HAN UNIVERSITY OF APPLIED SCIENCE

Interfaculty institute HAN Master's Programmes (HMP)
Master's Degree Programme of Control Systems Engineering

Cyprien Nsengimana
Student ID: 501516

INLINE MEASUREMENTS AND CONTROL OF PAPER WEB STRENGTH, ILSE

Company: Kenniscentrum Papier en Karton, The Netherlands
www.kcpk.nl

Examiners: ir. Edwin Tazelaar

Supervisors: : ir.Peter Ypma & ing. Chris van Beekum

ing. Bertjan Maasdam

ABSTRACT

The objective of this thesis was to investigate the feasibility of the patent [1] where the most defined input of the studied process is known as Elastic modulus. The implementation of inline measurement system would help a direct control of the paper quality online and therefore react to its instant changes by adjusting E-modulus online. To achieve this goal, different existing control techniques and assumptions were used on the process whereby the poperoller provides the energy to stretch the paper and the amount of stretch created into the paper is due to the speed difference between the last drying cylinder group and the poperoller. The simulation results proved that it is unsuitable to use the patent model whereby E-modulus is calculated based on theory of Young modulus, on the theory of stress to the strain. The quality specs that were used to train and validate the model are bursting strength; short compression test (SCTcd) in cross machine direction and concora medium test (CMT30). It was noted that the number of things might cause the low correlation or there may be the possibility of no-correlation at all. Indeed, it was found that the raw data give much more good prediction than E-modulus. This is because the information we normally want from the data is due to the contribution of individual variable in the process within the specified field of study. Therefore, it is undoubtedly known that good data gives good predictions, which result in the good model. One-month data was used for the modal analysis where two minutes was found to be the best time stamps to retrieve the online data in the system and therefore have the correlation with the measured paper quality in the laboratory. However there were no study done on the data storage system.The study trusted the storage system and proceeds to the analysis of the stored data. Since the result from data set are bad, it good to react on data storage in the paper factory and look on the sampling time that can be better than the chosen time before the end of Jumbo reel. Another important notice is about the direct use of raw materials usually used to increase the paper strength like starch. It is possible that the contribution of E-modulus will be assisted with other strength additives material for the recycled paper on the factory. Furthermore, the testing environments of the paper online are very different with the measured samples offline. This may cause low correlation together with the combination of other uncontrolled nonlinearities on the factory. To come up with those conclusive ideas; a research work was structured into three main parts. The first part was the identification of process variables and the feasible way of accessing the required data for process analysis. At this stage, a number of process variables, which are not directly measured, it was the time to investigate the way of measuring them during its second part of my research. The example was thickness. Referring to thickness testing, three types of sensors were reviewed but finally they were not implemented. Another technique with good assumptions was used due to its performance. It was a relation between already installed grammage sensor and thickness by assuming constant density of the paper. The final part of the thesis is devoted to PCA and PLS model design. The PCA model give both the correlation between process variables without taking into account the response variables and number of principal components, which represent the most of the variance in the data set and thus contribute to the dimension reduction of the process variables without losing too much information in the data set. The PLS simulations based on the PCA results proved the very low correlation between inputs where E-modulus is used and the qualities which thus makes the patent unsuitable for other paper mills.

PREFACE

This master's Thesis is a result of my studies at HAN University of Applied Sciences,
Interfaculty institute HAN master's Programmes (HMP)
Master's degree programme of Control Systems Engineering.
I would like to show my gratitude to several people who made this thesis possible.

My special thanks go to Netherlands Fellowship programmes (NFP), who gave me their full support during the entire course of my studies.

The studies presented in this Master thesis were carried out at KCPK and the work was carried out on the research patent project, NL Patent Pub.No.PCT/NL2011/050682,2011, sponsored by KCPK, Arnhem in The Netherlands.

I would express my sincere thanks to KCPK, Arnhem for providing a topic of my Master's thesis and all the supports they have offered to me in order to carry out this research.

I would like to thank SKRP, for its tremendous assistance in order to allow the feasible access to the practical part of this research. Your support was highly appreciated.

I would like to thank my supervisors: ir. Peter Ypma, ing. Chris van Beekum and ing. Bertjan Maasdam for their valuable support and advice during this work. I am also grateful to my lectures at HAN for their encouragement, patience, and extremely valuable advice during my studies. I also thank the courses coordinator, ir. Edwin Tazelaar for his encouragement.

I am also grateful for all the help and support I received from Smurfit Kappa Roermond papier, and especially the process technologists' staff Marcel Gelissen and Bart Stienen for leading me to the laboratory experiments. Further thanks go to Wim Muller for his support in paper quality measurements.

I would also express my gratitude to my friends for giving me a sense of balance that allowed me to define a reasonable scope for this thesis.

Cyprien Nsengimana

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLS AND ABBREVIATIONS

SYMBOLS

A	Area [m^2]
Bust	Bursting strength [kPa]
C	Correlation matrix [-]
C_E	Correlation matrix for 6-dimensional approach [-]
CMT30	Concora medium test [N]
C_{pv}	Correlation matrix for 10- dimensional approach [-]
dv	Speed difference between first and second roller [m/min]
E	X-block error matrix [-]
q_i	X-block error vector [-]
Emod	Elastic modulus [N/m^2]
F	Y-block error matrix [-]
F_t	Tensile force [N]
G	Grammage [g/m^2]
i	index variable [-]
M	Moisture [% wt]
M_p	Mass of specimen of paper
n	Number of observations in data matrix [-]
$P_{m \times k}$	Orthogonal loading matrix, X [-]
P_2	Power supplied to the web by poperoller [kW]
P_s	Power of paper web stretch [kW]
$Q_{m \times k}$	Orthogonal loading matrix, Y [-]
Q_0	Cost function [-]
r	Radius of the cylinder[m]
R^2	Coefficient of determination of fit [-]

SCTcd	Short compression test [kN/m]
$T_{n\,x\,k}$	Projections of X-scores [-]
t	latent factors for input matrix [-]
T	Thickness [µm]
T_p	Temperature of the paper [T_p]
$U_{n\,x\,k}$	Projections of Y-scores [-]
u	latent factors for output matrix Y [-]
V	Eigenvector of the correlation matrix, C [-]
V_1	Tangential speed of the first roller [m/min]
V_2	Tangential speed of the second roller [m/min]
w	PLS weight vector [-]
W	Work done by couple [J]
W_p	Width of paper web [m]
$X_{n\,x\,m}$	n x m data set consisting of n observations on m variables
x	input variable [-]
x’	Transpose of input variable [-]
$\bar{x}$	Mean value of x [-]
$\hat{x}$	Prediction of input variable [-]
x_i	i^{th} column of data set matrix, $X_{n\,x\,m}$ [-]
$Y_{n\,x\,q}$	n x q data set consisting of n observations on q variables
y	output or response variable [-]
$\hat{y}$	Output prediction [unit of variable]
β	regression coefficient [-]
$\hat{\beta}$	Regression coefficient of the predictors [-]
δ	stress [N/m²]
ϵ	Strain [-]
θ	Axial angle of rotation [rad]

ρ_p	paper density [kg/m^3]
σ	Standards deviation [unit of variable]
σ^2_{lab}	Variance due to the laboratory equipment [squared unit of variable]
$\sigma^2_{process}$	Variance of the process measurements [squared unit of variable]
σ^2_{sensor}	Variance of the sensor [squared unit of variable]
σ_{MD}	Standards deviation in MD direction [unit of variable]
σ_{res}	Residual standards deviation [unit of variable]
[%wt]	The percentage of wet water

ABBREVIATIONS

ILSE	Inline measurement sensor and E-modulus
BS EN	European standards adopted by UK Government
C.I.	Confidential interval
CD	Cross machine direction
DIN	German standards organization
IR	Infrared radiation
ISO	International Organization for standardization
KCPK	Kenniscentrum Papier en Karton
Lab	Laboratory
MD	Machine direction
MSE	Mean square errors
NCOMP	Number of principal components
PCA	Principal Component analysis
PC_i	i^{th} Principal component.
PCTVAR	Percentage of the variance
PLS(R)	Partial least squares (regression)
RF	Recycled fluting paper grade
SCAN	Scandinavian group of standards
SI	The international standard unit.
SKRP	Smurfit Kappa Roermond paper mill
SS_{err}	Sum of squares of Errors
SS_{reg}	sum of squares regression
SS_{tot}	Total sum of squares
TAPPI	Technical Association of the Pulp and Paper Industry
TL	Test liner paper grade
vs	Versus

1. INTRODUCTION

1.1 PROJECT BACKGROUND

ILSE means a real time prediction of quality of paper while it is still running on paper machine. Its acronym is 'Inline measurement with sensors and E-modulus '. It is a patent project owned by Bumaga.BV for KCPK, Daughter Company of Kenniscentrum Papier en Karton (KCPK). This project was conducted at Smurfit Kappa Roermond Paper Mill, SKRP. The pointed factory uses recycled papers as the main source of pulps which are processed with other additional chemicals such as starches, fillers etc. in order to make papers. Roermond paper hosted this project for testing and measurements because the successful implementation of the project would benefit their factory. The idea behind this project was to optimize the measurement process whereby the prediction of the paper quality would permit the paper maker to adjust the raw materials of the process inputs based on characteristics of the quality outputs. The knowledge of the quality prediction of the paper on or closer to the spec would avoid the unnecessarily higher amount of raw materials that is mostly used when there is no direct control of the paper quality online. The quality control online would satisfy the customers' needs and therefore boosts the welfare of the company in general.

1.2 PROBLEM DEFINITION

The problem addressed in this research topic is that the paper quality is not always on target spec. The higher spec product than necessary creates high investments in the purchase of raw materials and thus results in less profit. In the other way around, off spec product degrade the quality of paper and causes lesser consumer and less profit. It is noted from the patent [1] that when the paper is on the paper machine and running at high speed the specific strength variables we are able to measure can be related to several relevant final paper quality. The number of available sensors and actuators on factory for measurements are summarized in Table 1.1. The description of the main parts of the recycled paper industry is given in the Figure 1.2.
The major parts include the wire or forming section, the press section, the size press, and the drying section. As pointed out by Ajit K. Ghosh, 2001. The pulp slurry from the head box enters the wire section with 0.1-1% of dryness. At this point, the big part of water is removed by high gravity run of the wire section up to 25% of the paper dryness. The dryness of the paper increases to 40 - 55% during the wet pressing by mechanical compressions. The drying section is therefore designed to dewater the paper between 90 - 95% [2], [9], [10]. The thermal energy from steam-heated cylinders is used for the dewatering in the dryer section and the application of heat transfer is operational in this stage of papermaking process.

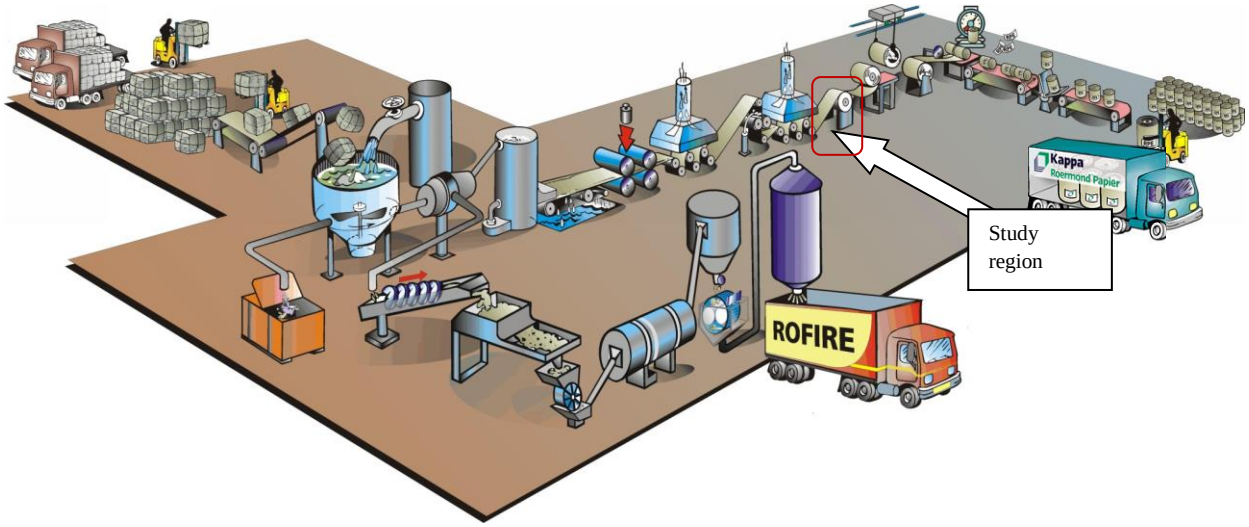
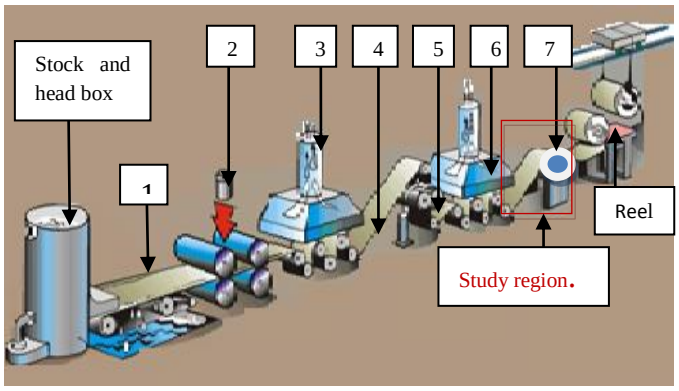


Figure 1. 1 Smurfit Kappa Roermond paper mill, SKRP.

The different parts for processing the recycle papers

The selected region on the Figure 1.2 describes the region where the paper has reached its end properties for the quality prediction.



Legend:
1: Wire or forming section
2: Press section
3: First group of drying cylinders
4: Paper web guided by dryer fabric
5: Size press section
6: Last group of drying cylinders.
7: Pope roller position.

Figure 1. 2 The recycled paper mill showing different sections.This figure was slightly modified from Figure1.1.

The part including the last drying cylinder speed referred to V_1 and the poperoller speed, referred to v_2 are the sections of the selected study region.
During the paper manufacturing the poperoller power, P_2 provide the energy that goes into the paper to pull it or to stretch it and the elongation which created in the pull of paper web is due to the speed difference between V_2 and V_1.

1) Goal(s) for control	1) Online prediction of the paper qualities based on E-modulus [1]
2) Process outputs (offline with Paper Lab machine)	1) Measured qualities of paper samples: (Lab paper testers: burst tester, compressive strength tester for SCT & CMT) Bursting strength CMT30 SCTcd
3) Control input(s) (with actuators& sensors)	1) The last drying cylinder speed,V_1 (Actuators: Steam control valves, condensate valves) 2) Poperoller speed , V_2 (Speed sensors: Tachometers) 3) Poperoller power , P_2 (Measured by Wattmeter) (Actuator: Poperoller motor drive) 4) Width of paper, W_P (IR Protagon sensors for grammage, moisture, width of paper) 5) Thickness of paper, T (Experimentally obtained from grammage sensor results) 6) Calculated E-modulus , E
4) Major disturbances	1) Temperature of the paper , T_P (Temperature sensors) 2) Moisture content ,M (Moisture sensor)

Table 1.1 Summary of the identified process inputs-outputs

1.3 OBJECTIVES

1. Identification of the process variables responsible for elastic modulus.
2. Identification of the paper quality parameters to be controlled.
3. Finding the correlation between elastic modulus and quality parameters
4. Make correction of correlation with moisture and temperature
5. Implementing the manual as the guidance model for the other paper mills.

1.4 PROJECT APPROACH

The research project is mainly focused on the description of the patent [1], and the samples for measurement were done based on TAPPI and ISO standards. In addition, the ideas from paper professionals and both the site and school supervisors were collected in order to effectively come up to its feasibility study. The thickness measurement was a tremendous challenge since it requires a lot of attention, accurate and speed sensors for measuring the paper.
The experimental sample tests in the laboratory and measurements of temperature of the paper on the factory gave much contribution in the data selection. Therefore, the data was collected for further analysis based on the identified process variables.
The collected data include both the online process data and offline quality measurements by paperlab machine. The measurement of paper samples in the laboratory set up were used for thickness measurement in order to identify the process variations that are occuring on the plant. Finaly, two techniques for dealing with multivariable regressions known as principal components analysis, PCA and partial least squares regressions, PLSR were adopted as the best methods for data analysis using the available software tools like matlab®.

1.5 OUTLINE OF THE THESIS

The report describes the method used to implement the inline measurement of paper web strength and control. The structure of this thesis is organized as follows:The first chapter discussed on the project introduction. The second chapter describes the literature survey and the paper review about different ISO standard for measurement methods that are used to define the stiffness of paper. The third chapter is reviewing the patent model to find elastic modulus and temperatures of the paper web. The fourth chapter discuss on the model development based on both PCA and PLSR results. The fifth chapter is devoted to the conclusions and recommendations of the investigation . The last two chapters will respectively discuss on the references and appendices.

2. REVIEW OF PAPER LITERATURE

2.1 PAPER DEFINITION

A paper is a fibrous material, i.e. it contains a large number of pores filled with air. The porous structure makes the paper stiff. The structure of the paper may differ from different manufacturers due to the methods and the materials used to make it. Generally, the paper may be considered as a composite containing fibers and a matrix consisting of filling materials and other paper-auxiliary agents. The strength of the well-bonded fibers determines the tensile strength of the sheet. The most appropriate fibers of paper production are fibers of length from 1-5 mm [3]. In many cases both in practical applications and in laboratory tests, paper is considered as an elastic material. Elastic modulus is one of standardized mechanical properties as determined for the paper. It is defined as the ratio of stress to the strain. The strain is referred to the change in some spatial dimension such as length, angle or volume and the maximum of its value determines the percentage of stretch that the paper can achieve [4], [5]. The stress is therefore referred to the force applied to a surface to cause the strain. The ability of paper to resist to an applied force is known as paper rigidity. The paper rigidity is different from paper stiffness as the ability of paper to support its own weight. The three basics types of stress create the corresponding three types of elastic moduli. The Tables 2.1 below summarizes the three types of moduli. The SI unit of stress is the Newton per square meter or Pascal [4], [6].

Modulus	Stress type	Strain type	Configuration change
Young's[E or Y]	Normal to opposite faces	Linear(length)	Longer & thinner or Shorter& fatter
Shear [S or G]	Tangential to opposite faces	Shear or tangent of angle	Rectangle becomes parallelogram
Bulk[B,K]	Normal to all faces	volume	Volume change but shape does not.

Table 2. 1 Different types of elastic modulus

2.2 PAPER FORMAT, THICKNESS, AND GRAMMAGE

Paper format, thickness, and grammage are the basics characteristics of paper and paperboard. The paperboard is defined as the paper with grammage higher than 224 [g/m²]. The maximum grammage paper produced at SKRP paper mill is 220 [g/m²].The grammage is defined as the weight in gram per unit area of the paper or paperboard. Grammage and thickness are significant properties in sale and use of the paper product for most of the paper mills.

The format of paper means its dimensions in terms of width and length. ISO 216 and literature [5] give more details about the paper format and size.
The vertical distance between the surfaces of paper defines the thickness.

The thickness is measured by micrometer in laboratory. Its measurement is a key characteristic in the assessment of the paper product quality.
The Figure 2.1 shows the paper strip with grammage and thickness measurement unit [4]

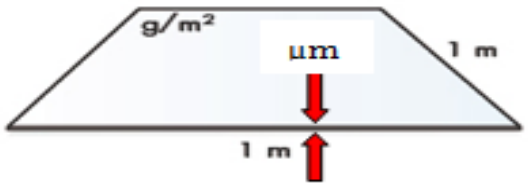


Figure 2. 1 Units of measurement of grammage and thickness. Figure is slightly modified from [6].

Standard methods	BS	ISO	TAPPI	OTHER
Grammage of paper and board	BS EN ISO 536	536	T410	SCAN P6; DIN 53104
Thickness of paper and board	BS EN ISO 534	534	T-411, T-511	SCAN P7 and P47; DIN 53105

Table 2. 2 Standard references for testing grammage and thickness of the paper [4]

2.3 TENSILE STRENGTH

TAPPI tests T-404 and T-494 describe the tensile strength properties of paper. The paper strength properties determine the permanence and durability of paper against its exposure to the environment deteriorations. The example of strength properties are tensile strength, tearing resistance and bursting strength [5], [7]

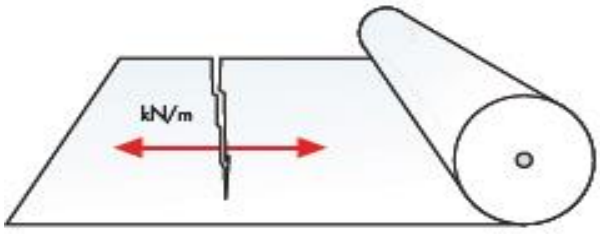


Figure 2. 2 Paper tensile breaking strength [5]

2.4 PAPER GRADES AND ELASTIC MODULUS

It has been found from [8] and [9] that the elastic modulus is higher in machine direction (MD) and is lower in cross machine direction (CD). The measured values of elastic modulus in the laboratory are shown from the listed literatures.

The Tables 2.3 gives the results computed from laboratory experiments with two types of papers that are tesliner and recycle fluting papers.

Those results was found by using nonlinear vibrations of beam like model of an axially moving paper web by Kapitaniak and Marynowski, 2007 [3], [8], [10].

Parameter, notation	Test liner paper	Recycled fluting paper
Thickness [mm]	0.347	0.241
Grammage [g/m^2]	220	123
Density [kg/m^3]	634.01	510.37
Young's modulus, Emod $_{MD}$ [N/m^2]	$5354x10^6$	$3373x10^6$
Young's modulus, Emod $_{CD}$ [N/m^2]	$2285x10^6$	$1863x10^6$

Table 2. 3 Physical paper parameters with corresponding elastic modulus. Results extracted from [3] by (Szewczyki, 2006).

2.5 MOISTURE CONTENT

The moisture content of pulps and paper is economically important for its end use of paper product, since the paper is sold on its air-dry basis. Researchers say that the moisture content affects many paper properties such as dimensional stability, physical strength, paper runnability and particularly in printability where printing is executed. Since moisture affects the strength property of paper [4], this makes it an important input for the model correction [7].

2.6 BURSTING STRENGTH

Bursting strength is the most measured strength property of a paper. Burst is the official test used for measuring the paper strength with thickness less than 0.6 mm. It is also known as Mullen Test. It is measured as the maximum hydrostatic pressure required for rupturing the paper sample by constantly increasing the pressure as per TAPPI T-403 om-02 standards [12]. In the laboratory, the burst is determined by holding the paper sample between two small circular clamps whereby it is subjected to an increasing pressure via rubber diaphragm that is expanded by hydraulic pressure at a controlled rate until the paper breaks or ruptures. The SI unit is kilopascals [kPa]. The instrument that commonly used to measure burst is called burst tester known as Mullen. Details in ISO2758.
Bursting strength was found to increase due to an increase of grammage, sheet densification, and degree of pulp refining [13].

2.7 SHOT SPAN COMPRESSION TEST (SCT)

It is a method developed by Christer Fellers in the late of 1970's [14]. This method has been found be useful to measure the stiffness of linerboards. Literature from ISO 9895 & TAPPI T-826 found that SCT differs from other crush tests such as ring crush test, in some crucial respects.
0Mostly, the test strip of 15 [mm] width is clamped at both ends, and the compression span is very short. The effect of buckling is more pronounced with lower basis weight of the sample.
It will measure the true compression strength. The SCT value is the maximum compression force per unit of width of the test strip and it is expressed in [kN/m].

