THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Investigation of Data Centric Diagnostic Techniques for Transformer Condition Assessment

Yi Cui

M.E (Electrical), B.E (Electrical)

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2016

School of Information Technology and Electrical Engineering

Abstract

Power transformer is one of the most important and expensive equipment in a power system. Its reliability directly affects a power system. To ensure the reliable operation of a power transformer, its condition needs to be continuously monitored and evaluated. Over the past two decades, a number of diagnostic techniques have been developed for transformer condition assessment such as dissolved gas analysis (DGA), degree of polymerization (DP) measurement, polarization and depolarization current (PDC) measurement, frequency domain spectroscopy (FDS), frequency response analysis (FRA), and partial discharge (PD) detection. However, the interpretations of measurement results acquired from these diagnostics are usually based upon the empirical models, which are sometimes inaccurate and incomplete especially in abnormal transformer operation scenarios. Therefore, accurate interpreting on the measurement data obtained by the above techniques and subsequently making explicit condition assessment of transformers is still a challenge task.

Nowadays, considerable efforts have been made in the field of transformer condition monitoring and assessment. Majority efforts are dedicated in developing accurate transformer models and reliable transformer fault diagnosis systems. After completing a comprehensive literature review on various diagnostic techniques for transformers condition assessment, this thesis focuses on three main aspects of transformer's health condition, including oil characteristics and dissolved gases in transformers, moisture concentration of oil-cellulose insulation and hot spot temperature of transformer windings.

Since there is a lack of common framework for applying pattern recognition algorithms (i.e. data centric approaches) to interpret oil characteristics and DGA data, this thesis firstly provides a critical review on various pattern recognition techniques for power transformer insulation diagnosis using DGA and oil characteristics datasets. A general pattern recognition application framework is then proposed. The important issues for improving the applicability of pattern recognition techniques for transformer insulation diagnosis are also discussed.

To improve the data quality of training database and enhance the diagnostic accuracy of pattern recognition algorithms, a hybrid algorithm, SMOTEBoost is proposed. It adopts Synthetic Minority Over-sampling Technique (SMOTE) to handle the class imbalance problem, where samples belonging to different fault types (insulation conditions) are unevenly distributed in the training database. By using the boosting approach for reweighting and grouping data points in the training database, the SMOTEBoost can facilitate pattern recognition algorithms consistently attaining desirable diagnosis accuracies.

To solve the intricate difficulties in moisture estimation of transformer oil-cellulose insulation system, this thesis introduces two modelling approaches, i.e., multi-physics finite element modelling (FEM) and particle tracing method, where the temperature dependent moisture dynamics in transformers' insulation system is taken into account. In multi-physics approach, moisture dynamics is modelled by coupling the electromagnetic, thermal, fluid flow and moisture migration physics simultaneously. In particle tracing method, moisture diffusion is formulated from a microscopic view of water molecules' motion. Especially, the transmission probability of water molecules (termed as particles in the paper) is employed to correlate the microscopic particles' motion with the macroscopic moisture distribution. Extensive ageing and moisture diffusion experiments have been conducted on a prototype transformer to verify the proposed modelling approaches for an accurate estimation of moisture in transformers.

This thesis also proposes a distributed parameter model to investigate the effect of moisture dynamics on dielectric response of a transformer's cellulose insulation. The correlation between moisture distribution (under non-equilibrium conditions due to thermal transients) and dielectric response parameters (dielectric losses and permittivity) of transformer cellulose insulation is revealed. The proposed methodology can help the proper interpretation of dielectric response measurement of field transformers under thermal transients.

To overcome the inaccuracy in empirical thermal dynamic models, in this thesis a moisture dependent thermal model (MDTM) is developed for estimating transformer's hot spot temperature. In this model, nonlinear thermal resistance is formulated by considering both oil and cellulose (paper and pressboard) of the transformer. Especially, the effect of moisture concentration and hot spot temperature on the thermal resistance of cellulose is taken into account. The proposed MDTM is verified by using historical data of moisture-in-oil and temperature measurements on an in-service vegetable oil-filled transformer.

To integrate every piece of data and information obtained from different transformer diagnostic measurements and subsequently evaluating the overall health condition of a transformer, this thesis proposes a data and information fusion framework based on Bayesian Network (BN). Within the Bayesian Network, Monte Carlo and Bootstrap methods are employed to extract the most informative characteristics regarding transformer condition from different diagnostic measurements. Results of case studies demonstrate that the proposed data and information fusion framework can evaluate the effectiveness of combinations of different diagnostic measurements and subsequently facilitate determining optimal diagnostic strategies involved in transformer condition assessment. It is expected that the data centric diagnostic approaches developed in this thesis can provide an accurate modelling and reliable assessment of transformer's health condition.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Peer-reviewed journal publications

1. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake, "Understanding Moisture Dynamics and Its Effect on Dielectric Response of Transformer Insulation," IEEE Transactions on Power Delivery, Vol. 30, Issue 5, pp. 2195- 2204, October, 2015.

2. **Yi Cui;** Hui Ma; Tapan Saha, "Improvement of Power Transformer Insulation Diagnosis Using Oil Characteristics Data Preprocessed By SMOTEBoost Technique," IEEE Transactions on Dielectrics and Electrical Insulation, Vol.21, Issue 5, pp. 2363 - 2373, October, 2014.

3. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Guangning Wu, "Multi-physics Modelling Approach for Investigation of Moisture Dynamics in Power Transformers," Paper accepted for publication in IET Generation Transmission & Distribution (accepted on February 6th, 2016).

4. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Daniel Martin, "Particle Tracing Modelling on Moisture Dynamics of Oil-impregnated Transformer," Paper accepted for publication in IET Science, Measurement & Technology (accepted on January 12th, 2016).

5. **Yi Cui;** Hui Ma; Tapan Saha, "Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 1: Framework, Literature, and Illustration", International Transactions on Electrical Energy Systems, Vol.25, Issue 10, pp. 2247-2259, October, 2015.

6. **Yi Cui;** Hui Ma; Tapan Saha, "Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 2: Implementation, Case Study, and Statistical Analysis", International Transactions on Electrical Energy Systems, Vol.25, Issue 10, pp. 2260 - 2274, October, 2015.

7. Daniel Martin; **Yi Cui;** Chandima Ekanayake; Hui Ma; Tapan Saha, "An Updated Model to Determine the Life Remaining of Transformer Insulation", IEEE Transactions on Power Delivery, Vol.30, Issue 1, pp. 395-402, January, 2015.

8. Atefeh Dehghani Ashkezari; Hui Ma; Tapan Saha; **Yi Cui,** "Investigation of Feature Selection Techniques for Power Transformer Insulation Condition Assessment," IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 21, Issue 2, pp. 836-844, April, 2014.

9. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Daniel Martin, " Moisture Dependent Thermal Modelling of Transformers Filled With Vegetable Oil," submitted to IEEE Transactions on Power Delivery, 2015.

Peer-reviewed conference publications

10. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Daniel Martin, "Time Domain Diffusion-Driven Dielectric Response Model for Investigation of Moisture Dynamics in Transformers Insulation," IEEE Power & Energy Society General Meeting (PESGM), July 26-30, 2015, Denver, Colorado, USA.

11. **Yi Cui;** Daniel Martin; Tapan Saha; Hui Ma; Chandima Ekanayake; Nick Lelekakis, "Multi-variable Approach for Evaluating Transformer Paper Insulation," International Conference on the Properties and Applications of Dielectric Materials (ICPADM), July 19-22, 2015, Sydney, Australia.

12. **Yi Cui;** Hui Ma; Tapan Saha, "Transformer Hot Spot Temperature Prediction Using a Hybrid of Support Vector Regression and Information Granulation," The 7th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), November 15-18, 2015, Brisbane, Australia.

13. Daniel Martin; Tapan Saha; Olav Krause; **Yi Cui;** Donald McPhail; Tony MacArthur; Dean Condon, "Effect of Rooftop-PV on Power Transformer Insulation and On-Load Tap Changer Operation," The 7th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), November 15-18, 2015, Brisbane, Australia.

14. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake, "A Diffusion-Driven Model for Investigating Moisture Effects on Dielectric Response Measurement of Transformer Insulation", IEEE Power & Energy Society General Meeting (PESGM), July 27-31, 2014, Washington DC, USA.

15. **Yi Cui;** Chandima Ekanayake; Tapan Saha; Peidong Du; Hui Ma, "Finite Element Method Modelled Dielectric Response for Condition Evaluation of Transformer Insulation", IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP) October 20-23, 2013, Shenzhen, China.

16. **Yi Cui;** Hui Ma; Tapan Saha, "Power Transformer Condition Assessment Using Support Vector Machine with Heuristic Optimization", Australasian Universities Power Engineering Conference (AUPEC), September 29-October 3, 2013, Hobart, TAS, Australia.

17. Daniel Martin; **Yi Cui;** Tapan Saha; Nick Lelekakis; Jaury Wijaya, "Life Estimation Techniques for Transformer Insulation", Australasian Universities Power Engineering Conference(AUPEC), September 29-October 3, 2013, Hobart, TAS, Australia.

Publications included in this thesis

1. **Yi Cui;** Hui Ma; Tapan Saha, "Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 1: Framework, Literature, and Illustration", International Transactions on Electrical Energy Systems, Vol.25, Issue 10, pp. 2247-2259, October, 2015.

– incorporated as Chapter 2.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Result interpretation and discussion (80%) Wrote the paper (70%)
Hui Ma	Experimental works (30%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (10%)

2. **Yi Cui;** Hui Ma; Tapan Saha, "Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 2: Implementation, Case Study, and Statistical Analysis", International Transactions on Electrical Energy Systems, Vol.25, Issue 10, pp. 2260 - 2274, October, 2015.

– incorporated as Chapter 2.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Result interpretation and discussion (80%) Wrote the paper (70%)
Hui Ma	Experimental works (30%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (10%)

3. **Yi Cui;** Hui Ma; Tapan Saha, "Improvement of Power Transformer Insulation Diagnosis Using Oil Characteristics Data Preprocessed By SMOTEBoost Technique", IEEE Transactions on Dielectrics and Electrical Insulation, Vol.21, Issue 5, pp. 2363 - 2373, October, 2014.

– incorporated as Chapter 3.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Result interpretation and discussion (80%) Wrote the paper (70%)
Hui Ma	Experimental works (30%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (10%)

4. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Guangning Wu, "Multi-physics Modelling Approach for Investigation of Moisture Dynamics in Power Transformers," Paper accepted for publication in IET Generation Transmission & Distribution (accepted on February 6th, 2016)

– incorporated as Chapter 4.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Result interpretation and discussion (70%) Wrote the paper (70%)
Hui Ma	Experimental works (30%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (10%)
Chandima Ekanayake	Discussion on results (5%)
Guangning Wu	Discussion on results (5%)

5. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Daniel Martin, "Particle Tracing Modelling on Moisture Dynamics of Oil-impregnated Transformer," Paper accepted for publication in IET Science, Measurement & Technology (accepted on January 12, 2016).

– incorporated as Chapter 4.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Result interpretation and discussion (70%) Wrote the paper (70%)
Hui Ma	Experimental works (30%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (10%)
Chandima Ekanayake	Discussion on results (5%)
Daniel Martin	Discussion on results (5%)

6. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake, "Understanding Moisture Dynamics and Its Effect on Dielectric Response of Transformer Insulation," IEEE Transactions on Power Delivery, Vol. 30, Issue 5, pp. 2195- 2204, October, 2015.

– incorporated as Chapter 5.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Result interpretation and discussion (70%) Wrote the paper (70%)
Hui Ma	Experimental works (30%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (10%)
Chandima Ekanayake	Discussion on results (10%)

7. **Yi Cui;** Hui Ma; Tapan Saha; Chandima Ekanayake; Daniel Martin, "Moisture Dependent Thermal Modelling of Transformers Filled With Vegetable Oil," submitted to IEEE Transactions on Power Delivery, 2015.

– incorporated as Chapter 6.

Contributor	**Statement of contribution**
Yi Cui	Experimental works (70%) Developed models (100%) Result interpretation and discussion (70%) Wrote the paper (70%)
Hui Ma	Experimental works (15%) Discussion on results (10%) Wrote and edited paper (20%)
Tapan Saha	Discussion on results (10%) Reviewed the paper (5%)
Chandima Ekanayake	Discussion on models (40%) Discussion on results (5%)
Daniel Martin	Experimental works (15%) Discussion on results (5%) Reviewed the paper (5%)

Contributions by others to the thesis

The transformer used in the ageing experiment was designed by Prof. Tapan Saha, Dr. Chandima Ekanayake and Dr. Hui Ma. Dissolved gas analysis was performed with help from Dr. Raj Jadav and Mr. Kapila Bandara. Moisture and DP measurements were conducted with assistance from Mr. Kapila Bandara.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

I wish to express my sincere acknowledgment and gratitude to those who have provided their various support during my PhD candidature. Even though I can only acknowledge a few here, I will be forever grateful to the others who are too important to be named.

I am greatly thankful to my principal advisor, Professor Tapan Kumar Saha, for his valuable guidance, in-depth technical advice, continuous encouragement and financial support during my PhD journey. I am grateful to him for introducing me to a wonderful research project and sharing his knowledge throughout my thesis. His supervision has made my PhD research a memorable experience.

Special thanks to my associate advisor Dr. Hui Ma for his valuable guidance in the form of technical and conceptual advice towards successful completion of my PhD study. Instead of simply answering my questions, he always encouraged and challenged me to form my own solutions. Thanks to Dr. Ma for the regular supervision and especially the freedom he gave. I am so blessed to have this outstanding advisory team!

I would like to sincerely thank Dr. Chandima Ekanayake for his support and technical advice related to experimental work during the research project. I also want to thank Mr. Steve Wright for his help and support in laboratory related works. I would also like to thank Dr. Daniel Martin, Mr. Shane Goodwin and other team members for their technical support related to experimental works during the project.

My sincere gratitude goes to Professor Simon Bartlett and Dr. Ruifeng Yan for their help, advice during my PhD journey. I will always remember their kind, generous and supportive nature. I would like to express my appreciations to colleagues working on transformer project: Ms. Atefeh Dehghani, Dr. Jeffery Chan, Dr. Mohd Fairouz Mohd Yousof and Mr. Junhyuck Seo, Mr. Kapila Bandara and Mr. Lakshitha Naranpanawe for their help and support during the research project. I also want to thank Ms. Yu Zhao, Mr. Huejie Gu, Mr. Licheng Wang and all other colleagues currently working at Power and Energy Systems Group for their help and support.

My thanks goes to Mrs. Maureen Shield and Mrs. Mandeep Waraich and other staff at the school of ITEE, University of Queensland for their support on financial and administrative matters.

I would like to thank all my family members for their encouragement, understanding and support. I would like to mention my special heartfelt thanks to my parents for their love, affection, motivation and moral support throughout all stage of my life. I would like to dedicate this thesis to my parents.

Finally, I would like to thank Dr. Rahul Sharma and all the anonymous referees for their impartial comments and suggestions on my work during my PhD candidature. Thanks to them for helping and suggesting my work so selflessly.

Keywords

transformer, condition assessment, insulation, dielectric response measurement, ageing, moisture diffusion, multi-physics, hot spot temperature, pattern recognition, data and information fusion

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 090607, Power and Energy Systems Engineering (excl. Renewable Power), 100%

Fields of Research (FoR) Classification

FoR code: 0906, Electrical and Electronic Engineering, 70%
FoR code: 0801, Artificial Intelligence and Image Processing, 30%

Table of Contents

List of Figures

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter VII

List of Tables

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter VI

Chapter VII

List of Abbreviations

FDS	Frequency Domain Spectroscopy
PDC	Polarization and Depolarization Current
FRA	Frequency Response Analysis
PF	Power Factor
2-FAL	2-Furaldehyde
2-FOL	2-Furfuryl Alcohol
5-H2F	5-Hydroxymethyl-2-Furaldehyde
5-M2F	5-Methyl-2-Furaldehyde
2-ACF	2-Acetyl Furan
HPLC	High Performance Liquid Chromatography
OLTC	On-Load Tap Changer
BDV	Breakdown Voltage
DDF	Dielectric Dissipation Factor
DGA	Dissolved Gas Analysis
PD	Partial Discharge
ANN	Artificial Neural Network
NN	Neural Network
BP	Back Propagation
RMP	Reduced Multivariate Polynomial
FLVQ	Fuzzy Learning Vector Quantization
SOM	Self-Organizing Map
FCM	Fuzzy *c*-Means
PSO	Particle Swarm Optimization
GP	Genetic Programming
MLP	Multiple Layer Perceptron
PNN	Probabilistic Neural Network
LVQ	Learning Vector Quantization
LVQ1	Type I Learning Vector Quantization
LVQ2	Type II Learning Vector Quantization
***k*NN**	*k*-Nearest Neighbor
EM	Expectation Maximization
SVM	Support Vector Machine
SVM_{lin}	Support Vector Machine with Linear kernel

SVM_{poly}	Support Vector Machine with Polynomial kernel
SVM_{gauss}	Support Vector Machine with Gauss kernel
IDA	Intelligent Decision Tree Algorithm
FSVM	Fuzzy Support Vector Machine
RBF	Radial Basis Function
GRNN	Generalized Regression Neural Network
SOTA	Self-Organizing Tree Algorithm
H_2	Hydrogen
CH_4	Methane
C_2H_6	Ethane
C_2H_4	Ethylene
C_2H_2	Acetylene
CD	Critical Difference
T1/T2	Thermal Faults Lower Than 700 Degree
T3	Thermal Faults Higher Than 700 Degree
AI	Artificial Intelligence
SMOTE	Synthetic Minority Over-sampling Technique
DP	Degree of Polymerization
CO	Carbon Monoxide
CO_2	Carbon Dioxide
HI	Health Index
DGAF	Dissolved Gas Analysis Factor
OQF	Oil Quality Factor
PIF	Paper Insulation Factor
PDIV	Partial Discharge Inception Voltage
3D	Three Dimensional
CFD	Computational Fluid Dynamics
DDP	Diamond Dotted Paper
MP	Mouldable Pressboard
SAR	Spacers All Round
2D	Two Dimensional
HV	High Voltage
LV	Low Voltage
FEM	Finite Element Modelling
DEM	Discrete Element Method

KFT	Karl Fischer Titration
HST	Hot Spot Temperature
SCADA	Supervisory Control And Data Acquisition
MDTM	Moisture Dependent Thermal Model
MSE	Mean Squared Error
BN	Bayesian Network
DAG	Directed Acyclic Graph
CPT	Conditional Probability Table
FSV	Fault State Vectors
SSV	Symptom State Vectors
SSES	Symptom State Evidence Sets
RM	Reliability Metric

List of Symbols

I_R	Resistive current
I_c	Capacitive current
$\tan\delta$	Dielectric dissipation factor
FAL	Furan concentrations
D	Dielectric displacement
E	Applied electric field
P	Vectors of electric field polarization
χ	Dielectric susceptibility
$g(t)$	Step function
$f(t)$	Dielectric response function
$\varepsilon^{'}$	Real part of complex permittivity
$\varepsilon^{''}$	Imaginary part of complex permittivity
ε_{oil}	Oil relative permittivity
ε_{spacer}	Relative permittivity of the spacers
$\varepsilon_{barrier}$	Relative permittivity of the barriers
$N^{'}_{total}$	Number of correctly classified samples in testing sub-dataset
N_{total}	Sample number in testing sub-dataset
x_i	*i-th* input
x_i^j	*j-th* element in x_i
w_i^{mj}	Weight of the *m-th* hidden node with respect to x_i^j
b_i^m	Bias
M	Hidden nodes number
w_i^{km}	Weight between the output node and the hidden node
K	Classes number
d	Dimension of the input
σ	Variance of Gaussian distribution
y_{ij}	Pattern node
C_i	Class *i*
N_i	Samples number in classes C_i

z	Predicted class
m	Classes number
m_k	Prototype vector
m_c	Winner unit
$\|.\|$	Euclidian distance operator
$\alpha\ (t)$	Learning rate
K_l	Frequency of the *l*-th class
$p(c_k)$	Class prior probability
$p(x)$	Normalization factor
θ	Number of Gaussian mixture elements
$p(x_i \mid j)$	Probability distribution of Gaussian mixture elements
$P(j)$	Weights
α_k	Lagrange multiplier
C	Penalty parameter
$K\left(\mathrm{x}_k, \mathrm{x}_j\right)$	Kernel function
γ	Variance
b	Bias term
d_i^{cen}	Euclidean distance
β	Decay steepness
φ_m	Basis functions
w^{km}	Output layer weights
C_k	Basis function
T	Number of classes
D_t	*t-th* dataset
r_j^i	Rank of the *j*-th algorithm on the *i*-th dataset
R_j	Averaged ranks for *j*-th algorithm
F_F	Friedman statistic
k	Number of algorithms to be compared
N	Number of datasets to be compared
S_i	Score of each DGA factor
W_i	Weight of each DGA factor

S_j	Score of each oil quality factor
W_j	Weight of each oil quality factor
$\mathbf{X}$	Sample instance in the minority class
$d(\mathbf{X},\mathbf{x}_2)$	Euclidean distance between $\mathbf{X}$ and $\mathbf{x}_2$
m_i	*i*-th synthesized new samples
m	Number of minority classes
$n_1,...,n_t,...,n_m$	Sample numbers in each majority class
n_{tr}	Desired sample numbers of each majority class
c_i	Centre of the class
$D_1, D_2, \cdots D_k$	*K* mutually exclusive subsets split from original dataset
$C(x,t)$	Moisture concentration of cellulose insulation at position *x* and time *t*
D	Diffusion coefficient
$\mathbf{A}$	Magnetic vector potential
λ_x , λ_y , λ_z	Permeability of material in three directions
σ_{dc}	Conductivity of copper conductors in winding
$\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$	Unit vectors
J_{sc}	Density of the current source
$\sigma(T)$	Conductivity at temperature *T*
$\sigma(T_0)$	Conductivity at reference temperature T_0
T_0	Reference temperature 293*K*
α_T	Temperature coefficient
V_{wdg-p}	Applied voltage on primary winding
V_{ind-p}	Induced voltage of primary winding
N_p	Turns number of the primary winding
A_w	Total cross-sectional area of the winding domain
R_{wdg-p}	Resistance of the primary winding
I^2R_{loss}	Joule loss
T	Temperature
k_x, k_y, k_z	Thermal conductivity of the medium in three directions
Q_{sc}	Total heat source

ρ_m	Mass density
c_m	Heat capacity
ϕ_{conv}	Heat flux of conductor's surface by convection
ϕ_{rad}	Heat flux of conductor's surface by radiation
h	Convection coefficient
ς	Surface emissivity
k_B	Stefan–Boltzmann constant
φ	Absorptivity
ξ	Irradiation
T_{amb}	Ambient temperature
N_u	Nusselt number
k_F	Fluid thermal conductivity
l	Geometry length
R_e	Rayleigh number
d_p	Distance between the plates
R_e^*	Rayleigh number for the symmetric isoflux
d_o	Width of oil ducts
$\phi_{rad,i}$	Net radiative heat transfer rate of *i*-th surface
V_{ij}	View factor from *i*-th surface to *j*-th surface
R_i	Radiosity of *i*-th surface
E_{bi}	Emissive power for *i*-th surface
ρ_F	Fluid density
$\mathbf{v}$	Fluid velocity
P	Pressure
g	Gravitational acceleration
μ	Dynamic viscosity
$\mathbf{F}$	Strain rate tensor of Newtonian fluids
$\mathbf{I}$	Unit tensor
$\mu_{\bar{\varepsilon}}$	Linearized strain rate tensor
E_a	Activation energy

ρ_{r0}	Resistivity at the reference temperature
α_t	Coefficient between resistivity and temperature
R_s	Resistor of secondary winding in transformer's equivalent circuit
$Q_s(t)$	Losses of different components in transformer at time instance *t*
Ω	Domain of each component
$q(x,y,z,t)$	Loss density of infinitesimal volume in each component
Q_{ST}	Average losses during one cycle
C_0	Moisture concentration of cellulose surface
P_v	Vapour pressure of water
V_c	Total bulk volume of the cellulose insulation structure
Ω_c	Cellulose domain
$c_{average}$	Average moisture concentration of the bulk cellulose insulation
$W(x,t)$	Moisture content of the cellulose insulation
m_p	Mass of the particle
$\mathbf{v}_p$	Particle's velocity vector
$\sum \mathbf{F}_p$	Collective effect of different forces imposed on the particle
$\mathbf{F}_g$	Gravity force
$\boldsymbol{g}$	Gravity vector
ρ_p	Density of particles
ρ	Density of the surrounding fluid
$\mathbf{F_D}$	Drag force due to the fluid flow
τ_p	Velocity response time constant
$\boldsymbol{v}_F$	Velocity vector of the fluid flow
d_p	Particle's diameter
Δt	Time step taken by the solver
r_p	Particle's radius
T_F	Fluid temperature
μ_F	Fluid viscosity
$\mathbf{F_b}$	Brownian force

ζ	Random number following normal distribution with zero mean and unit standard variation
$P(x,t)$	Particle distribution density
D_t	Interaction between spherical water particles and the cellulose medium
$\mathbf{F_V}$	Interactions between fluid flow and particles
δ	Delta function
$\mathbf{q}_i$	*i*-th particle position vector
$U(r)$	Intermolecular potential
ε	Interaction strength
r_c	Collision diameter
d	Distance between the particles
$\boldsymbol{F}_{Li}$	Lennard-Jones force
$w_i(x,t)$	Moisture concentration at the *i-th* layer of the pressboard
$P_t(\alpha)$	Transmission probability
$w_{ref}(t)$	Moisture concentration of the reference plane
ε_0	Permittivity of vacuum
ε_∞	High frequency relative permittivity of the material
$f(t)$	Response function of the material
$\hat{\chi}(\omega)$	Complex susceptibility
$\hat{j}(\omega)$	Total current density in the dielectric material
$Z(\omega)$	Impedance of the whole piece of the pressboard
$Z_i(\omega)$	Impedance of the *i*-th layer of the pressboard
$G_i(\omega)$	Admittance of the *i*-th layer of the pressboard
Δx	Thickness of each sliced layer of the pressboard
$\varepsilon_i^{'}$	Real part of complex permittivity of the *i*-th layer of pressboard
δ_i	Summation of conductive and polarization losses of the *i*-th layer of pressboard
$C_i^*(\omega)$	Complex capacitance of the pressboard
ω	Angular frequency
W_0	Moisture content in the whole bulk of un-aged pressboards

W_s	Concentration of moisture at the interface between pressboard and oil
$W(x, y, z)$	Moisture concentration at a particular location of the cellulose
$\alpha_1, \alpha_2, \beta_1, \beta_2$	Free parameters
ε_{rs}	Dielectric constant at zero excitation frequency or dc value
$\varepsilon_{r\infty}$	Dielectric constant at very high frequency
τ_h	Pre exponential factor
H	Activation energy
P_i	Total measurement duration at the *i*-th frequency
T_{ave}	Average temperature at a particular frequency
$\Delta\theta_{oil,I}$	Initial top oil temperature rise over ambient temperature
$\Delta\theta_{oil,U}$	Ultimate top oil temperature rise over ambient
$\Delta\theta_{hs,I}$	Initial hot spot temperature rise over top oil temperature
$\Delta\theta_{hs,U}$	Ultimate hot spot temperature rise over top oil temperature
q_{Fe}	Loss of transformer during no load test
q_{Cu}	Load loss of transformer
θ_{amb}	Ambient temperature
θ_{oil}	Top oil temperature
C_{insul}	Lumped thermal capacitance of cellulose
C_{oil}	Lumped thermal capacitance of oil
R_{insul}	Nonlinear thermal resistance of cellulose
R_{oil}	Nonlinear thermal resistance of oil convection
h	Convection heat transfer coefficient
A_{oil}, A_{insul}	Area of oil and cellulose
C_1 ,C_2	Constant
μ	Oil viscosity

n, m	Constant
l	Thickness of cellulose
k_{insul}	Thermal conductivity of cellulose
θ_{insul}	Cellulose temperature
Δw	Difference in water content
$\Delta\theta_{insul}$	Difference in cellulose temperature
w_0	Reference water content of cellulose
θ_0	Reference temperature of cellulose
w_{insul}	Water content of cellulose
P_v	Vapor pressure
A_w	Water activity in oil
θ_{hs}	Hot spot temperature
$R_{oil,R}$	Rated thermal resistance of oil
$R_{insul,R}$	Rated thermal resistance of cellulose
λ_R	Ratio between $R_{insul,R}$ and $R_{oil,R}$
μ_R	Oil viscosity at rated temperature
μ_{pu}	Ratio of oil viscosity between any temperature and rated top oil temperature
k_{pu}	Ratio of cellulose thermal resistance between any temperature and rated temperature
$\theta_{oil,R}$	Top oil temperature under rated load condition
$\theta_{insul,R}$	Cellulose temperature under rated load condition
$\theta_{hs,R}$	Hot spot temperature under rated load condition
$\Delta\theta_{oil}$	Top oil temperature rise over ambient temperature

$\Delta\theta_{oil,R}$	Rated top oil temperature rise over ambient temperature
$\Delta\theta_{hs}$	Hot spot temperature rise over top oil temperature
$\Delta\theta_{hs,R}$	Rated hot spot temperature rise over top oil temperature
$\tau_{oil,R}$	Oil time constant
R	Ratio between q_{Cu} and q_{Fe}
K	Ratio between load current and rated winding current
I	Load current of transformer winding
I_R	Rated current of transformer winding
$\tau_{wnd,R}$	Winding time constant
$P_{Cu,pu}$	Per unit value of copper loss
$P_{Cu,DC,pu}$	Per unit value of DC loss
$P_{Cu,Eddy,pu}$	Per unit value of eddy loss
$\hat{\theta}_i$	Estimated hot spot temperature at time instance i
θ_i	Measured hot spot temperature at time instance i
$\bar{\theta}$	Mean value of measured hot spot temperature
G	Sample size of hot spot temperature in comparison
X_1,X_2,X_3	Parent nodes of Bayesian network
X_4	Child node of Bayesian network
$P(X_1),P(X_2),P(X_3)$	prior probabilities of parent nodes
$P(X_4\mid X_1,X_2,X_3)$	conditional probabilities of the child node
n_{CPT}	number of elements for a particular child node
S_c	number of child node states
n_p	number of parent nodes associated with this child node

s_{pi}	states number of *i*-th parent node
P_0	leaky probability
P_i	conditional probability of the i-th parent node
μ_i	mean value
σ_i	variance
RM_j	reliability metric value for *j*-th symptom state evidence set

Chapter 1

Introduction

1.1 Background

Power transformer is one of the most crucial and expensive equipment in a power system. Its reliable operation can directly affect the security of the system [1]. During transformer's operation, its oil-cellulose insulation system can fail when the transformer is subjected to combined electrical, mechanical and thermal stresses. Such failures are sometimes catastrophic and may result in irreversible damages to the system [2]. To ensure a reliable operation of power system, power transformers' health condition must be continuously monitored and assessed for appropriate operation and maintenance decisions.

Generally, there are three main aspects which should be closely monitored to determine transformer's condition. These include oil characteristics and dissolved gases in transformer, moisture concentration of transformer's oil-cellulose insulation and hot spot temperature of transformer windings. Over the last several decades, various offline and online diagnostic techniques have been proposed to assess transformer's condition [3-7]. For example, oil characteristics test and dissolved gas analysis (DGA) are considered as the most established chemical technique to identify incipient faults in transformers [8-11]. Polarization based measurements including frequency domain spectroscopy (FDS) and time domain polarization and depolarization current (PDC) measurement are applied to estimate the moisture concentration of transformer's oil-cellulose system and evaluating its ageing condition [5, 12-17]. Thermal dynamic modelling has been proposed and utilized to estimate the hot spot temperature of transformer windings [18-19]. Frequency response analysis (FRA) is adopted to detect winding deformation and displacement in a transformer [20]. Partial discharge (PD) measurement is developed as a means of online/offline diagnostic technique to monitor transformer's insulation system [21]. A summary of different diagnostic techniques for transformer condition assessment is shown in Figure 1.1.

Nowadays, considerable efforts have been made for an effective and reliable transformer condition monitoring and assessment. Major research works are dedicated in developing transformer models and onsite transformer fault diagnosis systems. However, these approaches are usually based upon empirical models, which are sometimes inaccurate and incomplete especially in abnormal operation scenarios. The major drawbacks are rooted in lack of confidence in oil characteristics and dissolved gas analysis, intricate difficulties in moisture estimation and the inaccuracy in thermal dynamic models. In recent years, with the use of computational intelligence techniques, it becomes possible to analyze a large volume of data and get more in-depth understanding on the correlations between measurement data and transformer fault symptoms. This

is of great benefit for transformer condition assessment.

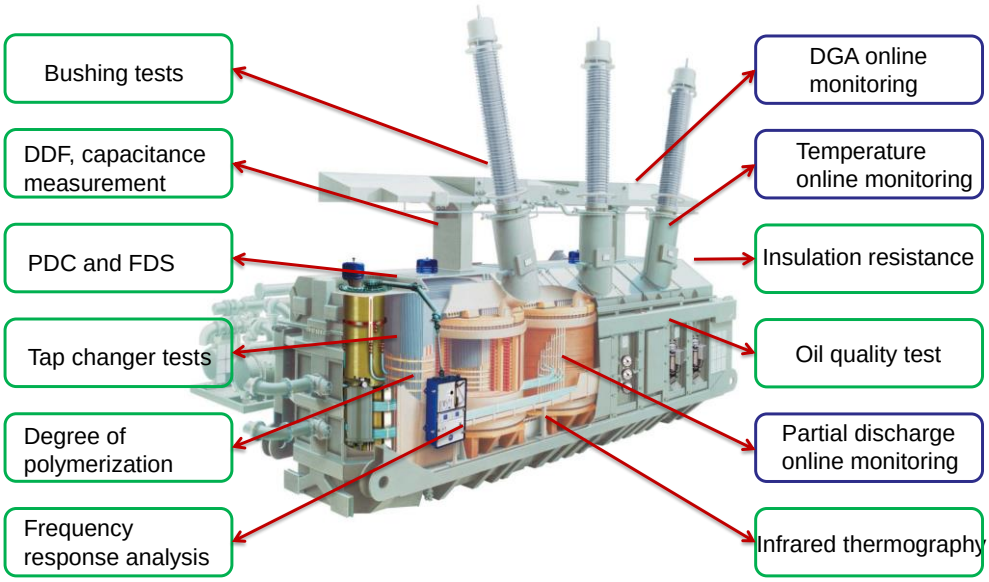


Figure 1.1 Typical transformer diagnostic techniques

This thesis is aimed at developing a set of advanced data centric diagnostic approaches for accurate modelling and reliable assessment of transformer's health condition. It presents three novel data centric AI based diagnostic approaches to deal with practical problems in transformer condition monitoring and assessment. These techniques include the hybrid Synthetic Minority Oversampling Technique and Boosting (SMOTEBoost) algorithm, multi-physics modelling approach and moisture dependent thermal modelling method, which are dedicated in improving DGA interpretations, evaluating moisture dynamics in oil-cellulose insulation and estimating hot spot temperature of transformer windings respectively. It also proposes a data and information fusion framework to integrate every piece of data and information obtained from different transformer diagnostic measurements and subsequently evaluating the overall health condition of a transformer.

1.2 Fundamentals of Transformer Condition Assessment

For an in-service power transformer, its insulation system is usually composited by oil and cellulose insulation material. Specifically, the cellulose insulation contains oil-impregnated papers and a number of pressboard barriers which are shown in Figure 1.2 and Figure 1.3. Oil-impregnated papers are normally wrapped on the surface of the conductors grouped in transformer windings and pressboard barriers are usually placed between the primary and secondary windings. Certain type of

fluid, such as mineral oil or natural ester, are then filled into power transformers. These liquids play a role as both dielectric insulating media and heat dissipation agent [22]. The most widely used insulating liquid for power transformers is mineral oil, though recently biodegradable oil appears as a promising alternate for liquid insulation medium [23-24]. During operation of a transformer, its oil-cellulose insulation degrades over time [2]. In the following section, fundamentals of condition assessment of transformer oil-cellulose insulation system are discussed and different conventional diagnostic techniques are reviewed.

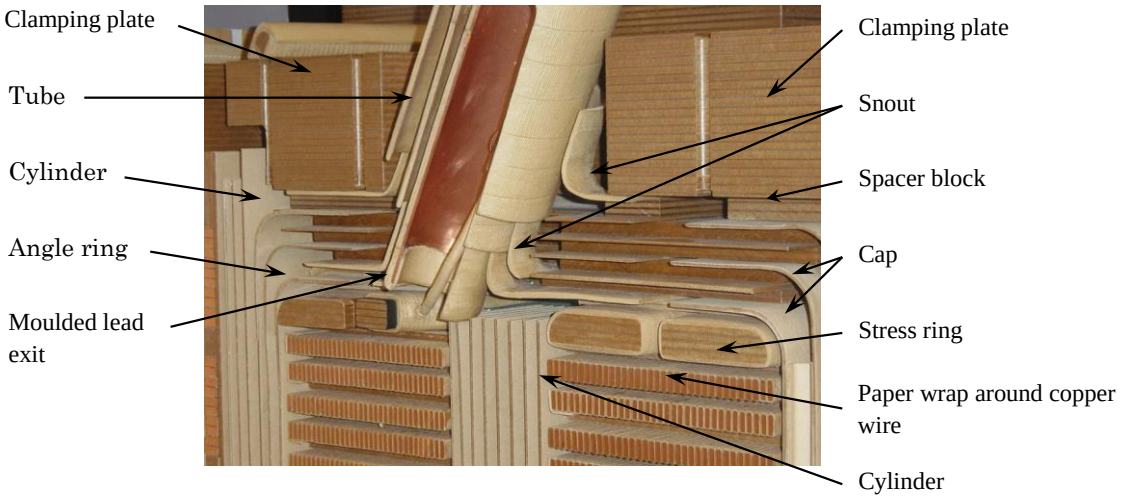


Figure 1.2 Cross section area of end insulation (400 kV transformer, 220 kV-side)[25]

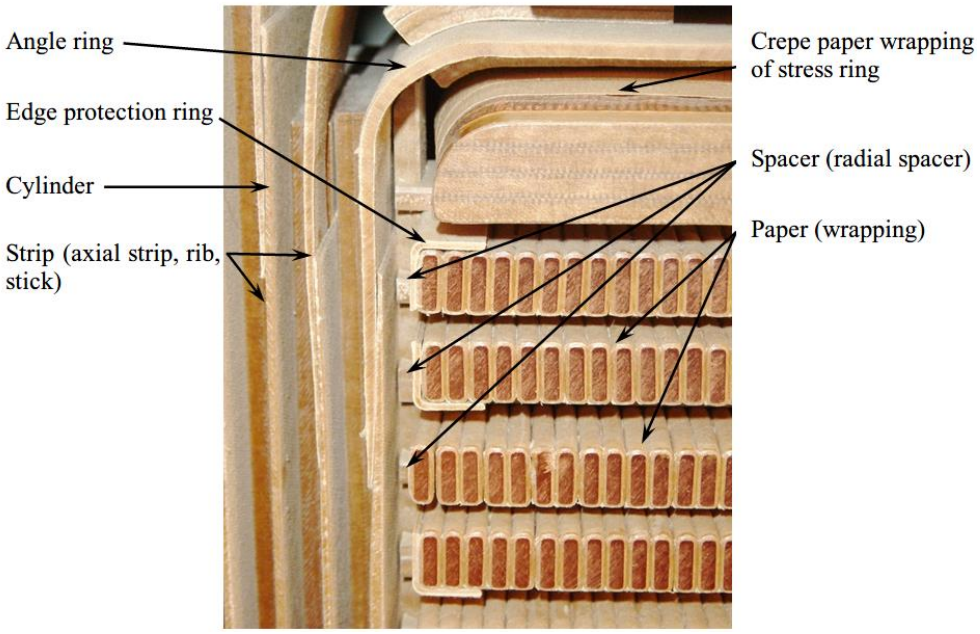


Figure 1.3 Cross section area of end insulation (detail)[25]

1.2.1 Oil characteristics test and dissolved gas analysis (DGA)

Table 1.1 summarizes the industry standards and related publications regarding standardized procedures of transformer oil characteristics tests. These tests are grouped into routine test, complementary test and special investigative test [26]. With the reference to the standards, this section focuses on dissolved gas measurements and a number of conventional oil characteristics tests, covering moisture content, acidity, dielectric dissipation factor, resistivity, breakdown voltage and 2-Furfural. By integrating the diagnostic results from both DGA measurements and oil characteristics tests, a comprehensive evaluation of the condition of transformer insulation system can be achieved[27-28].

Table 1.1 Summary of transformer oil characteristics test standards **[29]**

Test group	**Test name**	**Test standard**
Routine tests	Color and appearance	ISO 2049
	Breakdown voltage	IEC 60156
	Water content	IEC 60814
	Acidity (neutralization value)	IEC 62021
	Dielectric dissipation factor (DDF) or resistivity	IEC 60247
	Inhibitor content	IEC 60666
	Sediment and sludge	IEC 61125 method C
Complementary tests	Interfacial tension	ISO 6295
	Particles (particle count)	IEC 6097
Special investigative tests	Oxidation stability	IEC 61125
	Flash point	ISO 2719
	Compatibility	IEC 61125
	Pour point	ISO 3016
	Density	ISO 3675
	Viscosity	ISO 3104
	Polychlorinated biphenyls (PCBs)	IEC61619
	Corrosive sulphur	DIN 51353

1.2.1.1 Moisture content

Moisture in transformer may come either from the environment or be generated from the deterioration of insulating materials [30-31]. Figure 1.4 depicts the degradation mechanism of cellulose. From Figure 1.4 it can be seen that during thermal degradation of cellulose, carbon oxides

and moisture are the dominant by-products [32]. There are three main states of moisture existing in transformer oil. Large proportion of moisture is found to be dissolved in oil. Moisture can also exist in the oil tightly bound to oil molecules. Free water can be formed as droplet when moisture concentration exceeds the saturation level [33-34] at certain temperature. Moisture in cellulose insulation may also exist in three states: adsorbed to the cellulose surface, as free water in capillaries, and as imbibed free water [34].

Moisture in oil can be measured by chemical or electrical methods. Typical moisture measurement techniques include Karl Fischer Titration (KFT) [35], cellulose adsorption isotherms approach (also called equilibrium charts) [36-37] and dielectric response method [5, 12-17].

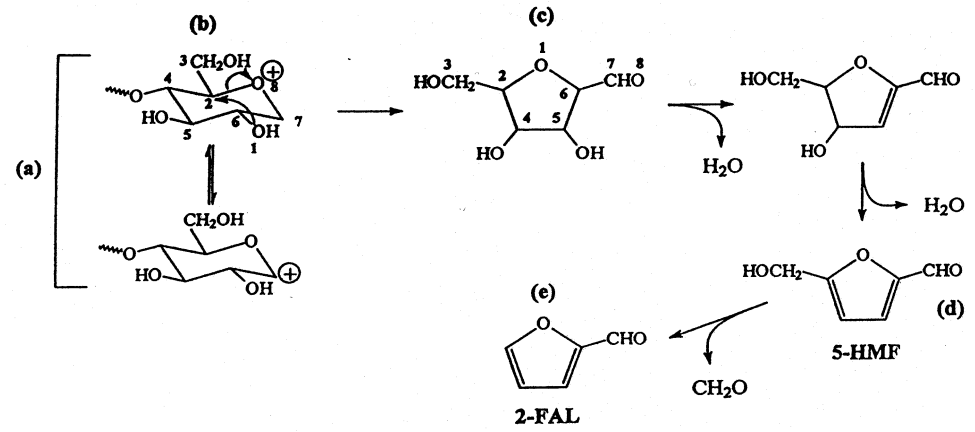


Figure 1.4 Mechanism of cellulose degradation [38]

1.2.1.2 Acidity

The hydrolysis of paper produces acids. The mechanisms have been well understood from the studies of laboratorial experiments [39-41]. Oxidation also produces acids. There may be free acids formed (e.g., small carboxylic acids) and acidic groups may also attach to the cellulose molecular chain[42]. In addition, acids can also be produced from atmospheric contamination [43].

The acids generated from paper and from oil may contain the same compounds[44]. Therefore, acidic oxidation products from the oil may also influence the degradation rate of paper. Most acids have a low molecular weight thus being volatile and having a low boiling point. They will also have a high polarity, meaning that they will be hydrophilic [42], tending to be dissolved well in paper. The total acidity of the insulation oil can be quantified by Electro-chemical technique, which measures the pH value of the solution at room temperature against the volume of the alcoholic potassium hydroxide that is used for the neutralization of the acids [39].

1.2.1.3 Dielectric dissipation factor

Oil-cellulose insulation is normally regarded as capacitive medium and it can be modelled as a simplified parallel circuit consisting capacitor and resistor (shown in Figure 1.5). If oil-cellulose insulation is imposed by an AC voltage, leakage current will be generated and flows through the insulating medium. The leakage current consists of two components, i.e., resistive current (I_R) and capacitive current (I_c). According to ASTM D 924 [45], the dielectric dissipation factor (DDF), which is also known as tanδ, can be computed as a dimensionless ratio of I_R to I_c in (1.1).

$$DDF = \tan\delta = \frac{I_R(f)}{I_c(f)} \tag{1.1}$$

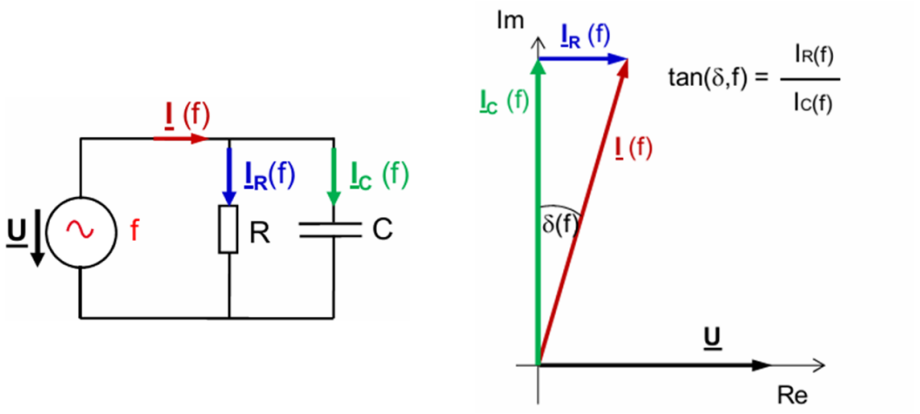


Figure 1.5 Vector diagram of dielectric dissipation factor subjected to AC voltage

Dissipation factor (tanδ) can be also derived through more widely used terminology of power factor (PF), which is expressed by (1.2)-(1.3).

$$\tan\delta = \frac{PF}{\sqrt{1-PF^2}} \tag{1.2}$$

$$PF = \frac{\tan\delta}{\sqrt{1+(\tan\delta)^2}} \tag{1.3}$$

As the dielectric dissipation factor is commonly expressed in the formation of decimal or percentage (%), if decimal value of DDF is below 0.05, tanδ and power factor values can be treated as approximately equivalent to each other. For most cases, these two terms may be considered

interchangeable [45] because the tanδ of insulating oil in good condition has decimal values less than 0.005. A low tanδ shows low AC dielectric losses of the insulating materials and it has been accepted as a reliable method to indicate variation in oil quality caused by deterioration and contamination in transformer operation [46].

1.2.1.4 Resistivity of oil

Resistivity is considered as one of the most sensitive properties of transformer oil. Oil resistivity is highly dependent on temperature. It would be desirable to maintain a high resistivity of oil through oil reclamation. The resistivity of oil can decrease considerably in the presence of solid contaminants, acids, moisture, etc[47]. A high resistivity of oil indicates low density of free ions or ion-forming contaminants and it also implies low concentration of conductive particles in oil [48].

1.2.1.5 Breakdown voltage

The dielectric strength of oil is considered to be the most significant parameter when choosing appropriate insulating liquid for oil-filled power transformers. The basic criterion of selecting oil is to assure a reliable dielectric function and a satisfied impregnation with different insulating components in transformers, e.g., cellulose paper/pressboard of winding insulation. A good impregnation between oil and solid insulation can eliminate the formation of air bubble and the dissolution of other gases, thus can avoid initiation of partial discharges and maintain a high dielectric strength [49-50].

The breakdown voltage (BDV) measures the capability of insulating oil to withstand electric stress without failure. BDV has a high sensitivity to the oil quality, and it is significantly affected by the presence of conductive and non-conductive particles, moisture and other dissolved gases or emulsions. BDV is commonly measured to compare the quality of different oils given the contamination in oil has been well controlled.

1.2.1.6 Furfural test

Transformer’s life expectancy is highly dependent on the condition of its cellulose insulation. Cellulose structure usually contains a number of glucose units which are chemically linked in a cellulose molecule (Figure 1.6). The average number of glucose units of cellulose is usually measured as the degree of polymerization (DP) value. During the ageing process, the cellulose will

degrade through oxidation, hydrolysis and pyrolysis, thus the number of glucose units in cellulose molecules will be reduced, which causes a reduction in mechanical strength of cellulose insulation. It is found both DP and tensile strength decreases exponentially during ageing at constant temperature [51]. The reduction in mechanical strength of cellulose will weaken transformer's capability to withstand stress caused by short-circuit faults. It has been accepted that a DP value less than 200 indicates the insulation material almost reaches its end of life [52].

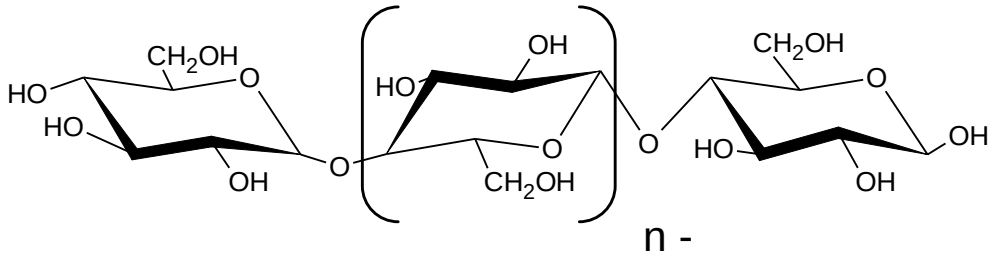


Figure 1.6 Molecular structure of cellulose

To evaluate DP of cellulose material in power transformers, furans measurement has been developed as a convenient method which can obtain the information on the degradation of solid insulating material. "Furans" contains a bunch of chemical composites which are the by-products from cellulose degradation. The most commonly measured furans compounds include 2-furaldehyde (2-FAL), 2-furfuryl alcohol (2-FOL), 5-hydroxymethyl-2-furaldehyde (5-H2F), 5-methyl-2-furaldehyde (5-M2F), and 2-acetyl furan (2-ACF) [53]. Similar to moisture, majority furans compounds resident in the pressboard/paper and only a minor proportion exists in oil [54]. To measure the furans concentration, High Performance Liquid Chromatography (HPLC) is usually adopted [55]. The condition of cellulose can be evaluated by measuring furan concentration in oil only when the dependency between furan content in the oil and DP is well established. Over the last several decades, a number of models have been proposed to explore the correlation between furan concentration and DP value of cellulose [51, 56-60]. The most commonly used models include Heisler's model [59], Chen's model [58] and DePablo model[51]. It should be notable that all the furan concentrations (FAL) in the above three models are measured in unit of *ug/g*.

Heisler and Banzer's model [59]

$$DP = 325 \times \left(\frac{19}{13} - \log_{10} 2FAL \right) \tag{1.4}$$

where $100 \le DP \le 900$.

Chen's model [58]

$$DP = \frac{1}{0.0035} \times \left(1.51 - \log_{10} FAL\right) \tag{1.5}$$

where $150 \le DP \le 1000$.

DePablo's model [51]

$$DP = \frac{1850}{2FAL + 2.3} \tag{1.6}$$

where $150 \le DP \le 600$.

The furans in the oil can be affected by various factors, including the degradation of cellulose, moisture and types of cellulose materials [61-62], oxygen levels, acidity of oil and transformer design, etc. Each furanic compound has different ratios between furans in paper and furans in oil (shown in Figure 1.7) and they are all additionally dependent on temperature [54]. Maintenance and replacement activities, including oil reclamations, will directly affect oil properties and, therefore, change furan content in it. This may impose some challenges on furan analysis for evaluating the ageing condition of cellulose in practice.

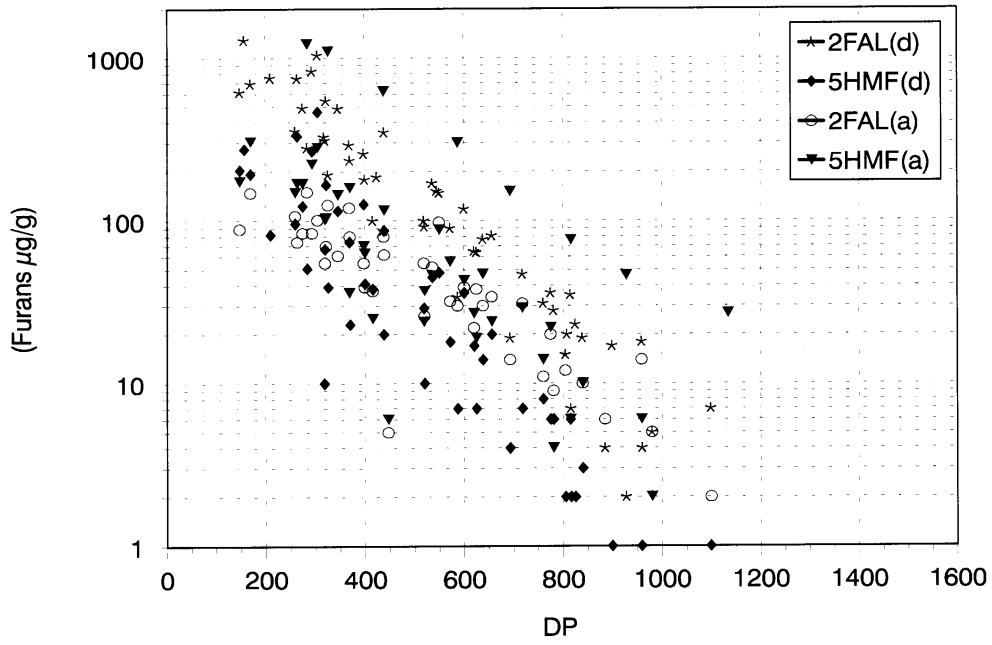


Figure 1.7 Correlation between different furanic compounds in oil and DP [51]

1.2.1.7 Dissolved gas analysis (DGA)

For an in-service oil-impregnated power transformer, insulation degradation will slowly produces certain amount of gases, dissolved in transformer oil under normal operations. The dissolved gases

can be formed at a higher rate than normal rate when certain faults occur in a transformer [10, 63].

The DGA measurement can reveal the incipient faults in transformers caused by electrical and thermal stresses. Dissolved gas measurement and its interpretation schemes are accepted as a convenient and reliable method to facilitate monitoring operation conditions of power transformers. A number of standards from different committees and organizations, such as IEC60559 [11], IEEE C57.104-2008 [10], and CIGRE TF 15.01.01 [22] have developed for DGA interpretation.

1.2.1.7.1 Key gas method

Key gas method (shown in Figure 1.8) measures the individual concentrations of the six dissolved gases (i.e., hydrogen-H_2, methane-CH_4, ethane-C_2H_6, ethylene-C_2H_4, acetylene-C_2H_2 and carbon monoxide-CO). Four types of incipient faulty conditions including overheating in oil and cellulose, partial discharge in oil and arcing can be diagnosed by computing the relative percentage of each gas.

The interpretation scheme of key gas method is originated from the practical experience of various experts. Though key gas method is easy to be implemented, it is normally not considered as a reliable method [64]. Some investigations reported that only 42% of DGA samples in a dissolved gases measurement dataset of inspected transformers can be correctly identified by key gas method [65].

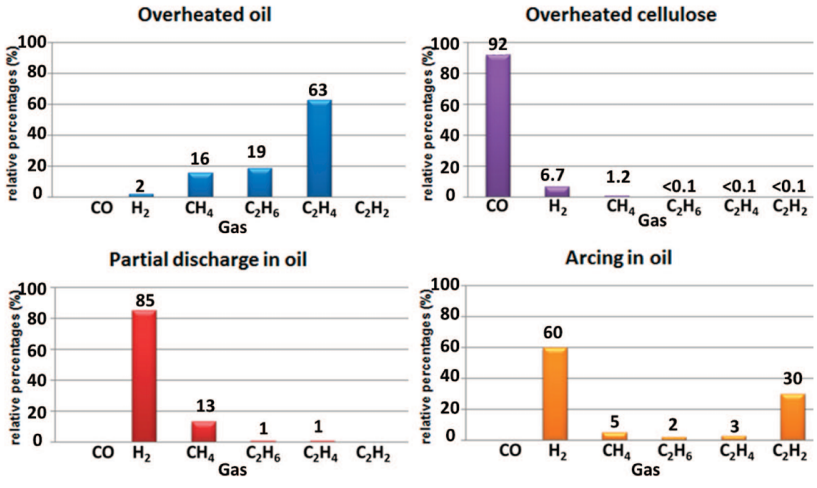


Figure 1.8 Key gases method chart [66]

1.2.1.7.2 Doernenburg ratio method

This method computes four types of gas ratios (C_2H_2/C_2H_4, C_2H_6/C_2H_2, C_2H_2/CH_4, and CH_4/H_2) among five gas concentrations to detect the faults of transformers [10, 67]. The fault types are

determined according to the predefined ranges of these four ratios, which are shown in Table 1.2 [10]. It should be notable that this method can be applied only if the concentration of minimal one key gas is above twice the threshold concentrations (shown in Table 1.3).

Table 1.2 Doernenburg ratio method for faults diagnosis

Fault identification	**CH_4/H_2**	**C_2H_2/C_2H_4**	**C_2H_2/CH_4**	**C_2H_6/C_2H_2**
Thermal decomposition	>1	<0.75	<0.3	>0.4
Partial discharge	<0.1	Not significant	<0.3	>0.4
Arcing	>0.1 to <1	>0.75	>0.3	<0.4

Table 1.3 Threshold of gas concentration in the Doernenburg ratio method

Dissolved gas	**Minimum concentrations (ppm)**
H_2	100
CH_4	120
CO	350
C_2H_2	35
C_2H_4	50
C_2H_6	65

1.2.1.7.3 Rogers ratio method

Rogers ratio method adopts similar interpretation of the aforementioned Doernenburg ratio method. However, compared to the Doernenburg ratio method which requires the minimal concentrations of the gases, Rogers ratio method can be implemented even when the concentrations do not exceed the threshold values in Table 1.3 [68].

The original Rogers ratio measures four gas ratios, i.e., C_2H_4/C_2H_6, CH_4/H_2, C_2H_2/C_2H_4, C_2H_6/CH_4 and it can recognize 12 incipient faults [69]. However, latter studies showed the ratio of C_2H_6/CH_4 appeared to not have a close relationship with the faults [64, 67] and thus it was removed in the updated IEEE Standard C57.104-1991[68]. Consequently, the original 12 faults were reduced to six, which included the normal health condition (Table 1.4 [10]). Similar to key gas method, some inconsistencies were found in Rogers ratio interpretation and the overall diagnostic accuracy of 58.9% was reported [70].

Table 1.4 Rogers ratio method for interpreting transformer faults

Fault identification	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
Unit normal	<0.1	>0.1 to <1	<1
Partial discharge	<0.1	<0.1	<1
Arcing	0.1 to 3.0	0.1to 1	>3
Low temperature	<0.1	>0.1to <1	1 to 3
Thermal <700 ℃	<0.1	>1	1 to 3
Thermal > 700 ℃	<0.1	>1	> 3

1.2.1.7.4 IEC ratio method

IEC ratio employs the same three gases ratios as in Rogers ratio method. However, the ranges of these three ratios and corresponding fault types are different. IEC codes and its interpretation are summarized in Table 1.5 [71]. In IEC ratio method, a new gas ratio C_2H_2/H_2 is introduced to identify faults caused by contamination from on-load tap changer (OLTC) [72]. Another enhancement made in IEC ratio is the use of graphical representation of ratio ranges. This would help to achieve a more reliable diagnose and identify the faults when the corresponding gas ratios are out of the predefined ranges [73].

Table 1.5 IEC codes and associated fault types

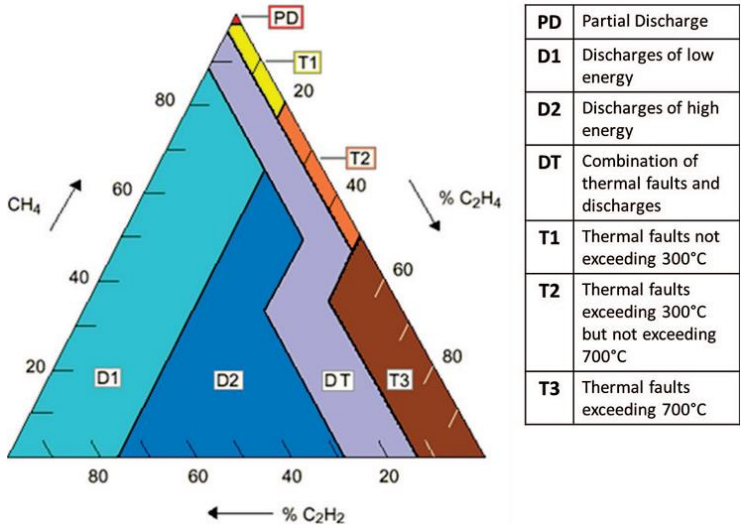
Fault identification	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
Partial discharge	NS[a]	<0.1	<0.2
Discharge of low energy	>1	0.1-0.5	>1
Discharge of high energy	0.6-2.5	0.1-1	>2
Thermal < 300 ℃	NS	>1 but NS	<1
Thermal: 300 ℃ < T < 700 ℃	<0.1	>1	1-4
Thermal >700 ℃	<0.2	>1	>4

a. Non-significant, whatever the value

1.2.1.7.5 Duval triangle

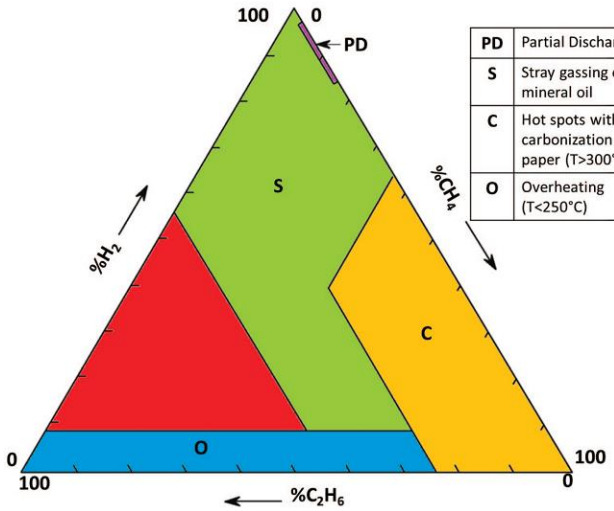
Duval triangle method was developed by integrating the aforementioned IEC ratio and IEC TC10 databases [71]. It adopts a graphical presentation to interpret DGA data. The Duval triangle method computes relative portion of three dissolved gases (as shown in Figure 1.9) in the forms of $\%C_2H_2=x/(x+y+z)$, $\%C_2H_4=y/(x+y+z)$, and $\%CH_4=z/(x+y+z)$, in which x, y, and z denote the concentrations of C_2H_2, C_2H_4, and CH_4, respectively. Seven fault zones are allocated in a triangle,

consisting partial discharge, thermal faults at different temperature level, and electrical arcing. Some researchers reported Duval triangle method outperforms other ratio based methods and it can provide more consistent diagnoses [67, 70, 74]. However, Duval triangle cannot be used to identify transformers without any faults since the fault-free zone is absent in the triangle. Other than Duval triangle shown in Figure 1.9, Duval also developed other types of triangle methods to detect faults in oil-type load tap changers and non-mineral oils filled equipment. Figure 1.10 and Figure 1.11 depict the complements of the original Duval triangle method. In Figure 1.10, H_2, C_2H_6 and CH_4, are used and in Figure 1.11 C_2H_6, CH_4, and C_2H_4 are employed. These two complements triangles can be applied only if T1 (< 300℃), T2 (> 300℃ but < 700℃) or PD are already recognized by original Duval triangle method.



PD	Partial Discharge
D1	Discharges of low energy
D2	Discharges of high energy
DT	Combination of thermal faults and discharges
T1	Thermal faults not exceeding 300°C
T2	Thermal faults exceeding 300°C but not exceeding 700°C
T3	Thermal faults exceeding 700°C

Figure 1.9 Duval triangle method and its interpretation [66]



PD	Partial Discharge
S	Stray gassing of mineral oil
C	Hot spots with carbonization of paper (T>300°C)
O	Overheating (T<250°C)

Figure 1.10 Complement of Duval triangle method and its interpretation (Type I) [66]

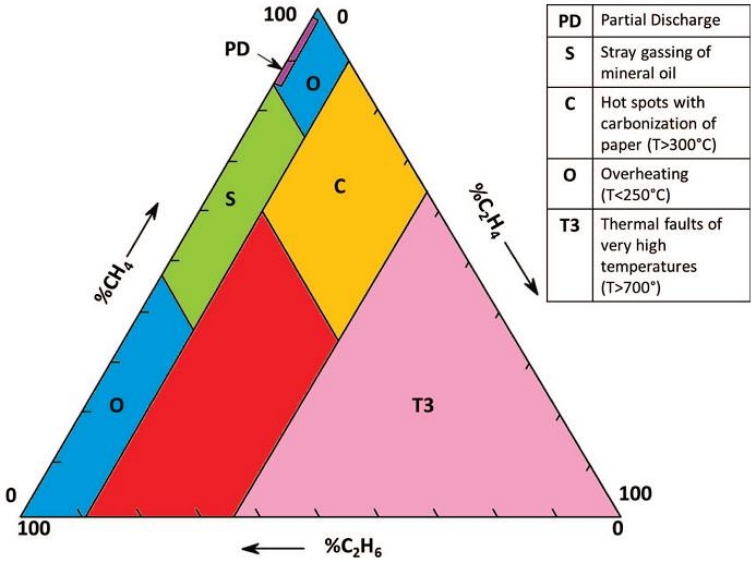


Figure 1.11 Complement of Duval triangle method and its interpretation (Type II) [66]

1.2.1.7.6 Duval pentagon method

Similar to Duval triangle method, another two graphical tools, i.e., Duval pentagons (type I and II) have been recently developed as new complementary tools to provide interpretation of DGA in mineral oil-filled transformers [75]. It adopts a pentagon coordinate where each axis denotes the relative portion of five dissolved gases (i.e., $\%H_2$, $\%CH_4$, $\%C_2H_6$, $\%C_2H_4$, and $\%C_2H_2$). The fault types can be determined by computing the geometrical center (red dot in Figure 1.12) of the polygon (red solid lines in Figure 1.12). Each type of Duval pentagon can identify seven potential faults and the difference between these two types pentagon are summarized in Table 1.6.

