Near infrared spectroscopy and machine learning classifier of crosslink density level of prevulcanized natural rubber latex

P Sirisomboon¹, C H Lim², J Posom³,*

¹Department of Agricultural Engineering, School of Engineering, King Mongkut’s Institute of Technology Ladkrabang, 1 Chalongkrung Road, Ladkrabang, Bangkok, 10520, Thailand
²Thai Rubber Latex Group Public Company Limited, Nongyai, Chonburi, 20190, Thailand
³Department of Agricultural Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen, 40002, Thailand

* Corresponding author: jetspo@kku.ac.th (Jetsada Posom)

Abstract. By toluene swell index for cross link density level of prevulcanized (PV) rubber latex knowledge, toluene swell of PV latex is measured for trading and production management. Therefore, aim of this research is to use the Fourier transform near infrared (FT-NIR) spectroscopy with machine learning to classify different cross link density levels by toluene swell index including, Unvulcanized (U) (> 160% swell), Lightly vulcanized (L) (100-160% swell), Moderately vulcanized (M) (80-100% swell), Fully vulcanized (F) (< 80% swell) of prevulcanized (PV) natural rubber latex of raw PV latex and 50% solids content PV latex (PV50). The result shows that toluene swell index of rubber prevulcanized latex could be 91.8% correctly classified into L group and M group using PV50 MSC pretreated spectra with PLS-DA classifier. Unfortunately, sample obtained for this experiment were loss of U and F groups. In future, to develop the robust model, the sample of all crosslink density levels should be collected.

1. Introduction

NIR spectroscopy have been studied for feasibility determination of cross link density of rubber latex such as cross link density determined by prevulcanisate relaxed modulus (PRM) test [1] using NIR Fourier Transform spectrometer (FT-NIR) and [2] using low cost spectrometer (LCS) and toluene swell test [3] using FT-NIR, when reference methods has been used in rubber latex factory. For prevulcanized latex an effective model was developed using partial least squares regression with preprocessing (first derivative + straight line subtraction method). The coefficient of determination ($R^2$), root mean square error of cross validation of the calibration set were 0.85 and 2.98% and coefficient of determination ($r^2$), root mean square error of cross validation and bias of the validation set were 0.71, 3.93% and -0.005%, respectively. By PRM test for PV 50% (prevulcanised latex of 50% total solids content), FT-NIR gave coefficient of determination of validation ($r^2$) and root mean square error of prediction (RMSEP) of 0.88 and 6.74×104 Nm$^{-2}$, respectively, while scanned PV50 latex in quartz cup [1] and LCS gave 0.70 and 8.13×104 Nm$^{-2}$, respectively, while dipped probe in PV50 latex in glass beaker [2]. For FT-NIR spectrometer to be used in prevulcanized latex production, Williams et al [4] indicated this model was for screening and some other approximate calibration ($r^2$ is between 0.66-0.81). Therefore, objective of this manuscript was to use FT-NIR spectroscopy and machine learning including partial least squares
discriminant analysis (PLS-DA) and artificial neural network (ANN) classifier to differentiate cross link density levels including Unvulcanized (U), Lightly vulcanized (L), Moderately vulcanized (M), Fully vulcanized (F) of prevulcanized (PV) natural rubber latex of raw PV latex and 50% solids content PV latex (PV50). Table 1 shows NIR spectroscopy has pro and con for property evaluation.

**Table 1.** NIR technique and toluene swell technique pros and cons.

| NIR technique | toluene swell technique |
|---------------|-------------------------|
| Fast (< 5 minutes) | Slow (2-3 h) |
| No chemical used, no sample preparation, no skilled analyst required | Make film. Use solvent toluene which is toxic |
| Equipment expensive to purchase | Tools used are simple (die cutter, petri dishes, roller). |
| Flexible as many constituents can be tested simultaneously | Can test only one parameter i.e., toluene swell |
| Separate calibration is needed for each parameter | No calibration needed |
| Lack of knowledge in its operation can affect its accuracy | Test is simple and reproducible |

2. **Experimental**

2.1. **Sample preparation**

The PV natural rubber latex samples for the experiment were provided by the factory of Thai Rubber Latex Group Public Company Limited in the Nongyai district, Chonburi province, Thailand. Typical formulation of PV natural rubber latex includes 20% KOH, 20% Ammonium laurate, 50% Sulphur, 50% ZnO and 0% ZDBC of 0.07, 0.08, 0.30, 0.20 and 0.40 Phr (Parts per hundred rubber), respectively. Totally seven 100-metric tonne storage tanks of PV latices were selected over a year. Each week, 3 samples were drawn from each tank at different levels (upper, middle and lower) for 8 consecutive weeks starting from the date of production (totally 9 times). The solids content of PV latices varied from batch to batch between 60.0 - 61.5%. For the purpose of comparison, the total solids content of PV latices should be the same for every batch. As such, the PV latices of all the batches were diluted to 50% solids content, hence PV50.