2.8 CONCORA MEDIUM TEST, CMT30

CMT-30 is used to measure the crushing resistance of the laboratory-fluted strip of corrugating medium. The fluted shape of the test piece is retained by adhering a length of pressure sensitive tape onto the flute tips in a special jig. This form a piece of single face corrugated board. It is placed on its face between the tester platens and crushed until failure. CMT value is the maximum force sustained by a test piece containing ten flute tips expressed in Newton, [N]. The compression may occur after the conditioning period of 30 min or immediately (CMT0) after corrugation. [SCAN-P27], [EN ISO 7263], TAPPI Test method T-809 om-06, Flash crush of corrugating medium (CMT test) [15].

3. DESIGN PROCEDURES TO FIND E-MODULUS

The design procedures used to find E-modulus are derived in the Appendix A and all derivations are based on the patent structure [1].The study region of the process as it is shown in the Figure 1.1 is extended below for analysis of the process variables.

3.1 PHYSICAL DESIGN OF THE PAPER WEB STRETCH

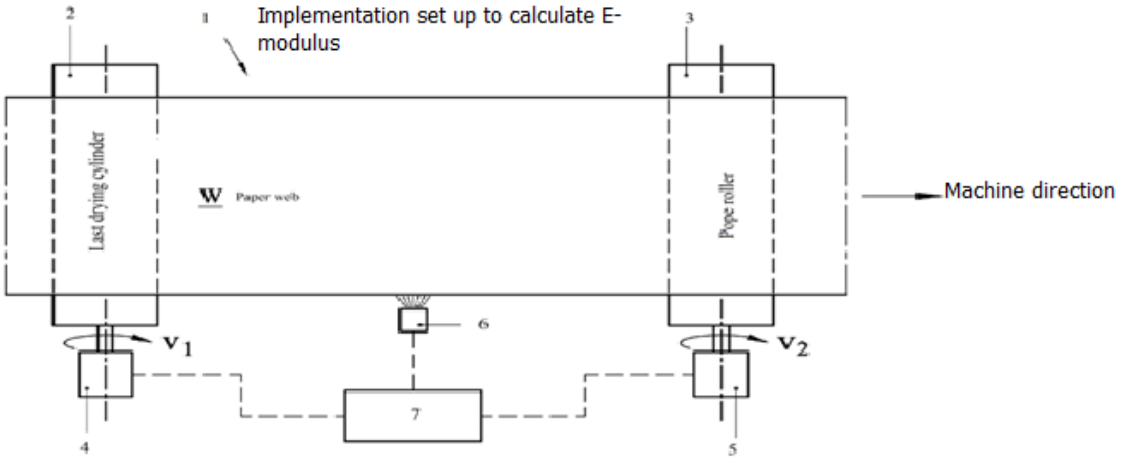


Figure 3. 1 Determination of Young modulus of paper web. Source of picture [1].

Legend:

1: Young modulus determination set up.
2: Control section of the first roller.
3: Control section of the second roller.
4: Tangential speed of the first roller, V_1
5: Tangential speed of the second roller, V_2 and P_2 is the power supplied to web by second roller.
6: Sensor to measure the thickness [µm] and width of the paper web [mm].
7: Controller designed to collect the said parameters and to determine that young modulus is within predetermined tolerance range.

The resulting equation for elastic modulus is given by $\text{Emod} = \frac{P_2.V_1}{A\,.\,V_2\,.\,(V_2-V_1)}\ [\text{N}/m^2]$ (3.1)

The equation (3.1) is the basis of my research, where its role to quality contribution has to be proven in practice. The proof starts by finding all the necessary measurements of the involved process variables.

3.2 SENSORS FOR THICKNESS MEASUREMENT

The E-modulus formula shown above in the equation (3.1) involves the area of the paper web, which is the product of paper width and thickness. In order to have its measurement a number of available sensors were discussed in order to provide the actual thickness of the paper web.
So far, three types of different thickness measurement sensors have been identified based on how accurate they would perform.

The first type of thickness sensor is working on a moving paper web whereby the paper web passes between two sensor's fingers on its both sides during the web run. One of the sensor sides has a coil that can send magnetic pulses through the paper.
The difference in heights, one height when there is no paper between, and another height when the paper passes between two sensor's fingers gives another frequency in the signal.
This can be done up to 2 [mm]. There will be an offset between the fingers of the sensor and the laboratory equipment because of the pressure in the sensor's fingers may be much lower than the one in the laboratory equipment.
More pressure on the paper web in the factory would grab the paper and cause a web break. This may be taken as its drawback that leads to another sensor's choice.

The second type of thickness sensor investigated through the application of the hovercraft principle. On one side, usually the bottom side, the paper goes over a metal plan with a vacuum that ensures that the paper will stay flat on the plate. A measuring head (same method as the finger sensor), which is a disk of 3 cm in diameter, has a small and constant flow of air coming on the sides which lets it float over the paper web. Because of the larger measuring surfaces in comparison with the fingers sensor, its accuracy is lower. The recycled paper could cause a build-up of dirty and stickiness of the bottom plate and this would cause a higher measured thickness. The biggest problem would be when a large bump in the paper goes under the sensor. It will lift the head on its one side, thus destroys the hovercraft principle, and presses the backside in the paper by causing a web break. This is the hovercraft principle drawback. The pros of the above two type of sensors are their relative cheap prices.

The third option is a laser sensor. The laser looks to the position of the reflected laser beam. If the angle is always the same, the distance of the laser to the paper determines the position of the reflection. To make sure the paper always starts at the same height a frame support on the horizontal plate would be installed with the vacuum distance. The reflection of the laser on the plate is the zero calibration.
The hardware structure is given in the Figure 3.2 below:

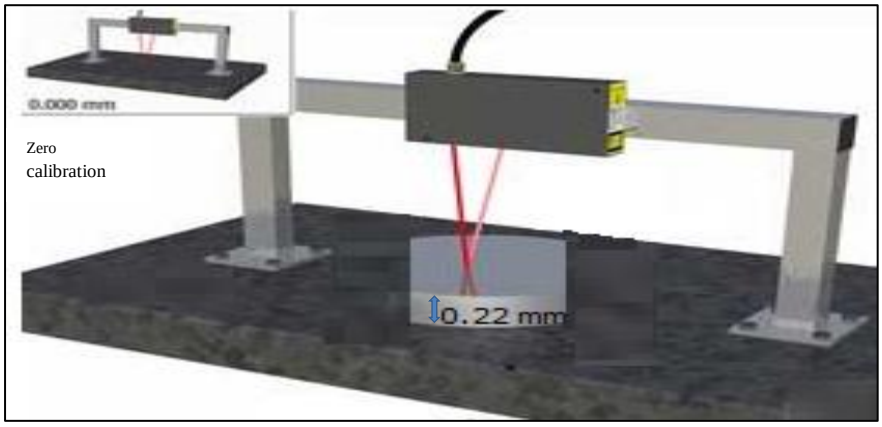


Figure 3. 2 Single sensor approach for laser thickness measurements (Source [16]).

The problem with this sensor is the same as other two sensor types. The recycled paper may have dirty and stickiness built up on the sensor's plate that is supporting the paper and therefore introduces false results.

The solution for that sensor is to use two opposing sensors to measure a single dimension because there is no reference surface for this case. The orientation of the laser sensor to the target surface of paper web must be aligned perpendicularly when integrating laser sensors. This is because the improper alignment will fail to measure a true differential thickness of the paper web. The Figure 3.3 shows the hardware structure where both sensors are synchronized together in order to measure the distance to both sides of the surface at the same time. Due to the vibrations of paper during manufacturing, the distance between the lasers is varying. It is therefore necessary to synchronize the laser sensors using hardware or software triggers to capture near – instantaneous distance readings on moving paper surfaces.

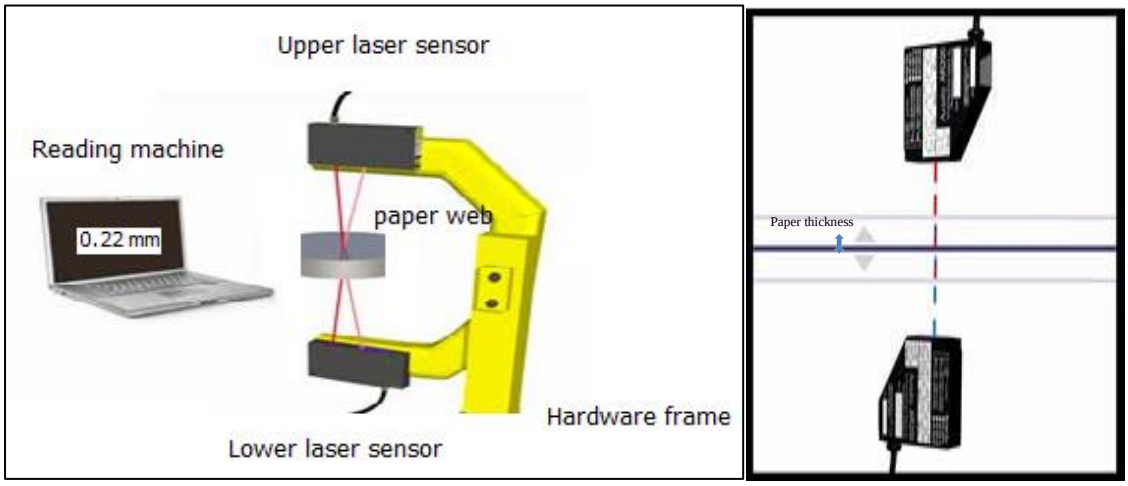


Figure 3. 3 Dual lasers approach for thickness measurement (source [16]).

3.3 ESTIMATION OF SENSOR ACCURACY OF THE THICKNESS MEASUREMENT

Thickness measurement of sheet of paper is really a challenging situation because its measurement size is small. It is in order of certain micro. This micro distance varies from one paper type to another. On the paper factory the papermaking process, need a thorough study and examination of the different variations of the paper along with paper machine direction (MD) and cross direction (CD).

In order to accumulate all the possible process variations on the paper machine. The expected deviation formula was developed. It summarizes all the possible variations one can meet on the factory by determining the thickness sensor accuracy .The formula was derived in order to determine how good or bad would be the thickness sensor online. The formula took into considerations the variability due to both the measured process variable such thickness and the measuring instruments.

The variance in the process as far as the thickness measurements is concerned is mainly influenced by the factory environmental conditions such as the moisture content ,the temperature change as well as the different procedural steps in the paper making process.

If one takes a single sample of paper, a rise question is what will be the chance that it would represent the average value of the whole jumbo? Since it is practically impossible to manually measure offline the thickness across the whole width of the manufactured Jumbo reel and therefore make an average that estimate the value that would be measured by the sensor while it is put on the factory, the developed formula is describing what one would expect for a picked up sample for thickness measurement. It is important to know how much is the process variations before choosing the particular type of sensor online. The simple way to know the process variations is to test samples in laboratory. Due to the accuracy of the laboratory equipment, the expected variation per given sample may result in the following formula:

$$\text{Expected standard deviation } \sigma \text{ per sample } = \sqrt{\sigma_{lab}^2 + \sigma_{process}^2} \quad (3.2)$$

where

σ_{lab}^2 represents the variance of the laboratory equipment.

$\sigma_{process}^2$ represents the variance of the process variable.

The laboratory measurements have good accuracy due to the standards environmental conditions that are stable within the measuring room. The maximum variance of the laboratory equipment is approximated to 1 μm by referring to deviation of the L & W micrometer for thickness measurement.

It is important for the process deviations to stay within the acceptable tolerance range with respect to the target production of the quality specs.

The expected deviation of the process together with the additional sensor deviations judge on how good thickness sensor would be on line. The expected deviation of the process online is given by the formula:

$$\text{Expected standard deviation on line is given by } \sigma = \sqrt{\sigma_{sensor}^2 + \sigma_{process}^2} \quad (3.3)$$

where σ_{sensor}^2 expresses the total variance of the sensor.

The online laser sensor is proposed for thickness measurement because of the lower sensitivity to dirt. The study proposes two sensors where their variance may be approximated to 2 [µm]. The upper sensor can estimate the distance from the sensor to the paper and the lower sensor estimate the other bottom distance from the paper to sensor and by knowing the total distance between sensors, which is determined by the sensor frame, the thickness of the paper, can be calculated.

The process may have some extra variations to the expected deviation between lab and online measurement. This is due to the possible appearance of unstable environmental conditions on the factory that are influenced by many factors including the temperature variations, moisture content and other process variations like raw material and fiber treatment. It was observed from the thickness measurements that the smaller process variations result in the better-expected measured quantity.

The study from thickness sensor proposes a fixed-point sensor somewhere around 50 [cm] from the backside of the machine due to the availability of small variations and the available space on the factory. The 2σ would be considered for each variation in order to estimate the 95% confidential interval for normal distribution of measurements samples.

The Table 3.1 shows the average summary of the thickness measurement across the width of two Jumbo reels with different grammages.

The width of the paper web is online measured and its maximum reaches up to 5 [m]. The width was split into horizontal four strips for easy measurements of samples and each strip size was around 5 [m] of length across the width of 20 [cm].

The first paper sort of grammage of 170 $[g/m^2]$ gives the overall average of thickness of 258.5[µm] with the average standard deviation of 5 [µm].

The measurements of samples were done by keeping a distance of 2 [cm] between each consecutive measurements until the four strips were done.

The second paper sort of grammage of 200 $[g/m^2]$ gives the overall average of thickness of 288.7 [µm] with the mean standard deviation of 8.3 [µm]. The measurements of samples were done by keeping a distance of 5 [cm] between each two consecutives measurements until the four strips of the Jumbo reel were done. The table below shows the results:

Paper sort 20.170.45	Measurements variables	Average value of strip 1	Average value of strip 2	Average value of strip 3	Average value of strip 4	Overall average.
Jumbo reel 1	T [µm]	259.01	259.4	259.2	256.2	258.5
170 $[g/m^2]$	σ [µm]	3.7	4.7	5.1	6.4	5.0
50.200.64						
Jumbo reel 2	T [µm]	289.17	287.9	291,64	285.2	288.7
200 $[g/m^2]$	σ [µm]	7.0	7.4	11.2	7.2	8.3

Table 3.1 CD Thickness measurements for two paper grades.

The main objectives of the CD measurements were both to identify the process variations across the width of the paper and to know the best position for MD samples. At this region is where the sensor would be installed due to the appearance of uniform process variations.

The Figure 3.4 below gives more details of the measured paper grades.

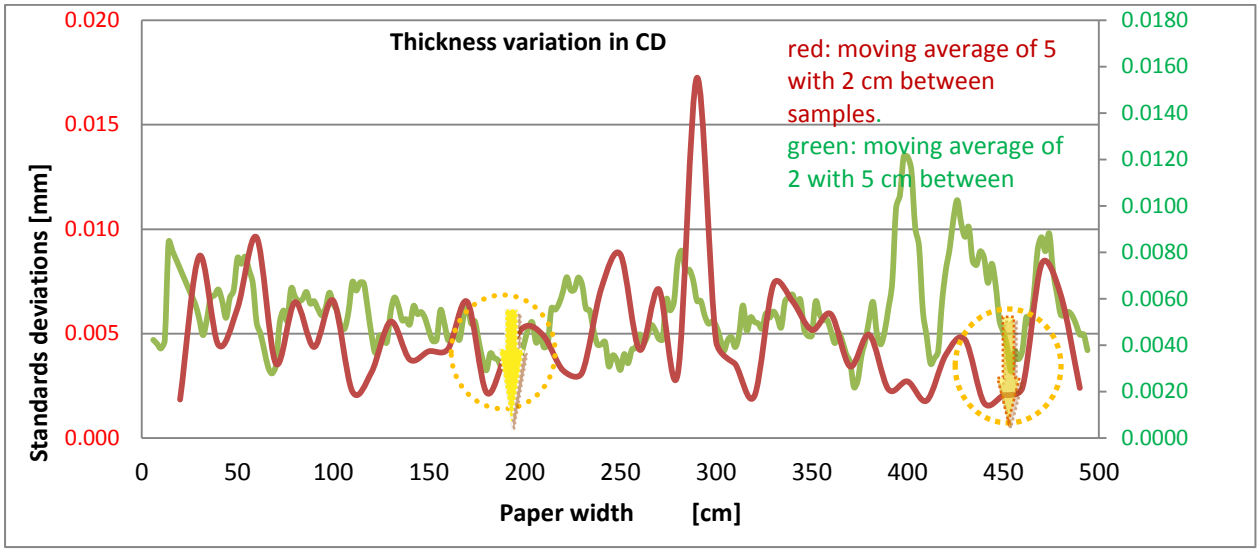


Figure 3. 4 CD thickness variations for two paper grades

The pointed regions are the uniform positions for the thickness sensors in CD direction.
The best position of MD variations on the machine would be at those two pointed regions.
The MD samples taken at presented region may lead to the uniform thickness variations.
Since it may be difficult to install the thickness sensor in the middle of the paper web, at 200 [cm] from the front side of the machine, the last point at 450 [cm] from the front side of the machine is practically a suitable position for the thickness sensor.

3.5 MD THICKNESS VARIATIONS OF THE PAPER WEB

The MD thickness measurements at the best point of CD variations were measured by using L&W micrometer. It is unsuitable to measure the whole reel of paper produced, which is around 5-10 tons. Yet, it is important to know how good the thickness variations on the chosen area would behave. Therefore, a rule of thumb was applied to measure MD thickness samples. The longest wet part of the machine felt was taken as a guide for choosing MD samples because it is the part where the papers can manifest lot of variations in the thickness. The total length of 70 [m] corresponding to 1.5 times the longest part of the felt was measured. The samples were measured by keeping 5 [cm] between every two consecutive sample points.

These leads to 1400 measurements. The results from the experimental data obtained during measurements are shown in the Figure 3.5.

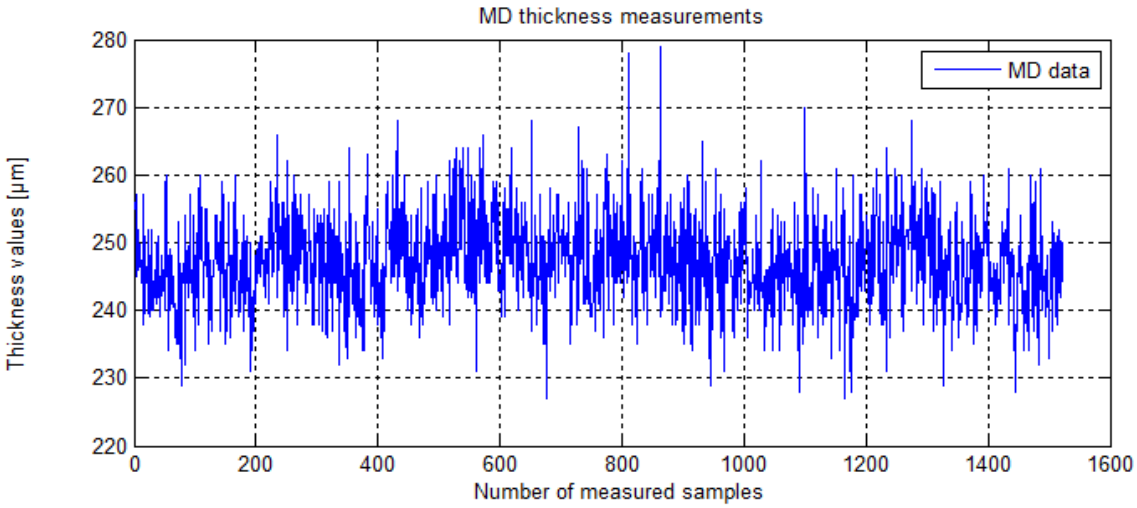


Figure 3. 5 MD thickness variations of paper web

If one is interested in the distribution of thickness data in MD direction the Figure 3.5 can also be represented in histogram by using matlab® command,>>hist(data,40) where 40 is the scalar showing the resolution used for better visualizations. The command above produces a histogram bar plot of the results as follows:

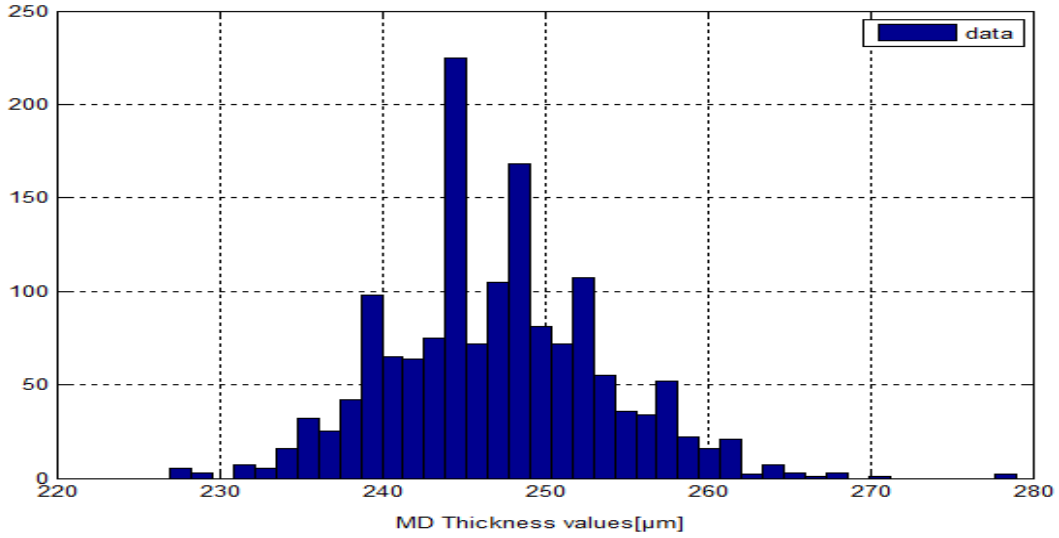


Figure 3. 6 MD thickness variations of paper web in a histogram bar plot.

The Figure 3.6 shows that there are many variations in the process. Those variations may be due to the dynamics in the factory. As pointed out in problem statement, the paper manufacturing process is a large dewatering process that involves lot of steps during paper drying process.
The paper web run at high-speed 900 $[m/_{min}]$ and it is possible to have many process variations across different paper stages. It is known that the force applied on the wet part of the machine cause deformation instead of elasticity.
A force when it is applied on the dry paper causes a stretch that can be undone if it is within the elasticity range. The latter is defined as the maximum stretch whereby the paper can return to its original shape without any change in its physical property.

-The following statistics summary is obtained from the loaded data in matlab® command window.
-The Figure 3.6 can show clearly the value of the minimum, maximum, and the mean without calculating them. This is the benefit of histogram plot.
-The minimum measured value (>>min (data)) was 227 [µm] and
the maximum value (>>max (data)) was 279 [µm].
-The mean of MD thickness measurement (>> mean(data)) was found to be 247[µm]
-The deviation of the process,σ or (>> std(data)) in MD direction was 6.64[µm]
-Therefore, two times σ gives 13.28 [µm], which represents 95% of C.I in the normal distributions.
-The total expected deviation of the process includes the accuracy of the lab equipment of $\sigma_{lab} = 1[\mu m]$.
This value is the acceptable deviation of the L&W micrometer for thickness measurements.
-Two σ_{MD} per one measured sample in the lab is given by equation (3.2) above and the result is shown below:

$$2\sigma_{MD} = \sqrt{1^2 + 13.28^2} = 13.40\ [\mu m] \text{ or } \sigma = 6.7\ [\mu m].$$

The resulting conclusion of MD thickness gives the maximum expected deviation that would be obtained when a sensor is installed online. We would expect it to have not more than 10% of the same deviations as in the lab experiment. The paper production that retains the production within maximum allowable deviation ensures that the quality of the paper stays within specification limit. It is undoubtedly known that the thickness always goes high for high grammage paper and goes low for low grammage of paper but the variations in process retains. Assuming the stable conditions on the factory, which may result in stable grammage and stable thickness a relationship existing between grammage and thickness, was derived. It is noted that since the thickness sensor were replaced by the use of grammage sensor that is already installed on the factory, their operating principle were not discussed in this report.