Table 1.6 Comparison of fault types of two types of Duval pentagon

Duval pentagon I	**Duval pentagon II**
PD: corona partial discharges	PD: corona partial discharges
D1: low energy discharges	D1: low energy discharges
D2: high energy discharges	D2: high energy discharges
T3: thermal faults >700 ℃	T3-H: thermal faults T3-H in oil only
T2: thermal faults of 300 to 700 ℃	C: thermal faults T3-C, T2-C, and T1-C with carbonization of paper
T1: thermal faults <300 ℃	O: overheating T1-O <250 ℃
S: stray gassing S of mineral oil at 120 and 200 ℃	S: stray gassing S of mineral oil at 120 and 200 ℃

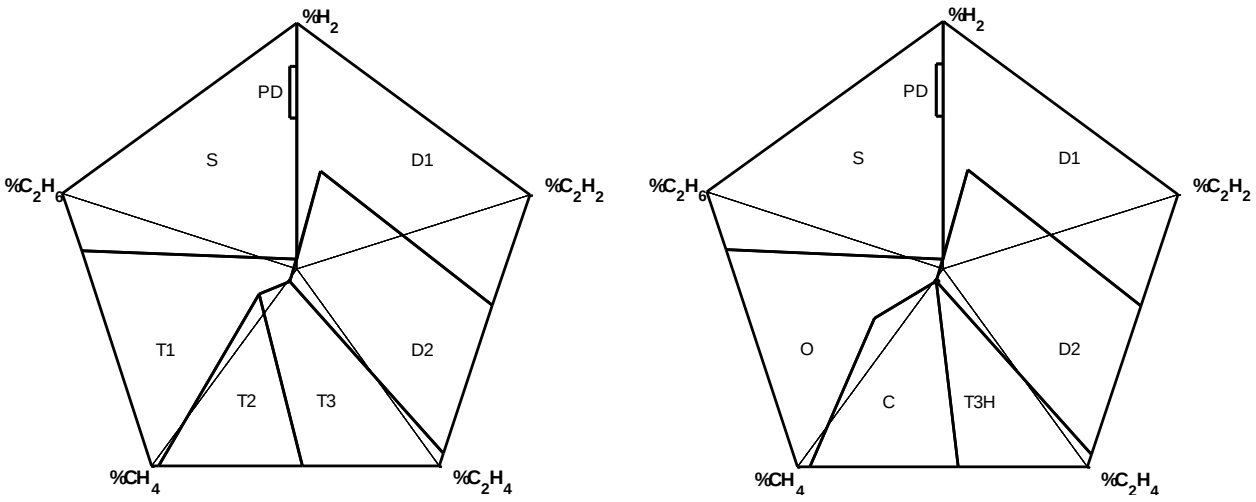


Figure 1.12 Coordinates and fault boundaries in the Duval pentagon method (a) Duval pentagon type I, (b) Duval pentagon type II

1.2.2 Water content estimation in transformer

Moisture is a dominant degradation product of paper and pressboard in transformers. It is formed by dehydration reactions following hydrolysis (which itself actually consumes moisture), but is also an end product in the oxidation of oil as well as that of paper [76-79]. The amount of moisture formed may constitute several percent by weight of solid insulation. Moisture in the insulation influences the remaining life of a transformer in many ways: accelerating ageing, increasing losses, reducing insulation strength and introducing the risk of bubble formation during overload [80-81]. To avoid premature ageing of the cellulose insulation, the transformer's moisture content must be kept to a minimum level when it is being commissioned. However, changes in the moisture content during cellulose degradation are inevitable and substantial (resulting from the cellulose degradation process) in the life span of a transformer. Moisture content in transformer insulation can be estimated through three approaches: cellulose adsorption isotherms approach, dielectric response measurement, and moisture diffusion modelling method.

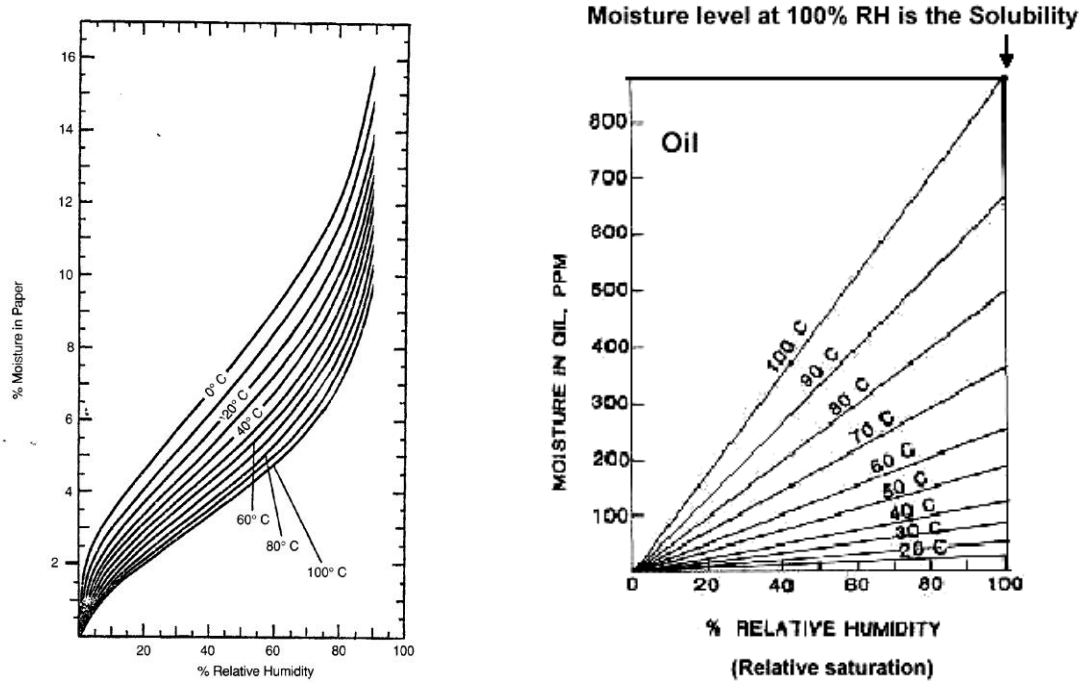
1.2.2.1 Cellulose adsorption isotherms approach

A large number of investigations have been performed to study a variety of phenomena involving cellulose adsorption isotherms of moisture between paper and oil [34, 82-87]. They involve the complex interpretation scheme to evaluate the humidity of cellulose insulation by using moisture-in-oil measurement, which is shown in Figure 1.13 [37, 88-89]. The knowledge of oil sampling

temperature and oil state (new or aged) is required to ensure an accurate moisture estimation.

Theoretical curves, known as Oommen curves, have been drawn to describe equilibrium conditions between new oil and solid insulation at different temperatures (shown in Figure 1.14) [37, 90]. The Oommen curves are theoretically derived from two experimental curves, describing the relation between moisture concentration in paper and relative water content of oil being in balance with that of air. These curves are in line with earlier direct measurement on unaged materials [34, 82]. To apply these curves, it is assumed that equilibrium of moisture between solid and liquid insulation can reach after a sufficient time period.

Equilibrium phenomena are dependent on temperature, thickness and humidity of the insulation material [34, 91-93]. The diffusion takes a long time especially at low temperatures and low water contents. Studies show that an equilibrium state may require more than three weeks [94-95]. In practice, the diffusion time can be even longer, since thick solid components (e. g. laminated wood) require long times to reach equilibrium [96].



(a) (b)

Figure 1.13 (a) Dependence between absolute water content in paper and relative humidity [97]. (b) Dependence between absolute water content in new oil and relative humidity according to [98].

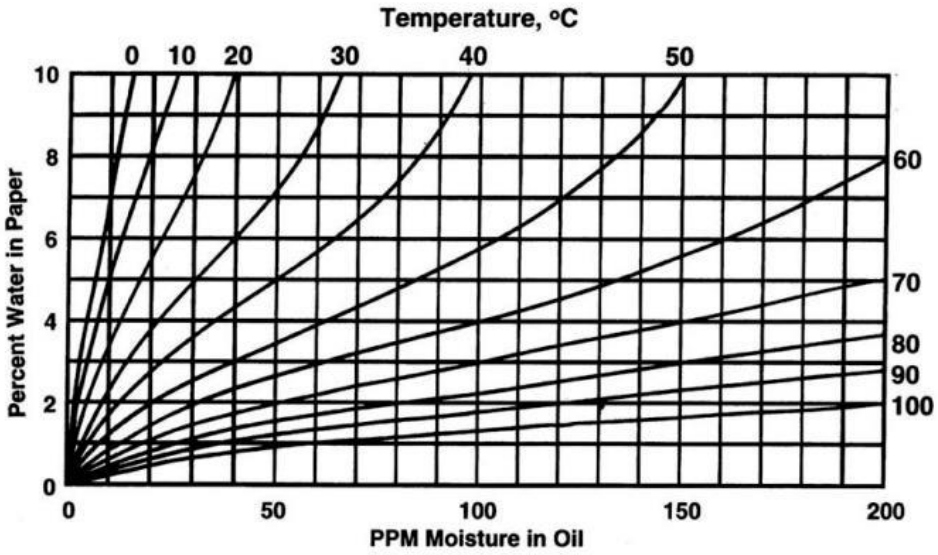


Figure 1.14 Oommen curves for water content distribution between new oil and paper [37]

1.2.2.2 Dielectric response measurement for moisture estimation

1.2.2.2.1 Dielectric response in time domain [3, 99-100]

Dielectric response measurement has been widely used for estimating moisture concentration of oil-cellulose insulation [100-102]. In insulating medium (dielectrics), which are in general isotropic and homogeneous, the vectors of electric field polarization, and dielectric displacement are of equal direction and they are correlated by (1.7) [103].

$$D = \varepsilon_0 E + P \tag{1.7}$$

where ε_0 denotes the permittivity of vacuum. Moreover, the dielectric displacement has a linear relationship with the applied field. The relation can be then expressed with a simple proportionality factor, ε, as shown in (1.8) .

$$D = \varepsilon\varepsilon_0 E \tag{1.8}$$

The factor ε denotes the relative permittivity and it describes the dielectric properties of the medium. Therefore, the polarization, P is also proportional to the field, E, as shown in (1.9) [103].

$$P = \chi\varepsilon_0 E = \varepsilon_0(\varepsilon - 1)E \tag{1.9}$$

where χ denotes the dielectric susceptibility of the medium.

The polarization in materials is a time dependent and therefore also a frequency dependent. This phenomenon can be caused by several mechanisms [103]: electronic, ionic (molecular), dipolar

(orientation), and interfacial. The first two types of polarization are extremely fast. For this reason, depending on the time scale used, the fast polarization mechanisms can together be considered as instantaneous, i.e. $P(t=t_0)=P_\infty$. Moreover, the polarization has to saturate, i.e. it must be finite at longer times, $P(t\to\infty)=P_s$. The time dependent polarization process is expressed in (1.10).

$$P(t)=P_\infty+(P_s-P_\infty)g(t-t_0) \tag{1.10}$$

where $g(t)$ is monotonically positively increasing with time

$$g(\mathrm{t})\geq 0,\frac{\partial g(\mathrm{t})}{\partial t}\geq 0 \qquad 0<t<\infty \tag{1.11}$$

And

$$g(\mathrm{t})=\begin{cases}0 & if \quad t\leq 0\\ 1 & if \quad t\to\infty\end{cases} \tag{1.12}$$

For an applied step-like constant electric field, and with the introduction of equation (1.9), the polarization can be written as (1.13).

$$P(t)=\varepsilon_0\left[(\varepsilon_\infty-1)+(\varepsilon_s-\varepsilon_\infty)g(t-t_0)\right]E \tag{1.13}$$

By applying superposition principle, one can show that for linear, homogeneous and isotropic dielectrics, the polarization under a time varying electric field, $E(t)$ can be written as (1.14).

$$P(t)=\varepsilon_0(\varepsilon_\infty-1)E(t)+\varepsilon_0\int_{-\infty}^{t}f(t-t_0)E(t_0)dt_0 \tag{1.14}$$

where $f(t)$ describes the dielectric response function and it is usually monotonically decreasing. The dielectric displacement $D(t)$ becomes (1.15).

$$D(t)=\varepsilon_0\varepsilon_\infty E(t)+\varepsilon_0\int_{-\infty}^{t}f(t-t_0)E(t_0)dt_0 \tag{1.15}$$

The first part of (1.15) represents the fast (instantaneous) polarization processes whereas the second part represents the slow (delayed) ones.

If a step electric field $E(t)$ is imposed on a dielectric medium, both free charges and dipoles will give rise to a current flowing through the medium. The movement of the free charges represents materials volume resistivity, whereas the bonded charges represent the dielectric displacement, including both polarization current and vacuum displacement current. The total current density is shown as (1.16).

$$\begin{aligned} j(t) &= \sigma_{dc}E(t) + \frac{\partial D(t)}{\partial t} = \sigma_{dc}E(t) + \varepsilon_0 \frac{\partial E(t)}{\partial t} + \frac{\partial P(t)}{\partial t} \\ &= \sigma_{dc}E(t) + \varepsilon_0 \left[\varepsilon_\infty \delta(t) + f(t)\right] E(t) \end{aligned} \tag{1.16}$$

The two asymptotic parts of the current, $j(t)$, are the instantaneous current density due to the capacitive component, $\varepsilon_0\varepsilon_\infty\delta(t)$ (where $\delta(t) = \frac{\partial E(t)}{\partial t}$ refers to the delta function), and the DC conductive current due to the conductivity of dielectric material, σ_{dc}, respectively. The current density caused by polarization is given by the dielectric response function, $f(t)$.

1.2.2.2.2 Dielectric response in frequency domain[100-101, 104]

Consider the properties of materials in electric fields which have an alternating time dependence that can be described as $\hat{E} = E_m e^{i\omega t}$, the complex dielectric displacement is written as (1.17).

$$\hat{D}(\mathrm{t}) = \varepsilon_0\varepsilon_\infty E_m e^{i\omega t} + \varepsilon_0 \int_{-\infty}^{t} f(\mathrm{t}-\mathrm{t}_0) E_m e^{i\omega t_0} dt_0 \tag{1.17}$$

It appears that the complex susceptibility, $\hat{\chi}(\omega)$ can be expressed as Fourier transformation of $f(t)$ as (1.18).

$$\hat{f}(\omega) = \int_0^\infty f(t) e^{-i\omega t} dt = \hat{\chi}(\omega) = \chi'(\omega) - j\chi''(\omega) \tag{1.18}$$

Both real and the imaginary parts of $\hat{\chi}(\omega)$ are not independent from each other since they are both generated by the same dielectric response function. They can be treated as the cosine and the sine transforms of the dielectric response function respectively.

The total current density $\hat{j}(\omega)$ in a dielectric material under $\hat{E}(\omega)$ excitation can therefore be expressed as (1.19).

$$\hat{j}(\omega) = i\omega\varepsilon_0 \left[\underbrace{\varepsilon_\infty + \chi'(\omega)}_{capacitive\ \ part} - \underbrace{i\left(\frac{\sigma_{dc}}{\varepsilon_0\omega} + \chi''(\omega) \right)}_{resistive\ \ part} \right] \hat{E}(\omega) \tag{1.19}$$

This expression shows that the current is composed of two components, one in phase with the applied field, $\hat{E}(\omega)$ - resistive part, and one that lies $90°$ before the driving field - capacitive part.

The part of the current in phase with the field represents the energy lost in the material. In fact, there are two different mechanisms contributing to this part, one is due to DC conduction (movement of free charges), and another due to relaxation losses (reorientation of bonded charges). The part of the current which lies 90 ° before the driving field is associated with the capacitance of the material.

In practices, complex permittivity is more often used instead of the complex susceptibility. The complex permittivity is expressed as (1.20).

$$\hat{j}(\omega) = i\omega\varepsilon_0 \left[\varepsilon'(\omega) - i\varepsilon''(\omega) \right] \hat{E}(\omega) \tag{1.20}$$

where $\varepsilon'(\omega) = \varepsilon_\infty + \chi'(\omega)$ and $\varepsilon''(\omega) = \frac{\sigma_{dc}}{\varepsilon_0 \omega} + \chi''(\omega)$.

The equations presented above show that the behaviour of dielectric material in frequency domain is dominated by the σ_{dc}, ε_∞, and $\hat{\chi}(\omega)$, which is similar in the time domain. If the dielectric is linear, homogeneous and isotropic material, the dielectric response embedded in either the time domain or frequency domain should be equivalent. The information found in one domain can be transformed to the other. If, however, the insulation system is composed of mixtures of several linear, homogeneous and isotropic materials, the total dielectric response is not a simple superposition of the individual contributions.

To make a precise moisture estimation on the oil-cellulose system in a transformer by using dielectric response measurements, a database containing dielectric properties ε_∞ , σ_{dc} and $f(t)$, of well characterized materials (oils and impregnated pressboards) at different temperature and humidity content is needed[5]. By comparing the measured dielectric frequency response of oil-cellulose system and the samples in the database, the moisture content of transformer insulation system and oil conductivity can be determined which should have the minimal mean squared errors between the two FDS measurement results. It should be notable that depending on the coupling of transformer's winding, different combinations of insulation may influence the dielectric response measurement. Therefore, information about the geometrical configuration of transformer's insulation system may also be needed for an accurate interpretation of dielectric response measurement results.

The main insulation of a core type transformer usually contains of a number of cylindrical shells of pressboard, separated by axial spacers. In dielectric response modelling, the complex insulation structure is usually simplified as *X*-*Y* model (Figure 1.15) [5]. In Figure 1.15, *X* denotes the proportion between the lumped thickness of all the barriers in the duct and the total duct's width, *Y* denotes the spacer coverage, and it can be calculated as the ratio between width summation of all

the spacers and the total length of duct's periphery.

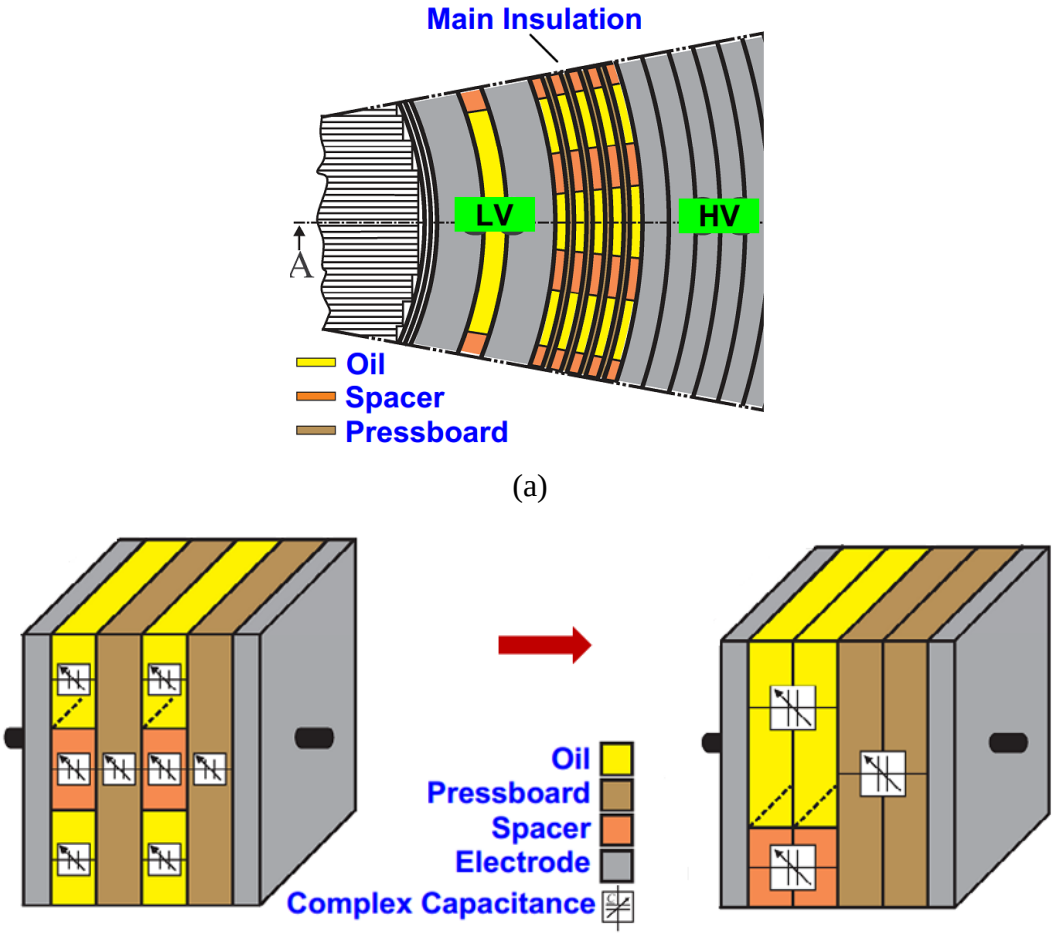


Figure 1.15 Schematic representation of the barrier content and the spacer coverage in the insulation duct (a) main insulation between HV-LV windings (b) *X*-*Y* model in homogenous electrical field

The range for the relative barrier amount X is typically from 0.15 to 0.5. The relative spacer coverage Y is typically from 0.15 to 0.25. In frequency domain (FDS method), the composite dielectric permittivity of *X*-*Y* model can be expressed as (1.21).

$$\varepsilon(\omega,\mathrm{T}) = \frac{Y}{\frac{1-X}{\varepsilon_{spacer}} + \frac{X}{\varepsilon_{barrier}}} + \frac{1-Y}{\frac{1-X}{\varepsilon_{oil}} + \frac{X}{\varepsilon_{barrier}}} \tag{1.21}$$

where ε_{oil} denotes the oil relative permittivity, ε_{spacer} and $\varepsilon_{barrier}$ denote the relative permittivity of the spacers and barriers. All the permittivities are complex values and they are influenced by temperature, moisture and frequency.

1.2.2.3 Moisture diffusion modelling approach

If all water molecules in a material are free to migrate, they tend to diffuse from the high moisture concentration region to the lower moisture concentration region, thereby minimizing the moisture gradient and equalizing the moisture concentration in the material [105]. Moisture migration can be modelled from a macroscopic perspective by solving the Fick's equation [84] in one dimension as (1.22).

$$\frac{\partial C(x,t)}{\partial t} = \frac{\partial}{\partial x}(D\frac{\partial C(x,t)}{\partial x}) \tag{1.22}$$

where $C(x,t)$ denotes the moisture concentration of cellulose insulation at position x and time t. D denotes the diffusion coefficient and is usually not a constant but influenced by both local moisture concentration and temperature.

The diffusion coefficient can be expressed as (1.23).

$$D = D_0 e^{[k\times c + E_a(\frac{1}{T_0}-\frac{1}{T})]} \tag{1.23}$$

Some researchers investigated moisture diffusion in transformers using Fick' second law and determined diffusion coefficients of cellulose [106-109]. In 1966, *Ast* [85] first proposed the diffusion coefficient of Kraft paper under different moisture concentrations and temperatures by using permeation method. Afterwards *Guidi* and *Fullerton* found an empirical relationship between the diffusion coefficients (moisture contents migrate to oil from cellulose) and the local moisture concentration and temperature [88, 110]. *Asem* determined diffusion coefficients for oil-immersed paper and non-immersed pressboard [83]. *Howe* investigated the diffusion coefficients for both Manila paper and pressboards [110]. *Foss* investigated a number of parameters of diffusion coefficients for both impregnated and non-impregnated Kraft paper based on other researchers' data [88].

By using dielectrometric sensors to measure moisture concentration in paper samples, *Du* calculated the diffusion coefficient for non-impregnated pressboards [33, 95]. A new mathematical tool to solve Fick's equation based on finite difference method was also developed. *Zhou* [111] proposed a model to dynamically determine the constants of moisture equilibrium under temperature transients and it allows estimating moisture content at any temperatures [112].

By taking into account the moisture distribution in oil-impregnated cellulose at both steady state and dynamic condition and moisture produced from insulation degradation, a new diffusion model was proposed by *Garcia* [86, 93, 113]. The proposed model was verified on a test transformer and the experimental results demonstrated that the increase in abnormal moisture can be predicted by comparing the estimated moisture from their developed model with the measured moisture in oil [87]. On the basis of [87], further studies were conducted to determine the formation of diffusion

coefficients of oil-impregnated pressboard by integrating the factor of pressboard thickness into conventional diffusion coefficient equation [91-92, 114]. It has showed a good agreement between the theoretic modelling and experimental results.

1.2.3 Thermal dynamic modelling on hot spot temperature

The deterioration of transformer's insulation is closely related to the thermal stress, which is determined mainly by the loading conditions and ambient temperature. A complete thermal model of transformer and an accurate estimation on hot spot temperature are also considered as one of the most essentials for transformer condition assessment. Different transformer loading guides are proposed to provide guidance on determining appropriate ratings and operation modes of transformer. This session provides a brief review of three existing thermal models, which are widely adopted in the estimation of transformer hot spot temperature.

1.2.3.1 IEEE model

IEEE loading guides [115] assumes a growth in the winding current will result in temperature rise in oil and winding. The top oil temperature rise is computed as an exponential response from the initial top oil temperature rise to the ultimate temperature rise as (1.24).

$$\tau_{oil,R}\frac{d\Delta\theta_{oil}}{dt}=\Delta\theta_{oil,U}-\Delta\theta_{oil,I} \tag{1.24}$$

where $\Delta\theta_{oil,I}$ denotes initial top oil temperature rise, $\Delta\theta_{oil,U}$ denotes ultimate top oil temperature rise, $\tau_{oil,R}$ denotes oil time constant, $\Delta\theta_{oil}$ denotes top oil temperature rise.

The ultimate temperature rise of top oil over ambient is calculated as (1.25).

$$\Delta\theta_{oil,U}=\Delta\theta_{oil,R}\cdot\left(\frac{1+R\cdot K^2}{1+R}\right) \tag{1.25}$$

where $\Delta\theta_{oil,R}$ denotes top oil temperature under rated load condition, R denotes ratio between load loss of transformer and loss of transformer during no load test, K denotes ratio between load current and rated winding current.

The winding hot spot temperature rise can be computed as an exponential response from the initial hot spot temperature rise ($\Delta\theta_{hs,I}$) to the ultimate temperature rise ($\Delta\theta_{hs,U}$) as (1.26).

$$\tau_{wnd,R}\frac{d\Delta\theta_{hs}}{dt}=\Delta\theta_{hs,U}-\Delta\theta_{hs,I} \tag{1.26}$$

where $\Delta\theta_{hs,I}$ denotes initial hot spot temperature rise, $\Delta\theta_{hs,U}$ denotes ultimate hot spot temperature rise over top oil, $\tau_{wnd,R}$ denotes winding time constant, $\Delta\theta_{hs}$ denotes hot spot temperature rise.

The ultimate temperature rise of hot spot temperature over top oil can be expressed as (1.27).

$$\Delta\theta_{hs,U}=\Delta\theta_{hs,R}\cdot\left(K^{2m}\right) \tag{1.27}$$

where $\Delta\theta_{hs,R}$ denotes hot spot temperature under rated load condition, m denotes nonlinear constant.

Finally hot spot temperature is calculated as the summation of ambient temperature, top oil temperature rise and hot spot temperature rise as (1.28).

$$\theta_{hs}=\theta_{amb}+\Delta\theta_{oil}+\Delta\theta_{hs} \tag{1.28}$$

where θ_{hs} denotes hot spot temperature, θ_{amb} denotes ambient temperature.

However, the above model does not consider the effect of ambient temperature variations on top oil temperature [116-118]. Lesieutre *et al.* [119] updated the model as (1.29).

$$\tau_{oil,R}\frac{d\Delta\theta_{oil}}{dt}=\Delta\theta_{oil,U}+\theta_{amb}-\Delta\theta_{oil} \tag{1.29}$$

Accordingly, the hot spot temperature is updated as (1.30).

$$\theta_{hs}=\theta_{oil}+\Delta\theta_{hs} \tag{1.30}$$

where θ_{oil} denotes top oil temperature.

1.2.3.2 Swift's model

Swift *et al.* [18, 120] adopted an equivalent circuit for transformer thermal modelling based on heat transfer theory. In Swift's model, hot spot temperature is computed by using nonlinear thermal resistance (Figure 1.16a). The differential equation for top oil temperature is expressed as (1.31).

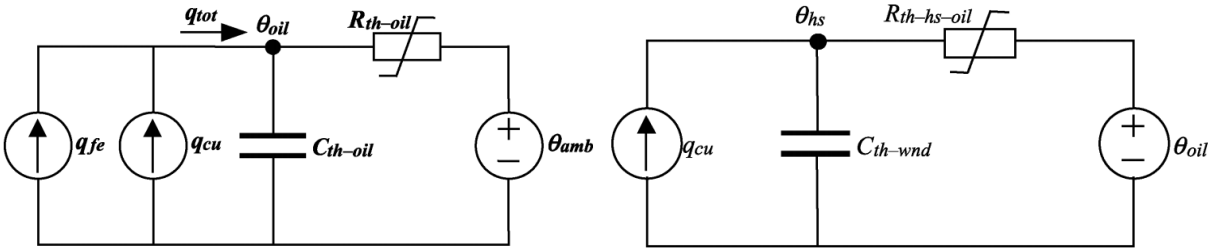
$$\frac{K^2R+1}{R+1}\cdot\Delta\theta_{oil,R}^{\frac{1}{n}}=\tau_{oil,R}\frac{d\theta_{oil}}{dt}+\left(\theta_{oil}-\theta_{amb}\right)^{\frac{1}{n}} \tag{1.31}$$

where n denotes nonlinear constant for top oil temperature.

The hot spot temperature is computed by using similar method of top oil temperature and its differential equation is expressed as (1.32).

$$K^2 \cdot \Delta\theta_{hs,R}^{\frac{1}{m}} = \tau_{wnd,R} \frac{d\theta_{hs}}{dt} + \left(\theta_{hs} - \theta_{oil}\right)^{\frac{1}{m}} \tag{1.32}$$

where m denotes nonlinear constant for hot spot temperature.



(a) Top oil temperature model (b) Hot spot temperature model

Figure 1.16 Swift's model based on electric thermal analogy [18]

1.2.3.3 Susa's model

Susa further improved Swift's model by considering the non-linear thermal resistance of mineral oil, which is caused by temperature dependent oil viscosity and loss variation [19, 121-122]. The top oil temperature of Susa's model can be expressed as (1.33).

$$\frac{K^2 R + 1}{R + 1} \cdot \Delta\theta_{oil,R} = \tau_{oil,R} \frac{d\theta_{oil}}{dt} + \frac{\left(\theta_{oil} - \theta_{amb}\right)^{1+n}}{\Delta\theta_{oil,R}^{n} \cdot \mu_{pu}^{n}} \tag{1.33}$$

where μ_{pu} denotes ratio of oil viscosity between certain temperature and rated top oil temperature.

The hot spot temperature is calculated as (1.34).

$$K^2 \cdot P_{Cu,pu} \cdot \Delta\theta_{hs,R} = \tau_{wnd,R} \frac{d\theta_{hs}}{dt} + \frac{\left(\theta_{hs} - \theta_{oil}\right)^{1+m}}{\Delta\theta_{hs,R}^{m} \cdot \mu_{pu}^{m}} \tag{1.34}$$

where $P_{Cu,pu}$ denotes per unit value of copper loss and it is calculated as (1.35).

$$P_{Cu,pu} = P_{Cu,DC,pu} \cdot \frac{235 + \theta_{hs}}{235 + \theta_{hs,R}} + P_{Cu,Eddy,pu} \cdot \frac{235 + \theta_{hs,R}}{235 + \theta_{hs}} \tag{1.35}$$

where $P_{Cu,DC,pu}$ denotes per unit value of DC loss, $P_{Cu,Eddy,pu}$ denotes Per unit value of eddy loss, $\theta_{hs,R}$ denotes hot spot temperature under rated load condition.

1.3 Challenges in Conventional Diagnostic Techniques

1.3.1 Uncertainty in dissolved gas analysis

Although DGA method has been widely used to identify incipient faults and assess health condition of power transformers, some key issues still need to be further investigated for utilities to perform a reliable diagnosis on power transformers. First, not all the combinations of gas ratios from measurement can be classified into certain fault type as defined in different DGA interpretation schemes (e.g., Doernenburg ratio, Roger's ratio and IEC ratio). Inconsistent diagnostic results may be produced by using different interpretation methods. Second, some DGA methods (e.g., Rogers ratio) may not be able to make diagnosis on certain faults (e.g., borderline cases). Thirdly, existing DGA interpretation schemes are mainly originated from statistical analysis on historical DGA data and experience of experts and engineers. The relationship between measured gas concentrations and the different fault types may not be fully explored and presented in the conventional interpretation schemes. Therefore, some intelligent interpretation techniques, such as expert system [6], ANNs [123] and fuzzy logic [124-125] are adopted to determine the correlation between dissolved gases and transformers' insulation conditions.

1.3.2 Intricate difficulties in moisture estimation

To ensure transformer's reliable operation, it is necessary to monitor the moisture in its oil-cellulose insulation system and perform a drying process when it is required. Moisture in oil can be directly measured from oil samples collected from the transformer. Karl Fischer Titration (KFT) method can be applied to measure moisture in cellulose. However, it requires collecting paper samples from winding insulation and it is not feasible for an operational transformer.

As an alternative, paper moisture can be indirectly determined through cellulose adsorption isotherms approach (also called equilibrium charts) [36-37] or certain mathematical equations [34] once the moisture in oil is obtained. However, this approach assumes an equilibrium state of moisture inside a transformer, which requires a constant temperature to be maintained in the transformer over a considerable long time. However, this is almost impossible for an operational transformer since it is normally subjected to continuous fluctuations of loading, temperature and other conditions. As such, moisture estimated from equilibrium charts might be erroneous.

For moisture estimation using Fick's diffusion equation, considerable difficulties still remain in applying the above diffusion coefficients to real transformers. This is because transformers may be

constructed by different cellulose materials and may have a heterogeneous temperature distribution during their operation. Moreover, Fick's law in one dimension (1D) may not be able to fully describe the diffusion process in the cellulose of a transformer. In addition, for 1D and 2D finite element modelling (FEM) of moisture distribution, the performance is normally satisfied and without the problem of numerical instabilities. However, the 3D FEM may encounter some difficulties. The first difficulty is related to the issues caused by Peclet number [126], which denotes the ratio between the energy transferred by the fluid convection and that by the fluid conduction. The second difficulty is that 3D FEM models require high level of details about transformer design.

1.3.3 Inaccuracy of existing thermal models

IEC and IEEE loading guides [115, 127] calculate the top oil temperature, bottom oil temperature and hot spot temperature of a transformer as an exponential response from the initial temperature rise to the ultimate temperature rise. However, later studies showed that the exponential response function may not be able to fully describe the hot spot temperature rise especially when the transformer is under temperature and load transients [128]. Moreover, some assumptions made in IEC and IEEE models, such as the top oil temperature has an instantaneous response to the variation in ambient temperature and the top oil temperature is identical to oil temperature in the cooling duct, may not be suitable for accurate hot spot temperature estimation [129].

For thermal dynamic models, some key parameters, such as oil time constant and winding time constant, need to be estimated from transformer's design parameters (e.g., mass of iron core and copper coil, weight of tank and volume of oil, etc.) for predicting hot spot temperature. These parameters may not be always available especially for old transformers. In addition, some investigations revealed that these models tend to produce a much faster initial temperature rise when transformer load increases and thus may produce a higher error in hot spot temperature compared to the measurement results [130]. Possible reason could be the overestimation of oil viscosity during the overload condition, thus the actual oil circulation in the transformer is at a much higher rate than it is presented in thermal dynamic models.

Thirdly, thermal dynamic models only consider the oil thermal properties by treating thermal transfer process as oil convection. None of these models consider cellulose thermal properties in determining hot spot temperature. This can cause a certain degree of uncertainties in temperature estimation. Therefore, it is of a highly demand for a practical thermal model to give reliable estimation on transformer thermal ratings.

1.4 Transformer Condition Assessment Using Data Centric Computational Intelligence

With the advancement of microelectronics and computer hardware, the diagnostics techniques presented in the precious sections have been widely applied in utilities. Consequently, huge volumes of raw data have been collected from these techniques in more efficient and inexpensive manners. However, it remains a challenging task to process raw data, extract information, integrate every piece and fragment of information, and transform them into a knowledge to assess conditions of transformers.

To overcome the above problems and provide a comprehensive assessment on transformer condition to facilitate the decision making of operation and maintenance, it is necessary to collect data from various online and offline measurements and historic database, and subsequently extracts useful information, and finally compute an index value regarding the health condition of the transformer [131].

Over the last two decades, various data mining techniques such as artificial neural networks (ANN) [123], fuzzy logic [132-133], neural-fuzzy system [125], and wavelet transform [134] have been applied to transformer insulation diagnosis using DGA or oil test results (e.g. oil characteristics). However, the data mining based interpretation scheme is still not a ready-to-use tool for power transformer condition assessment in utilities.

Firstly, there is lack of a common framework for defining the process of training, cross-validation, testing, and evaluation of data mining algorithms. Such framework is important to ensure generalization and applicability of various algorithms and provide performance comparisons amongst different algorithms. Secondly, considerable difficulties exist in constructing a statistically satisfied historic database for algorithm training. This is because the occurrence rate of some types of transformer faults is relatively low and thus the historic database may contain a small numbers of cases corresponding to these fault types. Trained by such database, the data mining algorithms may not be able to make reliable insulation diagnosis. Thirdly, oil sample collection, dissolved gas measurement, oil properties tests, and data analysis may vary among different utilities. These may influence the generalization ability of data mining algorithm, i.e. some algorithms trained on a "local" dataset may not be able to be readily applied "globally". This thesis will address the above issues and implement data mining algorithms, which can achieve consistent and desirable performance for transformer insulation diagnosis using DGA and/or oil characteristics.

Besides DGA, very few research works have been done to apply data centric techniques for interpreting dielectric response measurement results [135]. To investigate the moisture dynamics

and to accurately estimate moisture distribution in transformer's oil-cellulose system, two modelling approaches, i.e., multi-physics finite element modelling (FEM) and particle tracing method are proposed, which model the moisture diffusion from both macroscopic and microscopic levels. Based on the achievements obtained from these two modelling approaches, a dielectric frequency response model with distributed parameters is developed to reveal the effect of moisture diffusion on dielectric response measurement of transformer oil-cellulose system. The correlation between non-equilibrium moisture distribution and dielectric response parameters (dielectric losses and permittivity) of transformer cellulose insulation is explored. The proposed modelling methods can be used as an alternative to evaluate moisture concentration and improve the interpretation of dielectric response measurement of field transformers under thermal transients.

To provide an accurate estimation on hot spot temperature of transformer windings, this thesis developed a moisture dependent thermal model (MDTM) based on thermoelectric analogy theory. The proposed model formulates nonlinear thermal resistance by considering both oil and cellulose (paper and pressboard) of the transformer. The model also considers the effect of moisture concentration and hot spot temperature on the thermal resistance of cellulose. Based on the heat transfer principles and electric circuit laws, two sets of differential equations are derived to calculate the top oil and hot spot temperature of transformer when the transformer is under ambient temperature variation and load fluctuation. Finally, a practical tool is developed to estimate hot spot temperature of transformer windings by using the outputs of the two proposed thermoelectric analogy models.

By integrating every piece of data and information obtained from different transformer diagnostic measurements, a multi-source data and information fusion framework is developed based on Bayesian Network (BN) to facilitate determining optimal diagnostic strategies involved in transformer condition assessment.

1.5 Objectives

The overall aim of the research is to investigate and develop data centric diagnostic techniques for processing measurement data, recognizing various fault symptoms, and assessing the condition of transformers, which can provide an intelligent tool to help make informed decisions on transformer operation, maintenance and replacement strategies. To accomplish this aim, several objectives are addressed as follows:

(1) To apply data mining techniques for interpreting DGA and oil characteristics datasets and

diagnosing transformer insulation condition within a proposed framework of statistically satisfied training database construction, algorithms training, testing and validation, and performance evaluation.

(2) To develop novel interpretation schemes of dielectric response measurement for transformer condition assessment by employing data centric techniques with focus on water content estimation of oil and paper insulation of transformers.

(3) To develop a set of thermoelectric analogous thermal models to determine both top oil and hot spot temperature of the main parts in a transformer.

(4) To develop information fusion techniques to integrate diagnosis results obtained from various online and offline measurements and subsequently determine optimal diagnostic strategies for a transformer.

1.6 Thesis Overview

This thesis is based mostly on published and accepted papers. A small section of the thesis is prepared based on submitted papers, which are under review.

Chapter 2 provides a critical review on pattern recognition techniques for power transformer insulation diagnosis using dissolved gas and oil characteristics. A general pattern recognition application framework is outlined in this chapter and a comprehensive literature review on various pattern recognition techniques for transformer insulation diagnosis is provided. The important issues for improving the applicability of pattern recognition techniques for transformer insulation diagnosis are also discussed. To verify the applicability and generalization capability of different pattern recognition algorithms, 15 representative algorithms are implemented and extensive case studies are conducted on eight oil characteristics datasets collected from different utility companies. A statistical performance (in terms of classification accuracy) comparison amongst different pattern recognition algorithms for transformer insulation diagnosis using oil characteristics is also performed in this chapter.

Chapter 3 proposes a novel method for power transformer insulation assessment using oil characteristics. A hybrid algorithm, named as SMOTEBoost is implemented in the paper to improve the diagnosis accuracy and consistency. The SMOTEBoost can significantly enhance the generalization capability of artificial intelligence (AI) algorithms for transformer insulation diagnosis. This will provide important benefits for applying AI techniques in utility companies, i.e., an AI algorithm with its model built upon on a “local” dataset can be utilized “globally” to make transformer insulation diagnosis.

Chapter 4 proposes two modelling approaches for investigating moisture dynamics in power transformers. Firstly, a multi-physics finite element modelling (FEM) approach is developed where moisture dynamics is investigated by coupling the electromagnetic, thermal, fluid flow and moisture migration physics simultaneously. To overcome the numeric stability of FEM method in moisture dynamics modelling, an alternative particle tracing method is proposed which formulates moisture diffusion from a microscopic view of water molecules' motion. Especially, the transmission probability of water molecules (termed as particles in the paper) is employed to correlate the microscopic particles' motion with the macroscopic moisture distribution. The proposed modelling approaches are applied to model moisture dynamics in both pressboard specimens and cellulose insulation of a prototype transformer. Extensive ageing and moisture diffusion experiments have been conducted on a prototype transformer to verify the proposed modelling approaches.

Chapter 5 further investigates moisture dynamics and its effect on dielectric response of a transformer's cellulose insulation. It proposes a distributed parameter model to reveal the correlation between moisture distribution (under non-equilibrium conditions due to thermal transients) and dielectric response parameters (dielectric losses and permittivity) of cellulose insulation. It then estimates these parameters under moisture non-equilibrium conditions. The methodology developed in this chapter can help the proper interpretation of dielectric response measurement of field transformers under thermal transients.

Chapter 6 proposes a moisture dependent thermal model (MDTM) for estimating transformer hot spot temperature. In this model, nonlinear thermal resistance is formulated by considering both oil and cellulose (paper and pressboard) of the transformer. Especially, the effect of moisture concentration and hot spot temperature on the thermal resistance of cellulose is taken into account. The proposed MDTM is verified by using historical data of moisture-in-oil and temperature measurements on an in-service vegetable oil-filled transformer. Comparisons between the proposed MDTM and a number of existing thermal models are performed on the basis of adequacy and accuracy metrics.

Chapter 7 presents a multi-source data and information fusion framework for power transformer condition assessment. The proposed method adopts Bayesian Network (BN), which can integrate every piece of data and information obtained from different transformer diagnostic measurements. Within the Bayesian Network, Monte Carlo and Bootstrap methods are employed to extract the most informative characteristics regarding transformer condition from different diagnostic measurements. Reliability metrics are computed to evaluate the effectiveness of combinations of different type diagnostic measurements and subsequently facilitate determining optimal diagnostic strategies involved in transformer condition assessment.

Chapter 8 summarizes the contribution and draws conclusions from the research work. Suggestions for future works are also discussed in this chapter.

References

[1] M. Wang, A. J. Vandermaar and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electr. Insul. Mag.*, vol.18, Issue 6, pp. 12-25, 2002.

[2] L. E. Lundgaard, W. Hansen, D. Linhjell and T. J. Painter, "Aging of Oil-Impregnated Paper in Power Transformers," *IEEE Trans. Power Delivery*, vol.19, Issue 1, pp. 230-239, 2004.

[3] T. K. Saha and P. Purkait, "Investigation of polarization and depolarization current measurements for the assessment of oil-paper insulation of aged transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.11, Issue 1, pp. 144-154, 2004.

[4] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.10, Issue 5, pp. 903-917, 2003.

[5] C. Ekanayake, S. M. Gubanski, A. Graczkowski and K. Walczak, "Frequency response of oil impregnated pressboard and paper samples for estimating moisture in transformer insulation," *IEEE Trans. Power Delivery*, vol.21, Issue 3, pp. 1309-1317, 2006.

[6] C. E. Lin, J. M. Ling and C. L. Huang, "An expert system for transformer fault-diagnosis using dissolved-gas analysis," *IEEE Trans. Power Delivery*, vol.8, Issue 1, pp. 231-238, 1993.

[7] M. Duval, "Dissolved gas analysis: It can save your transformer," *IEEE Electr. Insul. Mag.*, vol.5, Issue 6, pp. 22-27, 1989.

[8] M. Duval and A. DePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electr. Insul. Mag.*, vol.17, Issue 2, pp. 31-41, 2001.

[9] M. Duval, "New techniques for dissolved gas-in-oil analysis," *IEEE Electr. Insul. Mag.*, vol.19, Issue 2, pp. 6-15, 2003.

[10] *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, C57.104, 2008.

[11] *Mineral Oil-impregnated Electrical Equipment In Service – Guide to the Interpretation of Dissolved and Free Gases Analysis*, IEC60599, 2007.

[12] T. K. Saha, P. Purkait and F. Muller, "Deriving an equivalent circuit of transformers insulation for understanding the dielectric response measurements," *IEEE Trans. Power Delivery*, vol.20, Issue 1, pp. 149-157, 2005.

[13] T. K. Saha and P. Purkait, "Investigation of an expert system for the condition assessment of transformer insulation based on dielectric response measurements," *IEEE Trans. Power Delivery*, vol.19, Issue 3, pp. 1127-1134, 2004.

[14] T. K. Saha, M. Darveniza, Z. T. Yao, D. J. T. Hill and G. Young, "Investigating the effects of oxidation and thermal degradation on electrical and chemical properties of power transformers insulation," *IEEE Trans. Power Delivery*, vol.14, Issue 4, pp. 1359-1367, 1999.

[15] T. K. Saha, M. Darveniza, D. J. T. Hill and T. T. Le, "Electrical and chemical diagnostics of transformer insulation - Part A: Aged transformer samples," *IEEE Trans. Power Delivery*, vol.12, Issue 4, pp. 1547-1554, 1997.

[16] T. K. Saha, M. Darveniza, D. J. T. Hill and T. T. Le, "Electrical and chemical diagnostics of transformer insulation - Part B: Accelerated aged insulation samples," *IEEE Trans. Power Delivery*, vol.12, Issue 4, pp. 1555-1561, 1997.

[17] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.10, Issue 5, pp. 903-917, 2003.

[18] G. Swift, T. S. Molinski and W. Lehn, "A Fundamental Approach to Transformer Thermal Modeling-Part I: Theory and Equivalent Circuit," *IEEE Trans. Power Delivery*, vol.16, Issue 2, pp. 171, 2001.

[19] D. Susa, M. Lehtonen and H. Nordman, "Dynamic Thermal Modelling of Power Transformers," *IEEE Trans. Power Delivery*, vol.20, Issue 1, pp. 197-204, 2005.

[20] Z. Wang, J. Li and D. M. Sofian, "Interpretation of transformer FRA responses-part I: influence of winding structure," *IEEE Trans. Power Delivery*, vol.24, Issue 2, pp. 703-710, 2009.

[21] R. Bartnikas, "Partial discharges their mechanism, detection and measurement," *IEEE Trans. Dielectr. Electr. Insul.*, vol.9, Issue 5, pp. 763-808, 2002.

[22] A. Mollmann and B. Pahlavanpour, "New guidelines for interpretation of dissolved gas analysis in oil-filled transformers," *Electra, CIGRE France*, vol.186, Issue30-51, 1999.

[23] T. V. Oommen, "Vegetable oils for liquid-filled transformers," *IEEE Electr. Insul. Mag.*, vol.18, Issue 1, pp. 6-11, 2002.

[24] M. Martins, "Vegetable Oils, an Alternative to Mineral Oil for power Transformers-Experimental Study of Paper Aging in Vegetable Oil Versus Mineral Oil," *IEEE Electr. Insul. Mag.*, vol.26, Issue 6, pp. 7-13, 2010.

[25] B. Sparling, "Assessing water content in solid transformer insulation from dynamic measurement of moisture in oil," in *Proceedings of IEEE PES Seminar, April 3*, 2008,Vancouver BC, pp. 1-70.

[26] P. Wiklund and B. Pahlavanpour, "Properties of Mineral Insulating Oils in Service," *J. Iranian Association of Electr. Electron. Eng*, vol.7, Issue 2, pp. 65-73, 2010.

[27] V. G. Arakelian, "Effective diagnostics for oil-filled equipment," *IEEE Electr. Insul. Mag.*, vol.18, Issue 6, pp. 26-38, 2002.

[28] M. Fischer, S. Tenbohlen, M. Schafer and R. Haug, "Determining Power Transformers' Sequence of Service in Power Grids," *IEEE Trans. Dielectr. Electr. Insul.*, vol.18, Issue 5, pp. 1789-1797, 2011.

[29] B. P. P. Wiklund, "Properties of Mineral Insulating Oils in Service," *CIGRE*, Issue 1, pp. 65-73, 2010.

[30] *Supervision and maintenance guide for mineral insulating oils in electrical equipment*, IEC 60422, 2005.

[31] Z. Xiang and E. Gockenbach, "Asset Management of Transformers Based on Condition Monitoring and Standard Diagnosis," *IEEE Electr. Insul. Mag.*, vol.24, Issue 4, pp. 26-40, 2008.

[32] S. Soares, G. Camino and S. Levchik, "Comparative study of the thermal decomposition of pure cellulose and pulp paper," *Polym. Degrad. Stab.*, vol.49, Issue 2, pp. 275-283, 1995.

[33] Y. Du, A. V. Mamishev, B. C. Lesieutre and M. Zahn, "Measurement of moisture diffusion as a function of temperature and moisture concentration in transformer pressboard," in *Proceedings of 1998. Annual Report of Conference on Electrical Insulation and Dielectric Phenomena, October 25-28*, 1998,Atlanta, USA, pp. 341-344.

[34] Y. Du, M. Zahn, B. C. Lesieutre, A. V. Mamishev and S. R. Lindgren, "Moisture equilibrium in transformer paper-oil systems," *IEEE Electr. Insul. Mag.*, vol.15, Issue 1, pp. 11-20, 1999.

[35] *Insulating Liquids—Oil-Impregnated Paper and Pressboard-Determination of Water by Automatic Coulometric Karl Fischer Titration*, IEC 60814 Ed. 2.0, 1997.

[36] Pahlavanpour, M. Martins and Eklund, "Study of moisture equilibrium in oil-paper system with temperature variation," in *Proceedings of International Conference on Properties and Applications of Dielectric Materials, June 1-5*, 2003,Nagoya, Japan, pp. 1124-1129.

[37] T. V. Oommen, "Moisture equilibrium charts for transformer insulation drying practice," *IEEE Transactions on Power Apparatus and Systems*, vol.103, Issue 10, pp. 3062-3067, 1984.

[38] A. M. Emsley, "The kinetics and mechanisms of degradation of cellulosic insulation in power transformers," *Polym. Degrad. Stab.*, vol.44, Issue 3, pp. 343-349, 1994.

[39] B. Pahlavanpour and I. A. Roberts, "Transformer Oil Condition Monitoring," in *Proceedings of Proceedings of the 1998 IEE Colloquium on Transformer Life Management, October 22*, 1998,London, UK, pp. 1-6.

[40] *Insulating liquids – Determination of acidity by automatic potentiometric titration*, IEC 62021, 2003.

[41] H. Gumilang, "Hydrolysis Process in PLN P3BJB Transformers as an Effect of Oil Insulation Oxidation," in *Proceedings of IEEE International Conference on Condition Monitoring and Diagnosis (CMD), September 23-27*, 2012,Bali, Indonesia, pp. 1147-1150.

[42] G. Ueta, T. Tsuboi, S. Okabe and T. Amimoto, "Study on Degradation Causing Components of Various Characteristics of Transformer Insulating Oil," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 6, pp. 2216-2224, 2012.

[43] D. Linhjell, L. E. Lundgaard and C. M. Selsbak, "Temperature-dependent Contamination Distribution between Oil, Paper and Pressboard," in *Proceedings of IEEE International Conference on Dielectric Liquids (ICDL), June 26-30*, 2011,Trondheim, Norway, pp. 1-4.

[44] N. Lelekakis, J. Wijaya, D. Martin and D. Susa, "The Effect of Acid Accumulation in Power-Transformer Oil on the Aging Rate of Paper Insulation," *IEEE Electr. Insul. Mag.*, vol.30, Issue 3, pp. 19-26, 2014.

[45] *Standard test method for dissipation factor (or power factor) and relative permittivity (dielectric constant) of electrical insulating liquids*, ASTM Standard D 924–04, 2004.

[46] *Measurement of relative permittivity, dielectric dissipation factor and DC resistivity of insulating liquids*, IEC 60247, 2004.

[47] M. Kohtoh, G. Ueta, S. Okabe and T. Amimoto, "Transformer Insulating Oil Characteristic Changes Observed Using Accelerated Degradation in Consideration of Field Transformer Conditions," *IEEE Trans. Dielectr. Electr. Insul.*, vol.17, Issue 3, pp. 808-818, 2010.

[48] S. Okabe, S. Kaneko, M. Kohtoh and T. Amimoto, "Analysis Results for Insulating Oil Components in Field Transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.17, Issue 1, pp. 302-311, 2010.

[49] M. Koch and S. Tenbohlen, "Evolution of bubbles in oil-paper insulation influenced by material quality and ageing," *IET Electr. Power Appl.*, vol.5, Issue 1, pp. 168-174, 2011.

[50] C. Y. Perkasa, N. Lelekakis, T. Czaszejko, J. Wijaya and D. Martin, "A comparison of the formation of bubbles and water droplets in vegetable and mineral oil impregnated transformer paper," *IEEE Trans. Dielectr. Electr. Insul.*, vol.21, Issue 5, pp. 2111-2118, 2014.

[51] A. De Pablo, "Interpretation of furanic compounds analysis-Degradation models, CIGRE WG D1.01.03, former WG 15-01, Task Force 03,", Paris, France, 1997.

[52] L. Cheim, D. Platts, T. Prevost and S. Z. Xu, "Furan Analysis for Liquid Power Transformers," *IEEE Electr. Insul. Mag.*, vol.28, Issue 2, pp. 8-21, 2012.

[53] M. Lessard, I. A. Höhlein, T. Buchacz, Luiz, L. Cheim and S. Eeckhoudt, "Furanic compounds for diagnosis," CIGRE Working Group, D1.01.13, 2010.

[54] D. Allan, "Recent advances in the analysis and interpretation of aged insulation from operating power transformers," in *Proceedings of The 5th International Conference on Properties and Applications of Dielectric Materials(ICPADM), May 25-30*, 1997,Seoul, Korea, pp. 202-205.

[55] R. Blue, D. Uttamchandani and O. Farish, "A novel solid-state material for furfuraldehyde detection," *IEEE Trans. Dielectr. Electr. Insul.*, vol.4, Issue 3, pp. 341-343, 1997.

[56] D. Urquiza, B. Garc á and J. C. Burgos, "Statistical Study on the Reference Values of Furanic Compounds in Power Transformers," *IEEE Electr. Insul. Mag.*, vol.31, Issue 4, pp. 15-23, 2015.

[57] A. M. Emsley, X. Xiao, R. J. Heywood and M. Ali, "Degradation of cellulosic insulation in power transformers. Part 2: Formation of furan products in insulating oil," *IEE Proceedings-Science Measurement and Technology*, vol.147, Issue 3, pp. 110-114, 2000.

[58] X. Chendong, "Monitoring Paper Insulation Ageing by measuring Furfural Contents in Oil," in *Proceedings of 7th International Symposium on High Voltage Engineering (ISH), August 26-30*, 1991,Dresden, Germany, pp. 1-6.

[59] T. Leibfried, M. Jaya, N. Majer, M. Schafer, M. Stach and S. Voss, "Postmortem Investigation of Power Transformers — Profile of Degree of Polymerization and Correlation With Furan Concentration in the Oil," *IEEE Trans. Power Delivery*, vol.28, Issue 2, pp. 886-893, 2013.

[60] A. B. Shkolnik and R. T. Rasor, "Statistical insights into furan interpretation using a large dielectric fluid testing database," in *Proceedings of IEEE PES Transmission and Distribution Conference and Exposition (T&D), May 7-10*, 2012,Orlando, FL, USA, pp. 1-8.

[61] A. J. Kachler and I. Hohlein, "Aging of cellulose at transformer service temperatures. Part 1: Influence of type of oil and air on the degree of polymerization of pressboard, dissolved gases, and furanic compounds in oil," *IEEE Electr. Insul. Mag.*, vol.21, Issue 2, pp. 15-21, 2005.

[62] I. Hohlein and A. J. Kachler, "Aging of cellulose at transformer service temperatures. Part 2. Influence of moisture and temperature on degree of polymerization and formation of furanic compounds in free-breathing systems," *IEEE Electr. Insul. Mag.*, vol.21, Issue 5, pp. 20-24, 2005.

[63] P. S. Pugh and H. H. Wagner, "Detection of incipient faults in transformers by gas analysis," *Power Apparatus and Systems, Part III. Transactions of the American Institute of Electrical Engineers*, vol.80, Issue 3, pp. 189-193, 1961.

[64] S. A. Ward, "Evaluating Transformer Condition Using DGA Oil Analysis," in *Proceedings of 2003 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, October 19-22*, 2003,Albuquerque, NM, USA, pp. 463-468.

[65] S. Corporation, "Serveron White Paper: DGA Diagnostic Methods,", 2007.

[66] N. Bakar, A. Abu-Siada and S. Islam, "A review of dissolved gas analysis measurement and interpretation techniques," *IEEE Electr. Insul. Mag.*, vol.30, Issue 3, pp. 39-49, 2014.

[67] R. R. Rogers, "IEEE and IEC Codes to Interpret Incipient Faults in Transformers, Using Gas in Oil Analysis," *IEEE Transactions on Electrical Insulation*, vol.13, Issue 5, pp. 349-354, 1978.

[68] *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, IEEE Std C57.104, 1991.

[69] *IEEE Guide for the Detection and Determination of Generated Gases in Oil-Immersed Transformers and Their Relation to the Serviceability of the Equipment*, ANSI/IEEE Std C57.104, 1978.

[70] A. Abu-Siada and S. Islam, "A New Approach to Identify Power Transformer Criticality and Asset Management Decision Based on Dissolved Gas-in-Oil Analysis," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 3, pp. 1007-1012, 2012.

[71] M. Duval, "The Duval Triangle for Load Tap Changers, Non-Mineral Oils and Low Temperature Faults in Transformers," *IEEE Electr. Insul. Mag.*, vol.24, Issue 6, pp. 22-29, 2008.

[72] M. Duval and A. DePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electr. Insul. Mag.*, vol.17, Issue 2, pp. 31-41, 2001.

[73] H. Sun, Y. Huang and C. Huang, "A Review of Dissolved Gas Analysis in Power Transformers," *Energy Procedia*, vol.14, Issue 1, pp. 1220-1225, 2012.

[74] V. Miranda and A. Castro, "Improving the IEC table for transformer failure diagnosis with knowledge extraction from neural networks," *IEEE Trans. Power Delivery*, vol.20, Issue 4, pp. 2509-2516, 2005.

[75] M. Duval and L. Lamarre, "The duval pentagon-a new complementary tool for the interpretation of dissolved gas analysis in transformers," *IEEE Electr. Insul. Mag.*, vol.30, Issue 6, pp. 9-12, 2014.

[76] A. M. Emsley and G. C. Stevens, "Review of chemical indicators of degradation of cellulosic electrical paper insulation in oil-filled transformers," *IEE Proceedings - Science, Measurement and Technology*, vol.141, Issue 5, pp. 324-334, 1994.

[77] L. E. Lundgaard, W. Hansen, D. Linhjell and T. J. Painter, "Aging of oil-impregnated paper in power transformers," *IEEE Trans. Power Delivery*, vol.19, Issue 1, pp. 230-239, 2004.

[78] J. Antonio Almendros-Ibanez, J. Carlos Burgos and B. Garcia, "Transformer Field Drying Procedures: A Theoretical Analysis," *IEEE Trans. Power Delivery*, vol.24, Issue 4, pp. 1978-1986, 2009.

[79] J. Fabre and A. Pichon, "Deteriorating processes and products of paper in oil. Application to transformers," *CIGRÉ paper 137*, Issue 1, pp. 1-18, 1960.

[80] W. A. Fessler, T. O. Rouse, W. J. McNutt and O. R. Compton, "A refined mathematical model for prediction of bubble evolution in transformers," *IEEE Trans. Power Delivery*, vol.4, Issue 1, pp. 391-404, 1989.

[81] H. Yoshida, Y. Ishioka, T. Suzuki, T. Yanari and T. Teranishi, "Degradation of insulating materials of transformers," *IEEE Transactions on Electrical Insulation*, vol.22, Issue 6, pp. 795-800, 1987.

[82] D. F. Garcia, B. Garcia and J. Burgos, "A review of moisture diffusion coefficients in transformer solid insulation-part 1: Coefficients for paper and pressboard," *IEEE Electr. Insul. Mag.*, vol.29, Issue 1, pp. 46-54, 2013.

[83] A. S. Asem and A. F. Howe, "Drying of power-transformer insulation," *IEE Proceedings - Generation, Transmission and Distribution*, vol.129, Issue 5, pp. 228-232, 1982.

[84] A. F. Howe, "Diffusion of moisture through power-transformer insulation," *Proc. Inst. Electr. Eng.*, vol.125, Issue 10, pp. 978-986, 1978.

[85] P. F. Ast, "Movement of moisture through A50P281 Kraft paper (dry and oil impregnated)," General Electric, 1966.

[86] B. Garcia, J. C. Burgos, A. M. Alonso and J. Sanz, "A moisture-in-oil model for power transformer monitoring - Part I: Theoretical foundation," *IEEE Trans. Power Delivery*, vol.20, Issue 2, pp. 1417-1422, 2005.

[87] B. Garcia, J. C. Burgos, A. M. Alonso and J. Sanz, "A moisture-in-oil model for power transformer monitoring - Part II: Experimental verification," *IEEE Trans. Power Delivery*, vol.20, Issue 2, pp. 1423-1429, 2005.

[88] S. D. Foss and L. Savio, "Mathematical and experimental-analysis of the field drying of power transformer insulation," *IEEE Trans. Power Delivery*, vol.8, Issue 4, pp. 1820-1828, 1993.

[89] A. M. Emsley and G. C. Stevens, "Kinetics and mechanisms of the low-temperature degradation of cellulose," *Cellulose*, vol.1, Issue 1, pp. 26-56, 1994.

[90] T. V. Oommen, "Moisture equilibrium charts for transformer insulation drying practice," in *Proceedings of IEEE/PES Transmission and Distribution Conference, April 29- May 4*, 1984,Kansas City, Missouri, USA, pp. 3062-3067.

[91] R. Villarroel, D. F. Garcia, B. Garcia and J. C. Burgos, "Diffusion coefficient in transformer pressboard insulation part 2: mineral oil impregnated," *IEEE Trans. Dielectr. Electr. Insul.*, vol.21, Issue 1, pp. 394-402, 2014.

[92] R. Villarroel, D. F. Garcia, B. Garcia and J. C. Burgos, "Diffusion coefficient in transformer pressboard insulation part 1: non impregnated pressboard," *IEEE Trans. Dielectr. Electr. Insul.*, vol.21, Issue 1, pp. 360-368, 2014.

[93] D. F. Garcia, "A new proposed moisture diffusion coefficient for transformer paper," *Int. J. Heat Mass Transfer*, vol.56, Issue 1-2, pp. 469-474, 2013.

[94] D. F. Garcia, R. Villarroel, B. Garcia and J. Burgos, "A review of moisture diffusion coefficients in transformer solid insulation - Part 2: Experimental validation of the coefficients," *IEEE Electr. Insul. Mag.*, vol.29, Issue 2, pp. 40-49, 2013.

[95] Y. Du, "Measurements and modeling of moisture diffusion processes in transformer insulation using interdigital dielectrometry sensors," PhD dissertation, Dept. Electr. Eng. Comp. Sci, Mass. Inst. Tech., Cambridge, MA, 1999.

[96] D. Martin, N. Lelekakis, J. Wijaya and K. Williams, "Water uptake rates of transformer paper insulation impregnated with vegetable oil," *IEEE Electr. Insul. Mag.*, vol.29, Issue 5, pp. 56-61, 2013.

[97] R. Jeffries, "The Sorption of Water by Cellulose and Eight Other Textile Polymers," *Journal of the Textile Institute Transactions*, vol.51, Issue 9, pp. T339-T340, 1960.

[98] *IEEE Guide for Acceptance and Maintenance of Insulating Oil in Equipment*, IEEE C57.106, 2002.

[99] W. S. Zaengl, "Dielectric spectroscopy in time and frequency domain for HV power equipment, Part I: Theoretical considerations," *IEEE Electr. Insul. Mag.*, vol.19, Issue 5, pp. 5-19, 2003.

[100] S. M. Gubanski, P. Boss, G. Csepes, V. Der Houbanessian, J. Filippini and P. Guuinic, *et al.*, "Dielectric response methods for diagnostics of power transformers," *IEEE Electr. Insul. Mag.*, vol.19, Issue 3, pp. 12-18, 2003.

[101] S. M. Gubanski, J. O. R. Blennow, B. Holmgren, M. Koch, A. Kuechler and R. Kutzner, *et al.*, "Dielectric response diagnoses for transformer windings," *CIGRE Task Force D1.01 Technical Brochure 414, Paris, 2010*, vol.1, Issue 1, pp. 1-58, 2010.

[102] W. S. Zaengl, "Applications of dielectric spectroscopy in time and frequency domain for HV power equipment," *IEEE Electr. Insul. Mag.*, vol.19, Issue 6, pp. 9-22, 2003.

[103] A. K. Jonscher, *Dielectric Relaxation in Solids*, London: Chelsea Dielectrics Press, 1983.

[104] M. Koch and T. Prevost, "Analysis of dielectric response measurements for condition assessment of oil-paper transformer insulation," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 6, pp. 1908-1915, 2012.

[105] R. B. Keey, T. A. G. Langrish and J. C. F. Walker, *Kiln-Drying of Lumber*, Berlin Heidelberg: Springer Verlag, 2000.

[106] R. J. Liao, Y. D. Lin, P. Guo, H. B. Liu and H. H. Xia, "Thermal Aging Effects on the Moisture Equilibrium Curves of Mineral and Mixed Oil-paper Insulation Systems," *IEEE Trans. Dielectr. Electr. Insul.*, vol.22, Issue 2, pp. 842-850, 2015.

[107] J. Li, Z. T. Zhang, S. Grzybowski and M. Zahn, "A New Mathematical Model of Moisture Equilibrium in Mineral and Vegetable Oil-Paper Insulation," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 5, pp. 1615-1622, 2012.

[108] J. Li, Z. T. Zhang, S. Grzybowski and Y. Liu, "Characteristics of Moisture Diffusion in Vegetable Oil-paper Insulation," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 5, pp. 1650-1656, 2012.

[109] B. Garcia, J. C. Burgos, A. M. Alonso and J. Sanz, "A Moisture-in-Oil Model for Power Transformer Monitoring-Part I: Theoretical Foundation," *IEEE Trans. Power Delivery*, vol.20, Issue 2, pp. 1417-1422, 2005.