2.2. **NIR spectral acquisition**

NIR spectral acquisition was followed [1] in Near Infrared Spectroscopy Research Center for Agricultural Product and Food (www.nirsresearch.com) at Department of Agricultural Engineering, School of Engineering, King Mongkut’s Institute of Technology Ladkrabang, Bangkok, Thailand. The Fourier transform near infrared (FT-NIR) spectrometer (MPA, Bruker, Ettlingen, Germany) was used to scan the samples at the wavenumber range of 12,500-3600 cm⁻¹. The gold plate was scanned internally for background reading before scanning the samples. For spectra of PV and PV50 latices, they were scanned by cup of 43 mm in diameter and 50 mm in height with the base area of 1453 mm² made of quartz material contained latex sample scanned through quartz bottom with diffuse reflectance mode at room temperature of 25±2°C. The apparent absorbance values were recorded using Log 1/R.

2.3. **Analysis of toluene swell by the reference method**

The toluene swell method is normally used for measuring the crosslink density of compounded or PV latex and toluene swell product. Analysis of toluene swell followed [3]. The percentage toluene swell was calculated from equation (1).

\[
\%\text{swell} = \frac{Y - X}{X} \times 100 \tag{1}
\]
where X is initial diameter and Y is swollen diameter read from the graph paper. The variation in %swell with crosslink density, for conventionally cured natural rubber in toluene, may be broadly described (Table 2) [5]. Latex concentrate before the additional of compounding chemicals has no crosslink and the toluene swell will exceed 160 % and the latex film test sample may even dissolve in the toluene rendering it immeasurable.

| % swelling       | Crosslink density level         |
|------------------|--------------------------------|
| > 160%           | Unvulcanized rubber            |
| 100-160%         | Lightly vulcanized             |
| 80-100%          | Moderately vulcanized          |
| < 80%            | Fully vulcanized               |

Table 2. Percentage swell and crosslink density level.

Figure 1 shows flow diagram for classification analysis. Detail of classification analysis is following:

Before the building of classification calibration models, principal component analysis (PCA) was used to detect outlier sample and no outlier was identified and removed. Before the calibration, by descendant running %swell data of 147 samples and assigning 3 samples in calibration set and 1 sample in prediction set (108 samples for developing a calibration equation and 39 of prediction set samples for validating the calibration equation). Absorbance spectra were pre-processed for optimal performance using multiplicative scatter correction and (MSC) standard normal variate (SNV). Principal component analysis (PCA) was performed on rubber latex spectra using maximum PC number determination on basis of error analysis described by Ni et al. [6] to get rid of risk to reserve noise and overfitting model if too many PC were used for modelling. PLS-DA and ANN were classifier for 4 levels of crosslink density of rubber latex using percentage toluene swell. However, only Lightly vulcanized (L) and Moderately vulcanized (M) levels were possible to be collected. The classification was performed on raw, pretreated spectra, and PCA pretreated spectra and the classification performance was assessed using sensitivity, specificity, and F1-score. Sensitivity indicated how accurate of classifier not to assign other group sample to be in model group. Specificity indicated how accurate of classifier to assign model sample to be in model group. F1 score indicated classifier either accurate of classifier not to assign other group sample to be in model group or accurate of classifier to assign model sample to be in model group. In our case, MATLAB R2021a (Mathworks Inc.) was used for spectral pretreatment, PCA, PLS-DA and ANN.
3. Results and discussion

Figure 2 shows a) PV and b) PV50 latex spectra scanned by FT-NIR spectrometer when water O-H bond vibration (1190, 1450 and 1940 nm) on 8285, 6888 and 5191 cm\(^{-1}\) (1207, 1452 and 1926 nm) were obviously illustrated and Para rubber polyisoprene absorbance band on 5924-5446 cm\(^{-1}\) (1688-1836 nm) and 4219 cm\(^{-1}\) (2370 nm) when compared to dried pure Para rubber sheet absorbance spectrum [7] where 1700-1884 nm absorbance band and 2278-2500 nm over absorbance band were illustrated.