3.5 RELATION BETWEEN GRAMMAGE AND THICKNESS

The linear relationship between grammage and thickness was experimentally derived from the paperlab results. By using the calibration settings, it was possible to linearly predict the thickness measurements based on historical data.
The following points have been examined during the experiment.
-Investigation of the relationship between grammage and thickness by processing offline measurements of the paper web samples.
-Finding the regression line and offset from the thickness and grammage measurements based on the paperlab average results.
-Calculating the residual standard deviation of both the paperlab machine and lab thickness results in order to see how accurate is the thickness measurements for that different equipment.
-Comparing the results from residual standard deviations from the paperlab machine and lab.
-The grammage of each sample was averaged over the three test specimens for 66 samples. The mass was measured by L&W balance, which is accurate to 0.001 mg and the thickness by L&W micrometer, which is accurate to 1µm. We used EN ISO 536 method to estimate the grammage. Samples for grammage were cut using the L&W manual circular sample cutter, model K211, and had circular area of 100 [cm²].
-The grammage was then calculated according to the formula:

$$G = \frac{M_p}{A} [g/m^2] \quad (3.4)$$

where G is the grammage, A is the specimen test area and M_p is the mass of the test specimen. The specimen is defined as the small piece for testing.

The results displayed on the Figure 3.7 below where the grammage is on the X-axis and the thickness is on the Y-axis.

The R^2 is the ratio of the explained variations to the total variations in the thickness measurements was determined. Its value is normally varying from 0 - 1 where 1 is an indication of the perfect fit between measured values and predictions.
The residual standard deviation (σ_{res}) give the difference in variability between observed(y) and predicted (ŷ) values:

$$\sigma_{res} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-2}} \quad (3.5)$$

Where the term $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ is called the sum of squared residuals and n-2 is the degree of freedom of the residual estimates [17]. Notice that for sample variance, n-1 is usually used for computing the variance in the part of the population.

After proving the linear relationship between grammage and thickness by manual measured samples, R^2= 0.96. It was necessary to consider a large number of samples measured by the paperlab machine where the number of samples was around 1523 during the chosen period.

The result from that experiment gives a better linear coefficient, R^2= 0.988 as it is shown on the Figure 3.7 below:

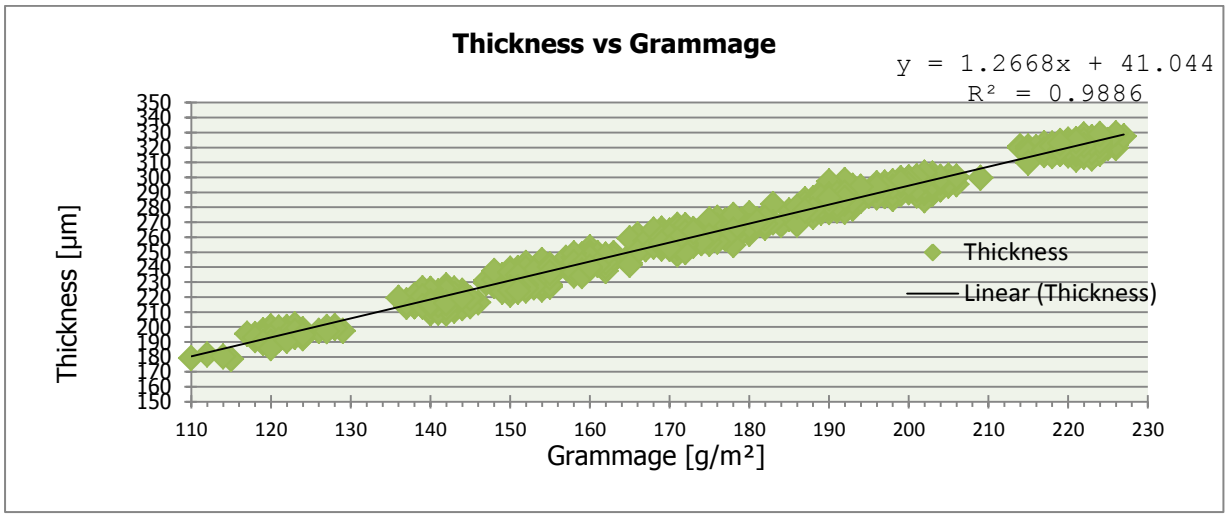


Figure 3. 7 Relationship between thickness and grammage.

A rather accurate linear relationship between thickness and grammage was found. Notice that it is valid under certain assumptions. It is not 100% correct, so minor errors are tolerated. It is assumed that the paper density is the same and the grammage is greater than [110 g/m²] .This choice is based on the type of papers, which are produced by SKRP. It is also noticed that the small deviations can occur from the thickness results and thus provide the wrong results [18]. As pointed out by Cantwell G. Carson and Roman E. Popil, the density (ρ) is related to the grammage (G) and thickness (T) by the following equation:

$$\rho = \frac{G}{T}\,[kg/m^3] \qquad (3.6)$$

As pointed in literatures [18], [19] when the paper density is not constant, the linearity between grammage and thickness is deteriorated or even disabled.

Notice again that this equation is restricted to typical paper produced at SKRP where the minimum grammage is 120 [g/m²] and the maximum grammage can reach up to 220 [g/m²].

The results comparison for manual measured and paperlab machine are summarized in Table 3.2

Number of measured samples	Linear regression	Residual standards deviation
Manually done in lab, n= 66.	Y = 1,486 x + 7,1 R^2 = 0,9648	8.02 [µm]
Measured by paperlab machine,= 66	Y=1,4155x+ 15,716 R^2 = 0,9806	5.38 [µm]
Measured by paperlab machine, n=1523.	y = 1,2668x + 41.0 R^2 = 0,9886	4.10 [µm]

Table 3. 2 Experimental results for thickness predictions

3.6 CONCLUSION ABOUT THICKNESS PREDICTION

In the laboratory, paper thickness is measured between two even plates, which are adjusted with a defined pressure. Because of the paper, compressibility and surface irregularities (inhomogeneity) several measurements have to carry out and the results are averaged.
The thickness of paper is dependent on grammage and specific volume or bulks of paper. The more details about thickness measurements are in ISO 534.
The thickness measurements results from paperlab machine are delayed by a half an hour and thus make it impossible to use them during the in line quality control basis.

Since there is grammage and moisture sensor already installed on the paper machine, it was necessary to investigate the relationship between thickness and grammage. A relation between grammage and thickness was found to be accurate by taking into account the linear regression between them and the residual standard deviation. Based on the linear regression line it was proposed to use a grammage sensor in order to predict the thickness measurements online. The residual standard deviation shows us how accurate is our predictions. If one measures 'y_1 [µm]' thickness value, the accurate value should be [$y_1 \pm 4.1$ µm] within 95% C.I.The thickness sensor is much more needed for the paper mills where the calendering is taken place. Calendering is known as a process of smoothing the surface of the paper by pressing it between cylinders or rollers. This is because smoothing the paper need results in different paper sizes and therefore the thickness also changes.
For that paper mills, Non-contact laser ultrasonic sensors are good due to its sensitivity to fiber-to-fiber bonding and internal drying stresses in the sheet. Optical sensors can also be used but they only look on the distribution and physical layout of fibers.
It has been found from David W. Vahey that sheets with the same fiber distributions with different bonding would measure the same optically but not ultrasonically [20-22].

3.7 TEMPERATURE MEASUREMENT BY HAND LASER SENSOR

The handheld laser temperature sensor was used to give an indication on how much is the paper temperature on the factory. The measurements were taken at a distance of 1 m from the jumbo reel. The same results were compared with the online temperature measurements of the poperoller cylinder that is continuously heated by the hot paper from the steam heated drying cylinder section. The random choices of five sample results are shown in the figure below:

Paper sort	Hand laser temperature measurements	Machine recorded values online
20.120.45	[70.0-72.5]°C	[65-66°C]
20.140.45	[69.4-74.5]°C	[60-66°C]
20.160.45	[70.4-72.5]°C	[67-70°C]
20.175.45	[70.6-73.3]°C	[64-67°C]
50.220.64	[69.4-75.5]°C	[70-72°C]

Table 3. 3 Temperatures results of both hand laser and online measured from the poperoller's drive.

Table 3.3 shows the value in range because temperature is a time varying variable, and each time a new value is obtained. This happen all the time for both hand laser sensor and for online temperatures measurements on the factory.

The hand laser sensor was able to demonstrate the variations of the temperature of the paper on the factory at the end of Jumbo reel. The temperature measurements in the Table 3.3 show the slightly difference between hand laser temperature measurement and online recorded temperatures for the poperoller drive cylinder. The difference is less than 10% and it is noticed that the calibration errors of the instrument were not considered. Those measurements were compared with the temperature in the system that comes from the heat exchanger from the drying steam cylinders to the paper web and from the paper web to the poperoller drive cylinder.

3.8 CONCLUSION ABOUT TEMPERATURES RESULTS

In order to know the influence of temperature of the paper while it is still on the paper mill, a temperature sensor is practically needed. Unfortunately, there was no temperature sensor dedicated to paper measurement during measurement period. Due to this reason, the handheld laser sensor was used to give some indication on the temperature difference between hand laser measurements that are slightly different but closer to the online temperature measurements of the pope roller drive cylinder, it was therefore necessary to take them into consideration in order check their influence on the manufactured paper.

This assumption is only true by assuming the steady state conditions. Those values were used as estimations since they do not exactly correspond to the paper web temperature itself. The measurement of the dryer surface temperatures is practically done to evaluate the efficiency of heat transfer as well as the performance of the steam and condensate system.

A difference of 10°C to 25°C between steam temperature at the operating pressure and the measured cylinder surface temperature is a typical show of the proper operation of the process'' Ajit K . Ghosh, 2011 [2].

Since the variation of temperatures in the paper obtained by hand laser sensor are much closer to the online measured temperatures of the surface of the poperoller cylinder as shown in the Table 3.3, it was quite reasonable to consider their measurements together with other process variables in order to extract their effect on the process.

3.9 THE EXPERIMENTAL FINDINGS OF THE ENERGY THAT UNIQUELY IS USED TO STRETCH THE PAPER WEB

The equation (3.1) derived from literature [1] was found to be true in paper application if one of the variable, P_2 is not the total power of the poperoller, but only the power that is used to stretch the paper (P_s).

The study proved that the total amount of actual power of the poperoller, P_2 in the storage PI system software is not fully used to stretch the paper because there are other losses that are associated with P_2 such as electrical losses, mechanical losses in the motor, gears and bearings, air friction losses, etc.

This reflects to the experimental proof of finding the exact amount of power that is really needed during paper manufacturing process. The idea is to find the feasible means of predicting the paper quality online based on the elastic modulus prediction together with other measurement process variables. The amount of the poperoller energy needed to stretch the paper (P_s) was found to be the difference between actual poperoller power, P_2 during manufacturing time and the power loss on the site, P_L. The poperoller power (P_2) versus poperoller speed (V_2) curve would be done to see the power effect regarding to the speed of the paper web. Since the paper web run at high speed (around 15 [m/s]) a step change of 25 [m/min] was used as the small resolution difference of the speed variations from its production range from $700[\mathrm{m/min}]$ to 950 $[\mathrm{m/min}]$. The relationship between poperoller power and its corresponding speed variations indicates how much is the stretch power needed in the paper production process.
The experiment proposes a way of finding the measurements of the power of the poperoller by varying its speed from 50 $[\mathrm{m/min}]$ up to 700 $[\mathrm{m/min}]$ by keeping a step change of 50 $[\mathrm{m/min}]$ and from 700 $[\mathrm{m/min}]$ up to 950 $[\mathrm{m/min}]$ in the step of 25 [m/min]. Note that the poperoller ran on its own mass and the mass of the Jumbo reel. The idea was to accumulate all the energy consumed in the system rather than pulling out the paper web by the poperoller. The different speeds of the poperoller were manually adjusted and the time for setting each value was approximated to 20[s]-30[s] and then we wait for a while like 2 [min] in order to allow the entire system to stabilize or to adapt to the new speed setting before we change to another speed value. After the system stabilization, one stable value of the power corresponding to the input speed was recorded. The experimental data were automatically saved in the PI system®, due to that facility, it was possible to record the data at the end of the experiment.
The table with the selected experimental data for the poperoller speed and poperoller power is shown below.

V_2 [m/min]	50	100	150	200	250	300	350	400	450	500	550	600
P^*_2 [kW]	1,36 6	2,840	3,518	4,760	6,200	8,123	10,511	12,237	13,375	15,090	16,710	18,615
V_2 [m/min]	650	700	725	750	775	800	825	850	875	900	925	950
P^*_2 [kW]	20,336	22,543	23,481	25,076	25,906	26,762	27,109	28,000	29,099	30,225	31,203	31,679

Table3. 4 Experimental data to find the energy delivered to the paper stretch.

P^*_2 is the special variable of the poperoller power during experiment.

Briefly, the actual poperoller power is including the paper stretch power, electrical losses due to heat and mechanical loss due to motor constructions and the frictions losses due to load it must drive.
By having, the table values between poperoller speed (V_2) and power (P_2^*), the graphs below show the regression line of power P_2 versus speed V_2 based on the obtained results.

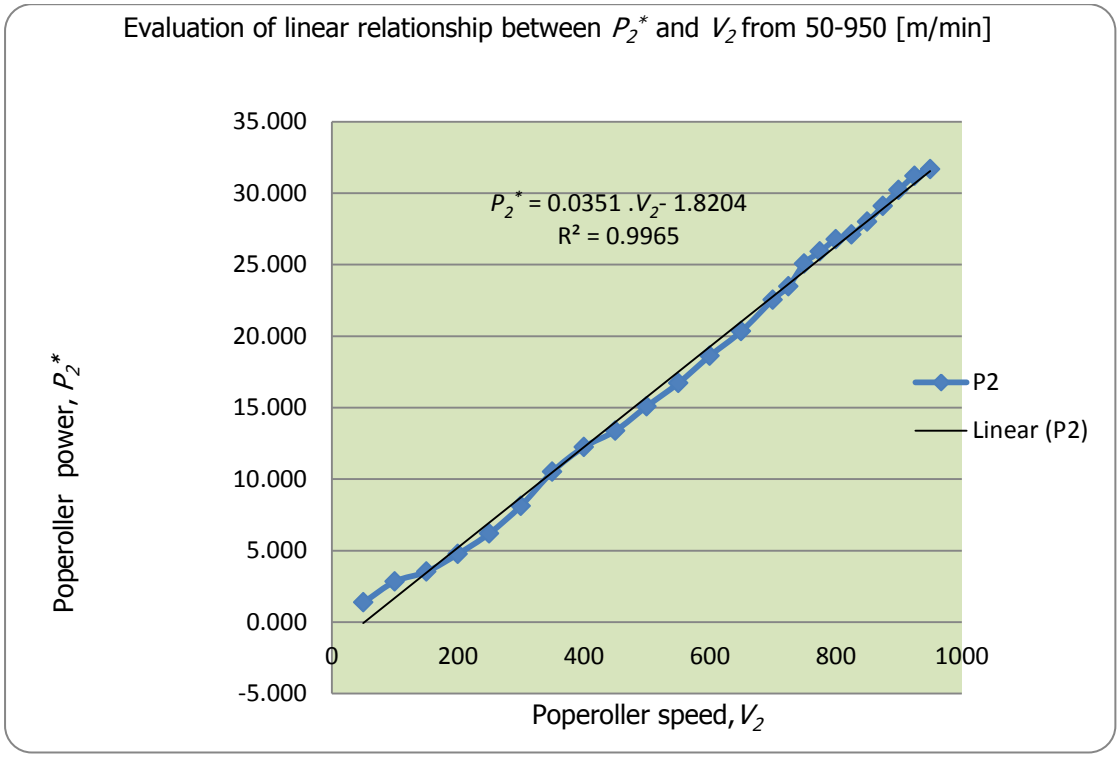


Figure 3. 8 Linearity evaluation of P^*_2 vs V_2 for the speed range (50-950) [m/min].

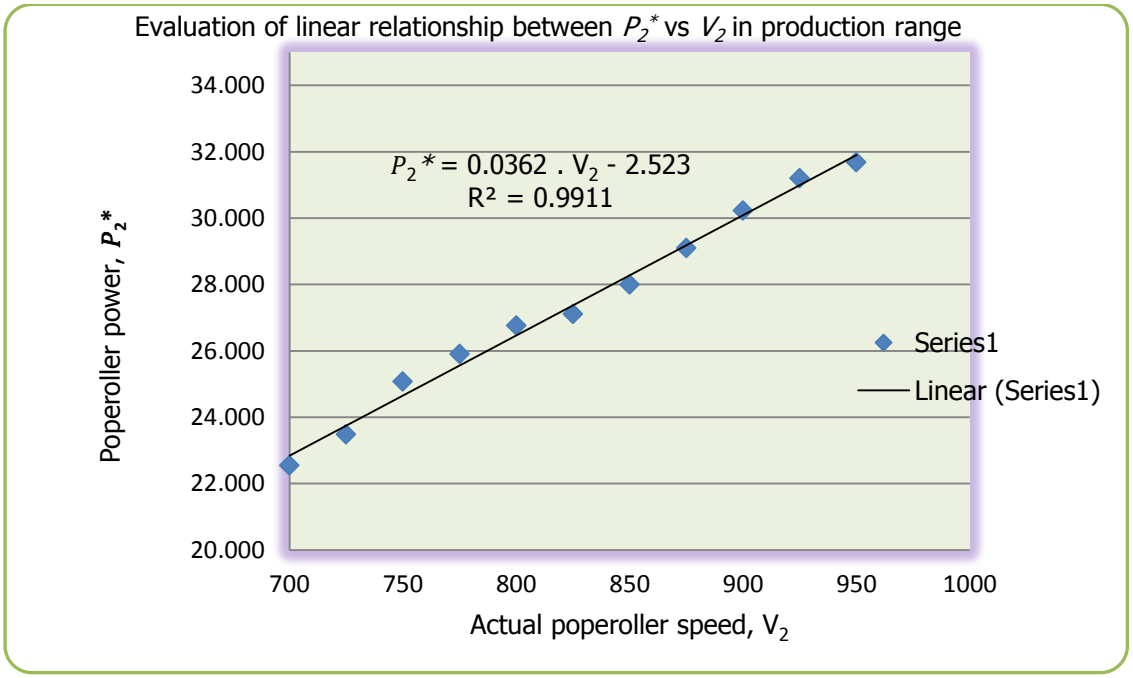


Figure 3. 9 Evaluation of linearity of P_2^* vs V_2 in production speed range (700-950) [m/min**]**.

3.10 CONCLUSION ON THE P_S RESULTS

There is a strong correlation between poperoller speed, V_2 and poperoller power during experiment, P_2*. This is the power losses across the poperoller, P_L. The first order linear relationship between P_2* and V_2 would be used to continuously predict the total energy loss to carry the poperoller load and run its mass.

The following equation was obtained for the speed range varying between (700 - 950) [m/min]:

$$P_2^* = 0.0362V_2 - 2.523 \text{ with } R^2 = 0.9911 \quad (3.7)$$

$$P_s = P_2 - P_L \quad (3.8)$$

where P_L is the same as P_2*, which is the predicted power loss and V_2 is the poperoller actual speed and P_2 is the poperoller power during paper manufacturing process.

The energy going into the paper stretch is now calculated from the difference between actual poperoller power and the predicted power using actual poperoller speed during production and the found linear line.

After finding the uniquely power used to stretch the paper, the equation (3.1) can now be corrected as follows:

$$\text{Emod} = \frac{P_s.V_1}{W_p.T.V_2.(V_2 - V_1)} \ [N/m^2] \quad (3.9)$$

where P_s is the Power to stretch the paper web, W_p is the width and T is the thickness of the paper web.

4. PROCESS MODELLING

This chapter presents the background of the data collection, model construction, and the specific analysis techniques needed to construct a statistical model that describes in details the prediction of the process outputs.

4.1 ONLINE DATA COLLECTIONS

Online data were collected from PI system®. It is Siemens software for storage of real process data during manufacturing process. The production time of one jumbo reel takes around 25 min. This is the time between two consecutives data points in the data collected online. The system facilitates the retrieving from its storage memory. Therefore, by using the time stamps, which is available at the end of reel production, it was possible to filter out the useful data for process analysis. Two minutes before the end of the production time was filtered out in order to avoid the switching to the next reel. The data collected with respect to the process variables that describe the paper manufacturing process based on the patent model [1]. The data used to model are split into two main groups.one group includes the directly measured data such as data for both speeds, V_1 and V_2.The data for paper width, W_p, moisture, M and Temperatures, T_p are also measured. The second group includes the experimentally calculated data, such as elastic modulus, E-modulus, Emod, and thickness of the paper web, T.

4.2 OFFLINE DATA COLLECTION AND SAMPLES CHARACTERISTICS

The offline data are measured by automatic lab machine and checked from laboratory for some unbalanced measurements. At the end of the paper production, a sample is directly sent to the paperlab machine for quality measurements. The sample that is sent to the paper lab machine is about 31 cm of width to (4.5 – 5) m long regarding to the width of paper web under production. According to the design of the paper lab machine, the distance between two measurements is 27 cm, the twelve measurements are taken, and their average is displayed and stored. Different measurements are offline taken but three among are important because they sell paper based on how good they behave. Those include SCT-cd [kN/m] and Bursting strength [kPa] for Testliner paper grade (TL) and CMT30 [N] for recycled fluting (RF). The difference between those two is based on their applications, which may be different. Another difference is their grammage quantity and slightly the amount of raw materials used during paper manufacturing such as synthetic glue and lignin.

The number of Jumbo reel produced online will be equal to the number of each quality-measured offline. This means that the time stamps to consider while collecting all data in the storage system will depend on the time of the Jumbo reel produced. To match the data time stamps between offline and online data, two minutes before the end of production of jumbo would be considered on the process data inputs online. The considered time will match the offline production time for quality measurement.

4.3 PROCESS MODEL CONSTRUCTION

The goal of this thesis is to find the correlation between process inputs and process output whereby the process inputs are fully explained in the patent Bumuga B.V[1].To achieve the goal a model was developed on the Figure 4.1 below.
The model provides the necessary input -output relations by using software tool like matlab®.
Two techniques are used to model and help to find the available correlation coefficient between the variables. The first technique is Principal component analysis; PCA that used to provide the necessary number of principal components that describes the most of the variance in the data set.
The second powerful technique is partial least squares, PLS that used to extract the maximum covariance between process inputs and process outputs and therefore find the existent correlation between them. The relationship study between predictors and responses using linear least squares regression is illustrated in the block diagram below. The predictors are known as the inputs of the process and the responses are represented by the qualities measured offline.