[110] A. F. Howe, "The diffusion of moisture through power transformer insulation," PhD dissertation, Dept. Electr. Electron. Eng., Univ. Nottingham, Nottingham, UK, 1975.

[111] L. J. Zhou, G. N. Wu and J. Liu, "Modeling of transient moisture equilibrium in oil-paper insulation," *IEEE Trans. Dielectr. Electr. Insul.*, vol.15, Issue 3, pp. 872-878, 2008.

[112] E. S. Mladenov, S. G. Staykov and G. S. Cholakov, "Water Saturation Limit of Transformer Oils," *IEEE Electr. Insul. Mag.*, vol.25, Issue 1, pp. 23-30, 2009.

[113] D. F. Garcia, B. Garcia and J. C. Burgos, "Modeling Power Transformer Field Drying Processes," *Drying Technol.*, vol.29, Issue 8, pp. 896-909, 2011.

[114] D. F. Garcia, B. Garcia, J. C. Burgos and N. Garcia-Hernando, "Determination of moisture diffusion coefficient in transformer paper using thermogravimetric analysis," *Int. J. Heat Mass Transfer*, vol.55, Issue 4, pp. 1066-1075, 2012.

[115] *IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators*, IEEE Std C57.91-2011 (Revision of IEEE Std C57.91-1995), 2012.

[116] O. A. Amoda, D. J. Tylavsky, G. A. McCulla and W. A. Knuth, "Acceptability of Three Transformer Hottest-Spot Temperature Models," *IEEE Trans. Power Delivery*, vol.27, Issue 1, pp. 13-22, 2012.

[117] L. Jauregui-Rivera and D. J. Tylavsky, "Acceptability of Four Transformer Top-Oil Thermal Models — Part I: Defining Metrics," *IEEE Trans. Power Delivery*, vol.23, Issue 2, pp. 860-865, 2008.

[118] L. Jauregui-Rivera and D. J. Tylavsky, "Acceptability of Four Transformer Top-Oil Thermal Models — Part II: Comparing Metrics," *IEEE Trans. Power Delivery*, vol.23, Issue 2, pp. 866-872, 2008.

[119] B. C. Lesieutre, W. H. Hagman and J. L. Kirtley, "An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system," *IEEE Trans. Power Delivery*, vol.12, Issue 1, pp. 249-256, 1997.

[120] G. Swift, T. S. Molinski, R. Bray and R. Menzies, "A Fundamental Approach to Transformer Thermal Modeling-Part II: Field Verification," *IEEE Trans. Power Delivery*, vol.16, Issue 2, pp. 176, 2001.

[121] D. Susa, J. Palola, M. Lehtonen and M. Hyvarinen, "Temperature Rises in an OFAF Transformer at OFAN Cooling Mode in Service," *IEEE Trans. Power Delivery*, vol.20, Issue 4, pp. 2517-2525, 2005.

[122] D. Susa, M. Lehtonen and H. Nordman, "Dynamic Thermal Modeling of Distribution Transformers," *IEEE Trans. Power Delivery*, vol.20, Issue 3, pp. 1919-1929, 2005.

[123] I. N. Da Silva, M. M. Imamura and A. N. de Souza, "Application of neural networks to the analysis of dissolved gases in insulating oil used in transformers," in *Proceedings of 2000 IEEE International Conference on Systems, Man and Cybernetics, October 8-11*, 2000,Nashville, TN, USA, pp. 2643-2648.

[124] Q. Su, C. Mi, L. L. Lai and P. Austin, "A fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer," *IEEE Trans. Power Syst.*, vol.15, Issue 2, pp. 593-598, 2000.

[125] R. Naresh, V. Sharma and M. Vashisth, "An integrated neural fuzzy approach for fault diagnosis of transformers," *IEEE Trans. Power Delivery*, vol.23, Issue 4, pp. 2017-2024, 2008.

[126] M. A. Tsili, E. I. Amoiralis, A. G. Kladas and A. T. Souflaris, "Power transformer thermal analysis by using an advanced coupled 3D heat transfer and fluid flow FEM model," *Int. J. Therm. Sci.*, vol.53, Issue 53, pp. 188-201, 2012.

[127] *Power transformers – Part 7:Loading guide for oil-immersed power transformers*, IEC 60076-7, 2005.

[128] H. Nordman and M. Lahtinen, "Thermal overload tests on a 400-MVA power transformer with a special 2.5-p.u. Short time loading capability," *IEEE Trans. Power Delivery*, vol.18, Issue 1, pp. 107-112, 2003.

[129] R. Godina, E. Rodrigues, J. Matias and J. Catalão, "Effect of Loads and Other Key Factors on Oil-Transformer Ageing: Sustainability Benefits and Challenges," *Energies*, vol.8, Issue 10, pp. 12147-12186, 2015.

[130] D. Susa and H. Nordman, "A Simple Model for Calculating Transformer Hot-Spot Temperature," *IEEE Trans. Power Delivery*, vol.24, Issue 3, pp. 1257-1265, 2009.

[131] A. Naderian, R. Piercy, S. Cress, J. Service and W. Fan, "An approach to power transformer asset management using health index," *IEEE Electr. Insul. Mag.*, vol.25, Issue 2, pp. 20-34, 2009.

[132] Z. Xu, "Multivalued logic and fuzzy logic - their relationship, minimization, and application to fault diagnosis.," *IEEE Trans. Comput.*, vol.C-33, Issue 7, pp. 679-681, 1984.

[133] Q. Su, C. Mi, L. L. Lai and P. Austin, "Fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer," *IEEE Trans. Power Syst.*, vol.15, Issue 2, pp. 593-598, 2000.

[134] W. Chen, C. Pan, Y. Yun and Y. Liu, "Wavelet Networks in Power Transformers Diagnosis Using Dissolved Gas Analysis," *IEEE Trans. Power Delivery*, vol.24, Issue 1, pp. 187-194, 2009.

[135] H. Ma, T. K. Saha and C. Ekanayake, "Statistical Learning Techniques and Their Applications for Condition Assessment of Power Transformer," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 2, pp. 481-489, 2012.

Chapter 2

A Critical Review on Pattern Recognition Techniques for Power Transformer Insulation Diagnosis Using Dissolved Gas and Oil Characteristics

Contribution of the Chapter

In this chapter, a general pattern recognition application framework is outlined and a comprehensive literature review on various pattern recognition techniques for transformer insulation diagnosis is provided. The important issues for improving the applicability of pattern recognition techniques for transformer insulation diagnosis are also discussed. To verify the applicability and generalization capability of different pattern recognition algorithms, fifteen representative algorithms are implemented and extensive case studies are conducted on eight oil characteristics datasets collected from different utility companies. Statistical performance (in terms of classification accuracy) comparison amongst different pattern recognition algorithms for transformer insulation diagnosis using oil characteristics is performed and presented in this chapter.

2.1 Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 1: Framework, Literature, and Illustration

Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 1: Framework, Literature, and Illustration

Yi Cui, Hui Ma, and Tapan Saha

The University of Queensland, Brisbane, Australia

Publication Journal: International Transactions on Electrical Energy Systems

Submitted: March 14, 2014

Revision Resubmitted: May 11, 2014

Accepted: 14 May, 2014

Published: 21 June, 2014

Author's Contributions:

Yi Cui: Theoretical model design, numeric experimental simulation, results analysis and manuscript preparation.

Hui Ma: Supervision on the numeric experiment, results discussion and manuscript revision.

Tapan Saha: Supervision on the numeric experiment, results discussion and manuscript revision.

Abstract

The condition of the insulation system of a power transformer has a significant impact on its overall reliability and serviceability. Transformer oil tests including breakdown voltage, acidity, dielectric dissipation factor, 2-furfuraldehyde, water content and dissolved gases analysis have been commonly performed in utility companies to provide information regarding the conditions of transformer insulation. Over the past two decades, various pattern recognition techniques are proposed to interpret the oil tests results and make diagnosis on transformer insulation. However, there are still considerable challenging issues to be investigated before the pattern recognition technique can become a "ready-to-use" tool at utility companies. This paper provides a comparative study of pattern recognition techniques for power transformer insulation diagnosis using oil tests results. A general pattern recognition application framework will be outlined in the paper. And a comprehensive literature review on various pattern recognition techniques for transformer insulation diagnosis will be provided in the paper. The important issues for improving the applicability of pattern recognition techniques for transformer insulation diagnosis will also be discussed. A case study will be presented to demonstrate the procedure of applying pattern recognition techniques to practical transformer insulation diagnosis using oil test results.

Index Terms: Dissolved gas analysis, insulation, oil characteristics, pattern recognition, power transformer

2.1.1 Introduction

Power transformer is one of the most crucial equipment in an electricity grid. Its serviceability has a significant influence on the reliable delivery of electricity. However, a power transformer's insulation can be eventually deteriorated since the transformer is continuously subjected to electrical, mechanical, and thermal stresses. Such insulation deterioration may lead to a disruptive failure of a transformer [1-3]. Therefore, a variety of techniques have been developed for transformer insulation diagnosis, including: (1) oil tests such as breakdown voltage (BDV), acidity, dielectric dissipation factor (DDF), resistivity, 2-furfuraldehyde, water content, and dissolved gases analysis (DGA); (2) dielectric response measurement consisting of polarization and depolarization

current (PDC) measurement and frequency dielectric spectroscopy (FDS); (3) frequency response analysis (FRA); and (4) partial discharge (PD) measurement [4-8].

Among all of the above techniques, transformer oil tests have been commonly adopted by utilities to reveal various chemical and physical properties of insulating oil. Through the measurement of 2-furfuraldehyde and carbon oxides, oil tests may also provide some indications of the condition of the solid (pressboard and paper) insulation of transformer. In the past twenty years, various interpretation methods have been proposed for analyzing oil tests results (oil characteristics) and detecting discharge and overheat faults occurring in transformer insulation system [9]. However, there still exist some limitations of these conventional schemes. For example, the diagnosis results might be inconsistent by using different industry standards. Sometimes the conventional schemes cannot produce the diagnosis results for every possible combination of dissolved gases' ratios [10].

To overcome the disadvantages of the conventional interpretation schemes, various pattern recognition techniques such as artificial neural networks (ANN) [11], fuzzy logic [12-13], neural-fuzzy system [14], and wavelet transforms [15] have been extensively investigated. These techniques have achieved some extent of success in transformer insulation diagnosis using oil characteristics. In contrast to several conventional interpretation schemes, pattern recognition techniques utilize not only the oil tests data obtained from the transformer of interest, but also the historical oil test datasets collected from other transformers. By using historic datasets, a pattern recognition algorithm learns the underlying relationship between oil characteristics and the transformers insulation and it then applies such knowledge to make a diagnosis on the insulation condition of the transformer of interest.

Nevertheless, there are still considerable challenging issues to be investigated before the pattern recognition based interpretation schemes can become a "ready-to-use" tool for transformer insulation diagnosis in utilities. This paper provides a comparative study of pattern recognition techniques and their applications in power transformer insulation diagnosis using oil characteristics. Starting with a brief review of oil tests and conventional interpretation schemes, this paper outlines the insulation diagnosis problem into a general pattern recognition framework. A comprehensive literature review on various pattern recognition algorithms for transformer insulation diagnosis is provided in this paper. The challenging issues of improving the applicability of pattern recognition techniques will also be discussed. A case study will be presented to demonstrate the procedure of applying pattern recognition techniques to practical transformer insulation diagnosis using oil characteristics. The detailed mathematic formulation, implementation, and statistical performance evaluation of 15 pattern recognition algorithms will be provided in the accompanying paper [16].

2.1.2 Transformer Oil Tests

2.1.2.1 Oil tests

Because of continuously operating under various stresses, it is possible that oil molecular bonds partially crack and free particles can be generated. The interactions among these particles or between these particles and external molecules may form by-products including water contents, dissolved gases, acid components and other types of contaminants in oil.

The transformer oil tests are able to detect the above by-products and subsequently can provide some insights regarding transformer insulation. DGA results can be useful to detect transformer incipient faults such as arcing, partial discharge, and thermal fault. Water contents play a key role in the transformer insulation ageing. The large increase of water contents in transformer oil can decrease the resistivity and electrical strength of transformer oil. In the presence of solid contaminant and water contents, the reverse impacts of acids on the dielectric properties of transformer oil may become significant. The DDF can reveal the contaminant alteration in transformer oil. Moreover, the resistivity test measures the degree of losses of transformer oil and the BDV indicates the dielectric strength of transformer oil.

2.1.2.2 Conventional interpretation schemes

The conventional interpretation schemes use the concentrations or the ratios of the dissolved gases to diagnose the transformer insulation faults. For example, IEC/IEEE and Rogers's schemes use three gas ratios of C_2H_2/ C_2H_4, CH_4/ H_2, and C_2H_4/ C_2H_6, while Doernenburg method uses four gas ratios of CH_4/H_2, C_2H_2/C_2H_4, C_2H_2/CH_4 and C_2H_6/C_2H_2. The ratios in these schemes are then compared with the threshold values provided by the relevant industry standards to diagnose insulation condition. On the other hand, the Duval triangle method uses relative portion of three dissolved gases in the forms of $\%C_2H_2=x/(x+y+z)$, $\%C_2H_4=y/(x+y+z)$, and $\%CH_4=z/(x+y+z)$, in which x, y, and z denote the concentrations of C_2H_2, C_2H_4, and CH_4, respectively. It can reveal multiple faults (e.g. thermal and discharge faults) that may simultaneously occur in a transformer.

Instead of using the above conventional schemes, this paper investigates pattern recognition

techniques and their applications for transformer insulation diagnosis using oil tests data (oil characteristics). After presenting a generic pattern recognition application framework for transformer insulation diagnosis, a comprehensive literature review on various pattern recognition techniques reported in the literature will be provided in the following sections.

2.1.3 Pattern Recognition Framework for Transformer Diagnosis

In pattern recognition, a computer algorithm is trained to distinguish the pattern of interest from the background and makes decisions on the category of this pattern.

Figure 2.1 depicts a generic pattern recognition application framework, which involves three steps: 1) data pre-processing, 2) feature extraction, and 3) classification.

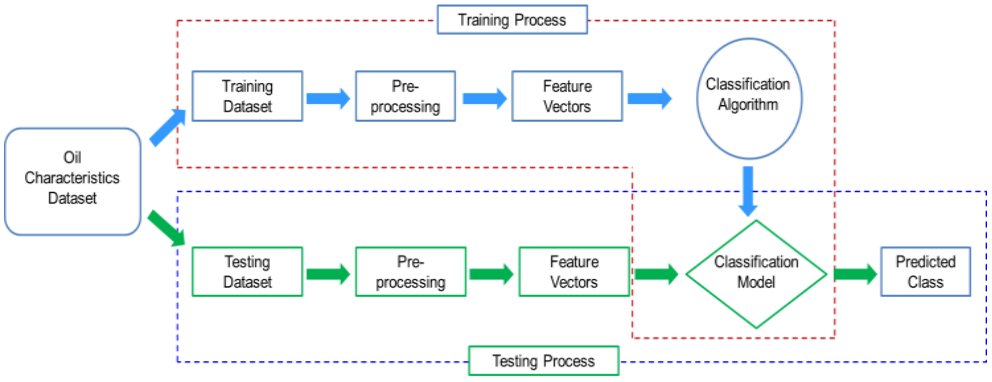


Figure 2.1 Framework of pattern recognition

2.1.3.1 Data pre-processing

The raw data might be inconsistent and noise-corrupted. Redundancies may also occur due to the integration of data from different sources. The data pre-processing step applies various techniques to process the raw data including: data cleaning for dealing with noise and removing redundant data; data transformation for converting the raw data into appropriate forms; data reduction for eliminating unnecessary attributes; and data discretization for reducing the number of levels of an attribute of data.

2.1.3.2 Feature extraction

Feature extraction aims to find the characteristic attributes (features) from the original data. This will enable pattern recognition algorithms to focus on those most relevant features for faults classification. In transformer insulation diagnosis using oil characteristics, the features may include the concentrations of dissolved gases and oil quality tests results such as acidity, BDV, DDF, 2-furfuraldehyde, resistivity, and water content.

2.1.3.3 Classification

In the classification step, the pattern recognition algorithm makes use of a historic oil tests dataset to construct a mathematical model that approximates the relationship between the oil characteristics (e.g. features) and the categories of transformer insulation condition or the types of incipient faults of transformer. Then the above model is used to make classification on the insulation condition or the incipient fault of the transformer of interest into one of the categories determined in the above historic dataset.

2.1.4 Literature Review on Pattern Recognition Techniques for Transformer Insulation Diagnosis

This section reviews various pattern recognition techniques for transformer insulation diagnosis. The techniques include artificial neural network (ANN), fuzzy logic, expert system, decision making algorithm, support vector machine (SVM), population based algorithms, and hybrid algorithms.

2.1.4.1 Artificial neural network (ANN)

Artificial neural network has been extensively investigated for transformers insulation diagnosis using oil characteristics [17-35]. During the training process, an ANN adjusts the weights between neurons and the thresholds of activation function of each neuron. In this way, the ANN constructs a

model describing the dependency between the input features and the fault types. Such model is applied to classify any new oil characteristics data into one of the transformer insulation conditions defined in the training process.

Zhang *et al.* proposed a two-stage ANN to diagnose transformer insulation condition using DGA data [17]. By including CO_2 as one of input features, the first stage network can distinguish the faults related to paper insulation from other types of transformer inception faults. Ten folds cross-validation was employed for finding the optimal number of neurons in the hidden layer of ANN. In [18], Vanegas *et al.* proposed a network with two sets of input features: the first set of features is the IEC gas ratios; and the second set is the concentrations of five dissolved gases. The authors demonstrated that the proposed network could attain better classification accuracy by using five gases concentrations as input features. In [19], Guardado *et al.* compared ANNs' efficiency for transformer insulation diagnosis. The neural network was trained based on the diagnosis criteria of five conventional interpretation schemes, i.e. Doernenburg, modified Rogers, Rogers, IEC, and CSUS (this diagnosis criterion uses the individual gas concentrations). The authors reported that a network having three layers with few neurons in the hidden layer could be suitable for transformer insulation incipient faults detection with overall accuracy above 87%. In [20], a multinomial logistic regression model and back-propagation (BP) neural networks were combined to determine fault types of power transformers. The multinomial logistic regression model was applied to find the compositions of the dissolved gases that were correlated to specific types of faults. The concentrations of these gas compositions were fed into BP neural network for training. With such approach, the diagnosis performance can be improved. Some other types of ANNs for transformer insulation diagnosis reported in the literature are: self-organizing polynomial network (SOPN), which heuristically formulated the problem into a hierarchical architecture having multiple layers of low-order polynomials [21]; and a reduced multivariate polynomial (RMP)-based network, in which the determination of its parameters only required an predefined RMP's order and no iterative procedure is needed [22].

To improve the adaptation capability and accelerate the re-training process when new samples (data) become available, the probabilistic neural network (PNN) was adopted in [23]. A fuzzy learning vector quantization (FLVQ) network was also proposed in [24]. The FLVQ network used a fuzzy classifier to segment historical DGA data into different categories of gas ratios. Then for each category of gas ratio, a learning vector quantization (LVQ) network was trained for classifying the types of transformer incipient faults.

For the normal ANNs, they may have some difficulties in determining the network architecture

and hidden layer neurons number. To tackle such limitations, the evolutionary algorithm was integrated into ANNs [25-26]. Such hybrid algorithms can calculate the optimal connection weights and bias terms of the networks simultaneously due to its excellent global search capability. The extension theory based clustering algorithm was also proposed for detecting transformer incipient faults [27-28]. Instead of using the Euclidean distance, the extension distance was used in this method for measuring the similarity between different data points. One significant advantage of this method is that it does not require the tuning of any particular artificial parameters and thus no learning process is needed.

To provide the data visualization capability, the self-organizing map (SOM) based network was proposed in [29]. The SOM can also provide the visualization on the evolution of an incipient fault by plotting the DGA trajectories using the data collected throughout many years. Similar to SOM, a set of auto-associative neural networks were implemented to provide visualization and clustering in transformer insulation diagnosis [30]. The simple *k*-Nearest Neighbour algorithm (*k*NN) and adaptive *k*NN algorithm were also applied for transformer insulation fault diagnosis [31-32]. Moreover, some researchers adopted Bayesian network for transformer fault diagnosis using dissolved gases analysis (DGA) measurement results [33].

The radial basis function (RBF) neural network has a number of advantages such as simpler network structure and better approximation ability. An integrated self-adaptive training based RBF for transformer diagnosis was developed in [34]. This RBF network was constructed by making use of fuzzy c-means (FCM) clustering and quantum-inspired particle swarm optimization (QPSO). Recently, the wavelet has also been incorporated into ANNs for analyzing DGA data. For example, a three layer structured ANN with an evolving wavelet networks (EWNs) [35] and a genetic algorithm tuned wavelet network (GAWN) [36]. In these two wavelet integrated networks, the network parameters and the weighting values were automatically tuned through an evolutionary algorithm based optimization process.

Though they have been widely applied to transformer insulation diagnosis using oil characteristics, ANNs still suffer some inherited drawbacks. For example, it has no explanation ability and requires a relatively large size of historic dataset for training the networks.

2.1.4.2 Fuzzy logic system and expert system

The fuzzy logic approach transforms the experience acquired by human experts into decision

rules and membership functions. The diagnosis on the insulation system of a transformer can be drawn by mapping the oil characteristics of this transformer to a set of rules. A number of researchers applied fuzzy logic techniques to power transformer fault diagnosis [37-40].

In [41], the authors developed a fuzzy logic system for diagnosing transformer insulation and providing recommendation on maintenance actions. In the system, the fuzzy set concept was adopted to manage the uncertainties in key gas analysis, thresholding and gas ratio boundaries. In [42], the authors introduced a framework for conducting transformer diagnostics by applying fuzzy information theory. The fuzzy relations were combined into a decision tree to give the diagnosis on transformer incipient faults. In [13], Su *et al.* proposed a fuzzy logic system to deal with the problem, in which multiple faults simultaneously occur in a transformer insulation system. The proposed system can also indicate the severity of each fault. In [43], the authors adopted the acceptable/unacceptable norms of both key gas concentrations and gas ratios in the implementation of the fuzzy logic system. Their method can interpret the boundary cases, in which a transformer has nearly equal probability of having two different types of faults. In [44], Duraisamy *et al* applied triangular, trapezoidal and Gaussian membership functions to a fuzzy logic system, which was then integrated with back propagation network for diagnosing transformer faults. The conventional IEC/IEEE DGA criteria and the gas concentrations were also used as references to build the fuzzy diagnosis system.

The rule-based expert system has also been applied to transformer diagnosis. Such expert system represents the human experts' knowledge into the forms of IF-THEN rules, which is applied for the evaluation of transformer insulation [45]. The rule-based expert system is also integrated with fuzzy logic to deal with uncertainties in the diagnosing process. For example, Flores *et al.* combined an expert system with a Type-2 fuzzy logic system for evaluating the transformer insulation by using oil characteristics [46]. This hybrid system was able to detect whether paper insulation was involved in any insulation faults and it also allowed other factors as inputs of the pattern recognition algorithm. In [47], Abu-Siada *et al* incorporated several DGA interpretation schemes into a single expert model to overcome individual expert's limitations. It was reported that such approach can improve the diagnosis performance and also pave the way for standardizing DGA interpretation schemes. The major limitations of fuzzy logic and expert system for transformer insulation diagnosis are that the performances are highly decided by the completeness of the pre-defined knowledge base. Neither fuzzy logic system nor expert system can automatically adjust the system parameters when new knowledge is incorporated.

2.1.4.3 Decision-making algorithms

Transformer insulation diagnosis is based on a variety of oil characteristics and is formulated as a multiple-attribute decision-making (MADM) problem. Two commonly adopted methods for solving this problem are evidence reasoning and grey theory.

Tang *et al.* adopted an evidence reasoning algorithm to combine evidences and deal with uncertainties in transformer condition assessment [48]. The algorithm provided the overall evaluation of the transformer condition and ranked several transformers based on their necessities of maintenance. In [49], a fuzzy set theory based algorithm was firstly employed, which provided the diagnosis results as a set of possible types of faults with probability to each type of fault. These diagnoses were then aggregated using an evidential reasoning algorithm. Based on an information fusion strategy, a multi-level and multi-aspect expert system was developed in [50].

The grey model can perform pattern recognition using a relatively small size dataset and without involving formal statistic process and inference. A number of approaches combining grey theory and extension theory were applied to predict the trend of dissolved gas in transformer oil [51-52].

2.1.4.4 Support vector machine and population based algorithms

There are also a number of other types of pattern recognition algorithms which have been applied for interpreting oil characteristics data. For example, the support vector machine (SVM) algorithm and the population based approaches such as particle swarm optimizer (PSO) and genetic programming (GP).

A Parzen windows-based classifier was proposed in [53]. This classifier integrated with a PSO to search for the optimal parameters for Parzen windows based classifier. In [54], genetic programming (GP) and bootstrap were implemented to deal with highly versatile DGA dataset. The bootstrap was used as a pre-processing to make the sample numbers of different fault types equal. Then a GP was applied to extract features, which were subsequently fed into classifiers for transformer insulation diagnosis.

Over the last two decades, support vector machines (SVMs) have been applied to various classification problems. SVM converts the input data from the original space to a higher dimensional space. In [55], the authors implemented a multi-layer SVM classifier and demonstrated

its applicability to fault diagnosis of power transformer. In [56], a clonal selection algorithm is adopted to select the optimal input features and appropriate parameters for a SVM algorithm. In [57], the SVM algorithm was integrated with GA to make the forecasting of gases concentrations in transformer oil. This hybrid algorithm can prevent from over-fitting or under-fitting with the proper selection of SVM parameters using GA. Particle swarm optimization (PSO) can also be integrated into SVM, in which PSO searches the optimal parameters for SVM [58-60]. Some modifications have been made by adopting time-varying acceleration coefficients for improving the PSO convergence in searching the optimal parameters in SVM [61]. Moreover, the artificial immune network classification algorithm (AINC) was also applied to transformer diagnosis [62].

2.1.4.5 Other hybrid algorithms

Attempts have been made to combine ANNs and fuzzy logics or expert systems in transformer insulation diagnosis. Such hybrid algorithm takes the advantages of the learning capability of ANNs, the knowledge formation of expert system, and the uncertainty representation of fuzzy logics. Xu *et al.* proposed a consultative mechanism to combine fuzzy logic and ANN [63]. A combined expert system and ANN was also proposed in [64]. In the above two hybrid algorithms, the knowledge base was derived from IEEE/IEC interpretation schemes and also included the experts' experiences. Miranda *et al.* proposed to combine a neural network into a fuzzy system for the extraction of rules [65]. Some other hybrid approaches proposed in the literature include: integration of conventional DGA interpretation schemes with ANNs and fuzzy logics [66], integrated neural fuzzy approach [67], association rule mining [68], hybrid of fuzzy approach and evidential reasoning based decision-making approach [69]. Other types of hybrid approaches have also been proposed in the literature. For example, in [70] the Dempster–Shafer evidential theory was combined with back propagation neural networks (BP-NN) and fuzzy logic. In [71], Neuro-fuzzy scheme and PNN were integrated. Table 2.1 summarizes the data configuration (total samples number, training dataset size, testing dataset size, number of input features, and the types of transformer insulation faults in terms of fault classes) as well as the classification accuracies of different pattern recognition algorithms presented in the literature. This table does not intend to supply an exhaustive survey but aims to provide a clear picture of the setup and performances of some state-of-the-art pattern recognition algorithms for transformer insulation diagnosis.

Table 2.1 Dataset configurations and results of some representative pattern recognition algorithms in the literature

Ref	Model	Samples	Training	Testing	# Features	# Classes	Accuracy (%)*
[17]	Two-step ANN (BP)	40	N/A	N/A	5	4	86-95
[18]	ANN	26	N/A	N/A	5/8/3	3	73-96
[19]	ANN- BP	150	117	33	3/5	8	87-100
[21]	Polynomial Networks (BP)	711	N/A	N/A	3/5	8	87-97
[22]	Multivariate polynomial neural network	167	156	11	6	N/A	100
[23]	PNN	503	497	4	4	8	100
[24]	Fuzzy LVQ	711	N/A	N/A	3	8	97
[25]	Genetic-based neural networks	630	N/A	N/A	3/5	5	91-95
[26]	Evolving neural nets	820	N/A	N/A	3/5	5	90-93
[28]	Extension theory and ANN	22	N/A	N/A	5	9	96-100
[30]	Auto associative neural networks	352	N/A	N/A	3	5	100
[41]	Expert system	101	N/A	N/A	9	3	93
[72]	Evolutionary fuzzy logic	711	N/A	N/A	3	8	92
[73]	Adaptive fuzzy logic	561	N/A	280	N/A	5	94
[43]	Fuzzy logic	20	N/A	N/A	9	9	100
[51]	Grey prediction	46	N/A	N/A	10	10	97
[53]	Parzen–Windows and PSO	168	N/A	N/A	3/3/8	4	80
[55]	SVM	75	50	25	5	4	100
[64]	ANN and ES	210	150	60	24	6	93-96
[65]	Knowledge Extraction and ANN	318	N/A	88	3	5	99
[74]	Clustering and extension theory	21	N/A	N/A	8	9	88
[75]	SVM and genetic algorithm	142	N/A	N/A	5	4	94

*Accuracy is computed as the ratio between the number of transformers that are correctly classified in the testing sub-dataset and the total number of transformers in the testing sub-dataset.

2.1.5 Key Issues of Developing the "Ready-To-Use" Pattern Recognition Algorithms for Transformer Diagnosis

A power transformer is constructed with the complex combination of different materials and its operation is extensively complicated, it is not a trivial task to make assessment on transformer insulation system using oil characteristics. Pattern recognition based interpretation scheme is still not a ready-to-use tool for utilities. There are several key issues need to be further investigated.

The first issue is the evaluation of the performance of pattern recognition algorithms in diagnosing transformer insulation faults. It can be observed from Table 2.1, there was lack of a

common framework for defining the process of training, cross-validation, testing, and evaluation to ensure the generalizability and applicability of various algorithms as well as for providing the statistical comparisons amongst different algorithms.

The second issue is how to build up a statistically satisfied training database. It is well known most pattern recognition algorithms require a large size of historic dataset for training. However, the occurrence of transformer fault is an event with relatively small probability. Therefore, the historic dataset may only consist of few records of certain fault types. If an algorithm is trained by such database, it cannot make reliable insulation diagnosis. In addition, the process of collecting oil samples, conducting dissolved gas measurement and interpreting the measured data may vary among different utilities. This may compromise the generalization capability of pattern recognition algorithms, i.e. some algorithms trained on a "local" dataset cannot readily applied "globally".

The third issue is the data quality issue. Because of the complex geometry of transformer insulation system, limitation of the measurement system, and possible presence of multiple faults in transformer insulation system, there may exist inconsistence in reaching diagnosis of transformer insulation.

The next section will presents an illustrative case study to address procedures of dataset preparation, training, cross-validation, and testing for applying pattern recognition algorithm for transformer insulation diagnosis. The statistical comparison of different pattern recognition algorithms will be addressed in the accompanying paper (i.e. Part 2 of this paper) through extensive case studies [16]. The research directions for solving the above second and third issues will also be discussed in the accompanying paper [16].

2.1.6 An Illustrative Case Study- By Generalized Regression Neural Network (GRNN)

In this case study, the generalized regression neural network (GRNN) is applied to make transformer insulation diagnosis. GRNN has been widely used in applications involving classification and predication. It can learn the underlying functions between input and output (e.g. class) from samples without having the prior knowledge of any specific function form between input and output [76]. Thus, GRNN has the advantages of having simple structure and less computation time.

In the case study, a database consisting of oil characteristics with the already-known insulation conditions of the corresponding transformers is used for training, validation and testing GRNN. Table 2.2 presents the configuration of such a database used in this case study. The original data in this database was obtained from a utility company. In Table 2.2, the features (input of algorithms) are the concentrations of dissolved gases (i.e. hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2)) and oil test results of water content, acidity, dielectric dissipation factor, 2-furfuraldehyde, resistivity and breakdown voltage. The insulation condition is categorized into four different classes (categories) by combining the diagnosis results obtained from fuzzy *c*-means clustering algorithm, Duval triangle method, and utility experts' assessment. The readers may refer to the authors' previous publication for more details of constructing this training database [77].

Table 2.2 Configuration of database using oil characteristics

Sample	**Feature**	**Class (condition of transformer insulation)**			
181	11	Excellent	Good	Fair	Poor
		80	50	21	30

The procedure of training, cross-validation, and testing of GRNN algorithm (other pattern recognition algorithms as well) using the database in Table 2.2 is as follows:

1) Normalizing all data into the range of [0, 1].

2) Randomly splitting database into a training sub-dataset and testing sub-dataset by 70% and 30% (this ratio can be adjusted based on the user's requirements).

3) Performing ten-fold cross-validation on training sub-dataset to obtain the optimal parameters of the algorithm.

4) Training the algorithm with the optimal parameters obtained in 3).

5) Testing the trained algorithm on the test sub-dataset (e.g. unseen data by the algorithm).

6) Obtaining the classification accuracy.

The above procedure will be repeated 50 times to obtain the statistical performance evaluation on the algorithm. The averaged classification accuracy on each class and the overall classification accuracy over 50 runs of the algorithm are adopted as the performance evaluation criteria.

Figure 2.2 depicts the samples distribution of the database as described in Table 2.2. It is obvious

that this database has unequal distributed samples in different classes, e.g. 80 samples belong to "Excellent" class while only 21 samples belong to "Fair" class. This will introduces significant difficulties for algorithms to make correct classification on the minority class. Some re-sampling or data balancing techniques can be applied to address this issue [10].

After establishing the training sub-dataset, ten-fold cross validation is applied to obtain the optimal parameters of GRNN. The procedure of the ten-fold cross validation is as follows: (1) the original training sub-dataset with N samples are divided into 10 groups of size N /10; (2) the algorithm is trained on nine groups and tested on one group; and (3) it is performed for 10 times and the mean accuracy is taken as the classification accuracy of the algorithm. The above procedure was repeated for each possible value of algorithm's parameters, and the parameters with the highest mean accuracy in the above step (3) are taken as the best parameters. After finding the optimal values of the above parameters, the algorithm is trained on the whole training sub-dataset. Subsequently, the trained algorithm is applied to evaluate the conditions of transformers in the testing sub-dataset. At this stage, the algorithm automatically interprets the oil characteristics data of the transformers in the testing sub-dataset and classifies these transformers into particular categories of insulation conditions such as Excellent, Good, Fair and Poor condition. The above optimal parameter acquisition, training and testing procedure is repeated for 50 times (runs). At each run, the classification accuracy is computed according to $N_{total}^{'}/N_{total}$, where $N_{total}^{'}$ denotes the number of transformers that are correctly classified and N_{total} denotes the total number of transformers in the testing sub-dataset. The overall accuracy is the average over 50 runs. Table 2.3 presents the classification accuracy (averaged over 50 runs) of the GRNN algorithm.

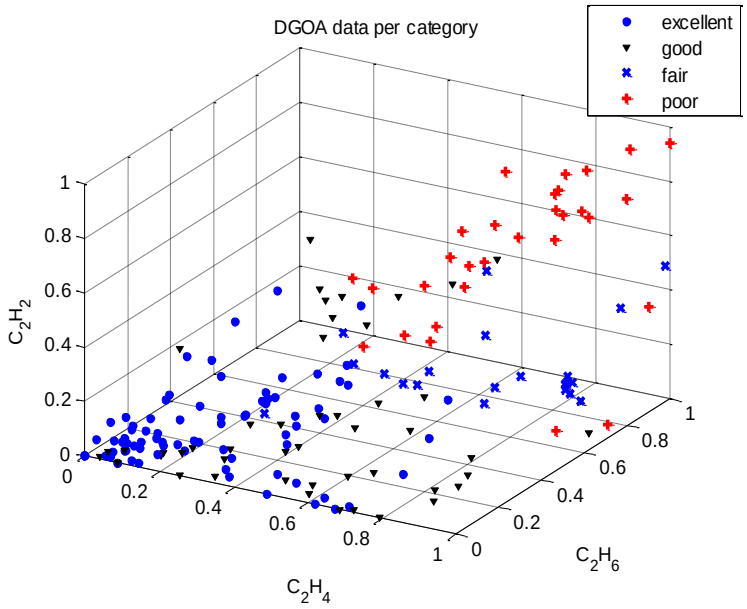


Figure 2.2 Samples distribution of the database in Table 2.2

Table 2.3 Classification accuracy of GRNN algorithm for transformer insulation diagnosis using oil characteristics

Excellent	Good	Fair	Poor	Overall
96%	90%	75%	83%	90%

It can be observed from Table 2.3 that the GRNN algorithm can correctly recognize faults occurring in most transformers in the testing dataset. The overall classification accuracy is 90% averaged over 50 runs. However, the fault classification accuracy of transformers with "Fair" insulation condition may not be satisfied (about 75%). This is because the original database (Table 2.2) is imbalanced, in which the transformers with "Fair" condition are outnumbered by the transformers with other three conditions (Excellent, Good and Poor). Trained by such an imbalanced database, the algorithm can lead to misclassification on the class of transformers with "Fair" condition. To deal with such problem, some pre-processing methods can be integrated with the algorithm to facilitate it achieving consistent desirable classification accuracy [10, 77].

2.1.7 Conclusion

This paper studied the pattern recognition techniques and their application for power transformer insulation diagnosis using oil characteristics. A general pattern recognition application framework was presented in the paper. A comprehensive literature review on the state-of-the-art pattern recognition techniques for transformer insulation diagnosis was also provided. Moreover, a case study was presented to illustrate the process of oil characteristics database preparation, training/testing datasets formation, cross-validation, testing, and evaluation in applying pattern recognition algorithm for diagnosing transformer insulation.

Acknowledgment

We gratefully acknowledge Australian Research Council, Powerlink Queensland, Energex, Ergon Energy, and TransGrid for providing supports for this work.

References

[1] M. Wang, A. J. Vandermaar and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electr. Insul. Mag.*, vol.18, Issue 6, pp. 12-25, 2002.

[2] A. Naderian, R. Piercy, S. Cress, J. Service and W. Fan, "An approach to power transformer asset management using health index," *IEEE Electr. Insul. Mag.*, vol.25, Issue 2, pp. 20-34, 2009.

[3] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.10, Issue 5, pp. 903-917, 2003.

[4] T. K. Saha and P. Purkait, "Investigation of polarization and depolarization current measurements for the assessment of oil-paper insulation of aged transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.11, Issue 1, pp. 144-154, 2004.

[5] M. Duval, "New techniques for dissolved gas-in-oil analysis," *IEEE Electr. Insul. Mag.*, vol.19, Issue 2, pp. 6-15, 2003.

[6] Z. Wang, J. Li and D. M. Sofian, "Interpretation of transformer FRA responses-part I: influence of winding structure," *IEEE Trans. Power Delivery*, vol.24, Issue 2, pp. 703-710, 2009.

[7] C. Ekanayake, S. M. Gubanski, A. Graczkowski and K. Walczak, "Frequency response of oil impregnated pressboard and paper samples for estimating moisture in transformer insulation," *IEEE Trans. Power Delivery*, vol.21, Issue 3, pp. 1309-1317, 2006.

[8] S. M. Strachan, S. Rudd, S. D. J. McArthur, M. D. Judd, S. Meijer and E. Gulski, "Knowledge-based diagnosis of partial discharges in power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.15, Issue 1, pp. 259-268, 2008.

[9] M. Duval and A. DePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electr. Insul. Mag.*, vol.17, Issue 2, pp. 31-41, 2001.

[10] H. Ma, C. Ekanayake and T. K. Saha, "Power transformer fault diagnosis under measurement originated uncertainties," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 6, pp. 1982-1990, 2012.

[11] I. N. Da Silva, M. M. Imamura and A. N. de Souza, "Application of neural networks to the analysis of dissolved gases in insulating oil used in transformers," in *Proceedings of 2000 IEEE International Conference on Systems, Man and Cybernetics, October 8-11*, 2000,Nashville, TN, USA, pp. 2643-2648.

[12] Z. Xu, "Multivalued logic and fuzzy logic - their relationship, minimization, and application to fault diagnosis.," *IEEE Trans. Comput.*, vol.C-33, Issue 7, pp. 679-681, 1984.

[13] Q. Su, C. Mi, L. L. Lai and P. Austin, "Fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer," *IEEE Trans. Power Syst.*, vol.15, Issue 2, pp. 593-598, 2000.

[14] R. Naresh, V. Sharma and M. Vashisth, "An integrated neural fuzzy approach for fault diagnosis of transformers," *IEEE Trans. Power Delivery*, vol.23, Issue 4, pp. 2017-2024, 2008.

[15] W. Chen, C. Pan, Y. Yun and Y. Liu, "Wavelet networks in power transformers diagnosis using dissolved gas analysis," *IEEE Trans. Power Delivery*, vol.24, Issue 1, pp. 187-194, 2009.

[16] Y. Cui, H. Ma and T. K. Saha, "Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 2: Implementation, Case Study, and Statistical Analysis," *Int. Trans. Electr. Energy. Syst.*, vol.25, Issue 10, pp. 2260-2274, 2015.

[17] Y. Zhang, X. Ding, Y. Liu and P. J. Griffin, "An artificial neural network approach to transformer fault diagnosis," *IEEE Trans. Power Delivery*, vol.11, Issue 4, pp. 1836-1841, 1996.

[18] O. Vanegas, Y. Mizuno, K. Naito and T. Kamiya, "Diagnosis of oil-insulated power apparatus by using neural network simulation," *IEEE Trans. Dielectr. Electr. Insul.*, vol.4, Issue 3, pp. 290-299, 1997.

[19] J. L. Guardado, J. L. Naredo, P. Moreno and C. R. Fuerte, "A comparative study of neural network efficiency in power transformers diagnosis using dissolved gas analysis," *IEEE Trans. Power Delivery*, vol.16, Issue 4, pp. 643-647, 2001.

[20] M. T. Yang and L. S. Hu, "Intelligent fault types diagnostic system for dissolved gas analysis of oil-immersed power transformer," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 6, pp. 2317-2324, 2013.

[21] H. T. Yang and Y. C. Huang, "Intelligent decision support for diagnosis of incipient transformer faults using self-organizing polynomial networks," *IEEE Trans. Power Syst.*, vol.13, Issue 3, pp. 946-952, 1998.

[22] H. Wu, X. Li and D. Wu, "RMP neural network based dissolved gas analyzer for fault diagnostic of oil-filled electrical equipment," *IEEE Trans. Dielectr. Electr. Insul.*, vol.18, Issue 2, pp. 495-498, 2011.

[23] W. M. Lin, C. H. Lin and M. X. Tasy, "Transformer-fault diagnosis by integrating field data and standard codes with training enhancible adaptive probabilistic network," *IEE Proceedings-Generation Transmission and Distribution*, vol.152, Issue 3, pp. 335-341, 2005.

[24] H. Yang, C. Liao and J. Chou, "Fuzzy learning vector quantization networks for power transformer condition assessment," *IEEE Trans. Dielectr. Electr. Insul.*, vol.8, Issue 1, pp. 143-149, 2001.

[25] Y. C. Huang, "Condition assessment of power transformers using genetic-based neural networks," *IEE Proceedings-Science Measurement and Technology*, vol.150, Issue 1, pp. 19-24, 2003.

[26] Y. C. Huang, "Evolving neural nets for fault diagnosis of power transformers," *IEEE Trans. Power Delivery*, vol.18, Issue 3, pp. 843-848, 2003.

[27] M. H. Wang, "Extension neural network for power transformer incipient fault diagnosis," *IEE Proceedings-Generation Transmission and Distribution*, vol.150, Issue 6, pp. 679-685, 2003.

[28] M. H. Wang, "A novel extension method for transformer fault diagnosis," *IEEE Trans. Power Delivery*, vol.18, Issue 1, pp. 164-169, 2003.

[29] K. F. Thang, R. K. Aggarwal, A. J. McGrail and D. G. Esp, "Analysis of power transformer dissolved gas data using the self-organizing map," *IEEE Trans. Power Delivery*, vol.18, Issue 4, pp. 1241-1248, 2003.

[30] V. Miranda, A. Castro and S. Lima, "Diagnosing faults in power transformers with autoassociative neural networks and mean shift," *IEEE Trans. Power Delivery*, vol.27, Issue 3, pp. 1350-1357, 2012.

[31] S. Sun and R. Huang, "An adaptive k-nearest neighbor algorithm," in *Proceedings of Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), August 10-12*, 2010,Yantai, Shandong, China, pp. 91-94.

[32] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol.13, Issue 1, pp. 21-27, 1967.

[33] A. Carita, L. C. Leite, A. Junior, R. B. Godoy and L. Sauer, "Bayesian networks applied to failure diagnosis in power transformer," *IEEE Lat. Am. Trans.*, vol.11, Issue 4, pp. 1075-1082, 2013.

[34] K. Meng, Z. Y. Dong, D. H. Wang and K. P. Wong, "A self-adaptive rbf neural network classifier for transformer fault analysis," *IEEE Trans. Power Syst.*, vol.25, Issue 3, pp. 1350-1360, 2010.

[35] Y. C. Huang and C. M. Huang, "Evolving wavelet networks for power transformer condition monitoring," *IEEE Trans. Power Delivery*, vol.17, Issue 2, pp. 412-416, 2002.

[36] Y. C. Huang, "A new data mining approach to dissolved gas analysis of oil-insulated power apparatus," *IEEE Trans. Power Delivery*, vol.18, Issue 4, pp. 1257-1261, 2003.

[37] Y. C. Huang and H. C. Sun, "Dissolved gas analysis of mineral oil for power transformer fault diagnosis using fuzzy logic," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 3, pp. 974-981, 2013.

[38] N. K. Dhote and J. B. Helonde, "Fuzzy algorithm for power transformer diagnostics," *Advances in Fuzzy Systems*, vol.2013, Issue 1, pp. 1-7, 2013.

[39] I. G. David, M. Rajaram and A. Gnanasaravanan, "Power transformer faults identification using fuzzy based dissolved gas analysis method," *Journal of Electrical Engineering*, vol.13, Issue 1, pp. 1-5, 2013.

[40] N. K. Dhote and J. B. Helonde, "Improvement in transformer diagnosis by DGA using fuzzy logic," *J. Electr. Eng. Technol.*, vol.9, Issue 2, pp. 615-621, 2014.

[41] C. E. Lin, J. M. Ling and C. L. Huang, "An expert system for transformer fault-diagnosis using dissolved-gas analysis," *IEEE Trans. Power Delivery*, vol.8, Issue 1, pp. 231-238, 1993.

[42] K. Tomsovic, M. Tapper and T. Ingvarsson, "A fuzzy information approach to integrating different transformer diagnostic methods," *IEEE Trans. Power Delivery*, vol.8, Issue 3, pp. 1638-1646, 1993.

[43] S. M. Islam, T. Wu and G. Ledwich, "Novel fuzzy logic approach to transformer fault diagnosis," *IEEE Trans. Dielectr. Electr. Insul.*, vol.7, Issue 2, pp. 177-186, 2000.

[44] V. Duraisamy, N. Devarajan, D. Somasundareswari, A. Vasanth and S. N. Sivanandam, "Neuro fuzzy schemes for fault detection in power transformer," *Appl. Soft Comput.*, vol.7, Issue 2, pp. 534-539, 2007.

[45] J. L. Kirtley, W. H. Hagman, B. C. Lesieutre, M. J. Boyd, E. P. Warren and H. P. Chou, *et al.*, "Monitoring the health of power transformers," *IEEE Comput. Appl. Power*, vol.9, Issue 1, pp. 18-23, 1996.

[46] W. C. Flores, E. E. Mombello, J. A. Jardini, G. Ratt á and A. M. Corvo, "Expert system for the assessment of power transformer insulation condition based on type-2 fuzzy logic systems," *Expert Syst. Appl.*, vol.38, Issue 7, pp. 8119-8127, 2011.

[47] A. Abu-Siada, S. Hmood and S. Islam, "A new fuzzy logic approach for consistent interpretation of dissolved gas-in-oil analysis," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 6, pp. 2343-2349, 2013.

[48] W. H. Tang, K. Spurgeon, Q. H. Wu and Z. J. Richardson, "An evidential reasoning approach to transformer condition assessments," *IEEE Trans. Power Delivery*, vol.19, Issue 4, pp. 1696-1703, 2004.

[49] K. Spurgeon, W. H. Tang, Q. H. Wu, Z. J. Richardson and G. Moss, "Dissolved gas analysis using evidential reasoning," *IEE Proceedings-Science Measurement and Technology*, vol.152, Issue 3, pp. 110-117, 2005.

[50] X. L. Wang, Q. M. Li, C. R. Li, R. Yang and Q. Su, "Reliability assessment of the fault diagnosis methodologies for transformers and a new diagnostic scheme based on fault info integration," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 6, pp. 2292-2298, 2013.

[51] M. H. Wang, "Grey-extension method for incipient fault forecasting of oil-immersed power transformer," *Electr. Pow. Compo. Sys.*, vol.32, Issue 10, pp. 959-975, 2004.

[52] M. H. Wang and C. P. Hung, "Novel grey model for the prediction of trend of dissolved gases in oil-filled power apparatus," *Electr. Power Syst. Res.*, vol.67, Issue 1, pp. 53-58, 2003.

[53] W. H. Tang, J. Y. Goulermas, Q. H. Wu, Z. J. Richardson and J. Fitch, "A probabilistic classifier for transformer dissolved gas analysis with a particle swarm optimizer," *IEEE Trans. Power Delivery*, vol.23, Issue 2, pp. 751-759, 2008.

[54] A. Shintemirov, W. Tang and Q. H. Wu, "Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol.39, Issue 1, pp. 69-79, 2009.

[55] G. Lv, H. Cheng, H. Zhai and L. Dong, "Fault diagnosis of power transformer based on multi-layer SVM classifier," *Electr. Power Syst. Res.*, vol.75, Issue 1, pp. 9-15, 2005.

[56] M. Cho, T. Lee, S. Gau and C. Shih, "Power transformer fault diagnosis using support vector machines and artificial neural networks with clonal selection algorithms optimization," in Knowlegde-Based Intelligent Information and Engineering Systems, vol.4251, B. Gabrys, R. J. Howlett and L. C. Jain Ed. Berlin: Springer Berlin Heidelberg, 2006, pp. 179-186.

[57] S. Fei and Y. Sun, "Forecasting dissolved gases content in power transformer oil based on support vector machine with genetic algorithm," *Electr. Power Syst. Res.*, vol.78, Issue 3, pp. 507-514, 2008.

[58] R. Liao, H. Zheng, S. Grzybowski, L. Yang, C. Tang and Y. Zhang, "Fuzzy information granulated particle swarm optimisation-support vector machine regression for the trend forecasting of dissolved gases in oil-filled transformers," *IET Electr. Power Appl.*, vol.5, Issue 2, pp. 230-237, 2011.

[59] T. Lee, M. Cho and F. Fang, "Features selection of SVM and ANN using particle swarm optimization for power transformers incipient fault symptom diagnosis," *International Journal of Computational Intelligence Research*, vol.3, Issue 1, pp. 60-65, 2007.

[60] T. Lee, M. Cho, C. Shieh, H. Lee and F. Fang, "Particle swarm optimization-based SVM for incipient fault classification of power transformers," in Foundations of Intelligence Systems, vol.4203, F. Esposito, Z. W. Ras, D. Malerba and G. Semeraro Ed. Berlin: Springer Berlin Heidelberg, 2006, pp. 84-90.

[61] R. J. Liao, H. B. Zheng, S. Grzybowski and L. J. Yang, "A multiclass SVM-based classifier for transformer fault diagnosis using a particle swarm optimizer with time-varying acceleration coefficients," *Int. Trans. Electr. Energy. Syst.*, vol.23, Issue 2, pp. 181-190, 2013.

[62] H. Xiong and C. Sun, "Artificial immune network classification algorithm for fault diagnosis of power transformer," *IEEE Trans. Power Delivery*, vol.22, Issue 2, pp. 930-935, 2007.

[63] W. Xu, D. Wang, Z. Zhou and H. Chen, "Fault diagnosis of power transformers: Application of fuzzy set theory, expert systems and artificial neural networks," *IEE Proceedings-Science Measurement and Technology*, vol.144, Issue 1, pp. 39-44, 1997.

[64] Z. Wang, Y. Liu and P. Griffin, "A combined ANN and expert system tool for transformer fault diagnosis," *IEEE Trans. Power Delivery*, vol.13, Issue 4, pp. 1224-1229, 1998.

[65] V. Miranda and A. Castro, "Improving the IEC table for transformer failure diagnosis with knowledge extraction from neural networks," *IEEE Trans. Power Delivery*, vol.20, Issue 4, pp. 2509-2516, 2005.

[66] D. Morais and J. Rolim, "A hybrid tool for detection of incipient faults in transformers based on the dissolved gas analysis of insulating oil," *IEEE Trans. Power Delivery*, vol.21, Issue 2, pp. 673-680, 2006.

[67] R. Naresh, V. Sharma and M. Vashisth, "An integrated neural fuzzy approach for fault diagnosis of transformers," *IEEE Trans. Power Delivery*, vol.23, Issue 4, pp. 2017-2024, 2008.

[68] Z. Yang, W. Tang, A. Shintemirov and Q. Wu, "Association rule mining-based dissolved gas analysis for fault diagnosis of power transformers," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol.39, Issue 6, pp. 597-610, 2009.

[69] R. Liao, H. Zheng, S. Grzybowski, L. Yang, Y. Zhang and Y. Liao, "An integrated decision-making model for condition assessment of power transformers using fuzzy approach and evidential reasoning," *IEEE Trans. Power Delivery*, vol.26, Issue 2, pp. 1111-1118, 2011.

[70] D. Bhalla, R. K. Bansal and H. O. Gupta, "Integrating AI based DGA fault diagnosis using dempster-shafer theory," *Int. J. Electr. Power Energy Syst.*, vol.48, Issue31-38, 2013.

[71] H. Malik, A. K. Yadav, S. Mishra and T. Mehto, "Application of neuro-fuzzy scheme to investigate the winding insulation paper deterioration in oil-immersed power transformer," *Int. J. Electr. Power Energy Syst.*, vol.53, Issue 1, pp. 256-271, 2013.

[72] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Networks*, vol.2, Issue 6, pp. 568-576, 1991.

[73] A. D. Ashkezari, H. Ma, T. K. Saha and C. Ekanayake, "Application of fuzzy support vector machine for determining the health index of the insulation system of in-service power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 3, pp. 965-973, 2013.

[74] Y. C. Huang, H. T. Yang and C. L. Huang, "Developing a new transformer fault diagnosis system through evolutionary fuzzy logic," *IEEE Trans. Power Delivery*, vol.12, Issue 2, pp. 761-767, 1997.

[75] H. T. Yang and C. C. Liao, "Adaptive fuzzy diagnosis system for dissolved gas analysis of power transformers," *IEEE Trans. Power Delivery*, vol.14, Issue 4, pp. 1342-1350, 1999.

[76] M. Wang, Y. Tseng, H. Chen and K. Chao, "A novel clustering algorithm based on the extension theory and genetic algorithm," *Expert Syst. Appl.*, vol.36, Issue 4, pp. 8269-8276, 2009.

[77] S. Fei and X. Zhang, "Fault diagnosis of power transformer based on support vector machine with genetic algorithm," *Expert Syst. Appl.*, vol.36, Issue 8, pp. 11352-11357, 2009.

2.2 Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 2: Implementation, Case Study, and Statistical Analysis

Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 2: Implementation, Case Study, and Statistical Analysis

Yi Cui, Hui Ma, and Tapan Saha

The University of Queensland, Brisbane, Australia

Publication Journal: International Transactions on Electrical Energy Systems

Submitted: March 14, 2014

Revision Resubmitted: May 22, 2014

Accepted: May 26, 2014

Published: July 17, 2014

Author's Contributions:

Yi Cui: Theoretical model design, numeric experimental simulation, results analysis and manuscript preparation.

Hui Ma: Supervision on the numeric experiment, results discussion and manuscript revision.

Tapan Saha: Supervision on the numeric experiment, results discussion and manuscript revision.

Abstract

Transformer oil tests such as breakdown voltage, resistivity, dielectric dissipation factor, water content, 2-furfuraldehyde, acidity and different dissolved gases have been adopted in utility companies for evaluating the conditions of transformer insulation. Over the past twenty years, various pattern recognition techniques have been applied for power transformer insulation diagnosis using oil tests results (oil characteristics). This paper investigates a variety of state-of-the-art pattern recognition algorithms for transformer insulation diagnosis. To verify the applicability and generalization capability of different pattern recognition algorithms, this paper implements 15 representative algorithms and conducts extensive case studies on eight oil characteristics datasets collected from different utility companies. A statistical performance (in terms of classification accuracy) comparison amongst different pattern recognition algorithms for transformer insulation diagnosis using oil characteristics is also conducted in the paper.

Index Terms: Dissolved gas analysis, insulation, oil characteristics, pattern recognition, power transformer

2.2.1 Introduction

Over the past ten years, a variety of pattern recognition techniques have been applied to power transformer insulation diagnosis using oil characteristics, which consists of oil tests results (i.e. breakdown voltage, resistivity, dielectric dissipation factor, water content, 2-furfuraldehyde, and acidity) and the concentrations of dissolved gases (i.e. acetylene, ethylene, methane, ethane, hydrogen, carbon monoxide and carbon dioxide). In the previous paper [1], a comparative literature survey of various pattern recognition techniques with the emphasis on their applications in transformer insulation diagnosis is presented. The important issues of making the pattern recognition techniques as practical tools for utilities were also addressed in the previous paper. The three most challenging issues are: (1) defining a common framework for training, cross-validation, testing, and statistical performance comparisons of different pattern recognition algorithms for their applicability in transformer insulation diagnosis; (2) constructing a statistically satisfied training database to improve the generalizability and applicability of the pattern recognition algorithms; and (3) investigating the effective methods for handling the data quality problem, in which the samples in one category of insulation conditions are significantly outnumbered by those in another category

or the samples are contaminated by noise [1-2].

This paper addresses the above issues through the implementation of 15 state-of-the-art pattern recognition algorithms. The mathematical formulation of these pattern recognition algorithms will be presented in Section 2.2.2. The approaches of improving the data quality in training dataset will be discussed in Section 2.2.3. The case studies and the discussion on results are provided in Section 2.2.4. The statistical comparisons amongst different algorithms are also given in Section 2.2.4. The conclusion is provided in Section 2.2.5.

2.2.2 Pattern Recognition Algorithms for Transformer Insulation Diagnosis

This section provides a brief review of 15 pattern recognition algorithms, which will be applied for transformer insulation diagnosis in this paper. These algorithms are: multiple layer perceptron (MLP), probability neural network (PNN), learning vector quantization (LVQ), *K*-nearest neighbors (*K*NN), Bayesian classifier, support vector machine (SVM) with three different types of kernels, decision tree, fuzzy support vector machine (FSVM), radial basis function (RBF) network, generalized regression neural network (GRNN), fuzzy logic, and self-organizing tree algorithm (SOTA).

2.2.2.1 Multiple layer perceptron (MLP)

MLP networks have been adopted to various applications. In a MLP network, the input and hidden nodes are connected by one set of weights while the hidden nodes and outputs are also connected by another set of weights [3]. Figure 2.3 depicts the architecture of MLP.

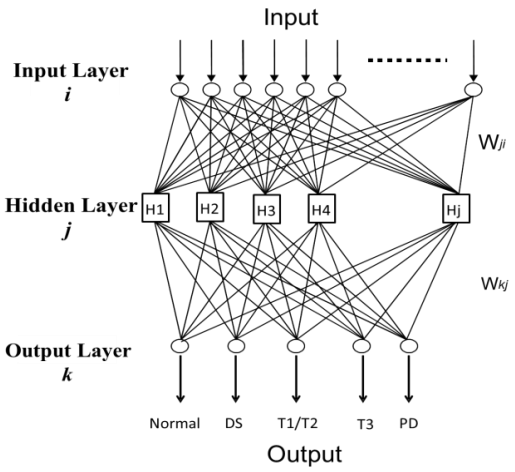


Figure 2.3 Architecture of MLP network

A tanh activation function is adopted to determine the values of hidden layer nodes. This function takes effect on *M* linear combinations of the inputs as:

$$y_i^m = \tanh(\sum_{j=1}^{d} w_i^{mj} x_i^j + b_i^m), i = 1,\ldots N; m = 1\ldots M \tag{2.1}$$

where x_i^j denotes the *j*-th element in x_i, which is the *i*-th input, w_i^{mj} is the weight of the *m*-th hidden node with respect to x_i^j, b_i^m denotes the bias, and *M* is hidden nodes number. A softmax activation function will then be used to transform y_i^m to the final output:

$$z_k = \frac{\exp(\sum_{m=1}^{M} w_i^{km} y_i^m + b_i^k)}{\sum_{k'} \exp(\sum_{m=1}^{M} w_i^{k'm} y_i^m + b_i^{k'})}, k = 1,\ldots K \tag{2.2}$$

where w_i^{km} is the weight between the output node and the hidden node, and *K* denotes classes number.

The back-propagation technique is adopted in MLP training, which computes the derivatives for the error function regarding the weights and biases in (2.1) and (2.2) [3].

2.2.2.2 Probability neural network (PNN)

The architecture of PNN is shown in Figure 2.4 [4]. Upon receiving the input $\mathrm{x} = [x_1 \ldots x_d]$, the pattern node produces its output as

$$\vartheta(\mathrm{x}) = \frac{1}{(2\pi)^{d/2}\sigma^d} \exp\left[-\frac{(\mathrm{x}-\mathrm{y}_{ij})^T(\mathrm{x}-\mathrm{y}_{ij})}{2\sigma^2}\right] \tag{2.3}$$

where *d* denotes dimension of the input, σ is the variance of Gaussian distribution, and y_{ij} is the pattern node.

The summation node adds up and averages the outputs of all hidden nodes to obtain the probability of *X* being classified into one of the classes C_i as

$$p_i(x) = \frac{1}{(2\pi)^{d/2}\sigma^d} \frac{1}{N_i} \sum_{i=1}^{N_i} \exp\left[-\frac{(x-y_{ij})^T(x-y_{ij})}{2\sigma^2}\right] \tag{2.4}$$

where N_i is the samples number in classes C_i .

Finally, the output node makes classification according to the results from the summation nodes

$$z(x) = \arg\max\{ p_i(x) \}, \qquad i = 1,2,...,m \tag{2.5}$$

where z is the predicted class for input *x* and *m* is the classes number in the original training database.

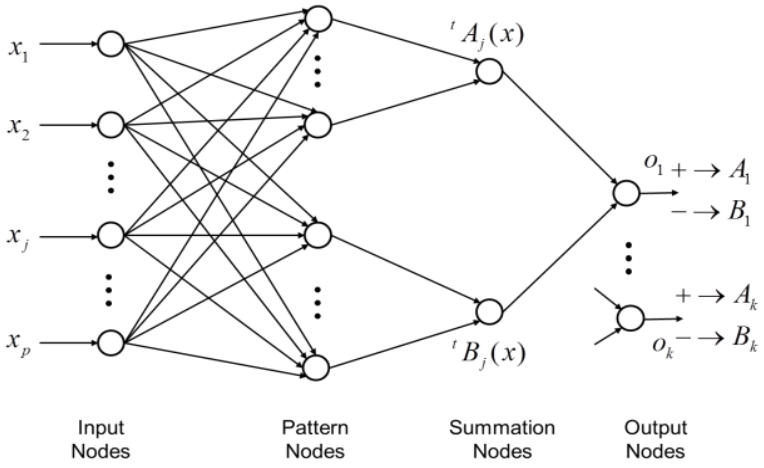


Figure 2.4 Architecture of PNN network

2.2.2.3 Learning vector quantization (LVQ)

LVQ combines competitive learning with supervision. Instead of computing the probability of the input with respect to different classes, LVQ constructs the discrimination function by computing the distances between the input vectors and the prototype vectors. The input *x* is assigned to a class based on the label of its closest prototype vector [5].

At each training step, the input sample x_i is randomly drawn from the input dataset. The distance between x_i and the prototype vector m_k is computed using some distance measure (for example the Euclidian distance). The LVQ is trained iteratively to update a winner unit m_c , which is

$$m_c = arc\min\{ \| x_i - m_k \| \} \tag{2.6}$$

where $\|.\|$ denotes the Euclidian distance. The update equation for the above winner unit is

$$m_c(t+1) = m_c(t) \pm \alpha\ (t)\ \ [\ x_i(t) - m_c\ (t)] \tag{2.7}$$

where $\alpha(t) \in [0, 1]$ is the learning rate and it is a monotonically decrease function, and the sign is decided by the data classification results. For different samples selected from the training dataset, the above procedure is repeated until the convergence is attained.

The above iteration algorithm is known as LVQ1. Instead of adopting one prototype vector, another type of LVQ algorithm, LVQ2 adopts two prototype vectors and updates these two prototype vectors.

2.2.2.4 K-nearest neighbors (KNN)

The *K*NN algorithm firstly computes the distances between a new data point **y*** to each data point in the training dataset [3]. Then *K*NN assigns **y*** to the class *t*, which has the largest frequency in S_k, the subset consisting of the *K* closest data points

$$K_t > K_l, \text{ for } \quad l = 1, 2, \ldots T \tag{2.8}$$

where K_l denotes the frequency of the *l*-th class in S_k.

2.2.2.5 Bayesian classifier

Bayesian classifier computes the probability of a sample x_i with regard to a class $C_k, P(C_k \mid x_i)$ through the Bayes' theorem [3]:

$$P(C_k \mid x_i) = \frac{p(x_i \mid c_k) p(c_k)}{p(x)} \tag{2.9}$$

where $p(c_k)$ is the class prior probability, $p(x) = \sum_{k=1}^{K} p\{x_i \mid C_k\} P(C_k)$ is a normalization factor, and $p(x_i \mid C_k)$ is computed through [3]:

$$p(x_i \mid C_k) \approx \sum_{j=1}^{\theta} p(x_i \mid j) P(j) \tag{2.10}$$

where θ is the number of Gaussian mixture elements, $p(x_i \mid j)$ is the probability distribution of these mixture elements that are Gaussian (mean u_j and covariance $\sum j$, and $P(j)$ are the weights

and satisfy $0 < P(j) < 1$ and $\sum_{j=1}^{\theta} P(j) = 1$.

The parameters $\theta_j = \{\mu_j, \sum j, P(j)\}$ are computed through the expectation maximization (EM) algorithm using the training samples [3]. Once the above parameters are decided, any new sample is assigned to the class, to which it attains the highest probability.

2.2.2.6 Support vector machine (SVM)

Since it has excellent generalization capability, SVM algorithms are applied to solve many practical problems [3], [6]. During the training of SVM, the samples are mapped to a higher dimensional space. In that space, a hyperplane is established to maximize the distances between samples in different classes and minimize the classification error.

The SVM training is performed by solving a dual optimization problem as

Maximize

$$\sum_{k=1}^{N} \alpha_k - \frac{1}{2} \sum_{k=1}^{N} \sum_{j=1}^{N} \alpha_k \alpha_j y_k y_j K\left(\mathrm{x}_k, \mathrm{x}_j\right) \tag{2.11}$$

Subject to

$$\begin{aligned} &\sum_{k=1}^{N} \alpha_k y_k = 0 \\ &0 \le \alpha_k \le \rho_k C, k = 1, ..., N \end{aligned} \tag{2.12}$$

where α_k is the Lagrange multiplier, C is the penalty parameter used to make the tradeoff between the margin maximization and misclassification, and $K\left(\mathrm{x}_k, \mathrm{x}_j\right)$ is the kernel function in the form of $K\left(\mathrm{x}_k, \mathrm{x}_j\right) = \varphi\left(\mathrm{x}_k\right)^T \varphi\left(\mathrm{x}_j\right)$.

