This was demonstrated with a newly developed transmission-line sensor. Many other arrangements are conceivable which have greater practicability than the open-ended-coaxial line.

Acknowledgement. The contributions and many discussions with Dr. Mike Kent are greatly appreciated. This study has been carried out with financial support from the Commission of the European Communities, Agriculture and Fisheries (FAIR) specific RTD program FAIR CT97-3020 Added Water in Food Products. It does not necessarily reflect the Commission's views and in no way anticipates its future policy in this area.

References

- Archibald, D. D., Trabelsi, S., Kraszewski, A. W., and Nelson, S. O.: Regression Analysis of Microwave Spectra for Temperature-Compensated and Density-Independent Determination of Wheat Moisture Content, *Applied Spectroscopy*, 52, 11, 1998.
- Gajda, G. and Stuchly, S.: An Equivalent Circuit of an Open-Ended Coaxial Line, *IEEE Transactions on Instrumentation and Measurement*, IM-32, 4, pp. 506–508, 1983.
- Grant, E. H., Sheppard, R. J., and South, G. P.: Dielectric Behaviour of Biological Molecules in Solution, Oxford University Press, Oxford, 1978.
- Johnson, R. A. and Wichern, D. W.: Applied Multivariate Statistical Analysis, Prentice-Hall, New Jersey, 1992.
- Kent, M. and Anderson, D.: Dielectric studies of added water in poultry meat and scallops, *Journal of food engineering*, 28, pp. 239–259, 1996.
- Kent, M.: Simultaneous Determination of Composition and Other Material Properties by Using Microwave Sensors, *Sensors Update*, 7, pp. 4–25, Wiley-VCH, Weinheim, Germany, 1999.
- Mathworks: Neural Network Toolbox, User's Guide, 4, The Mathworks Inc., 2000.
- Marsland, T. P. and Evans, S.: Dielectric Measurement with an Open-Ended Coaxial Probe, *IEE Proceedings*, 134, 4, pp. 341–349, 1987.
- Martens, H. and Naes, T.: Multivariate Calibration, John Wiley and Sons, Chichester, 1989.
- Patterson, D.: Künstliche neuronale Netze, Prentice Hall Verlag, Haar, 1997.
- Stuchly, M., Brady, M., Stuchly, S., and Gajda, G.: Equivalent Circuit of an Open-Ended Coaxial Line in a Lossy Dielectric, *IEEE Transactions on Instrumentation and Measurement*, IM-31, 2, pp. 116–119, 1982.