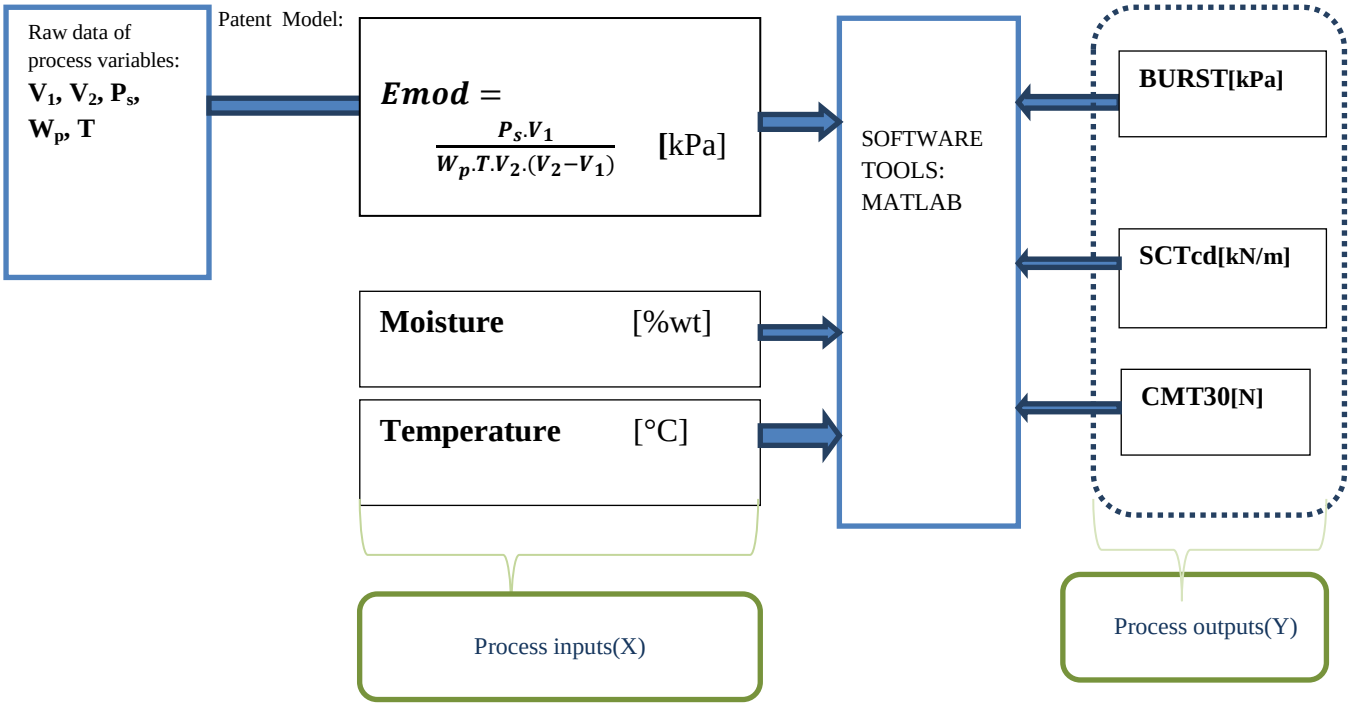


Figure 4. 1 Modeling block diagram scheme

4.4 DESCRIPTION OF THE DATA

Consider the data set, $X_{n\,x\,m}$ consisting of n observations and m-measured variables. The n observations equal to 1911 data points and m equals to six was used to model the process for six dimensional approach.
The procedure starts with autoscaling the data set. This is a typical preprocessing step whereby each measured variable, x_i is mean centered, $\bar{x}_i$ and divided by its standard deviation, σ_i. The mean of x_i is given by the sum of all x_i divided by how many they are in a data set.

The equation (4.1) describes the auto scaling process.

$$X_i = \frac{x_i - \bar{x}_i}{\sigma_i} \tag{4.1}$$

Where i varies from 1 to 6 for six dimensional approach.

This standardization method is used to prevent variables from appearing dominant in the analysis due to solely to their relative magnitude with respect to the other variables. Once the data have been autoscaled, the n x m correlation matrix, C_E can be given by:

$$C_E = \frac{X'X}{n-1} \tag{4.2}$$

where n-1 means observed samples for model training.

Applying equation (4.2) with the help of matlab® into the current autoscaled X variables, we get:

>>C = (X'. X)/ (1274 - 1) % define the correlation matrix of the predictive variables. Where X has n = 1274 rows and m = 6 columns of the data set for the training model. Here the ratio of 2:1 was used with 2 data for training and 1 data for validating the model within 1911 observations.

The above correlation matrix can also be given by the matlab® command: >> corrcoef (X)

The correlation matrix with the corresponding predictive variables for the E-modulus is as follows:

$$C_E = \begin{matrix} \\ \text{Emod} \\ T_p \\ \text{M} \\ \text{Burst} \\ \text{CMT30} \\ \text{SCTcd} \end{matrix} \begin{matrix} \begin{matrix} \text{Emod} & T_p & \text{M} & \text{Burst} & \text{CMT30} & \text{SCTcd} \end{matrix} \\ \begin{bmatrix} 1.0000 & 0.0170 & -0.0782 & 0.0616 & 0.0116 & 0.0729 \\ 0.0170 & 1.0000 & -0.1061 & 0.4759 & 0.1863 & 0.4766 \\ -0.0782 & -0.1061 & 1.0000 & -0.3714 & -0.2445 & -0.4299 \\ 0.0616 & 0.4759 & -0.3714 & 1.0000 & 0.6413 & 0.9660 \\ 0.0116 & 0.1863 & -0.2445 & 0.6413 & 1.0000 & 0.6089 \\ 0.0729 & 0.4766 & -0.4299 & 0.9660 & 0.6089 & 1.0000 \end{bmatrix} \end{matrix} \tag{4.3}$$

The correlation matrix (C_E) gives a measure of the correlation or relationship of the data set elements to one another. As it can be seen, Equation 4.3 shows that Burst and SCTcd are extremely high correlated and correlated with other two variables. Both Burst and SCTcd are the quality specs of the paper product. While E-modulus is poorly correlated with all the other variables, the moisture is negatively correlated with all of them. This means that moisture has negative impact on the other variables.

4.5 PRINCIPAL COMPONENTS ANALYSIS FOR SIX DIMENSIONAL APPROACH

The principal components are obtained by eigenvalue decomposition of the covariance or correlation matrix of the predictive variables under considerations. The decomposition of eigen-values can easily be computed by most of the statistical software for the specified data set and variables with which we want to make the principal components. In this thesis, matlab® with its statistics toolbox can provides the outputs of linear coefficients for eigenvectors along with mean and standard deviations of each predictive variable.

These can be used to compute the principal components on the data set for the analysis. Therefore, the following matlab® command is used to obtain the eigenvector, V_E corresponding to each eigenvalue, λ_E.

$$>> [V_E,\lambda_E] = \text{eig}(C_E) \quad (4.4)$$

Where C_E is the symmetric, square, and exclusively populated with real values and

λ_E is the diagonal matrix in decreasing order. The subscripts letter E is there only to specify the calculations done for E-modulus.

The symmetric matrix has important properties that describe the subspaces in principal components.

In other words, a non-zero column vector, V_E is a right eigenvector of a matrix C_E. if and only if there exists a number, λ_E] such that the equation (4.5) holds true.

$$C_E . V_E = \lambda_E . V_E \quad (4.5)$$

Where λ_E is called the eigenvalue corresponding to eigenvector, V_E.

The set of all eigenvectors of matrix C_E, each paired with its corresponding eigenvalue is called the eigen system of that matrix [23], [24].

The eigenvectors of a real symmetric matrix are orthogonal. This allows the eigenvector of the correlation matrix to be used as the basis (i.e coordinate system) for the new subspace. Since the eigenvalues of the real symmetric matrix is also real. The magnitude of each eigenvalue describes the variance in the direction of its corresponding eigenvector. The largest variation in the data is referred to as the first principal component (PC_1), the next largest variation is the second principal component (PC_2) and so forth. Eventually, the higher components represent mainly noise. This is one of the reasons of using PCA in the first place. Since the noise is regarded to the higher-level components, it is not present from the first few components. Note that all components are orthogonal to each other, which means that they are statistically independent or uncorrelated.

The equation (4.5) represents a standard eigenvalue problem. To solve this eigenvalue problem, several techniques can be used. The complete solution set of eigenvectors and eigenvalues is desired in principal components analysis.

The computation results of Jordan decomposition of the correlation matrix C_E with help of matlab® are shown below:

Therefore, the results from the equation (4.4) are expressed as follows:

$$\lambda_E = \begin{bmatrix} 0.0309 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.3637 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.7462 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.8935 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.0146 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.9510 \end{bmatrix} \quad (4.6)$$

The equation (4.6) represents the eigenvalues matrix in descending order. It is noted that the sum of eigenvalues correspond to the total number of six variables required to make correlation matrix. The corresponding eigenvectors are given in the equation (4.7) bellow:

$$V_E = \begin{bmatrix} 0.0043 & 0.0408 & -0.1675 & 0.3150 & 0.9316 & 0.0558 \\ 0.0006 & 0.4193 & 0.4311 & 0.7010 & -0.1976 & 0.3287 \\ -0.0506 & -0.2235 & -0.6492 & 0.5929 & -0.2891 & -0.3016 \\ 0.7074 & -0.4197 & -0.1072 & 0.0418 & -0.0515 & 0.5546 \\ -0.0387 & 0.6347 & -0.5936 & -0.2366 & -0.0798 & 0.4254 \\ -0.7039 & -0.4400 & -0.0291 & 0.0149 & -0.0210 & 0.5562 \end{bmatrix} \quad (4.7)$$

The eigenvalues are now found from the covariance matrix. The next step is to order them by eigenvalue, from the highest to the lowest. This gives us the components in order of significance. Then, it is possible to ignore the components with lesser significance. This will make you lose some information, but if the eigenvector are small, you don't lose much. If you leave out some components, the final data set will have fewer dimensions than the original. To be precise, the remaining dimension will be equal to the first retained eigen vectors of your data set. The Table 4.1 below shows the results of the sorted eigenvalues of our data set.

Eigenvalues	2.9510	1.0146	0.8935	0.7462	0.3637	0.0309
Proportional of total variance	0.4918	0.1691	0.1489	0.1244	0.0606	0.0052
Cumulative proportional of total variance	0.4918	0.6609	0.8098	0.9342	0.9948	1.0000

Table 4. 1 Cumulative eigenvalues for six dimensional approach.

The same results can also be given by using matlab® function ">> pareto (percent_explained) " where "percent_explained" is the symbol used to specify the variance explained by the six dimensional variables in matlab® codes as it can be seen in the Appendix G.

It is noted that 95% of the cumulative distribution of the variance is displayed in descending order. The cumulative variance and the respective associated original principal components because of pareto function in matlab® are shown on the Figure 4.2 below:

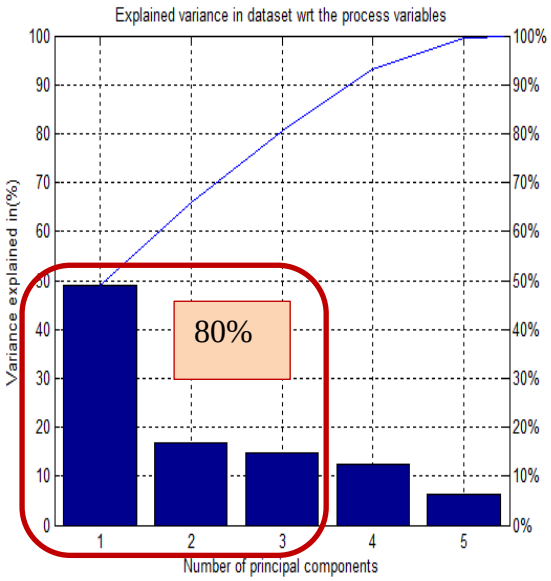


Figure 4.2.a

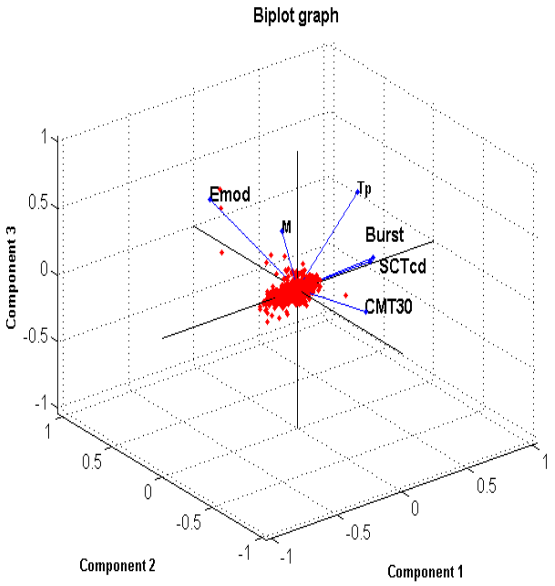


Figure 4.2.b

Figure 4. 2 Variance explained in data set for six dimensional approach (Figure 4.2.a & Figure 4.2.b)

Notice from the Figure 4.2.a that each bar labels the associated value of the variance from the autoscaled data matrix by the number of principal components. The pareto function represents only 95% of the cumulative distribution of the variances in the data set. Due to that reason, the variable with less than 5% of the variance is not displayed. The Figure 4.2.b shows the use of matlab® function biplot. A biplot allows us to visualize the magnitude and the sign of each variable's contribution to the first three principal components and how each observation is represented in terms of those three components. Notice from the Figure 4.2.b that E-modulus (Emod) has the largest magnitude and moisture (M) is with opposite sign to the temperatures (Tp). It is said that they are negatively correlated since they are in different planes. This is realistic because the increase in temperature of the paper causes the decrease in its moisture. Burst and SCTcd are the paper qualities and are much more correlated. This has been proved also by the equation (4.3) of the covariance matrix where the correlation between those two qualities is [0.96].

As we can see from both the Table 4.1 and the Figure 4.2.a, the first three eigenvalues capture about 80% of the total energy in the process. Accepting the inaccuracy of 20%, it is clear that the most three principal eigenvectors which are corresponding to the maximum three eigenvalues can estimate the total variance in the entire process for the case of six dimensional approach.

Since the first three eigenvalues capture the most of the variance in the data set, we can compress our data into three-dimensional space without losing too much of information in our data set.

This technique will help us to construct the new predictor variables called components, which are in the linear combination with original predictors. Therefore, the first three principal components can now be computed as follows:

PC_1 = 0.0558 Emod*+ 0.3287 T_p* - 0.3016 M* +0.5546 Burst* + 0.4254 CMT30* + 0.5562 SCTcd*

PC_2 = 0.9316 Emod* - 0.1976 T_p* - 0.2891 M* - 0.0515 Burst* - 0.0798 CMT30* - 0.0210 SCTcd*

PC_3 = 0.3150 Emod*+ 0.7010 T°* + 0.5929 M* + 0.0418 Burst* - 0.2366 CMT30* + 0.0149 SCTcd*

(4.8)

Where (*) means that each variables should be mean subtracted and divided by its standards deviation. The mean and standards deviation of each x-variable are calculated with matlab® by using the respective functions as follows >> mean (x) and >> std (x) and the results are shown below:

Process variables	Emod	T_p	M	Burst	CMT30	SCTcd
Mean (x1.0e+004)	1.8183	0.0070	0.00075	0.0388	0.0218	0.0009
Standard deviation (x1.0e+004)	3.1498	0.0003	0.00005	0.0079	0.0022	0.0002

Table 4. 2 Mean and standard deviation for each process variable for six dimensional approach.

Note that each variable is autoscaled while computing the principal components and the three principal components as shown in Table 4.1 represent 80% of the total variance in the data set.

It is obvious that three out of six components can be used to represent the model without losing so much information in original data set. This is how PCA contribute to the dimension reduction by reducing multidimensional data sets to lower dimensions for analysis.

The retained three dimensions can now be described by using the data of Table 4.2 into equation (4.8) as follows:

$PC_1 = 0.0558\left(\frac{Emod-18183}{31498}\right) + 0.3287\left(\frac{T_p-70}{3}\right) - 0.3016\left(\frac{M-7.5}{0.5}\right) + 0.5546\left(\frac{Burst-388}{79}\right) + 0.4254\left(\frac{CMT30-218}{22}\right) + 0.5562\left(\frac{SCTcd-9}{2}\right).$

$PC_2 = 0.9316\left(\frac{Emod-18183}{31498}\right) - 0.1976\left(\frac{T_p-70}{3}\right) - 0.2891\left(\frac{M-7.5}{0.5}\right) - 0.0515\left(\frac{Burst-388}{79}\right) - 0.0798\left(\frac{CMT30-218}{22}\right) - 0.0210\left(\frac{SCTcd-9}{2}\right).$

$PC_3 = 0.3150\left(\frac{Emod-18183}{31498}\right) + 0.7010\left(\frac{T_p-70}{3}\right) + 0.5929\left(\frac{M-7.5}{0.5}\right) + 0.0418\left(\frac{Burst-388}{79}\right) - 0.2366\left(\frac{CMT30-218}{22}\right) + 0.0149\left(\frac{SCTcd-9}{2}\right).$

(4.9)

Simplifying the equation (4.9) result in the following three principal components:

$PC_1 = 1.77\text{x}10^{-6}\ \text{Emod} + 0.1095\text{T}_\text{p} - 0.6032\ \text{M} + 7.02\text{x}10^{-3}\text{Burst} + 0.019\ \text{CMT30} + 0.2781\ \text{SCTcd} - 12.6$

$PC_2 = 2.9576\text{x}10^{-5}\ \text{Emod} - 0.0658\text{T}_\text{p} - 0.5782\ \text{M} + 6.5189\text{x}10^{-4}\ \text{Burst} - 3.6272\text{x}10^{-3}\ \text{CMT30} - 0.0105\ \text{SCTcd} + 9.54$

$PC_3 = 1.0\text{x}10^{-5}\ \text{Emod} + 0.2336\text{T}_\text{p} + 1.1858\ \text{M} + 5.2911\text{x}10^{-4}\ \text{Burst} - 0.0107\ \text{CMT30} + 7.45\text{x}10^{-3}\ \text{SCTcd} - 23.34$

(4.10)

From equation (4.10), it clear that the contribution of Emod is considerably small, this may result in poor correlation with other process values and the qualities as well. However, PCA help in determination of the available components required for the model and the amount of energy expected by each component. However, it does not look on the causes and effect of the prediction of output from the inputs. The extracted energy within those three new PCs with respect to the output quality is shown by using PLSR method.

4.6 PLSR FOR A SIX DIMENSIONAL APPROACH

Principal component analysis, PCA and Partial least squares PLS are two techniques used for multivariate data analysis. While PCA looks only on the correlation between independent variables or predictors, PLS take into account both independent variables, X and dependent variables, Y for regression of highly correlated process data.

Scores also called principal components and are extracted while explaining the decreasing amount of variations obtained within a group of variables. The fact is due to the presence of majority of variations in few process variables that can explain the whole process without considerably losing too much of the information within the data set [25].

4.6.1 DESCRIPTION OF PLS TECHNIQUE

The description of PLS technique is shown by referring to the work of Saikat Maitra and Jun Yan [26]. Let us define the independent variables or predictors, $X_{n \times m}$ matrix and dependent variables or responses or again target to achieve, $Y_{n \times q}$ matrix. The PLS technique works by successively extracting factors from both X and Y such that covariance between the extracted factors (i. e $T_{n \times k}$ and $U_{n \times k}$) are maximized.
The idea is to find the linear decomposition such that outer relation for input matrix(X) and outer relation for output matrix(X) are written as follow:

$$\begin{cases} X = TP' + E = t_1 p_1' + t_2 p_2' + \cdots + t_k p_k' + E \\ Y = UQ' + F = u_1 q_1' + u_2 q_2' + \cdots + u_k q_k' + F \end{cases} \quad (4.11)$$

where $T_{n \times k}$ and $U_{n \times k}$ are matrices that are respectively projections of X (X-scores, components or factor matrix) and projections of Y (Y-scores).
$P_{m \times k}$ and $Q_{q \times k}$ are orthogonal loading matrices.
$E_{n \times m}$ and $F_{n \times q}$ are respectively X-residuals and Y-residuals or in general are called errors terms.
In order to use PLS technique, different algorithms exist to estimate the factor and loading matrices. Indeed, all the techniques may use much iterative process to extract the X-scores and Y-scores. The factors or scores for both X and Y are extracted iteratively and the number of factors (k) depends on the rank of X and Y matrices.

4.6.2 EIGEN VALUE DECOMPOSITION ALGORITHM FOR PLS.

The eigenvalue decomposition is similarly done as in the equation (4.4) by considering both input matrix and output matrix. By using eigenvalue decompositions method X-scores are in linear combinations of predictors.

Let's extract iteratively the X-scores and for the first case t = X. w where w is the eigenvector corresponding to the first eigenvalue $X^{'}YY^{'}X$.

The same as above for Y-scores, the first case is given by u =Y. v, where v is the eigenvector corresponding to the first eigenvalue of $Y^{'}XX^{'}Y$.

Here the covariance of predictors(X) and predicted variables (Y) is denoted as $X^{'}Y$.

Here is the method to extract the first PLS factor:

$$\begin{cases} X_1 = X - tt'X \\ Y_1 = Y - uu'Y \end{cases} \quad (4.12)$$

The second PLS factors are iteratively extracted the equation (4.12) as follows:

$$\begin{cases} X_2 = X_1 - tt'X_1 \\ Y_2 = Y_1 - uu'Y_1 \end{cases} \quad (4.13)$$

The next PLS components are thus obtained by following the same procedures as shown above until all latent factors t and u have been extracted. The number of latent factors extracted depends on the rank of X when the last ones are reduced to null matrix during factors extractions [25] [26].

4.6.3 PLS REGRESSION AND PARAMETERS ESTIMATION

A regression and least squares model relates the response variables (Y) to a function of the predictor variables (X) and parameter beta, β.

The values of β's contribute to the model fit.

The equation of simple linear regression model for a data set is as follows:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ with i = 1, 2, 3…, n.} \quad (4.14)$$

where Y_i is the response (dependent variable) for 'ith' observation and X_i is the value of the independent variable or predictor variable or repressor variable for the ith observation .

β_0 , β_1 are the unknown parameters and they are also called regression coefficients to be estimated.

ε_i is the random error for the ith observation. It is assumed to have normal distribution with the mean zero and the constant variance, σ^2.

4.6.4 PARAMETERS DERIVATION OF LINEAR REGRESSION METHOD

The regression coefficients β_0 and β_1, are traditionally estimated using least squares [27]

Such estimators, usually noted by $\hat{\beta}_0$ and $\hat{\beta}_1$are defined as the minimizer of the following cost function as follows:

$$Q_0(\beta_0, \beta_1) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 \quad (4.15)$$

By differentiating equation (4.15) with respect to β_0 and β_1 and then setting the obtained partial derivatives to zero, we get the following equations:

$$\left\{\frac{\partial Q_0}{\partial \beta_0} = 0\right\} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) \quad (4.16)$$

$$\left\{\frac{\partial Q_0}{\beta_1} = 0\right\} = -2\sum_{i=1}^{n} x_i \,.\, (Y_i - \beta_0 - \beta_1 x_i) \quad (4.17)$$

Solving equations (4.16) with respect to β_0 and (4.17) with respect to β_1 yields the least square estimates below:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (4.18)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (4.19)$$

with $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ are the sample mean of the x_i's and y_i's respectively and n is the number of observations.

The equation (4.18) determines the intercept (offset) of the regression line while equation (4.19) is known as the slope of the regression line or the fitting line ($\hat{y}$), which can be written as follow:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (4.20)$$

where the indices "i" for $\hat{y}_i$ is the i^{th} observation value and $\hat{y}_i$ estimates the prediction of y_i for a given value of x_i.

The equation (4.20) is the regression line or line of fit that minimizes the sum of squared errors, SS_{err}. The variability of the data set is measured through different sums of squares:

The sum of squares of residuals (SS_{err}), also known as the residual sum of squares is given by:

$$SS_{err} = \sum \varepsilon_i{}^2 = \sum (y_i - \hat{y}_i)^2 \quad (4.21)$$

The regression sum of squares (SS_{reg}), also known as the explained sum of squares is given by:

$$SS_{reg} = \sum (\hat{y}_i - \bar{y}_i)^2 \quad (4.22)$$

where the term ε_i is called the residual or unexplained deviation and it is known as the difference between the observed (y_i) and the fitted value ($\hat{y}_i$).
The total sum of squares (SS_{tot}) that is proportional to the sample variance and is represented by:

$$SS_{tot} = \sum (y_i - \bar{y}_i)^2 \quad (4.23)$$

The coefficient of determination, R^2 is indication measure of fit of the model. Its expression is given by:

$$R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \quad (4.24)$$

It is noted that the better is the line of fit of the data in linear regression, the closer value of R^2 is to 1.