Once an optimal separating hyperplane is found, SVM can assign a new sample x^* with a class label as

$$y = \mathrm{sgn}\left[\sum_{k=1}^{N} \alpha_k y_k K\left(\mathrm{x}_k, \mathrm{x}^*\right) + b \right] \tag{2.13}$$

where N is the total support vectors number, b denotes the bias term.

This paper adopts three types of kernels including linear $k(x_i, x_j) = x_i^T x_j$, polynomial $k(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$, and radial basis function (RBF) kernel $k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$, in which γ, *r*, and *d* are related parameters for the above kernels.

2.2.2.7 Decision tree

In a decision tree, each branch node is a choice between several input samples and each leaf denotes a class of these samples. The most widely used algorithms of decision tree are ID3, intelligent decision tree algorithm (IDA), and C4.5 [7].

2.2.2.8 Fuzzy support vector machine (FSVM)

In the original SVM all training samples are treated uniformly. In contrast, in the fuzzy support vector machine (FSVM), the samples have different weights; the sample near to the class centre is given a higher weight [8].

In FSVM, the centre of a class is the Euclidean mean of the samples in that class and the sample weight $\rho(\boldsymbol{x}_i)$ is calculated as [9]:

$$\rho(\mathbf{x}_i) = \frac{2}{1 + \exp(\beta d_i^{cen})}, \beta \in [0,1] \tag{2.14}$$

where $d_i^{cen} = \|x_i - c_i\|^{1/2}$ is the Euclidean distance x_i and class centre c_i , β describes the decay steepness.

2.2.2.9 Radial basis function (RBF)

In contrast to MLP, the RBF network adopts a non-linear activation function, which is decided by the distance between the input and a weight vector. The RBF network derives an input-output mapping using *M* radial basis functions [3]

$$z(y) = \sum_{m=1}^{M} w^{km} \varphi_m(y) \qquad k = 1,...,T \tag{2.15}$$

where φ_m are the basis functions and w^{km} are the output layer weights. The posterior probability of class membership is:

$$p(C_k|y) = \sum_{m=1}^{M} w^{km} \varphi_m(y) \tag{2.16}$$

where $C_k \in \{1,...,T\}$, T is the number of classes, and the basis function ϕ_m are given by

$$\varphi_m(y) = \frac{p(y|m)\,p(m)}{\sum_{m=1}^{M} p(y|m')\,p(m')} \tag{2.17}$$

and the weights are given by

$$w^{km} = \frac{p(m|C_k)\,p(C_k)}{p(m)} = p(C_k|m) \tag{2.18}$$

In the RBF network, the parameters for the radial basis functions are determined through an unsupervised learning procedure. The computation of output layer weights is formulated as a polynomial optimization problem [3].

2.2.2.10 Generalized regression neural network (GRNN)

Generalized Regression Neural Networks (GRNN) can provide estimation of continuous variables and converges to the optimal linear or nonlinear surface [10]. This algorithm can learn the underlying functions between input and output (e.g. class) from samples without having the prior knowledge of any specific function form between input and output.

2.2.2.11 Fuzzy logic classifier

The fuzzy logic classifier is trained using an iterative method, which consists of three steps [11]. In the first step, for a new input sample residing in the support-region of an existing if-then rule, the core-region of this rule is extended to include this new sample. In the second step, for a new input sample does not reside in any existing regions of if-then rules, a new fuzzy rule is generated for it. In the third step, for a new sample that is incorrectly classified by an if-then rule, the region of this rule

shrinks to exclude this sample.

2.2.2.12 Self-organizing tree algorithm (SOTA)

SOTA is based on self-organized map (SOM) and hierarchical clustering [12]. It adopts the growing cell structures, in which the network topology is built up incrementally. The training procedure of SOTA is quite similar to that of the learning vector quantization (LVQ) [13].

2.2.3 Improving Data Quality of Training Dataset for Pattern Recognition Algorithms

As it is mentioned in the previous paper [1], the pattern recognition algorithms are trained to learn the mathematic model from a training dataset (oil characteristics and the insulation conditions of the corresponding transformers) and then make classification on any new data samples, which are not included in the training dataset. Therefore, the quality of training dataset may significantly influence the performance of a pattern recognition algorithm.

A transformer is a complex machine system and involves complicated electrical and mechanical processes during its operation. Moreover, due to the limitations of the sensors and measurement systems, the extensive environmental interference and noise, and the possible misinterpretation on the measurement data, there might be inaccuracy and uncertainty in the oil characteristic dataset used for training pattern recognition algorithms. Furthermore, the possibility of power transformer having certain types of faults is not high. This implies that the samples belonging to these types of faults are not sufficient and can be significantly outnumbered by other types of faults. Trained by such imbalanced database, the algorithms may not be able to attain desirable performance. To attain the consistent classification accuracy in transformer insulation diagnosis, it is necessary to effectively improve the quality of training dataset.

In the previous publications of the authors of this paper [2, 14], the hybrid FSVM and fuzzy *c*-means clustering algorithms were proposed to reduce the effects of outlier and noise of training samples. A number of weighting and sampling methods were also implemented for handling the class imbalance problem during the algorithms' training process [2].

2.2.4 Case Studies

This section presents the case studies of applying the above 15 pattern recognition algorithms for transformer insulation diagnosis using oil characteristics. The dataset preparation, numeric simulation setup, the results analysis, and the statistical comparison among different algorithms will be detailed in this section.

2.2.4.1 Case study datasets

Eight oil characteristics datasets are used for evaluating the pattern recognition algorithms implemented in this paper. Table 2.4 summarizes the configurations of these eight datasets. In Table 2.4, Datasets 1, 2, 3, 5, and 6 were collected from different utility companies and Dataset 4 was used from the literature [15]. Moreover, Dataset 7 is the combination of Datasets 1, 2, 3, and 4 with a few incompatible records removed. In Dataset 8, about two third samples are from Dataset 5 and about one third records are from the original Dataset 6. The purpose of combining different datasets collected from different sources is to evaluate the pattern recognition algorithms, which are trained on one dataset but tested on another dataset. The algorithm, which is trained on a "local" dataset but can achieve consistent classification accuracy on "global" dataset, is regarded as having higher generalization capability.

Table 2.4 Configuration of eight oil characteristics datasets

Datasets	Source	Samples	Number of features
1	Utility	390	5 dissolved gases concentrations
2	Utility	992	5 dissolved gases concentrations
3	Utility	181	5 dissolved gases concentrations
4	Publications	206	5 dissolved gases concentrations
5	Utility	181	5 dissolved gases concentrations and 6 oil tests results
6	Utility	882	5 dissolved gases concentrations and 6 oil tests results
7	Utility & Publications	1542	5 dissolved gases concentrations
8	Utility	875	5 dissolved gases concentrations and 6 oil tests results

In Table 2.4, the features (e.g. the attributes of data) of the samples in Datasets 1, 2, 3, 4, and 7 are the concentrations of five dissolved gases (e.g. C_2H_4, C_2H_2, C_2H_6, H_2, CH_4) and the insulation

conditions (e.g. classes) include normal, discharge fault, partial discharge fault, and thermal faults. The data attributes of the samples in Datasets 5, 6 and 8 include the concentrations of five dissolved gases and six oil tests results such as acidity, breakdown voltage, dielectric dissipation factor, 2-furfuraldehyde, resistivity, and water content. Moreover, in these three datasets, the transformer insulation condition is categorized into four classes including "excellent", "good", "fair", and "poor" condition [14]. The relationship between typical faults diagnosed by DGA from IEC standard and insulation condition can be mapped as Table 2.5.

Table 2.5 Relationship between typical faults in IEC standard and insulation condition

Insulation Condition	**Faults Determined in IEC standard**
Excellent	No faults
Good	Discharge of low energy
Fair	Discharge of high energy Thermal faults less than 300 ℃
Poor	Partial discharge Thermal faults between 300 ℃ and 700 ℃ Thermal faults higher than700 ℃

2.2.4.2 Numeric experiments set-up

To evaluate the performance of the 15 pattern recognition algorithms as presented in Section 2.2.2, each dataset in Table 2.4 is randomly split into a training sub-dataset (70% samples) and a testing sub-dataset (30% samples). Some algorithms' parameters to be determined in the training are: the hidden nodes number in MLP, RBF, and GRNN, the nearest neighbors number for *K*NN, the learning rate of LVQ, the mixture elements number in Bayesian classifiers, the penalty parameter C and the variance γ of the RBF kernel in SVM and FSVM.

To decide the optimal values of the above parameters, tenfold cross validation was firstly performed on the training dataset for each of the algorithms. For k fold cross-validation, the original database D is randomly split into k mutually exclusive subsets $D_1, D_2, \cdots D_k$ with approximately equal size. The algorithm is then trained on $D_1 \cdots D_{t-1}$ and validated on D_t, $t \in \{1, 2, \cdots, k\}$. Total k trained models can be obtained. The parameters of the trained model, which has the highest classification accuracy, are chosen to build the final mathematic model for performing classification on the testing dataset. Finally, the trained algorithms are applied to categorize the samples in the testing sub-dataset into different conditions (excellent, good, fair, poor) of transformer insulation or different types of transformer

incipient faults (discharge fault, partial discharge, thermal fault, normal operating condition).

The above procedure of dividing training/testing sub-datasets, conducting tenfold cross validation, and performing algorithms' testing is repeated 50 times for each pattern recognition algorithm. It is also worthy to mention that the all training samples are normalized to [0, 1] interval before any process in the above procedure.

Some available software routines have been modified and extended for implementing the 15 pattern recognition algorithms presented in this paper [16-17] .

2.2.4.3 Results analysis and performance evaluation

For each of the eight datasets descried in Table 2.4, total 15 pattern recognition algorithms including PNN, MLP, LVQ1, LVQ2, *K*NN, SVM with linear kernel, SVM with polynomial kernel, SVM with Gaussian kernel, decision tree, GRNN, FSVM, Bayesian classifier, RBF, fuzzy logic, and SOTA have been applied to make classifications on transformer insulation condition. Due to the space limitation, this section only presents the classification results of the above algorithms on Datasets 1, 3, 5, 7 and 8. The classification accuracy on each class and the overall classification accuracy for the above five datasets are presented in Tables 2.6 to 2.10.

Tables 2.6 shows the classification results on Dataset 1, of which the input features are five dissolved gasses concentrations and the corresponding transformer insulation conditions are normal, discharge fault, thermal fault, and partial discharge. From Table 2.6 it can be observed that most algorithms can attain reasonable overall classification accuracy. However, the classification accuracy on the samples belonging to discharge fault is relatively low. For example, SVMs with polynomial kernel attained the classification accuracy less than 50%, and LVQ 1 and LVQ 2 even could not make classification. This due to the fact that the samples number of discharge fault class (12 samples) is rather low compared to the samples number in other three classes in Dataset 1 (201 samples in normal class, 145 samples in thermal fault class, and 32 samples in PD class).

Table 2.6 Classification accuracy of 15 algorithms on Dataset 1 (results are in percentage)

Algorithm	**Classification Accuracy of Original Dataset**				
	Normal condition	**Discharge fault**	**Thermal fault**	**Partial discharge**	**Overall**
MLP	90	83	92	78	89
PNN	86	67	95	78	86
LVQ1	85	0	91	0	78
LVQ2	90	0	93	0	79
KNN	86	97	91	78	87
Bayesian classifier	83	15	93	66	83
SVM_{lin}	88	50	95	84	89
SVM_{poly}	94	47	93	68	90
SVM_{gauss}	90	60	93	76	89
Decision tree	85	67	87	70	82
FSVM	97	17	81	67	87
RBF	87	67	95	84	89
GRNN	87	67	92	84	87
Fuzzy Logic	98	60	88	60	90
SOTA	96	45	92	66	90

Tables 2.7 and 2.8 present the classification results on Datasets 3 and 5 respectively. It needs to mention that Dataset 3 and 5 are constructed by using the records of the same batch of 181 transformers. In Dataset 3, the records only consist of the concentrations of five dissolved gases and the conditions of the corresponding transformers are grouped into four classes with the conditions of discharge fault, partial discharge, thermal fault and normal operating condition. In Dataset 5, the records consist of the concentrations of five dissolved gas and six oil tests results. The conditions of the corresponding transformers are categorized into four classes including “excellent”, “good”, “fair”, and “poor” condition. It can be observed from both Table 2.7 and 2.8 that algorithms (except Bayesian classifier) can make classification on two datasets. However, by comparing the results shown in Tables 2.7 and 2.8, it can be seen that with more features (Table 2.7 uses 5 features while Table 2.8 uses 11 features) the algorithms can attain slightly better classification accuracy. This is because oil test results can also reveal some aspect of the insulation conditions of transformers, which may not be revealed by dissolved gases alone.

Table 2.7 Classification accuracy of 15 algorithms on Dataset 3 (results are in percentage)

Algorithm	**Classification Accuracy of Original Dataset**				
	Excellent	**Good**	**Fair**	**Poor**	**Overall**
MLP	60	67	75	78	67
PNN	92	37	92	89	76
LVQ1	77	40	75	67	63
LVQ2	77	33	92	78	67
KNN	71	43	75	78	65
Bayesian classifier	77	31	20	91	59
SVM_{lin}	83	33	75	83	69
SVM_{poly}	83	33	58	61	63
SVM_{gauss}	85	50	58	83	72
Decision tree	56	50	83	61	58
FSVM	98	27	67	67	69
RBF	81	33	83	67	66
GRNN	79	33	83	78	67
Fuzzy Logic	76	52	82	61	67
SOTA	77	43	64	78	66

Table 2.8 Classification accuracy of 15 algorithms on Dataset 5 (results are in percentage)

Algorithm	**Classification Accuracy of Original Dataset**				
	Excellent	**Good**	**Fair**	**Poor**	**Overall**
MLP	88	83	75	72	82
PNN	94	83	67	78	85
LVQ1	88	80	92	67	82
LVQ2	77	83	100	67	76
KNN	83	93	67	100	87
Bayesian classifier	58	49	34	93	58
SVM_{lin}	83	83	83	67	81
SVM_{poly}	94	87	50	78	84
SVM_{gauss}	94	90	83	94	92
Decision tree	83	70	58	78	76
FSVM	100	70	67	56	81
RBF	83	87	75	61	80
GRNN	92	94	95	95	94
Fuzzy Logic	100	93	80	44	86
SOTA	88	57	71	56	72

Table 2.9 Classification accuracy of 15 algorithms on Dataset 7 (results are in percentage)

Algorithm	**Classification Accuracy of Original Dataset**				
	Normal condition	**Discharge fault**	**Thermal fault**	**Partial discharge**	**Overall**
MLP	70	78	86	76	80
PNN	30	87	91	82	75
LVQ1	4	91	87	84	69
LVQ2	2	95	63	0	46
KNN	54	84	87	84	79
Bayesian classifier	0	11	98	77	49
SVM_{lin}	39	92	83	84	76
SVM_{poly}	21	94	75	68	68
SVM_{gauss}	72	86	88	84	84
Decision tree	54	75	71	85	71
FSVM	10	90	78	82	67
RBF	71	85	88	86	84
GRNN	36	88	89	84	76
Fuzzy Logic	76	86	84	84	83
SOTA	33	57	72	71	60

Table 2.10 Classification accuracy of 15 algorithms on Dataset 8 (results are in percentage)

Algorithm	**Classification Accuracy of Original Dataset**				
	Normal condition	**Discharge fault**	**Thermal fault**	**Partial discharge**	**Overall**
MLP	88	92	89	91	90
PNN	92	96	94	95	95
LVQ1	0	96	86	81	81
LVQ2	0	100	0	0	36
KNN	92	92	97	99	95
Bayesian classifier	0	94	68	95	77
SVM_{lin}	79	98	77	95	89
SVM_{poly}	75	98	86	93	91
SVM_{gauss}	83	96	97	94	95
Decision tree	85	96	92	96	94
FSVM	98	91	91	93	92
RBF	90	94	96	95	95
GRNN	92	95	94	97	95
Fuzzy Logic	89	97	93	91	94
SOTA	85	85	87	91	87

Table 2.10 shows the classification results on Dataset 8, of which the input features are five dissolved gasses concentrations and six oil tests results. The conditions of the corresponding transformers are grouped into excellent, good, fair, and poor. The samples in Dataset 8 were taken from both Dataset 5 and Dataset 6. From Table 2.10 it can be concluded that the most algorithms (except LVQ 2) can attain reasonable classification accuracy. This demonstrates the generalization capability of the pattern recognition algorithms, which are trained by utilizing one dataset but later tested over another dataset.

The pattern recognition algorithms are also trained and tested on other Dataset 2, 4, 6 and the results show that most of the algorithms can attain consistent classification accuracy on different dataset. However, it is also found that it is not a trivial task to make comparisons amongst different algorithms. For example, the SVM with Gaussian kernel is the top performing algorithm on Dataset 7. However, its classification accuracy is not better than that of SVM with polynomial kernel function on Dataset 1. Next section will present an approach for providing the statistical comparison among different pattern recognition algorithms when they are applied for transformer insulation diagnosis.

2.2.4.4 Statistical comparison of different pattern recognition algorithms for transformer insulation diagnosis

To evaluate the performance of various algorithms over multiple datasets, the Friedman test, a statistical test is performed. The procedure of this test is as follows. Firstly, the algorithms for each dataset are ranked individually. The algorithm having the highest classification accuracy is ranked as one and the second best algorithm is ranked as two and so on. Then the averaged rank of each algorithm is computed as

$$R_j = \frac{1}{N}\sum_{i=1}^{N} r_i^j \tag{2.19}$$

where r_i^j denotes the rank of the j-th algorithm (total k=15 algorithms) on the i-th dataset (total N=8 datasets). By computing the averaged rank of each algorithm over eight datasets, it appears that the top five best performing classification algorithms are: fuzzy logic, PNN, *K*NN, SVM (Gaussian kernel), and FSVM (refer to Table 2.11).

To formally state that fuzzy logic is the best among all 15 algorithms, it is necessary to reject the null-hypothesis, which states that, all the algorithms are equivalent so that any differences among

their averaged ranks R_j are merely random. This can be done by computing the Friedman statistic as follows [18]

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1)-\chi_F^2} \tag{2.20}$$

$$\chi_F^2 = \frac{12N}{k(k+1)}\left[\sum_{j=1}^{k} R_j^2 - \frac{k(k+1)^2}{4}\right] \tag{2.21}$$

Based on the averaged ranks of 15 algorithms, the F_F is calculated as 8.34, which is greater than the critical value of $F(14, 98)=2.345$ at significance level $\alpha = 0.01$. Therefore, the above null-hypothesis can be rejected. And this indicates that the performances of 15 algorithms are not equivalent at the significance level of $\alpha = 0.01$.

Table 2.11 Averaged scores of 15 algorithms over eight datasets

Algo.	**Score**	**Algo.**	**Score**	**Algo.**	**Score**
MLP	7.5	Bayesian	9.1	FSVM	6.6
PNN	5.0	SVM_{lin}	8.1	RBF	12.5
LVQ1	11.3	SVM_{poly}	7.9	GRNN	6.9
LVQ2	11.5	SVM_{gauss}	6.3	Fuzzy Logic	3.6
KNN	5.0	Decision Tree	8.9	SOTA	9.9

Then, Benferroni–Dunn test is used to compare the fuzzy logic to the other algorithms. In this test, a critical difference (CD) value is defined as $CD = q_\alpha\sqrt{\frac{k(k+1)}{6N}}$ where q_α value is 2.055 at $\alpha = 0.05$ (95% confidence coefficient). The corresponding CD value is 4.59. If the rank difference between fuzzy logic and another algorithm is at least this critical difference value, then fuzzy logic is significantly superior to that algorithm.

The differences between fuzzy logic and LVQ1, LVQ2, decision tree, Bayesian classifier, RBF, and SOTA are greater than the above CD value. Thus it is able to claim that fuzzy logic is significantly better than these six algorithms for transformer insulation diagnosis. However, the difference between fuzzy logic and PNN, MLP, three SVMs, GRNN, and FSVM are less than the above CD value. Therefore we cannot claim that fuzzy logic is significantly superior to the above seven algorithms.

2.2.4.5 Discussions

In utilities, the faults occurrence rate of transformers is rather low and this implies that the records of some types of faults are very limited. In the datasets used in this section, there are even no records for some types of faults. Therefore, the records in the above datasets are only grouped into four categories. However, the algorithms can also classify the transformer conditions with more categories. For example, the records in Dataset 1 are re-grouped into five categories including normal, discharge, low and medium temperature overheat, high temperature overheat and partial discharge) and numerical experiments are then performed on the extended Dataset 1. As it is shown in Table 2.12, compared to the overall classification accuracy on the original Dataset 1 (with four categories, Table 2.6), the overall classification accuracy on this extended Dataset 1 (with five categories, Table 2.12) is relatively lower. In Table 2.12, the thermal faults of transformers are further divided into low and medium temperature overheat (T1/T2, lower than 700 degree) and high temperature thermal fault (T3, higher than 700 degree). It can be observed that in Table 2.12 the classification accuracy on "T3" class is very low. Actually, most algorithms failed to make classification due to small sample size of "T3" class compared with other classes. To deal with such a problem, some pre-processing methods can be integrated with the algorithm to facilitate it in achieving consistent desirable classification accuracy [2, 14].

Table 2.12 Classification accuracy of 15 algorithms on Dataset 1extended with five classes

Algorithm	**Classification Accuracy of Original Dataset**					
	Normal condition	**Discharge fault**	**T1/T2**	**T3**	**Partial discharge**	**Overall**
MLP	89	83	91	0	72	86
PNN	86	83	95	17	89	87
LVQ1	83	0	80	17	0	72
LVQ2	91	0	95	0	0	80
KNN	85	67	98	33	72	87
Bayesian classifier	83	15	93	33	66	82
SVM_{lin}	96	43	93	0	68	89
SVM_{poly}	94	33	93	0	76	88
SVM_{gauss}	91	33	91	33	79	87
Decision tree	82	17	88	67	84	82
FSVM	97	17	81	33	67	85
RBF	87	67	95	33	84	88
GRNN	87	80	95	0	80	87
Fuzzy Logic	98	60	88	33	60	89
SOTA	96	45	92	33	66	89

In the paper, the purposes of evaluating 15 algorithms are two-fold. Firstly, through the evaluation process, it aims to standardize the procedure of training, cross-validation, testing and evaluation of algorithms for power transformer condition assessment. This can provide a common ground for evaluating the performances of different algorithms. Secondly, these algorithms are evaluated on a number of datasets collected from different utility companies and each dataset consists of a reasonable large number of records. This can facilitate the assessment on the generalization capability of the algorithms, i.e. an algorithm trained on a local dataset can be applied "globally" for another new dataset.

The performance of a pattern recognition algorithm is influenced by a number of factors such as the structure of its mathematic model, parameters of its model and approach of training. Given many different permutations of these factors, it is not an easy task to evaluate and compare the performance of different algorithms. Therefore, the comparison made in this paper provides some indications on the classification accuracies of the 15 studied algorithms.

In this paper, these 15 algorithms adopted their basic and simplest structures. In ANN algorithm, three layers structure (input, one hidden and output layer) was adopted and the tanh activation function was used. Ten-fold cross-validation was also applied to determine the optimal parameters of the algorithms, including the hidden nodes number in MLP, RBF, and GRNN, the nearest neighbors' number for *K*NN, the learning rate of LVQ, the mixture elements number in Bayesian classifiers, the penalty parameter C and the variance γ of the RBF kernel in SVM and FSVM. Moreover, for training the supervised learning algorithms, a conventional scheme for transformer incipient fault identification (IEC 60599) was employed for Datasets 1, 2, 4 and 7 while a hybrid interpretation scheme (combination the interpretation results obtained from related industry standards, human experts' judgments, a conventional scheme and a clustering algorithm) was employed for Datasets 3, 5, 6 and 8. It is expected that the numerical experiments on the 15 algorithms configured with the above simplified structures, few optimal parameters and training arrangements can still reveal the baseline performance of these algorithms.

2.2.5 Conclusion

This paper investigated a variety of state-of-the-art pattern recognition algorithms for transformer insulation diagnosis. The key advantages of pattern recognition algorithms over the conventional interpretation schemes are their capabilities in exploring the correlations between oil characteristics

and insulation condition by using historic dataset. Such knowledge will greatly facilitate evaluating insulation condition of any transformer of interest. To verify the applicability and generalization of pattern recognition algorithms, extensive case studies using eight oil characteristics datasets were conducted in this paper. To compare the performance (in terms of classification accuracy) of 15 algorithms implemented in this paper, a statistical comparison approach was adopted. Although this approach may not be able to provide exclusive ranking on all the algorithms presented in this paper since the dataset number (total eight datasets) and dataset size (from 180 to 1600 records) in this paper are limited, it does provide an avenue for comparing the performance of various pattern recognition algorithms.

Acknowledgment

We gratefully acknowledge Australian Research Council, Powerlink Queensland, Energex, Ergon Energy, and TransGrid for providing supports for this work.

References

[1] Y. Cui, H. Ma and T. K. Saha, "Pattern Recognition Techniques for Power Transformer Insulation Diagnosis - A Comparative Study Part 1: Framework, Literature, and Illustration," *Int. Trans. Electr. Energy. Syst.*, vol.25, Issue 10, pp. 2247-2259, 2015.

[2] H. Ma, C. Ekanayake and T. K. Saha, "Power transformer fault diagnosis under measurement originated uncertainties," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 6, pp. 1982-1990, 2012.

[3] C. M. Bishop, *Pattern Recognition and Machine Learning*, New York: Springer, 2006.

[4] D. F. Specht, "Probabilistic neural networks," *Neural Networks*, vol.3, Issue 1, pp. 109-118, 1990.

[5] J. Hollmén, V. Tresp and O. Simula, "A learning vector quantization algorithm for probabilistic models," in *Proceedings of 10th European Signal Processing Conference, September 4-8*, 2000,Tampere, Finland, pp. 721-724.

[6] N. Cristianini and J. Shawe-Taylor, *An Introduction of Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, 2001.

[7] J. Quinlan, *C4.5: Programs for Machine Learning*, San Mateo: CA:Morgan Kaufman, 1993.

[8] C. Lin and S. Wang, "Fuzzy support vector machines," *IEEE Trans. Neural Networks*, vol.13, Issue 2, pp. 464-471, 2002.

[9] R. Batuwita and V. Palade, "FSVM-CIL: Fuzzy support vector machines for class imbalance learning," *IEEE Trans. Fuzzy Syst.*, vol.18, Issue 3, pp. 558-571, 2010.

[10] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Networks*, vol.2, Issue 6, pp. 568-576, 1991.

[11] T. R. Gabriel and M. R. Berthold, "Influence of fuzzy norms and other heuristics on “Mixed fuzzy rule formation”," *Int. J. Approximate Reasoning*, vol.35, Issue 2, pp. 195-202, 2004.

[12] J. Herrero, A. Valencia and J. I. N. Dopazo, "A hierarchical unsupervised growing neural network for clustering gene expression patterns," *Bioinformatics*, vol.17, Issue 2, pp. 126-136, 2001.

[13] T. Kohonen, *Self-organizing maps*, Springer, 1995.

[14] A. D. Ashkezari, H. Ma, T. K. Saha and C. Ekanayake, "Application of fuzzy support vector machine for determining the health index of the insulation system of in-service power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 3, pp. 965-973, 2013.

[15] M. Duval and A. DePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electr. Insul. Mag.*, vol.17, Issue 2, pp. 31-41, 2001.

[16] I. T. Nabney, *Netlab Algorithms for Pattern Recognition*, Springer, 1995.

[17] C. C. Chang and C. J. Lin,"LIBSVM: a library for support vector machines" available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[18] J. Demsar, "Statistical comparisons of classifiers over multiple data sets," *J. Machine Learning Research*, vol.1, Issue 1, pp. 1-30, 2006.

Chapter 3

Improvement of Power Transformer Insulation Diagnosis Using Oil Characteristics Data Preprocessed by SMOTEBoost Technique

Contribution of the Chapter

This chapter proposes a novel method for power transformer insulation assessment using oil characteristics. A hybrid algorithm, named as SMOTEBoost is implemented in the chapter to improve the diagnosis accuracy and consistency. The SMOTEBoost can significantly enhance the generalization capability of artificial intelligence (AI) algorithms for transformer insulation diagnosis. This will provide important benefits for applying AI techniques in utility companies, i.e., an AI algorithm with its model built upon on a “local” dataset can be utilized “globally” to make transformer insulation diagnosis.

Improvement of Power Transformer Insulation Diagnosis Using Oil Characteristics Data Preprocessed by SMOTEBoost Technique

Yi Cui, Hui Ma, and Tapan Saha

The University of Queensland, Brisbane, Australia

Publication Journal: IEEE Transactions on Dielectrics and Electrical Insulation

Submitted: January 3, 2014

Revision Resubmitted: April 17, 2014

Accepted: April 29, 2014

Published: October 23, 2014

Author's Contributions:

Yi Cui: Theoretical model design, numeric experimental simulation, results analysis and manuscript preparation.

Hui Ma: Supervision on the numeric experiment, results discussion and manuscript revision.

Tapan Saha: Supervision on the numeric experiment, results discussion and manuscript revision.

Abstract

This paper proposes a novel method for power transformer insulation assessment using oil characteristics. A hybrid algorithm, named as SMOTEBoost is implemented in the paper to improve the diagnosis accuracy and consistency. The SMOTEBoost can significantly enhance the generalization capability of artificial intelligence (AI) algorithms for transformer insulation diagnosis. This will provide important benefits for applying AI techniques in utility companies, i.e., an AI algorithm with its model built upon on a "local" dataset can be utilized "globally" to make transformer insulation diagnosis. The SMOTEBoost adopts Synthetic Minority Over-sampling Technique (SMOTE) to handle the class imbalance problem, in which data points belonging to different fault types or insulation conditions are unevenly distributed in the training dataset. By using this boosting approach for reweighting and grouping data points in the training dataset, the SMOTEBoost facilitates AI algorithms consistently attaining desirable diagnosis accuracy. To verify the advantages of SMOTEBoost algorithm, it is integrated with a number of representative AI algorithms including support vector machine (SVM), C4.5 decision tree, radial basis function (RBF) network and *k*-nearest neighbor (*K*NN) to make transformer insulation diagnosis using various oil characteristic datasets collected from different utility companies. A statistical performance comparison amongst these algorithms is presented in the paper.

Index Terms: Dissolved gas analysis (DGA), insulation, oil characteristics, power transformer, support vector machine

3.1 Introduction

Within the designed life span of a power transformer, its oil-paper insulation system may eventually degrade due to electrical, mechanical and thermal stresses. Therefore, different techniques are explored for transformer diagnosis [1-2]. Amongst these techniques, oil tests have been widely performed in utility companies, which include the measurement of acidity, water content, breakdown voltage (BDV), dielectric dissipation factor (DDF) and 2-furfuraldehyde as well as dissolved gas analysis (DGA).

To diagnose the incipient fault and assess the insulation condition of a transformer, various

interpretation methods have been developed, such as key gas method, different gas ratios (e.g. IEEE/IEC, Rogers and Doernenburg) and Duval triangle [3-4]. Over the past two decades, artificial intelligence (AI) techniques including support vector machine (SVM), artificial neural network (ANN), fuzzy logic, neural fuzzy system and wavelet network have been extensively investigated and applied to transformer insulation diagnosis [5-9]. The advantage of an artificial intelligence (AI) algorithm is that it utilizes not only the oil characteristics obtained from the oil tests on the transformer of interest, but also the historical oil characteristics dataset collected from other transformers. Learning from this historic dataset, an AI algorithm can determine the correlation between oil characteristics and transformers' insulation condition and then make diagnosis on the transformer of interest [5].

However before AI techniques can be adopted by utilities as a practical tool for transformer insulation diagnosis, two major challenging issues need to be investigated. The first key issue is to build up a statistically satisfied training dataset, which plays a decisive role for the performance of an AI algorithm [10]. A statistically significant training dataset needs to include a sufficient number of samples, which can be used by the AI algorithms to exploit the correlations between oil characteristics and fault types or insulation condition (i.e. classes) of the corresponding transformers. Moreover, the samples need to be evenly distributed amongst different classes to avoid causing any bias in fault classification. The second key issue is the generalization capability of AI algorithms. During the construction of a training dataset, the fault types of the transformers are labeled based on utility expert judgments, which is made by using conventional methods (i.e. IEC/IEEE ratio, Duval Triangle) and expert's experience. Due to the heuristic nature of expert judgements and the possible presence of multiple faults in a transformer, there may exist some degree of inconsistency between datasets collected from different utility companies. Such data inconsistency can jeopardize an algorithm's generalization capability, i.e. an algorithm trained by the oil test dataset provided by one utility can be used by other utility companies (refer to Section 3.5.4 for further discussion).

To address the above issues, this paper implements a novel algorithm, SMOTEBoost for power transformer insulation diagnosis using oil characteristics. The SMOTEBoost can improve the data quality of a training dataset and enhance the diagnosis accuracy and generalization capability of any AI algorithms that adopt it as a data pre-processing step for transformer insulation diagnosis. In the paper, the SMOTEBoost is integrated with a number of representative AI algorithms, i.e. support vector machine (SVM) with Gauss kernel function, C4.5 decision tree, radial basis function (RBF) network and *k*-nearest neighbor (*K*NN). Then these algorithms are applied to transformer insulation diagnosis, i.e. transformer incipient faults or transformer insulation condition classification using

eight oil test datasets collected from different utilities. A statistical performance comparison amongst these AI algorithms is also presented.

3.2 Brief Review of Oil Tests and Training Dataset Construction

This section provides a brief review on transformer oil tests. It also describes the procedure of applying AI algorithms to transformer insulation diagnosis and details the construction of DGA and oil characteristic (acidity, water content, BDV, DDF, 2-furfuraldehyde and various dissolved gases) datasets for training AI algorithms.

3.2.1 Transformer Oil Tests

Due to a variety of stresses and ageing, the molecular bonds of transformer oil can be broken and particles and fragments can be generated within the oil. Eventually, water, acid compound and dissolved gases are formed in the transformer oil. To assess the condition of a transformer's insulation system a variety of tests are performed on samples of its oil [11-13].

Among oil tests, Dissolved Gas Analysis (DGA) can provide information on transformer incipient fault conditions, including Partial Discharge (PD), thermal fault and discharge [3-4]. The conventional DGA interpretation schemes include ratio based method (IEC, Roger and Doernenburg Ratio) and Duval's Triangle. However this paper applies AI algorithms to interpret DGA data and classify transformer incipient faults.

Besides DGA, other physical and chemical tests can also reveal the condition of transformer insulation. Acids are produced from either the oxidation products in transformer oil or atmospheric contamination. The acid number can indicate the extent of oil corrosiveness on the paper insulation and metallic parts inside the transformer tank. Dielectric dissipation factor (DDF) is measured to monitor the ageing condition of transformer oil since it is sensitive to the presence of ageing products and soluble polar contaminants. Breakdown voltage (BDV) of the oil must remain high to withstand electrical stress without failure and a low BDV may imply the formation of contaminants in oil. The 2-furfuraldehyde value can be used to estimate the condition of the paper insulation since it has a correlation with the degree of polymerization (DP) of the paper. The concentrations of carbon oxides (CO and CO_2) can also be used to assess the condition of paper. Moreover, due to the deterioration of paper insulation or water absorption from the atmosphere, free water can be formed

when water content is beyond the oil saturation level. As a result, the electrical strength of insulating oil will decrease. By making use of oil test data, this paper applies AI algorithms to evaluate the condition of a transformer's insulation and determine a health index for the transformer insulation system.

3.2.2 Procedure of AI algorithms for transformer insulation diagnosis

AI algorithms based transformer insulation diagnosis involves three steps. They are: (1) training dataset construction - a dataset is constructed by using oil test records collected from utilities; (2) training - algorithms are trained to learn the relationships between DGA (oil characteristics) data and the transformer incipient faults (conditions of transformer insulation); and (3) classification - the trained algorithms classify the types of incipient faults using DGA data or computes health indexes (i.e. insulation condition) for transformers under investigation. Note that these transformers are not included in the training database.

3.2.3 Training dataset construction

The performance of AI algorithms relies on the model trained using a training dataset. Two types of training dataset are constructed in this paper. One is the DGA dataset for transformer incipient faults classification and another is the oil characteristics dataset for transformer insulation condition (health index) evaluation.

In the DGA dataset, each record consists of the concentrations of five dissolved gases, i.e. hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4) and ethane (C_2H_6) as well as the incipient faults of the corresponding transformers. These fault diagnosis results are determined upon industrial standards, experts' knowledge and onsite inspections.

In the oil characteristics dataset, each record consists of 12 oil characteristics (H_2, CH_4, C_2H_2, C_2H_4, C_2H_6, carbon monoxide (CO), carbon dioxide (CO_2), acidity, water content, BDV, DDF and 2-furfuraldehyde) of a transformer and its health index level. The calculation approach of health index level is recapitulated as follows. Readers may refer to the authors' previous publication [14] for more details.

(1) Determining health index (HI_1) based on industry standards - Both DGA factor (DGAF) and oil quality factor (OQF) are defined as:

$$DGAF = \sum_{i=1}^{7} S_i \times W_i \Big/ \sum_{i=1}^{7} W_i \tag{3.1}$$

$$OQF = \sum_{j=1}^{4} S_j \times W_j \Big/ \sum_{j=1}^{4} W_j \tag{3.2}$$

where i and j denote dissolved gases and other oil characteristics respectively. S_i and S_j are the scores based on test results. W_i and W_j are weighting factors of tests. A paper insulation factor (PIF) is decided by the amount of 2-furfuraldehyde. DGAF, OQF, and PIF are combined and normalized into the health index HI_1.

(2) Determining health index (HI_2) based on expert judgements - additionally to oil test records, comments from utility experts are also taken into consideration. Their comments include: (1) transformer condition is acceptable; (2) further investigation required; (3) DGA caution and repeat oil tests; and (4) DGA danger, abnormal or serious situation. Based on these comments, the condition of transformer insulation is quantified into health index HI_2.

(3) Determining health index (HI_3) based on Duval's Triangle - by using the faults identified from Duval's Triangle, the corresponding transformer insulation condition is quantified into health index HI_3.

(4) Determining health index (HI_4) based on a fuzzy *c*-means (FCM) algorithm - FCM assigns an oil characteristic record with a weight (possibility) of belonging to a particular health index. It helps to take the uncertainty and inaccuracy in deciding health index into account in dataset construction.

(5) Assigning final health index and forming the dataset - after obtaining the above four sets of health indexes a majority vote strategy is adopted for deciding the final health indexes of insulation systems for the transformers.

3.3 SMOTEBoost for Improving Data Quality of Training Dataset

At utility companies, the faulty rate of power transformers is rather low and few transformers' insulation systems are in critical condition. This implies that the records (samples) of some types of incipient faults or insulation conditions (e.g. transformer insulation with critical condition) are extremely insufficient. Thus the samples in the training datasets discussed in Section 3.2 are unevenly scattered over the data space, in which samples in some classes (minority classes) are significantly outnumbered by those in other classes (majority classes). Trained by such an unbalanced dataset, an

AI algorithm will be in favour of the majority classes which can lead to misclassification on the minority classes.

A resampling method might be a possible solution for the above class imbalanced problem. Resampling reconstructs the training dataset to reduce the degree of imbalance. Among resampling techniques, over-sampling duplicates samples in the minority class but has the possibility of over fitting; while the under-sampling method eliminates samples in the majority class but has the risk of losing information. As a consequence, simply applying resampling techniques may not be able to effectively improve the diagnosis accuracy and guarantee the generalization capability of an AI algorithm in transformer insulation diagnosis.

This paper applies SMOTEBoost to facilitate AI algorithm consistently achieving desirable diagnosis accuracy and generalization capability. In SMOTEBoost, the synthetic over-sampling technique (SMOTE) is implemented for dealing with the class imbalanced problem while the boosting approach is adopted for adaptively reweighting and grouping data points in the training dataset. Moreover, the bootstrap method is also utilized to statistically equalize the samples number after executing SMOTE.

Starting with a brief review on SMOTE technique, this section details the principles and formulations of SMOTEBoost. The theory of bootstrap is also briefly introduced. Finally, the implementation of the hybrid of SMOTEBoost and bootstrap is presented.

3.3.1 Synthetic minority oversampling technique (SMOTE)

Figure 3.1 depicts the schematic diagram of the SMOTE algorithm. For a particular sample in the minority class, SMOTE generates synthetic samples along the sections of the line, which joins the k nearest neighbors of this sample [10].

SMOTE is implemented in the following two steps [10]:

(1) Calculating the Euclidean distance between $\mathbf{X}$ and its k closest neighboring samples in the minority class. If k is set to a large value, it may increase the possibility of including outliers in the training dataset and lead to large error rate. By contrast, if k is set to a very small value, the number of generated synthetic samples may not be enough to ensure the diversity of the minority class. By using a trial and error method, k is chosen to be five in this paper.

(2) Determining the number of new samples to be synthesized (over-sampling rate) based on the

imbalance ratio of the original dataset. Suppose $\mathbf{x}_2$ is one of the closest neighbors of sample $\mathbf{X}$ and the Euclidean distance between $\mathbf{X}$ and $\mathbf{x}_2$ is $d(\mathbf{X},\mathbf{x}_2)$, then the synthetic samples are:

$$\mathbf{m}_i = \mathbf{X} + rand(0,1) \times d(\mathbf{X},\mathbf{x}_2) \tag{3.3}$$

where rand (0,1) is a random number in the range of [0, 1].

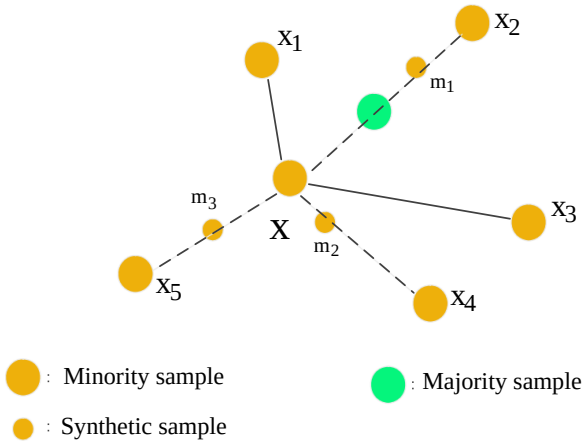


Figure 3.1 Schematic diagram of SMOTE algorithm (X denotes a sample and $X_1 \ldots X_5$ are its five nearest neighbours. SMOTE generates new data m_i along the line between $\mathbf{X}$ and one of five neighbours)

However, the above standard SMOTE still poses some limitations. Since new samples are synthetized by the random interpolation it cannot guarantee that all these synthesized samples are distributed within the center area of the minority class. Some samples may locate in the centre area of the minority class while others may become outliers, which locate remotely from the centre of the minority class. Such sample distribution may still introduce difficulties for an AI algorithm to make classification on fault types (or insulation conditions). To handle such limitations, a boosting procedure is proposed in this paper, which is integrated with SMOTE.

3.3.2 Integrating Boosting to SMOTE

Boosting is a meta-learning technique that “boosts” the performance of AI algorithms. Its basic premise is that multiple under performance algorithms (termed as weak learners) can be combined to construct a high performance ensemble (termed as a strong learner) [16]. During the initial

training of the boosting all samples are assigned with an equal weight. Then the incorrectly classified samples are given a higher weight to redirect the subsequent training to focus on these samples. The training iterations will continue until the ensemble of weak learners attains the desirable classification accuracy. The limitation of boosting is the possibility of over-fitting. Thus it generally requires large amounts of training samples [15]. Therefore, this paper integrates boosting into SMOTE to form the SMOTEBoost algorithm [16], which utilizes SMOTE for improving the classification accuracy of the AI algorithm on the minority classes while it adopts boosting to enhance the performance over the entire training dataset.

In SMOTEBoost, SMOTE is executed in each iteration of boosting to enable each classifier (AI algorithm) to attain high classification accuracy on the minority class. It should be noted that SMOTE is only invoked for processing the samples in the minority class. This implicitly increases these samples' weights. The synthesized samples will later be discarded after completing the training at iteration *t*, i.e. they are not added to the training dataset at iteration *t+1*. Such an arrangement has a significant benefit of increasing the data diversity for classifiers as different sets of synthetized samples are provided to the classifiers at different iterations.

3.3.3 Hybrid of SMOTEBoost and bootstrap

Figure 3.2 depicts the architecture of SMOTEBoost and bootstrap hybrid algorithm. In this hybrid algorithm, SMOTE is firstly used to synthesize new samples for the minority classes in the training dataset. SMOTE may change the original sample distribution by synthesizing higher cost samples. The number of synthetic samples is determined by SMOTE until the appearance of different samples are proportional to their costs [17]. After SMOTE, sample numbers in minority and majority classes may still not be equal. To statistically equalize the sample numbers amongst majority and minority classes, bootstrap is then adopted to redraw the samples in the majority class [18-19] in the training dataset.

Bootstrap is a resampling technique, which can be used to approximately equalize the sample numbers in each class of a dataset. Supposing after the SMOTE process, *n* samples are distributed in m majority classes and $n_1,...,n_t,...,n_m$ are the sample numbers in each majority class. Let n_{tr}, $t=1,...,m$ be the desired sample numbers of each majority class. Bootstrapping is performed for each class according to the ratio of n_{tr}/n_t, $t=1,...,m$ to redraw new samples from the original majority class. Then these samples are used to form the decreased size majority classes. Finally,

these majority samples are combined with the minority samples generated by SMOTE to construct a class balance training dataset, where different classes have equal numbers of samples. This training dataset will be used for algorithm training at the current iteration. It needs to be mentioned that the samples (in the majority classes) discarded by bootstrap are put back into the dataset after completing the training at the current iteration and can be used at the next iteration.

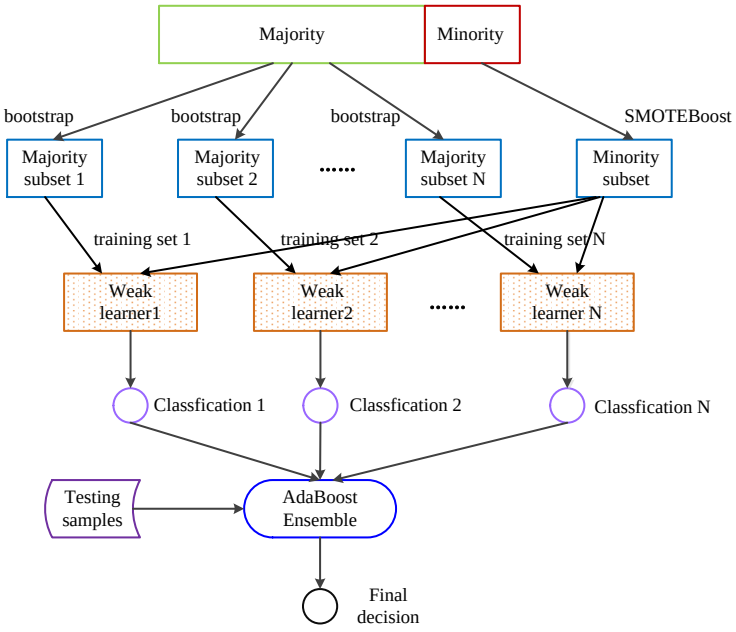


Figure 3.2 Architecture of hybrid SMOTEBoost and Bootstrap algorithm

After being processed by the SMOTE and bootstrap, both majority and minority classes now have statistically equalized samples and the original dataset becomes class balanced. This will help improve the classification performance of an AI algorithm which integrates SMOTE and bootstrap as a pre-processing step. After such integration a boosting approach will be invoked to enable the above AI algorithm to make a classification on the fault types or insulation condition of the transformer of interest.

In the boosting approach, during each iteration a number of training subsets will be established by combining the samples from the majority class and minority class. By using these training subsets a number of weak learners (e.g. any classifiers that only attain low classification accuracy during the training process) can be obtained at the current iteration. And then a strong learner can be obtained by combining these weak leaners after certain iterations (i.e. classifiers ensemble that attains high classification accuracy) [17]. Subsequently, this strong learner is applied to make classification on the fault types or the conditions of the transformer insulation in the testing dataset. These testing

samples are previously unknown for the strong learner. The above process (iterations) will be continued until a satisfied performance is attained by the strong learner. In this paper, SMOTE, boosting and bootstrap are integrated with a number of representative AI algorithms including C4.5, *k*NN, RBF network, and SVM for transformer incipient insulation condition fault classification. For the sake of integrity, the next section will provide a brief review of these four AI algorithms.

3.4 AI Algorithms for Transformer Insulation Diagnosis

3.4.1 Support vector machine (SVM)

To separate the samples in the original data space, SVM maps these samples into a high dimensional hyperspace. In that hyperspace, SVM finds a hyperplane which can keep the separation of samples of different classes maximized while minimize the classification error [20]. To obtain the above hyperplane, the following optimization problem needs to be solved [23].

Minimize

$$\sum_{k=1}^{N}\alpha_k - \frac{1}{2}\sum_{k=1}^{N}\sum_{j=1}^{N}\alpha_k\alpha_j y_k y_j K\left(\mathrm{x}_k,\mathrm{x}_j\right) \tag{3.4}$$

Subject to

$$\begin{aligned}&\sum_{k=1}^{N}\alpha_k y_k = 0\\ &0 \le \alpha_k \le \rho_k C,\ k = 1,...,N\end{aligned} \tag{3.5}$$

where α_k is the Lagrange multiplier, C is the regularization parameter to make the balance between the margin maximization and classification error and $K\left(\mathrm{x}_k,\mathrm{x}_j\right)$ is the kernel function formulated as $K\left(\mathrm{x}_k,\mathrm{x}_j\right)=\varphi\left(\mathrm{x}_k\right)^T\varphi\left(\mathrm{x}_j\right)$ [18]. After obtaining the above hyperplane, SVM determines the label of the class to which a new sample x^* belongs by using the equation of $y=\mathrm{sgn}\left[\sum_{k=1}^{N}\alpha_k y_k K\left(\mathrm{x}_k,\mathrm{x}^*\right)+b\right]$ (N is the support vectors number and b is the bias).

3.4.2 K-nearest neighbour (KNN) algorithm

In *K*NN, the Euclidean distances between sample **x*** to each sample in the training dataset are

computed. Let *Sk* denotes a subset consisting of the *K* closest data points and *Kl* denotes the frequency of the *l*-th class in *Sk*. Then **x*** is assigned to a certain class which has the largest frequency in *Sk* [21-22].

3.4.3 C4.5 decision tree

A decision tree is a tree like graph which consists of a number of decisions and the corresponding outcomes. In the graph, the internal nodes of the tree represent tests on input data and the leaf nodes represent the classes (fault types or insulation condition). The C4.5 decision tree adopts the information gain I(Y |X) to measure the reduction in uncertainty about the class label Y of a sample X given the value of this sample [23]. The uncertainty that attributes to the class label Y is calculated with its entropy, H(Y) and the uncertainty about Y given the value of X is denoted by the conditional entropy H (Y |X). C4.5 also adopts the gain ratio, which is (Y|X) = I(Y|X)/H(X) to measure the information gain with respect to the raw information contained in X distribution.

3.4.4 Radial basis function (RBF) network

The architecture of radial basis function network is presented in Figure 3.3 and it belongs to a family of artificial neural networks. Generally, the RBF network contains input, hidden and output layers.

In a RBF network, the interpolation function is in the form of $f(\|x-c_i\|)$, where $\|.\|$ is a distance measure and c_i is the center of the class to which the sample x belongs. The Gaussian function $f(x)=\exp(-\frac{\|x-c_i\|^2}{2\sigma^2})$ is adopted in many applications. For the large enough neurons in the hidden layer, it can prove that it can be sufficiently approximated by the summation of RBFs, where each is located on a different point in the space [21].

$$y(x)=w_0+\sum_{i=1}^{k} w_i \exp(-\frac{(x-c_i)^T(x-c_i)}{2\sigma^2}) \tag{3.6}$$

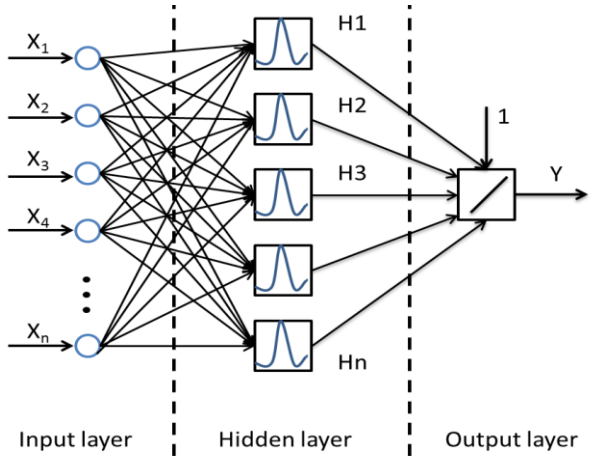


Figure 3.3 Schematic of radial basis function network

3.5 Case Studies and Analysis

3.5.1 Numeric experiment setup

In this paper, eight oil characteristic datasets are used for verifying the performance of the above four AI algorithms with integration of SMOTEBoost and bootstrap. The construction of these training datasets follows the methods discussed in Section 3.2. Table 3.1 summarizes the configurations of these eight datasets. Among these datasets, datasets 1, 2, 3, 5, and 6 were collected from different utility companies and dataset 4 was digested from the literature [4]. Dataset 7 is the combination of datasets 1, 2, 3 and 4 with a few incompatible records removed. In dataset 8, about two thirds of the samples are from dataset 5 and one third of the records are from dataset 6. The purpose of combining datasets collected from different sources is to evaluate the generalization capability of AI algorithms. The algorithm which is trained on one dataset but can achieve consistent classification accuracy on another dataset is regarded as having higher generalization capability.

In Table 3.1, the features (input data of AI algorithms) of datasets 1, 2, 3, 4 and 7 are the concentrations of dissolved gases (i.e. hydrogen (H_2), methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2)) and the transformer conditions include normal deterioration, discharge, partial discharge and thermal faults (refer to Section 3.2). The features in datasets 5, 6 and 8 include the concentrations of dissolved gases, i.e. C_2H_4, C_2H_2, C_2H_6, H_2, CH_4, carbon monoxide (CO) and carbon dioxide (CO_2) and oil test data of water content, acidity, dielectric dissipation factor, 2-furfuraldehyde, resistivity and breakdown voltage. In these three datasets, transformer insulation conditions are categorized into four levels of “excellent”, “good”, “fair” and “poor” condition (refer to Section 3.2) [14]. The reason for selecting the gas concentration rather than the gas ratio as input

features is that in some cases it may not be possible to calculate if the concentration of certain gas equals zero.

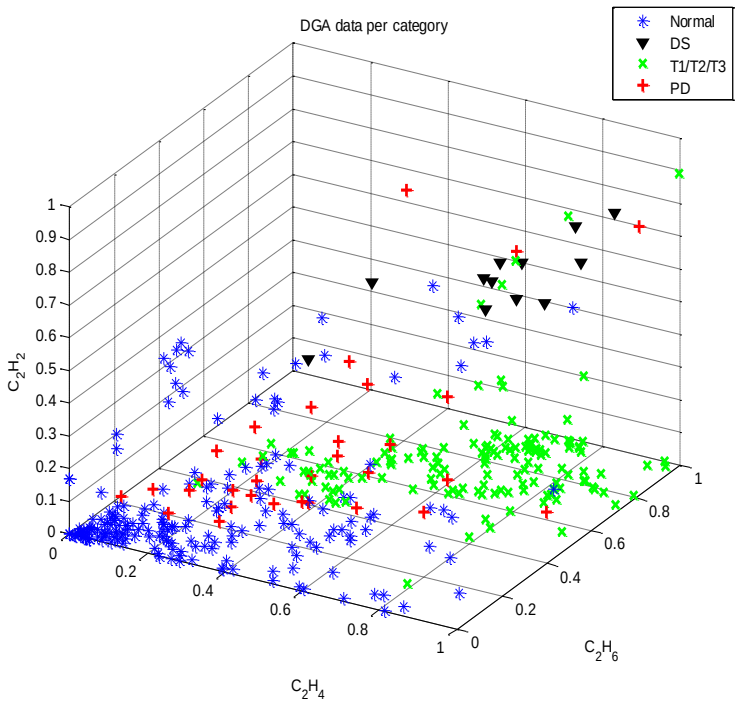
Table 3.1 Configuration of eight datasets

Dataset	Source	Samples	Features
1	Utility	390	dissolved gases concentrations
2	Utility	992	dissolved gases concentrations
3	Utility	181	dissolved gases concentrations
4	Publications	206	dissolved gases concentrations
5	Utility	181	dissolved gases + oil tests
6	Utility	882	dissolved gases + oil tests
7	Utility & Publications	1542	dissolved gases concentrations
8	Utility	875	dissolved gases + oil tests

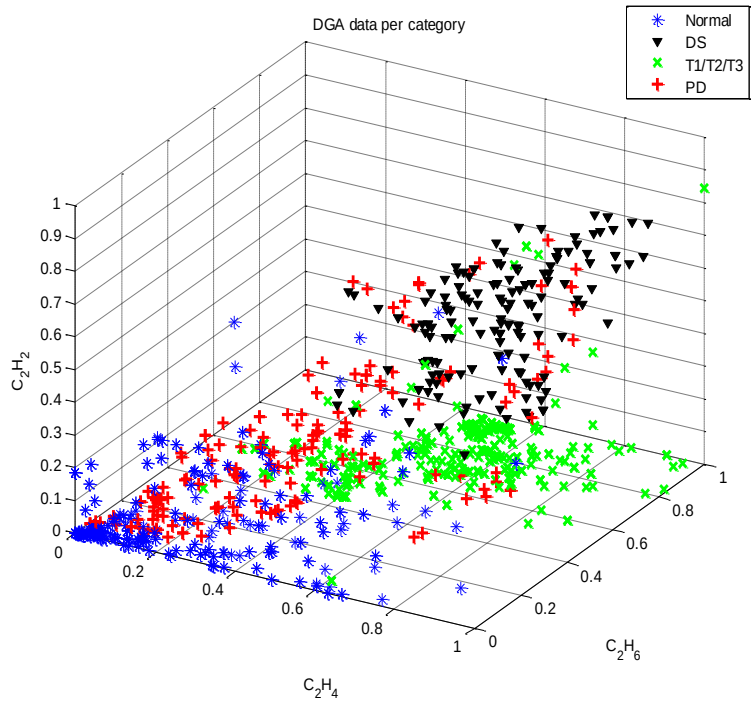
Table 3.2 describes the samples distribution of the original eight datasets. To exemplify the effect of SMOTE and bootstrap, Figure 3.4 provides the visualizations of sample distributions of Dataset 1 before and after applying SMOTE. It can be seen that the original Dataset 1 is class imbalanced, in which certain classes significantly outnumber other classes. After performing SMOTE, the sample numbers in different classes tend to be equal (normal deterioration-252, discharge-253, thermal fault-250, partial discharge-258). It can be observed from Figure 3.4(b) that the SMOTE does not replicate the samples in the majority class but synthetizes new samples for the minority class to achieve approximate equal distribution of samples throughout the whole dataset. Table 3.3 presents the sample distributions of eight datasets after applying SMOTE and bootstrap on the original datasets.

Table 3.2 Sample distribution of original dataset

Dataset	Normal deterioration	Discharge	Thermal	Partial Discharge
1	201	12	145	32
2	169	227	295	301
3	21	30	50	80
4	50	83	56	17
7	248	440	523	331
Dataset	**Excellent**	**Good**	**Fair**	**Poor**
5	80	50	21	30
6	165	188	266	263
8	80	316	284	195



(a)



(b)

Figure 3.4 Samples distribution of dataset 1 in Table 3.1 (a) original dataset; (b) after processed by SMOTE and bootstrap. (DS-discharge faults, OT-thermal faults, PD-partial discharge, Normal deterioration)

To evaluate the performance of four AI algorithms (C4.5, *k*NN, RBF, and SVM) with and without integrating SMOTEBoost and bootstrap, the following training and testing procedure is followed. Firstly, each dataset of Table 3.1 is divided into a training sub-dataset (70% samples) and a testing sub-dataset (30% samples). The samples in both datasets are normalized within [0, 1]. Then tenfold

cross validation is performed on the training dataset to determine the optimal values of the above four algorithms, including the hidden neurons number in RBF, the neighbors number for *k*NN, the parameters *C* and γ in SVM. For *k* fold cross-validation, the original dataset *D* is randomly split into *k* mutually exclusive subsets $D_1, D_2, \cdots D_k$ with approximately equal size. The algorithm is then trained on $D_1 \cdots D_{t-1}$ and validated on $D_t, t \in \{1, 2, \cdots, k\}$. Total *k* trained models can be obtained. The parameters of the trained model, which has the highest classification accuracy, are chosen to build the final model for performing classification on the testing dataset.

Table 3.3 Sample distribution after SMOTE and Bootstrap

Dataset	**Normal deterioration**	**Discharge**	**Thermal**	**Partial Discharge**
1	252	253	250	258
2	411	410	412	411
3	116	118	117	115
4	99	100	98	99
7	729	728	728	730
Dataset	**Excellent**	**Good**	**Fair**	**Poor**
5	107	108	107	108
6	378	379	377	378
8	436	435	437	435

After finding the optimal values of the above parameters, AI algorithms are trained on the whole training sub-dataset. Subsequently, the trained algorithms are applied to label each sample of the testing sub-dataset into particular transformer insulation conditions (excellent, good, fair and poor) or incipient fault types (discharge, partial discharge and thermal fault). The above procedure is repeated 50 times (runs). In each run, the classification accuracy is computed according to $N_{total}^{'} / N_{total}$ ($N_{total}^{'}$ denotes the number of samples which are correctly classified and N_{total} denotes the total number of samples of the whole dataset). The overall accuracy is the average over 50 runs. The software routines in [24-26] are tailored for the algorithmic implementation in this section.

3.5.2 Results and discussions

Table 3.4 presents the classification performance of SVM algorithm with and without integrating SMOTEBoost and bootstrap on the eight datasets. In this table, the overall classification accuracy

and the classification accuracy on each individual class (fault types or insulation conditions) are provided. This will help to verify the capability of SMOTEBoost in facilitating AI algorithm achieving consistent performance on both majority and minority classes.

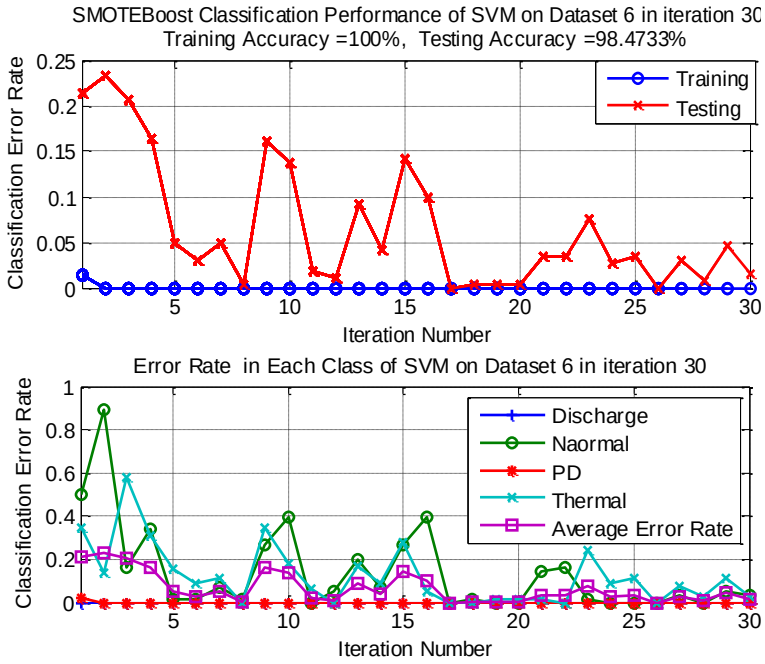


Figure 3.5 Classification error rate of the SVM with the integration with SMOTEBoost using oil characteristics dataset 6

Table 3.4 Comparison of classification accuracy of SVM over eight datasets (in percentage)

Accuracy	**Without integrating SMOTEBoost and bootstrap**					**Integrating SMOTEBoost and bootstrap**				
	Class 1	**Class 2**	**Class 3**	**Class 4**	**Overall**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Overall**
Dataset 1	85	67	87	70	82	100	100	100	100	100
Dataset 2	60	63	97	89	88	100	100	100	93	97
Dataset 3	56	50	83	61	58	100	100	100	100	100
Dataset 4	83	88	88	60	84	100	100	100	100	100
Dataset 5	83	70	58	78	76	100	100	100	100	100
Dataset 6	43	81	98	84	74	100	96	100	97	98
Dataset 7	54	75	71	85	71	100	86	90	91	91
Dataset 8	85	96	92	96	94	100	98	100	94	98

From Table 3.4 it can be seen that the original SVM algorithm (e.g. without integration) can attain reasonable overall classification accuracy over most datasets. To make a comparison of diagnosis accuracy between AI algorithms and conventional diagnostic methods, diagnosis

accuracy by using Duval triangle method is also calculated on Dataset 5 where the overall accuracy is 52%. This demonstrates the effectiveness of AI algorithms on improving insulation condition diagnosis of power transformers.

From Table 3.5 it can be seen that the classification accuracy of SVM on the samples belonging to the minority class is rather low. For datasets 3, 6 and 7, the original SVM attains less than or slightly above 50% accuracy on class 1. This is because these three datasets are class imbalance datasets, in which class 1 has less samples compared to other classes. In contrast, the SVM algorithm integrating SMOTEBoost and bootstrap consistently attains higher overall classification accuracy and accuracy on each class as well. For datasets 3, 6 and 7, this hybrid algorithm can achieve close to 100% accuracy for class 1. This demonstrates the advantages of SMOTEBoost and bootstrap in dealing with class imbalance problems and improving SVM classification accuracy. Figure 3.5 illustrates the classification error rate of the SVM with the integration of SMOTEBoost and bootstrap on dataset 6 over 30 iterations.

From Figure 3.5 it can be seen that after five iterations the overall error rate decreases to 5% (0.05, top graph) while the error rate on the minority class has changed to 15% (0.15, bottom graph). At the 30^{th} iteration, the overall error rate and the error rate of each individual class are reduced to around 2% (0.02) In Figure 3.5, the oscillation of error rate is due to the over fitting at some iterations in the training stage. Therefore, properly selecting the iteration number is very important. In the paper, the iteration number is set to 30 based on the tradeoff between computational cost and error rate.

The classification results of the other three algorithms, i.e. *K*NN, RBF and C4.5 with and without the integration of SMOTEBoost and bootstrap over the eight datasets are shown in Tables 3.5, 3.6 and 3.7. It can be observed that these three algorithms with the integration of SMOTEBoost and bootstrap can achieve higher classification accuracy compared to those without the integration. The only exception is the *K*NN algorithm with the integration of SMOTEBoost and bootstrap algorithm on dataset 8 (Table 3.5). The error rate of this hybrid algorithm on dataset 8 over 30 iterations is presented in Figure 3.6. It can be observed from Figure 3.6 that the *K*NN after integrating SMOTEBoost and bootstrap experiences a more serious over fitting problem for the minority class ("thermal" in the bottom graph) at the 30^{th} training iteration. The error rate of the minority class is increased to 30% (0.3) at the 30^{th} iteration from less than 5% (0.05) at the 29^{th} iteration.