Statistical data of crosslink density of rubber latex samples indicated by dried rubber sheet % swelling in toluene where maximum, mean, minimum and standard deviation value were 120, 92.81, 82 and 8.58%, respectively. There was no sample of unvulcanized latex due to our sample was prevulcanized latex. Crosslinking starts immediately upon additional of the compounding chemicals. For even distribution of chemicals to ensure complete reaction, the latex compound has to be gently stirred for 24 h. As such, on the first measurement of the crosslink density, a certain amount of crosslinks had already occurred. By the 8\(^{th}\) week, which was the upper boundary of our sample collection, most of the compounding chemicals would have been used up and the reaction continued at a very slow rate. However, upon further storage (about 3 months), where no sample collected, the toluene swell would reduce further to < 80% first by high modulus (HM), then middle modulus (MM) and low modulus (LM) as HM has a higher level of vulcanising chemicals. The cohesive mechanism of the polymer becomes stronger as the crosslinking increases, whereas the solvating power of the solvent remains constant. Therefore, a crosslinked high polymer substance is not dissolved but merely swelled, and the swelling decreases with increasing crosslinking. Thus, HM has more crosslinks than MM, and MM has more cross links than LM. The speed of crosslink formation was influenced by the external temperature of the environment.

Models from PV50 latex spectra was slightly better than from PV latex due to the effect of different total solids content in latex was suppressed. Among ANN, PCA-ANN and PLS-DA algorithm and different pretreatment methods (SNV and MSC), PLS-DA with MSC provided highest performance for classification with sensitivity, specificity and F1 score of prediction set of 87.5, 96.6 and 91.8% for lightly vulcanized level latex and 96.6, 87.5 and 91.8% for moderately vulcanized level latex, indicating PLS-DA classifier for lightly vulcanized level latex had 12.5% error in wrong assigning moderately vulcanized group and had 3.4% error in wrong assigning lightly vulcanized latex group and for moderately vulcanized level latex had 3.4% error in wrong assigning lightly vulcanized group and had 12.5% error in wrong assigning moderately vulcanized latex group and both classifier had 91.8% accuracy in classification. There was no literature reporting on classification of crosslink density of rubber latex. NIR spectroscopy exhibits good sensitivity offering detection limits on the order of 0.1% (w/w) concentration of constituent in quantitative analysis.[8] However, in qualitative analysis for group classification in case of vulcanized level by toluene swelling of 4 groups (> 160%, 100-160%, 80-100%, < 80%) in Table 2 where range of percentage of any one group is wide. Therefore, it is possible to classify into different 4 groups. Group classify detection limit by NIR spectroscopy though has been not
specified depends on how different of spectrum structure in different groups and how wide group reference properties for classification.

4. Conclusions
Crosslink density by toluene swell index of rubber prevulcanized latex could be 91.8% correct classified into lightly vulcanized latex group (100-160%swell) and moderately vulcanized latex group (80-100%swell) using FT-NIR PV50 MSC pretreated spectra with PLS-DA classifier. However, sample collected in this experiment were lack of unvulcanized rubber (> 160%swell) and fully vulcanized (< 80%swell). Therefore, more sample have to be collected to complete crosslink density range for develop applicable model to be used in prevulcanized rubber latex and product industries.

Acknowledgement
The authors would like to express their sincere thanks to Thailand Research Fund for the financial support (Grant number RDG5950011).

References
[1] Lim CH and Sirisomboon P 2017 J Near Infrared Spectrosc. 25 407–415.
[2] Lim CH and Sirisomboon P 2021 Industrial Crops and Product 159 1130162021.
[3] Lim CH and Sirisomboon P 2018 Journal of Near Infrared Spectroscopy 26(3) 159-168.
[4] Williams P, Antoniszn J and Manley M 2019 Near-infrared Technology: Getting the best out of light, AFRICAN SUN MeDIA, Stellenbosch, South Africa, 301 pages.
[5] Mausser RF 1987 (ed) The Vanderbilt latex handbook. 3rd ed. Norwalk, CT, USA, R.T. Vanderbilt Company Inc.
[6] Ni L-J, Zhang LG, Xie J, Luo JQ 2009 Analytica Chimica Acta. 633 43-50.
[7] Kaewkaebthong A 2012 Master thesis of Faculty of Engineering, King Mongkut’s Institute of Technology Ladkrabang, Bangkok, Thailand.
[8] Talebi M, Armstrong DW 2020 Water determination in Specification of Drug Substances and Products (Second Edition) Edited by: Christopher M. Riley, Thomas W. Rosanske and George Reid, Elsevier Ltd.