In some cases, the SS_{tot} is equal to the sum of the SS_{err} and SS_{reg} [29], [30], [31].

4.6.5 PLS REGRESSION AND PARAMETER ESTIMATION BY MATLAB®

The partial least squares(PLS) regression role is to find the latent variables that capture the variance in the data and therefore achieves the maximum correlation between predicted variables (Y) and predictors variables (X).
If the both inputs (X) and outputs(Y) for "n" data sets are known, a number of available software tools can help solving the problem. The matlab® command below can be used to approximate parameter β and other parameters elements need to visualize the responses based on the available predictors. Literatures [28] and matlab® command window (>>help plsregress) give more details about PLSR method.
The following abbreviations were used in matlab® >> help plsregress.

[XL, YL, XS, YS, BETA, PCTVAR, MSE, STATS] = PLSREGRESS(X, Y, NCOMP) (4.25)

Equation (4.20) returns a structure that contains fields explained below:

X_o = original predictors or independent variables
Y_o = original predicted or response variables
X_S = X - scores ($T_{n\,x\,k}$) and Y_S = Y- scores ($U_{n\,x\,k}$)
X_L= X - loadings ($P_{m\,xk}$) and Y_L= Y- loadings ($Q_{q\,x\,k}$)
w is from the relation of P-by-NCOMP matrix of PLS weights, i.e., $X_s = X_o .w$
T^2 statistic, abbreviated as STATS is calculated for each point in X-scores
X residuals means the predictor residuals, i.e. $X_o - X_S.X_L'$
Y residuals means the response residuals, i.e. $Y_o - Y_S.Y_L'$ where

The idea behind the software tools is to estimate the important parameter; β that estimates y from x. Refer to equation (4.14) the linear regression is a model to predict the value of one variable from another. Multiple regressions are natural extension of linear regression. They are used whenever there is prediction of values of an outcome from several predictors.

4.6.6 APPLICATION OF LINEAR REGRESSION METHOD FOR SIX DIMENSIONAL APPROACH

Let us assume that the relationship between predictor variables and their responses is linear and static. By using the same principle as in linear regression model, the example of multiple regression expression with the covariates: E-modulus, Temperature, and Moisture can be written in matrix form as shown below:

$$\begin{bmatrix} y_1^{Burst} \\ y_2^{CMT30} \\ y_3^{SCTcd} \end{bmatrix} = \begin{bmatrix} \beta_1^{Burst} & \beta_2^{Burst} & \beta_3^{Burst} \\ \beta_1^{BCMT30} & \beta_2^{BCMT30} & \beta_3^{BCMT30} \\ \beta_1^{SCTCd} & \beta_2^{SCTCd} & \beta_3^{SCTCd} \end{bmatrix} \begin{bmatrix} \text{Emod} \\ T_P \\ \text{M} \end{bmatrix} + \begin{bmatrix} \beta_0^{Burst} \\ \beta_0^{BCMT30} \\ \beta_0^{SCTCd} \end{bmatrix} \quad (4.26)$$

Using equation (4.20) and equation (4.21) in our data set for 6-dimensional approach, we get:

$$\begin{bmatrix} y_1^{Burst} \\ y_2^{CMT30} \\ y_3^{SCTcd} \end{bmatrix} = \begin{bmatrix} 0.039 & 0.4644 & -0.3152 \\ -0.0068 & 0.1921 & -0.1936 \\ 0.0499 & 0.4576 & -0.3781 \end{bmatrix} \begin{bmatrix} \text{Emod} \\ T_P \\ \text{M} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (4.27)$$

It was observed from the equation (4.27) that all the coefficients of Emod are negligibly small for the three independent outputs. This is shows the lack of Emod to predict the quality. It is also an indication of bad fit as it is shown for the PLS simulations below.

The temperature and moisture have high impact on each quality and they are negatively correlated. This corresponds to the normal operation of the paper process because if one goes high, the other goes low.

4.6.7 PLS SIMULATION RESULTS FOR PREDICTING THE PAPER QUALITY IN SIX DIMENSIONAL APPROACH

As discussed earlier, PLS was computed in order to predict the output variables (y) for the given input variables (x).The number statistics used to describe the variance of one variable that is accounted for (predicted) by the other variable is known as the coefficient of determination, usually noted as R^2.

The values of R^2 are always positive and vary from 0 to 1, where 0 indicates poor accountability from one variable to another and 1 suggest the perfect prediction. If one variable can consistently predict the value of the other variable, then a high degree of correlation exist between them. It is noted that high correlation between variables does not mean that one variable causes the other. High correlation suggests that a causal relationship might exist. No correlation or low correlation assumes no causal relationship exists between two variables. However, lack of correlation may be due to other factors such as, poor measurements, restricted range, non-linear relationships, or other extraneous factors that mask the true relationship. The PLSR was computed using developed codes in matlab® software with its built in statistics toolbox in matlab®, plsregress, and only the results are shown below. As previously discussed, the variables were first standardized and then proceed to regression analysis. Regression analysis is concern with finding a formula that represents the relationship between variables to find an approximate value of one variable from the values of others.

The PLS R simulations results are describe below for both training and validation of the model.

Since we have the three different outputs due to three different qualities, the results will show case by case the three different trends for quality prediction. It is noted from all the predictions graphs that the small part of simulations is shown for easy demonstration of the trends.

Case1: Bursting predictions:

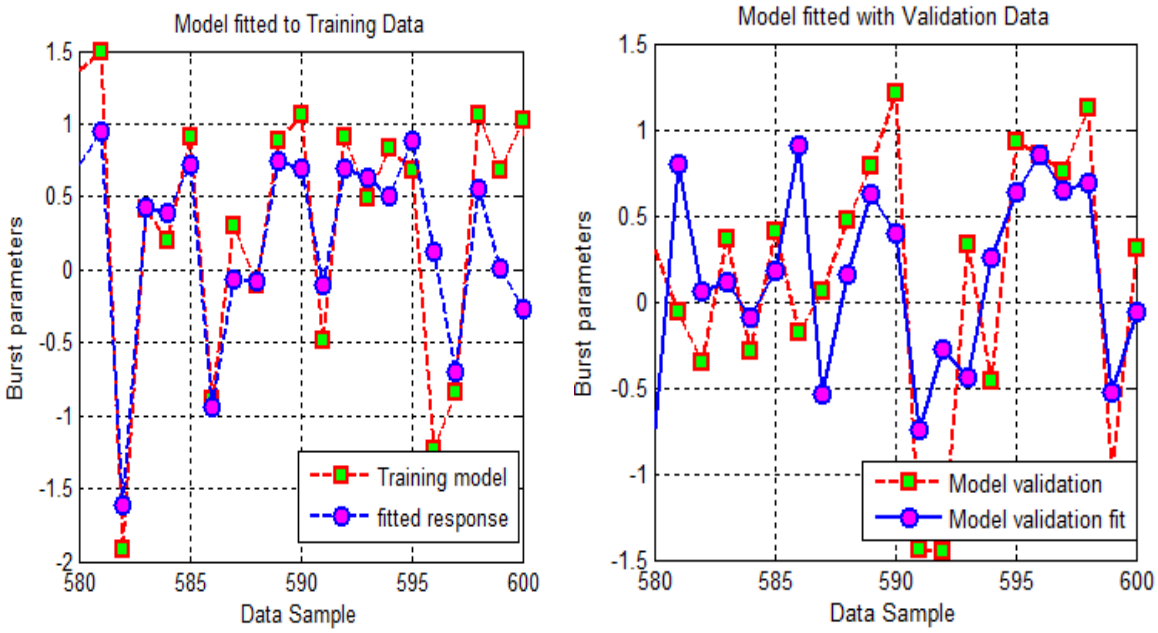


Figure 4. 3 Predictions of Burst trends (Training and validation)

The Figure 4.3 shows the output results of the prediction of burst; the linear regression coefficient, R^2 that quantifies goodness of fit of the burst quality from the six dimensional approach is given by:

- R^2_{Burst} training = 0.29
- R^2_{Burst} validation = 0.27

The Burst model response cannot fit well to the training model. The similar case is obtained from validation of the fitting data. The fitting response does not follow the training trend. This is a result of weak correlation between inputs with Emod and burst as the output quality.

Case2: CMT30 prediction:

The second output quality is CMT30.Notice that the order of the outputs is specified based on how they are arranged in the model (see Figure 4.1). The prediction graphs for both training model and validation of the model are shown on the Figure 4.4. Each graph shows the comparison plot of the actual measured values and the fit line of prediction for both training model and validation of the model with the new data.

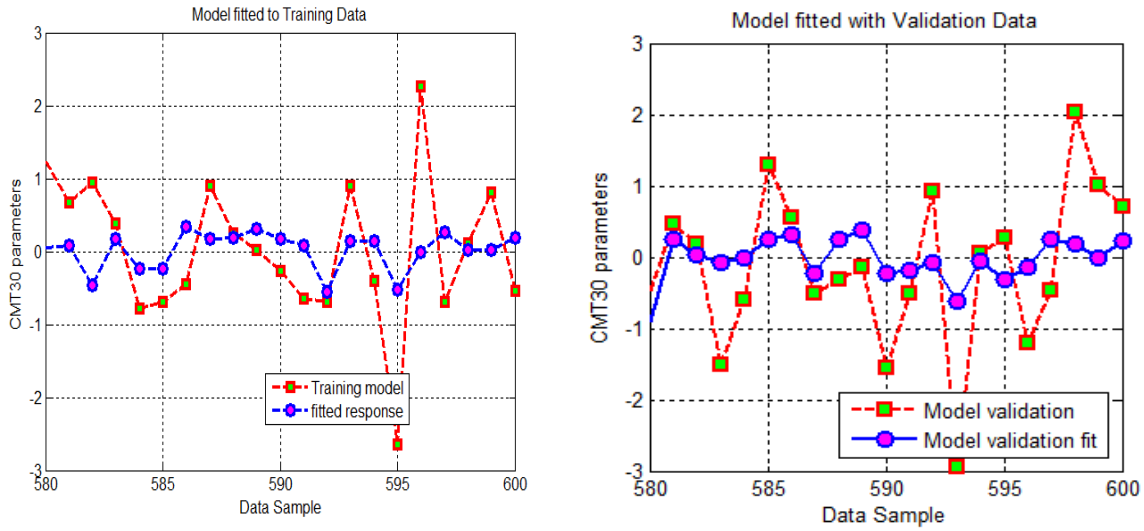


Figure 4. 4 Predictions of CMT30 trends (Training and validation)

The figure 4.4 shows the output results of the prediction of CMT30; the linear regression coefficient, R^2 that quantifies the goodness of fit for CMT30 from the six dimensional approach is given by:

- R^2_{CMT30} training = 0.28
- R^2_{CMT30} validation = 0.25

The CMT30 model response cannot fit well to the training model. The similar case is obtained from the validation of model fitting. The prediction for this paper quality is quite random.
No fitting at all as you can see on the predictions trends below:
The last output for this dimensional approach is SCTcd. The model prediction is as follows:

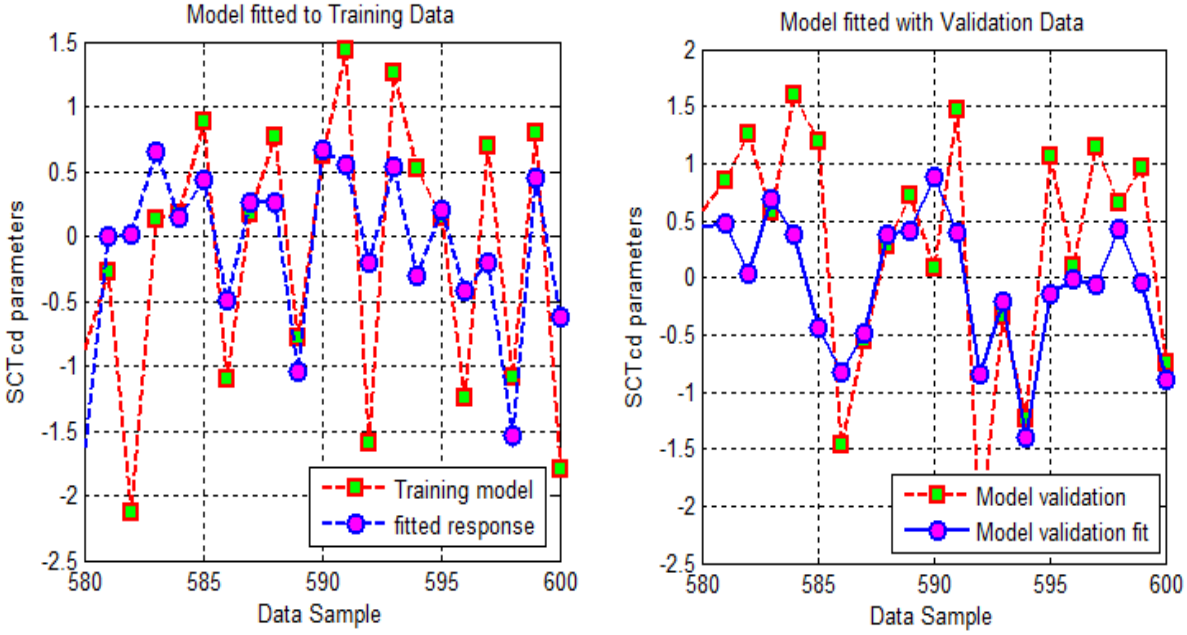


Figure 4. 5 Predictions of SCTcd trends (Training and validation)

The Figure 4.5 shows the output results of the prediction of SCTcd; the linear regression coefficient, R^2 that quantifies goodness of fit for SCTcd from the six dimensional approach is given by:

- R^2_{SCTcd} training = 0.28
- R^2_{SCTcd} validation = 0.31

The SCTcd model response cannot fit well to the training model.

4.6.8 SUMMARY RESULTS ABOUT PLS SIMULATIONS FOR SIX DIMENSIONAL APPROACH

It is obvious that none of the three qualities can be predicted by refering to the current six dimensional approach. The model response for the quality always give the unsatisfactory results. This is a conclusive research of my thesis. It is now clear that based on the results on the Figure 4.3 till the Figure 4.5, it comes to the conclusion of not relying on the patent [1] for the quality prediction online where the Emod would be adjusted for keeping constant the quality during the production.

However it is interesting to see what would happen if direct raw data of process variables are used for predicting the quality.The following chapter discuss the details.

4.7 PRINCIPAL COMPONENTS ANALYSIS FOR TEN DIMENSIONAL APPROACH

What will happen when we apply the raw data from the process variables required to make E-modulus? Applying equations (4.1) and (4.2) on the process variables, we get now ten dimensional space that is given by ten process variables.

Let X_{asct} be the autoscaled data matrix and X_{asct}' be its transpose matrix of a 10-dimensional space whereby ten process variables are standardized and C_{pv} is the covariance matrix that is calculated using equation (4.2) as follows:

$$C_{pv} = \frac{X_{asct}' * X_{asct}}{n-1}$$

Where n is the total number of samples used to model the process, in this case n=1274 observations.

The result from equation (4.6) is given below:

$$C_{pv}=\begin{array}{c}\\ P_s\\ V_2\\ V_1\\ dV\\ A\\ T_p\\ M\\ Burst\\ CMT30\\ SCTcd\end{array}\begin{array}{c}\begin{array}{cccccccccc} P_s & V_2 & V_1 & dV & A & T_p & M & Burst & CMT30 & SCTcd \end{array}\\ \begin{bmatrix} 1.0000 & -0.1276 & -0.1287 & 0.0944 & 0.2672 & 0.5910 & -0.1983 & 0.3090 & 0.2542 & 0.3452 \\ -0.1276 & 1.0000 & 1.0000 & 0.2628 & -0.9003 & -0.4642 & 0.3068 & -0.8043 & -0.1459 & -0.7698 \\ -0.1287 & 1.0000 & 1.0000 & 0.2546 & -0.9003 & -0.4653 & 0.3054 & -0.8043 & -0.1458 & -0.7696 \\ 0.0944 & 0.2628 & 0.2546 & 1.0000 & -0.2286 & 0.0057 & 0.2377 & -0.2119 & -0.0432 & -0.2178 \\ 0.2672 & -0.9003 & -0.9003 & -0.2286 & 1.0000 & 0.5306 & -0.3207 & 0.8978 & 0.2224 & 0.8708 \\ 0.5910 & -0.4642 & -0.4653 & 0.0057 & 0.5306 & 1.0000 & -0.1104 & 0.5054 & 0.1836 & 0.5046 \\ -0.1983 & 0.3068 & 0.3054 & 0.2377 & -0.3207 & -0.1104 & 1.0000 & -0.3687 & -0.2208 & -0.4351 \\ 0.3090 & -0.8043 & -0.8043 & -0.2119 & 0.8978 & 0.5054 & -0.3687 & 1.0000 & 0.5801 & 0.9650 \\ 0.2542 & -0.1459 & -0.1458 & -0.0432 & 0.2224 & 0.1836 & -0.2208 & 0.5801 & 1.0000 & 0.5543 \\ 0.3452 & -0.7698 & -0.7696 & -0.2178 & 0.8708 & 0.5046 & -0.4351 & 0.9650 & 0.5543 & 1.0000 \end{bmatrix}\end{array} \quad (4.28)$$

The correlation matrix of the process variables in equation (4.10) shows that there is a high correlation between speeds, V_1 and V_2 and this is because both speeds have slightly small difference. There is collinearity between V_1 and V_2 (Appendix E).

The perfect multicollinearty among independent variables in the process data may cause the covariance matrix to be invertible and thus creates wrong results in data analysis. The equation (4.28) shows that there is higher correlation between cross sectional area of paper and the speeds v_1 and $v_{2,}$ which is [-0.9]. Burst and SCTcd are much stiffer related and their stiffness is related to the paper thickness and paper width. The correlation between area and burst as well as SCTcd is [0.89] and [0.87] respectively. Temperature (T_p) and moisture (M) have higher and opposite effect on the paper process. The effect is slightly higher for Temperatures than Moisture.

This is because the increase of either causes the decrease of the other. It is noted from the equation (4.10) that correlation of P_s with other explanatory variables is low. The same case happens for equation (4.28) too .This argument strengthens the use of directly measured data for statically analysis purpose.

The eigenvalues and the corresponding eigenvectors are computed in matlab® using equation (4.4).

The results are as follows:

Eigenvalues	5.2907	1.4450	1.1396	0.8785	0.6990	0.3087	0.1715	0.0494	0.0176	-0.000
Proportional of total	0.5291	0.1445	0.1140	0.0878	0.0699	0.0309	0.0172	0.0049	0.0017	-0.000
Cumulative proportional of total	0.5291	0.6736	0.7876	0.8754	0.9453	0.9762	0.9934	0.9983	1.000	1.000

Table 4.3 Cumulative eigenvalues with respect to the total eigenvectors for ten dimensional approach

The eigenvalues table shows that even if it is difficult to find the relationship or correlations between variables within a high dimensional data set, principal components analysis makes it easier by using eigenvalues and eigenvectors decompositions.

It is obvious that five out of 10 eigenvalues contain 95% of the total energy in the system. This means that we can compress our dimensional data set into a 5-dimensional with inaccuracy of 5% .This is the main advantage of using PCA. The analysis of PCA graph with matlab® together with the computational mathematics shows that five principals components are enough to describe all the information with 10 dimensional approach.

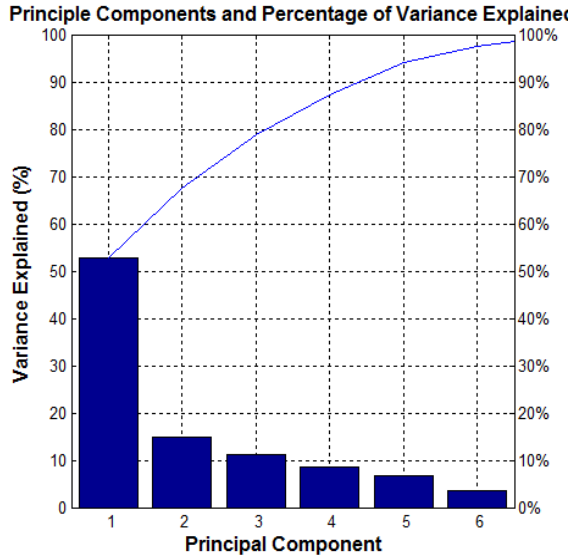


Figure 4.6.a

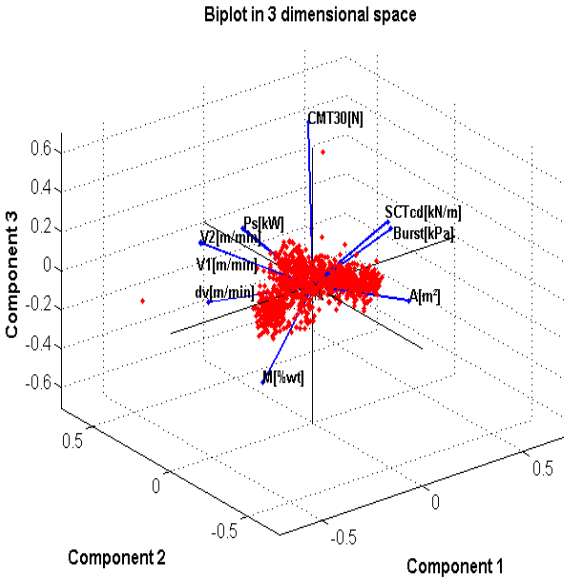


Figure 4.6.b

Figure 4. 6 Variance explained in data set for ten dimensional approach (Figure 4.6.a & Figure 4.6.b)

The eigenvectors corresponding to each of the eigenvalues from equation (4.4) are given in matrix form below:

$$V_{pv} = \begin{bmatrix} 0.1663 & 0.6277 & -0.0583 & 0.4290 & 0.0885 & -0.5900 & -0.1858 & -0.0072 & 0.0212 & -0.0000 \\ -0.3898 & 0.2522 & 0.2300 & 0.0490 & 0.0816 & 0.0381 & 0.4603 & 0.0889 & 0.0132 & -0.7079 \\ -0.3896 & 0.2488 & 0.2333 & 0.0533 & 0.0876 & 0.0379 & 0.4614 & 0.0891 & 0.0133 & 0.7063 \\ -0.1188 & 0.4454 & -0.3268 & -0.4804 & -0.6703 & 0.0236 & -0.0030 & -0.0036 & -0.0022 & 0.0062 \\ 0.4083 & -0.1139 & -0.1734 & -0.0446 & -0.0418 & -0.1793 & 0.4092 & 0.6445 & -0.4142 & -0.0000 \\ 0.2681 & 0.3928 & -0.3487 & 0.2562 & 0.2253 & 0.7297 & 0.0579 & 0.0036 & -0.0010 & 0.0000 \\ -0.1945 & 0.0514 & -0.4874 & -0.4926 & 0.6520 & -0.2187 & 0.0208 & -0.0758 & -0.0173 & 0.0000 \\ 0.4158 & 0.0292 & 0.1128 & -0.2208 & 0.0420 & -0.0698 & 0.2573 & 0.1069 & 0.8249 & 0.0000 \\ 0.1896 & 0.3302 & 0.5991 & -0.4483 & 0.2216 & 0.1549 & -0.3581 & 0.2101 & -0.2228 & -0.0000 \\ 0.4114 & 0.0493 & 0.1470 & -0.1411 & -0.0108 & -0.0847 & 0.4184 & -0.7123 & -0.3117 & 0.0000 \end{bmatrix}$$

Eigenvectors to retain or to keep are V_{pv} (:, 1:5) and eigenvectors to discard are V_{pv}(:,6:10)

(4.29)

Since we now have the idea on how many components need to represent the most of the variance in our data set, we can use partial least squares regression, PLSR with only 5PCs to find the correlation between inputs and outputs.