This may be due to the fact that the performance of *K*NN is normally determined by the nearest neighbors of the samples rather than the samples themselves. At each iteration SMOTE generates new samples for the minority class and these new samples can be influenced by their neighbors,

which might be embedded with noise. This can compromise the classification accuracy of *K*NN on the minority class and in turn on the overall classification accuracy.

Table 3.8 presents the classification accuracy of the C4.5 decision tree algorithm integrating either SMOTEBoost or bootstrap alone. It can be seen that SMOTE and bootstrap alone can facilitate the C4.5 algorithm improving classification performance to some extent. However, their performance is still outperformed by the hybrid of SMOTEBoost and bootstrap.

Table 3.5 Comparison of classification accuracy of *K*NN over eight datasets (in percentage)

Accuracy	**Without integrating SMOTEBoost and bootstrap**					**Integrating SMOTEBoost and bootstrap**				
	Class 1*	**Class 2***	**Class 3***	**Class 4***	**Overall**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Overall**
Dataset 1	86	97	91	78	87	100	100	100	89	99
Dataset 2	70	56	98	89	87	100	88	100	94	96
Dataset 3	71	43	75	78	65	100	93	100	89	96
Dataset 4	77	94	75	80	83	100	92	94	80	93
Dataset 5	83	93	67	100	87	100	100	100	89	98
Dataset 6	80	53	97	94	82	100	93	98	88	95
Dataset 7	54	84	87	84	79	100	84	86	91	89
Dataset 8	92	92	97	99	95	100	95	69	98	88

*For Dataset 1, 2, .3, 4,7, Class 1 to Class 4 denote "Normal condition", "Discharge fault", "Thermal fault" and "Partial discharge fault" respectively. For Dataset 5, 6, 8, Class 1 to Class 4 denote transformer insulation condition of "Excellent", "Good", "Fair" and "Poor".

Table 3.6 Comparison of classification accuracy of RBF over eight datasets (in percentage)

Accuracy	**Without integrating SMOTEBoost and bootstrap**					**Integrating SMOTEBoost and bootstrap**				
	Class 1	**Class 2**	**Class 3**	**Class 4**	**Overall**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Overall**
Dataset 1	87	67	95	84	89	100	100	100	100	100
Dataset 2	93	31	41	65	74	100	100	100	84	95
Dataset 3	81	33	83	67	66	100	87	100	100	96
Dataset 4	80	85	28	90	69	100	100	88	100	97
Dataset 5	83	87	75	61	80	100	93	100	100	98
Dataset 6	83	49	99	93	82	100	100	100	96	99
Dataset 7	71	85	88	86	84	97	89	83	92	89
Dataset 8	90	94	96	95	95	100	97	94	95	96

Table 3.7 Comparison of classification accuracy of C4.5 decision tree over eight datasets (in percentage)

Accuracy	Without integrating SMOTEBoost and bootstrap					Integrating SMOTEBoost and bootstrap				
	Class 1	**Class 2**	**Class 3**	**Class 4**	**Overall**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Overall**
Dataset 1	90	60	93	76	89	100	100	100	67	97
Dataset 2	62	32	99	88	82	100	46	98	100	87
Dataset 3	85	50	58	83	72	100	93	100	100	98
Dataset 4	80	90	50	30	72	100	96	100	100	98
Dataset 5	94	90	83	94	92	100	73	100	100	93
Dataset 6	7	64	99	97	63	100	98	94	100	98
Dataset 7	72	86	88	84	84	100	89	83	92	91
Dataset 8	83	96	97	94	95	100	97	94	95	96

Table 3.8 Comparison of classification accuracy improvement with SMOTE and Bootstrap (C4.5 decision tree)

Accuracy	Integrating SMOTE only					Integrating bootstrap only				
	Class 1	**Class 2**	**Class 3**	**Class 4**	**Overall**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Overall**
Dataset 1	90	65	92	80	89	90	63	92	78	89
Dataset 2	63	35	98	90	84	65	33	99	87	85
Dataset 3	87	65	65	83	75	86	55	60	80	72
Dataset 4	83	91	60	35	76	80	93	55	32	74
Dataset 5	94	91	85	90	92	92	90	84	94	91
Dataset 6	50	66	98	98	82	37	65	97	97	78
Dataset 7	78	88	88	85	85	73	86	88	85	84
Dataset 8	88	96	97	92	95	84	95	93	93	93

Besides the classification accuracy, precision and recall are also performance indicators. Table 3.9 presents the results of precision and recall for the hybrid of *K*NN and SMOTEBoost and bootstrap algorithm. It can be observed that the precisions and recall of this hybrid algorithm over each class are quite high (except Class 3 of Dataset 8, the reason has been explained in Table 3.5). This indicates *K*NN and SMOTEBoost and bootstrap hybrid algorithm can attain desirable accuracy.

To further evaluate the performance of SVM, *K*NN, RBF and C4.5 with the integration of SMOTEBoost and bootstrap, a statistical comparison is performed in this section. Firstly, the four algorithms are ranked for each dataset: the algorithm having the highest classification accuracy is ranked as one and the second best algorithm is ranked as two and so on. Then the averaged rank of

each algorithm is computed as $R_j = \frac{1}{N}\sum_{i=1}^{N} r_i^j$, where r_i^j denotes the rank of the *j*-th algorithm (total $k = 4$ algorithms) on the *i*-th dataset (total $N = 8$ datasets). The average rank is SVM 1.50, RBF 2.63, C4.5 2.88 and *K*NN 3.

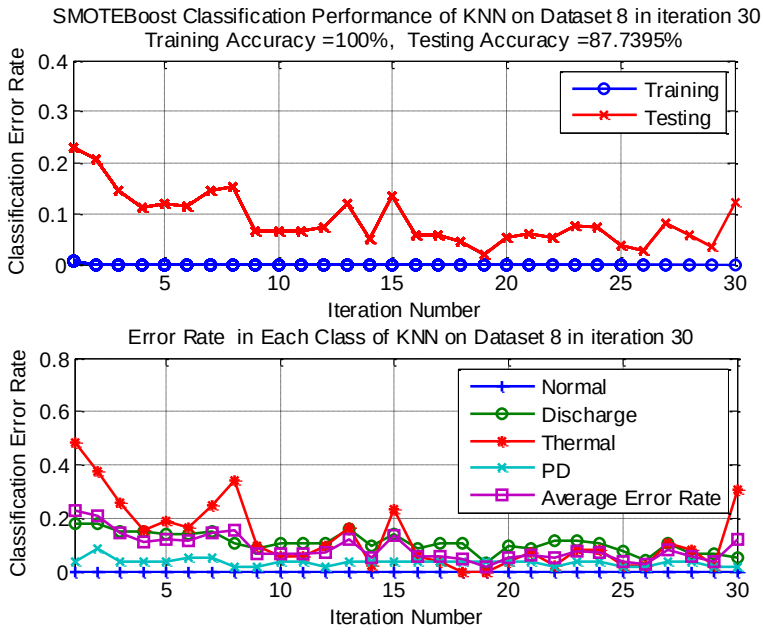


Figure 3.6 Classification error rate of the KNN classifier with the integration of SMOTEBoost and bootstrap

Table 3.9 Precision and recall of *K*NN over eight datasets (in percentage)

Precision/Recall	Class 1	Class 2	Class 3	Class 4
Dataset 1	100/100	100/100	100/100	100/89
Dataset 2	100/100	100/88	100/100	100/94
Dataset 3	100/100	100/93	100/100	100/89
Dataset 4	100/100	100/92	100/94	100/80
Dataset 5	100/100	100/100	100/100	100/89
Dataset 6	100/100	100/93	100/98	100/88
Dataset 7	100/100	100/84	100/86	100/91
Dataset 8	100/100	100/95	100/69	100/98

To formally judge whether SVM is the best among all four hybrid algorithms the Benferroni–Dunn test is used. In this test, a critical difference (*CD*) value is defined as $CD = q_\alpha\sqrt{\frac{k(k-1)}{6N}}$ where

k=4 (the number of algorithms) and *N*=8 (the number of datasets). q_α value is 2.128 at $\alpha = 0.1$ (90% confidence coefficient). The corresponding CD value is 1.3736. Since the rank difference between SVM and *K*NN and C4.5 is larger than this CD value, it is able to claim that the SVM integrating with SMOTEBoost and bootstrap is significantly better than the C4.5 and *K*NN integrating with SMOTEBoost and bootstrap for transformer insulation diagnosis using oil characteristics. Moreover, the SVM integrating with SMOTEBoost and bootstrap is better than RBF integrating with SMOTEBoost and bootstrap.

3.5.3 Generalization capability validation

To evaluate the generalization ability of the above four algorithms with and without the integration of SMOTEBoost and bootstrap, these algorithms are trained on one dataset and tested on another dataset. The two datasets are collected from two different sources (e.g. utility companies).

Tables 3.10 and 3.11 present the results of such cross training and testing. In Table 3.10 the four algorithms are trained on dataset 4 but tested on dataset 1 (Table 3.1). In Table 3.11, the four algorithms are trained on dataset 6 but tested on dataset 5.

Table 3.10 Comparison of generalization ability of different AI algorithms (trained on Dataset 4, tested on Dataset 1)

Accuracy (in percentage)		**Normal**	**DS**	**OT***	**PD**	**Overall**
Original	C4.5	3	58	6	84	13
	*K*NN	33	100	62	82	50
	RBF	31	92	68	74	50
	SVM	35	100	68	79	52
SMOTE-Boost	C4.5	100	100	99	97	99
	*K*NN	100	100	92	97	97
	RBF	75	92	97	88	85
	SVM	100	100	100	100	100

*OT: thermal fault in transformer insulation

From Tables 3.10 and 3.11 it can be observed that the original algorithms (i.e. not integrating with SMOTEBoost and bootstrap) fail to make classification on the testing dataset. This indicates the generalization capability of the four original algorithms is limited. This is because there exists

inconsistency among data in the training datasets collected from different utilities (refer to the next section). The four algorithms are unable to deal with such inconsistency. However, by integrating SMOTEBoost and bootstrap these algorithms can attain much higher classification accuracy on the testing dataset. The comparison results indicate that SMOTEBoost and bootstrap can facilitate an AI algorithm improving its generalization capability for power transformer diagnosis using oil test data.

Table 3.11 Comparison of generalization ability of different AI algorithms (trained on Dataset 6, tested on Dataset 5)

Accuracy		Excellent	Good	Fair	Poor	Overall
Original	C4.5	33	0	0	0	14
	*K*NN	29	31	95	0	33
	RBF	29	11	29	2	20
	SVM	3	4	100	0	14
SMOTE-Boost	C4.5	100	100	100	97	98
	*K*NN	100	84	95	35	84
	RBF	100	58	100	74	84
	SVM	100	100	100	90	98

3.5.4 Remarks on training dataset construction

The construction of a training dataset is crucial for ensuring AI algorithms to consistently attain desirable classification accuracy. In constructing the training dataset, it needs to use conventional methods to interpret DGA and other test data and then determine incipient fault types or insulation conditions for the transformers included in the training dataset. However, inconsistencies can arise from using these conventional methods.

(1) Different conventional methods utilise different key gas ratios and the utility companies have to adapt these ratio-based methods heuristically.

(2) The empirical nature of the conventional methods leads to the discrepancies in interpretation. For the same test data, different methods may produce different diagnoses.

(3) The conventional methods may not be able to provide the interpretation for every possible combination of gas ratio values and thus it cannot make diagnoses in these cases.

(4) On some occasions, the interpretation of test data may rely on expert judgement. However,

each expert may have his/her own idea on what is happening inside the transformer based on the provided data and information.

(5) In the presence of multiple faults, some test records can be fully assigned to one of the fault types while others may not be assigned exactly to one of the fault types.

As a result, the training datasets built upon the data collected from different utility companies may exhibit inconsistency and discrepancy. The general AI algorithms lack the capability of handling such inconsistency in a training dataset. SMOTEBoost involves multiple iterations to train an AI algorithm and it redirects the training iteration of the algorithm to focus on the samples incorrectly classified previously. Moreover, SMOTEBoost can effectively deal with the class imbalance problem. Therefore, AI algorithms with the integration of SMOTEBoost can consistently attain better classification accuracy.

3.6 Conclusion

This paper proposed a novel data pre-processing approach, the hybrid of SMOTEBoost and bootstrap, for enhancing AI algorithms to consistently achieve desirable classification accuracy in power transformer diagnosis using oil test data. The SMOTEBoost and bootstrap were integrated with a number of AI algorithms to classify the types of incipient faults or the conditions of the transformer insulation system on eight oil test datasets collected from different utility companies. The results demonstrated the proposed approach can improve the generalization capability of AI algorithms when these algorithms are applied for transformer diagnosis.

Acknowledgment

The authors gratefully acknowledge the contributions of Australian Research Council, Powerlink Queensland, Energex, Ergon Energy and TransGrid.

References

[1] Z. Xiang and E. Gockenbach, "Asset Management of Transformers Based on Condition Monitoring and Standard Diagnosis," *IEEE Electr. Insul. Mag.*, vol.24, Issue 4, pp. 26-40, 2008.

[2] M. Wang, A. J. Vandermaar and K. D. Srivastava, "Review of condition assessment of power transformers in service," *IEEE Electr. Insul. Mag.*, vol.18, Issue 6, pp. 12-25, 2002.

[3] M. Duval, "New techniques for dissolved gas-in-oil analysis," *IEEE Electr. Insul. Mag.*, vol.19, Issue 2, pp. 6-15, 2003.

[4] M. Duval and A. DePablo, "Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases," *IEEE Electr. Insul. Mag.*, vol.17, Issue 2, pp. 31-41, 2001.

[5] H. Ma, C. Ekanayake and T. K. Saha, "Power transformer fault diagnosis under measurement originated uncertainties," *IEEE Trans. Dielectr. Electr. Insul.*, vol.19, Issue 6, pp. 1982-1990, 2012.

[6] D. R. Morais and J. G. Rolim, "A hybrid tool for detection of incipient faults in transformers based on the dissolved gas analysis of insulating oil," *IEEE Trans. Power Delivery*, vol.21, Issue 2, pp. 673-680, 2006.

[7] W. Chen, C. Pan, Y. Yun and Y. Liu, "Wavelet Networks in Power Transformers Diagnosis Using Dissolved Gas Analysis," *IEEE Trans. Power Delivery*, vol.24, Issue 1, pp. 187-194, 2009.

[8] R. Naresh, V. Sharma and M. Vashisth, "An integrated neural fuzzy approach for fault diagnosis of transformers," *IEEE Trans. Power Delivery*, vol.23, Issue 4, pp. 2017-2024, 2008.

[9] Q. Su, C. Mi, L. L. Lai and P. Austin, "A fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer," *IEEE Trans. Power Syst.*, vol.15, Issue 2, pp. 593-598, 2000.

[10] N. V. Chawla, F. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol.16, Issue 16, pp. 321-357, 2002.

[11] *IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers*, C57.104, 2008.

[12] *Mineral Oil-impregnated Electrical Equipment In Service – Guide to the Interpretation of Dissolved and Free Gases Analysis*, IEC60599, 2007.

[13] *IEEE Guide for Acceptance and Maintenance of Insulating Oil in Equipment*, IEEE Std C57.106-2006, 2006.

[14] A. D. Ashkezari, H. Ma, T. K. Saha and C. Ekanayake, "Application of fuzzy support vector machine for determining the health index of the insulation system of in-service power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 3, pp. 965-973, 2013.

[15] G. Ratsch, T. Onoda and K. R. Muller, "Soft margins for AdaBoost," *Mach. Learn.*, vol.42, Issue 3, pp. 287-320, 2001.

[16] N. V. Chawla, A. Lazarevic, L. O. Hall and K. W. Bowyer, "SMOTEBoost: Improving prediction of the minority class in boosting," in *Proceedings of Proceedings of Principle of Knowledge Discovery in Databases, September 22-26*, 2003,Croatia, pp. 107-119.

[17] M. T. Kai, "An instance-weighting method to induce cost-sensitive trees," *IEEE Trans. Knowl. Data Eng.*, vol.14, Issue 3, pp. 659-665, 2002.

[18] A. M. Zoubir and B. Boashash, "Bootstrap and its application in signal processing," *IEEE Signal Process. Mag.*, vol.15, Issue 1, pp. 56-76, 1998.

[19] B. Efron and R. Tibshirani, *An Introduction to the Bootsrap*, London: U.K.: Chapman & Hall, 1993.

[20] N. Cristianini and J. Shawe-Taylor, *An Introduction of Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, 2001.

[21] C. M. Bishop, *Pattern Recognition and Machine Learning*, New York: Springer, 2006.

[22] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol.13, Issue 1, pp. 21-27, 1967.

[23] A. Navada, A. N. Ansari, S. Patil and B. A. Sonkamble, "Overview of use of decision tree algorithms in machine learning," in *Proceedings of 2011 IEEE Control and System Graduate Research Colloquium, ICSGRC 2011, June 27 - June 28*, 2011,Shah Alam, Malaysia, pp. 37-42.

[24] I. T. Nabney, *Netlab Algorithms for Pattern Recognition*, Springer, 1995.

[25] C. C. Chang and C. J. Lin,"LIBSVM: a library for support vector machines" available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[26] J. Demsar, "Statistical comparisons of classifiers over multiple data sets," *J. Machine Learning Research*, vol.1, Issue 1, pp. 1-30, 2006.

Chapter 4

Modelling Approach for Investigation of Moisture Dynamics in Power Transformers

Contribution of the Chapter

This chapter introduces two modelling approaches for investigating moisture dynamics in power transformers. Firstly, a multi-physics finite element modelling (FEM) approach is developed where moisture dynamics is investigated by coupling the electromagnetic, thermal, fluid flow and moisture migration physics simultaneously. To overcome the numeric stability of FEM method in moisture dynamics modelling, an alternative particle tracing method is proposed which formulates moisture diffusion from a microscopic view of water molecules' motion. The proposed modelling approaches are applied to model moisture dynamics in both pressboard specimens and cellulose insulation of a prototype transformer. Extensive ageing and moisture diffusion experiments have been conducted on a prototype transformer to verify the proposed modelling approaches.

4.1 Multi-physics Modelling Approach for Investigation of Moisture Dynamics in Power Transformers

Multi-physics Modelling Approach for Investigation of Moisture Dynamics in Power Transformers

Yi Cui[1], Hui Ma[1], Tapan Saha[1], Chandima Ekanayake[1] and Guangning Wu[2]

1. The University of Queensland, Brisbane, Australia

2. Southwest Jiaotong University, Chengdu, China

Publication Journal: IET Generation Transmission & Distribution

Submitted: September 8, 2015

Revision Resubmitted: December 1, 2015

Accepted: February 6, 2016

Author's Contributions:

Yi Cui: Primary model design, theoretical simulation, experimental validation, measurement results analysis and manuscript preparation.

Hui Ma: Supervision on the experimental measurement, results discussion and manuscript revision.

Tapan Saha: Supervision on the experimental measurement, results discussion and manuscript revision.

Chandima Ekanayake: Results discussion and manuscript revision.

Guangning Wu: Results discussion.

Abstract

The presence of moisture in a power transformer is harmful to its oil-cellulose system and can also accelerate the ageing. Therefore, it is necessary to estimate the moisture content inside the transformer. Due to the variation of loading conditions and ambient temperatures, moisture is neither in a static state nor uniformly distributed in transformers. Moisture migrates between oil and cellulose and moves inside oil and cellulose of transformers. Moreover, given the complexity of a transformer's structure, moisture distribution is also a function of mass and dimensions of transformer's oil-cellulose system. This paper adopts a multi-physics modelling approach to investigate moisture dynamics in transformers. In this approach, moisture dynamics is investigated by coupling electromagnetic, thermal, fluid flow and moisture migration physics simultaneously. An accurate estimation on moisture distribution in transformers can then be obtained. Extensive ageing and moisture diffusion experiments have been conducted on a prototype transformer to verify the proposed approach.

Index Terms: Cellulose, diffusion, insulation, moisture, multi-physics, oil, transformer

4.1.1 Introduction

Moisture in transformers has numerous adverse effects on transformers' oil-cellulose system [1]. Moisture breaks celluloses chain and reduces cellulose dielectric and mechanical strength, decreases Partial Discharge Inception Voltage (PDIV), accelerates cellulose ageing and leads to bubble formation in oil at elevated temperatures [2-3]. Thus, accurately evaluating moisture in a transformer is of great concern for utilities.

Due to variations of ambient temperatures and loading conditions, moisture in transformer is normally not equilibrium but migrates between cellulose and oil. Cellulose absorbs water from oil with the decrease of temperature or desorbs water to oil with the increase of temperature. Moisture also moves inside cellulose, which mainly involves a diffusion process decided by moisture gradients. Furthermore, moisture distribution is a function of mass and dimensions of transformer's insulation system. To accurately estimate moisture concentration in a transformer, it is necessary to investigate moisture dynamics by considering various factors especially temperature and interactions between different physics that influence water movement.

Moisture in oil can be directly measured through moisture-in-oil sensor and then moisture in cellulose can be estimated through equilibrium charts [4-6]. This approach assumes moisture equilibrium in a transformer, which requires a constant temperature to be maintained over a considerable long time. However, this is almost impossible for an operational transformer. As such, moisture estimated from equilibrium charts might be erroneous [7-8]. Some researchers investigated moisture diffusion in transformers using Fick's law and determined diffusion coefficients of cellulose. However, these investigations were mainly based on the experimental results of pressboard samples under well-controlled conditions. They may not be readily applicable for estimating moisture content in transformers.

This paper proposes a multi-physics approach to investigate moisture dynamics in transformers' oil-cellulose system. It considers the effects of electromagnetic, thermal, fluid flow and moisture migration physics on moisture dynamics in transformers. Especially it takes into account the coupling and interactions of these physics as they collectively influence the moisture dynamics inside a transformer. By using this approach, three dimensional (3D) complex temperature driven moisture dynamics in oil and cellulose can be revealed. Extensive accelerated ageing and moisture diffusion experiments are conducted on a prototype transformer to demonstrate the advantages of multi-physics approach in estimating moisture content in transformers.

4.1.2 Modelling Moisture Dynamics in Transformers

For an operational transformer, moisture dynamics can be modelled as a mass transfer and diffusion process. Since it is difficult to model moisture dynamics at a microscopic scale, therefore, in this paper, moisture dynamics is modelled as a collective diffusion process at the macroscopic scale by solving Fick's equation as (4.1) [7].

$$\frac{\partial C(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial C(x,t)}{\partial x}\right) \tag{4.1}$$

where $C(x,t)$ denotes the moisture concentration of cellulose at position x and time t. D denotes the diffusion coefficient, which is usually not a constant but influenced by local moisture concentration and temperature.

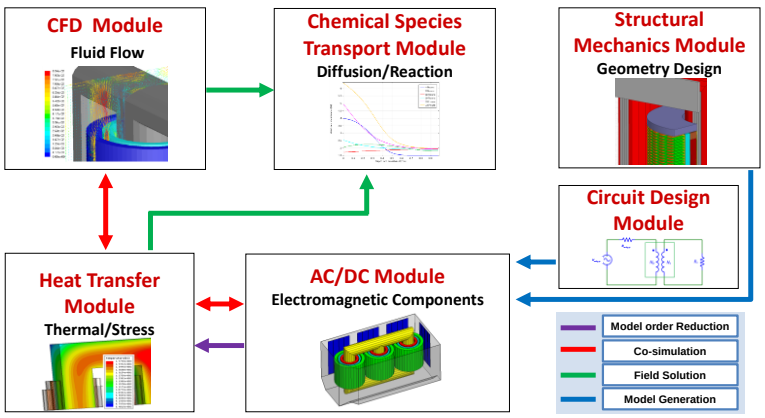
Extensive studies have been performed to understand diffusion mechanism and determine diffusion coefficients. Ast investigated the diffusion coefficient of Kraft paper under different moisture concentrations and temperatures by using permeation method [9]. Guidi and Fullerton

found an empirical relationship between the diffusion coefficient and the local moisture concentration and temperature [10-11]. Howe investigated the diffusion coefficients for both Manila paper and pressboard [11]. Asem determined diffusion coefficients for oil-immersed paper and non-immersed pressboard [12]. Foss determined a number of diffusion coefficients for both impregnated and non-impregnated Kraft paper by using other researchers' data [10]. By using dielectrometric sensors and mathematical tools to solve Fick's equation, Du calculated the diffusion coefficient for non-impregnated pressboard [4, 13]. Over the past several years, García reported a series of diffusion coefficients [14-16], where the thickness of the oil-impregnated pressboard was taken into consideration.

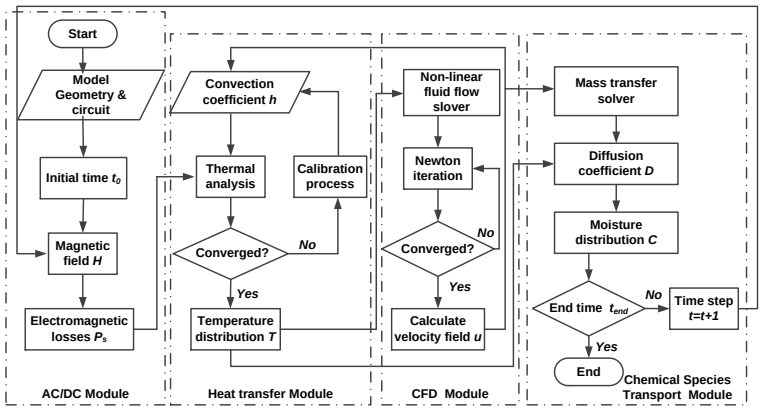
Though the above moisture diffusion approach gains some success in estimating moisture contents in power transformers, they still suffer some limitations. Especially, the above temperature dependent diffusion coefficients are estimated by conducting laboratory experiments on pressboard/paper samples, which is highly dependent on the experimental setup. This may also introduce some discrepancies when applying these diffusion coefficients to the transformers constructed with different cellulose materials and heterogeneous temperature distribution. Moisture diffusion in oil-cellulose system of a transformer involves different physics. Therefore, multiple physics modelling should be conducted to provide accurate moisture estimation.

4.1.3 Multi-physics Modelling of Moisture Dynamics

This paper adopts COMSOL Multiphysics software package to investigate different physics and their coupling mechanism governing the moisture distribution inside a transformer. Figure 4.1a presents the schematic diagram of multiple physics considered in this paper.



(a)



(b)

Figure 4.1 Coupling mechanism and flowchart of multi-physics model (a) Coupling mechanism among different physics including electromagnetic, thermal, fluid flow and moisture migration (b) Flowchart of implementing multi-physics model

In model generation procedure (Figure 4.1a), structural mechanics module constructs transformer's geometry while circuit design module configures an external circuit for transformers ("model generation"). Then AC/DC module, heat transfer module and computational fluid dynamics (CFD) module are invoked to co-simulate the electromagnetic field, thermal distribution and fluid flow of transformer simultaneously using iterative numeric computation ("co-simulation"). Certain simplifications are made in thermal modelling to reduce the computation load given the accuracy of thermal analysis can still be ensured ("model order reduction"). Finally all above results are fed into chemical species transport module for studying the evolution of moisture migration by diffusion and convection at each time instance ("field solution"). The theoretical foundations of the above physics as well as their coupling mechanisms are reviewed in the remaining of this section.

4.1.3.1 Electromagnetic Field Analysis

For an operational transformer, its magnetic field distribution can be derived from the Maxwell's equations as (4.2) [17].

$$\frac{\partial}{\partial x}\left(\frac{1}{\lambda_x}\frac{\partial \mathbf{A}}{\partial x}\right)\mathbf{x}+\frac{\partial}{\partial y}\left(\frac{1}{\lambda_y}\frac{\partial \mathbf{A}}{\partial y}\right)\mathbf{y}+\frac{\partial}{\partial z}\left(\frac{1}{\lambda_z}\frac{\partial \mathbf{A}}{\partial z}\right)\mathbf{z}=\sigma_{dc}\frac{\partial \mathbf{A}}{\partial t}-J_{sc} \tag{4.2}$$

where $\mathbf{A}$ denotes the magnetic vector potential, λ_x , λ_y and λ_z represent the material permeability in three directions, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ denote the unit vectors, σ_{dc} denotes the conductivity of copper conductors in winding. The temperature dependent conductivity is shown as (4.3) [18].

$$\sigma(T) = \sigma(T_0) - \alpha_T (T - T_0) \tag{4.3}$$

where $\sigma(T)$ and $\sigma(T_0)$ denote conductivity at temperature T and reference temperature T_0. $\sigma(T_0) = 5.96 \times 10^7 (S/m)$, $T_0 = 293K$, α_T=3.862×10^{-3} denotes the temperature coefficient.

In (4.2), J_{sc} denotes the current density in transformer winding. In electromagnetic field simulation, primary winding with voltage $V_{wdg-p} = \sqrt{2} \times 2000 \sin(314t)(V)$ is set as the excitation and the current density of the winding can be calculated as (4.4).

$$J_{sc} = \frac{N_p \left(V_{wdg-p} + V_{ind-p}\right)}{A_w R_{wdg-p}} \tag{4.4}$$

where V_{wdg-p} denotes the applied voltage on primary winding, V_{ind-p} denotes the induced voltage, which is calculated by integrating the electric field along the winding. N_p=765 denotes the turns number of the primary winding, A_w denotes the total cross-sectional area of the winding domain, R_{wdg-p} denotes the resistance of the primary winding. The values of A and R_{wdg-p} at each time instance are calculated by COMSOL, which considers the temperature dependent conductivity and arrangement of conductors in the primary winding.

4.1.3.2 Thermal Analysis

During the operation of a transformer, various power losses such as joule loss ($I^2 R_{loss}$), stray loss, and eddy loss occur in its winding. These losses will elevate temperature inside the transformer. Among these losses, stray loss is not considered in thermal analysis since it cannot directly increase the temperature of the winding [19]. By contrast, eddy loss is considered as it can influence the distribution of current density and power loss in the winding, and subsequently affect the temperature rise under particular conditions [20].

The thermal transfer in solid material can be calculated by as (4.5) [21].

$$k_x \frac{\partial^2 T}{\partial x^2} + k_y \frac{\partial^2 T}{\partial y^2} + k_z \frac{\partial^2 T}{\partial z^2} + Q_{sc} = \rho_m c_m \frac{\partial T}{\partial t} \tag{4.5}$$

where T denotes temperature and it is a function of location (x, y, z) and time t , k_x , k_y and k_z denote the thermal conductivity in three directions, Q_{sc} denotes the total heat source, ρ_m denotes

mass density and c_m denotes the heat capacity.

The boundary conditions of heat convection and radiation along the surface of the winding conductor are determined by (4.6) and (4.7).

$$\phi_{conv} = h\left(T - T_{amb}\right) \tag{4.6}$$

$$\phi_{rad} = \varsigma k_B T^4 - \varphi\xi \tag{4.7}$$

where ϕ_{conv} and ϕ_{rad} denote the heat flux of conductor's surface by convection and radiation respectively, h denotes the convection coefficient, ς denotes the surface emissivity, k_B denotes the Stefan–Boltzmann constant, φ denotes the absorptivity, ξ denotes the irradiation and T_{amb} denotes the ambient temperature.

4.1.3.2.1 Determination of convection coefficient: The convection coefficient (*h*) can be computed according to Nusselt number (N_u), i.e., the ratio of transferred heat by convective process and by conductive process across the boundary. In transformer, vertical parallel plate model is employed to model the geometry of oil ducts between windings. It is assumed that compared with the distance between two plates (d_p), the vertical lengths (l) of the plates are much larger.

For such symmetric isoflux, Nusselt number N_u can be calculated as (4.8) [21].

$$N_u = \frac{hl}{k_F} = \sqrt{0.144 \times R_e^* \times \frac{d_p}{d_o}} \tag{4.8}$$

where k_F denotes the fluid thermal conductivity, l denotes the geometry length, R_e^* denotes the Rayleigh number for the symmetric isoflux. d_p denotes the distance between two plates and d_o denotes the width of oil ducts.

4.1.3.2.2 Determination of thermal radiation: The heat exchange by radiation can be expressed as surface radiosity as (4.9), which indicates the total radiant energy transferring from one surface to another [19].

$$\phi_{rad,i} = \sum_{j=1}^{N} \frac{R_i - R_j}{\left(S_i V_{ij}\right)^{-1}} = \frac{E_{bi} - R_i}{\left(1-\varsigma_i\right)\varsigma_i S_i} \tag{4.9}$$

where $\phi_{rad,i}$ denotes a net radiative heat transfer rate of the *i*-th surface, S_i denotes the *i*-th surface area, V_{ij} denotes view factor from the *i*-th surface to the *j*-th surface, R_i and R_j denote the radiosity of the *i*-th surface and the *j*-th surface respectively, and E_{bi} is the emissive power for the *i*-th

surface. The above equation is used to describe the radiation between winding surfaces.

4.1.3.3 Fluid Flow Analysis

In fluid flow analysis mineral oil is treated as viscous fluid and modelled by Navier- Stokes equation as (4.10)[22].

$$\rho_F\left(\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\cdot\mathbf{v}\right)=-\nabla\cdot P+\mu\nabla^2\mathbf{v}+\mathbf{F} \tag{4.10}$$

where ρ_F denotes the fluid density, $\mathbf{v}$ denotes the fluid velocity, P denotes the pressure, and μ denotes dynamic viscosity and $\mathbf{F}$ denotes the strain rate tensor of Newtonian fluids, which can be expressed as (4.11).

$$\mathbf{F}=2\mu_{\bar{\varepsilon}}-\frac{2}{3}\mu(\nabla\cdot\mu)\mathbf{I}-P\mathbf{I} \tag{4.11}$$

where $\mathbf{I}$ denotes the unit tensor and $\mu_{\bar{\varepsilon}}$ denotes the linearized strain rate tensor. In case of an incompressible material, P is considered as constant [22]. The detailed parameters for thermal and fluid flow modelling are provided in Section 4.1.4.3.

4.1.3.4 Moisture Migration Between Oil and Cellulose System in Transformers

If all water molecules in a material are free to migrate, they tend to diffuse from the region of high moisture concentration to the region of lower concentration, thereby reducing the moisture gradient and equalizing the moisture concentration [4]. Chemical species transport module (Figure 4.1) is used for studying the evolution of moisture transported by diffusion and convection as well as migration due to a fluid flow field. As moisture concentration in oil-cellulose system is usually less than 10%, moisture dynamics is treated as diluted species transport. The driving forces for such moisture dynamics are diffusion (governed by Fick's law) and convection (due to coupling to fluid flow). The moisture diffusion and convection can be expressed as (4.12) [23].

$$\frac{\partial C(x,t)}{\partial t}=\frac{\partial}{\partial x}\left(D_t\frac{\partial C(x,t)}{\partial x}\right)-\mathbf{v}\cdot\frac{\partial C(x,t)}{\partial x} \tag{4.12}$$

where $\mathbf{v}$ denotes the velocity vector of fluid flow of mineral oil.

The diffusion coefficient in (4.12) can be expressed as(4.13).

$$D = D_0 e^{[k \times c + E_a (\frac{1}{T_0} - \frac{1}{T})]} \tag{4.13}$$

where $D_0 = 1.34 \times 10^{-13} m^2 / s$ and $k = 0.5$, E_a is the activation energy which equals to 8074 *kJ/mol* [16], T is the measured temperature in Kelvin and T_0 equals to 298 K.

The left side of (4.12) corresponds to the accumulation of the transported moisture. The first term on the right side of (4.12) accounts for the moisture transported through diffusion. It can be determined once the diffusion coefficient *D* is known. The second term on the right hand side of (4.12) describes the moisture transported by convection due to a velocity field ($\mathbf{v}$) of mineral oil. This field can be computed by using COMSOL's CFD module.

4.1.3.5 Coupling Between Electromagnetic, Thermal, Fluid Flow and Moisture Migration Models

Electromagnetic field distribution of transformers can be regarded as thermal dependent since the resistivity of the medium in transformers (iron core, cellulose and mineral oil) are influenced by temperature. The correlation between resistivity and temperature can be expressed in (4.14).

$$\rho_r (T) = \rho_{r0} \left[1 + \alpha_t (T - T_0) \right] \tag{4.14}$$

where ρ_{r0} denotes the resistivity at the reference temperature T_0 , and α_t denotes coefficient between resistivity and temperature.

The heat sources as described in (4.5) are also temperature dependent. This implies a strong link between the electromagnetic physics and heat transfer phenomenon. A number of oil characteristics such as viscosity, mass density, heat conductivity and internal pressure are all influenced by temperature. On the other hand, these characteristics also have considerable effect on transformer's temperature distribution. Therefore, the thermal analysis should be accomplished by coupling both heat transfer and fluid flow field. Moreover, as mentioned earlier, moisture dynamics is highly determined by temperature and thus it should also be coupled with thermal analysis.

The procedure of implementing the above multi-physics model starts with the establishment of transformer's geometry using the structural mechanics module in COMSOL (Figure 4.1b). A circuit design module is adopted to configure an appropriate circuitry representation of the transformer. After the transformer's geometry and circuitry models are constructed, they are then fed into the

electrical-magnetic solver (AC/DC module) to obtain the electromagnetic field distribution of the transformer. Losses induced by this electromagnetic field are then computed and exported to the heat transfer module as heat sources.

Fluid (oil) flow field is also studied by using CFD module and modelled results are subsequently exported to heat transfer module. Simultaneously analysing the electromagnetic field and fluid flow distribution can facilitate thermal analysis attaining satisfied accuracy regarding temperature distribution in a transformer. Finally, the moisture diffusion is modelled by diluted species transport module (refer to Section 4.1.3.4) and the moisture distribution in the transformer can be obtained.

4.1.4 Experiments and Results Analysis

4.1.4.1 Prototype Transformer Configuration

Extensive moisture diffusion experiments were performed on a prototype transformer (5 kVA, 240/2000 V), which is a scaled down version of a real power transformer. The prototype transformer is a shell type unit. The solid insulation consists of Kraft paper, mouldable pressboard and spacers all round. Three k-type thermocouples at axial heights of 190, 270 and 370 mm respectively have been fitted internally to measure the temperature variation within the transformer. Figure 4.2 shows the geometry of the transformer. Table 4.1 summarizes the thickness and outer perimeter of the cellulose for different components in the transformer.

Since the losses due to the current flowing through the transformer's winding may not generate the required temperature, a heater was installed on the bottom of the prototype transformer. A monitoring and control system was also developed, which consists Programmable Logic Controller (PLC), peripheral circuits, sensors, and software for achieving control and recording voltage, current, load and temperature of the transformer. With these setups, the temperature of the prototype transformer can be controlled to simulate different thermal conditions.

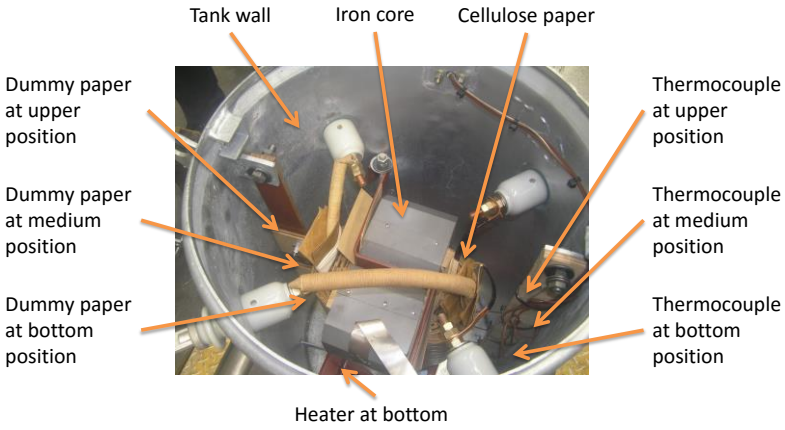


Figure 4.2 Geometry of the prototype transformer

Table 4.1 Geometry information of the prototype transformer

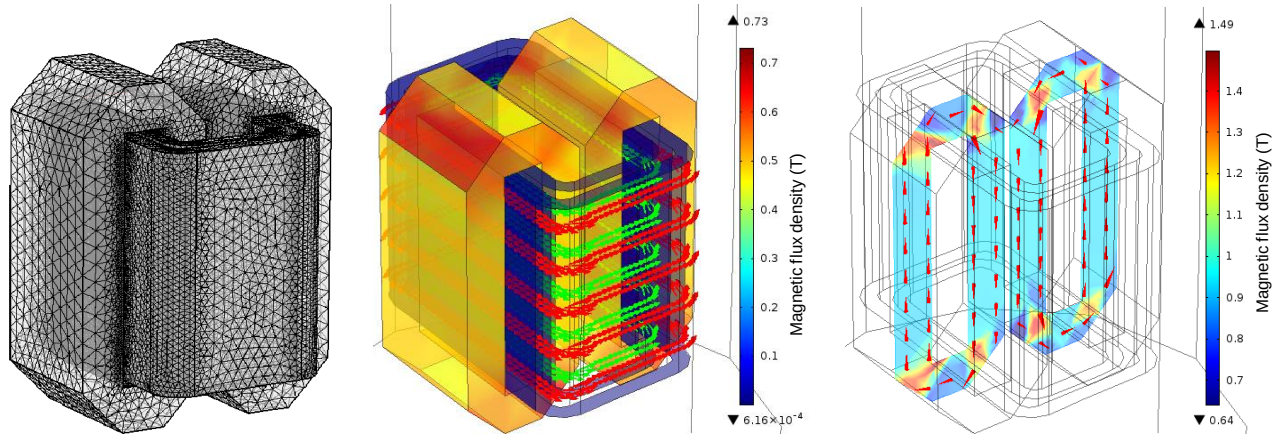
Thickness of core - LV insulation (mm)	**Thickness of HV-LV insulation (mm)**	**Thickness of HV-tank insulation (mm)**
1.75	9.7	2.75
Outer perimeter of core - LV insulation (mm)	**Outer perimeter of HV-LV insulation (mm)**	**Outer perimeter of HV-LV insulation (mm)**
396	527	624

4.1.4.2 Electromagnetic Field Simulation

The electromagnetic field distribution of the prototype transformer's iron core and windings is shown in Figure 4.3. The primary and secondary windings were made of thin wires and had multiple turns. It is assumed the wire diameter is less than the skin effect depth. Therefore, these windings are modelled as multi-turn coils. In electromagnetic field analysis, an equivalent circuit is created which consists a primary winding, connecting with a resistor (R_{wdg-p}) and an AC voltage source (V_{wdg-p}) and the secondary winding connecting with the load bank (R_s =8.8 Ω).

In the multi-physics modelling, self-adaptive meshing is employed. The fineness of the meshing is automatically tuned, considering transformer's geometry, the convergence performance and computation load. Figure 4.3a shows the meshing of the prototype transformer. The numbers of elements for each type of mesh are: tetrahedral - 258.8×10^3, triangular - 193×10^3, edge - 5.7×10^3 and vertex - 128. Figure 4.3b plots the magnetic flux density distribution and the currents in primary winding (red arrows) and secondary winding (green arrows) at the peak voltage. From Figure 4.3b it can be seen that the magnetic flux density of iron core surface is maximized at 0.7 T when the transformer is operated under the rated load. The induced currents are uniformly distributed in the conductors of primary and secondary windings. Figure 4.3c shows the slice plot of the magnetic

flux density of iron core. As can be seen from Figure 4.3c, most magnetic flux resides in the iron core and the leakage magnetic flux in the surroundings of the windings is minimal. The maximum magnetic flux density is located at the sharp joints of the iron core with 1.5 T. Therefore, the internal loss caused by the eddy currents and joule loss may be higher for these areas.



(a) (b) (c)

Figure 4.3 Meshing and magnetic flux density of transformer at peak voltage (a) Meshing for prototype transformer (b) Magnetic flux density and currents in the windings, red arrows and green arrows denote current in HV/LV windings (c) Magnetic flux density of cross-section of transformer core

Figure 4.4a shows the current density of transformer windings at rated load. The current density in the conductors of HV and LV windings are 1.2 A/mm^2 and 1.5 A/mm^2 respectively. Figure 4.4b shows the loss distribution in the transformer. Three heat sources, including loss dissipation in transformer core, Joule losses in the windings (both primary and secondary) and eddy current loss are calculated from electromagnetic analysis and considered as input for thermal modelling. The stray loss of the transformer is ignored.

The loss dissipation in transformer core is calculated by the magnetic flux distribution in the core area. The Joule losses in the transformer windings are computed based on the winding current density. The magnetic losses and Joule losses of different components (i.e., core, primary/secondary winding) at different time instances are computed as (4.15) and the average losses during one cycle are computed as (4.16).

$$Q_s(t) = \iiint_{\Omega} q(x, y, z, t)\,dxdydz \tag{4.15}$$

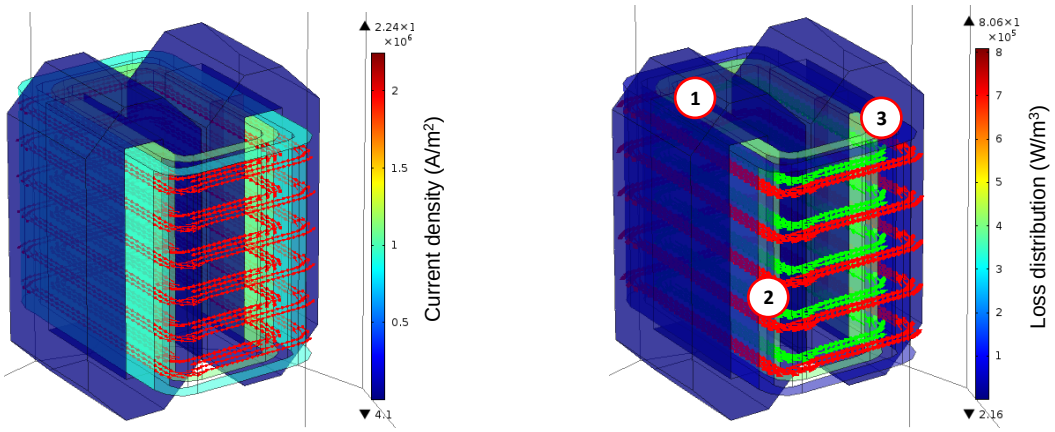
$$Q_{ST} = \frac{1}{T}\int_0^T Q_s(t)\,dt \tag{4.16}$$

where $Q_s(t)$ denotes the losses of different components in transformer at time instance t. Ω denotes the domain of each component. $q(x,y,z,t)$ denotes the loss density of infinitesimal volume in each component. Q_{ST} denotes the average losses during one cycle and $T = 0.02s$ denotes the period of each cycle.

The eddy current loss of the winding is provided by the transformer manufacturer. Table 4.2 summarizes the locations and heat sources of the prototype transformer.

Table 4.2 Locations and heat sources of prototype transformer

Core (W)	**HV - Joule (W)**	**HV - eddy (W)**	**LV - Joule (W)**	**LV - eddy (W)**
26.9	61	38.1	423	34
Location 1	Location 2	Location 2	Location 3	Location 3



(a) (b)

Figure 4.4 Current density and loss distribution in the transformer at rated load (a) Current density (b) Loss distribution, denotations 1, 2 and 3 indicate the iron core, primary winding and secondary winding area

4.1.4.3 Thermal And Fluid Flow Field Simulation

Thermal distribution of the prototype transformer is computed by coupling the conduction-convection physics and the non-isothermal fluid flow. To reduce the computation burden, the original 3D transformer's geometry is simplified as 2D geometry and the conductor arrangement of transformer winding is shown in Figure 4.5.

In Figure 4.5, three radial sections are formed, which include the iron core (left part), HV and LV windings (rectangles in central area) and the boundary of tank (right boundary). In the central area, conductors with different turns in windings form a radial layered structure and the axially multiple turns of conductors are grouped as ten blocks (coils) with oil ducts in between. The oil ducts are also located in the LV-HV gap. The detailed geometry of each component in thermal modelling is provided in Table 4.3.

Table 4.3 Geometry configuration of thermal modelling

LV conductors (mm)	**HV conductors (mm)**	**Radial space for each conductor (mm)**	**Vertical space for each conductor (mm)**
1.75×15	1.5×15	2.5	5
Group number of conductors	**Iron core - LV (mm)**	**HV-tank (mm)**	**Oil duct between HV- LV (mm)**
10	5	70	9.7

For thermal and fluid flow modelling, temperature dependent oil properties [24] are used as inputs which are summarized in Table 4.4. The values of surface emissivity of transformer core, winding conductors and tank wall are determined by the material properties and temperature. For the prototype transformer, its core, winding conductor and tank are made of silicon iron, copper, and stainless steel, respectively. Their surface emissivity at 27℃ are 0.6, 0.78 and 0.85.

Table 4.4 Temperature dependent oil characteristics for thermal and fluid flow model

Temperature (K)	**Viscosity (m^2/s)**	**Density (kg/m^3)**	**Thermal conductivity (W/m ℃)**	**Specific heat (J/kg ℃)**
258	4.5×10^{-4}	9.02×10^{2}	0.134	1.7×10^{3}
268	1.8×10^{-4}	8.96×10^{2}	0.133	1.74×10^{3}
278	8.5×10^{-5}	8.9×10^{2}	0.132	1.785×10^{3}
288	4.5×10^{-5}	8.84×10^{2}	0.131	1.825×10^{3}
298	2.7×10^{-5}	8.79×10^{2}	0.131	1.87×10^{3}
308	1.7×10^{-5}	8.73×10^{2}	0.13	1.91×10^{3}
318	1.15×10^{-5}	8.67×10^{2}	0.129	1.95×10^{3}
328	8.2×10^{-6}	8.61×10^{2}	0.128	1.995×10^{3}
338	6.1×10^{-6}	8.55×10^{2}	0.128	2.04×10^{3}
348	4.7×10^{-6}	8.49×10^{2}	0.127	2.08×10^{3}
358	3.8×10^{-6}	8.44×10^{2}	0.126	2.12×10^{3}
373	3×10^{-6}	8.35×10^{2}	0.125	2.18×10^{3}

The windings of the prototype transformer are immersed in the mineral oil with constant temperature of 55 ℃. The boundary conditions for fluid flow simulation are set as follows: At the bottom of transformer tank, the initial fluid velocity is set to zero $v|_{z=0}=0(mm/s)$ since no pump is installed in the transformer. At the top of the transformer, the internal pressure $P=1.01\times10^5(Pa)$ is kept as constant and the initial radial oil flow velocity is zero. Along the tank wall and winding surfaces, there is no slip between the fluid and solid component. Conjugate heat transfer with laminar flow is used to model the temperature distribution in the transformer when the cellulose is treated as being immersed in a slow-moving fluid flow of oil.

Figure 4.5 depicts the steady state temperature distribution of transformer windings at rated load. From Figure 4.5 it can be seen that there is not significant difference in the temperature at different sections of the windings. Most area has quite similar temperatures around 55 ℃. This is due to the fact that not enough losses are generated to elevate windings' temperature. The maximum temperature (hot spot) is 56 ℃, which is located at the top inner (LV) winding.

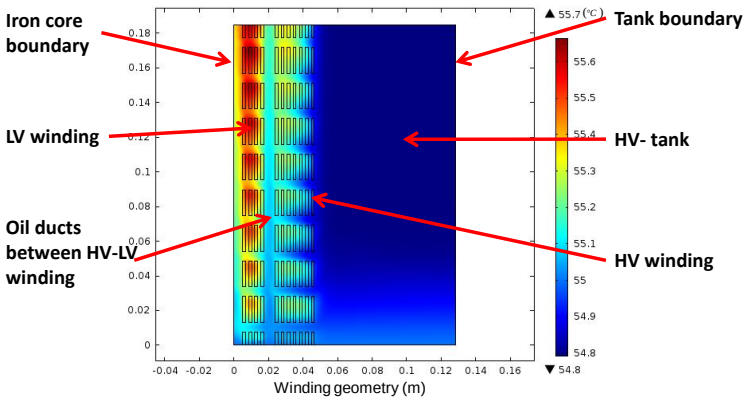


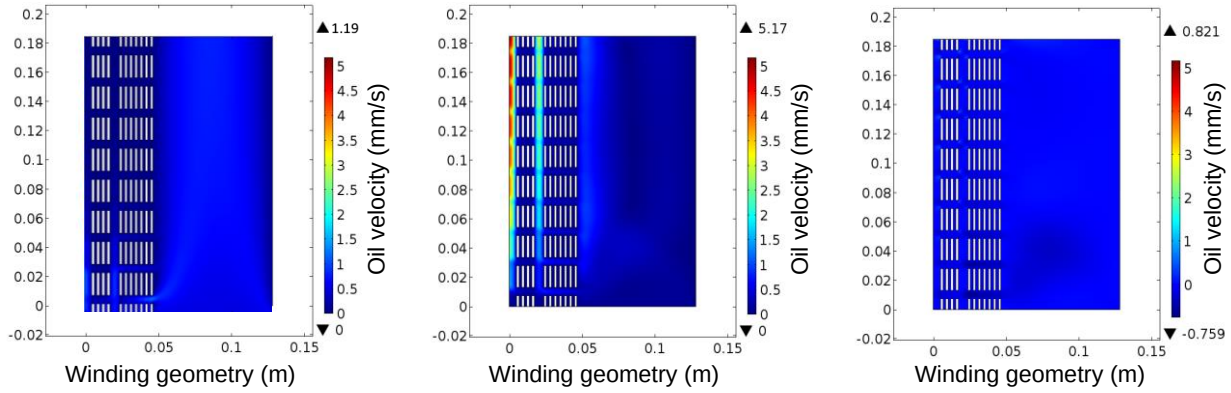
Figure 4.5 Temperature distribution in the transformer cross section

From Figure 4.5 it can be observed for LV winding, the temperature distribution of conductors in the middle layers is quite uniform. Therefore, the radiation heat exchange between the surfaces of these conductors is insignificant due to the slight temperature difference in these areas. However, due to the convection cooling of oil flow, the temperatures of surface between iron core and LV insulation and the surface between the LV-HV winding insulation are lower than that of the middle layer of LV windings.

For HV windings, the temperature of the surface between LV-HV winding insulation is higher than that of layers facing the tank. This is due to the fact that more heat is generated in this area

caused by a large number of turns of conductors. Another possible reason could be that heat is transferred by the oil from the internal area to the tank surface with strong convection cooling capability. This will be beneficial for heat ejection into the ambience. Similarly, the top of HV windings are hotter than that of the bottom since the effects of convective heat transfer are more dominated in the bottom area.

Figure 4.6 presents the oil flow field of the prototype transformer when it is not energized (Figure 4.6a) and when it is operated at rated load (Figure 4.6b). When the transformer is not energized, the temperature distribution and the fluid flow are only influenced by the heater installed at the centre bottom of the transformer's tank.



(a) (b) (c)

Figure 4.6 Fluid flow field in the transformer cross section area (a) Transformer was not energized but the heater (installed at the bottom of tank) is in operation (b) Transformer was operated at rated load (c) Radial component of the oil flow field

From Figure 4.6a it can be obtained that the oil velocity is higher close to the centre of the tank's bottom when the transformer is not energized. This is because high temperature (the heater installed at the bottom of tank) causes reduction in oil viscosity. Along the space between HV windings and the tank, oil velocity increases from bottom to top and reaches maximum of 2 mm/s. The possible reason for the high velocity in this area is that very few solid barriers (i.e., winding coils, cellulose paper and spacers) are placed and this is beneficial for oil movement. In addition, the buoyancy effects of the oil can also contribute to some extent. In contrast, the oil velocity is more uniform in the vertical oil ducts between LV-HV windings, which indicates small temperature difference between these two windings.

From Figure 4.6b it can be seen that the oil velocity in the vertical oil ducts between LV-HV windings increases to 5 mm/s when the transformer is energized. The oil attains maximum velocity

in the top oil ducts. This is because the temperature of top windings is relatively higher than that of the bottom when the transformer is energized (shown in Figure 4.6b). Similar to Figure 4.6a, in the space between HV windings and the tank, the oil velocity also increases from bottom to the top. Figure 4.6c shows the radial oil velocity (under rated load) in the oil ducts. From Figure 4.6c it can be seen that there exists a radial oil velocity between HV and LV windings. However, it only accounts for a rather small proportion (less than 1 mm/s) of the total oil velocity. The maximum radial oil velocity is at the corner between LV windings facing the iron core and oil ducts as well as inter-winding areas.

4.1.4.4 Moisture Distribution

After obtaining transformer's thermal distribution and the fluid flow field in Section 4.1.4.3, 3D moisture distribution (coupled with modelled thermal and fluid flow field) in the transformer is calculated as follows:

(a) Determining diffusion coefficients: The temperature distribution of oil-cellulose system at each time instance *t* from previous thermal analysis is imported into chemical species transport module to determine the diffusion coefficients for each infinitesimal volume of cellulose. Diffusion coefficients calculated by Foss's equation (4.17) are assigned to the cellulose paper in moisture diffusion modelling [10].

(b) Configuring boundary condition: A moisture-in-oil sensor [25] was installed in the prototype transformer. The sensor's tip was close to the cellulose paper of winding. By using this setup, temperature, water activity (defined as the ratio of the partial pressure of water in the material and the saturated vapor pressure of pure water at the same temperature) and moisture in oil can be continuously monitored. As most moisture in transformer resides in cellulose insulation, for flowing oil, the distribution of water content in oil is considered as relatively uniform. The moisture concentration at the interface of cellulose paper (contacting with oil) can be obtained by using the readings from the sensor and Fessler's equation [26-27] as shown in (4.17). It is validated only when the moisture in oil is below moisture saturation level of mineral oil. It showed that at the start of moisture diffusion experiments, the moisture dissolved in oil was 27 ppm and the moisture concentration in cellulose was 1.5%. The above measured and derived moisture concentrations in oil and cellulose are used as boundary conditions while the moisture concentrations of iron core and tank wall are set to zero.

$$C_0 = 2.173\times10^{-7}\times P_v^{0.6685}\times e^{\frac{4275.6}{T}} \tag{4.17}$$

where C_0 is the moisture concentration of cellulose surface, P_v denotes the vapour pressure of moisture and it can be derived from moisture in oil sensor [27], T denotes the temperature.

(c) Coupling with fluid flow field: The fluid flow field of mineral oil at each time instance *t* is imported into chemical species transport module. Then the moisture concentration of each infinitesimal volume of oil/cellulose can be calculated as (4.12).

(d) Calculating average moisture concentration: Based on the moisture distribution in (c), the average moisture concentration of the bulk cellulose is calculated as (4.18).

$$W_{average} = \frac{1}{V_c} \iiint_{\Omega_c} c(x, y, z) dV \tag{4.18}$$

where V_c denotes the total bulk volume of the cellulose, Ω_c denotes cellulose domain for integration and $c(x, y, z)$ denotes the moisture concentration of infinitesimal volume in cellulose.

The above mentioned approaches are implemented to model moisture dynamics in the prototype transformer. Usually, moisture diffusion in cellulose is a rather slow process. Therefore, a sinusoidal temperature profile is applied to the prototype transformer to accelerate the moisture transfer between cellulose and oil. The moisture diffusion experiments were conducted following three steps:

(1) The transformer was kept at a constant temperature (55 ℃) by using the installed heater for one week to attain moisture equilibrium.

(2) The transformer was subjected to a sinusoidal temperature profile (Figure 4.7) for one week. As oil viscosity is higher at low temperature and its mobility may be limited inside the transformer, the highest temperature for moisture diffusion is set to 80 ℃ and the lowest temperature is set to 30 ℃. The period of one cycle sinusoidal temperature is 24 hours. This is also aimed at simulating the load and temperature variation of a real transformer.

(3) The transformer was kept at a constant temperature (55 ℃) for another one week to attain moisture equilibrium.

Figure 4.7 presents one complete cycle of moisture measurement under a sinusoidal temperature. In Figure 4.7, the temperature (in green), water activity (in blue) and moisture in oil (in red) were directly obtained from the above moisture-in-oil sensor. When the temperature of transformer oil increases, the moisture saturation level of mineral oil will increase significantly which results in a decrease in the water activity. Therefore, water vapor pressure (defined as the partial pressure exerted by water vapor) will also decrease and moisture tends to move into oil from cellulose

insulation. This will result in an increase in water content of transformer oil. By contrast, when temperature of transformer oil decreases, moisture will return to cellulose insulation from transformer oil and the water content of cellulose will also increase.

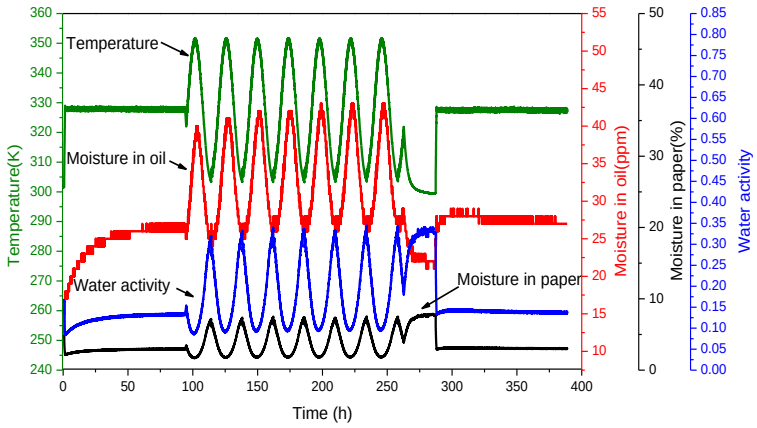
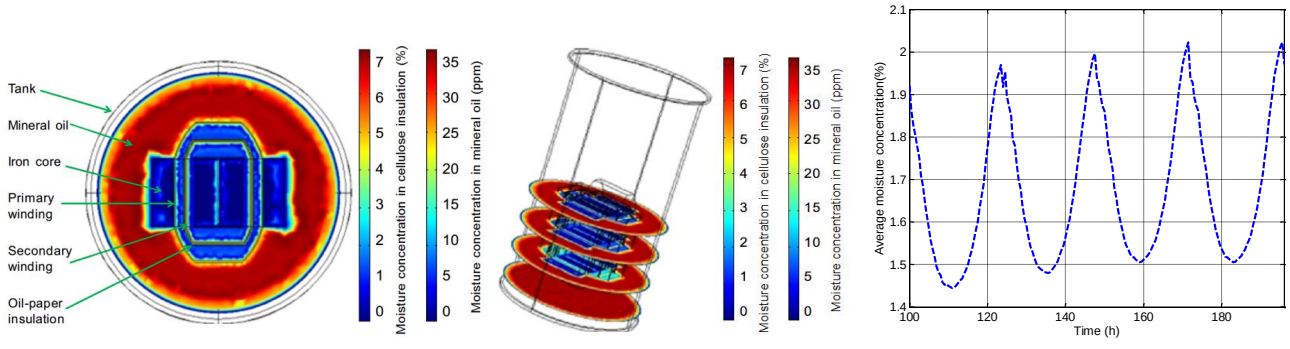


Figure 4.7 Sinusoidal variation of temperature (green), water activity in oil (blue), moisture in oil (red) and water content in cellulose paper (black) of prototype transformer

The modelled 3D moisture distribution of the prototype transformer is shown in Figure 4.8 (at time instance t = 172 hours). From Figure 4.8a it can be obtained that the moisture concentration is not uniformly distributed inside the cellulose. Large proportional of the cellulose has a relative low moisture (less than 1.5%) while the moisture of the oil-cellulose interfaces and their surroundings may reach up to 4% - 7%. From Figure 4.8b it can be seen that even the moisture along the transformer's height direction is not uniformly distributed either. This is due to the non-uniform thermal distribution and fluid flow field inside the transformer. The calculated average moisture concentration of cellulose at different time is shown in Figure 4.8c.



(a) (b) (c)

Figure 4.8 Moisture distribution of the prototype transformer (a) Moisture distribution of cross section area of transformer (at time instance t = 172 hours) (b) 3D moisture distribution at time instance t = 172 hours (c) Moisture variation of cellulose under a sinusoidal temperature

As can be seen from Figure 4.8c average moisture concentration varies in a sinusoidal shape between 1.45%-2.3%. The overall moisture in the cellulose of the prototype transformer is 1.7 % by taking the average of data in Figure 4.8c. To verify the modelling results, Karl Fischer Titration (KFT) method [28] was employed to measure water content in paper samples which were located close to the winding inside the prototype transformer. Before sample collection, the transformer was maintained at 55℃ for seven days to facilitate moisture equilibrium between oil and cellulose. The paper samples were collected from three locations corresponding to upper, medium and bottom height of winding (Figure 4.2). This takes the non-uniform moisture distribution into consideration. It demonstrates the modelled moisture has a good agreement with the testing results from KFT, which indicates 1.5% moisture in cellulose specimen collected from the prototype transformer.

4.1.4.5 Experimental Validation

To further verify the proposed multi-physics model for transformer moisture estimation, a series of experiments have been carried out on the prototype transformer at different stages.

After commissioning the prototype transformer, it was subjected to an equivalent 35-day accelerated thermal and electrical ageing, in which the transformer was kept at 110℃ with 30A load current. The Degree of Polymerization (DP) of cellulose was reduced to 364 at the final stage of the accelerated ageing experiment.

After the accelerated ageing, moisture diffusion experiments were conducted. It consisted of five stages and each stage took 21 days as aforementioned in Section 4.1.4.4 (Figure 4.7). At the end of each stage, dummy paper samples close to the transformer winding were collected and KFT method was used for measuring moisture. Each paper sample was divided into two adjacent sessions and KFT measurements were conducted consecutively on these two sessions. The final moisture concentration of the paper sample was the average of the two measurements. The accuracy of KFT equipment is 0.01% and the standard deviation is 0.03%. Table 4.5 summarizes the average moisture contents estimated by the multi-physics model and KFT method and degree of polymerization at different ageing stages.

From Table 4.5 it can be obtained that the moisture content calculated by the multi-physics modelling is quite close to those measured by the KFT method with some small variation for 21 and 28 days ageing samples. The difference could be due to the inconsistent sample collection practices and change in the collection conditions.

Table 4.5 Comparison of average moisture and degree of polymerization of prototype transformer

Time (days)	Modelling Results (%wt)	KFT (%wt)	Degree of Polymerization
0	1.55	1.52	364
7	1.63	1.60	547
14	1.65	1.37	266
21	1.66	1.73	374
28	1.66	1.25	390
35	1.70	1.55	387

4.1.5 Conclusion

This paper proposed a multi-physics modelling approach to reveal the temperature dependent moisture dynamics of transformer's oil-cellulose system. Four dominated physics were considered simultaneously to model moisture distribution of transformer under temperature transients. Losses of different components of a transformer were calculated and used as heat sources in thermal modelling. Temperature dependent oil characteristics and diffusion coefficients were subsequently determined to solve fluid flow field of mineral oil and moisture concentration. The multi-physics modelling approach was implemented on a prototype transformer. The results from extensive accelerated ageing and moisture diffusion experiments were presented to verify this approach. It is expected that the proposed multi-physics modelling approach can provide a better understanding on the complicated heat and moisture transfer process in power transformer, and will help to estimate moisture in transformer more accurately.

Acknowledgments

We highly appreciate the supports provided by Australian Research Council, Powerlink Queensland, Energex, Ergon Energy, and TransGrid.

References

[1] L. E. Lundgaard, W. Hansen, D. Linhjell and T. J. Painter, "Aging of oil-impregnated paper in power transformers," *IEEE Trans. Power Delivery*, vol.19, Issue 1, pp. 230-239, 2004.

[2] T. Toudja, H. Moulai, A. Nacer, A. Beldjilali, I. Khelfane and A. Debche, "Moisture and electrical discharges effect on naphthenic mineral oil properties," *IET Sci. Meas. Technol.*, vol.8, Issue 6, pp. 588-594, 2014.