We use the number of components found from PCA that represent the most of the variance, here 94.5% are represented only by 5 PCs. This is the benefit of PCA where we can reduce multidimensional data sets to lower dimensions for making good predictive models and making easier for analysis [28].

The first 5 PCs can be computed using the similar procedure as done from equation (4.7) to equation (4.9) as follows:

$$PC_1 = 0.1663\ P_s\ *- 0.3898\ V_2\ *\ -\ 0.3896\ V_1\ *\ -\ 0.1188\ dV*\ + 0.4083\ A * +0.2681 T_p * -0.1945\ M * +0.4158\ Burst*\ +0.1896\ CMT30* + 0.4114\ SCTcd*$$

$$PC_2 = 0.6277\ P_s\ *+ 0.2522\ V_2\ *\ + 0.2488\ V_1\ *\ + 0.4454\ dV*\ - 0.1139\ A * +0.3928 T_p + 0.0514\ M + 0.0292\ Burst*\ +0.3302\ CMT30* + 0.0493\ SCTcd*$$

$$PC_3 = -0.0583\ P_s\ *+ 0.2300\ V_2\ *\ + 0.2333\ V_1\ * -0.3268\ dV* -0.1734\ A * -0.3487 T_p - 0.4874\ M + 0.1128\ Burst* + 0.5991\ CMT30* + 0.1470\ SCTcd*$$

$$PC_4 = 0.4290\ P_s\ *+ 0.0490\ V_2\ *\ + 0.0533\ V_1 * -0.4804\ dV* -0.0446\ A * 0.2562\ T_p * -0.4926\ M * -0.2208\ Burst* -0.4483\ CMT30* -0.1411\ SCTcd*$$

$$PC_5 = 0.0885\ P_s\ *+ 0.0816\ V_2\ *\ + 0.0876\ V_1 * -0.6703\ dV* -0.0418\ A * +0.2253\ T_p * +0.6520\ M * +0.0420\ Burst*\ + 0.2216\ CMT30* -0.0108\ SCTcd*$$

(4.30)

The table that contains the mean and standards variables computed for each for each x- variable by using mean(x) and std(x) matlab® functions and the results are given below:

Process variables	P_s [kW]	V_2 [m/min]	V_1 [m/min]	dV [m/min]	A [m²]	T_p [°C]	M [%wt]	Burst [kPa]	CMT30 [N]	SCTcd [kN/m]
Mean	18.99	818.92	817.68	1.23	0.0013	70.14	7.58	385.62	218.16	9.4
Standard deviation	2.14	68.14	67.96	0.62	0.0002	3.17	0.45	80.43	23.35	1.86

Table 4. 4 Mean and standard deviation for each process variable for 10-dimensional approach

Using the mean and standard deviations data in Table 4.3, we can simplify 5 five retained PCs in equation (4.12) as follows:

$$PC_1 = 0.0777\ P_s\ -5.73x\ 10^{-3}\ V_2\ -\ 5.732x\ 10^{-3}\ V_1\ -\ 0.1916\ dV\ + 2.041.5\ A + 0.0845 Tp - 0.4322\ M + 5.1697\ Burst\ +8.12x10^{-3}\ CMT30 + 0.2211\ SCTcd\ -\ 9.564.$$

$$PC_2 = 0.2933\ P_s + 3.7x10^{-3}\ V_2\ + 3.66\ x10^{-3}\ V_1 + 0.718\ dV - 569.5\ A + 0.124 T_p + 0.1142\ M + 3.63\ x10^{-4}\ Burst + 0.0141\ CMT30 + 0.026\ SCTcd\ -\ 24.761$$

$$PC_3 = -0.272 P_s + 3.375\ x10^{-3}\ V_2\ + 3.433\ x10^{-3}\ V_1 - 52.7 dV - 867\ A - 0.11 T_p - 1.083 M + 1.402\ Burst\ + 0.0256\ CMT30 + 0.079 SCTcd + 31.6541.$$

$PC_4 = 0.2\ P_s + 7.19\,x10^{-4}\,V_2 + 7.8428\,x10^{-4}\,V_1 - 0.7748\ dV - 223\,A + 0.0815T_p - 1.095\,M - 9.456\,x10^{-3}$ Burst -0.241 CMT30 -0.0758 SCTcd+55.825.

$PC_5 = 0.041\ P_s + 1.1975\,x10^{-3}\,V_2 + 1.288\,x10^{-3}\,V_1 - 1.489\ dV - 209\,A + 0.071\,T_p + 1.448\,M + 5.22\,x10^{-4}$ Burst $+9.49\,x10^{-3}$ CMT30$-5.8\,x10^{-3}$ SCTcd - 19.0076.

(4.31)

4.8 PLS RESULTS FOR TEN DIMENSIONAL APPROACH

The simulation for 10-dimensional approach is quite better than the one from the six dimensional approaches. The reason is due to the use of much more raw data rather than using experimentally manipulated data. It is quite clear that data measured from the plant can be reliable for analyzing the process performance. The principal components analysis proved the number of PCs, which are required to maximize the variability in the data set. Refer to the equation (4.20), the number of 5 PCs that represent 95% of the total variances was used during simulations. Using the same formula as in equation (4.21) extended for 10 dimensional, we get the following PLS regression line:

$$\begin{bmatrix} y_1^{Burst} \\ y_2^{CMT30} \\ y_3^{SCTcd} \end{bmatrix} = \begin{bmatrix} 0.1218 & -0.0362 & -0.0360 & -0.0279 & 0.7972 & -0.0293 & -0.0472 \\ 0.1988 & 0.0767 & 0.0772 & -0.0364 & 0.3242 & -0.031 & -0.1365 \\ 0.1753 & -0.0194 & -0.019 & -0.0464 & 0.7732 & -0.039 & -0.1009 \end{bmatrix} \begin{bmatrix} P_s \\ V_2 \\ V_1 \\ dV \\ A \\ Tp \\ M \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

(4.32)

The equation (4.32) represents the prediction of the paper qualities such as Burst, CMT30, and SCTcd for the case where the raw data of same the process variables was used to calculate E-modulus .The idea was to see the effect of combining many variables into one main variable, called in this research Emodulus.

The simulations graphs showing how good is the prediction equation (4.14) are described below:

Case 1: Bursting strength prediction for ten dimensional approach

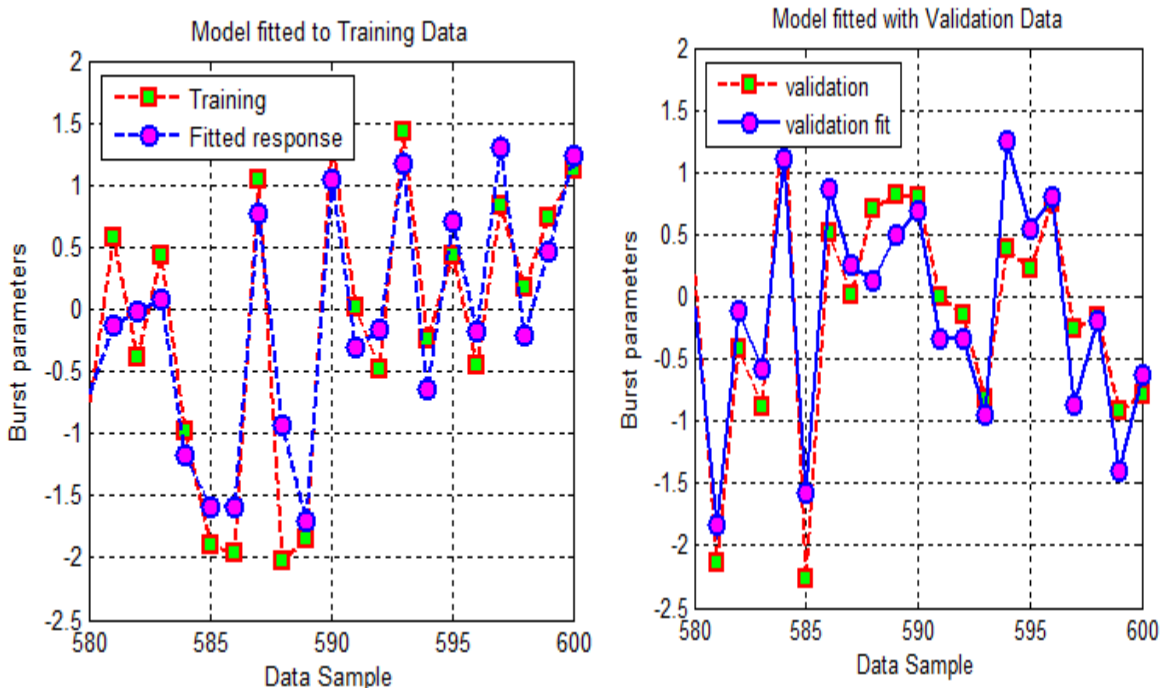


Figure 4. 7 Prediction trend for bursting strength (Model train and validation)

Summary of simulation of PLS results about bursting prediction

The figure 4.7 shows the output results of the model for the bursting prediction. Both training and validation trends are fitting at the degree evaluated by following parameters:

- R^2 Training = 0.58
- R^2 Validation = 0.78

The Bursting model response fit moderately well to the training model. The trends for training and validation data are almost similar (Figure 4.7). This is an indication of how the future inputs would fit to the process. The residual is the difference of the actual value from the predicted value.

The residual computed for our model is negligibly small. This is an indication of good predictions. However, it is very important to use the model for the purpose it has been designed for. Therefore, a follow up study of its feasibility for the practice is advised.

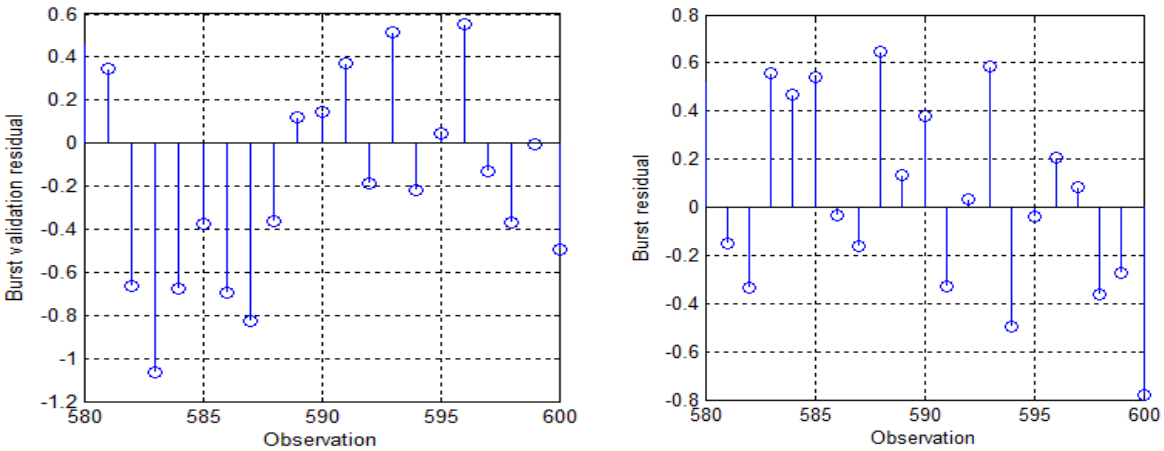


Figure 4. 8 Residual for the prediction of burst (Training and validation error)

Case 2: CMT30 prediction:

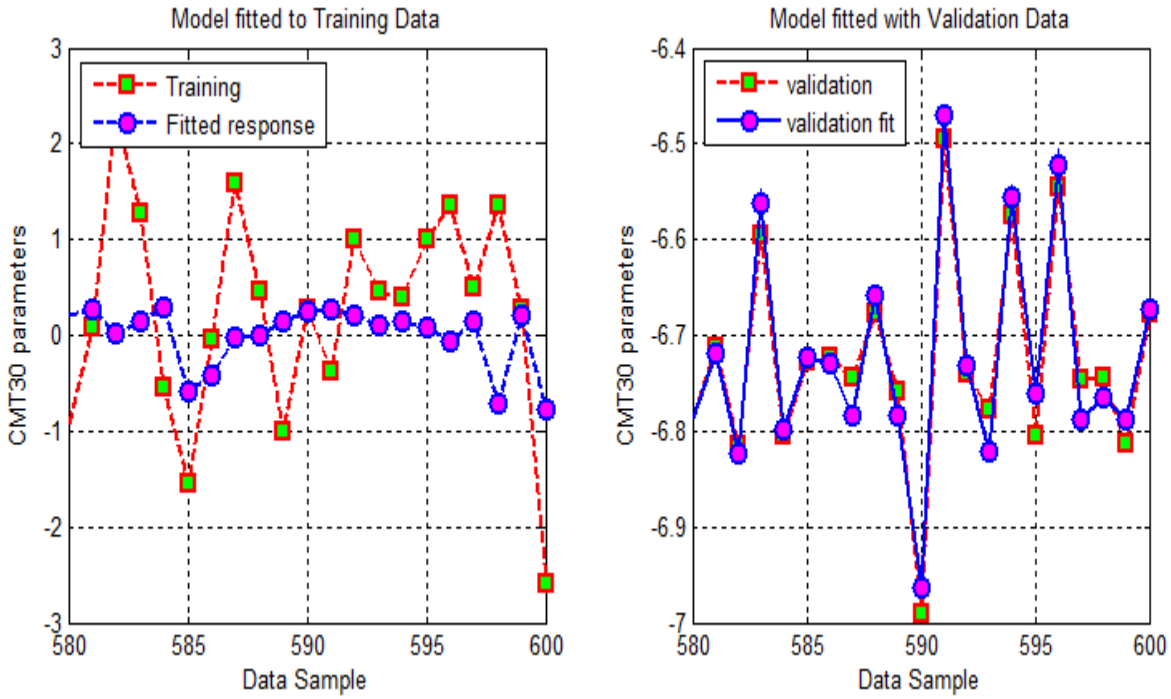


Figure 4. 9 The prediction trends of CMT30 (Model train& Validation)

Summary of simulation results for CMT30

The figure 4.9 shows the output results of the model for the CMT30 prediction. Both training and validation trends results are shown below:

- R^2 Training = 0.58
- R^2 validation= 0.81

The CMT30 training model is not good but mode validation response fit well to actual measured data. This is quite strange; it may be due to multicollinearity between the process variables.
The prediction errors resulting from the predicted quality are shown below:

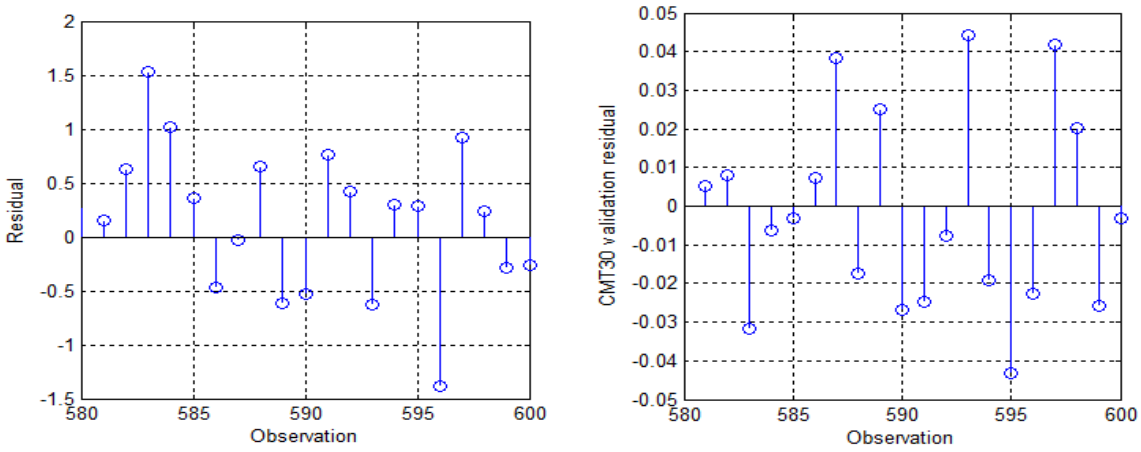


Figure 4. 10 Residual from CMT30 prediction.

The prediction errors for CMT30 prediction are higher for the model train and lower for the model validation. However, the model can predict the quality at reasonable predictions. The model validation errors are less than 10%. This is a perfect fit of the model validation. However due to quality figures accuracy, a little improvement is needed before it is implemented in practice.

CASE3: SCTcd prediction:

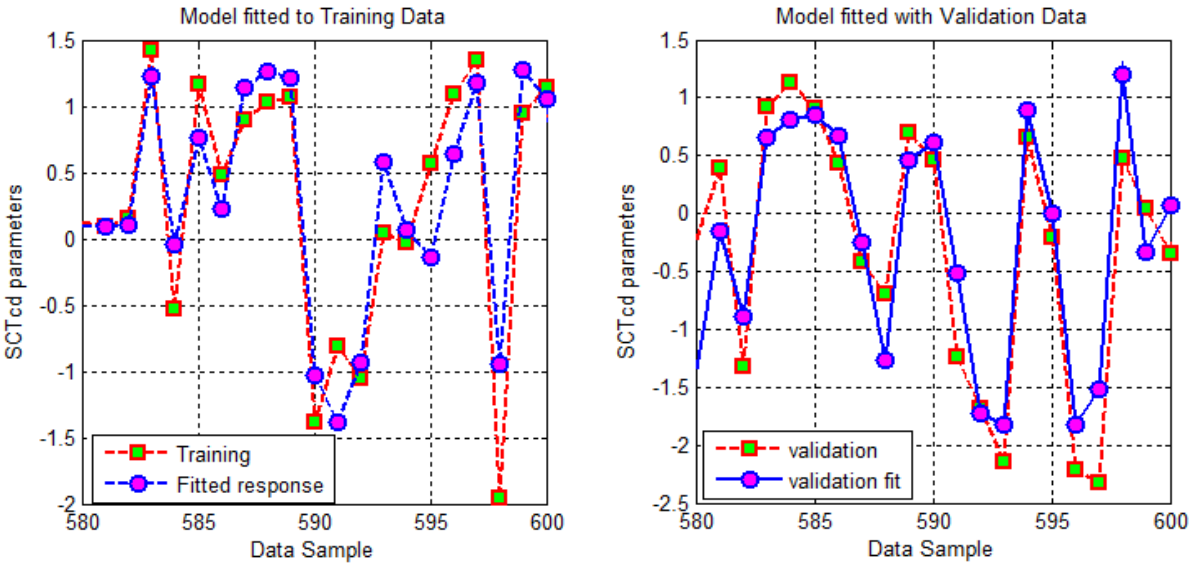


Figure 4. 11 Prediction trend of SCTcd quality (Training and validation)

Summary of simulation results about SCTcd prediction

The figure 4.11 shows the output results of the model for the SCTcd prediction. Both training and validation trends results are shown below:

- R^2 Training = 0.580
- R^2 validation= 0.804

The SCTcd model response fit well to the training model. The prediction values from the actual data are good. This means that the predictions from the given model by using process variables come up to moderate expectations. They are moderate because of the highest accuracy of the quality figures.

The difference of the line of prediction and the actual measured values results in the residual trend for both training model and validation of the model. The Figure 4.11 shows the residual for training and validation of model. The residual trend for some values is near to zero and for others values are even more. The average and absolute value of residuals is 0.6. This error cannot be tolerated for SCTcd prediction because the quality figures are estimated at 10% of its margin (see the Appendix F). This means that the model for ten dimensional approach would be improved before it is reliable for practical purpose.

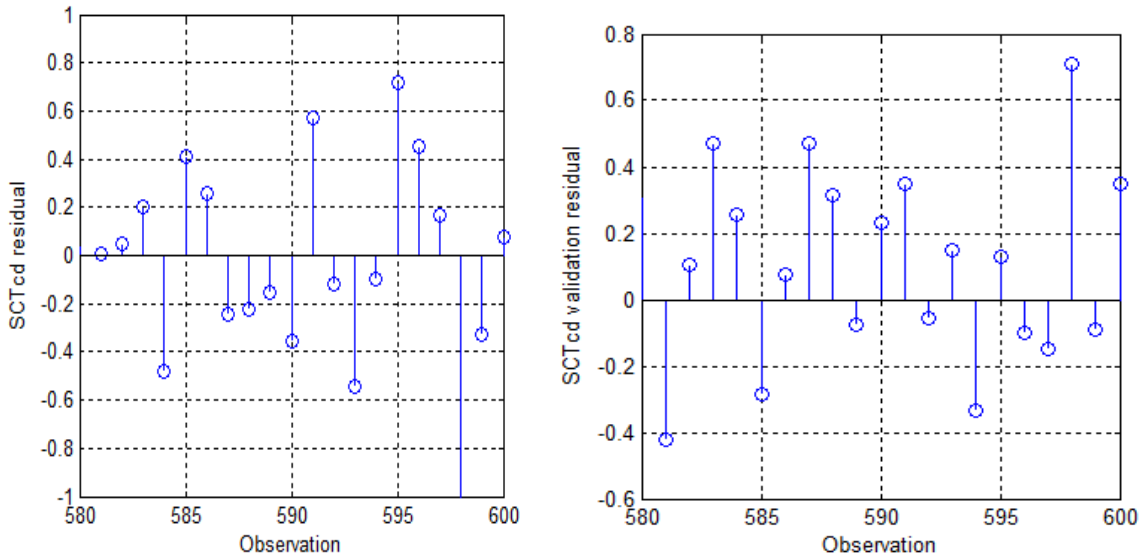


Figure 4. 12 Residual from SCTcd prediction (Training and validation error)

4.9 SUMMARY RESULTS AND DISCUSSIONS

The design model assumes static linear model. The E-modulus results gives the unexpected correlation as shown by the deterministic correlation coefficient (R^2) which is varying between 0.26 and 0.29.

The Appendix D shows model trials for static non linearities tests but unfortunatly the simulations results did not find any improvement for the quality prediction ,the trials result in non correlations.No correlation was found.

Furthermore, the process variables was also used to find how good they are correlated to the quality prediction. This is seen for ten dimensional approach.The results show that they are better than elastic modulus results. However by refering to the quality figures in Appendix F, it is obvious that the minimum correlation desired for simulations of the paper quality online with respect to the input variables of the process would be at least 0.9 based on the target value of the quality .The reason is due to the market standards and ISO standards that a recycled paper should have before it is sold. Since the quality figures are much more sensitives to the target values, it is advisable to optimize the process variables for ten dimensional approach by looking on other valuable inputs which are not specified in the patent [1] . The elastic modulus theory is correct but the simulations show the incapacity of its implementation on inline measurements of paper web and control.

5. CONCLUSIONS AND RECOMENDATIONS

5.1 CONCLUSIONS

The goal of this thesis is mainly devoted to the practical feasibility of the patent [1].The investigation about the study of this thesis used the statistical analysis that leads to the predictions of the paper quality. The analysis of mathematical derivations together with the matlab® simulations was done so far and gave the unexpected results. The results are very bad for the case where E-modulus is used as the basis for the quality predictions. Referring to the following: equations (4.3) of the correlation matrix, C_E and equation (4.10) for PCA analysis, equation (4.27) for the PLS regression together with matlab® simulations from the Figure 4.3 up to the Figure 4.5, it is obviously concluded that the use of the patent structure cannot help for the quality predictions.

Due to low correlation of quality with E-modulus, the following trials were done in order to probably get the better correlation. Indeed, the inverse, square and square root of E-modulus were calculated and then they were correlated to the three qualities. The idea was to check for the presence of static non linearities that can be compensated each by their inverse functions. However all the steps done for elastics modulus give unsatisfactory results.The results from those trials are shown in the Appendix D.
The next and the last steps was to look on the other possible causes of low correlation.
The following points list the investigated possible causes of low correlations:
1. **Data inaccuracy due to sampling time:** it is impossible to get the paper quality at the exact time of production. It takes almost half an hour to get the quality from the laboratory. To estimate the corresponding data online we took data two minutes before the end of jumbo reel to approximate the matching time between real production and the quality measurements. This time was considered in order to avoid the case of switching to the data of the next reel of production.

2. **Inaccuracy of data due to either the data compression in the PI storage system or time dependency of online measurement:** Every data in the database is compressed with a minimum of 0.01% .This small compression may hide some important variations in the process. The more impact may be big for the speed difference where the difference between V_1 and V_2 is small. The simulation results showed that at certain data observed, there are good estimates of the quality especialy when raw data for the process variables are considered for simulations [See Figure 4.7 & 4.11]. The figure in Appendix E shows that V_1 and V_2 are more correlated,but in the process definition they must be always different in order to pull out the paper from one position to an other.Since they are much more closer, the small change happen to one can affect the other .

3. **Important process variables may not be specified in the patent model**: Examples are starch or other additional chemicals needed for the recycle papers.
5**. Lack of use of raw data** i.e data measured from the plant without experimentally monitoring them can give good results because they contain direct information from the process.
Example is the comparison plots between 6-dimensional approach and 10-dimensional approach.
6. **Non correlation at all**: It may be practically possible for not having any correlation by following the proposed patent model [1]. The Appendix D also proves random distribution of data.

5.2 RECOMMENDATIONS

1. Analysis of sampling time.
   It is possible for selecting a data online which does not correspond exactly to the paper quality or which does not contain enough information about the quality due to the wrong choice of sampling time (Notice from this thesis that a data before two minutes was chosen before the end of jumbo), it would be adviced to make further analysis about the choice of sample data online.It was also noted that the sampling frequency of the sample data online is 1Hertz.This is because each second a new data is stored in the system.Therefore the analysis about sampling frequency is highly needed.

2. The research about the correlation coeffcient for ten dimensional approach shows that raw data for process variables perform well.Unfortunately they can not reach to the required quality figures which requires at least 90% of the quality prediction as it is specified in Appendix F.To be specific, let us take an example of paper grade known as testliner 150 g/m² where the following quality figures are mentioned in the highlighted circle in Appendix F.
   -Target value is compared with actual value in the process: 2.85 [kN/m]
   -Margin is compared with minimum acceptable prediction value : 2.69 [kN/m]
   -The difference between target value and margin value is equivalent to the residual error of prediction which is equal to 0.16 .
   -It is obvious that in order to have this residual prediction, the coefficient of determination should be at least 0.84.This is almost 90% of the total information in the data set.To achieve this correlation requires to have a perfect model of fit.Then the data should much more accurate with negligeable noises.This is quite challenge in real paper process where temperatures and moisture are not exactly controlled by the paper maker.
   -Due to the requirements of high correlation coefficient, it is clear that the process variables should be optimized in order to be used for quality prediction online.In this case, important inputs which has not been specified in the patent[1] would be checked. The example given is starch in this case.

3. The research also looked at the static non-linarites on the E-modulus, which can be compensated by their inverse functions. However square & square root of Emodulus, E-modulus and its inverse did not prove the coefficient of determination (see Appendix D the plot of E-modulus vs quality). It is therefore requested to look on the dynamics effect of the plant to see if they have an impact on the paper quality.

4. The last but not the least is to check the data storage system and the status of the measuring machine in the laboratory. This will ensure the storage of good data, which can help in the quality prediction.

APPENDIX A: SUBSTITUTION OF E-MODULUS FORMULAS

The E-modulus (Emod) relates the linear stress (δ) to the strain (∈) for an elastic material and its expression is derived as shown below.

The tensile stress, denoted as δ, is the force (F_t) per unit of initial transverse cross sectional area (A) within a portion of the web through which the force acts and it is calculated as follows:

$$\delta = \frac{F_t}{A} \ [N/m^2] \qquad (A.1)$$

The uniaxial strain, denoted as ∈, is a measure of the geometrical deformation of the portion of web due to the applied stress. It may thus be defined as the ratio of the deformation (ΔL) of the portion of the web to the initial dimension (L).

$$\in = \frac{\Delta L}{L} \ [-] \qquad (A.2)$$

Accordingly, using the equations (A.1) and (A.2), the young's modulus may be defined as:

$$Emod = \frac{\delta}{\in} = \frac{\frac{F_t}{A}}{\frac{\Delta L}{L}} \ [N/m^2] \qquad (A.3)$$

Its standards units are also [kPa] or [kN/m²] or [N/mm²]

In fact, the Young modulus is the stress needed to get 100% strain in elastic material. In papermaking, 100% stress is imaginary because long before this elongation the paper will be torn.

The present invention may use the fact that, the uniaxial stress δ in a fibrous web may be inferred from the power required to transport the web between two of the rollers i.e. the last drying cylinder and the second roller, i.e poperoller while the strain of the web ∈ may be inferred from the difference in rotational velocity between them.

The power P_2 employed to drive the second roller is calculated as follows:

$P_2 = F_t.V_2$ where F_t is the tensile force applied to the web by the second roller and V_2 is the velocity of the web at the second roller .Similarly, $P_2 = \delta. A. V_2$

Assuming the actual cross sectional surface area of the web, A at point between the two rollers is at good approximation just equal to the web initial cross sectional surface area, the stress may be written as follows:

$$\delta = \frac{P2}{A.V2} \qquad (A.4)$$

The strain of the paper web in between the two rollers may be expressed as

$$\in = \frac{V_2 - V_1}{V_1} [-] \qquad (A.5)$$

Equations A.4 and A.5 may be combined to give the resulting Young modulus of the paper web material as follows:

$$Emod = \frac{P_2 . V_1}{A . V_2 . (V_2 - V_1)} \ [N/m^2] \qquad (A.6)$$

APPENDIX B: RELATIONSHIP BETWEEN PAPER WEB STRETCH AND SPEED DIFFERENCE OF THE ROLLERS.

The equation (A.5) can also be rewritten in terms of elongation as follows:

$$\text{Strain } (\epsilon) = \frac{\Delta l}{L} = \frac{V_2 - V_1}{V_1} \; [-] \qquad \text{(B.1)}$$

Notes; this is shown for example when both last drying cylinders speed (V_1) are exactly equal to poperoller speed (V_2), it sound to be without stretch in the paper web.
Let's take example where V_2 speed is twice the first roller speed (V_1),

Strain $(\epsilon) = \frac{\Delta l}{L} = \frac{(2v_1) - v_1}{v_1} = 100\%$ stretch in paper! This is a theoretical assumption and it cannot happen for a real process.

APPENDIX C: RELATIONSHIP BETWEEN POPEROLLER POWER (P_2) SUPPLIED TO THE PAPER WEB AND ITS MOMENT OR COUPLE.

A couple such as represented in Figure C.1 consists of two equal but opposite forces whose lines of action do not coincide. A couple tends to produce rotation. The product of the force and the distance gives the work done.

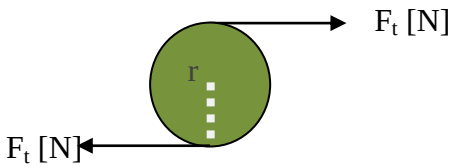


Figure C. 1 Poperoller drive with its moment forces

For single rotation, the distance travelled is equal to the perimeter of the circle (diameter.π) and the work done is given by the same force times the distance travelled.
Similarly, the work done in rotational case is given by torque times the angle turned.
In general, the work done by couple is expressed by:

$$W = 2.\, F_t.\, r.\, \theta \; [J] \qquad \text{(C.1)}$$

where:

W = the work done [J]
F_t = the force [N] applied to the paper web by poperoller.
D = 2.r is the axial distance between couple
θ = the axial angle of rotation by poperoller.
For 'k' rotations in 1sec, the work performed during that period of duration is now,

$$W = F_t.\, (2.\, r)\, .\, \pi\, .\, k \; [J] \qquad \text{(C.2)}$$

Therefore, the power as the rate at which the stretch energy is transferred into the paper is given by:

$$P_2 = \frac{W}{t} = 2\pi . \frac{F_t . r . k}{1s} \text{ [J/s]} \quad \text{(C.3)}$$

where k is the number of rotations made by the poperoller.

From this expression of power, the poperoller speed is given by

$$V_2 = \frac{2\pi . r . k}{1s} \text{ [m/s]} \quad \text{(C.4)}$$

This is the same as $P_2 = F_t.V_2$ where V_2 is the speed of the poperoller.
Then after the stress is thus defined by:

$$\frac{F_t}{A} = \frac{P_2}{V_2 . A} \text{ [N/m²]} \quad \text{(C.5)}$$

and strain is defined by the equation (B.1).

Finally, the elastic modulus is defined by the ratio of stress over strain and is thus given by equation (A.6)

APPENDIX D: STATIC NON-LINEARITIES TRIALS FOR E-MODULUS

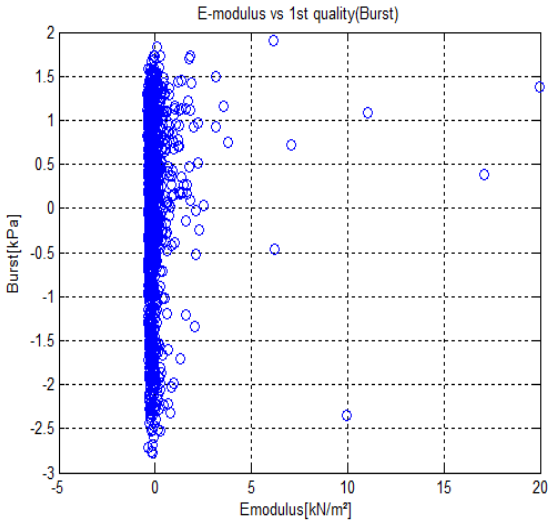


Figure D.1: E-modulus vs Burst.

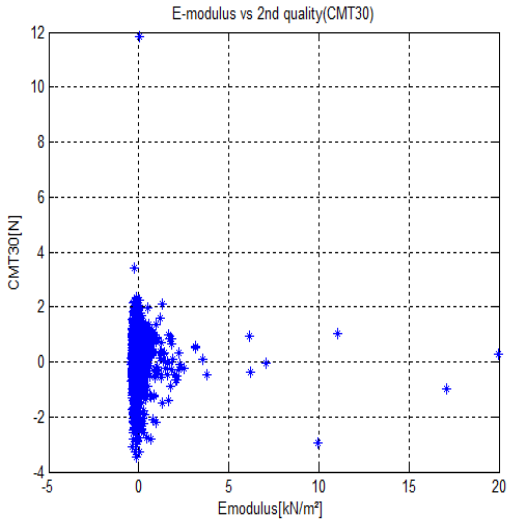


Figure D.2: E-modulus vs CMT30.