[3] J. Gielniak, A. Graczkowski, H. Moranda, P. Przybylek, K. Walczak and Z. Nadolny, *et al.*, "Moisture in cellulose insulation of power transformers - statistics," *IEEE Trans. Dielectr. Electr. Insul.*, vol.20, Issue 3, pp. 982-987, 2013.

[4] Y. Du, M. Zahn, B. C. Lesieutre, A. V. Mamishev and S. R. Lindgren, "Moisture equilibrium in transformer paper-oil systems," *IEEE Electr. Insul. Mag.*, vol.15, Issue 1, pp. 11-20, 1999.

[5] Pahlavanpour, M. Martins and Eklund, "Study of moisture equilibrium in oil-paper system with temperature variation," in *Proceedings of International Conference on Properties and Applications of Dielectric Materials, June 1-5*, 2003,Nagoya, Japan, pp. 1124-1129.

[6] T. V. Oommen, "Moisture equilibrium charts for transformer insulation drying practice," *IEEE Transactions on Power Apparatus and Systems*, vol.103, Issue 10, pp. 3062-3067, 1984.

[7] D. F. Garcia, B. Garcia and J. Burgos, "A review of moisture diffusion coefficients in transformer solid insulation-part 1: Coefficients for paper and pressboard," *IEEE Electr. Insul. Mag.*, vol.29, Issue 1, pp. 46-54, 2013.

[8] D. F. Garcia, R. Villarroel, B. Garcia and J. Burgos, "A review of moisture diffusion coefficients in transformer solid insulation - Part 2: Experimental validation of the coefficients," *IEEE Electr. Insul. Mag.*, vol.29, Issue 2, pp. 40-49, 2013.

[9] P. F. Ast, "Movement of moisture through A50P281 Kraft paper (dry and oil impregnated)," General Electric, 1966.

[10] S. D. Foss and L. Savio, "Mathematical and experimental-analysis of the field drying of power transformer insulation," *IEEE Trans. Power Delivery*, vol.8, Issue 4, pp. 1820-1828, 1993.

[11] A. F. Howe, "The diffusion of moisture through power transformer insulation," PhD dissertation, Dept. Electr. Electron. Eng., Univ. Nottingham, Nottingham, UK, 1975.

[12] A. S. Asem and A. F. Howe, "Drying of power-transformer insulation," *IEE Proceedings - Generation, Transmission and Distribution*, vol.129, Issue 5, pp. 228-232, 1982.

[13] Y. Du, "Measurements and modeling of moisture diffusion processes in transformer insulation using interdigital dielectrometry sensors," PhD dissertation, Dept. Electr. Eng. Comp. Sci, Mass. Inst. Tech., Cambridge, MA, 1999.

[14] D. F. Garcia, "A new proposed moisture diffusion coefficient for transformer paper," *Int. J. Heat Mass Transfer*, vol.56, Issue 1-2, pp. 469-474, 2013.

[15] R. Villarroel, D. F. Garcia, B. Garcia and J. C. Burgos, "Diffusion coefficient in transformer pressboard insulation part 1: non impregnated pressboard," *IEEE Trans. Dielectr. Electr. Insul.*, vol.21, Issue 1, pp. 360-368, 2014.

[16] R. Villarroel, D. F. Garcia, B. Garcia and J. C. Burgos, "Diffusion coefficient in transformer pressboard insulation part 2: mineral oil impregnated," *IEEE Trans. Dielectr. Electr. Insul.*, vol.21, Issue 1, pp. 394-402, 2014.

[17] J. Driesen, G. Deliege, R. Belmans and K. Hameyer, "Coupled thermo-magnetic simulation of a foil-winding transformer connected to a nonlinear load," *IEEE Trans. Magn.*, vol.36, Issue 4, pp. 1381-1385, 2000.

[18] D. Giancoli, "Electric Currents and Resistance," in Physics for Scientists and Engineers with Modern Physics, 4th ed.Ed. New Jersey: Jocelyn Phillips, 2009, pp. 658.

[19] L. W. Pierce, "Transformer design and application considerations for nonsinusoidal load currents," *IEEE Trans. Ind. Appl.*, vol.32, Issue 3, pp. 633-645, 1996.

[20] B. S. Ram, "Loss and current distribution in foil windings of transformers," *IEE Proceedings - Generation, Transmission and Distribution*, vol.145, Issue 6, pp. 709-716, 1998.

[21] F. P. Incropera and D. P. Dewitt, *Fundamentals of Heat and Mass Transfer*, New York: Wiley, 2011.

[22] M. A. Tsili, E. I. Amoiralis, A. G. Kladas and A. T. Souflaris, "Power transformer thermal analysis by using an advanced coupled 3D heat transfer and fluid flow FEM model," *Int. J. Therm. Sci.*, vol.53, Issue 53, pp. 188-201, 2012.

[23] D. E. Rosner, *Transport processes in chemically reacting flow systems*, New York: Dover Publications, 2012.

[24] D. Susa, M. Lehtonen and H. Nordman, "Dynamic Thermal Modelling of Power Transformers," *IEEE Trans. Power Delivery*, vol.20, Issue 1, pp. 197-204, 2005.

[25] V. Oyj,"MMT330 Moisture and Temperature Transmitter Series for Oil" available at http://www.vaisala.com/en/industrialmeasurements/products/moistureinoil/Pages/MMT330.aspx.

[26] W. A. Fessler, T. O. Rouse, W. J. McNutt and O. R. Compton, "A refined mathematical model for prediction of bubble evolution in transformers," *IEEE Trans. Power Delivery*, vol.4, Issue 1, pp. 391-404, 1989.

[27] D. Martin, C. Perkasa and N. Lelekakis, "Measuring paper water content of transformers: a new approach using cellulose isotherms in nonequilibrium conditions," *IEEE Trans. Power Delivery*, vol.28, Issue 3, pp. 1433-1439, 2013.

[28] *Insulating Liquids—Oil-Impregnated Paper and Pressboard-Determination of Water by Automatic Coulometric Karl Fischer Titration*, IEC 60814 Ed. 2.0, 1997.

4.2 Particle Tracing Modelling on Moisture Dynamics of Oil-impregnated Transformer

Particle Tracing Modelling on Moisture Dynamics of Oil-impregnated Transformer

Yi Cui, Hui Ma, Tapan Saha, Chandima Ekanayake and Daniel Martin

The University of Queensland, Brisbane, Australia

Publication Journal: IET Science, Measurement & Technology

Submitted: September 4, 2015

Revision Resubmitted: December 4, 2015

Accepted: January 12, 2016

Author's Contributions:

Yi Cui: Primary model design, numeric simulation, experimental validation, results analysis and manuscript preparation.

Hui Ma: Supervision on the experimental measurement, results discussion and manuscript revision.

Tapan Saha: Supervision on the experimental measurement and manuscript revision.

Chandima Ekanayake: Results discussion.

Daniel Martin: Results discussion.

Abstract

An accurate estimation of moisture in oil-impregnated cellulose of a transformer is difficult due to the complex moisture dynamics inside the transformer, which is highly influenced by the temperature and dimension of the transformer. In this paper, a novel particle tracing method is proposed for estimating the moisture in transformers. Different from the conventional approaches using Fick's diffusion law, the proposed method formulates moisture diffusion from a microscopic view of water particles' motion. Especially, the transmission probability of water particles is employed to correlate the microscopic particles' motion with the macroscopic moisture distribution. The proposed particle tracing method is applied to model moisture dynamics in both pressboard specimens and a prototype transformer. Extensive moisture diffusion and ageing experiments are carried out to verify the proposed method.

Index Terms: Cellulose, insulation, moisture, particle, transformers

4.2.1 Introduction

Life expectancy of a power transformer is highly dependent on the condition of its cellulose materials [1]. However, cellulose can eventually degrade due to electrical, thermal and mechanical stresses. Moisture is one of the most harmful by-products of cellulose degradation. It can reduce the electrical and mechanical strength of the cellulose and also can further accelerate the cellulose degradation. Therefore, it is necessary to accurately estimate the moisture content in transformers and help utilities make informed decisions on their transformers' operation and maintenance schedules.

During transformer operations, load condition and ambient temperature is always changing. This leads to moisture dynamics in the transformer, including (1) moisture exchanges at the interface between the cellulose and oil due to the vapor pressure difference and temperature variation; and (2) moisture moves inside the cellulose caused by moisture gradient. Moisture dynamics is a complex heat transfer and mass diffusion process. However, due to the time constant of heat transfer is considerably smaller than that of diffusion, moisture dynamics is usually modelled as a diffusion process.

Many researchers studied moisture diffusion using Fick's diffusion law and determined diffusion

coefficients by conducting experiments on pressboard samples [2-9]. Instead of adopting Fick's diffusion law, this paper proposes a novel particle tracing method to estimate moisture of oil-impregnated cellulose in transformers. In particle tracing method, moisture diffusion is modelled from a microscopic perspective of water particles' movements. Water particle is a conceptual entity for describing collective and dynamic behaviour of bunch of water molecules. It contains a huge number of water molecules; however, its physical size is small enough that can be treated at microscopic level.

In particle tracing method, particles' motion is governed by certain motion principle and the trajectory of each particle is estimated and directly visualized. This method is exemplified in estimating moisture of both pressboard specimens and oil-cellulose system in a prototype transformer. The estimation results are validated by extensive moisture diffusion and accelerated ageing experiments.

This paper is organized as follows. Theories of moisture diffusion are briefly reviewed in Section 4.2.2. Section 4.2.3 details the formulation of the particle tracing method. Section 4.2.4 applies the proposed method to estimate moisture distribution in both pressboard specimens and prototype transformer. This paper is concluded in Section 4.2.5.

4.2.2 Moisture Diffusion in Oil-Impregnated Cellulose of Transformers

Moisture diffusion is a water molecule movement process governed by the random motion of individual molecule. In cellulose material, the bindings of water molecules and cellulose cell-walls are different and they are affected by local moisture concentration. Therefore, in cellulose water molecules are not equally free to migrate.

In the literature, moisture dynamics is commonly modelled by Fick's law as (4.19)[2].

$$\frac{\partial W(x,t)}{\partial t} = D \times \frac{\partial^2 W(x,t)}{\partial x^2} \tag{4.19}$$

where $W(x,t)$ is the moisture content of the cellulose at position x and the t-th time step. D denotes the diffusion coefficient [2] which can be expressed as (4.20).

$$D = D_0 \times e^{[k \times W + E_a \times (\frac{1}{T_0} - \frac{1}{T})]} \tag{4.20}$$

where $D_0 = 1.34 \times 10^{-13} m^2 / s$ and $k = 0.5$, E_a is the activation energy which equals 8074 *kJ/mol* [2],

T denotes the measured temperature in Kelvin and T_0 equals 298 K.

Extensive studies have been carried out to determine the diffusion coefficients of cellulose (paper and pressboard). Ast adopted permeation method for determining the diffusion coefficients of Kraft paper at different moisture levels and temperatures [3]. Guidi and Fullerton formulated a relationship between the diffusion coefficients and the local water content and temperature [4-5]. Howe investigated the diffusion coefficients for different types of cellulose materials [5]. Asem found the diffusion coefficients for oil-impregnated pressboard [6]. Foss verified the parameters of diffusion coefficients for both impregnated and non-impregnated Kraft paper by using other researchers' data [4]. By using dielectrometric sensors, Du studied the diffusion coefficients for non-impregnated pressboard [7]. Recently a series of diffusion coefficients considering pressboard thickness was reported by Garc á [8-9].

However, considerable difficulties still remain in applying the above results to transformers. This is because transformers may be constructed by different cellulose materials and may have a heterogeneous temperature distribution during their operation. Moreover, Fick's law in one dimension (1D) may not be able to fully describe the diffusion process in the cellulose of a transformer. In addition, for 1D and 2D finite element modelling (FEM) of moisture distribution, the performance is normally satisfied and without the problem of numerical instabilities. However, the 3D FEM may encounter some difficulties. The first difficulty is related to the Peclet number (refer to Section 4.2.3.1). The second difficulty is that 3D FEM models require high level of details about transformer design.

4.2.3 Particle Tracing Method for Modelling Moisture Diffusion

4.2.3.1 Formulation of Particle Tracing Method

Particle tracing method employs numerous small quantities (particles) and their collective dynamic behavior to model the mass transfer and diffusion process [10]. Compared to FEM method, it has several advantages. Firstly, it is not sensitive to the Peclet number in modelling mass transfer process. Peclet number is the ratio between the energy transferred by the fluid convection and that by the fluid conduction. If Peclet number is large (generally above 1000) the FEM may not be applicable to model mass transfer. The Peclet number may reach 10^7 in 3D modelling of moisture dynamics in transformers [11]. Moreover, particle tracing method does not require a pre-defined

mesh as FEM does. Instead, it utilizes discrete element method (DEM) [12] to calculate the position and velocity of each particle at each time instance. With the DEM approach, particle tracing method can avoid numerical instabilities associated with the continuum approach (e.g. FEM). Furthermore, by using particle tracing method, the trajectory of each particle can be visualized, which provides a complementary tool to interpret the modelled results and help to understand moisture diffusion in power transformers.

In particle tracing method, the transferable quantities (e.g. moisture) are discretised as certain amount of particles from the microscopic perspective. The position and velocity of each particle is computed iteratively by using the corresponding motion equation. The macroscopic distribution of the transferrable quantities can be approximated by the state space equation of the particles. To determine the particle's motion, Newtonian equation is adopted as (4.21).

$$\frac{d\left(m_p \boldsymbol{v}_p\right)}{dt} = \sum \boldsymbol{F}_p \tag{4.21}$$

where m_p denotes particle's mass, $\boldsymbol{v}_p$ is particle's velocity vector and $\sum \boldsymbol{F}_p$ is the collective effect of different driving forces imposed on the particle.

These driving forces include the gravity force of the particles, the drag force from the fluid flow of oil, the diffusion force from the moisture gradient, Brownian force and interaction forces [13]. In this paper both particle-particle interactions and particle-fluid flow (transformer oil) interactions are considered. Table 4.6 summarizes the driving forces imposed on the water particles.

Table 4.6 Domains of different driving forces imposed on the water particles

Domain	Mineral oil	Interface between cellulose and oil	Cellulose bulk
Gravity force	√	√	√
Drag force	√	√	
Brownian force	√	√	√
Diffusion force	√	√	√
Particle- fluid flow interactions	√	√	
Particle-particle interactions	√	√	√

The gravity force $\boldsymbol{F}_g$ is expressed as (4.22). For submicron or even smaller size particles, the drag force from fluid flow and other external forces dominate while gravity has little effect on particle's trajectories.

$$\boldsymbol{F}_g = \frac{m_p \boldsymbol{g}\left(\rho_p - \rho\right)}{\rho_p} \tag{4.22}$$

where $m_p = 3.75\times10^{-18} kg$ denotes the mass of particles (moisture), $\rho_p = 10^3 kg/m^3$ denotes the density of particles and $\rho = 861 kg/m^3$ denotes the density of the surrounding fluid (transformer oil), $\boldsymbol{g} = 6.67\times10^{-11} m^3/\left(kg\cdot s^2\right)$ denotes the gravity vector.

The drag force $\boldsymbol{F}_D$ due to the fluid flow is expressed as (4.23).

$$\boldsymbol{F}_D = \frac{1}{\tau_p} m_p \left(\boldsymbol{v}_F - \boldsymbol{v}_p\right) \tag{4.23}$$

where τ_p is the velocity response time constant, i.e. the time for a particle to respond to a change in the flow velocity of the carrier fluid [14]. $\boldsymbol{v}_F$ is the fluid flow velocity vector.

The velocity response time for spherical particles in a laminar flow is defined as (4.24).

$$\tau_p = \frac{\rho_p d_p^2}{18\mu_F} \tag{4.24}$$

where $\mu_F = 8.2\times10^{-6} m^2/s$ denotes the oil viscosity at 55℃. $d_p = 2\times10^{-7} m$ denotes the diameter of spherical moisture particles in both oil and cellulose at submicron scale and it is determined by considering the model accuracy and the computational cost.

In particle tracing method, Brownian force as (4.25) is applied to the particles to take account for diffusion of suspended water particles in a fluid. In Brownian forces, each particle is imposed with a force at each time step to describe particles' random movements.

$$\boldsymbol{F}_b = \zeta\sqrt{\frac{12\pi r_p k_B \mu_F T_F}{\Delta t}} \tag{4.25}$$

where Δt is the time step, $r_p = 10^{-7} m$ is the particle's radius, T_F is the fluid temperature, k_B is Boltzmann's constant and ζ is a random number, which follows normal distribution (zero mean, unit standard variation). In the calculation, ζ is chosen in all directions in space. At each time step, a unique value of ζ is created for each particle.

In particle tracing modelling, diffusion force is treated as a nominal force that depends on the gradient of moisture concentrations as (4.26).

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(D_t \frac{\partial P(x,t)}{\partial x} \right) - \mathbf{v} \cdot \frac{\partial P(x,t)}{\partial x} \tag{4.26}$$

where $P(x,t)$ denotes the particle distribution density at position x and time t, $D_t = k_B T / (6\pi\eta r_p)$, representing the interaction between spherical water particles and the cellulose medium. η denotes a scale factor which is determined by $P(x,t)$. $\boldsymbol{v}$ denotes the velocity vector of fluid flow of mineral oil.

The interactions between fluid flow and particles are considered in such a way that the particles are imposed by drag force from the fluid and in turn particles exert a momentum force on the fluid. It allows performing an accurate coupling between the water particles' movements and the velocity field of mineral oil. The mineral oil is treated as continuous medium rather than particles. Oil characteristics are temperature dependent [15]. They are used as the input in fluid flow modelling (Table 4.7). The total volume force exerted by all particles on the fluid at position $\boldsymbol{r}$ can be expressed as (4.27).

$$\boldsymbol{F}_V(\boldsymbol{r}) = -\sum_{i=1}^{n} \boldsymbol{F}_D \delta(\boldsymbol{r} - \boldsymbol{q}_i) \tag{4.27}$$

where δ is the delta function, $\boldsymbol{q}_i$ is i-th particle position vector, and the sum is taken over all particles.

Table 4.7 Temperature dependent oil characteristics for coupling fluid flow with particle tracing modelling

Temperature (K)	Viscosity (m^2/s)	Density (kg/m^3)	Thermal conductivity (W/m ℃)	Specific heat (J/kg ℃)
258	4.5×10^{-4}	9.02×10^{2}	0.134	1.7×10^{3}
268	1.8×10^{-4}	8.96×10^{2}	0.133	1.74×10^{3}
278	8.5×10^{-5}	8.9×10^{2}	0.132	1.785×10^{3}
288	4.5×10^{-5}	8.84×10^{2}	0.131	1.825×10^{3}
298	2.7×10^{-5}	8.79×10^{2}	0.131	1.87×10^{3}
308	1.7×10^{-5}	8.73×10^{2}	0.13	1.91×10^{3}
318	1.15×10^{-5}	8.67×10^{2}	0.129	1.95×10^{3}
328	8.2×10^{-6}	8.61×10^{2}	0.128	1.995×10^{3}
338	6.1×10^{-6}	8.55×10^{2}	0.128	2.04×10^{3}
348	4.7×10^{-6}	8.49×10^{2}	0.127	2.08×10^{3}
358	3.8×10^{-6}	8.44×10^{2}	0.126	2.12×10^{3}
373	3×10^{-6}	8.35×10^{2}	0.125	2.18×10^{3}

For particle-particle interactions, Lennard-Jones force and universal gravitation between two particles are taken into consideration [10]. The Lennard-Jones interaction is expressed as (4.28), representing the intermolecular potential function to determine the transport property of a particle.

$$U(r) = 4 \times \varepsilon \times \left[\left(\frac{r_c}{d} \right)^{12} - \left(\frac{r_c}{d} \right)^{6} \right] \tag{4.28}$$

where d denotes the distance between the particles, ε denotes the interaction strength, and r_c denotes the collision diameter.

Based on (4.28) the Lennard-Jones force of i-th particle $\boldsymbol{F}_{Li}$ can be computed as (4.29).

$$\boldsymbol{F}_{Li} = -\nabla U \tag{4.29}$$

4.2.3.2 Boundary Conditions

In particle tracing method, the solid material of cellulose insulation is modelled as continuous medium at the macroscopic level. Two boundary conditions are considered to simulate the effect of cellulose properties (e.g. relative magnitude of the paper fibers and the voids between the fibers) on the moisture diffusion as follows:

1. Stick boundary. It is assumed that all the water particles contacting the surfaces of the cellulose (boundaries) will be forced to stay steadily upon the boundaries without any movement. Consequently, the velocities of these particles are set to zero and their positions are kept unchanged (Figure 4.9a).

2. Diffuse-reflection boundary. When large amount of water particles move towards the boundaries, some proportion of water particles will be bounced back. If there is no energy loss in the reflection, such bounce back of water particles is described by specular reflection. If energy loss is considered, there will be a reduction in velocities of particles and Knudsen's cosine law is employed [10] (Figure 4.9b).

(a) (b)

Figure 4.9. Two boundary conditions. In the figure, dots denote particles and solid lines denote particles' trajectories. (a) Stick boundary, (b) Diffuse- reflection boundary.

During moisture diffusion both two boundary conditions can occur with equal possibility. Thus, a probability fraction of 50% is assigned to the water particles reaching the boundaries to determine which boundary condition governs the further motion of these particles. When modelling particle movement in different materials and conditions, this fraction can be altered. Moreover, for diffuse-reflection boundary, the probability that particles can bounce back from walls of cellulose can be also changed. The reflection can be done either by specular reflection or according to Knudsen's cosine law. Different from FEM method, the permittivities and the diffusion coefficients of the oil-cellulose interfaces need to be continuous (smoothed using linear interpolation) in particle tracing method.

4.2.3.3 Initializations

After determining the driving forces and boundary conditions, the system is initialized by releasing certain amount of water particles from the cellulose surface, which is in contact with oil. The initial velocities of these particles are set to zero.

The number of the released water particles can be a random integer and it is independent from the flow pattern and mass of the fluid. If this number is too small, the diversity of particles movement cannot be guaranteed. Then it is possible that eventually all the particles stick on the boundaries without moving further. Consequently this leads to an incorrect moisture distribution. On the other hand, if the number of the released water particles is too large, more computation resources are required. It is found the accuracy of modelled moisture almost reach saturation level if the number of particles exceeds certain threshold value. After careful comparisons of the calculated moisture distribution using different number of initial water particles, 3000 water particles are chosen to be released with uniform distribution on the cellulose surface (Figure 4.10). It should be notable that the optimal number of particles is not unique, which can vary depending on the geometry, boundary

conditions and physics to be modelled.

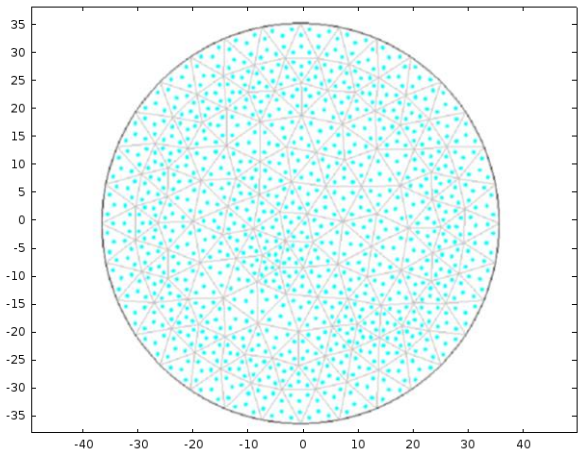


Figure 4.10. Initialization of particles' density and positions (3000 particles are depicted in blue dots). In the figure, x and y axis denote the geometry of the cellulose in mm. The grey circle denotes the outer edge of the pressboard (boundary) while the grey triangles denote the mesh for guaranteeing a uniform releasing of the particles at time instance t = 0 hour.

4.2.3.4 Post-Processing and Evaluations

After being released from the cellulose surface, all water particles are excited by the combined driving forces and move along the depth of the cellulose to its drier portion. The movement of water particles from wet area to dry area is ruled by both driving forces and boundary conditions. The transmission probability (can be regarded as an indicator of moisture distribution) of water particles inside the cellulose can be computed by counting the number of particles in the cellulose and dividing it by the total particles (distributed in both cellulose and oil). Finally, moisture distribution of the cellulose can be computed. The detailed procedures for deriving the moisture distribution from the transmission probability are discussed in the next section.

4.2.4 Results Analysis and Discussions

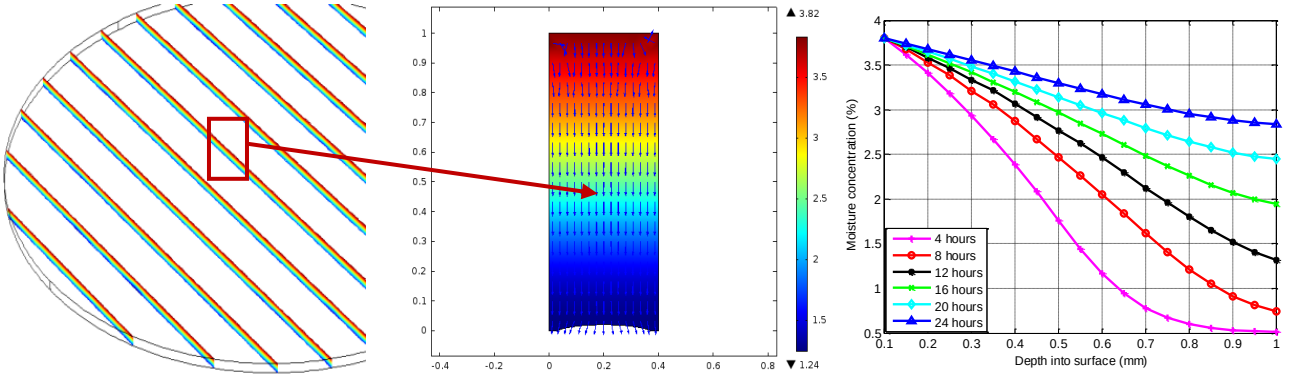
4.2.4.1 Application to Pressboard Specimens

Oil-impregnated pressboard specimens with 1 mm thickness are chosen in this study. The density

of the pressboard is 1.04 g/mm^3.The pressboard is brand new without being subjected to ageing before conducting this study. Necessary pre-treatment procedures were performed on the pressboard specimens including vacuum drying at 100℃ inside an oven for 24 hours. After the above treatments, pressboard samples have an initial moisture concentration less than 0.5%. Then, these specimens were immersed in mineral oil at a constant temperature of 25 ℃ for 72 hours to ensure complete impregnation.

The prepared oil-impregnated pressboards were put into a specially designed chamber with proper seals [16]. The surrounding of oil-impregnated pressboard is air with constant humidity maintained by saturated salt solution [17]. Temperature dependent water solubility in oil may affect the boundary conditions and in this study only one surface (top surface) of the pressboard was exposed to the moisture source with a relative humidity of 3.8%. The other surfaces of the pressboard surface were sealed. The temperature of moisture diffusion was kept as constant of 50 ℃.

Moisture distribution of the pressboard along its thickness is estimated by using both Fick's law and the particle tracing method for comparison. Figure 4.11a presents the 3D moisture distribution of the pressboard after 12 hours diffusion by solving Fick's law as in (4.19) and (4.20) [2]. To provide a better illustration, in Figure 4.11a the pressboard is sliced into a number of thin pieces so that the moisture distribution of the whole cross-section area can be visualized. Figure 4.11b presents the moisture gradient at the local cross-section area (blue arrows). As it can be seen from Figure 4.11b, there is a decreasing trend in the moisture gradient when moisture migrates into the depth of the pressboard.

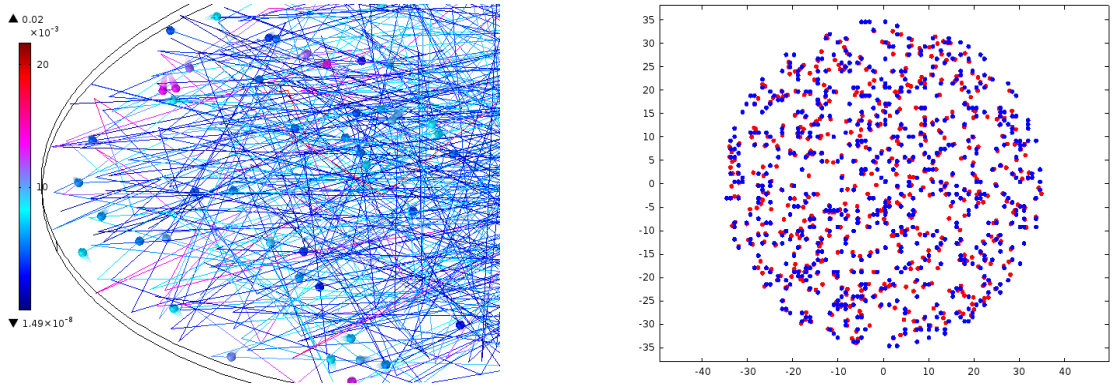


(a) (b) (c)

Figure 4.11 Moisture distribution of the pressboard after 12 hours diffusion at T=50℃, (a) Moisture distribution of the pressboard bulk (b) Moisture gradient in the pressboard (as indicated by the blue color arrows). Calculated using Fick's diffusion law, (c) Moisture distribution of the pressboard along its depth (thickness) at different time instances at T=50℃.Calculated using Fick's diffusion law.

Figure 4.11c shows the moisture distribution of the pressboard at different time instances. From Figure 4.11c it can be observed the moisture migration inside the cellulosic is a quite slow process. Even at a constant temperature of 50 ℃, the moisture of another surface of the pressboard only reaches 1.5% after 12 hours diffusion.

Moisture diffusion of the above pressboard is also modelled by using particle tracing method. The trajectories of water particles during the diffusion process are calculated as shown in Figure 4.12a (only 24 hours diffusion is presented). In Figure 4.12a, each water particles is defined as a sphere with certain mass and radius. The trajectory of each water particles is represented as solid lines and the moving direction can be recognized from the "tail of the comet".



(a) (b)

Figure 4.12 Modelled trajectories of water particles in the pressboard. (a) Trajectories of water particles in the pressboard during diffusion process (3000 particles) after 24 hours diffusion at T=50℃. (b) Poincare map of water particles of pressboard's top and bottom surfaces after 24 hours diffusion at T=50℃. In the figure, x and y axis denote the geometry of the pressboard in mm.

Particle tracing method describes hypothetical trajectories of water. It is worth mentioning that it is hard to exactly capture the particles' exact positions since moisture migration exhibits continuous behaviour, which is highly influenced by moisture and temperature gradients. In particle tracing method the predication on water particles' positions is largely dependent on the fineness of the time steps in the calculation. In Figure 4.12a, the velocity of each particle during diffusion is quantified by a color bar on the left (in mm/s). The maximum velocity of water particles is 15 mm/s. These particles have enough energy to move and they are more likely to leave the pressboard and enter the oil. By contrast, certain amounts of particles are bonded on the surface of the pressboard and they cannot move further.

To further investigate the moisture distribution in the above pressboard, Poincare map is adopted in the particle tracing method. Poincare map preserves many properties of periodic and quasi-periodic orbits of the particles and has a lower-dimensional state space. Thus, it is often used to provide an insight into particles' movements [10]. To construct a Poincare map, several observation planes (Poincare sections) are predefined and a reference plane with already known moisture concentration needs to be selected. They can be located at any position in the coordinate system. In this paper, the observation planes are paralleled with the pressboard surface and they are located at different depth from the pressboard surface. Therefore, the pressboard is "sliced" into several layers by the observation planes. The reference plane is chosen as the pressboard surface which is in contact with oil.

At a particular time instance, if a water particle passes one of the above observation planes, a dot will be recorded on this plane to track the movement of the particles. By collecting the number of particles penetrating this plane, the transmission probability of the particles can be computed. Subsequently, the moisture concentration at this observation plane (e.g. a particular depth of the pressboard) can be calculated as (4.30).

$$W_i(x,t) = \frac{W_{ref}(t) \times P_t(\alpha)}{1 - P_t(\alpha)} \tag{4.30}$$

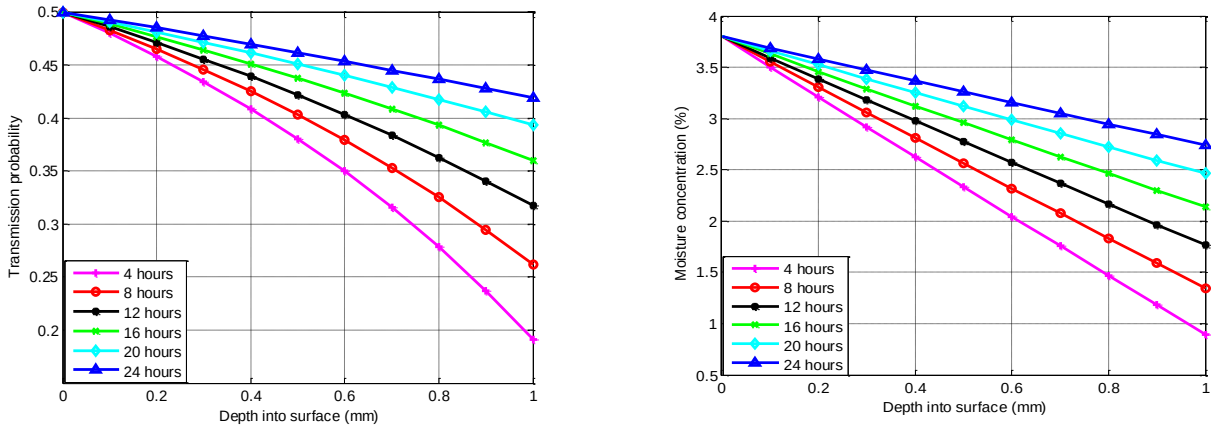
where $W_i(x,t)$ denotes the moisture concentration at the *i-th* layer of the pressboard with the distance of x to the reference plane, $P_t(\alpha)$ denotes the transmission probability (aforementioned in Section 4.2.3.4) and $W_{ref}(t)$ denotes the moisture concentration of the reference plane. In the above calculation, if the layer number *i* is large enough (the pressboard is sliced as infinite thin), the overall moisture distribution of the whole bulk volume of the pressboard can be obtained.

The Poincare map of water particles during the diffusion (after 24 hours) is shown in Figure 4.12b. In Figure 4.12b an observation plane and a reference plane are presented. The observation plane is located at the bottom of the pressboard (this side is in touch with the electrode, drier portion). The reference plane is with horizontal axis (denoting vertical Z direction) of zero. The particle traces penetrating the reference plane are presented as blue dots, while the red dots denote the water particles that penetrate the observation plane. From Figure 4.12b it can be observed that all the water particles diffuse within the pre-defined boundaries of the pressboard specimen. After 24 hours diffusion, most of the water particles have penetrated the pressboard and reached the drier part of the pressboard.

Figure 4.13a presents the transmission probability of the water particles along the depth of the

pressboard. Here, the pressboard is sliced into ten layers and the transmission probability for each layer is computed by counting the number of particles penetrating each layer (observation planes) and dividing it by the number of total released particles. Figure 4.13b shows the calculated moisture distribution of the pressboard at different diffusion time by using the particle tracing method.

Table 4.8 compares the calculated average moisture concentration of the pressboard based on Fick's law and particle tracing method. Since moisture diffusion coefficients have a significant influence on the moisture distribution, a comparison of the moisture distribution for oil-impregnated pressboard by using Foss, Guidi [2] and García's diffusion coefficients [8] is provided in Table 4.8.



(a) (b)

Figure 4.13 Transmission probability and moisture distribution of the pressboard calculated using particle tracing method. (a) Transmission probabilities of water particles in the pressboard during diffusion process at T=50℃. (b) Moisture distribution of pressboard along its depth (thickness) direction derived by particle tracing method at T=50℃.

From Table 4.8 it can be seen García's coefficient produces the lowest moisture content in the pressboard. Guidi's diffusion coefficient produces a relatively lower moisture content compared to Foss's coefficient. The moisture content calculated by Foss's coefficient is the closest to that obtained by particle tracing method. Thus Foss's coefficient is adopted to solve Fick's equation in the remaining of this paper.

From Figure 4.11c, Figure 4.13b and Table 4.8 it can be concluded that with the increase in the diffusion time and moisture content of the pressboard, both Fick's law and particle tracing method give quite similar values of moisture concentration (after 12 hours diffusion). However, if the pressboard is not wet enough or the moisture is not sufficiently diffused, particle tracing method tends to overestimate moisture of the pressboard (e.g. moisture distribution at 4 hours). This may be

due to the boundary selection in particle tracing method under low moisture circumstance. In this situation, the possibility fraction between the stick boundary and reflection boundary may not be kept as constant of 50% as previously set in the paper (Section 4.2.3.2).

Table 4.8 Comparison of calculated average moisture (%wt) in pressboard

Time (hours)	Foss	Guidi	Garc á	Particle tracing
4	1.77	1.37	0.83	2.34
8	2.24	1.83	0.94	2.57
12	2.58	2.15	1.03	2.78
16	2.85	2.4	1.1	2.96
20	3.07	2.58	1.16	3.12
24	3.26	2.72	1.22	3.26

4.2.4.2 Application to a Prototype Transformer

Particle tracing method is also applied to estimate the moisture content in a prototype transformer (5 kVA, 240/2000 V). The transformer was subjected to accelerated ageing and moisture diffusion experiments. The prototype transformer was designed and manufactured by ABB and its ratio between the paper and oil in the transformer was maintained at a level similar to a real transformer (represented by a *X*-*Y* model [18] with ratios $X = 38.32\%, Y = 16.45\%$). A heater was installed on the bottom of the transformer to control the temperature for ageing and moisture diffusion experiments. Table 4.9 summarizes prototype transformer's geometry. The solid cellulose materials consist of Kraft paper, mouldable pressboard and spacers.

Table 4.9 Geometry information of the model transformer

LV conductors (mm)	HV conductors (mm)	LV windings	HV windings
1.6×7.1	1.4×1.4	22 turns/layer, 4 layers	110 turns/layer, 7 layers
Thickness of layer insulation(mm)	**Thickness of core - LV insulation (mm)**	**Thickness of HV-LV insulation (mm)**	**Thickness of HV-tank insulation (mm)**
0.25	1.75	9.7	2.75

Due to the complex geometry of the transformer, multi-physics modelling is adopted to model 3D moisture dynamics in the transformer. The model integrates the effects of electromagnetic, thermal,

fluid flow and moisture migration physics on the moisture diffusion. Especially it considers the coupling and interactions of these physical phenomena [11]. The details of multi-physics modelling for moisture estimation will be provided in another paper.

After the commission of the prototype transformer, it was subjected to both electrical and thermal loading to attain a certain degree of ageing of its cellulose and oil (equivalent to 34 years of life consumption based on the degree of polymerization (DP) measurement of paper samples collected from this transformer). Electrical loading was provided by using the load bank with maximum power capacity of 6 kW. The prototype transformer was kept at 110 ℃ (using the abovementioned heater) with 30A load current for a time period equivalent of 35 days (the transformer was kept at 50 ℃ during night and weekend).

After 35 days accelerated ageing on the prototype transformer, experiments were arranged to study the moisture diffusion in the transformer. A sinusoidal temperature profile was imposed on the transformer by using the heater to simulate the operating conditions of a field transformer. A capacitive moisture-in-oil sensor (Vaisala MMT 330) [19] was installed in the transformer. The sensor was inserted into the transformer through a valve on the lid and its tip is close to the cellulose of the winding. The sensor is used to continuously measure the water content in oil (in ppm). Then, water content in cellulose can be derived from the measured moisture concentration in oil based on cellulose isotherms [20]. By using this setup, the moisture content of the transformer could be continuously monitored.

The moisture diffusion experiments were conducted in the following steps. Firstly, the transformer was heated up and maintained at 55 ℃ for seven days to attain equilibrium status of moisture. Then a sinusoidal temperature profile was applied to the transformer for seven days. One cycle of this temperature profile was 24 hours and the temperature variations were between 30 ℃ and 80 ℃. After that the transformer was again kept at 55 ℃ for seven days to let it reach moisture equilibrium status.

Figure 4.7 shows one complete cycle of the moisture measurement under the sinusoidal temperature. In Figure 4.7, the temperature (in green) and moisture in oil (in red) were directly recorded from the above moisture-in-oil sensor. The water content variation of cellulose interface (in black) was calculated by Fessler equation [21] as (4.31). The average water content in cellulose is calculated by averaging the data shown in black curve.

$$W = 2.173 \times 10^{-7} \times P_v^{0.6685} \times e^{\frac{4275.6}{T}} \tag{4.31}$$

where W denotes the water content, P_v denotes the vapor pressure of water, and T denotes the temperature.

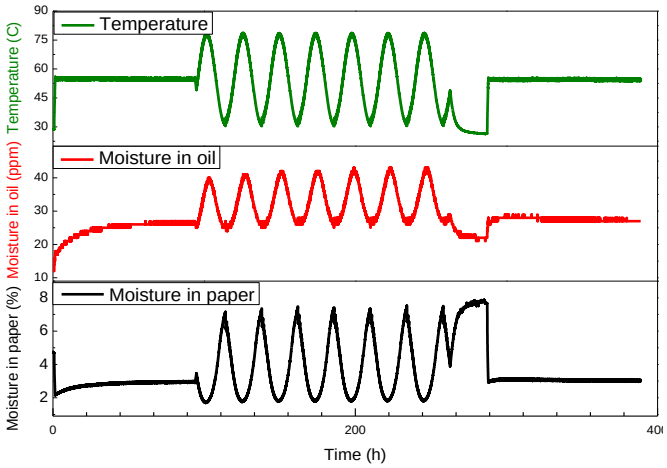


Figure 4.14 Sinusoidal variation of temperature (green), moisture in oil (red) and water content of cellulose surface contacting with oil (black) of the prototype transformer

By using COMSOL Multiphysics software, 3D moisture distribution of the prototype transformer under the above sinusoidal temperature was obtained as shown in Figure 4.8 (t = 172 hours). From Figure 4.8 it can be seen that moisture is not evenly distributed in the transformer. Most part of the cellulose attains a relatively low moisture level (less than 1.5%). However, the moisture of cellulose surface in contact with oil may reach up to 4%-7%. Along the transformer's axial direction, the moisture also presents a non-uniform distribution. This is caused by the variation in the temperature and oil flow inside the transformer.

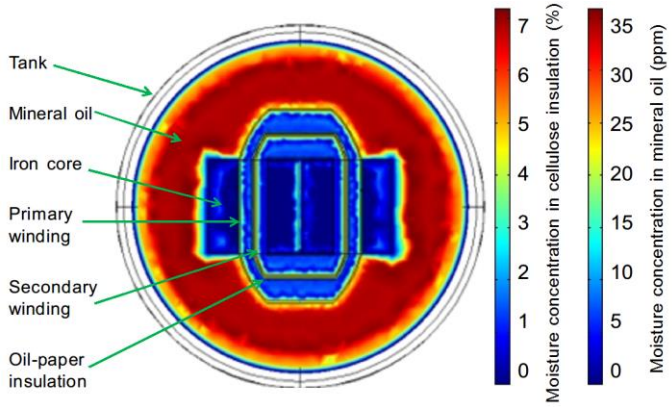


Figure 4.15 Moisture distribution of the prototype transformer under sinusoidal temperature (at time instance t = 172 hours). The results are obtained by multi-physics modelling method.

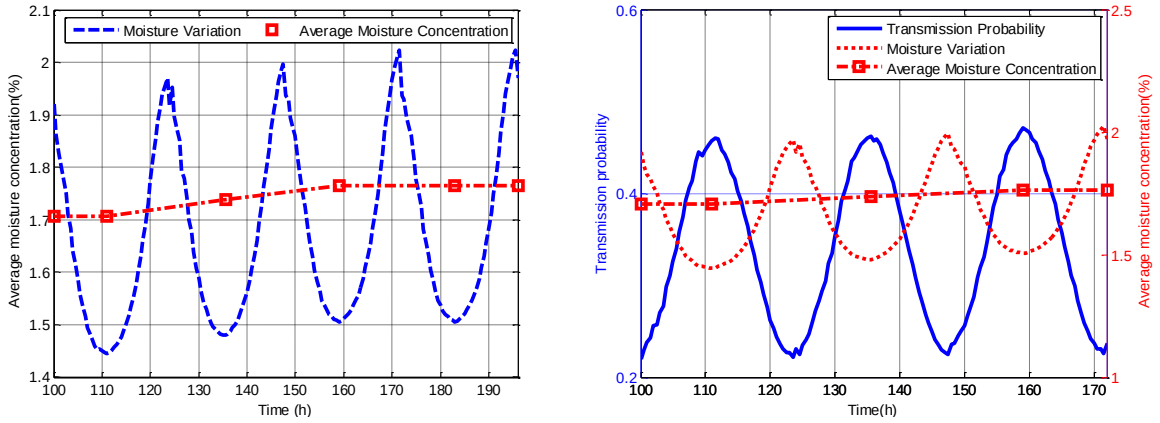
Based on the estimated moisture in Figure 4.8, the overall moisture of the whole cellulose bulk can be obtained as (4.32).

$$W_{average} = \frac{1}{V_c}\iiint_{\Omega} W(x, y, z)dV \tag{4.32}$$

where V_c is the volume of cellulose bulk.

The calculated overall moisture of the cellulose bulk at different diffusion time is shown in Figure 4.16a. As can be seen from Figure 4.16a that the overall moisture varies as the sinusoidal shape within the range from 1.45%-2.3%. The average moisture of the transformer's cellulose is 1.76 % by taking the average of data presented in Figure 4.16a.

Particle tracing method is also applied to estimate the moisture content in the transformer. The transmission probability of water particles and average moisture concentration of the transformer are calculated as shown in Figure 4.16b.



(a) (b)

Figure 4.16 Comparison of moisture concentration between multi-physics model and particle tracing method. (a) Average moisture concentration of the cellulose in the prototype transformer at different time under sinusoidal temperature profile shown in Figure 4.7. The results are obtained by multi-physics modelling method. (b) Average moisture concentration of the cellulose in the prototype transformer at different time under sinusoidal temperature profile shown in Figure 4.7. The results are obtained by particle tracing method. The water particles' transmission probability is also shown (in blue color).

From Figure 4.16b it can be seen that there is a high correlation between the moisture distribution inside the cellulose and the particles transmission behavior. When the moisture diffusion is excited

by the sinusoidal temperature, the particles transmission probability exhibits a sinusoidal variation. The moisture in the cellulose also shows a sinusoidal variation but in the opposite direction of particles transmission probability.

The above phenomenon in Figure 4.16b can be explained as follows. Particles transmission probability changes with the variations in the temperature. When the temperature reaches the peak value of 80 ℃ (112 hours diffusion in Figure 4.16b), the maximum transmission probability rises to 0.46 and this implies that the water particles have the highest action energy at this moment. With the highest action energy, the water particles are more prone to leave the host (cellulose) and enter oil or reflect at the boundaries rather than to stay steadily. Consequently, the lowest moisture concentration in cellulose occurs at this time. When the temperature drops, water particles have less action energy and most of them tend to reside inside the cellulose without movement. Therefore, the overall moisture contents of the bulk cellulose material increases. By comparing Figure 4.16a and Figure 4.16b it can be seen that multi-physics modelling and particle tracing method show good agreement in estimating moisture content in the cellulose of the transformer.

To further verify the particle tracing method, extensive moisture diffusion experiments have been performed on the transformer. The diffusion experiments were divided into five stages and each stage took 21 days, including: (1) the transformer was maintained at 55 ℃ for seven days to attain moisture equilibrium; (2) the sinusoidal temperature profile (the same as aforementioned earlier in this section, refer to Figure 4.7) was applied to the transformer for seven days; and (3) the transformer was kept at 55℃ for one week to facilitate moisture equilibrium. After the above procedures, dummy paper samples were collected from three locations corresponding to upper, medium and bottom height of transformer winding. This takes the non-uniform moisture distribution into consideration. Karl Fischer Titration (KFT) [22] was used to validate the calculation results from both multi-physics modelling and particle tracing method. Table 4.10 summarizes the modelled and KFT measured moisture at different stages.

Table 4.10 Results comparison of moisture concentration (%wt) in model transformer

Time (days)	Multi-physics	Particle	KFT
0	1.55	1.60	1.52
7	1.63	1.68	1.60
14	1.65	1.71	1.37
21	1.66	1.73	1.73
28	1.66	1.73	1.25
35	1.70	1.75	1.55

From Table 4.10 it can be seen that the moisture in transformer's cellulose estimated by particle tracing method and multi-physics are quite close. These results are also similar to that measured by KFT except some discrepancies at 14 days and 28 days of ageing experiments. This is mainly because the procedures of collecting pressboard specimens for KFT measurement may not be consistent at different collections. The above results demonstrate particle tracing method can provide an alternative for estimating the moisture in the transformers from the microscopic view.

4.2.4.3 Discussion on the Complexity of Particle Tracing Method

The challenges of applying particle tracing method for estimating moisture concentration of transformers lay in two aspects:

(1) Comprehensive geometric information of transformer's insulation system. If the geometry of the transformer insulation is known, particle tracing method can be applied to estimate the moisture distribution in the cellulose of the transformer. However, if the detailed geometry information of the transformers is not available, it will cause certain difficulties in its implementation. In this case, some approximation needs to be made. It is assumed that the transformer's insulation construction follows common engineering practice and meets certain criteria (standards) to withstand electromagnetic, thermal, and mechanical stresses during the transformer's operation. Thus, by using the nameplate information of the transformer (e.g. rating, core type, temperature rise, oil volume etc.), the prototype insulation geometry of the transformer will be used. Based on the approximated geometry, the particle tracing method can then be applied for moisture estimation.

(2) Complex relationship between oil/cellulose properties and simulation configuration. Moisture migration between oil and cellulose insulation is a complex process. Moisture diffusion can be highly influenced by the properties of oil (density, viscosity, thermal conductivity, etc.) and cellulose (density, ageing condition, etc.) medium. Changes in oil/cellulose characteristics will result in variation of different driving forces and boundary conditions in particle tracing modelling.

Take boundary configuration as an example, the stick boundary describes whether the water particles can move further when they enter the cellulose while the diffuse-reflection boundary models the water particles' reflection when they collide with the cellulose. When different cellulose materials are used in transformers or the cellulose becomes degraded, the possibility fraction between the stick boundary and reflection boundary needs to be altered. When the cellulose is aged, the possibility fraction of stick boundary needs to be increased to more than 50%. This implies that

the water particles are more prone to stay with the cellulose instead of being bounced back from cellulose into oil.

Moreover, for diffuse-reflection boundary, the probability that water particles can bounce back from cellulose also needs to be adjusted based on the properties of oil and cellulose. Therefore, further studies are required to investigate the complex relationship between oil/cellulose properties and simulation configuration to provide an accurate moisture estimation of transformers.

4.2.5 Conclusion

This paper proposed particle tracing method to estimate the moisture contents of the oil-impregnated cellulose in transformers. Particle tracing method can avoid convergence difficulties in the FEM method. Particle trajectories are computed in a Lagrangian reference frame, removing the restriction on the ranges of the Peclet number. In particle tracing method, water particles' motion can be visualized and the moisture distribution inside the cellulose can be derived. The experiment results of moisture diffusion on the pressboard samples and a prototype transformer verified the particle tracing method. The particle tracing method could be used as a complement to conventional methods (Fick's law, FEM method) for moisture estimation in transformer oil-cellulose system, especially when conventional methods encounter numeric instabilities.

Acknowledgments

Supports from Australian Research Council, and industry partners Powerlink Queensland, Energex, Ergon Energy, and TransGrid are gratefully acknowledged.

References

[1] L. E. Lundgaard, W. Hansen, D. Linhjell and T. J. Painter, "Aging of oil-impregnated paper in power transformers," *IEEE Trans. Power Delivery*, vol.19, Issue 1, pp. 230-239, 2004.

[2] D. F. Garcia, B. Garcia and J. Burgos, "A review of moisture diffusion coefficients in transformer solid insulation-part 1: Coefficients for paper and pressboard," *IEEE Electr. Insul. Mag.*, vol.29, Issue 1, pp. 46-54, 2013.

[3] P. F. Ast, "Movement of moisture through A50P281 Kraft paper (dry and oil impregnated),", 1966.

[4] S. D. Foss and L. Savio, "Mathematical and experimental-analysis of the field drying of power transformer insulation," *IEEE Trans. Power Delivery*, vol.8, Issue 4, pp. 1820-1828, 1993.

[5] A. F. Howe, "The diffusion of moisture through power transformer insulation," PhD dissertation, Dept. Electr. Electron. Eng., Univ. Nottingham, Nottingham, UK, 1975.

[6] A. S. Asem and A. F. Howe, "Drying of power-transformer insulation," *IEE Proceedings - Generation, Transmission and Distribution*, vol.129, Issue 5, pp. 228-232, 1982.

[7] Y. Du, "Measurements and modeling of moisture diffusion processes in transformer insulation using interdigital dielectrometry sensors," PhD dissertation, Dept. Electr. Eng. Comp. Sci, Mass. Inst. Tech., Cambridge, MA, 1999.

[8] R. Villarroel, D. F. Garcia, B. Garcia and J. C. Burgos, "Diffusion coefficient in transformer pressboard insulation part 2: mineral oil impregnated," *IEEE Trans. Dielectr. Electr. Insul.*, vol.21, Issue 1, pp. 394-402, 2014.

[9] B. Garcia, J. C. Burgos, A. M. Alonso and J. Sanz, "A moisture-in-oil model for power transformer monitoring - Part I: Theoretical foundation," *IEEE Trans. Power Delivery*, vol.20, Issue 2, pp. 1417-1422, 2005.

[10] "Particle Tracing Module-For Studying the Interaction Between Particles and Fields" available at http://www.comsol.com/particle-tracing-module.

[11] M. A. Tsili, E. I. Amoiralis, A. G. Kladas and A. T. Souflaris, "Power transformer thermal analysis by using an advanced coupled 3D heat transfer and fluid flow FEM model," *Int. J. Therm. Sci.*, vol.53, Issue1, 188-201, 2012.

[12] Y. T. Feng, K. Han, C. F. Li and D. R. J. Owen, "Discrete thermal element modelling of heat conduction in particle systems: Basic formulations," *J. Comput. Phys.*, vol.227, Issue 10, pp. 5072-5089, 2008.

[13] R. B. Keey, T. A. G. Langrish and J. C. F. Walker, *Kiln-Drying of Lumber*, Berlin Heidelberg: Springer Verlag, 2000.

[14] C. T. Crowe, M. Sommerfeld and Y. Tsuji, *Multiphase Flows with Droplets and Particles*, CRC Press, 1997.

[15] D. Susa, M. Lehtonen and H. Nordman, "Dynamic Thermal Modelling of Power Transformers," *IEEE Trans. Power Delivery*, vol.20, Issue 1, pp. 197-204, 2005.

[16] R. B. Jadav, C. Ekanayake and T. K. Saha, "Impact of moisture and ageing on the dielectric response of transformer insulation," in *Proceedings of 22nd Australasian Universities Power Engineering Conference(AUPEC),September 26 - 29,* 2012, Bali, India, pp. 1-6.

[17] *Standard Practice for Maintaining Constant Relative Humidity by Means of Aqueous Solutions*, ASTM E104-02 (2007), 2002.

[18] C. Ekanayake, S. M. Gubanski, A. Graczkowski and K. Walczak, "Frequency response of oil impregnated pressboard and paper samples for estimating moisture in transformer insulation," *IEEE Trans. Power Delivery*, vol.21, Issue 3, pp. 1309-1317, 2006.

[19] V. Oyj,"MMT330 Moisture and Temperature Transmitter Series for Oil" available at http://www.vaisala.com/en/industrialmeasurements/products/moistureinoil/Pages/MMT330.aspx.

[20] D. Martin, C. Perkasa and N. Lelekakis, "Measuring paper water content of transformers: a new approach using cellulose isotherms in nonequilibrium conditions," *IEEE Trans. Power Delivery*, vol.28, Issue 3, pp. 1433-1439, 2013.

[21] W. A. Fessler, T. O. Rouse, W. J. McNutt and O. R. Compton, "A refined mathematical model for prediction of bubble evolution in transformers," *IEEE Trans. Power Delivery*, vol.4, Issue 1, pp. 391-404, 1989.

[22] *Insulating Liquids—Oil-Impregnated Paper and Pressboard-Determination of Water by Automatic Coulometric Karl Fischer Titration*, IEC 60814 Ed. 2.0, 1997.

Chapter 5

Understanding Moisture Dynamics and Its Effect on Dielectric Response of Transformer Insulation

Contribution of the Chapter

This chapter investigates moisture dynamics and its effect on dielectric response of a transformer's cellulose insulation. It proposes a distributed parameter model to reveal the correlation between moisture distribution (under non-equilibrium conditions due to thermal transients) and dielectric response parameters (dielectric losses and permittivity) of cellulose insulation. It then estimates these parameters under moisture non-equilibrium conditions. The accelerated ageing and moisture diffusion experiments are conducted on a prototype transformer to verify the proposed model. The methodology developed in this chapter can help the proper interpretation of dielectric response measurement of field transformers under thermal transients.

Understanding Moisture Dynamics and Its Effect on Dielectric Response of Transformer Insulation

Yi Cui, Hui Ma, Tapan Saha, and Chandima Ekanayake

The University of Queensland, Brisbane, Australia

Publication Journal: IEEE Transactions on Power Delivery

Submitted: September 25, 2014

Revision Resubmitted: December 11, 2014

Accepted: April 22, 2015

Published: April 24, 2015

Author's Contributions:

Yi Cui: Primary model design, theoretical simulation, experimental validation, measurement results analysis and manuscript preparation.

Hui Ma: Supervision on the experimental measurement, results discussion and manuscript revision.

Tapan Saha: Supervision on the experimental measurement and manuscript revision.

Chandima Ekanayake: Results discussion and manuscript revision.

Abstract

Dielectric response measurement has recently been adopted by utilities for evaluating moisture content in cellulose insulation (paper and pressboard) of transformers. Moisture distribution is highly dependent on temperature. Since the temperature inside a transformer may change during the dielectric response measurement, the moisture in the transformer's cellulose and oil insulation can hardly attain an equilibrium state. Instead, moisture dynamics exist inside the transformer: (1) cellulose absorbs (desorbs) moisture from (to) oil with the changes in temperature; and (2) moisture migrates inside cellulose due to a moisture gradient. This paper investigates moisture dynamics and its effect on dielectric response of a transformer's cellulose insulation. It proposes a distributed parameter model to reveal the correlation between moisture distribution (under non-equilibrium conditions due to thermal transients) and dielectric response parameters (dielectric losses and permittivity) of cellulose insulation. It then estimates these parameters under moisture non-equilibrium conditions. The accelerated ageing and moisture diffusion experiments are conducted on a prototype transformer to verify the proposed model. The methodology developed in this paper can help the proper interpretation of dielectric response measurement of field transformers under thermal transients.

Index Terms: Cellulose, dielectric response, insulation, moisture diffusion, moisture dynamics, oil, transformer.

5.1 Introduction

Life expectancy of a power transformer is largely determined by the ageing condition of its cellulose insulation. Moisture is one of the most harmful agents for cellulose insulation. It can accelerate the cellulose ageing rate and reduce both dielectric and mechanical strength of cellulose insulation [1]. Therefore, it is of great interest for utilities to estimate moisture content in cellulose insulation of transformers.

Moisture content in cellulose can be directly measured by the Karl Fischer Titration (KFT) method. However, it requires collecting paper samples from a transformer's winding, which is difficult in practice. An alternative approach is through measuring the moisture content of an oil sample and then determining the moisture content in the cellulose from equilibrium charts [2-4]. This approach assumes an equilibrium state of moisture distribution in the transformer. However, it is hard to attain

an equilibrium state in the transformer and any variation in temperature tends to change the moisture distribution in cellulose and oil [5]. This can lead to an inaccurate estimation of moisture in the cellulose [6].

The utilities have widely adopted the dielectric response method for estimating moisture content in transformers' cellulose insulation. During dielectric response measurement, the transformer is removed from the grid and thus the temperature inside the transformer continuously drops. Consequently, during the course of the dielectric response measurement, moisture migrates between cellulose and oil. This will influence the interpretation of the results of dielectric response measurement and affect the accuracy of moisture estimation in the transformer's cellulose insulation. As an example, the authors made two dielectric measurements on one utility's transformer. During the first measurement a moisture content of 2.9 % at 30 ℃ top-oil temperature was recorded. The second measurement commenced immediately after, and recorded 3.6 % at 26.5 ℃. Therefore, for an accurate estimation of moisture in cellulose insulation based on dielectric response measurement, it is necessary to investigate the correlations between moisture dynamics and dielectric response parameters (i.e. dielectric loss and permittivity) of a transformers' insulation.

This paper is aimed at understanding temperature dependent moisture dynamics and its effect on the dielectric response of the cellulose insulation of transformers. A distributed parameter model is proposed to explore correlations between moisture distribution under non-equilibrium conditions and dielectric response of cellulose insulation. By using this model, dielectric losses and permittivity of cellulose insulation are estimated when a transformer is under temperature variation and consequent moisture non-equilibrium. To verify the methodology developed in this paper, extensive ageing and moisture diffusion experiments are performed on a prototype transformer.

5.2 Moisture in Transformer's Insulation System

5.2.1 Moisture dynamics in cellulose and oil insulation

Due to moisture ingress from the environment and cellulose degradation, moisture is present in a transformer. During the normal operation of a transformer, most moisture is affiliated with its cellulose insulation. Water generally has low solubility in transformer oil. With an increase in temperature, the water solubility in oil can be significantly increased. On the other hand, free water can be formed if the moisture in oil exceeds the saturation level.

For an in-service transformer, it can absorb water from environment and its oil can contain some amount of water. The transformer's oil-cellulose insulation system is in an equilibrium state if the partial pressure of water in oil equals the partial pressure of water in cellulose. When the partial pressures in oil and cellulose become different, cellulose absorbs/desorbs water from/to oil to maintain the equilibrium. Moreover, moisture diffusion occurs inside cellulose insulation due to a moisture gradient.

5.2.2 Estimation of moisture contents in cellulose insulation

Over the past twenty years, extensive studies have been performed for estimating moisture content in cellulose insulation. Most of these studies are focused on moisture diffusion using Fick's second law [7] in one dimension as (5.1)

$$\frac{\partial W(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial W(x,t)}{\partial x} \right) \tag{5.1}$$

where $W(x,t)$ is the moisture content of cellulose at position x and time t. D denotes the diffusion coefficient and is usually not a constant but depends on both moisture concentration and temperature. By conducting laboratory experiments, researchers have derived diffusion coefficients for different types of cellulose (i.e. paper or pressboard, non-impregnated or impregnated, aged or non-aged, paper/pressboard with different thickness) under different local moisture concentration and temperatures [2-9]. However, certain variations may exist in these derived diffusion coefficients since the experiments were conducted with different setups and conditions. Such variations can pose difficulties for utilities on selecting the correct coefficient for a moisture content estimation in their transformers.

García *et al.* proposed a moisture-in-oil model for power transformer monitoring [8]. The model considered the equilibrium relations between oil and cellulose, the moisture dynamics before attaining the steady-state equilibrium and the increase of the amount of moisture due to cellulose ageing. However, it is not clear whether moisture in cellulose can be derived from this model.

Originally proposed by Piper and Fessler [9], the isothermal approach has also attracted the attention of some researchers [10]. This approach uses temperature and vapor pressure to derive the local moisture content in cellulose from the measurement of moisture-in-oil sensor. However, the sensor's location has a large influence on the moisture estimation. Given the complexity of a transformer's construction and different types of cellulose used at different layers and locations inside

the transformer, a sufficient number of moisture-in-oil sensors need to be installed to provide an accurate estimation of moisture in the cellulose insulation of a transformer.

An alternative technique for estimating moisture content in cellulose is the dielectric response method. It measures several dielectric response parameters, which are influenced by temperature, moisture content and the ageing condition of the cellulose insulation in a transformer. By using a database built based on well-defined pressboard samples with different temperatures and moisture contents, it is possible to estimate the moisture content in cellulose insulation in a transformer. The next section will provide a brief review on dielectric response technique for moisture content estimation.

5.3 Dielectric Response of Transformer Insulation

5.3.1 Dielectric response in frequency domain

When a dielectric material is imposed by an alternating electric field $\hat{E} = E_m e^{i\omega t}$, the complex dielectric displacement in this material becomes (5.2)

$$\hat{D}(\mathrm{t}) = \varepsilon_0 \varepsilon_\infty E_m e^{i\omega t} + \varepsilon_0 \int_{-\infty}^{t} f(\mathrm{t}-\mathrm{t}_0) E_m e^{i\omega t_0} dt_0 \tag{5.2}$$

where ε_0 denotes the permittivity of vacuum, ε_∞ denotes the high frequency relative permittivity of the material and *f(t)* is the response function of the material which monotonically decreases with time.

The Fourier transform of the response function $f(t)$ yields the complex susceptibility, $\hat{\chi}(\omega)$

$$\hat{f}(\omega) = \int_0^\infty f(t) e^{-i\omega t} dt = \hat{\chi}(\omega) = \chi'(\omega) - j\chi''(\omega) \tag{5.3}$$

The real and the imaginary parts of the complex susceptibility are not independent from each other since they are both generated by the same response function $f(t)$. They can be regarded as the cosine and sine transforms of the response function respectively.

The total current density $\hat{j}(\omega)$ in the dielectric material under $\hat{E}(\omega)$ excitation can therefore be expressed as

$$\hat{j}(\omega) = i\omega\varepsilon_0[\underbrace{\varepsilon_\infty + \chi'(\omega)}_{capacitive\ \ part} - \underbrace{i(\sigma_{dc}/\varepsilon_0\omega + \chi''(\omega))}_{resistive\ \ part}]\hat{E}(\omega) \tag{5.4}$$

where σ_{dc} denotes dc conductivity of the material. This expression shows that the current is composed of a resistive part and a capacitive part. The resistive part represents the energy losses in the material, which are dominated by two different mechanisms, one is due to dc conduction (movement of free charges) and the other is because of relaxation losses (re-orientation of bonded charges). The capacitive part is associated with the capacitance of the material.

In many situations, it is more convenient to use the complex permittivity instead of the complex susceptibility. It can be defined as follows

$$\hat{j}(\omega) = i\omega\varepsilon_0[\varepsilon'(\omega) - i\varepsilon''(\omega)]\hat{E}(\omega) \tag{5.5}$$

where $\varepsilon'(\omega) = \varepsilon_\infty + \chi'(\omega)$ and $\varepsilon''(\omega) = \sigma_{dc}/\varepsilon_0\omega + \chi''(\omega)$.

The above equations show that the dc conductivity σ_{dc}, the high frequency component of the relative permittivity ε_∞, and the complex dielectric susceptibility $\hat{\chi}(\omega)$, characterise the dielectric material in the frequency domain. It is possible to determine these parameters by measuring the magnitude and phase angle of the resultant currents when the material is subjected to an alternating voltage under different frequencies.

5.3.2 Dielectric response measurement for moisture estimation

For estimating moisture content in a transformer, the geometric information of the winding insulation of the transformer under investigation is needed. For a core type transformer the main insulation usually consists of a number of cylindrical shells of pressboard barriers, separated by axial spacers. A so-called *X-Y* model is widely used to represent such a structure [6]. By making use of a database and *X-Y* model, an algorithm can be implemented to find the best fit between the response of the model and the measured response of the transformer. This result can then be used for moisture estimation of a transformer's cellulose insulation.

The above moisture estimation assumes the transformer under test is kept at a constant temperature and attains moisture equilibrium. However, equilibrium is difficult to attain in the transformer during measurement. It is therefore necessary to investigate the correlation between temperature dependent moisture dynamics and the dielectric response of cellulose insulation. This will pave a way for properly evaluating moisture in cellulose insulation of the transformer onsite, which is under a

temperature transient.