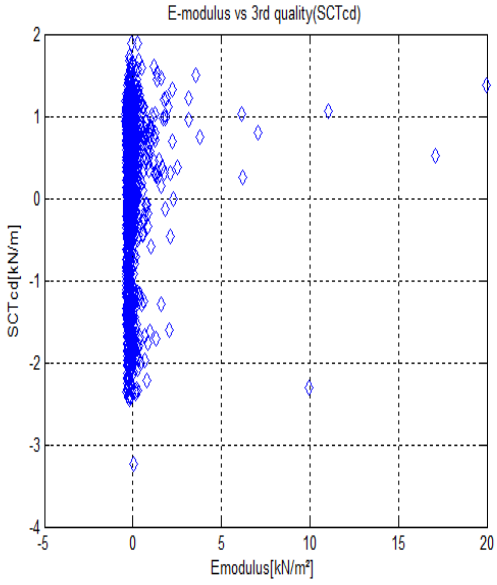


Figure D.3: E-modulus vs SCTcd

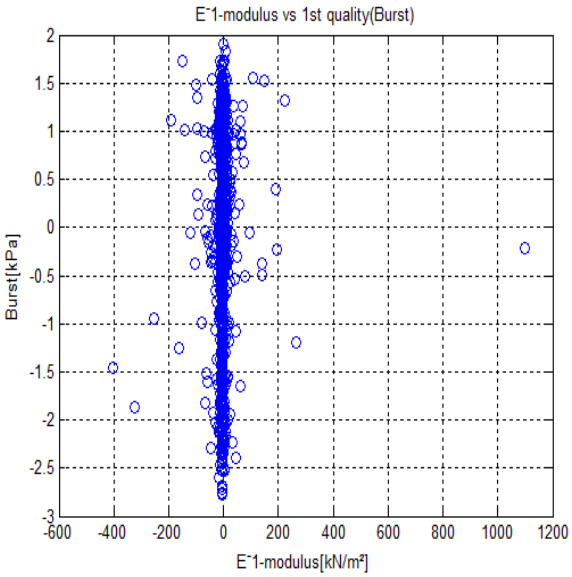


Figure D.4: E-1 modulus vs Burst

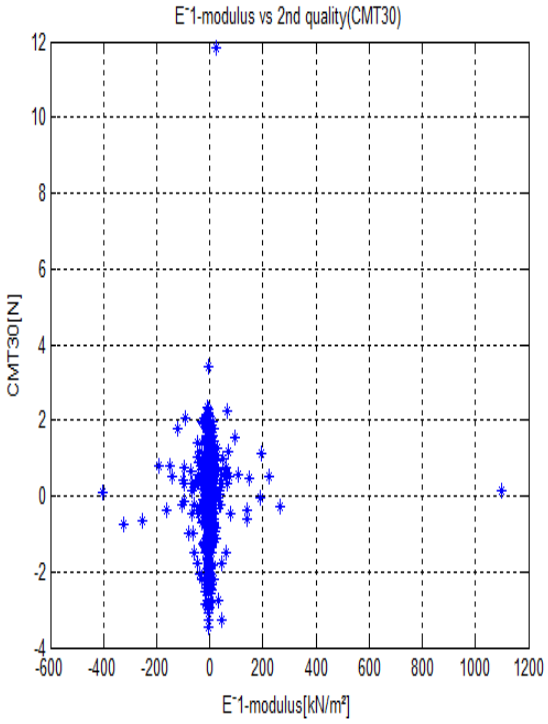


Figure D.5: E^{-1}-modulus vs CMT30

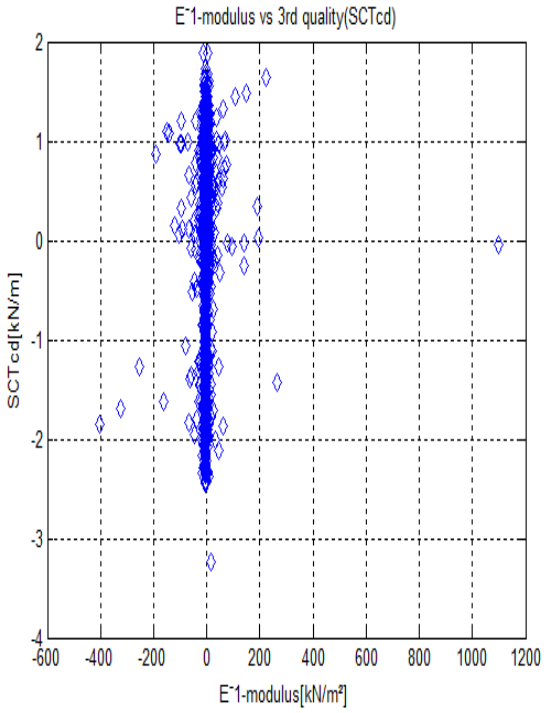


Figure D.6: E^{-1}-modulus vs SCTcd

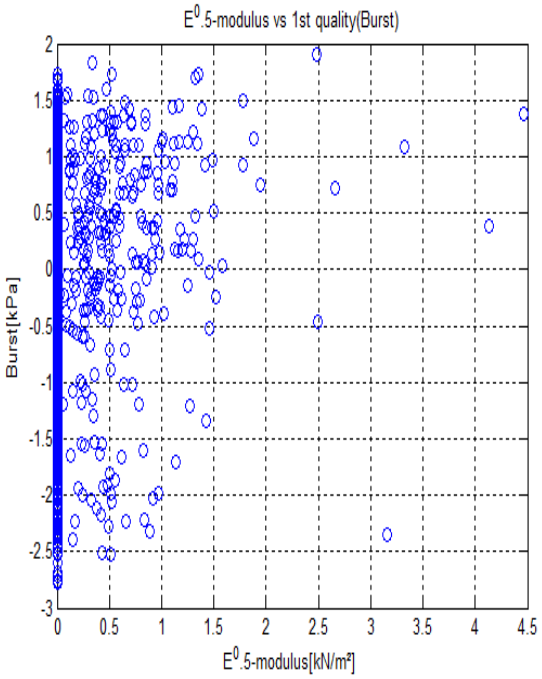


Figure D.7: $E^{1/2}$-modulus vs Burst

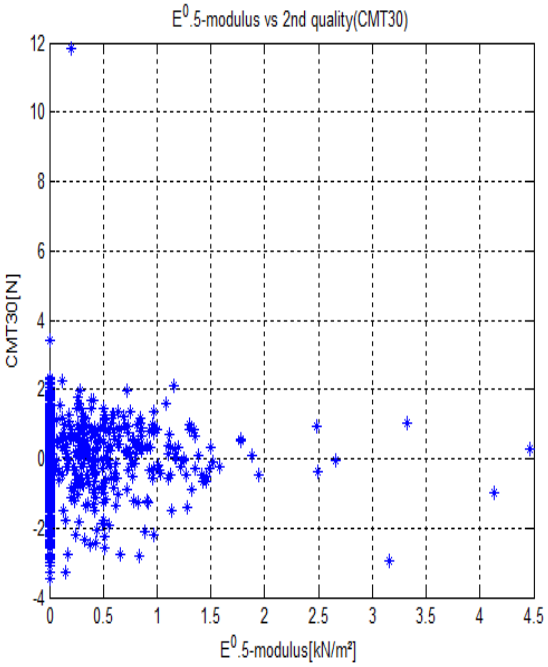


Figure D.8: $E^{1/2}$-modulus vs CMT30

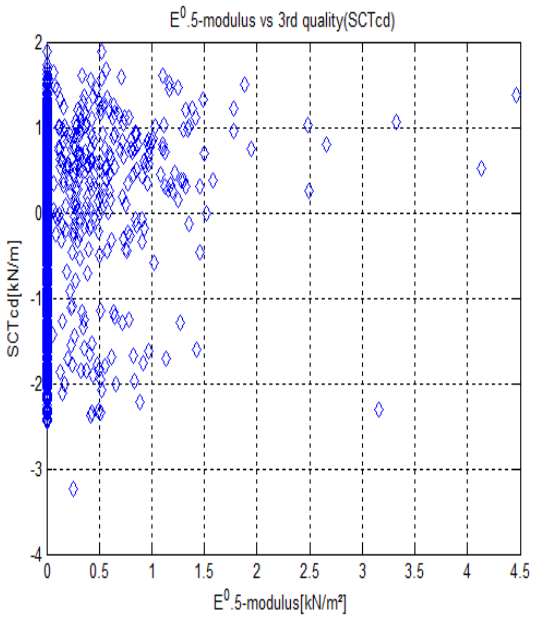


Figure D.9: $E^{1/2}$-modulus vs SCTcd

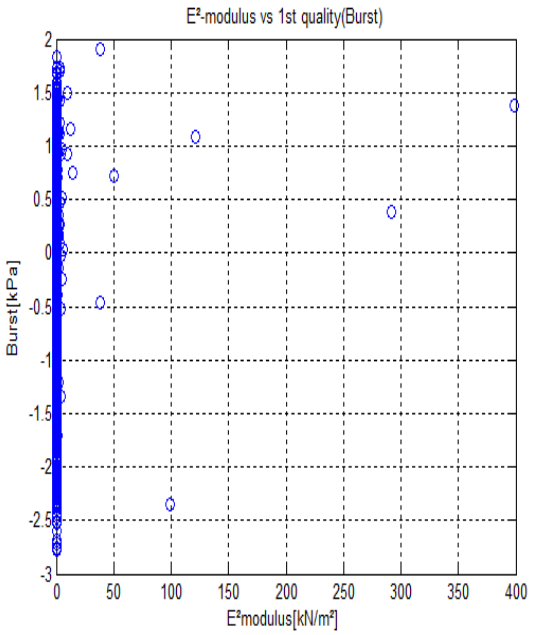


Figure D.10: E^2-modulus vs Burst

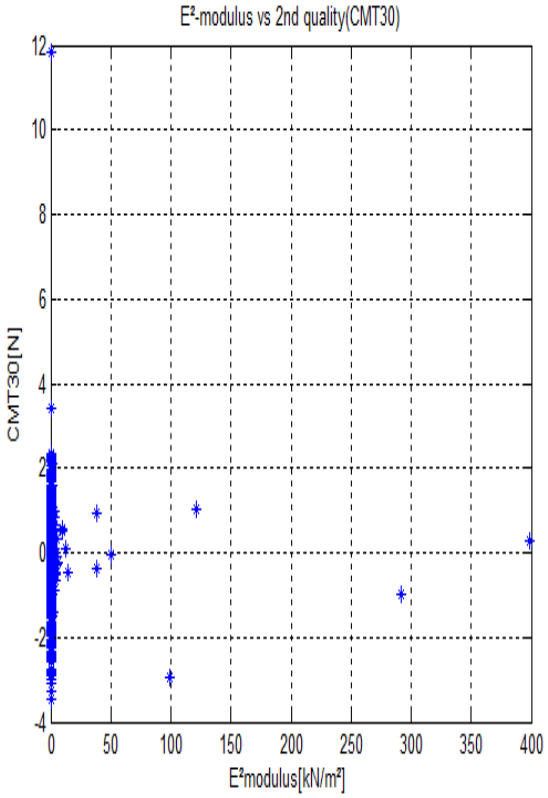


Figure D.11: E^2-modulus vs CMT30

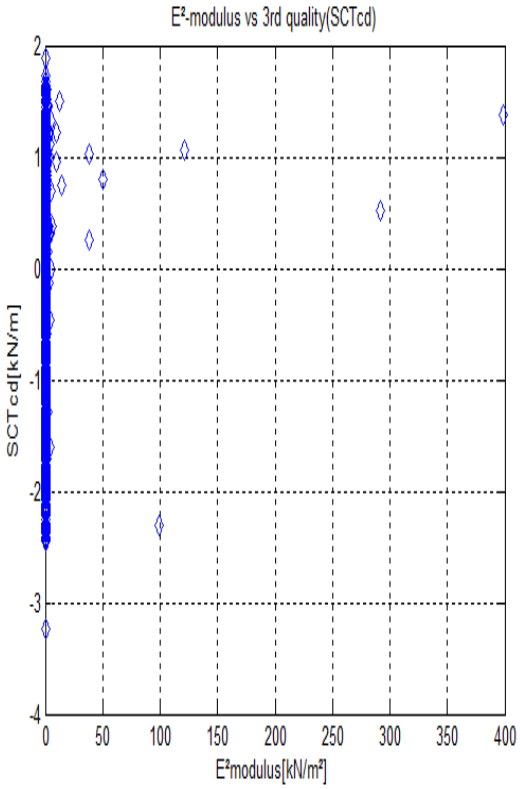


Figure D.12: E^2-modulus vs SCTcd

Figure D. 1 Static Non-linearity trial

APPENDIX E: FIGURES FOR PVs IN SAMPLED TIME

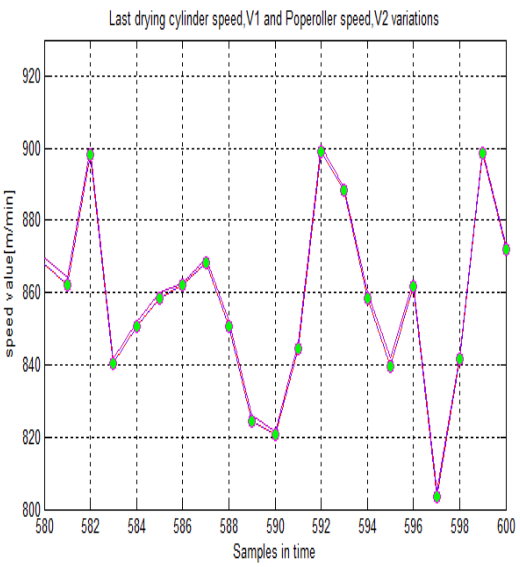


Figure E.1: V_1 & V_2 vs sampled time

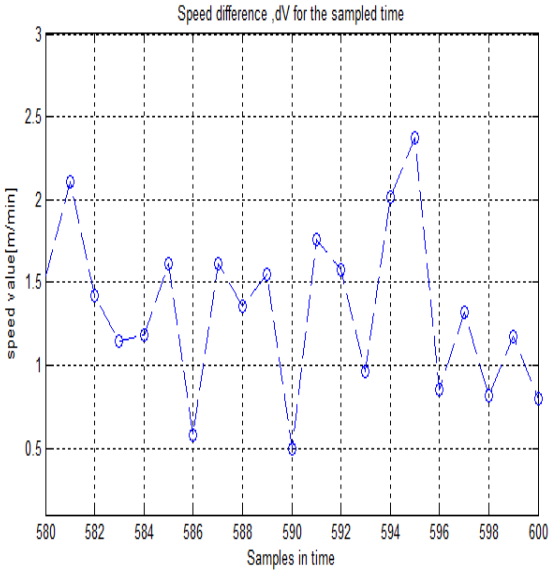


Figure E.2: dV vs sampled time

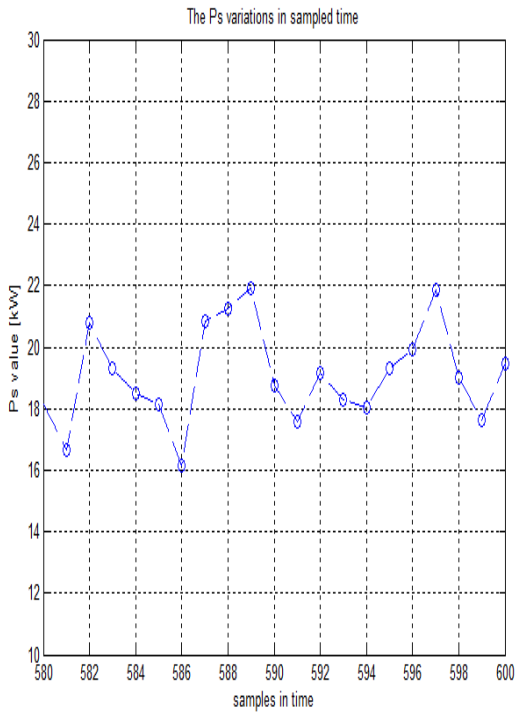


Figure E.3: P_s vs sampled time

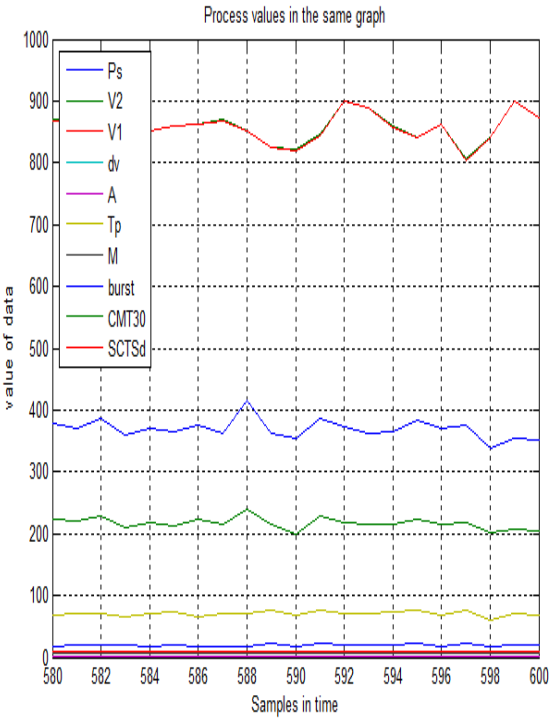


Figure E.4: PVs P_s vs sampled time

Figure E. 1 PV in sampled time

APPENDIX F: EXAMPLES OF PAPER SPECIFICATIONS FOR SKRP.

Smurfit Kappa
Roermond Papier

Product specificaties

Papiersoort Recycled Fluting 1
Soortnummer 20
Receptcode 45

Gramgewicht	gr/m²	nominaal	90	100	105	110	112	115	120	125	130	135	140	150	160	170	175
gemiddeld	gr/m²	maximum	92,7	103,0	108,2	113,3	115,4	118,5	123,6	128,8	133,9	139,1	144,2	154,5	164,8	175,1	180,3
	gr/m²	minimum	87,3	97,0	101,9	106,7	108,6	111,6	116,4	121,3	126,1	131,0	135,8	145,5	155,2	164,9	169,8
piek tot piek 1)	gr/m²	maximum	5,4	6,0	6,3	6,6	6,7	6,9	7,2	7,5	7,8	8,1	8,4	9,0	9,6	10,2	10,5
Berstdruk	kPa	nominaal															
	kPa	herkeur															
CMT langs	N	nominaal	140	165	175	*185*	190	*200*	210	*220*	*230*	*235*	245	260	*270*	*275*	280
	N	herkeur	135	160	*170*	179	*185*	188	197	206	214	222	*230*	245	253	267	*275*
SCT dwars	kN/m	nominaal	1,50	1,65	1,75	*1,80*	1,85	1,90	2,00	*2,05*	2,15	*2,20*	2,30	*2,45*	*2,65*	*2,80*	2,90
	kN/m	herkeur	1,42	1,58	1,66	*1,72*	1,77	1,81	1,89	*1,95*	2,05	*2,10*	2,21	*2,35*	*2,52*	2,68	2,77

Papiersoort Testliner 3 vanaf 150 grs
Soortnummer 50
Receptcode 64

Gramgewicht	gr/m²	nominaal	160	170	180	190	200	210	220	230
gemiddeld	gr/m²	maximum	164,8	175,1	185,4	195,7	206,0	218,4	228,8	239,2
	gr/m²	minimum	155,2	164,9	174,6	184,3	194,0	201,6	211,2	220,8
piek tot piek	gr/m²	maximum	9,6	10,2	10,8	11,4	12,0	16,8	17,6	18,4
Berstdruk	kPa	nominaal	*360*	*375*	*400*	*420*	*440*	*455*	*470*	*490*
	kPa	herkeur	*340*	*355*	*375*	*395*	*410*	*425*	*440*	*460*
CMT langs	N	nominaal								
	N	herkeur								
SCT dwars	kN/m	nominaal	*2,55*	*2,70*	*2,85*	*3,00*	*3,15*	*3,25*	*3,40*	*3,55*
		herkeur	*2,40*	*2,53*	*2,69*	*2,84*	*2,99*	*3,12*	*3,26*	*3,41*

Papiersoort Testliner 3 100 t/m 150 gr
Soortnummer 50
Receptcode 64

Gramgewicht	gr/m²	nominaal	100	110	120	125	130	135	140	145	150
gemiddeld	gr/m²	maximum	103,0	113,3	123,6	128,8	133,9	139,1	144,2	149,4	154,5
	gr/m²	minimum	97,0	106,7	116,4	121,3	126,1	131,0	135,8	140,7	145,5
piek tot piek 1)	gr/m²	maximum	6,0	6,6	7,2	7,5	7,8	8,1	8,4	8,7	9,0
Berstdruk	kPa	nominaal	*220*	*240*	*260*	*275*	*285*	*300*	*310*	*325*	*340*
	kPa	herkeur	*210*	*225*	*250*	*265*	*270*	*290*	*300*	*315*	*325*
CMT langs	N	nominaal									
	N	herkeur									
SCT dwars	kN/m	nominaal	*1,60*	*1,75*	*1,90*	*1,95*	*2,00*	*2,10*	*2,15*	*2,25*	*2,35*
	kN/m	herkeur	*1,47*	*1,63*	*1,79*	*1,84*	*1,90*	*2,00*	*2,05*	*2,15*	*2,26*

Papiersoort Testliner 2 tot 190 grs
Soortnummer 60
Receptcode 74

Gramgewicht	gr/m²	nominaal	125	*130*	135	140	145	150	160	170	175	180
gemiddeld	gr/m²	maximum	128,8	*133,9*	139,1	144,2	149,4	154,5	164,8	175,1	180,3	185,4
	gr/m²	minimum	121,3	*126,1*	131,0	135,8	140,7	145,5	155,2	164,9	169,8	174,6
piek tot piek 1)	gr/m²	maximum	7,5	*7,8*	8,1	8,4	8,7	9,0	9,6	10,2	10,5	10,8
Berstdruk	kPa	nominaal	310	*320*	335	350	365	380	400	420	430	445
	kPa	herkeur	290	*310*	320	332	345	360	385	405	410	420
CMT langs	N	nominaal										
	N	herkeur										
SCT dwars	kN/m	nominaal	2,15	*2,20*	2,30	2,39	2,47	2,55	2,73	2,90	3,00	3,07
	kN/m	herkeur	2,05	*2,15*	2,21	2,28	2,37	2,47	2,61	2,78	2,85	2,94

Note: The specifications of the paper standards above were kept in the language of their original copies.
SCTcd means SCT dwards;
Bursting strength means Berstdrunk
CMT30 means CMT Langs.

APPENDIX G: PCA&PLS R MATLAB FILES FOR SIX DIMENSIONAL APPROACH:

```
%ILSE ONE MONTH DATA FROM SKRP, ROERMOND (BY USING E-MODULUS)

clc
clear all
format long

txt =  [' Emodulus[kN/m²]     ';...
        ' Temperature[°C]     ';...
        ' Moisture[%wt]       ';...
        ' Burst[kPa]          ';...
        ' CMT30[N]            ';...
        ' SCTcd[kN/m]         '];

load ilse0012.mat % load the data set in matlab®

ilse12 =data(randperm(1911),:);%Randomize data to evenly distribute the data
%Inputs variables
x1 = ilse12(:,1);x2 = ilse12(:,2);x3 = ilse12(:,3);x4 = ilse12(:,4);
x5 = ilse12(:,5);x6 = ilse12(:,6);x=[x1 x2 x3 x4 x5 x6] ;
figure(1)
[n,p] = size(data);
% Define the x-values
t = 1:n;
% Plot the variables in sampled time
plot(t,data)
grid on
legend('Emod','Tp','M','burst','Cmt30','SCTcd',...
'Location','northwest')
xlabel('Samples in time'), ylabel('value of data')% at avg of 22.6 min for
each sample

% choice of data with 2 data for training,one data for validation

trainingdata   = ilse12(1:1274,:);% total of 1274 observations(2/3).
validationdata = ilse12(1275:1911,:);% 637 observations(1/3).

% column separation

x1t = trainingdata (:,1);x2t = trainingdata (:,2);x3t = trainingdata (:,3);
x4t = trainingdata (:,4);x5t = trainingdata (:,5);x6t = trainingdata (:,6);
% Autoscaling process/standardization to eliminate the effect of magnitude
% of data set.
x1S = std(x1t);x1M = mean(x1t);xasc1  =(x1t-x1M)/x1S;
x2S = std(x2t);x2M = mean(x2t);xasc2 = (x2t-x2M)/x2S;
x3S = std(x3t);x3M = mean(x3t);xasc3 = (x3t-x3M)/x3S;
x4S = std(x4t);x4M = mean(x4t);xasc4 = (x4t-x4M)/x4S;
x5S = std(x5t);x5M = mean(x5t);xasc5 = (x5t-x5M)/x5S;
x6S = std(x6t);x6M = mean(x6t);xasc6 = (x6t-x6M)/x6S;

x1v = validationdata(:,1);x2v = validationdata (:,2);
x3v = validationdata(:,3);x4v = validationdata (:,4);
x5v = validationdata(:,5);x6v = validationdata (:,6);

xasc1v  = (x1v-x1M)/x1S;xasc2v = (x2v-x2M)/x2S;xasc3v = (x3v-x3M)/x3S;
xasc4v = (x4v-x4M)/x4S;xasc5v = (x5v-x5M)/x5S;xasc6v = (x6v-x6M)/x6S;
xascv = [xasc1v xasc2v xasc3v xasc4v xasc5v xasc6v];
```

```
% PCA -to obtain information on variance in input variables
Xasct = [xasc1 xasc2 xasc3 xasc4 xasc5 xasc6];
[coefs,scores,variances,tsquare]= princomp(Xasct);% contains the coefficients
%of the linear combinations of the original variables that generate the PCs.
% T2 is a statistical measure of the multivariate distance
%of each observation from the center of the data set

variances;% square of standards deviations
percent_explained = 100*variances/sum(variances);
corrcoef(Xasct);%calculates a matrix R of correlation coefficients for
%an array Xasct, in which each row is an observation and each column is a
%variable.

figure(2)
vlabs = {'Emod','Tp','M','Burst','CMT30','SCTcd'};
biplot(coefs(:,1:3), 'scores',scores(:,1:3), 'varlabels',vlabs);
%Use the biplot function to help visualize both the PrinComp coefficients for
%each variable and the PrinComp scores for each observation in a single plot.
%the direction and length of the vector indicates how each variable...
%contributes to the three principal components in the plot.
grid on;
title('Biplot in 3D')
figure(3)
pareto(percent_explained);%Pareto charts display the values
%in the vector Y as bars drawn in descending order.
xlabel('Number of principal components')
ylabel('Variance explained in(%)')
grid on
title('Explained variance in dataset wrt the process variables')
%PLSR Modelling
figure(4)  %inpts vs outputs----------
Xmod=[xasc1 xasc2 xasc3];%define inputs data
Ymod=[xasc4 xasc5 xasc6];%define output data
[XL,YL,XS,YS,beta,PCTVAR,MSE,stats] = plsregress(Xmod,Ymod,3);
Yfitmod = [ones(size(Xmod,1),1) Xmod]*beta;
plot(Ymod,Yfitmod,'o');
xlabel('scatter plot -observed Response');
ylabel('model Response');
grid on
figure(5)
Xv=[xasc1v xasc2v xasc3v];Yv=[xasc4v xasc5v xasc6v];
[XLo,YLo,XSo,YSo,betao,PCTVARo,MSEo,statso] = plsregress(Xv,Yv,3);
Yfitval = [ones(size(Xv,1),1) Xv]*betao;
plot(Yv,Yfitval,'*');
xlabel('scatter plot-new observed data');
ylabel('model validaion');
grid on
% Model parameter check
TSS = sum((Ymod-repmat(mean(Yfitmod),1274,1)).^2);
RSS = sum((Ymod-Yfitmod).^2);
Rsquared = 1-(RSS/TSS);
TSSv = sum((Yv-repmat(mean(Yfitval),637,1)).^2);
RSSv = sum((Yv-Yfitval).^2);
Rsquaredv = 1-(RSSv/TSSv);
figure(6)
plot(1:3,cumsum(100*PCTVAR(2,:)),'-bo')%we can easily calculate the
% of the total variability explained by each PC.
xlabel('Number of PLS components');
ylabel('Percent Variance Explained in y');
grid on
```

```
title('Percent of Variance Explained by the model w.r.t three PLS
Components')
legend('PLSR','location','NW');
% The 1st row of PCTVAR contains the% of variance explained in X(Predictors)
%by each PLS component&The 2nd contains the% of VAR explained in
Y(Predicted).
rmse = sqrt(mean(Ymod-Yfitmod).^2);
rmsev = sqrt(mean(Yv-Yfitval).^2);
figure(7)
%Compute the fitted model response and display the residuals:
Yfitmod = [ones(size(Xmod,1),1) Xmod]*beta;
residuals = Ymod(:,1)-Yfitmod(:,1);
stem(residuals)%STEM(Y) plots the data sequence Y as stems from the x axis
% terminated with circles for the data value. If Y is a matrix then
% each column is plotted as a separate series.
xlabel('Observation');
ylabel('Residual');
grid on

% Plot Training model
figure(8)
plot(Ymod(:,1), '--rs','LineWidth',2,...
                 'MarkerFaceColor','g',...
                'MarkerSize',8)
legend('Training Data')
hold on

plot(Yfitmod(:,1),'--bo','LineWidth',2,...
                  'MarkerFaceColor','m',...
                  'MarkerSize',8);
hold off
title('Model fitted to Training Data')
xlabel('Data Sample')
ylabel('Quality parameters')
grid on
legend('Training model','fitted response',2);

% Plot Validation model
figure(9)
plot(Yv(:,1),'--rs','LineWidth',2,...
                  'MarkerFaceColor','g',...
                  'MarkerSize',8);
legend('validation Data')
hold on
plot(Yfitval(:,1),'-bo','LineWidth',2,...
                 'MarkerFaceColor','m',...
                 'MarkerSize',8)
hold off
title('Model fitted with Validation Data')
xlabel('Data Sample')
ylabel('Quality parameters')
grid on
legend('Model validation ','Model validation fit',2);
```

APPENDIX H: GLOSSARY OF PAPER AND PAPER BOARDS [SCAN-P: 75]

1. **Relative humidity**: ratio of the partial pressure of the water vapor to its saturation pressure at the same temperature. This ratio always is expressed in [percentage].

2. **Sample conditioning**:

In moisture equilibrium with a conditioning atmosphere: a sample of paper or board is conditioned and said to be in equilibrium with the conditioning atmosphere when it has been exposed to the atmosphere long enough for the results of two weighing of at least 1hour, not to differ by more than 0.25 per cent

3**. Conditioning atmosphere**:

For general purposes the conditioning atmosphere in testing paper and board shall be as follow
Temperature: 23±1°C and Relative humidity: 50%±2%

4. **The moisture content** of the paper is defined as the loss of weight of a sample, dried under specified conditions to constant weight at a temperature of 103±2°C and is expressed as percentage of the weight of the moist sample.

5. **Grammage** is defined as the mass per unit area. Its unit is [g/m^2].

6. **Thickness**: is defined as the distance between one surface of a sheet of paper or board and the other.

7. **The machine direction** (MD) is the dimension of a paper corresponding to the direction of flow of the paper web on the paper machine. Most of the fibres, especially on the wire side, are oriented in the machine direction.

8. **The cross direction** (CD) is at right angle to the machine direction.

9. Tensile strength and stretch

9.1 **A tensile strength** is defined as the maximum tensile force per unit width that a test piece of a paper or board will stand before breaking in a tensile test.

9.2 Tensile index: tensile strength divided by the grammage.

9.3 **Stretch at rupture**: is defined as the ratio of the increase in length of a test piece of a paper or board at the time of maximum tensile force during a tensile test, to its length before the test. The principle of measurement is of the pendulum type where the test piece (paper strip of 15 mm wide) is stretched so that the rupture occurs within a specified period. The maximum tensile force the test piece will stand before it breaks and the elongation at the time of rupture are measured.

10. **Tensile energy absorption**: The amount of energy per unit area of paper absorbed during straining until the onset of rupture, the moment of maximum tensile force, in a tensile test.

11. **Tensile energy absorption index**: The amount of energy per unit of mass of paper absorbed during straining until the onset of rupture in a tensile test. For more information including calculation, methods see [SCAN-P 44:81].

BIBLIOGRAPHY

[1] BUMAGA B.V., Apparatus and method for determining Young's modulus of a web, NL Patent Pub.No.PCT/NL2011/050682,2011 or W02012/047105 A1.

[2] Ajit K. Ghosh, 2011. Fundamentals of Paper Drying – Theory and Application from Industrial Perspective, Evaporation, Condensation and Heat transfer, Dr. Amimul Ahsan (Ed.), ISBN: 978-953-307-583-9.

[3] Szewczyk W., Marynowski K., Tarnawski W.,2006, The analysis of Young's modulus distribution in paper plane, Fibers and Textiles in Eastern Europe, 58, 91-94

[4] Pita, www.pita.co.uk/pdf/TestMethods.pdf, December,2012

[5] Testing paper, www.europapier.com/service/knowhow/testingpaper/basic-properties , November, 2012

[6] Elasticity, www.physics.info/elasticity.com, November, 2012

[7] Caulfield D. F., Gunderson D. E, Paper testing and strength characteristics,TAPPI proceedings of the 1988, paper preservation symposium: 1988 October19-21: Washington, DC. Atlanta and Ga: Tappi press: 31-40.

[8] Krzysztof Marynowski, 2008. Non-linear vibrations of the axially moving paper web. Technical University of Łódź, Department of Dynamics of Machines, Poland.

[9] Lauri Salminen, 2003. Aspects of fracture processes in paper. Dissertations of Laboratory of Physics, Helsinki University of Technology, 2003.

[10] M. Marynowski, 2000. Analysis of the mechanical properties of paper on the basis of two dimensional rheological model' (in Polish), PhD. Thesis, Technical University of Lodz, 2000.

[11] Kristian Saliminen, 2010. *The effect of some Furnish and paper Structure Related factors on wet Web Tensile and Relaxation Characteristics*, PhD. Thesis, Lappeenranta ,2010

[12] Tappi T-403, www.tappi.org/content/tag/t403.pdf

[13] Burst Tester, www.techlabsystems.com/en/paper_testing.html , November 2012

[14] SCT,www.ipst.gatech.edu/faculty/popil_roman/newsletters/newsletter0910.pdf,November2012

[15] TAPPI Test method T-809 om-06, Flash crush of corrugating medium(CMT test)

[16] Laser sensor, www.acuitylaser.com/pdf/synchronizing-laser-sensors-thickness measurements-synchronizing dual laser sensors, January 2013.

[17] M. H. Kutner, C. J. Nachtsheim, and J. Neter (2004), "Applied Linear Regression Models", 4th ed., McGraw-Hill/Irwin, Boston (p. 25).

[18] Cantwell G. Carson and Roman E. Popil. Examining interrelationships between caliper, bending, and tensile stiffness of paper in testing validation'.

[19] C.T.J Dodson, Y. Oba, and W.W. Sampson, 2000. On the distribution of mass, thickness, and density in paper.

[20] GA Baum, and CC Habeger, Orthotropic elastics constants of paper, TAPPI, V64, n8, pp97-101, 1981

[21] GA Baum, DC Brennan, CC Habeger, online measurement of paper mechanical properties, TAPPI, V63, n7, pp63-67, 1980.

[22] David W.Vahey, John M.Considine, Andy Kahra and Mark Scotch, comparisons of fiber Orientation and Tensile stiffness orientation measurements in paper.

[23] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery (2007), Numerical Recipes: The Art of Scientific Computing, Chapter 11: Eigensystems, pages = 563-597. Third edition, Cambridge University Press.

[24] P.A.C Ypma, i. Lectures notes on Functions, Spaces, Orthogonality and convexity. January 2013, HAN University of Applied sciences.p.8

[25] Prof.dr.ir. Brian Roffel. Development of black box mathematical models for process variable prediction and statistical models for process analysis. December 2012.HAN university of Applied sciences.p.14.

[26] Saikat Maitra and Jun, 2008. Principle Component Analysis and Partial Least Squares: Two Dimension Reduction Techniques for Regression; Casualty Actuarial Society, 2008 Discussion Paper Program.

[27] Dr.Michael Porter, 2008. Regression and least squares: A Matalb Tutorial. Department of Statistics. North Carolina state University and SAMSI

[28] PLS,www.mathworks.nl/help/stats/examples/partial-least-squares-regression-and-principal-components-regression.html, November 2012

[29] Kenney, J. F. and Keeping,E. S. "Linear Regression and Correlation." Ch.15 in *Mathematics of Statistics, Pt. 1, 3rd ed.*

[30] Princeton, NJ: Van Nostrand, pp. 252-285, 1962Weisberg, S. (1980) .Applied Linear Regression. New York: John Wiley.

[31] Kenney, J. F. and Keeping, E. S. "Linear Regression and Correlation." Ch. 15 in *Mathematics of Statistics, Pt. 1, 3rd ed.* Princeton, NJ: Van Nostrand, pp. 252-285, 1962.