In the next section a distributed parameter model is proposed for studying the dielectric behavior of oil and cellulose insulation under non-equilibrium and non-uniform moisture distribution in a transformer.

5.4 Distributed Parameter Dielectric Response Model

The proposed distributed model is shown in Figure 5.1. In this model, the pressboard with thickness l is sliced into N layers ($i = 1, 2, ..., N$). Assuming moisture in the transformer is in a non-equilibrium state, the moisture is not uniformly distributed along the thickness of the cellulose insulation. Consequently, each layer of pressboard has a different dielectric loss δ_i and permittivity ε_i.

The impedance $Z(\omega)$ of the whole piece of the pressboard in the frequency domain can be written as (5.6)

$$Z(\omega) = 1/G(\omega) = \sum_{i=1}^{N} Z_i(\omega) = \sum_{i=1}^{N} 1/G_i(\omega) \tag{5.6}$$

where N denotes the total number of sliced layers of the pressboard, $Z_i(\omega)$ and $G_i(\omega)$ are the impedance and admittance of the *i*-th layer of the pressboard respectively.

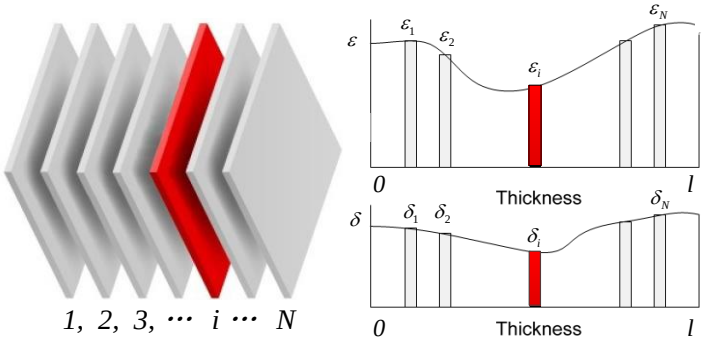


Figure 5.1 Distributed model for modelling dielectric response of pressboard

In Figure 5.1, each layer of pressboard can be represented by a parallel circuit consisting of a capacitor and a resistor. For the *i*-th layer of the pressboard, its admittance can be written as

$$G_i^*(\omega) = (\delta_i + j\omega\varepsilon_0\varepsilon_i^{'})S/\Delta x \tag{5.7}$$

According to frequency domain dielectric spectroscopy theory [11], (5.6) can be rewritten as (5.8)

$$Z^*(\omega) = \sum_{i=1}^{N} \Delta x / S\left(\delta_i + j\omega\varepsilon_0\varepsilon_i^{'}\right) \tag{5.8}$$

where S denotes the total area of the cylindrical pressboard, Δx is the thickness of each sliced layer of the pressboard, $\varepsilon_i^{'}$ denotes the real part of complex permittivity of the *i*-th layer of pressboard and δ_i represents the summation of conductive and polarization losses of the *i*-th layer of pressboard (the loss of the whole piece of pressboard is shown in (5.5)).

Assuming that the above pressboard can be sliced infinitely thin, (5.8)can be rewritten as (5.9)

$$Z^*(\omega) = \frac{1}{S}[\int_0^l \frac{\delta_x}{\delta_x^2 + (\omega\varepsilon_0\varepsilon_x^{'})^2} dx - j\int_0^l \frac{\omega\varepsilon_0\varepsilon_x^{'}}{\delta_x^2 + (\omega\varepsilon_0\varepsilon_x^{'})^2} dx)] \tag{5.9}$$

In dielectric response measurement, the phases and magnitudes of the applied AC voltage and resultant current can be measured. Subsequently, the impedance $Z^*(\omega)$ of the pressboard can be obtained. Then the complex capacitance of the pressboard can be calculated by using (5.10)

$$Z^*(\omega) = 1/G^*(\omega) = 1/j\omega C^*(\omega) = 1/\ j\omega C_0\varepsilon^*(\omega) \tag{5.10}$$

Combining (5.7) and(5.10), the complex capacitance of the pressboard is expressed in (5.11)

$$C_i^*(\omega) = G_i^*(\omega)/\ j\omega = (\varepsilon_0\varepsilon_i^{'} - j\delta_i/\omega)S/\Delta x \tag{5.11}$$

On the other hand, from (5.9) and (5.10) the complex capacitance of the pressboard can also be directly written as

$$C_i^*(\omega) = \varepsilon_i^*(\omega)C_0 = \varepsilon_0 S\left(\varepsilon_x^{'} + j\varepsilon_x^{''}\right)/\Delta x \tag{5.12}$$

By comparing (5.11) and(5.12), it can be expressed as

$$\varepsilon_x^{'} = \varepsilon_i^{'}, \quad \delta_x = \delta_i = \varepsilon_0\omega\varepsilon_i^{''} \tag{5.13}$$

Eq. (5.13)reveals that $\varepsilon_x^{'}$ in (5.8) equals the real part of permittivity $\varepsilon_i^{'}$ while δ_x in (5.8) can be calculated from the imaginary part of permittivity $\varepsilon_i^{''}$. Both the real and imaginary parts of permittivity of pressboard can be obtained from the results of dielectric frequency response measurements. The above distributed model will be used to investigate the moisture dynamics effect on dielectric response measurement in the following sections.

5.5 Experimental Setup

5.5.1 Prototype transformer configuration

A prototype transformer is used for experimental study in this paper. It is a single phase transformer rated at 5 kVA with 240 V secondary and 2.2 kV primary. Since the losses due to current flowing through the transformer's winding may not generate the required temperature, a heater was installed at the bottom of the transformer. This also helped to control the prototype transformer's temperature to simulate different thermal conditions for the accelerated ageing and moisture diffusion experiments. Figure 5.2 shows the geometry of the insulation of the prototype transformer.

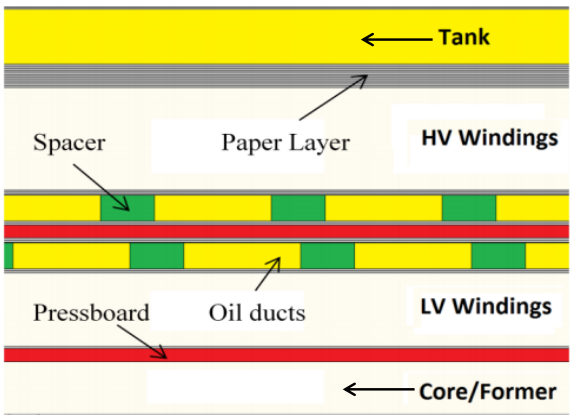


Figure 5.2 Insulation geometry of the prototype transformer

In the prototype transformer, the cellulose insulation includes hybrid Diamond Dotted Paper (DDPP), Mouldable pressboard (MPB) and Spacers All Round (SAR) and they are immersed in transformer oil (Nynas Libra). The temperature class of the transformer's paper insulation is A. The ratio in *X*-*Y* model is calculated as $X = 38\%, Y = 16\%$, where *X* is the lumped sum of the thickness of all barriers in the duct divided by the duct width; *Y* is the total width of all spacers divided by the total length of the periphery of the duct [6].

5.5.2 Moisture dynamic experiments

After the commission of the prototype transformer, it was subjected to both electrical and thermal loading to attain a certain degree of ageing of its cellulose insulation (equivalent to 34 years of life

consumption based on the degree of polymerization (DP) measurement of this transformer). Electrical loading was imposed by using a load bank with maximum power capacity of 6 kW. The prototype transformer was kept at 110 ℃ (using the abovementioned heater) with 30A load current for an effective time period equal to 35 days (the transformer was kept at 50 ℃ during the night and weekends).

After 35 days of accelerated electrical and thermal ageing of the prototype transformer, experiments were arranged to study the moisture dynamics and its effect on dielectric response of the cellulose insulation of the transformer. In the experiments, a sinusoidal temperature profile was imposed on the prototype transformer by using the above mentioned heater. The prototype transformer is sealed from the atmosphere. A moisture-in-oil sensor (Vaisala MMT 330 [12]) was installed in the prototype transformer. The sensor's tip was close to the cellulose insulation of the winding. By using this setup, the moisture at the interface between oil and cellulose insulation could be continuously monitored. Temperature was also measured by this sensor.

The moisture dynamics experiments were conducted in three steps:

(1) The transformer was kept at a constant temperature (55 ℃) for one week to let it attain moisture equilibrium.

(2) The transformer was subjected to a sinusoidal temperature profile (Figure 5.3) for one week. The period of one cycle sinusoidal temperature was 24 hours with the highest temperature of 80 ℃ and the lowest temperature of 30 ℃.

(3) The transformer was kept at a constant temperature (55 ℃) for another one week to let it attain moisture equilibrium. Dielectric response measurement was then performed.

Figure 5.3 presents one complete cycle of moisture measurements under sinusoidal temperature. In the figure, the temperature (in green), moisture in oil (in red) and water activity (in blue, is defined as the ratio of the partial pressure of water in the material and the saturated vapor pressure of pure water at the same temperature [10]) were directly obtained from the above moisture-in-oil sensor.

The moisture concentration at the cellulose surface in contact with oil (in black in Figure 5.3) was obtained using Fessler [9] equation as

$$W = 2.173 \times 10^{-7} \times P_v^{0.6685} \times e^{4275.6/T} \tag{5.14}$$

where W is the concentration of absorbed water as the ratio of the mass of water to the mass of dry cellulose, P_v is the vapour pressure of water in atm and T is the temperature in Kelvin.

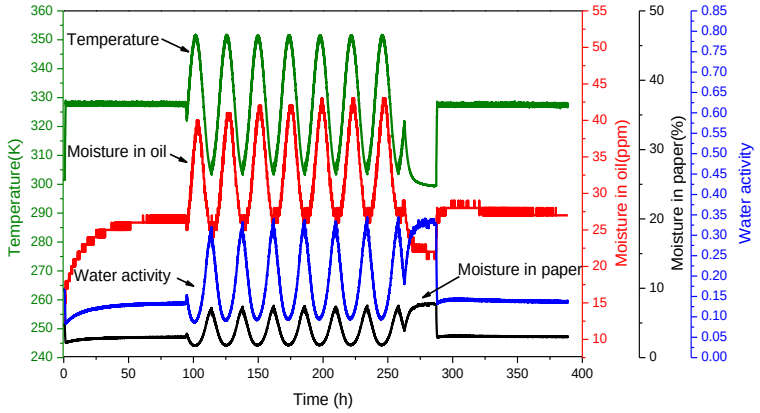


Figure 5.3 Sinusoidal variation of temperature (green), water activity in oil (blue), moisture in oil (red, in ppm) and water content at the cellulose surface in contact with oil n (black, in percentage) of prototype transformer

5.6 Modelling of Moisture Dynamics in Cellulose

This section investigates the above sinusoidal temperature driven moisture dynamics in cellulose. The time of moisture diffusion inside cellulose is much larger than that of moisture migration at the cellulose's surface contacting oil. Therefore, this section mainly focuses on moisture diffusion in cellulose.

This section firstly models moisture diffusion in un-aged pressboards. Then it investigates moisture diffusion in the prototype transformer based on the experiments in Section 5.5.

5.6.1 Modelling moisture diffusion in pressboard

It is assumed an un-aged dry pressboard (1 mm in thickness with less than 0.5% water content) has one side of its surface contacting with a large volume of oil, which contains a sufficient amount of water. When the pressboard is immersed into the oil, the partial pressure of water in oil is significantly larger than that in pressboard. Thus, the pressboard absorbs water from the oil when it is just immersed into the oil.The other side of the pressboard surface is sealed by an electrode for dielectric response measurement. The whole test cell is considered as a closed system. Under sinusoidal temperature variation, the moisture content at this side of the surface follows the curve in black shown in Figure 5.3.

The initial values of moisture content and boundary condition of this pressboard is

$$W(x,t=0)=W_0,\quad W_0<0.5$$
$$W(x=0,t)=W_s,\frac{\partial W(x=l,t)}{\partial x}=0 \tag{5.15}$$

where W_0 is the moisture content in the whole bulk of un-aged pressboards and is assumed less than 0.5% initially. W_s is the concentration of moisture at the interface between pressboard and oil and it varies as sinusoidal shape (black curve in Figure 5.3).

By solving Fick's second law (5.1) and (5.15), the moisture distribution along the thickness of the pressboard under sinusoidal temperature profile from the time instances t = 1 to 72 hours are obtained and depicted in Figure 5.4.

It is assumed that at t = 0 when the dry and un-aged pressboard sample was put into the oil tank, there exists a significant moisture gradient between the oil and the pressboard and at their interface vapor pressure of the oil is larger than that of the pressboard. Therefore, at t = 0 hour, the pressboard absorbs water from the oil and eventually water moves towards the depth of the pressboard. At the time instance t = 1 hour, there exists a large moisture gradient in the pressboard: moisture concentration at the pressboard surface is 5.7% and only 0.5% at other locations away from the surface (Figure 5.4(b)). This moisture gradient dominates the water migration and water will continuously move into the pressboard along its thickness for some time though temperature starts to increase.

At the time instances t = 4 and 8 hours, the moisture moves further into the depth of the pressboard. The moisture at the pressboard surface is decreased from 4% to 2.1% while the moisture in most positions of the pressboard increased to above 1% from 0.5% (Figure 5.4(b)). At the time instance t = 12 hours, the moisture at the surface of the pressboard almost attains uniformity of 1.75%. As can be observed from Figure 5.4(a), at t = 12 hours the temperature starts to decrease from 80 ℃. The water in the oil (close to the oil-pressboard interface) starts to move towards the pressboard. Therefore, at the time instances t = 16 and 20 hours, the moisture concentration at the pressboard surface (in contact with oil) increased. The moisture at the pressboard's surface is 2.5%, 5.0% for t= 16 and 20 hours respectively (Figure 5.4(c)). The moisture at the pressboard surface reaches its peak of 6.4% at the time instance t = 24 hours. Also the moisture inside the pressboard is larger compared to previous time instances in this first temperature cycle (Figs. 4(b) and 4(c)).

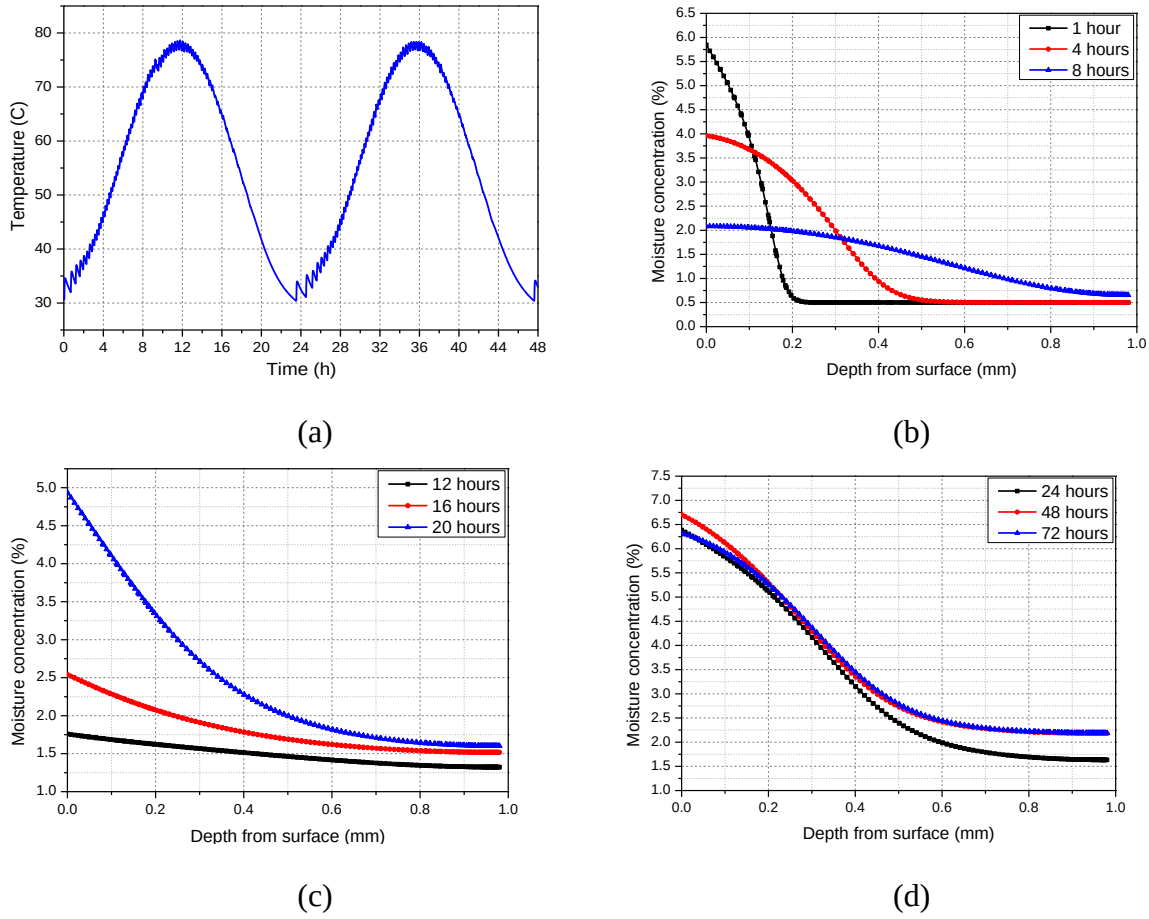


Figure 5.4. Simulation results of moisture distribution in an un-aged pressboard (one dimension moisture diffusion) (a) temperature variation at different time instances (b)-(d) moisture distribution at different time (x=0 refers to the pressboard surface in contact with oil).

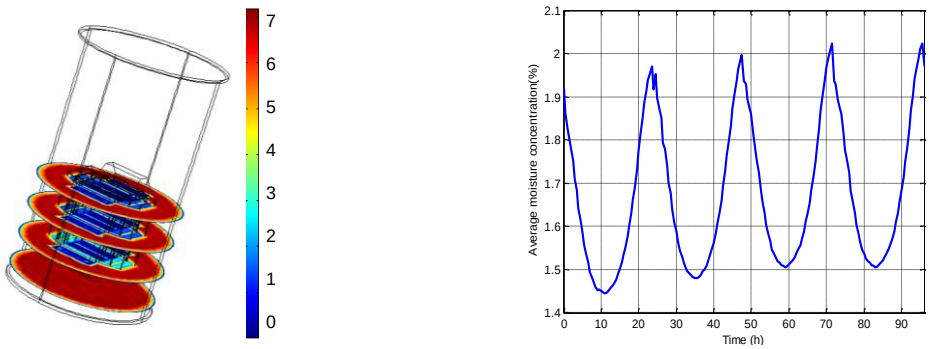
The above moisture dynamics are repeated for the next 24 hours sinusoidal temperature cycle. It can be seen from Figure 5.4(d), the moisture distribution at the surface of the pressboard at the time instances t = 24, 48 and 72 hours are very close (the difference is less than 0.2%). The difference of moisture in the depth of the pressboard between t = 24, 48 and 72 hours is about 0.5%. Such a difference can be explained as follows.

After several cycles of diffusion, the pressboard becomes humid and water migration is mainly dominated by temperature. At low temperature (30 ℃ at time instances t = 24, 48, 72 hours) the slow water molecules' mobility and strong bond between cellulose molecules and water molecules imply that the moisture movement is not able to change its directions immediately when temperature changes. Therefore, a certain amount of moisture moves "back and forth" at a particular position of pressboard and a "standing wave" of moisture distribution can be formed.

5.6.2 Modelling moisture diffusion in prototype transformer

Moisture distribution is not only a function of temperature and moisture gradients but also a function of mass and dimensions of the transformer's insulation system. Therefore, multi-physics modeling [13] is adopted to investigate moisture distribution in three dimensions (3D) of the prototype transformer as shown in Figure 5.5. It considers the effects of electromagnetic, thermal, fluid flow and moisture migration physics on moisture dynamics in transformers. It especially takes into account the coupling and interactions of these physics as they collectively influence the moisture dynamics and moisture distribution inside the transformer. The details of multi-physics modeling will be provided in a future paper.

Figure 5.5(a) presents the moisture distribution in the prototype transformer at the time instance t = 72 hours. Based on the 3D moisture distribution results, the average moisture concentration of the whole bulk of cellulose insulation can be calculated as $W_{average} = \iiint_{\Omega} W(x, y, z)dV/V$, where V is the total bulk volume of the cellulose insulation and $W(x, y, z)$ is the moisture concentration at a particular location of the cellulose insulation. Figure 5.5(b) presents the calculated average moisture concentration of the cellulose insulation at different time instances. It can be observed from Figure 5.5(b) that moisture variations in sinusoidal shape are within the range from 1.45% - 2.0%. The overall moisture of the cellulose insulation of the prototype transformer is 1.7 % by taking the average of data in Figure 5.5 (b). It has good agreement with the results from the Karl Fischer Titration method, which indicates 1.5% moisture content in a cellulose specimen collected from the prototype transformer.



(a) (b)

Figure 5.5 Moisture distribution of the prototype transformer (a) 3D moisture distribution (at time instance t = 72 hours); (b) averaged moisture concentration at different time instances. Transformer was under thermal transients as shown in Figure 5.3.

Figure 5.6 shows the moisture distribution in the prototype transformer's whole bulk cellulose insulation along its radial centerline under different time instances (temperature varies according to Figure 5.3). In Figure 5.6 the legend of the vertical axis is in ppm for moisture concentration in oil and it is in percentage for moisture concentration in cellulose insulation. From Figure 5.6, it can be seen that the moisture is not uniformly distributed inside the cellulose insulation of the prototype transformer. A large proportion of the cellulose has a relatively low moisture level (less than 3%) while the moisture of the cellulose surface in contact with oil (x = 0 and 0.4 m) may reach up to 5% - 7%. The moisture distribution along the transformer height direction also exhibits gradients (Figure 5.5(a)). This is due to the non-uniform thermal distribution along the height direction and the existence of a fluid flow field inside the prototype transformer.

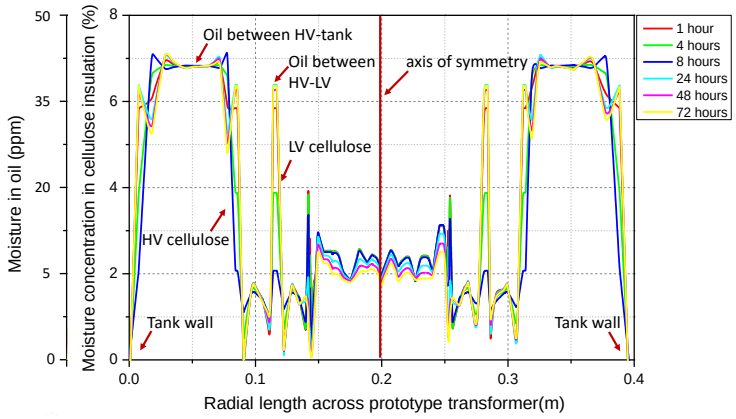


Figure 5.6 Moisture distribution of whole bulk cellulose insulation in the prototype transformer (d = 0.2 m denotes the symmetric line of the cellulose insulation). Temperature varies according to Figure 5.3.

5.7 Analysis of Moisture Dynamics' Effects on Dielectric Response of Cellulose Insulation

Based on the experiments and modeling results provided in Sections 5.5 and 5.6, this section applies the distributed parameter model proposed in Section 5.4 to study the moisture dynamics' effect on dielectric response of cellulose insulation of the prototype transformer. The study is carried out in two steps:

(1) By using a database built upon the dielectric response measurements on the prototype transformer under constant temperature, the distributed model is applied to explore the correlations

between the cellulose's position along its thickness and the cellulose's dielectric response parameters (loss and permittivity) under uniform moisture distribution conditions.

(2) Based on the analysis in step 1, the dielectric response of the cellulose insulation in the prototype transformer under moisture dynamics and temperature transient is estimated. The estimation is verified by the experiments under sinusoidal temperature profile. It needs to be mentioned that the analysis in step 1 does not include the influence of the temperature on the dielectric response. Therefore, the temperature corrections will be conducted (detailed in Section 5.7.3).

5.7.1 Dielectric loss of cellulose insulation vs. moisture

After the prototype transformer attains moisture equilibrium and the moisture is uniformly distributed (Section 5.5.2, Step 3), dielectric frequency response measurement is performed from 1 kHz to 1 mHz. Figure 5.7 presents both real and imaginary parts of permittivity of the cellulose insulation at 50 ℃ with four different moisture contents, i.e. 1.08%, 1.37%, 1.61% and 3.23% in the prototype transformer. It can be seen from Figure 5.7 that the imaginary part of permittivity is influenced by moisture concentration. At a particular frequency, the imaginary part of permittivity increases when more moisture resides in the cellulose insulation.

The dielectric loss of cellulose insulation under each frequency with different moisture contents can be obtained from Figure 5.7 and (5.13), which is shown in Figure 5.8 (data points). Suppose the dielectric loss and the moisture can be modelled as

$$\delta = \alpha_1 e^{\beta_1 \cdot W} \tag{5.16}$$

where W denotes the moisture concentration in the cellulose insulation, α_1 and β_1 are parameters, which can be computed using the least squares method. Figure 5.8 presents the modelled curves of dielectric loss with respect to moisture contents in the cellulose insulation under uniform moisture distribution at 50 ℃.

Once the relationship between dielectric loss and moisture content is obtained by using Figure 5.6 and (16), the dielectric loss of the cellulose insulation at different diffusion positions (x axis in Figure 5.6) at any specific time instance can be estimated. The estimated values at time instance t = 6 hours are shown in Figure 5.9 (denoted as data points). Due to the axis symmetry, only half of the curve (x axis from 0 - 0.2 m) in Figure 5.6 is presented.

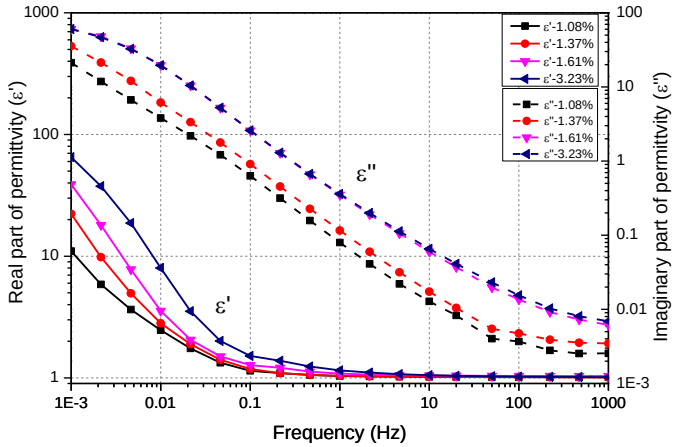


Figure 5.7 Real (ε') and imaginary (ε") parts of permittivity of the prototype transformer insulation under uniform moisture distribution at T = 50 ℃

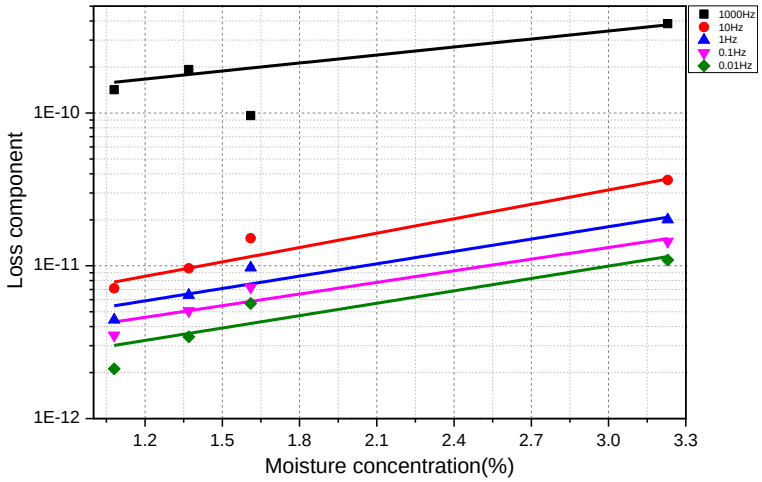


Figure 5.8 Dependency between dielectric loss and moisture contents under uniform moisture distribution condition. T = 50 ℃.

Suppose that the dielectric loss and diffusion position of the cellulose has an exponential relationship

$$\delta = \alpha_2 e^{\beta_2 \cdot x} \tag{5.17}$$

where x denotes the position of cellulose insulation, α_2 and β_2 are parameters which can be computed using the least squares method. The modelled dielectric loss with respect to the position of cellulose at the time instance t = 6 hours is presented in Figure 5.9 (denoted as curves, T = 50 ℃). It should be mentioned that the curves in Figure 5.9 are applied to non-equilibrium and non-uniform moisture distribution in the cellulose insulation.

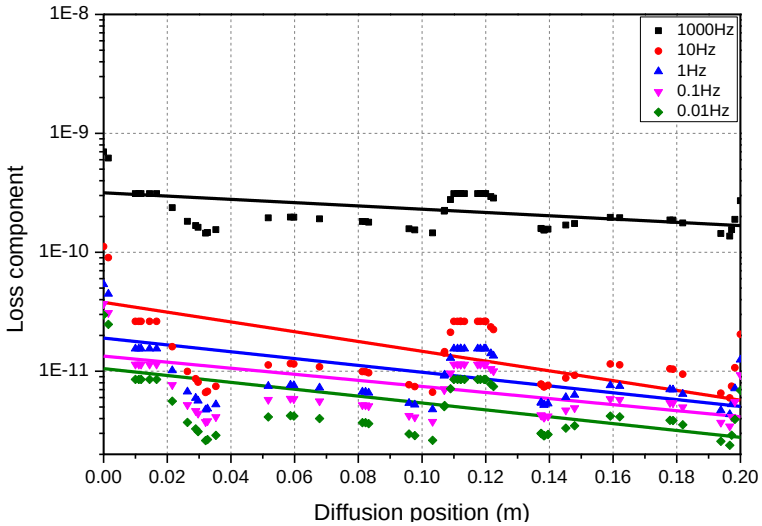


Figure 5.9 Dependency between loss and the depth of cellulose insulation (at time instance t = 6 hours diffusion) Data points denote the values obtained from (5.16) and Figure 5.6; Curves are obtained using (5.17). T = 50 ℃.

5.7.2 Permittivity of cellulosic insulation vs. moisture

The above approaches are also applied to investigate the correlation between the real part of permittivity of cellulose insulation and moisture content. Figure 5.10 depicts the modelled curves of the real permittivity with respect to moisture concentrations under different frequencies at T = 50 ℃. Figure 5.11 depicts the modelled real permittivity with respect to the depth of cellulose insulation at T = 50 ℃. In Figure 5.11, the modelled real part of permittivity is for the time instance t = 6 hours.

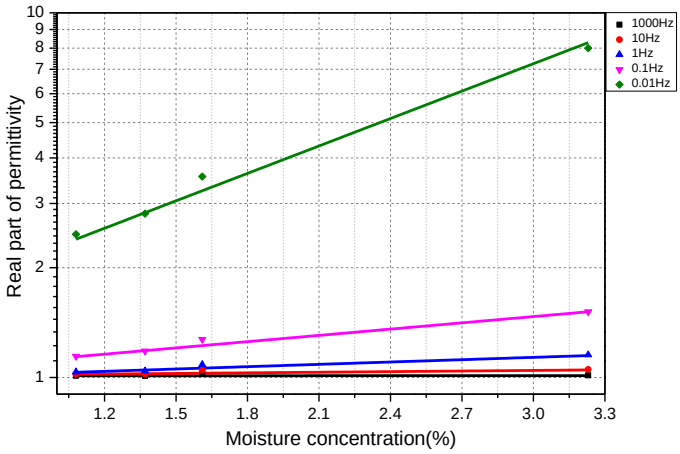


Figure 5.10 Dependency between real part of permittivity and moisture contents under uniform distribution condition at T = 50 ℃.

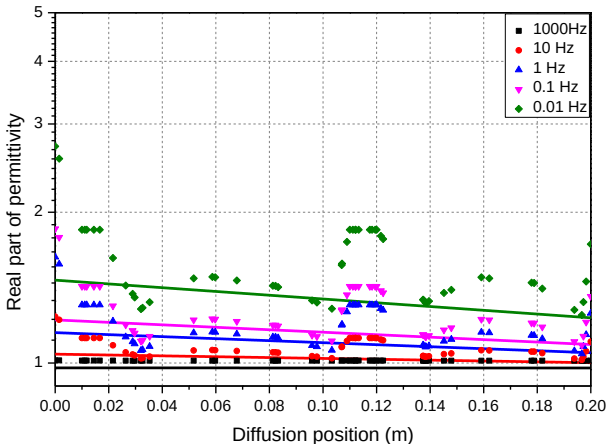


Figure 5.11 Dependency between real part of permittivity and the depth of cellulose insulation (at time instance t = 6 hours). T = 50 ℃.

5.7.3 Validation of distributed parameter model

After obtaining the complex permittivity through the procedures described in Section 5.7.1 and 7.7.2, dielectric response measurement is performed on the prototype transformer. The comparison between the measured and the estimated complex permittivities is then made. The comparison provides a validation for the proposed distributed parameter model.

The dielectric response measurement was repeated every hour for 24 hours while the sinusoidal temperature profile (the same as shown in Figure 5.3) was imposed on the prototype transformer. During measurement, sinusoidal temperature variation can significantly influence the dielectric response of cellulose due to temperature dependent moisture distribution and charge carriers, especially for low frequencies because it takes a relatively longer time to take measurement at low frequencies. The proposed method does not considered the effects of dynamic temperature on moisture and charge carriers separately. By contrast, it takes account the effect of temperature on dielectric response (caused by moisture distribution and charge carriers) simultaneously.

The dependency between permittivity and temperature [11] can be expressed as (5.18)

$$\varepsilon'_T = \varepsilon_{r\infty} + \frac{\varepsilon_{rs} + \varepsilon_{r\infty}}{1 + \omega^2 \tau_h{}^2 e^{2H/kT}}, \quad \varepsilon''_T = \frac{(\varepsilon_{rs} - \varepsilon_{r\infty})\omega\tau_h e^{H/kT}}{1 + \omega^2 \tau_h^2 e^{2H/kT}} \tag{5.18}$$

where ε_{rs} is the dielectric constant at zero excitation frequency or dc value, $\varepsilon_{r\infty}$ is the dielectric constant at very high frequency, τ_h is a pre exponential factor, H is the activation energy, T is the absolute temperature and k is the Boltzmann constant.

As the dependency between dielectric response and diffusion positions in Section 5.6 is obtained

under a constant temperature of 50℃, necessary temperature corrections are needed to compensate the influence of the temperature on the complex permittivity in dielectric response modelling when sinusoidal temperature is applied. The temperature correction is presented in Figure 5.12.

In Figure 5.12, for each round of dielectric response measurements, the average temperature at a particular frequency is corrected by $T_{ave} = (1/P_i)\int_0^{P_i} T dt$, where P_i is the total measurement duration at the *i*-th frequency and T is the temperature at different time instances as depicted in Figure 5.4 (a). By substituting T_{ave} into T of (5.18), the estimated real and imaginary part of permittivity at each frequency under sinusoidal temperature can be corrected. The estimated and measured real and imaginary parts of permittivity of cellulose insulation after temperature correction under non-equilibrium and non-uniform moisture distribution at different times are drawn in Figure 5.13.

From Figure 5.13 it can be observed there is a similar trend in the estimated and measured dielectric responses at different time instances. Some degree of discrepancy between the estimated and measured permittivities can be observed: (1) at low frequencies, which may be due to the modeling errors can be observed in Figure 5.9 and Figure 5.11; and (2) at time instance t = 6 hours, which may be due to the lack of data for high moisture content in modeling (Figure 5.8 and Figure 5.10).

From Figure 5.13 it can also be seen that there is a decreasing trend in both estimated and measured real permittivities at different time instances (i.e. different moisture distribution in the cellulose insulation). When the moisture concentration increases, the real permittivity also increases at each frequency. The imaginary permittivity increases throughout the whole frequency range when more moisture moves into cellulose insulation. However, the increase of the imaginary part of permittivity at low frequencies is more dominant.

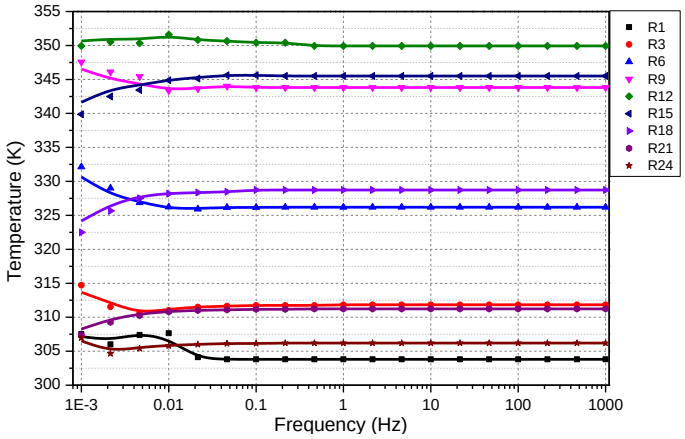


Figure 5.12. Temperature correction (R_j denotes the *j*-th round of dielectric response measurement, data points denote the maximum temperature at each frequency, solid lines denote the corrected temperature at each frequency)

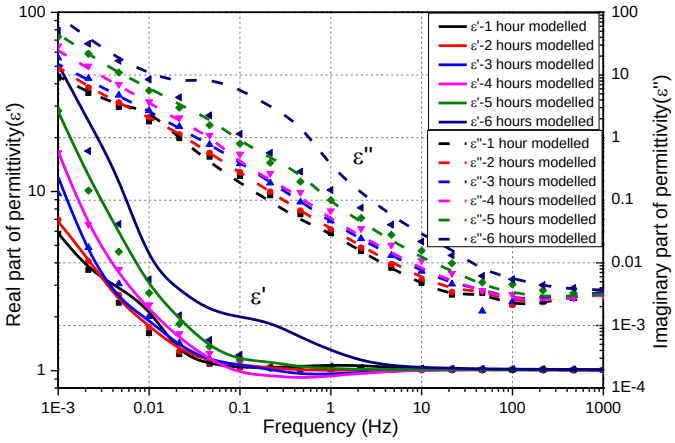


Figure 5.13. Comparison between measured and estimated dielectric response (temperature varies according to Figure 5.3, from T = 30 ℃ to T = 80 ℃. Lines denote the modelled data and dots denote the measured data)

5.8 Discussion On The Applicability Of Modelling Method To Field Transformer

Other than KFT and oil-paper moisture equilibrium chart, the methods for estimating moisture in field transformers can be categorized into two major groups. The first group is based on online/offline measurements including dielectric response measurement and cellulose isotherm method. Dielectric response measurement can provide an estimation of average moisture content in transformer's cellulose insulation. However, it requires removing the transformer from service during the measurement. The cellulose isotherm method derives the moisture content in cellulose from the measurement of moisture-in-oil sensor. However, given the complexity of a transformer's construction and the use of different types of cellulose at different layers and locations in the transformer, a sufficient number of moisture-in-oil sensors need to be installed before an accurate estimation of moisture in cellulose can be achieved.

The second group of the estimation of moisture is through numeric modelling approaches such as multi-physics FEM and particle tracing method. The modelling approach requires the geometric information of transformer's insulation. If the insulation geometry of the transformer is known, modelling approach can be applied to estimate the moisture distribution in the cellulose insulation of the transformer. If the detailed geometry information of the transformers is not available, some approximation needs to be made. It is assumed that the transformer's insulation construction follows

common engineering practice and meets certain criteria (standards) to withstand electromagnetic, thermal, and mechanical stresses during the transformer's operation. Thus, by using the nameplate information of the transformer (e.g. rating, core type, temperature rise, oil volume etc), it is possible to deduce the insulation geometry of the transformer. Based on the approximated geometry, the modelling approach can then be applied for moisture estimation.

To improve the applicability of the proposed method, a software tool is implemented, which can be used by practicing engineers as an alternative to current dielectric response measurement analysis software available with the commercial products. Figure 5.14 illustrates the major components of this software tool, including:

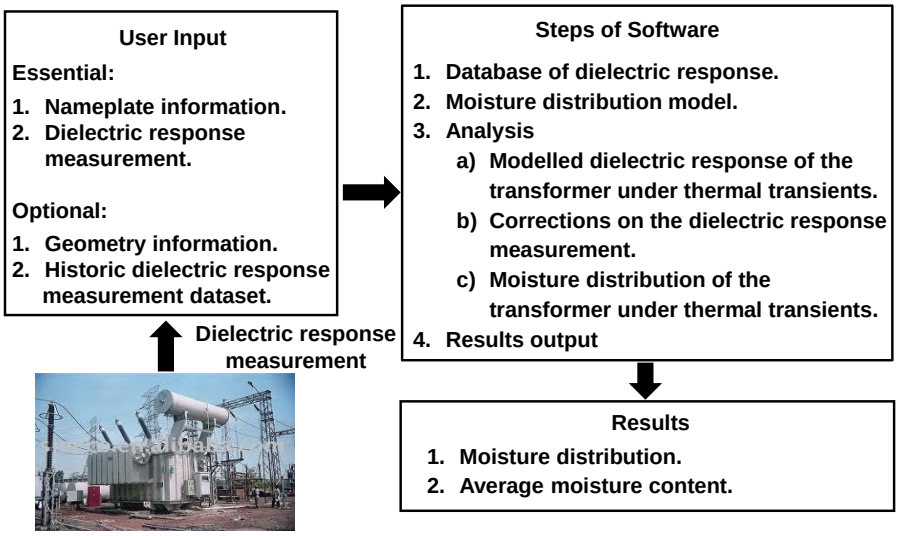


Figure 5.14. Components of software tool for moisture estimation

(1) Database of dielectric response of transformers: This database consists of dielectric response measurement from various transformers of different manufacturers. Users can also expand the database by including dielectric response measurement results of their transformers. If the transformer of interest is not included in the database and the user doesn't have any historic measurement dataset of this transformer, the datasets of other transformers with similar rating, insulation geometry and age will be utilized by the software in its calculation.

(2) Moisture distribution model of transformers: The software tool consists of moisture distribution models of many transformers, which are developed from transformers' geometrical information of different manufacturers. Multi-physics modelling and particle tracing method are integrated into the software to determine moisture distribution of oil-cellulose insulation system of a transformer. For the transformer of which the geometry information is not available, the software will approximate its

geometry by using the available information of the transformer (e.g. rating, core type, temperature rise, oil volume etc.). It is assumed that the moisture dissolved in oil is measured (by either water-in-oil probes or DGA/oil test of the transformer). Based on the above approximated transformer geometry and the measured moisture in oil, the software will produce a moisture distribution model for the transformer of interest.

(3) Analysis: Users conduct dielectric response measurement on the transformer of interest and input the dielectric response data to the software. The software will automatically perform the following calculations:

(a) Combine the above database and moisture distribution model to compute the dielectric response of the transformer considering the effect of temperature variations.

(b) Apply the computed dielectric response in (a) to make corrections on the measured dielectric response of the transformer. Such correction can be done by adjusting the parameters (α and β in Equations 5.16 and 5.17) in the model until the error between computed and measured dielectric response is minimized.

(c) Compute the moisture distribution inside the transformer's cellulose insulation under thermal transients. The computed moisture distribution will be combined with the results from multi-physics modelling and particle tracing method to provide an accurate moisture estimation on moisture concentration of transformer's cellulose insulation.

(4) Output: The software will provide the following information.

(a) Moisture distribution of transformer's cellulose insulation.

(b) Average moisture content in transformer's cellulose insulation.

With the advancement of sensor technology, more sensors can be installed at critical positions for continuously monitoring the temperature and moisture in transformer oil. By integrating the online sensor based monitoring, modeling approach, and dielectric response measurement, the reliability of estimation of moisture and its distribution in transformer insulation system can be improved by the proposed method described in this paper.

5.9 Conclusion

This paper developed a distributed parameter model to interpret the dielectric response of cellulose insulation of transformers under thermal and moisture transients. By using the methodologies

proposed in this paper, the parameters of dielectric response (dielectric losses and permittivity) of a transformers' cellulose insulation under thermal transients and moisture non-equilibrium condition can be estimated. It is expected that the proposed methodology can provide proper interpretation of dielectric response measurement and accurate moisture estimation for transformers at substations.

Acknowledgment

The authors gratefully acknowledge Australian Research Council, and industry partners Powerlink Queensland, Energex, Ergon Energy, and TransGrid for supporting this work.

References

[1] L. E. Lundgaard, W. Hansen, D. Linhjell and T. J. Painter, "Aging of oil-impregnated paper in power transformers," *IEEE Trans. Power Delivery*, vol.19, Issue 1, pp. 230-239, 2004.

[2] Pahlavanpour, M. Martins and Eklund, "Study of moisture equilibrium in oil-paper system with temperature variation," in *Proceedings of International Conference on Properties and Applications of Dielectric Materials, June 1-5*, 2003,Nagoya, Japan, pp. 1124-1129.

[3] T. V. Oommen, "Moisture equilibrium charts for transformer insulation drying practice," *IEEE Transactions on Power Apparatus and Systems*, vol.103, Issue 10, pp. 3062-3067, 1984.

[4] Y. Du, M. Zahn, B. C. Lesieutre, A. V. Mamishev and S. R. Lindgren, "Moisture equilibrium in transformer paper-oil systems," *IEEE Electr. Insul. Mag.*, vol.15, Issue 1, pp. 11-20, 1999.

[5] B. Buerschaper, O. Kleboth-Lugova and T. Leibfried, "The electrical strength of transformer oil in a transformerboard-oil system during moisture non-equilibrium," in *Proceedings of Annual Conference on Electrical Insulation and Dielectric Phenomena, October 19-22*, 2003,Albuquerque, NM, United states, pp. 269-272.

[6] C. Ekanayake, S. M. Gubanski, A. Graczkowski and K. Walczak, "Frequency response of oil impregnated pressboard and paper samples for estimating moisture in transformer insulation," *IEEE Trans. Power Delivery*, vol.21, Issue 3, pp. 1309-1317, 2006.

[7] A. F. Howe, "Diffusion of moisture through power-transformer insulation," *Proc. Inst. Electr. Eng.*, vol.125, Issue 10, pp. 978-986, 1978.

[8] B. Garcia, J. C. Burgos, A. M. Alonso and J. Sanz, "A moisture-in-oil model for power transformer monitoring - Part I: Theoretical foundation," *IEEE Trans. Power Delivery*, vol.20, Issue 2, pp. 1417-1422, 2005.

[9] W. A. Fessler, T. O. Rouse, W. J. McNutt and O. R. Compton, "A refined mathematical model for prediction of bubble evolution in transformers," *IEEE Trans. Power Delivery*, vol.4, Issue 1, pp. 391-404, 1989.

[10] D. Martin, C. Perkasa and N. Lelekakis, "Measuring paper water content of transformers: a new approach using cellulose isotherms in nonequilibrium conditions," *IEEE Trans. Power Delivery*, vol.28, Issue 3, pp. 1433-1439, 2013.

[11] A. K. Jonscher, *Dielectric Relaxation in Solids*, London: Chelsea Dielectrics Press, 1983.

[12] V. Oyj,"MMT330 Moisture and Temperature Transmitter Series for Oil" available at http://www.vaisala.com/en/industrialmeasurements/products/moistureinoil/Pages/MMT330.aspx.

[13] M. A. Tsili, E. I. Amoiralis, A. G. Kladas and A. T. Souflaris, "Power transformer thermal analysis by using an advanced coupled 3D heat transfer and fluid flow FEM model," *Int. J. Therm. Sci.*, vol.53, Issue 53, pp. 188-201, 2012.

Chapter 6

Moisture Dependent Thermal Modelling of Transformers Filled With Vegetable Oil

Contribution of the Chapter

This chapter introduces a moisture dependent thermal model (MDTM) for estimating transformer hot spot temperature. In this model, nonlinear thermal resistance is formulated by considering both oil and cellulose (paper and pressboard) of the transformer. Especially, the effect of moisture concentration and hot spot temperature on the thermal resistance of cellulose is taken into account. The proposed MDTM is verified by using historical data of moisture-in-oil and temperature measurements on an in-service vegetable oil-filled transformer. Comparisons between the proposed MDTM and a number of existing thermal models are performed on the basis of adequacy and accuracy metrics.

Moisture Dependent Thermal Modelling of Transformers Filled With Vegetable Oil

Yi Cui, Hui Ma, Tapan Saha, Chandima Ekanayake and Daniel Martin

The University of Queensland, Brisbane, Australia

Publication Journal: IEEE Transactions on Power Delivery

Submitted: July 9, 2015

Author's Contributions:

Yi Cui: Primary model design, theoretical simulation, experimental validation, measurement results analysis and manuscript preparation.

Hui Ma: Supervision on the experimental measurement, results discussion and manuscript revision.

Tapan Saha: Supervision on the experimental measurement, results discussion and manuscript revision.

Chandima Ekanayake: Initial concept design, supervision on the experimental measurement and results discussion.

Daniel Martin: Supervision on the experimental measurement and results discussion.

Abstract

Hot spot temperature (HST) is recognized as one of key factors that affect the life expectancy of a transformer. In this study, a moisture dependent thermal model (MDTM) for estimating transformer hot spot temperature is proposed. In this model, nonlinear thermal resistance is formulated by considering both oil and cellulose (paper and pressboard) of the transformer. Especially, the effect of moisture concentration and hot spot temperature on the thermal resistance of cellulose is taken into account. The proposed MDTM is verified by using historical data of moisture-in-oil and temperature measurements on an in-service vegetable oil-filled transformer. Comparisons between the proposed MDTM and a number of existing thermal models are performed on the basis of adequacy and accuracy metrics.

Index Terms: hot spot temperature, moisture, top oil temperature, transformer

6.1 Introduction

Transformer hot spot temperature (HST) is regarded to be one of the most significant parameters in determining its load capability [1-2]. On the other hand, cellulose close to the hot spot in a transformer would experience thermal degradation and inevitably become weak links. This can eventually cause significant damage on the transformer's cellulose (paper and pressboard) and shorten its lifespan [3]. Therefore, it becomes increasingly important to determine transformer hot spot temperature under variable loading conditions and different ambient temperature.

Transformer hot spot temperature can be measured by fiber optic sensors, which are attached on the outer surface of the cellulose wrapped around winding conductors [4-5]. However, such direct temperature measurements are at an additional cost to utilities and an agreement with transformer manufacturers may need to be made in advance. For those in-service transformers without optic sensors, it might not be feasible to make installation and upgrade SCADA (Supervisory Control and Data Acquisition) system accordingly. Therefore, thermal dynamic models have been developed to estimate transformer's hot spot temperature [1-9].

In thermal dynamic models, a commonly adopted approach is based on an electrical-thermal equivalent circuit originated from heat transfer theory [9]. In this approach, the thermal behavior of a transformer was described by electrical thermal analogy and the transformer's nonlinear thermal

resistance was formulated by considering oil convention inside oil ducts. To obtain the model parameters in this approach, a "heat run test" on the transformer before commission is required [10].

Existing thermal dynamic models only consider thermal resistance of oil in transformer hot spot calculation. However, transformer cellulose's thermal properties are influenced by both temperature and moisture [11-13]. When transformer cellulose is subjected to high moisture and temperature, its thermal conductivity may increase by 50% and cannot be ignored in thermal modelling [11]. Therefore, it is necessary to take the thermal resistance of cellulose into account in determining transformer hot spot temperature, especially when transformer is under moisture dynamics and temperature transients.

This paper is aimed at deriving a moisture dependent thermal model (MDTM), which takes into account the effect of moisture concentration in cellulose on transformer hot spot temperature calculation. Compared to existing thermal models, it is expected that the proposed MDTM can consistently attain high accuracy in hot spot temperature estimation when the transformer is under load variation and moisture dynamics. The validation of the proposed model is performed on a vegetable-oil filled transformer. The statistical metrics are calculated to evaluate differences between hot spot temperatures estimated by different thermal models and those measured by sensors. The results demonstrate the acceptability of the proposed MDTM over several existing models such as IEEE model, Swift's model and Susa's model.

6.2 Brief Review on Thermal Modelling and Moisture Dynamics of Transformer

6.2.1 Transformer thermal modelling

This session provides a brief review on three existing thermal models, which are widely adopted in estimating transformer hot spot temperature.

(1) IEEE model [1]

An increase in the winding current of a transformer will result in temperature rise in oil and winding. The top oil temperature rise can be computed as an exponential response from the initial top oil temperature rise ($\Delta\theta_{oil,I}$) to the ultimate temperature rise ($\Delta\theta_{oil,U}$) as (6.1) and (6.2).

$$\tau_{oil,R}\frac{d\Delta\theta_{oil}}{dt} = \Delta\theta_{oil,U} - \Delta\theta_{oil,I} \tag{6.1}$$

$$\Delta\theta_{oil,U} = \Delta\theta_{oil,R} \cdot \left(\frac{1 + R \cdot K^2}{1 + R} \right) \tag{6.2}$$

The winding hot spot temperature rise can be computed as an exponential response from the initial hot spot temperature rise ($\Delta\theta_{hs,I}$) to the ultimate temperature rise ($\Delta\theta_{hs,U}$) as (6.3) and (6.4).

$$\tau_{wnd,R} \frac{d\Delta\theta_{hs}}{dt} = \Delta\theta_{hs,U} - \Delta\theta_{hs,I} \tag{6.3}$$

$$\Delta\theta_{hs,U} = \Delta\theta_{hs,R} \cdot \left(K^{2m} \right) \tag{6.4}$$

Finally hot spot temperature is calculated as the summation of ambient temperature θ_{amb}, top oil temperature rise $\Delta\theta_{oil}$ and hot spot temperature rise $\Delta\theta_{hs}$ as (6.5).

$$\theta_{hs} = \theta_{amb} + \Delta\theta_{oil} + \Delta\theta_{hs} \tag{6.5}$$

However, the above model does not consider the effect of ambient temperature variations on top oil temperature. Lesieutre *et al.* [7] updated the model (6.1) as (6.6) .

$$\tau_{oil,R} \frac{d\Delta\theta_{oil}}{dt} = \Delta\theta_{oil,U} + \theta_{amb} - \Delta\theta_{oil} \tag{6.6}$$

Accordingly, the hot spot temperature is updated as (6.7).

$$\theta_{hs} = \theta_{oil} + \Delta\theta_{hs} \tag{6.7}$$

(2) Swift's model

Swift *et al.* adopted an equivalent circuit in transformer thermal modelling [9, 14] based on heat transfer theory. In Swift's model, hot spot temperature is computed by using nonlinear thermal resistance. The differential equation for top oil temperature can be expressed as (6.8).

$$\frac{K^2 R + 1}{R + 1} \cdot \Delta\theta_{oil,R}^{\frac{1}{n}} = \tau_{oil,R} \frac{d\theta_{oil}}{dt} + \left(\theta_{oil} - \theta_{amb} \right)^{\frac{1}{n}} \tag{6.8}$$

The differential equation of hot spot temperature can be written as (6.9).

$$K^2 \cdot \Delta\theta_{hs,R}^{\frac{1}{m}} = \tau_{wnd,R} \frac{d\theta_{hs}}{dt} + \left(\theta_{hs} - \theta_{oil} \right)^{\frac{1}{m}} \tag{6.9}$$

(3) Susa's model

Susa *et al.* extended Swift's model by considering the non-linear thermal resistance of mineral oil, which is caused by temperature dependent oil viscosity and loss variation [8]. In Susa's approach,

top oil temperature is calculated as (6.10).

$$\frac{K^2R+1}{R+1}\cdot\Delta\theta_{oil,R}=\tau_{oil,R}\frac{d\theta_{oil}}{dt}+\frac{\left(\theta_{oil}-\theta_{amb}\right)^{1+n}}{\Delta\theta_{oil,R}^{n}\cdot\mu_{pu}^{n}} \tag{6.10}$$

The hot spot temperature is computed as (6.11).

$$K^2\cdot P_{Cu,pu}\cdot\Delta\theta_{hs,R}=\tau_{wnd,R}\frac{d\theta_{hs}}{dt}+\frac{\left(\theta_{hs}-\theta_{oil}\right)^{1+m}}{\Delta\theta_{hs,R}^{m}\cdot\mu_{pu}^{m}} \tag{6.11}$$

The above three models only consider the oil thermal properties by treating thermal transfer process in transformers as oil convection. None of these models consider cellulose thermal properties in determining hot spot temperature. In this paper moisture dependent thermal resistance of cellulose will be integrated into the updated thermal model. A brief review on moisture in transformer and moisture measurement is provided in the next section.

6.2.2 Moisture dynamics in transformer

Due to the low moisture affinity of insulating oil, in an in-service transformer moisture mainly resides in its cellulose. Moisture can exist in cellulose in three states: adsorbed to the cellulose surface, as free water attached on capillaries, or as imbibed free water [15].

Moisture in oil can be directly measured through moisture-in-oil sensor and then moisture in cellulose can be estimated through cellulose adsorption isotherms approach [15]. This approach assumes an equilibrium state of moisture inside a transformer, which requires a constant temperature to be maintained in the transformer over a considerable long time. However, this is almost impossible for an operational transformer since the transformer is normally subjected to continuous fluctuations of loading, temperature and other conditions. Moisture in cellulose can be also estimated by formulating it as a mass transfer process, which can be solved by using Fick's law [16].

With the advancement of sensor technology, the latest moisture in oil sensor can measure the vapor pressure of moisture dissolved in oil and temperature. The measurement results can then be used to determine the moisture in cellulose through cellulose isotherms approach [17-18]. This approach is adopted in this paper.

6.3 Moisture Dependent Thermal Model (MDTM)

6.3.1 Heat transfer in transformer

In this paper, thermal modelling is performed on a 50 MVA vegetable oil-filled transformer. Figure 6.1 shows the construction of this transformer [19]. The transformer used FR3 vegetable oil and is cooled by the forced oil circulation with directed flow. Two external liquid-to-water heat exchangers were also installed for circulating water over the transformer's cooling surface [20]. The detailed parameters of the transformer will be provided in Section 6.4.

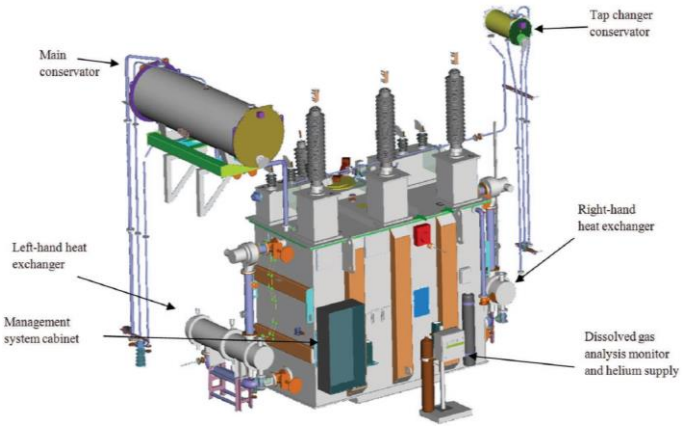


Figure 6.1 Vegetable oil filled transformer under investigation [19]

During transformer's operation, certain losses (copper loss, iron core loss and stray loss) are generated and lead to a temperature rise in core and windings. Since core and windings are mainly made of metallic materials with high thermal conductivity, their temperature can increase instantly once transformer is just energized or load current is increased [21]. Most heat is transferred to the surface of the core and winding conductors through thermal conduction. This results in temperature difference between conductors' surface and their surroundings (cellulose and oil). The heat is subsequently transferred by oil convection and thermal conduction of cellulose. Eventually, the heat is dissipated to the ambient mainly through external coolers. A small portion of heat may also be dissipated through the tank wall.

6.3.2 Moisture dependent thermal resistance of cellulose

For a core type transformer the typical insulation structure between HV and LV windings is considered as a number of cylindrical shells of pressboard barriers, separated by axial spacers. The

heat generated by transformer windings will be transferred through cellulose (paper) wrapped around conductors and cellulose (pressboards) laid between high voltage and low voltage windings. The heat flow paths of cross section of core type transformer winding are shown in Figure 6.2.

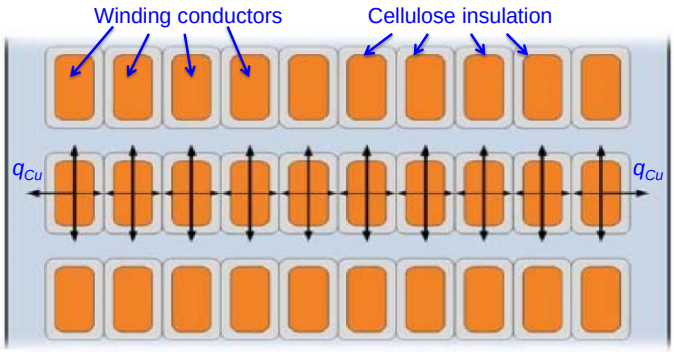


Figure 6.2 Heat flow paths of cross section of core-type transformer winding

Based on heat transfer theory, the thermal resistance of transformer cellulose can be expressed as (6.12) [22].

$$R_{insul} = l / \left(k_{insul} \cdot A_{insul} \right) \tag{6.12}$$

where A_{insul} denotes the cellulose area, l denotes the thickness of cellulose and k_{insul} denotes the thermal conductivity of cellulose.

Lopatkiewicz *et al.* measured the thermal conductivity of oil-impregnated cellulose paper with different moisture concentrations and temperatures [11-13]. They reported that if the radiation heat transfer is ignored, a linear correlation between thermal conductivity of cellulose and moisture can be found as in (6.13)-(6.15).

$$k_{insul} = 0.176 \times \left[1 + 3.6 \times 10^{-3} \cdot \Delta w + 7.3 \times 10^{-4} \cdot \Delta \theta_{insul} \right] \tag{6.13}$$

$$\Delta w = w_{insul} - w_0 \tag{6.14}$$

$$\Delta \theta_{insul} = \left(\theta_{insul} - \theta_0 \right) \tag{6.15}$$

where $w_0 = 1.03$, $\theta_0 = 60°C$, denoting the reference moisture and temperature.

To determine the moisture concentration of cellulose, online moisture-in-oil measurement is performed on the vegetable oil-filled transformer shown in Figure 6.1. Moisture concentration in cellulose is derived by cellulose isotherms method as in (6.16) and (17) [17-18].

$$w_{insul} = 2.173\times10^{-7}\cdot P_v^{\,0.6685}\cdot e^{\frac{4275.6}{\theta_{insul}}} \tag{6.16}$$

$$P_v = 6.1121\times e^{\frac{17.502\times(\theta_{insul}+273)}{240.97+\theta_{insul}+273}}\cdot A_w\times9.869\times10^{-4} \tag{6.17}$$

where P_v denotes vapor pressure, which is dependent on cellulose temperature θ_{insul} and water activity A_w in oil. The water activity A_w is the relative ratio of the partial pressure of moisture in cellulose and the saturation vapor pressure of pure moisture under the same temperature [23]. It can be directly measured from moisture-in-oil sensor.

Due to the non-uniform temperature distribution of cellulose in transformer, average cellulose temperature θ_{insul} is calculated as average of hot spot temperature (θ_{hs}) and top oil temperature (θ_{oil}) as (6.18).

$$\theta_{insul} = \left(\theta_{hs} + \theta_{oil}\right)/2 \tag{6.18}$$

6.3.3 Top oil thermal model

The above moisture dependent thermal resistance of cellulose is incorporated into the top oil thermal equivalent circuit, which was adopted in Swift's and Susa's models [8-9].

The thermal dynamic model for top oil temperature is shown in Figure 6.3. Two current sources are used to represent the internal heat sources of no load loss (q_{Fe}) and load loss of transformer (q_{Cu}) [9]. A voltage source (θ_{amb}) is used to model the ambient temperature. Different from the Swift's and Susa's models, in Figure 6.3 the nonlinear thermal resistance also consists of thermal conduction resistance of cellulose (R_{insul}) other than only the convection resistance of transformer oil (R_{oil}). In Figure 6.3, C_{insul} and C_{oil} denote the lumped thermal capacitance of cellulose and oil.

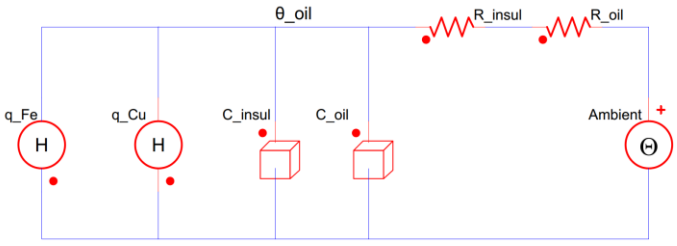


Figure 6.3 Dynamic model of top oil temperature, R_{insul}, C_{insul} denote the thermal conduction resistance and capacitance of cellulose.

In Figure 6.3, the top oil temperature is calculated by a differential equation in (6.19).

$$q_{Fe} + q_{Cu} = \left(C_{insul} + C_{oil}\right)\frac{d\theta_{oil}}{dt} + \frac{\theta_{oil} - \theta_{amb}}{R_{insul} + R_{oil}} \tag{6.19}$$

The thermal resistance of oil convection can be expressed by (6.20) [22].

$$R_{oil} = 1/\left(h \cdot A_{oil}\right) \tag{6.20}$$

where h denotes convection heat transfer coefficient, A denotes the area.

The convection heat transfer coefficient of oil can be further simplified as (6.21) according to fluid dynamics theory [24].

$$h = C_1 \cdot \theta_{oil}^n / \mu^n \tag{6.21}$$

where μ denotes oil viscosity, n is a constant, denoting the nonlinearity of oil thermal resistance, C_1 is a constant which has the same form as in [8].

The constant *n* represents nonlinearity in top oil thermal model for different cooling types and its typical values are given in Table 6.1[25].

Table 6.1 Constant *N* representing nonlinearity in top-oil thermal model

n	**With coolers**		**Without coolers**
	OFAF /ONAF	ONAN	ONAN
Cold start	0.5	0	0.25
Transformer energised	0.2	0.25	0.25

The viscosity of vegetable oil is about five times higher than that of mineral oil at the same temperature [26]. The dependency between vegetable oil viscosity and temperature can be modelled as (6.22) [8].

$$\mu = 7.29 \times 10^{-5} \cdot e^{\frac{4030.5}{\theta_{oil}+273}} \tag{6.22}$$

By substituting (6.12) and (6.20) into (6.19), (6.19) can be rewritten as (6.23).

$$\left(q_{Fe} + q_{Cu}\right) = \cdot\left(C_{oil} + C_{insul}\right) \cdot \frac{d\theta_{oil}}{dt} + \frac{\theta_{oil} - \theta_{amb}}{\left[\frac{1}{AC_1}\frac{\mu^n}{\Delta\theta_{oil}} + \frac{l}{AC_2}\frac{1}{0.9963 + \alpha \cdot e^{\frac{4275.6}{\theta_{insul}+273}} + \beta \cdot \left(\theta_{insul} - 60\right)}\right]} \tag{6.23}$$

where $C_2 = 0.176$, α=$7.8228\times10^{-10}\cdot P_v^{0.6685}$, β=7.3×10^{-4} .

With the following definitions

$$R_{oil,R} = \mu_R^n / \left(A \cdot C_1 \cdot \Delta\theta_{oil,R}^n \right) \tag{6.24}$$

$$\mu_{pu}^n = \mu^n / \mu_R^n = e^{\frac{4030.5}{\theta_{oil}+273}} / e^{\frac{4030.5}{\theta_{oil,R}+273}} \tag{6.25}$$

$$R_{insul,R} = \frac{l}{AC_2} \cdot \frac{1}{0.9963 + \alpha \cdot e^{\frac{4275.6}{\frac{\theta_{oil,R}+\theta_{hs,R}}{2}+273}} + \beta \cdot \left(\frac{\theta_{oil,R}+\theta_{hs,R}}{2} - 60 \right)} \tag{6.26}$$

$$k_{pu} = \frac{R_{insul}}{R_{insul,R}} = \frac{0.9963 + \alpha \cdot e^{\frac{4275.6}{\theta_{insul,R}+273}} + \beta \cdot \left(\theta_{insul,R} - 60 \right)}{0.9963 + \alpha \cdot e^{\frac{4275.6}{\theta_{insul}+273}} + \beta \cdot \left(\theta_{insul} - 60 \right)} \tag{6.27}$$

$$\tau_{oil,R} = \left(R_{oil,R} + R_{insul,R} \right) \cdot \left(C_{oil} + C_{insul} \right) \tag{6.28}$$

$$\Delta\theta_{oil,R} = \left(q_{Fe,R} + q_{Cu,R} \right) \left(R_{oil,\mathrm{R}} + R_{insul,R} \right) \tag{6.29}$$

$$R = q_{Fe} / q_{Cu} \tag{6.30}$$

$$K = I / I_R \tag{6.31}$$

$$\lambda_R = R_{insul,R} / R_{oil,R} \tag{6.32}$$

Eq (6.23) can be rewritten as (6.33), which is the dynamic thermal equation for calculating top oil temperature.

$$\frac{K^2 R + 1}{R + 1} \cdot \Delta\theta_{oil,R} = \tau_{oil,R} \frac{d\theta_{oil}}{dt} + \frac{\left(\theta_{oil} - \theta_{amb} \right)^{1+n}}{\Delta\theta_{oil,R}^n} \cdot \frac{1 + \lambda_R}{\mu_{pu}^n + \lambda_R \cdot k_{pu} \cdot \left(\frac{\Delta\theta_{oil}}{\Delta\theta_{oil,R}} \right)^n} \tag{6.33}$$

In the above derivations, λ_R denotes the ratio between thermal resistance of cellulose conduction and oil convection. If λ_R=0, i.e. the thermal resistance of cellulose conduction is ignored, top oil temperature equation (6.33) is the same as (6.10) in Susa's model. When cellulose insulation degrades, the water content of the cellulose may increase compared with that of new cellulose at the same temperature. This will result in a decrease in the value of λ_R.

The value of λ_R can be obtained by comparing the top oil temperature computed from (6.33) with

that from measurements. Once the ratio λ_R is determined in top oil thermal modelling, it will be used for calculating hot spot temperature in the next section.

6.3.4 Hot spot thermal model

The moisture dependent thermal resistance of cellulose is also incorporated into hot spot thermal equivalent circuit, which was adopted in Swift's and Susa's models [8-9]. The thermal dynamic model for hot spot temperature is shown in Figure 6.4. Different from the Swift's and Susa's models, in Figure 6.4 the nonlinear thermal resistance consists of thermal conduction resistance of cellulose (R_{insul}).

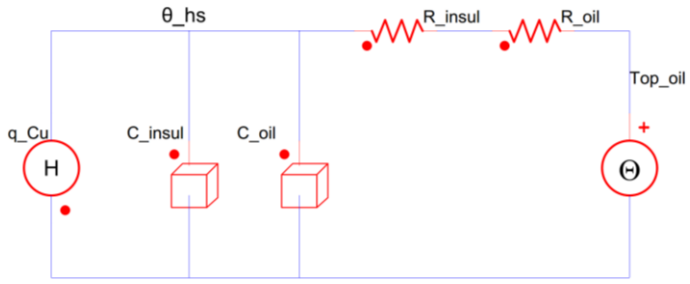


Figure 6.4 Dynamic model of hot spot temperature, R_{insul}, C_{insul} denote the thermal conduction resistance and capacitance of cellulose.

The differential equation for modelling hot spot temperature can be expressed as (6.34).

$$q_{Cu} = \left(C_{oil} + C_{insul}\right)\frac{d\theta_{hs}}{dt} + \frac{\theta_{hs} - \theta_{oil}}{R_{insul} + R_{oil}} \tag{6.34}$$

The derivations of thermal resistance of oil convection and cellulose conduction for modelling hot spot temperature are similar to those in modelling top oil temperature [22].

$$R_{oil} = \mu^m / \left(A \cdot C_1 \cdot \Delta\theta_{hs}^m\right) \tag{6.35}$$

The typical values of constant m representing nonlinearity of the above hot spot model for different cooling types are given in Table 6.2[25].

Table 6.2 Constant *M* representing nonlinearity in hot spot thermal model

M	**With coolers**	**Without coolers**
	OFAF/ONAF/ONAN	ONAN
Cold start	2	0.25
Transformer energised	0.2	0.25

By substituting (6.12) and (6.35) into (6.34), (6.34) can be rewritten as (6.36).

$$q_{Cu} = \cdot\left(C_{oil} + C_{insul}\right) \cdot \frac{d\theta_{hs}}{dt} + \frac{\theta_{hs} - \theta_{oil}}{\left[\frac{1}{AC_1}\frac{\mu^m}{\Delta\theta_{hs}} + \frac{l}{AC_2}\frac{1}{0.9963 + \alpha \cdot e^{\frac{4275.6}{\theta_{insul}+273}} + \beta \cdot \left(\theta_{insul} - 60\right)}\right]} \tag{6.36}$$

With the following definitions

$$R_{oil,R} = \mu_R^m / \left(A \cdot C_1 \cdot \Delta\theta_{hs,R}^m\right) \tag{6.37}$$

$$\mu_{pu}^m = \mu^m / \mu_R^m \tag{6.38}$$

$$\tau_{\mathrm{wnd},R} = \left(R_{oil,R} + R_{insul,R}\right) \cdot \left(C_{oil} + C_{insul}\right) \tag{6.39}$$

$$P_{Cu,pu} = P_{Cu,DC,pu} \cdot \frac{235 + \theta_{hs}}{235 + \theta_{hs,R}} + P_{Cu,Eddy,pu} \cdot \frac{235 + \theta_{hs,R}}{235 + \theta_{hs}} \tag{6.40}$$

In MDTM hot spot temperature can be computed as (6.41).

$$K^2 \cdot P_{Cu,pu} \cdot \Delta\theta_{hs,R} = \tau_{wnd,R} \frac{d\theta_{hs}}{dt} + \frac{\left(\theta_{hs} - \theta_{oil}\right)^{1+m}}{\Delta\theta_{hs,R}^m} \cdot \frac{1 + \lambda_R}{\mu_{pu}^m + \lambda_R \cdot k_{pu} \cdot \left(\frac{\Delta\theta_{hs}}{\Delta\theta_{hs,\mathrm{R}}}\right)^m} \tag{6.41}$$

6.4 Results and Discussions

This section verifies the MDTM developed in Section 6.3 by comparing the calculated hot spot temperature with the measured value of a vegetable oil-filled transformer.

6.4.1 Transformer configuration

Table 6.3 summarizes the parameters for thermal modelling on the vegetable oil-filled transformer. An online moisture-in-oil monitoring system was installed on the transformer. It consists of a number of capacitive moisture-in-oil sensors (MMT330 manufactured by Vaisala [27]) placed at different locations inside the transformer. Temperatures and moisture concentrations of tank's top and bottom oil, cooler's top and bottom oil and ambient can be obtained in every one minute. The hot spot temperature of the transformer is also recorded by eight fibre optic probes, which are placed at different locations of inter-disc insulation of the transformer.

Table 6.3 Parameters for modelling hot spot temperature of transformer

Rated power (MVA)	50	**$\tau_{wnd,R}$ (min)**	6.5
Rated voltage (kV)	132/11/11*	**$\tau_{oil,R}$ (min)**	90
Average oil temperature rise (℃)	24.2	**$\Delta\theta_{hs,R}$ (℃)**	45.4
Winding gradient (℃)	14.4	**$\Delta\theta_{oil,R}$ (℃)**	24.2
Winding temperature rise (℃)	38.6	**$\Delta\theta_{amb}$ (℃)**	16.5
Hot spot gradient (℃)	18.8	**Tank oil volume (liter)**	21850
Cooling type	KDWF	**Cooler oil volume (liter)**	185

* Slash (/) denotes the parameters for HV/LV windings respectively.

6.4.2 Calculation of moisture in cellulose

As discussed in Section 6.3.2, moisture in cellulose is computed through moisture-in-oil measurement. Figure 6.5 presents one week's moisture concentration in transformer's cellulose when the transformer was under heavy load condition. In Figure 6.5a, the temperature of cellulose was calculated from thermal modelling (6.18). In Figure 6.5b, the water activity was directly obtained from a moisture-in-oil sensor located at the top of the tank. The moisture variation in cellulose (Figure 6.5c) is calculated from cellulose isotherms method (6.16).

It can be seen that the average temperature of cellulose varies periodically within the range from 40℃ to 58℃ in one week (Figure 6.5a). The measured water activity in oil also cycles in the inverse proportion of the temperature (Figure 6.5b). The computed moisture concentration of cellulose is highly dependent on temperature and water activity, which is shown as dash line in black in Figure 6.5c.

In Figure 6.5c, the daily moisture concentration of cellulose is calculated by averaging the maximum and minimum moisture during one day interval. So the average moisture concentration of

cellulose for the whole week is depicted as nine red squares (at 0, 12, 36, 60, 84, 108, 132, 156 and 168 hours) in Figure 6.5c. It can be observed from Figure 6.5c that most moisture concentration in cellulose is less than 1.5% which implies the transformer cellulose is relatively dry.

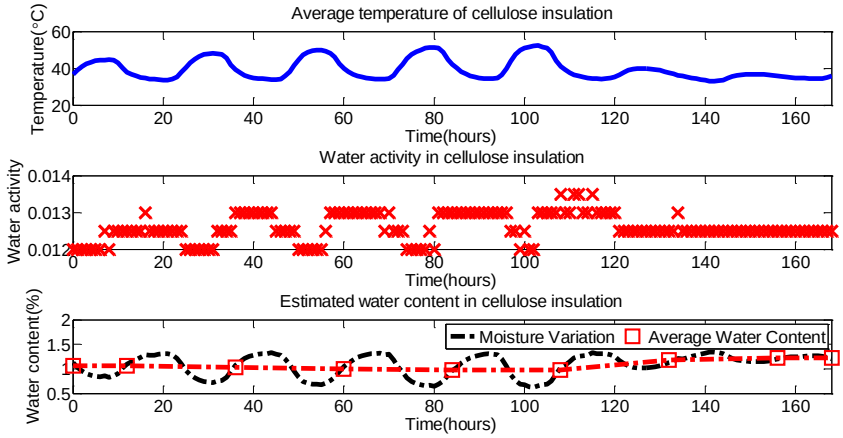


Figure 6.5 Moisture concentration estimation of transformer cellulose

6.4.3 Estimating hot spot temperature for heavy load season

Figure 6.6 presents one week loading profile of the transformer and the ambient temperature during heavy load season (the same as that in Figure 6.5). Figure 6.7 shows the calculated hot spot and top oil temperature by using different thermal models. The measured hot spot and top oil temperatures are also shown in Figure 6.7. From Figure 6.6 and Figure 6.7 it can be seen that the during heavy load season, the daily peak load of the transformer is above 65% of its full load capability. The maximum hot spot temperature and top oil temperature may reach above 71℃ and 58℃ under some circumstances.

From Figure 6.7 it can be observed that the proposed model and Susa's model show good agreement with the measured hot spot temperature. However, IEEE model and Swift's model have a relatively large error in hot spot temperature estimation, which agrees with the results reported in [28].

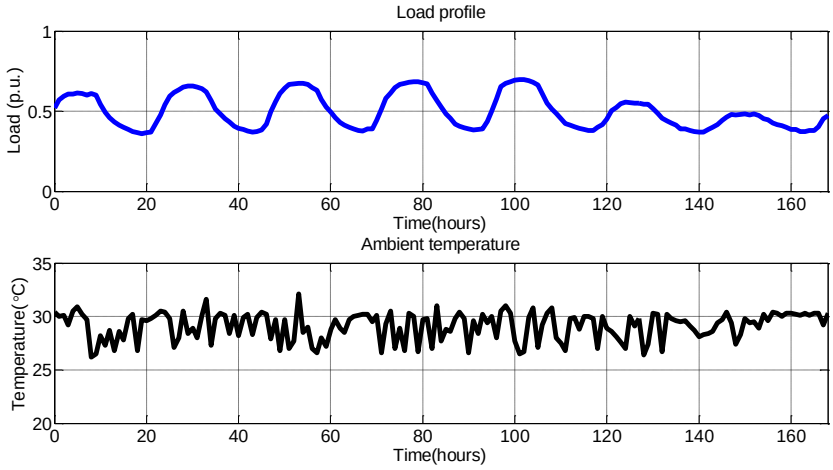


Figure 6.6 Loading profile and ambient temperature during peak load season

To provide a better comparison among hot spot temperatures estimated by different models, a magnification of modelled and measured hot spot and top oil temperature from 90 hours to 124 hours in Figure 6.7 is provided in Figure 6.8. It can be seen that the hot spot temperatures estimated by IEEE, Swift, Susa and MDTM models are close when the fluctuation of load is small, i.e. during the time period of 90 hours to 96 hours and 107 hours to 124 hours in Figure 6.8. However, when the fluctuation of load is large, i.e. during the time period of 97 hours to 106 hours in Figure 6.8, IEEE, Swift and Susa models produce relatively higher hot spot temperature estimation. In contrast, the hot spot temperatures estimated by the proposed MDTM are much close to the measurements during this time period.

The improvement of MDTM over other models is due to the integration of cellulose thermal resistance, which is both temperature and moisture dependent. The transformer load fluctuation leads to changes in cellulose temperature and moisture. These changes cause variations in cellulose thermal conductivity. As the remaining life of the cellulose in transformer is halved with every 6℃ increase in hot spot temperature, the proposed MDTM can help utilities achieve the maximum load capability of transformer while ensuring its operational reliability.

Further statistics comparison among three existing thermal models and the proposed MDTM is made by computing adequacy and accuracy metrics [28] on the estimated hot spot temperature with reference to the measured. The coefficient of determination (R^2) is calculated as adequacy metrics (6.42) and mean squared errors (MSE) is computed as accuracy metrics (6.43).

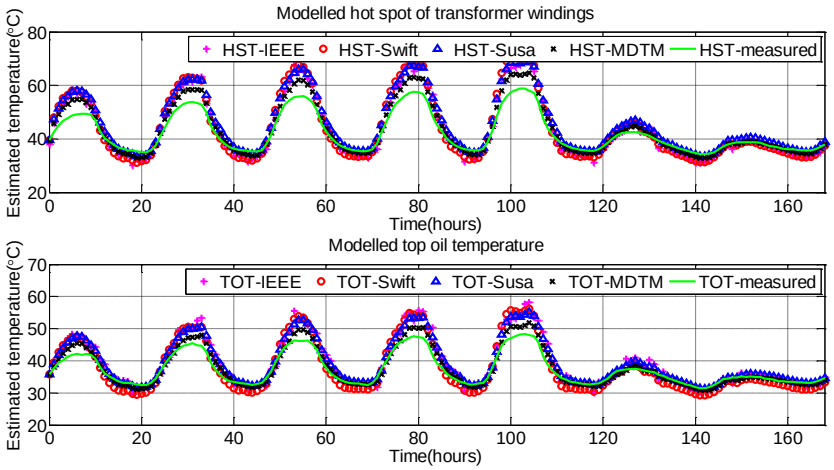


Figure 6.7 Comparison between modelled and measured hot spot and top oil temperature of transformer

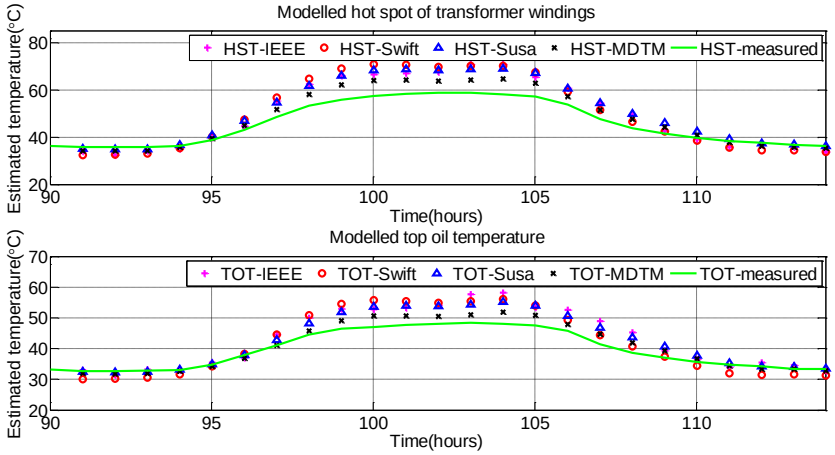


Figure 6.8 Magnification of modelled and measured hot spot and top oil temperature of transformer (from 90 hours to 124 hours in Figure 6.7)

$$R^2 = \frac{\sum_{i=1}^{G}\left(\hat{\theta}_i - \bar{\theta}\right)^2}{\sum_{i=1}^{G}\left(\theta_i - \bar{\theta}\right)^2} \tag{6.42}$$

$$MSE = \frac{1}{G}\sum_{i=1}^{G}\left(\theta_i - \hat{\theta}_i\right)^2 \tag{6.43}$$

where $\hat{\theta}_i$ and θ_i denote the estimated and measured hot spot temperature at time instance i, $\bar{\theta}$ denotes the mean value of measured hot spot temperature, G denotes the sample size of hot spot temperatures in comparison.

The adequacy metrics (coefficient of determination) evaluates the capability of a model to predict new samples, which are not used for establishing the model. R^2 that is close to 1 is desired.

Accuracy metrics measures the average of squares of errors between the modelled temperature and the measured results. MSE that is close to zero is preferred. The results of the two metrics are provided as in Table 6.4 and Table 5 respectively.

From Table 6.4 it can be seen the proposed MDTM attains highest coefficient of determination (0.86, close to 1) among the four thermal models. This demonstrates that MDTM has a satisfied adequacy and a good capability in further forecasting hot spot temperatures. From Table 5 it can be observed that the proposed MDTM has the lowest mean squared error. This again demonstrates that MDTM can consistently attain highest accuracy in estimating the transformer hot spot temperature.

Table 6.4 Adequacy metrics of dynamic thermal models (correlation of determination)

R^2	IEEE	Swift	Susa	MDTM
Top oil temperature	0.49	0.57	0.67	0.91
Hot spot temperature	0.45	0.48	0.57	0.86

Table 6.5 Accuracy metrics of dynamic thermal models (mean squared error)

MSE	IEEE	Swift	Susa	MDTM
Top oil temperature	12.2	10.4	8.1	2.2
Hot spot temperature	29.2	27.5	22.7	7.4

6.4.4 Estimating hot spot temperature for light load season

Another case study is also performed when the transformer was operated in light load seasons. Figure 6.9 presents one week's load profile of the transformer and the recorded ambient temperature in a light load season. From Figure 6.9 it can be seen during the light load season, the transformer normally provides less than 50% of its load capability. Figure 6.10 shows the moisture-in-oil measurement results. From Figure 6.10 it can be observed that the average moisture in cellulose is higher than that in heavy load seasons. This is because the temperature of the transformer oil-cellulose system becomes lower and more moisture is adsorbed in cellulose.

The estimated hot spot and top oil temperature by using different thermal models are shown in Figure 6.11. From Figure 6.11 it can be observed that the proposed MDTM still outperforms other three models. Compared to the results obtained when the transformer was heavily loaded (Figure 6.7), the hot spot temperatures estimated by the proposed MDTM are more close to those measured

when the transformer was lightly loaded (Figure 6.11). The possible reason is that the influence of moisture on thermal conductivity of cellulose becomes more dominant when more moisture resides in the cellulose during light load season.

The adequacy metrics and accuracy metrics among different thermal models under light load condition are computed and the results are shown in Table 6.6 and Table 6.7 respectively. It demonstrates the proposed model consistently outperforms existing models in estimating transformer hot spot temperature.

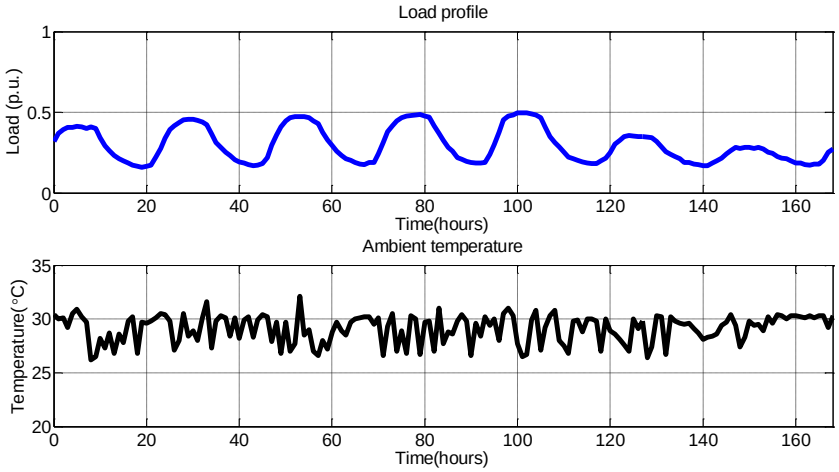


Figure 6.9 Loading profile and ambient temperature during light load season

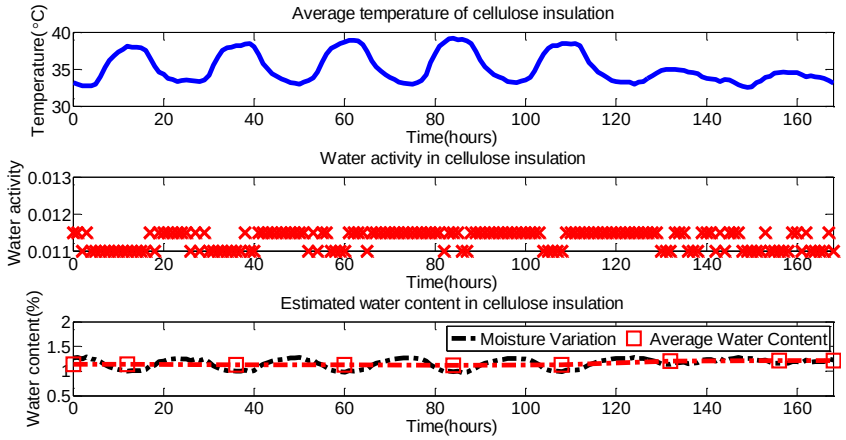


Figure 6.10 Moisture concentration estimation of cellulose

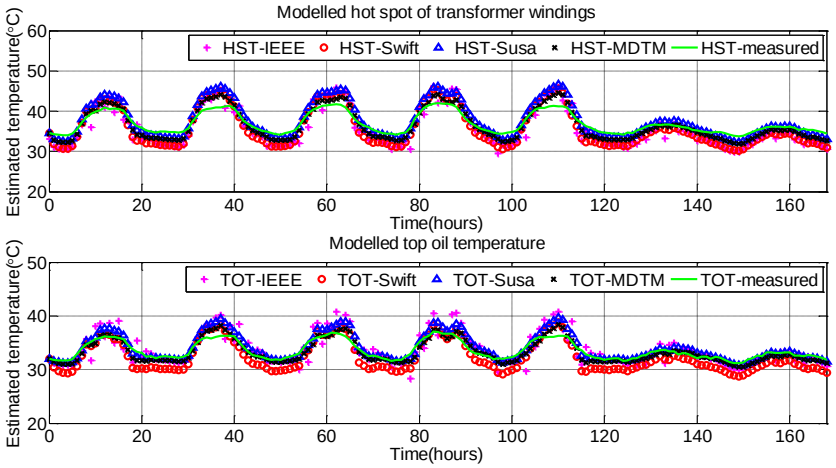


Figure 6.11 Comparison between modelled and measured hot spot and top oil temperature of transformer

Table 6.6 Adequacy metrics of dynamic thermal models (correlation of determination)

R^2	IEEE	Swift	Susa	MDTM
Top oil temperature	0.59	0.61	0.62	0.78
Hot spot temperature	0.53	0.60	0.62	0.69

Table 6.7 Accuracy metrics of dynamic thermal models (mean squared errors)

MSE	IEEE	Swift	Susa	MDTM
Top oil temperature	3.2	2.9	1.1	0.6
Hot spot temperature	6.4	5.7	3.9	2.1

6.5 Conclusion

This paper proposed an improved thermal dynamic model MDTM for computing transformer hot temperature, which considers moisture and temperature effect on the thermal resistance of transformer cellulose. Case studies on a vegetable oil filled transformer demonstrated that the proposed model consistently outperformed existing IEEE model, Swift's model and Susa's model. The proposed model can also be applied to mineral oil-filled transformer with slight modification in oil thermal resistance calculation.

Acknowledgment

The authors gratefully acknowledge Australian Research Council, and industry partners Powerlink Queensland, Energex, Ergon Energy, and TransGrid for supporting this work.

References

[1] *IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators*, IEEE Std C57.91-2011 (Revision of IEEE Std C57.91-1995), 2012.

[2] *Power transformers – Part 7:Loading guide for oil-immersed power transformers*, IEC 60076-7, 2005.

[3] L. W. Pierce, "An investigation of the thermal performance of an oil filled transformer winding," *IEEE Trans. Power Delivery*, vol.7, Issue 3, pp. 1347-1358, 1992.

[4] H. Nordman and M. Lahtinen, "Thermal overload tests on a 400-MVA power transformer with a special 2.5-p.u. Short time loading capability," *IEEE Trans. Power Delivery*, vol.18, Issue 1, pp. 107-112, 2003.

[5] Z. Radakovic and K. Feser, "A new method for the calculation of the hot-spot temperature in power transformers with ONAN cooling," *IEEE Trans. Power Delivery*, vol.18, Issue 4, pp. 1284-1292, 2003.

[6] M. Schafer and F. K, "Thermal monitoring of large power transformers," *PowerTech Budapest 99. Abstract Records. (Cat. No.99EX376)*, Issue98, 1999.

[7] B. C. Lesieutre, W. H. Hagman and J. L. Kirtley, "An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system," *IEEE Trans. Power Delivery*, vol.12, Issue 1, pp. 249-256, 1997.

[8] D. Susa, M. Lehtonen and H. Nordman, "Dynamic Thermal Modelling of Power Transformers," *IEEE Trans. Power Delivery*, vol.20, Issue 1, pp. 197-204, 2005.

[9] G. Swift, T. S. Molinski and W. Lehn, "A Fundamental Approach to Transformer Thermal Modeling-Part I: Theory and Equivalent Circuit," *IEEE Trans. Power Delivery*, vol.16, Issue 2, pp. 171, 2001.

[10] D. Susa and H. Nordman, "IEC 60076-7 loading guide thermal model constants estimation," *Int. Trans. Electr. Energy. Syst.*, vol.23, Issue 7, pp. 946-960, 2013.

[11] R. Lopatkiewicz, Z. Nadolny and P. Przybylek, "The influence of water content on thermal conductivity of paper used as transformer windings insulation," in *Proceedings of 2012 IEEE*

10th International Conference on the Properties and Applications of Dielectric Materials (ICPADM), July 24-28, 2012,Bangalore, India, pp. 1-4.

[12] R. Lopatkiewicz, Z. Nadolny and P. Przybylek, "Influence of water content in paper on its thermal conductivity," *PRZEGLAD ELEKTROTECHNICZNY*, vol.86, Issue 11B, pp. 55-58, 2010.

[13] R. Lopatkiewicz, Z. Nadolny, P. Przybylek and W. Sikorski, "The influence of chosen parameters on thermal conductivity of windings insulation describing temperature distribution in transformer," *PRZEGLAD ELEKTROTECHNICZNY*, vol.88, Issue 11B, pp. 126-129, 2012.

[14] G. Swift, T. S. Molinski, R. Bray and R. Menzies, "A Fundamental Approach to Transformer Thermal Modeling-Part II: Field Verification," *IEEE Trans. Power Delivery*, vol.16, Issue 2, pp. 176, 2001.

[15] Y. Du, M. Zahn, B. C. Lesieutre, A. V. Mamishev and S. R. Lindgren, "Moisture equilibrium in transformer paper-oil systems," *IEEE Electr. Insul. Mag.*, vol.15, Issue 1, pp. 11-20, 1999.

[16] D. F. Garcia, B. Garcia and J. Burgos, "A review of moisture diffusion coefficients in transformer solid insulation-part 1: Coefficients for paper and pressboard," *IEEE Electr. Insul. Mag.*, vol.29, Issue 1, pp. 46-54, 2013.

[17] W. A. Fessler, T. O. Rouse, W. J. McNutt and O. R. Compton, "A refined mathematical model for prediction of bubble evolution in transformers," *IEEE Trans. Power Delivery*, vol.4, Issue 1, pp. 391-404, 1989.

[18] D. Martin, C. Perkasa and N. Lelekakis, "Measuring paper water content of transformers: a new approach using cellulose isotherms in nonequilibrium conditions," *IEEE Trans. Power Delivery*, vol.28, Issue 3, pp. 1433-1439, 2013.

[19] D. Martin, N. Lelekakis, W. Guo and Y. Odarenko, "Further studies of a vegetable-oil-filled power transformer," *IEEE Electr. Insul. Mag.*, vol.27, Issue 5, pp. 6-13, 2011.

[20] *IEEE Standard Terminology for Power and Distribution Transformers*, IEEE Std C57.12.80, 2010.

[21] D. Susa, "Dynamic Thermal Modelling of Power Transformers," Dissertation for the degree of Doctor of Science in Technology dissertation, Department of Electrical and Communications Engineering, Helsinki University of Technology, 2005.

[22] F. P. Incropera and D. P. Dewitt, *Fundamentals of Heat and Mass Transfer*, New York: Wiley, 2011.

[23] "Oil moisture expressed as water activity (aw)" available at http://www.vaisala.com /Vaisala%20Documents/Application%20notes/OilMoistureExpressedasWaterActivity_B210806EN-A.pdf.

[24] B. R. Munson, D. F. Young and T. H. Okiishi, *Fundamentals of Fluid Mechanics*, New York: Wiley, 2009.

[25] D. Susa, "Dynamic Thermal Modelling of Power Transformers," Doctoral dissertation, Department of Electrical and Communications Engineering, Helsinki University of Technology, Helsinki, 2005.

[26] "Envirotemp FR3 Oil Testing Guide," Cooper Power Systems, WI: Cooper Industries Inc, Waukesha, 2004.

[27] V. Oyj,"MMT330 Moisture and Temperature Transmitter Series for Oil" available at http://www.vaisala.com/en/industrialmeasurements/products/moistureinoil/Pages/MMT330.aspx.

[28] O. A. Amoda, D. J. Tylavsky, G. A. McCulla and W. A. Knuth, "Acceptability of Three Transformer Hottest-Spot Temperature Models," *IEEE Trans. Power Delivery*, vol.27, Issue 1, pp. 13-22, 2012.

Chapter 7

Multi-Source Information Fusion for Power Transformer Condition Assessment

Contribution of the Chapter

This chapter presents a multi-source data and information fusion framework for power transformer condition assessment. The proposed method adopts Bayesian Network (BN), which can integrate every piece of data and information obtained from different transformer diagnostic measurements. Within the Bayesian Network, Monte Carlo and Bootstrap methods are employed to extract the most informative characteristics regarding transformer condition from different diagnostic measurements. Reliability metrics are computed to evaluate the effectiveness of combinations of different type diagnostic measurements and subsequently facilitate determining optimal diagnostic strategies involved in transformer condition assessment. Theories, implementations, and results of the proposed method are presented using case studies in this chapter.

7.1 Introduction

Power transformers are one of the most critical and expensive assets in a power system. Its reliability has a significant influence on the reliable operation of a power system [1]. Therefore, it is of great importance to continuously monitor and evaluate transformer's health condition to provide an appropriate advice on transformer operation and maintenance.

Over the decades, a number of diagnostic techniques have been investigated for transformer condition assessment such as dissolved gas analysis (DGA), degree of polymerization (DP) measurement, polarization and depolarization current (PDC) measurement, frequency domain spectroscopy (FDS), frequency response analysis (FRA), and partial discharge (PD) detection. [1]. However, accurate interpreting the measurement data obtained by the above techniques for transformer condition assessment still remains a challenging task.

In this chapter, a Bayesian Network (BN) based multi-source data and information fusion method is proposed. It is aimed at evaluating the overall health condition of a transformer using data and information obtained from some transformer diagnostic measurements. In the proposed method, statistical approaches including Monte Carlo simulation and Bootstrap are employed for extracting most relevant but not redundant information from multiple diagnostic measurements. Then the information extracted from different measurements is aggregated to provide transformer health condition. A driver for this research is to answer an important question that would it be possible to evaluate the overall health condition of a transformer using the limited types of diagnostic measurements. If it is possible, then how to configure the optimal diagnosis strategies, which only require limited types (number) of diagnostic measurements but still can provide reliable condition assessment results.

7.2 Brief Review on Bayesian Network

Bayesian Network (BN) has been commonly used for evaluating equipment condition and identifying potential faults [2-4]. Its implementation involves network structure construction and network parameter configuration. The network structure qualitatively represents the potential relationship between different components in a network while network parameters quantitatively

evaluate how these components are correlated through probabilities.

The structure of a BN is usually represented as a Directed Acyclic Graph (DAG), which is shown in Figure 7.1. In the figure, X_1, X_2, X_3 denote the parent nodes and X_4 denotes the child node. The association between parent nodes and the child node is presented by arcs connecting each other. The network parameters consist of prior probabilities of parent nodes and the conditional probabilities of child nodes. The prior probability represents the occurrence possibility of an event before new information or evidence is available. It can be obtained from statistical interpretation of historical measurement results and experts' experiences. A high prior probability of a parent node implies the event is more likely to happen. The conditional probability denotes the possibility that an event may happen, given the occurrence of certain events are already known. In Figure 7.1, the prior probabilities of parent nodes X_1, X_2 and X_3 are denoted as $P(X_1), P(X_2)$ and $P(X_3)$, and the conditional probabilities of the child node X_4 is denoted as $P(X_4 \mid X_1, X_2, X_3)$ .If conditional probabilities of each child node are determined, the joint probability distribution can be calculated as (7.1) [5].

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2)P(X_3)P(X_4 \mid X_1, X_2, X_3) \tag{7.1}$$

For a child node, its associated components may not be all discovered and subsequently cannot be represented as parent nodes. Therefore, a leaky node is used to represent such component (node X_L in Figure 7.1) [6-7]. A leaky node describes a child node, of which the probability can still be one even the values of all its parent nodes are zero.

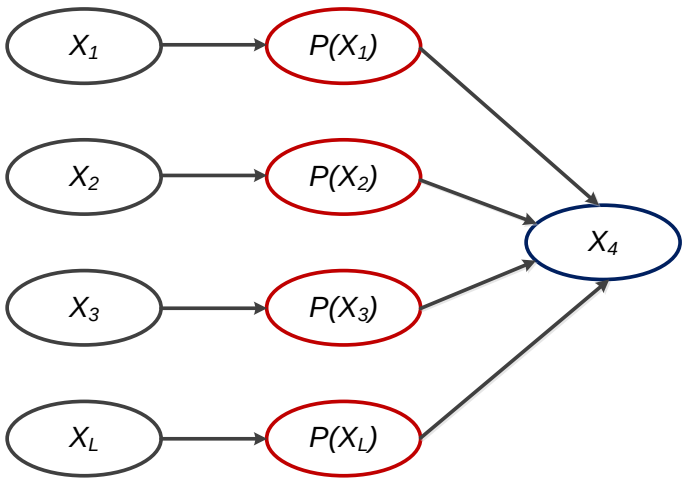


Figure 7.1 Bayesian network with leaky node

7.3 Implementation of Bayesian Network for Multi-Source Data and Information Fusion

7.3.1 Network structure

Table 7.1 summarizes the components and denotations in the Bayesian Network developed for multi-source data and information fusion in the context of transformer condition assessment. Ten potential faults (denoted as F_1 to F_{10} in Table 7.1) such as core multi-point earth fault, insulation ageing, overheating by leakage flux, etc., are considered as causes [8]. These faults cover various faulty conditions that may occur in a transformer. The consequences of these faults are considered as fault evidence, which are represented as nine fault symptoms (denoted as S_1 to S_9 in Table 7.1) and three abnormal operation conditions (denoted as C_1 to C_3 in Table 7.1).

From the literature and experience of utility engineers [8-9], the causal relationship between different fault types and fault evidence are obtained and used to establish the network (Figure 7.2). The Bayesian network contains two layers: fault diagnosis layer and fault evidence layer.

Table 7.1 Denotation and description of nodes in network

Denotation		Description
Fault types	F_1	Core multi-point earth fault
	F_2	Insulation ageing
	F_3	Overheating by leakage flux
	F_4	Winding short circuit
	F_5	Moistened insulation
	F_6	Tap-changer failure
	F_7	Floating potential discharge
	F_8	Discharge in barrier
	F_9	Winding deformation
	F_{10}	Discharge in oil
Fault symptoms	S_1	Core earthing current
	S_2	Overheating indicated by IEC code
	S_3	Unbalanced DC resistance of winding
	S_4	Moisture in oil
	S_5	Discharge indicated by IEC code
	S_6	Deviation of winding ratio
	S_7	Partial discharge

	S_8	Dissolved volume ratio CO/CO_2
	S_9	Polarization index
Abnormal operation conditions	C_1	Abnormal overload
	C_2	External short circuit
	C_3	Lightening

7.3.2 Network parameter

To infer transformer condition, parameters need to be assigned to each node in the above network. These parameters consist of two groups, i.e., prior probabilities of parent nodes (nodes in fault diagnosis layer) and conditional probabilities of child nodes (nodes in fault evidence layer). For transformer condition assessment, the prior probabilities of each fault are originated from CIGRE report [9-10] and reference [11] and they are listed in the first row of Table 7.2.

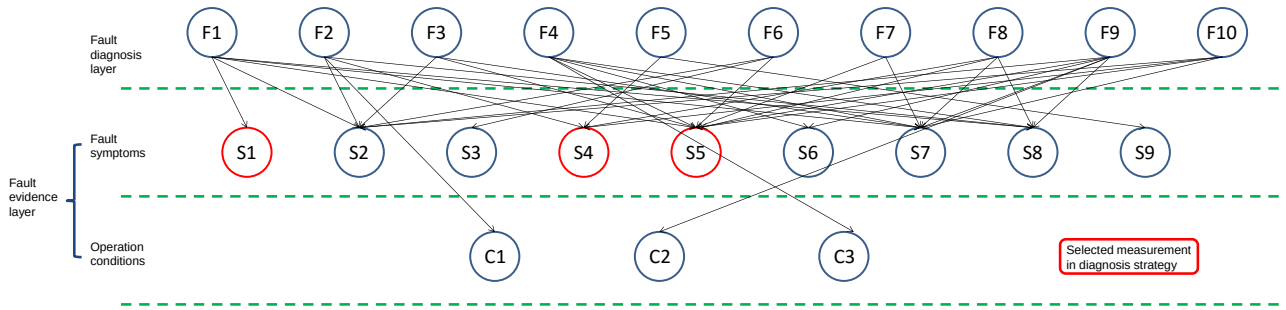


Figure 7.2 Multiple source Bayesian network for power transformer condition assessment

The conditional probabilities of each child node are normally presented in a tabulate form, known as Conditional Probability Table (CPT). The major difficulty in implementing CPT is the exponential increase in the elements number when the network contains a large number of child nodes. The elements number for a particular child node can be calculated as (7.2).

$$n_{CPT} = (s_c - 1)\prod_{i=1}^{n_p} s_{p_i} \tag{7.2}$$

where n_{CPT} denotes the number of elements for a particular child node, s_c denotes the number of child node states, n_p denotes the number of parent nodes associated with this child node, and s_{p_i} denotes the states number of *i*-th parent node.

Suppose each node in the network has binary states (i.e., present and absent), Noisy-OR model [6] is used to compute the conditional probabilities of CPT as shown in Table 7.2. Consider there are *m*

causes $X_1, X_2, \ldots, X_m$ of event Y, in Noisy-OR model, the full conditional probabilities [6] of Y can be calculated as(7.3).

$$P\left(Y \mid X_1, X_2, \cdots, X_m\right) = 1 - \left[\left(1 - P_0\right) \prod_{1 \le i \le m} \left(1 - P_i\right)\right] \tag{7.3}$$

where P_0 denotes the leaky probability, and P_i denotes the conditional probability of the i-th parent node.

Table 7.2 Correlation between fault types and symptoms and operation conditions

Fault types (prior probability)	**F_1 (0.45)**	**F_2 (0.11)**	**F_3 (0.13)**	**F_4 (0.12)**	**F_5 (0.1)**	**F_6 (0.26)**	**F_7 (0.16)**	**F_8 (0.28)**	**F_9 (0.24)**	**F_{10} (0.14)**	**Leaky node***
S_1	0.9										0.01
S_2	0.818	0.219	0.713			0.674			0.149	0.2	0.01
S_3						0.87					0.01
S_4		0.267			0.718			0.416		0.6	0.01
S_5	0.189		0.289	0.515		0.231	0.863	0.879	0.618	0.7	0.01
S_6				0.8					0.8		0.01
S_7	0.3		0.35	0.9			0.9	0.9	0.75	0.9	0.01
S_8	0.2	0.816		0.681				0.759	0.721		0.01
S_9					0.75						0.01
C_1		0.02									0.01
C_2									0.24		0.01
C_3				0.55							0.01

* Leaky probability describes the probability of a leaky node, which can still be one even the values of all parent nodes of this leaky node are zero.

7.4 Case Studies and Results Analysis

This section presents case studies to demonstrate how to determine optimal diagnosis strategies by using the proposed Bayesian network given limited types of diagnostic measurements. The flow chart for determining optimal diagnosis strategies is shown in Figure 7.3.

In the case study, it is assumed that there are three measurements that can provide evidence of three fault types: core earthing current (S_1), moisture in oil (S_4), and discharge indicated by IEC code (S_5). One measurement corresponds to one transformer diagnostic technique. The original

(non-optimized) diagnosis strategy consists of all the three measurements. The optimal diagnosis strategies would be different combinations of S_1, S_4 and S_5, which involves less number of measurements but still can provide reliable diagnosis. Table 7.3 summarizes the measurement configurations for different diagnosis strategies. Four strategies are investigated, including: the non-optimized diagnosis strategy including all three measurements $[s_1, s_4, s_5]$; and three optimized strategy candidates with each only consisting of two measurements (i.e., $[s_1, s_4]$,$[s_4, s_5]$ and $[s_1, s_5]$). It is worth mentioning that the selection of fault evidences in original diagnosis strategy (i.e., S_1, S_4 and S_5) is not exclusive. Actually it can be any combinations of certain number child nodes (from S_1 to C_3) in fault evidence layer in Figure 7.2.

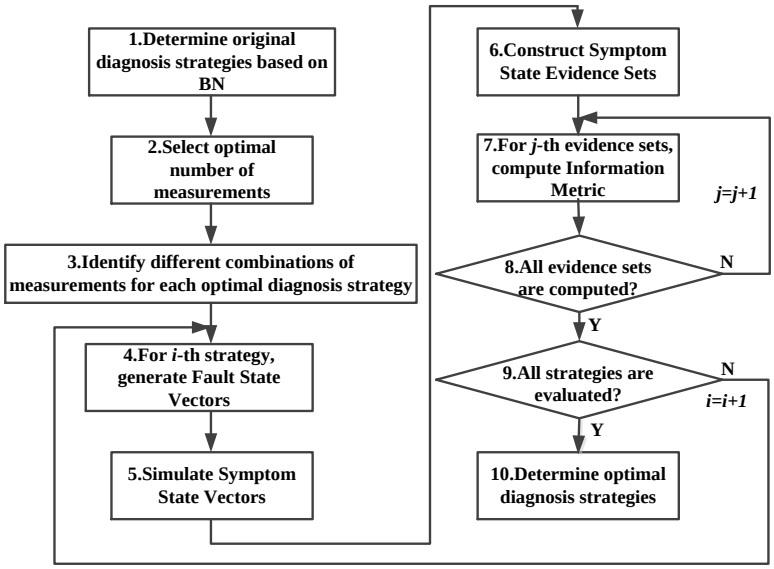


Figure 7.3 Flow chart of optimal diagnosis strategies determination

Table 7.3 Diagnosis strategies under investigation

Strategies index	Strategies type	Measurement configuration
1	Non-optimized strategy	S_1, S_4, S_5
2	Optimized strategy	S_1, S_4
3	Optimized strategy	S_4, S_5
4	Optimized strategy	S_1, S_5

7.4.1 Fault state vectors (FSV) generation

All possible combinations of parent nodes' states in fault diagnosis layer of Figure 7.2 are defined

as Fault State Vectors (FSV). To generate FSV from the network, Monte Carlo and Bootstrap methods are employed [12-14]. In Monte Carlo method, uniform distribution is applied. First, a vector containing ten random numbers $[x_1, x_2, \cdots, x_{10}]$ within range $[0,1]$ is generated and converted to binary states by (7.4). This vector denotes the combination of occurrence of F_1 to F_{10} and it is used to generate the first fault state vector (FSV_1) in Table 7.4. Then the above procedure is repeated by using bootstrap for a large number of times to create a set of fault state vectors. A publicly available bootstrap toolbox for MATLAB [15] has been used in this research. Since each parent node has binary states and ten parent nodes are included in the network, the pattern of FSV may reach up to 2^{10}=1024. The occurrence probability of each pattern can be calculated as a ratio of occurrence number of each FSV over the total number of vectors (shown in Table 7.4).

$$f(x_i) = \begin{cases} \Pr esent & 0.5 \le x_i \le 1 \\ Absent & 0 \le x_i < 0.5 \end{cases} \tag{7.4}$$

Table 7.4 Sorted fault state vectors and probability

Fault state vectors	**F_1**	**F_2**	**F_3**	**F_4**	**F_5**	**F_6**	**F_7**	**F_8**	**F_9**	**F_{10}**	**Prob**
FSV_1	Abs	Abs	Abs	Abs	Abs	Abs	Abs	Abs	Abs	Abs	2.18×10^{-3}
FSV_2	Abs	Abs	Abs	Abs	Abs	Abs	Abs	Abs	Abs	**Pres**	1.93×10^{-3}
…	…	…	…	…	…	…	…	…	…	…	…
FSV_{1024}	**Pres**	**Pres**	**Pres**	**Pres**	**Pres**	**Pres**	**Pres**	**Pres**	**Pres**	**Pres**	3.6×10^{-4}

Pres: Present, Abs: Absent, Prob: Occurrence probability

7.4.2 Symptom state vectors (SSV) simulation

Each FSV generated from Table 7.4 is then fed into the proposed Bayesian network. In turn corresponding probabilities of child nodes in fault evidence layer are computed as outputs. Each set of output probabilities at the selected child nodes in Strategy 1 (S_1, S_4 and S_5 in Table 7.3) are recorded and used to generate a set of Symptom State Vectors (SSV). The generated SSV for Strategy 1 are shown in Table 7.5.

From Table 7.5 it can be seen that in Strategy 1, the whole set of SSV which contains three measurements with binary states, may have 2^3=8 patterns (SSV_1 to SSV_8). The occurrence probability of each pattern can be calculated as a ratio of occurrence number of each SSV over the total number of vectors. The occurrence probabilities for each SSV in Strategy 1 are shown in the last row of Table 7.5.

Table 7.5 Symptom state vectors and probability in configuration 1

	SSV_1	**SSV_2**	**SSV_3**	**SSV_4**	**SSV_5**	**SSV_6**	**SSV_7**	**SSV_8**
S_1	Pres	Pres	Pres	Pres	Abs	Abs	Abs	Abs
S_4	Pres	Pres	Abs	Abs	Pres	Pres	Abs	Abs
S_5	Pres	Abs	Pres	Abs	Pres	Abs	Pres	Abs
Prob	0.381	0.006	0.09	0.007	0.4	0.009	0.098	0.009

7.4.3 Symptom state evidence sets (SSES) construction

Based on SSV generated in Table 7.5, a large number of Symptom State Evidence Sets (SSES) are established by utilizing Monte Carlo simulation. During each simulation, three SSV are randomly drawn from Table 7.5 to construct a symptom state evidence set. The Monte Carlo simulation is repeated for 10^5 times and the symptom state evidence sets are shown in Table 7.6.

Table 7.6 Symptom state evidence sets and probability in configuration 1

	Member 1	**Member 2**	**Member 3**	**Prob**	**RM***
$SSES_1$	SSV_5	SSV_5	SSV_1	0.18	0.48
$SSES_2$	SSV_5	SSV_1	SSV_1	0.17	0.27
…	…	…	…	…	…
$SSES_{80}$	SSV_3	SSV_3	SSV_2	0.0001	0.98

*RM: Reliability Metric value

7.4.4 Reliability metric (RM) evaluation

For each set of symptom state evidence in Table 7.6, Reliability Metric (RM) is computed to evaluate the reliability for different diagnosis strategies [16] by the following steps.

(i) For the first evidence set ($SSES_1$ in Table 7.6) of Strategy 1, the first member (i.e., SSV_5) is used as input $\left[\overline{s}_1, s_4, s_5\right]$ to instantiate the Bayesian network and in turn the output probabilities of ten faults in fault diagnosis layer are recorded (shown in Table 7.7). The values in the second row of Table 7.7 denote the occurrence probabilities of different faults given the fault symptoms of S_1 is absent while S_4 and S_5 are presented.

(ii) Repeat step (i) for the rest two members (SSV_5 and SSV_1) in the first evidence set $SSES_1$.

Consequently, two more sets of occurrence probabilities for different faults can be obtained and they are shown in Table 7.7.

(iii) For each parent node (F_1 to F_{10}), mean value (μ_i) and variance (σ_i) of three sets of output probabilities obtained from step (ii) are calculated. The calculation results are summarized in Table 7.8. Then the ratios of σ_i / μ_i for all parent nodes are used to form a vector $\left[\frac{\sigma_1}{\mu_1}, \frac{\sigma_2}{\mu_2}, \ldots \frac{\sigma_{10}}{\mu_{10}}\right]$.

(iv) The summation of all elements in the vector of step (iii) are computed as $\left(\frac{\sigma_1}{\mu_1} + \frac{\sigma_2}{\mu_2} + \ldots + \frac{\sigma_{10}}{\mu_{10}}\right) = 0.48$. This value is used as the reliability metric (RM_1) value for the first SSES of diagnosis Strategy 1.

(v) The occurrence probability of the first evidence set is computed as the ratio between the occurrence number of the first evidence set and the total number of generated evidence sets. The occurrence probability is list in the column "Prob" of Table 7.6.

(vi) The same procedures are repeated for the rest of SSES in Table 7.6 (up to $SSES_{80}$). Then the reliability metric for diagnosis Strategy 1 can be calculated as (7.5) using Table 7.6.

$$RM_{DS1} = \sum_{j=1}^{k} RM_j \times \text{Pro}bability_j = 0.48 \times 0.18 + 0.27 \times 0.17 + \cdots + 0.98 \times 10^{-4} = 0.4 \quad (7.5)$$

where $k = 80$ denotes the total number of symptom state evidence sets for diagnosis Strategy 1, RM_j denotes the reliability metric value for *j*-th symptom state evidence set.

Table 7.7 Output probabilities of faults in fault diagnosis layer for input $SSES_1$

Input of BN	**F_1**	**F_2**	**F_3**	**F_4**	**F_5**	**F_6**	**F_7**	**F_8**	**F_9**	**F_{10}**
SSV_5	0.08	0.16	0.14	0.14	0.22	0.27	0.19	0.63	0.28	0.38
SSV_5	0.08	0.16	0.14	0.14	0.22	0.27	0.19	0.63	0.28	0.38
SSV_1	0.99	0.16	0.14	0.13	0.23	0.27	0.19	0.61	0.27	0.37

Table 7.8 Mean value and variance of posterior probability of parent nodes in $SSES_1$ for strategy 1

		F_1	**F_2**	**F_3**	**F_4**	**F_5**	**F_6**	**F_7**	**F_8**	**F_9**	**F_{10}**	**RM_1**	**Prob**
$SSES_1$	μ_i	0.38	0.16	0.14	0.14	0.22	0.27	0.19	0.63	0.28	0.38	0.48	0.18
	σ_i	0.18	6×10^{-6}	9×10^{-7}	3×10^{-6}	4×10^{-5}	2×10^{-6}	10^{-5}	10^{-4}	2×10^{-5}	2×10^{-5}		

The reliability metric values for other diagnosis strategies in Table 7.3 are calculated by using the same procedures. The comparison of reliability metric values among four diagnosis strategies is shown in Figure 7.4.

From Figure 7.4 it can be seen that the reliability metric value for Strategy 1 attains the highest among all diagnosis strategies since all fault symptoms are considered and all relevant measurements results are integrated in this strategy. For diagnosis Strategies 2, 3 and 4, only two fault symptoms are considered. An intuition would be a significant reduction in the reliability of diagnostic results for these three strategies. However, from Figure 7.4 it can be seen Strategy 2 can achieve similar diagnostic performance as Strategy 1 though it only requires two measurements. Therefore, this diagnosis strategy is considered as the optimal measurement configuration. For Strategy 3, the RM value attains the lowest among four strategies. This implies that insufficient information regarding transformer faults detection can be provided and difficulties exist in obtaining reliable diagnostic results by only revealing fault symptoms of S_4 and S_5.

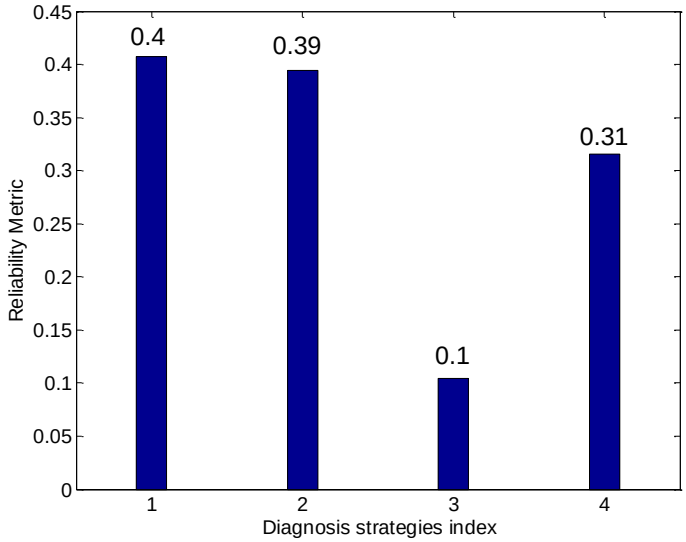


Figure 7.4 Comparison of reliability metrics for different strategies

7.5 Conclusion

In this chapter, a multi-source information fusion approach is developed, which aims at helping utilities to decide the most reliable diagnosis strategies for transformer condition assessment. The utilization of Bayesian Network can facilitate inferring potential faults of a transformer with limited types of diagnostic measurements. The proposed method can provide an efficient way to configure

the diagnostic measurements, which can provide adequate but non redundant information for revealing transformer condition. Future work will focus on collecting more datasets from different diagnostic measurements and conducting extensive case studies. It is expected that this work can lead to a practical tool, which can help utilities better determine their condition monitoring and diagnosis strategies in health management of their transformer fleet.

References

[1] A. E. B. Abu-Elanien and M. M. A. Salama, "Asset management techniques for transformers," *Electr. Power Syst. Res.*, vol.80, pp. 456-464, 2010.

[2] J. Seshadrinath, B. Singh and B. K. Panigrahi, "Incipient Interturn Fault Diagnosis in Induction Machines Using an Analytic Wavelet-Based Optimized Bayesian Inference," *IEEE Trans. Neural Networks Learn. Syst.*, vol.25, pp. 990-1001, 2014.

[3] B. Cai, Y. Liu, Q. Fan, Y. Zhang, Z. Liu and S. Yu, *et al.*, "Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network," *Appl. Energy*, vol.114, pp. 1-9, 2014.

[4] D. Kateris, D. Moshou, X. Pantazi, I. Gravalos, N. Sawalhi and S. Loutridis, "A machine learning approach for the condition monitoring of rotating machinery," *J. Mech. Sci. Technol.*, vol.28, pp. 61-71, 2014.

[5] G. F. Cooper and E. Herskovits, "A Bayesian Method for the Induction of Probabilistic Networks from Data," *Mach. Learn.*, vol.9, pp. 309-347, 1992.

[6] A. Zagorecki and M. J. Druzdzel, "Knowledge Engineering for Bayesian Networks: How Common Are Noisy-MAX Distributions in Practice," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.43, pp. 186-195, 2013.

[7] F. J. D Ez and S. F. Gal N, "Efficient computation for the noisy MAX," *Int. J. Intell. Syst.*, vol.18, pp. 165-177, 2003.

[8] W. Gao, C. Bai and T. Liu, "A Dynamic Integrated Fault Diagnosis Method for Power Transformers," *The Scientific World Journal*, vol.2015, pp. 1-8, 2015.

[9] A. Bossi, "An international survey on faults in large power transformers in service," *Final Report of CIGRE Working Group 12.05*, pp. 22-48, 1983.

[10] J. A. Lapworth, "Transformer Reliability Surveys," *CIGRE 2006 Paris Symposium*, 2006.

[11] L. Wu, "Research on assessment method of transformer," PhD dissertation, Department of Electrical Engineering, North China Electrical Power University, Baoding, 2005.

[12] A. M. Zoubir and B. Boashash, "The bootstrap and its application in signal processing," *IEEE*

Signal Process. Mag., vol.15, pp. 56-76, 1998.

[13] A. M. Zoubir and D. R. Iskandler, "Bootstrap methods and applications," *IEEE Signal Process. Mag.*, vol.24, pp. 10-19, 2007.

[14] D. Zhou, Z. Wang, P. Jarman and C. Li, "Data Requisites for Transformer Statistical Lifetime Modelling—Part II: Combination of Random and Aging-Related Failures," *IEEE Trans. Power Delivery*, vol.29, pp. 154-160, 2014.

[15] A. M. Zoubir and D. R. Iskander,"Bootstrap Matlab Toolbox" available at http://www.csp.curtin.edu.au/downloads/bootstrap_toolbox.html.

[16] M. Pourali and A. Mosleh, "A Functional Sensor Placement Optimization Method for Power Systems Health Monitoring," *IEEE Trans. Ind. Appl.*, vol.49, pp. 1711-1719, 2013.

Chapter 8

Conclusions and Recommendations for Future Work

8.1 Conclusions

The main contribution of this thesis is the development of a suite of advanced data-centric approaches for appropriately extracting information from measurement data, accurately modelling key parameters, and reliably evaluating transformer's condition. By exploiting the correlations between measurement data and transformer fault symptoms using historic datasets and conducting coupled multi-physics modeling using transformer design data, the data-centric approaches can reveal the current health status of a transformer and predict its future condition. The data-centric approaches have been successfully applied to three main aspects concerning transformer's health condition, including oil characteristics and dissolved gases in transformers, moisture dynamics between oil and cellulose insulation, and hot spot temperature of transformer winding.

After conducting a thorough literature survey on different online/offline condition monitoring and fault diagnostic strategies of transformers, the major limitations of existing transformer diagnostic techniques are identified. These limitations are: (1) the lack of framework and proper techniques in constructing training database, improving data quality, recognizing different types of faults, and evaluating the performance of algorithms in applying data-centric techniques to oil characteristics and dissolved gas analysis; (2) the lack of comprehensive multi-physics modelling on complex temperature dependent moisture dynamics between oil and cellulose system and interpretation of its effect on dielectrics response measurement; (3) the lack of accurate thermal dynamic models for estimating hot spot temperature of transformer windings considering the effect of moisture concentration and hot spot temperature on the thermal resistance of transformer's oil-cellulose system.

To improve the interpretation of dissolved gas analysis for transformer fault identification, comparative studies on a variety of state-of-the-art pattern recognition algorithms for transformer insulation diagnosis using oil characteristics are conducted. A variety of representative pattern recognition algorithms are implemented by using eight DGA and oil characteristics datasets obtained from different utilities companies. The procedures of algorithms training, cross-validation, testing and evaluation are standardized, which can provide a common ground for evaluating the performances of different algorithms. A statistical performance (in terms of classification accuracy) comparison amongst different pattern recognition algorithms is conducted for DGA interpretation. The diagnostic results from different algorithms show that the performance of the algorithms highly depends on the quantity and quality of the training datasets. Among 15 algorithms, fuzzy logic attains the highest classification accuracy since it can soften fault decision boundaries in traditional DGA methods and explicitly display expert knowledge that cannot be directly extracted from the

DGA data.

To overcome the lack of samples of particular faults (class imbalance) during algorithm training process, a hybrid data pre-processing algorithm SMOTEBoost and bootstrap is proposed in this thesis. In the proposed algorithm, SMOTE is implemented for dealing with the class imbalanced problem while the boosting approach is adopted for adaptively reweighting and grouping data points in the training dataset. Moreover, the bootstrap method is also utilized to statistically equalize the samples number after executing SMOTE. After being processed by SMOTEBoost and bootstrap, the original DGA datasets become balanced and then they are employed as training datasets for a number of representative pattern recognition algorithms including C4.5, *k*NN, RBF network, and SVM to perform multi-category fault classifications. Extensive numeric experiments are performed to validate the generalization capability of the SMOTEBoost algorithm. The simulation results show that the integration of SMOTEBoost and bootstrap can significantly improve the fault diagnosis accuracy. The highest accuracy of generalization capability may reach up to 98% by using C4.5 decision tree and SVM classifiers integrated with SMOTEBoost, which is much higher than that obtained without using the proposed hybrid data pre-processing techniques.

Other than DGA and oil characteristics, moisture is another key factor that can significantly affect transformer insulation condition. This thesis proposes a multi-physics modelling approach to reveal the temperature dependent moisture dynamics in transformer's insulation system. In this approach, moisture dynamics is modelled by coupling a number of physics simultaneously including the electromagnetic, thermal, fluid flow and moisture migration. Finite element method (FEM) is then utilized to model the moisture distribution of a prototype transformer under load variation and thermal transients.

As the finite element method is quite sensitive to the Peclet number for modelling the mass transfer process, this thesis proposes an alternative population based approach, i.e., particle tracing method for estimating moisture of cellulose insulation in transformers. Different from the conventional approaches that are based on Fick's diffusion law, the population based method formulates moisture diffusion from a microscopic view of water molecules' motion. Especially, the transmission probability of water molecules (termed as particles in the thesis) is employed to correlate the microscopic particles' motion with the macroscopic moisture distribution. Particle tracing method can avoid convergence difficulties often encountered in the conventional FEM method. Particle trajectories are computed in a Lagrangian reference frame, removing the restriction on the ranges of the Peclet number. The particle tracing method is implemented for estimating moisture distribution of both oil-impregnated pressboard specimens and oil-cellulose insulation of a prototype transformer. The comparison of estimated moisture among the proposed particle tracing

method, multi-physic approach and KFT is conducted. It demonstrated both two modelling techniques can be used to evaluate the moisture content with satisfied accuracy.

To evaluate moisture content in transformers, dielectric response measurements are investigated in the thesis. It is found that considerable difficulties still existed in the interpretation of dielectric response measurement of oil-cellulose insulation of the transformer since it is largely dependent on moisture concentration and temperature. On the basis of achievements obtained from moisture dynamics modelling techniques, further investigations on moisture dynamics and its effect on dielectric response of a transformer's cellulose insulation are conducted. This investigation develops a distributed parameter model where the oil-cellulose insulation is sliced into a number of pieces of paper and each layer of paper has a different permittivity and conductivity. The correlation between moisture distribution (under non-equilibrium conditions due to thermal transients) and dielectric response parameters (dielectric losses and permittivity) of cellulose insulation is modelled as exponential function. Two case studies are performed to evaluate the effect of non-uniform moisture distribution on the dielectric response of transformer insulation. The methodology developed in this thesis can help the proper interpretation of dielectric response measurement of field transformers under moisture and thermal transients.

Conventional IEC and IEEE loading guides provide some practical and convenient ways to calculate hot spot temperature of transformer windings. However, certain errors can be observed due to inaccuracies of these empirical thermal models. In this thesis, a moisture dependent thermal model for estimating transformer hot spot temperature is proposed. Two electrical-thermal equivalent circuits are developed for modelling top oil and hot spot temperature of power transformers. In these models, nonlinear thermal resistance is formulated by considering both oil and cellulose (paper/pressboard) of the transformer. Especially, the effect of moisture concentration and hot spot temperature on the thermal resistance of cellulose is taken into consideration. The proposed dynamic thermoelectric models are employed to calculate hot spot temperatures of oil-cellulose system of an in-service vegetable oil-filled power transformer under ambient temperature variation and load fluctuation. Considering the configuration of transformer's cooling system and the feasibility of the optic fibers sensor installment, two temperatures, i.e., top oil and hot spot temperature, are chosen as the model outputs. The modelled results are verified by collecting the historical data from online moisture-in-oil and temperature measurements of the transformer. The comparison between the proposed thermal dynamic model and a number of existing thermal models is performed on the basis of statistical adequacy and accuracy metrics. The comparison results indicate that the proposed model can accurately reveal thermal dynamics of oil-immersed transformers.

To provide an accurate condition assessment of transformer, it is necessary to integrate every piece of data and information obtained from different transformer diagnostic measurements. This thesis proposes a multi-source data and information fusion framework. In the proposed framework, Bayesian Network (BN) is constructed to reveal the correlations between fault symptoms and fault types of a transformer. The elements in the Bayesian Network including fault state vectors, symptom state vectors and symptom state evidence sets are generated by using Monte Carlo and Bootstrap methods. Then, the most informative but not redundant characteristics regarding transformer condition are extracted from different diagnostic measurements. Reliability metrics are subsequently computed to evaluate the effectiveness of combinations of different type diagnostic measurements and subsequently facilitate determining optimal diagnostic strategies involved in transformer condition assessment.

8.2 Recommendations for future works

Various issues regarding development of data-centric diagnostic approaches have been investigated in this thesis. Nevertheless, the investigation can be continued in the future to address other issues. These are as follows:

(1) The performance of a pattern recognition algorithm is determined by a number of factors such as the model structure, model parameters and training method. Given many different combinations of these factors, it is not an easy task to evaluate and compare the performance of different algorithms. More numeric experiments could be performed by using different parameters in various algorithms to provide a common ground for evaluating the performances of different algorithms.

(2) When applying pattern recognition techniques to assess the health condition of transformer insulation, data quality issues could be further investigated. As DGA and oil characteristics database normally has a non-uniform sample distribution, some data-quality control techniques could be implemented in future to identify those data, which can degrade the reliability of the algorithms and then remove them from the datasets.

(3) Moisture diffusion measurement could be further conducted on pressboard samples at different humidity and temperature levels to help understand moisture diffusion process in more detail. In addition, moisture diffusion measurements on pressboard samples with different thickness will provide further understandings. Moisture estimated from dielectric response measurement needs to be validated with other alternative methods. More measurements on pressboard samples at different humidity and temperatures could be useful to determine moisture diffusion coefficients.

(4) Additional works are recommended to further understand the nature of polarization based dielectric response measurements. Future works could be directed towards the development of improving interpretation schemes of PDC and FDS and understanding the impact of different parameters including moisture, ageing of pressboard sample, ageing products such as acids on dielectric response measurements.

(5) Hot spot temperature estimation could be further studied and more field measurements would help to verify proposed moisture dependent thermal dynamic models.

(6) Additional techniques such as frequency response analysis, partial discharge measurements, acoustics and vibration experiments also need to be integrated into the data and information fusion framework to provide a more accurate and overall health condition of transformers.